INNOCENZO GASPARINI INSTITUTE FOR ECONOMIC RESEARCH

Institutional Members: CEPR, NBER and Università Bocconi

WORKING PAPER SERIES

Robust Mean-Variance Approximations

Simone Cerreia-Vioglio, Fabio Maccheroni, Massimo Marinacci

Working Paper n. 689

This Version: December, 2022

IGIER – Università Bocconi, Via Guglielmo Röntgen 1, 20136 Milano –Italy
http://www.igier.unibocconi.it

Robust Mean-Variance Approximations*

Simone Cerreia-Vioglio, Fabio Maccheroni, Massimo Marinacci

Bocconi University and IGIER

December 2022

Abstract

We study mean-variance approximations for a large class of preferences. Compared to the standard mean-variance approximation that only features a risk variability term, a novel index of variability appears. Its neglect in an empirical estimation may result in puzzling inflated risk terms of standard mean-variance approximations.

1 Introduction

Rational preferences form a large class of monotone preferences over monetary acts f that are represented by a decision criterion V that can be decomposed as

$$V(f) = I(u \circ f)$$

where u is a monotone utility function over monetary outcomes and I is a suitable monotone functional, like for example the (Choquet) integral functional in (Choquet) subjective expected utility. In this paper we study the mean-variance approximation for the certainty equivalent functional

$$c(f) = u^{-1}(V(f))$$

We show that, when I and u are suitably differentiable, the functional c admits a unique quadratic approximation of the form

$$c(w+h) \approx w + \mathbb{E}(h) - \underbrace{\frac{\lambda_u(w)}{2}\mathbb{V}(h)}_{\text{risk factor}} - \underbrace{\frac{u'(w)}{2}\Phi_w(h)}_{\text{ambiguity factor}}$$

*We thank Ales Cerny for insightful comments. In particular, Proposition 19 was developed with him.

where w is the wealth of the decision maker, λ_u is the Arrow-Pratt index of risk aversion and Φ_w is a quadratic index of variability (Proposition 4). This index, which can be interpreted as an ambiguity factor (Propositions 5 and 6), is the novel term relative to the traditional quadratic approximation of the subjective expected utility case. If neglected, this factor would conflate in the risk factor. This is best seen by rewriting the mean-variance approximation as

$$c(w+h) \approx w + \mathbb{E}_{Q_w}(h) - \frac{1}{2}\lambda_u(w)(1+\theta_u^w(h))\mathbb{V}_{Q_w}(h)$$

with an ambiguity wedge θ_u^w given by the ratio

$$\theta_u^w(h) = \frac{u'(w)\Phi_w(h)}{\lambda_u(w)\mathbb{V}_{Q_w}(h)}$$

between the ambiguity and risk factors (provided the denominator is not zero). An empirical analysis that overlooks the role of ambiguity, so this wedge, might well result in a puzzlingly inflated estimated risk factor. This wedge has been originally proposed by Hara (2022) as a measure of ambiguity aversion, with an insightful analysis that complements ours.

We characterize the quadratic index of variability Φ_w for several classes of preferences, in particular for smooth ambiguity and variational preferences (Section 4.3). In particular, we show that the analysis of Maccheroni et al. (2013) fits into our framework. For variational preferences, the quadratic index of variability is especially well behaved because it does not depend on wealth levels and shares several properties of the standard variance (which emerges under probabilistic sophistication, as Proposition 7 shows).

2 Mathematical preliminaries

A probability space (S, Σ, P) underlies our analysis. It consists of a space S endowed with a σ-algebra of subsets Σ, with typical elements E and F, on which a countably additive probability measure P is defined. We denote by L^∞ the space of Σ-measurable and essentially bounded functions, with typical element $\xi : S \to \mathbb{R}$.[1] With a standard abuse of notation, we denote by k both the scalar and the constant function $k1_S$. Unless otherwise specified, we endow L^∞ with the topology induced by the (*essential*) *supnorm* $\|\xi\|_\infty = \operatorname{ess\,sup}|\xi|$.

As well-known, the supnorm dual of L^∞ can be identified with the space *ba* of all bounded finitely additive measures $Q : \Sigma \to \mathbb{R}$ such that, for each $E \in \Sigma$, $P(E) = 0$

[1]As usual, we view and treat the elements of L^∞ as functions, despite them being equivalence classes. Equalities and inequalities between elements of L^∞ hold almost surely with respect to P.

implies $Q(E) = 0$. Thus, P identifies the null sets. We endow the dual space ba with the total variation norm $\|\cdot\|_*$. We denote by $\langle \ , \ \rangle : ba \times L^\infty \to \mathbb{R}$ the duality map

$$\langle Q, \xi \rangle = \int f dq = \mathbb{E}_Q(\xi)$$

An important subset of ba is the space ca of all its countably additive elements. We denote by ba_+ and ca_+ the cones of the positive elements of ba and ca, respectively. In particular, we denote by Δ and Δ^σ the sets of the probability measures of ba_+ and ca_+, respectively. We endow Δ with the weak*-topology.

Let C be an open interval, possibly unbounded, of the real line. Rather than the whole space L^∞, we often consider its open and convex subset

$$L^\infty(C) = \{\xi \in L^\infty : [\operatorname{ess\,inf} \xi, \operatorname{ess\,sup} \xi] \subseteq C\}$$

of the functions in L^∞ with "extended" images contained in the interval C. We say that a functional $T : L^\infty(C) \to \mathbb{R}$ is:

(i) *normalized* if $T(k) = k$ for all $k \in C$;

(ii) *monotone* if

$$\xi \geq \xi' \Longrightarrow T(\xi) \geq T(\xi')$$

for all $\xi, \xi' \in L^\infty(C)$;

(iii) *constant additive* if

$$T(\xi + k) = T(\xi) + k$$

for all $\xi \in L^\infty(C)$ and $k \in C$ with $\xi + k \in L^\infty(C)$.

Clearly, a constant additive functional is normalized. Next we relate a functional with the underlying probability measure P. To this end, given any two maps $\xi, \xi' \in L^\infty$, we write $\xi \overset{d}{\sim} \xi'$ when they have the same distribution under P. We say that a functional $T : L^\infty(C) \to \mathbb{R}$ is:

(iv) *law invariant* if

$$\xi \overset{d}{\sim} \xi' \Longrightarrow T(\xi) = T(\xi)$$

for all $\xi, \xi' \in L^\infty(C)$.

The import of this property is especially strong when the probability measure P is *nonatomic*, i.e., when for each $P(E) > 0$ there exists $F \subseteq E$ such that $0 < P(F) < P(E)$.

A functional $T : L^\infty(C) \to \mathbb{R}$ is *continuous* when it is continuous under the sup-norm, i.e.,

$$\|\xi_n - \xi\|_\infty \to 0 \Longrightarrow T(\xi_n) = T(\xi)$$

for all sequences $\{\xi_n\}$ in $L^\infty(C)$. A stronger notion of continuity is, however, natural in the law invariant case.[2] Denoting by $\rightsquigarrow$ convergence in law, we say that a functional $T : L^\infty(C) \to \mathbb{R}$ is:

(v) (*uniformly*) *continuous in law* if

$$\xi_n \rightsquigarrow \xi \Longrightarrow T(\xi_n) = T(\xi)$$

for all uniformly bounded sequences $\{\xi_n\}$ in $L^\infty(C)$.

When a law invariant functional is continuous in law we say that it is *continuously law invariant.*

Finally, let M be a non-empty subset of Δ^σ, endowed with the smallest σ-algebra $\mathcal{M}$ that makes the map $Q \mapsto \mathbb{E}_Q(\xi)$ measurable for all $\xi \in L^\infty$. A countably additive probability measure $\mu : \mathcal{M} \to [0,1]$ is called *prior.* We thus have a second-order probability space $(M, \mathcal{M}, \mu)$. On it, the space L^∞_M is defined in the obvious way. The prior μ induces a *predictive* distribution $\bar{\mu} \in \Delta^\sigma$ given by $\bar{\mu}(E) = \int Q(E)\, d\mu(Q)$.

Given any $\xi \in L^\infty$, the $\mathcal{M}$-measurable map $Q \mapsto \mathbb{E}_Q(\xi)$ is denoted by

$$\mathbb{E}_{(\cdot)}\xi : M \to \mathbb{R}$$

Since M is $\|\cdot\|_*$-bounded, it holds that $|\mathbb{E}_Q(\xi)| \leq \|\xi\|_\infty$ for all $Q \in M$. The map $\mathbb{E}_{(\cdot)}\xi$ thus belongs to L^∞_M.

3 Decision-theoretic setup

In decision-theoretic terms, we interpret S as a state space, Σ as an event σ-algebra, C as a monetary consequence space and P as a reference probability measure that identifies the null events. Here the elements of $L^\infty(C)$ are denoted by $f : S \to \mathbb{R}$ and called *monetary acts.* They associate to each state $s \in S$ an amount of money $f(s)$.

In our approximation exercise we consider directly a decision criterion $V : L^\infty(C) \to \mathbb{R}$, leaving in the background the preferences over acts that they represent.

Definition 1 *A decision criterion* $V : L^\infty(C) \to \mathbb{R}$ *is* rational *if it can be decomposed as*

$$V(f) = I(u \circ f)$$

[2]This natural status is substantiated in Appendix A.1, where a few notions of continuity are considered.

where the utility function $u : C \to \mathbb{R}$ is strictly increasing and continuous, and the functional $I : L^{\infty}(\operatorname{Im} u) \to \mathbb{R}$ is normalized and monotone.

Most decision criteria under ambiguity have this form, as discussed in Cerreia-Vioglio et al. (2011).[3] This decomposition is mathematically useful and conceptually significant. Indeed, if we enrich the setting with lotteries, so it takes an Anscombe-Aumann form, one can show that u models risk attitudes and I ambiguity attitudes.

Variational preferences are, for example, represented by a rational decision criterion featuring

$$I(\xi) = \min_{Q \in \Delta} \mathbb{E}_Q(\xi) + c(Q) \tag{1}$$

where $c : \Delta \to [0, \infty]$ is a grounded (i.e., its infimum value is zero), lower semicontinuous and convex index of ambiguity aversion. Thus,

$$V(f) = I(u \circ f) = \min_{Q \in \Delta} \mathbb{E}_Q(u \circ f) + c(Q)$$

For our analytic purposes, next we define variational decision criteria through the properties that, as proved in Maccheroni et al. (2006), characterize the functionals I of the form (1).

Definition 2 *A rational decision criterion V is* variational *if its functional I is constant additive and concave.*

The classic *maxmin* decision criterion features a variational functional I given by

$$I(\xi) = \min_{Q \in \mathcal{C}} \mathbb{E}_Q(\xi)$$

where $\mathcal{C}$ is a weak*-compact and convex set of probability measures (so, here the index c is two-valued, equal to 0 on $\mathcal{C}$ and to $+\infty$ otherwise). Formally, the maxmin criterion features a constant additive and superlinear functional I, as proved in the seminal Gilboa and Schmeidler (1989).

The smooth ambiguity model (Klibanoff et al., 2005) is another instance of a rational decision criterion, with

$$I(\xi) = \phi^{-1}\left(\mathbb{E}_{\mu}\left(\phi\left(\mathbb{E}_{(\cdot)}\xi\right)\right)\right)$$

where $\phi : \operatorname{Im} u \to \mathbb{R}$ is strictly increasing and continuous. Thus,

$$V(f) = I(u \circ f) = \phi^{-1}\left(\mathbb{E}_{\mu}\phi\left(\mathbb{E}_Q(u \circ f)\right)\right)$$

Next we introduce a class of rational decision criteria, due to Machina and Schmeidler (1992).

[3] See also the Omnibus Theorem stated in Cerreia-Vioglio et al. (2022).

Definition 3 *A rational decision criterion V is* probabilistically sophisticated *if its functional I is law invariant.*

In other words, probabilistic sophistication requires acts f and f' which are equally distributed under P to be equally ranked, i.e.,

$$f \overset{d}{\sim} f' \Longrightarrow V(f) = V(f')$$

When the law invariant I is also continuous in law, we say that V is *continuously probabilistically sophisticated.*

The *multiplier* decision criterion of Hansen and Sargent (2001, 2008) is an important instance of a continuously probabilistically sophisticated variational criterion. It features an entropic index $c(\cdot) = \theta R(\cdot||P)$,[4] with $\theta \in (0, \infty)$, and admits a dual formulation

$$I(\xi) = \min_{Q \in \Delta^{\sigma}} \mathbb{E}_Q(\xi) + \theta R(Q||P) = -\theta \log \mathbb{E}_P(e^{-\frac{\xi}{\theta}})$$

that clarifies its continuous probabilistically sophisticated nature.

We close with a notion of comparative ambiguity aversion of Ghirardato and Marinacci (2002). In our analytic setting, it takes the following form.

Definition 4 *Given any two rational decision criteria V_1 and V_2, we say that V_1 is* more ambiguity averse *than V_2 if $u_1 = u_2$ and $I_1 \leq I_2$.*

The requirement that the two criteria share the same utility is a *ceteris paribus* assumption that factors out risk attitudes.[5] Next we report the absolute notion that results from this comparative notion, once an expected utility neutrality benchmark is posited.

Definition 5 *A rational decision criterion is* ambiguity averse *if it is more ambiguity averse than a subjective expected utility criterion.*

Formally, ambiguity aversion amounts to the existence of a probability measure Q such that

$$I(\xi) \leq \mathbb{E}_Q(\xi) \qquad \forall \xi \in L^{\infty}(\operatorname{Im} u) \tag{2}$$

It is easy to see that variational decision criteria are ambiguity averse because of the concavity of I. To allow for non-averse ambiguity attitudes, next we introduce a natural enlargement of the variational class in which we drop the concavity requirement.

Definition 6 *A rational decision criterion V is* invariant *if its functional I is constant additive.*

The invariant biseparable criterion of Ghirardato et al. (2004) is the special case when I is also positively homogeneous.

[4] Here $R(\cdot||\cdot)$ denotes the relative entropy.

[5] As discussed at length in Ghirardato and Marinacci (2002).

4 Analysis

4.1 Measuring variability

A key element of our mean-variance approximation is a measure of variability. Here we study a general notion of this kind. To this end, we say that a functional $\Phi : L^\infty \to \mathbb{R}$ is:

(i) *homogeneous of degree* 2 if $\Phi(\lambda\xi) = \lambda^2 \Phi(\xi)$ for all $\lambda \in \mathbb{R}$ and all $\xi \in L^\infty$;

(ii) *quadratic* if
$$\Phi(\xi + \xi') + \Phi(\xi - \xi') = 2\Phi(\xi) + 2\Phi(\xi') \tag{3}$$
for all $\xi, \xi' \in L^\infty$;

(iii) *constant invariant* if $\Phi(\xi + k) = \Phi(\xi)$ for all $\xi \in L^\infty$ and all $k \in \mathbb{R}$.

Equality (3) is the abstract parallelogram law used by Jordan and von Neumann (1935) to characterize normed vector spaces with inner product structure. Its basic version for scalars
$$(x + y)^2 + (x - y)^2 = 2x^2 + 2y^2$$
characterizes the quadratic function $g(x) = g(1)x^2$ among continuous functions on the real line.[6] This observation justifies the "quadratic" terminology.

The next result, essentially due to Kurepa (1959), sheds light on this notion.

Proposition 1 *A quadratic functional $\Phi : L^\infty \to \mathbb{R}$, continuous at 0, is:*

(i) continuous and homogeneous of degree 2;

(ii) convex if and only if it is positive.[7]

With this, next we introduce an index of variability.

Definition 7 *A functional $\Phi : L^\infty \to \mathbb{R}$ is a* quadratic index of variability *if it is quadratic and continuous at 0, with $\Phi(1) = 0$.*

By Proposition 1, a quadratic index of variability Φ is homogeneous of degree 2. Thus, it holds $\Phi(k) = k^2\Phi(1) = 0$ for all $k \in \mathbb{R}$ and so Φ is null on constant functions.

Among quadratic indexes of variability, the following class is especially important.

[6] Kurepa (1959) p. 57 shows that this characterization actually extends well beyond continuous functions.

[7] Positive quadratic functionals are traditionally called semidefinite positive.

Definition 8 *A quadratic index of variability is a* protovariance *if it is constant invariant.*

By Proposition 1, a protovariance is convex if and only if it is positive. The main instance of a positive protovariance is, of course, the variance. For simplicity, next we characterize this case for the reference probability measure P.

Proposition 2 *Let P be adequate. The following conditions are equivalent:*

(i) $\Phi : L^\infty \to \mathbb{R}$ *is a continuously law invariant quadratic index of variability;*

(ii) there exists $a \in \mathbb{R}$ such that

$$\Phi(\xi) = a\mathbb{V}_P(\xi) \tag{4}$$

for all $\xi \in L^\infty$.

Moreover, Φ is positive if and only if $a \geq 0$.

Up to a multiplicative constant, the variance is thus a quadratic index of variability anchored to a probability measure, here P, via continuous law invariance. Adequacy ensures the existence of an event E with $P(E) = 1/2$. Consider the "white noise" function $\eta \in L^\infty$ which is equal to 1 on E and to -1 otherwise. We can write (4) as

$$\Phi(\xi) = \Phi(\eta)\,\mathbb{V}_P(\xi)$$

In particular, we have $\Phi(\xi) = \mathbb{V}_P(\xi)$ if and only if $\Phi(\eta) = 1$. The variance can thus be seen as a continuously law invariant quadratic index of variability normalized through this white noise.

Variances are the most important instances of protovariances but, of course, there exist "genuine" protovariances that do not reduce to variances. A basic instance are convex combinations

$$\sum_{i=1}^{n} \alpha_i \mathbb{V}_{Q_i} \tag{5}$$

of variances. They are easily seen to be protovariances but, as we show next, in general they are not variances.

Proposition 3 *Let $\{Q_i\}_{i=1}^n$ be a finite collection of $n \geq 2$ distinct countably additive nonatomic probability measures. Given n distinct weights α_i, the convex combination (5) of their variances is a protovariance but not a variance.*

4.2 Mean-variance approximation

In our analysis we denote the elements of C by the letter w, and refer to them as wealth levels, because applications of quadratic approximations often involve these levels.[8]

Definition 9 *A rational decision criterion V is* analytical *if u is twice continuously differentiable, with $u' > 0$, and I is twice continuously Gateaux differentiable.*

For an analytical rational decision criterion, we denote by $\lambda_u : C \to \mathbb{R}$ the Arrow-Pratt index of u given by $\lambda_u = -u''/u'$.

Differentiability is a non-trivial requirement. For instance, in the variational case it requires an (essentially) strictly convex function c, which implies *inter alia* that the maxmin criterion is not Gateaux differentiable, so not analytical.[9]

The *monetary certainty equivalent functional* $c : L^\infty(C) \to C$ is defined by

$$c(f) = u^{-1}(V(f)) = u^{-1}(I(u \circ f)) \tag{6}$$

It assigns to each act its monetary certainty equivalent. This functional is the protagonist of our analysis. Next we establish a general mean-variance approximation for it.

Proposition 4 *Let V be an analytical rational decision criterion. At each $w \in C$, there exists a unique reference probability Q_w and a unique quadratic index of variability Φ_w such that*

$$c(w+h) = w + \mathbb{E}_{Q_w}(h) - \frac{\lambda_u(w)}{2}\mathbb{V}_{Q_w}(h) - \frac{u'(w)}{2}\Phi_w(h) + R(h) \tag{7}$$

where $R(th) = o(t^2)$. If, in addition, P is nonatomic and V is probabilistically sophisticated, then $Q_w = P$ for all $w \in C$.

In this mean-variance approximation, it is natural to interpret the term

$$\frac{1}{2}\lambda_u(w)\mathbb{V}_{Q_w}(h)$$

as a risk factor. This suggests to view the term

$$\frac{1}{2}u'(w)\Phi_w(h)$$

as an ambiguity factor. The next result justifies this interpretation.

Proposition 5 *Let V_1 and V_2 be any two analytical rational decision criteria. If V_1 is more ambiguity averse than V_2, then $Q_w^1 = Q_w^2$ and $\Phi_w^1 \geq \Phi_w^2$ for all $w \in C$.*

[8] In Appendix A.2 we detail the notions of differentiability for functionals that we use.

[9] See Maccheroni et al. (2006) pp. 1465-1467. Cf. also Corollary 1 in Appendix A.2.

A higher ambiguity aversion thus corresponds to a higher Φ_w, which can be regarded as an index of ambiguity aversion within the mean-variance approximation.[10] The next result corroborates this interpretation of Φ_w by showing that it is negative under ambiguity aversion.

Proposition 6 *Let V be an analytical rational decision criterion. If V is ambiguity averse, then $Q_w = Q$ and $\Phi_w \leq 0$ for all $w \in C$.*[11]

Thus, ambiguity aversion renders the reference probability independent of the wealth level and, perhaps more importantly, makes positive the ambiguity factor, i.e.,

$$\frac{1}{2}u'(w)\,\Phi_w(h) \geq 0$$

Under aversion to both risk and ambiguity,[12] we can thus decompose the mean-variance approximation

$$c(w+h) \approx w + \mathbb{E}_Q(h) - \underbrace{\frac{\lambda_u(w)}{2}\mathbb{V}_Q(h)}_{\text{risk factor}} - \underbrace{\frac{u'(w)}{2}\Phi_w(h)}_{\text{ambiguity factor}}$$

in negative risk and ambiguity factors. These factors reinforce each other and are both needed for a proper quantitative account of the negative impact of uncertainty on monetary certainty equivalents. If either is omitted, the remaining factor may need to be overstretched. This is best seen by writing the general approximation (7) as

$$c(w+h) \approx w + \mathbb{E}_{Q_w}(h) - \frac{1}{2}\lambda_u(w)\left(1+\theta_u^w(h)\right)\mathbb{V}_{Q_w}(h)$$

where the map θ_u^w is given by

$$\theta_u^w(h) = \frac{u'(w)\,\Phi_w(h)}{\lambda_u(w)\,\mathbb{V}_{Q_w}(h)}$$

This map represents an *ambiguity wedge* defined as the ratio between the ambiguity and risk factors. If neglected, this wedge would conflate in the risk factor that, when calibrated, may then assume implausible values. This wedge has been originally proposed by Hara (2022) as a measure of ambiguity aversion, with an insightful analysis that extends some classic arguments of Pratt (1964) and Arrow (1971).

Under probabilistic sophistication the reference probability Q_w reduces to P for all wealth levels, as seen in Proposition 4. Next we show that under continuity also

[10]Formally, Φ_w is a second-order differential at w computed along the diagonal, as detailed in the appendix.

[11]Recall that Q is the probability measure in (2), unique because of differentiability, that makes V ambiguity averse.

[12]As usual, risk aversion corresponds to a concave u, with $u'' \leq 0$ and so $\lambda_u \geq 0$ (as $u' > 0$).

the index of variability Φ_w simplifies as it reduces, up to a wealth-dependent multiplicative constant, to the variance under P. Thus, the mean-variance approximation substantially simplifies in the continuous probabilistically sophisticated case.

Proposition 7 *Let V be an analytical rational decision criterion. If P is nonatomic and V is continuously probabilistically sophisticated, then at each $w \in C$ there exists $b_w \in \mathbb{R}$ such that*

$$c(w+h) = w + \mathbb{E}_P(h) - \frac{\lambda_u(w) + b_w u'(w)}{2} \mathbb{V}_P(h) + R(h)$$

where $R(th) = o(t^2)$.

In view of Proposition 6, we have $b_w \geq 0$ under ambiguity aversion.

4.3 Special cases

We now turn to the form that the general approximation (7) takes for some specific decision criteria. The smooth ambiguity model is analytical when u and ϕ are both twice continuously differentiable, with $u', \phi' > 0$. The analysis of Maccheroni et al. (2013) shows that, for each $w \in C$,

$$Q_w = \bar{\mu} \quad \text{and} \quad \Phi_w(h) = \lambda_\phi(u(w)) \mathbb{V}_\mu\left(\mathbb{E}_{(\cdot)} h\right)$$

That is,

$$c(w+h) \approx w + \mathbb{E}_{\bar{\mu}}(h) - \frac{\lambda_u(w)}{2} \mathbb{V}_{\bar{\mu}}(h) - \frac{u'(w)}{2} \lambda_\phi(u(w)) \mathbb{V}_\mu\left(\mathbb{E}_{(\cdot)} h\right) \tag{8}$$

The reference probability is independent of the wealth level w and is given by the predictive distribution $\bar{\mu}$. The index Φ_w, instead, depends on w via the Arrow-Pratt index of the "second-order" utility function ϕ. It has the nature of a proper variance that accounts for the variability of the expectation $\mathbb{E}_{(\cdot)} h$, which is in turn implied by model ambiguity.[13]

Using the function $v = \phi \circ u$, which accounts for attitudes toward model ambiguity,[14] we can rewrite the quadratic approximation (8) as

$$c(w+h) \approx w + \mathbb{E}_{\bar{\mu}}(h) - \frac{\lambda_u(w)}{2} \mathbb{V}_{\bar{\mu}}(h) - \frac{\lambda_v(w) - \lambda_u(w)}{2} \mathbb{V}_\mu\left(\mathbb{E}_{(\cdot)} h\right)$$

This version of the quadratic approximation permits to write the ambiguity wedge as

$$\theta_u^w(h) = \frac{\lambda_v(w) - \lambda_u(w)}{\lambda_u(w)} \frac{\mathbb{V}_\mu\left(\mathbb{E}_{(\cdot)} h\right)}{\mathbb{V}_{\bar{\mu}}(h)} \tag{9}$$

[13] As discussed at length in Maccheroni et al. (2013).

[14] See, e.g., Maccheroni et al. (2013) p. 1082.

The first factor on the r.h.s. is a taste component, the ratio between model ambiguity and risk attitudes, while the second factor is an information component, the ratio between perceived model ambiguity and risk.[15] The wedge is higher when either ratio is higher, that is, when there is either a higher aversion or a higher perception of model ambiguity.

The analysis of Maccheroni et al. (2013) thus fits into our own analysis. Next we give a novel mean-variance approximation for variational and, more generally, for invariant decision criteria.

Proposition 8 *Let V be an analytical rational decision criterion. If V is invariant, then there exists a unique probability Q and a unique protovariance Φ such that, for each $w \in C$,*

$$c(w+h) = w + \mathbb{E}_Q(h) - \frac{\lambda_u(w)}{2}\mathbb{V}_Q(h) - \frac{u'(w)}{2}\Phi(h) + R(h) \tag{10}$$

where $R(th) = o(t^2)$. If, in addition, V is variational, then Φ is positive (so, convex).

In the invariant case the general mean-variance approximation (7) takes a more global flavor with both the reference probability and the index of variability uniquely pinned down by the decision criterion. Here the ambiguity wedge is

$$\theta_u^w(h) = \frac{u'(w)\,\Phi(h)}{\lambda_u(w)\,\mathbb{V}_Q(h)}$$

Inspired by (10), we can define a robust mean-variance decision criterion $c : L^\infty(C) \to \mathbb{R}$ by

$$c(f) = \mathbb{E}_Q(f) - \frac{\lambda}{2}\mathbb{V}_Q(f) - \frac{\rho}{2}\Phi(f) \tag{11}$$

where $\lambda, \rho > 0$ and Φ is a protovariance. Interestingly, this decision criterion is invariant and, when Φ is positive, it is variational. Thus, it inherits some key behavioral features of a rational decision criterion that provides a theoretical underpinning for it as a quadratic approximation. This robust criterion can be used to explore, for instance, portfolio allocation problems.

A Mathematical analysis

Throughout this appendix O denotes an open and convex subset of L^∞, with generic element ξ and generic constant element k. A main example for our purposes is, of course, the set $L^\infty(C)$. The notions earlier defined (Section 2) for functionals T defined on $L^\infty(C)$ immediately extend to functionals defined on a generic set O.

[15] Under Savage's reduction of uncertainty to risk (cf. Marinacci, 2015).

A.1 Continuity

Denoting by $\xrightarrow{P}$ convergence in probability, we say that a functional $g : O \to \mathbb{R}$ is:

(i) *Lebesgue continuous* if

$$\xi_n \to \xi \quad P\text{-a.e.} \Longrightarrow g(\xi_n) \to g(\xi)$$

for all uniformly bounded sequences $\{\xi_n\}$ in O.[16]

(ii) (*uniformly*) *continuous in probability* if

$$\xi_n \xrightarrow{P} \xi \Longrightarrow g(\xi_n) \to g(\xi)$$

for all uniformly bounded sequences $\{\xi_n\}$ in O.

Lebesgue continuity is stronger than supnorm continuity and is equivalent to the apparently stronger notion of continuity in probability.

Proposition 9 *A functional $g : O \to \mathbb{R}$ is Lebesgue continuous if and only if it is continuous in probability.*

Proof We prove the "if" as the converse trivially holds because almost sure convergence implies convergence in probability (see, e.g., Billingsley, 1995, p. 330). So, assume *per contra* that g is not probabilistically continuous. Then, there exists a uniformly bounded sequence $\{\xi_n\} \subseteq O$ such that $\xi_n \xrightarrow{P} \xi \in O$ but $\lim_n g(\xi_n) \neq g(\xi)$. Thus, there exists $\bar{\varepsilon} > 0$ and a uniformly bounded subsequence $\{\xi_{n_k}\} \subseteq \{\xi_n\}$ such that, for each k,

$$|g(\xi_{n_k}) - g(\xi)| \geq \bar{\varepsilon} > 0 \tag{12}$$

Since $\xi_{n_k} \xrightarrow{P} \xi$, there exists a uniformly bounded subsubsequence $\left\{\xi_{n_{k_l}}\right\} \subseteq \{\xi_{n_k}\}$ that converges almost surely to ξ with $\left\|\xi_{n_{k_l}}\right\|_\infty \leq M$ for all l. Since g is Lebesgue continuous, $\lim_l \xi_{n_{k_i}} = g(\xi)$, which contradicts (12). ■

In general, continuity in law is stronger than Lebesgue continuity. Remarkably, in the law invariant case they become equivalent when P is nonatomic.

Proposition 10 *Let P be nonatomic. A law invariant $g : L^\infty(C) \to \mathbb{R}$ is continuous in law if and only if it is Lebesgue continuous.*

[16] This form of continuity, named after Lebesgue's Dominated Convergence Theorem, is with respect to (uniform) bounded pointwise convergence. For properties of this notion, see Jouini et al. (2006) and Cerreia-Vioglio et al. (2011).

Proof We prove the "if" as the converse trivially holds because almost sure convergence implies convergence in law (see, e.g., Billingsley, 1995, p. 330). So, assume that g is Lebesgue continuous. Let $\{\xi_n\}_{n\geq 1}$ be a uniformly bounded sequences in $L^\infty(C)$ that converges in law to $\xi_0 \in L^\infty(C)$. We want to show that $\lim_n g(\xi_n) = g(\xi_0)$. Set $p_n = P \circ \xi_n^{-1}$ for each $n \geq 0$. By definition of convergence in law, p_n weakly converges to p_0. Let $((0,1), \mathcal{B}, \lambda)$ be the measurable space consisting of the open unit interval with its Borel σ-algebra $\mathcal{B}$ and Lebesgue measure λ. By a classic result of Skorokhod (1956), there exists a sequence $\{\chi_n\}$ of Borel measurable functions $\chi_n : (0,1) \to \mathbb{R}$ that converges λ-a.s. to a Borel measurable function $\chi_0 : (0,1) \to \mathbb{R}$, with $p_n = \lambda \circ \chi_n^{-1}$ for each $n \geq 0$ (cf. Billingsley, 1995, p. 333). As P is nonatomic, there exists a Σ-measurable function $\varphi : S \to (0,1)$ such that $P \circ \varphi^{-1} = \lambda$. Set $\zeta_n = \chi_n \circ \varphi : S \to \mathbb{R}$ for each $n \geq 0$. The function ζ_n is Σ-measurable for each $n \geq 0$, with

$$P \circ \zeta_n^{-1} = P \circ \left(\varphi^{-1} \circ \chi_n^{-1}\right) = \left(P \circ \varphi^{-1}\right) \circ \chi_n^{-1} = \lambda \circ \chi_n^{-1} = p_n = P \circ \xi_n^{-1}$$

for each $n \geq 0$. Thus, $\zeta_n \overset{d}{\sim} \xi_n$ for each $n \geq 0$, with

$$[\operatorname{ess\,inf} \zeta_n, \operatorname{ess\,sup} \zeta_n] = [\operatorname{ess\,inf} \xi_n, \operatorname{ess\,sup} \xi_n] \subseteq C$$

So, $\{\zeta_n\}_{n\geq 0} \subseteq L^\infty(C)$. Since there exists $M > 0$ such that $p_n([-M, M]) = 1$ for all $n \geq 0$, we also have $(P \circ \zeta_n^{-1})([-M, M]) = 1$ for all $n \geq 0$. Thus, the sequence $\{\zeta_n\}$ is uniformly bounded in $L^\infty(C)$. Moreover, $\{\zeta_n\}$ converges P-a.s. to ζ_0 because $\{\chi_n\}$ converges λ-a.s. to χ_0. By the Lebesgue continuity and law invariance of g, we have

$$\lim_n g(\xi_n) = \lim_n g(\zeta_n) = g(\zeta_0) = g(\xi_0)$$

as desired. ■

A.2 Gateaux Differentiability

Let E be a normed vector space with norm $\|\cdot\|$. An operator $T : O \to E$ is *Gateaux differentiable* at a point $\xi \in O$ if there exists a continuous linear operator $\ell_\xi : L^\infty \to E$ such that, for each direction $h \in L^\infty$,

$$\lim_{t\to 0} \frac{\|T(\xi + th) - T(\xi) - \ell_\xi(th)\|}{t} = 0 \tag{13}$$

If it exists, ℓ_ξ is unique; it is called the *Gateaux differential* of T at ξ and denoted by $dT(\xi)$. Thus, for each direction $h \in L^\infty$,

$$dT(\xi)(h) = \lim_{t\to 0} \frac{T(\xi + th) - T(\xi)}{t}$$

We say that T is *Gateaux differentiable* if it is Gateaux differentiable at each point ξ in O.

When T is Gateaux differentiable, we denote by[17]

$$dT : O \times L^\infty \to L\left(L^\infty, E\right)$$

the *differential map* that, to each point $\xi \in O$ and direction $h \in L^\infty$, associates the Gateaux differential $dT\left(\xi\right)\left(h\right)$ of T at the point ξ in the direction h. We say that T is *continuously Gateaux differentiable* when this map is jointly continuous.

We denote T by g when it is a functional, that is, when $E = \mathbb{R}$. In this case the differential map

$$dg : O \times L^\infty \to \left(L^\infty\right)^*$$

associates the Gateaux differential $dg\left(\xi\right)\left(h\right)$ of T at the point ξ in the direction h. In particular, the Gateaux differential $dg\left(\xi\right) : L^\infty \to \mathbb{R}$ is an element of the supnorm dual $\left(L^\infty\right)^*$, so there exists a unique element in ba, denoted by $\nabla g\left(\xi\right)$ and called *Gateaux derivative*, that represents it, i.e.,

$$dg\left(\xi\right)\left(h\right) = \left\langle \nabla g\left(\xi\right), h\right\rangle \qquad \forall h \in L^\infty$$

When g is Gateaux differentiable, we thus have a well-defined map $\nabla g : O \to ba$. This map is $\left\|\cdot\right\|_\infty$-$\left\|\cdot\right\|_*$ continuous when g is continuously Gateaux differentiable.

Proposition 11 *Let $g : O \to \mathbb{R}$ be normalized and monotone. If g is Gateaux differentiable at $k \in O$, then $\nabla g\left(k\right)$ is a probability measure.*

Proof. Let $k \in O$ and set $\ell\left(\cdot\right) = \left\langle \nabla g\left(k\right), \cdot\right\rangle$. Since g is monotone, it follows that $\nabla g\left(k\right) \in ba_+$. Consider $\left\{t_n\right\} \subseteq \mathbb{R}\setminus\left\{0\right\}$ such that $k + t_n \in O$ for all n and $t_n \to 0$. Since g is normalized and Gateaux differentiable at k,

$$\begin{aligned} 0 &= \lim_{n\to\infty} \left| \frac{g\left(k+t_n\right) - g\left(k\right) - \ell\left(t_n\right)}{t_n} \right| = \lim_{n\to\infty} \left| \frac{g\left(k+t_n\right) - g\left(k\right)}{t_n} - \ell\left(1\right) \right| \\ &= \lim_{n\to\infty} \left| \frac{k+t_n-k}{t_n} - \ell\left(1\right) \right| = \lim_{n\to\infty} \left|1 - \ell\left(1\right)\right| = \left|1 - \ell\left(1\right)\right| \end{aligned}$$

Thus, $\nabla g\left(k\right)$ is a probability. ■

We say that a Gateaux differentiable functional $g : O \to \mathbb{R}$ is *twice Gateaux differentiable* at a point $\xi \in O$ if, for each direction $h \in L^\infty$, there exists a continuous bilinear operator $b_\xi : L^\infty \times L^\infty \to \mathbb{R}$ such that, for all directions $h, \eta \in L^\infty$,

$$\lim_{t\to 0} \frac{\left|dg\left(\xi + t\eta\right)\left(h\right) - dg\left(\xi\right)\left(h\right) - b_\xi\left(h, t\eta\right)\right|}{t} = 0$$

[17]We denote by $L\left(L^\infty, E\right)$ the dual space of the continuous linear operators $\ell : L^\infty \to E$. In the special case $E = \mathbb{R}$, we denote it by $\left(L^\infty\right)^*$.

If it exists, b_ξ is unique; it is called the *second-order Gateaux differential* of g at ξ and denoted by $d^2 g(\xi)$. Thus,

$$d^2 g(\xi)(h,\eta) = \lim_{t\to 0} \frac{dg(\xi + t\eta)(h) - dg(\xi)(h)}{t}$$

We can represent $d^2 g(\xi) : L^\infty \times L^\infty \to \mathbb{R}$ as

$$d^2 g(\xi)(h,\eta) = \left\langle \nabla^2 g(\xi)(h), \eta \right\rangle$$

where $\nabla^2 g(\xi) : L^\infty \to ba$ is a linear operator.

We say that g is *twice Gateaux differentiable* if it is Gateaux differentiable at each point ξ in O. In particular, g is *twice continuously Gateaux differentiable* if the second-order differential map

$$d^2 g : O \times L^\infty \times L^\infty \to \mathbb{R}$$

is jointly continuous. In this case, the second-order Gateaux differential $d^2 g(\xi)$ is symmetric, i.e., $d^2 g(\xi)(h,\eta) = d^2 g(\xi)(\eta, h)$ (see Hamilton, 1982, p. 81) and the linear operator $\nabla^2 g(\xi) : L^\infty \to ba$ is continuous, the so-called *Hessian operator*.[18]

Set $\varphi(t) = g(\xi + th)$. When φ is twice differentiable, it holds $\varphi'(0) = dg(\xi)(h)$ and $\varphi''(0) = d^2 g(\xi)(h,h)$. So, for a twice continuously Gateaux differentiable $g : O \to \mathbb{R}$ the following Taylor formula holds, at ξ in O,

$$g(\xi + th) = g(\xi) + \langle \nabla g(\xi), h\rangle t + \frac{1}{2}\left\langle \nabla^2 g(\xi)(h), h\right\rangle t^2 + o\left(t^2\right) \tag{14}$$

that is,

$$g(\xi + h) = g(\xi) + \langle \nabla g(\xi), h\rangle + \frac{1}{2}\left\langle \nabla^2 g(\xi)(h), h\right\rangle + R(h) \tag{15}$$

where $R(th) = o(t^2)$.

Given a twice continuously Gateaux differentiable $g : O \to \mathbb{R}$, for each $\xi \in L^\infty$ we define $\Gamma_\xi : L^\infty \to \mathbb{R}$ by

$$\Gamma_\xi(h) = d^2 g(\xi)(h,h)$$

It is easy to check that Γ_ξ is a quadratic functional. Next we consider its properties on the diagonal – i.e., for all $k \in O$. To this end, for a normalized $g : O \to \mathbb{R}$, let

$$\operatorname{core} g = \left\{\ell \in (L^\infty)^* : \forall \xi \in O, g(\xi) \le \ell(\xi) \text{ and } \ell(1) = 1\right\}$$

be the collection of all continuous and normalized linear functionals that dominate g. When $O = L^\infty$, this set is the superdifferential of g at 0. With this, we say that a normalized $g : O \to \mathbb{R}$ is *balanced* if $\operatorname{core} g \neq \emptyset$.

[18]Though in the paper we focus on Gateaux differentiability, note that twice continuous Gateaux differentiability amounts to twice Frechet differentiability.

Proposition 12 *Let $g : O \to \mathbb{R}$ be balanced.*

(ii) If g is Gateaux differentiable on the diagonal,[19] *then* $\operatorname{core} g = \{\ell\}$ *with $\ell = dg(k)$ for all $k \in O$.*

(iii) If g is twice continuously Gateaux differentiable, then $\Gamma_k(h) = \Gamma_k(h + k')$ for all $k, k', h \in O$.

Point (i) says that $\operatorname{core} g$ is a singleton consisting of the Gateaux differential $dg(k)$ of each $k \in O$. Thus, these Gateaux differentials are all each other equal.

Proof (i) Let $k \in O$ and set $\tilde{\ell}(\cdot) = \langle \nabla g(k), \cdot \rangle$. Let $\ell \in \operatorname{core} g$. Let $h \in L^\infty$ and consider $\{t_n\} \subseteq \mathbb{R}_+ \setminus \{0\}$ such that $k + t_n h \in O$ for all n and $t_n \to 0$. Since g is balanced and Gateaux differentiable at k, we have:

$$\begin{aligned}
0 &= \lim_{n\to\infty} \frac{g(k+t_nh) - g(k) - \tilde{\ell}(t_nh)}{t_n} = \lim_{n\to\infty} \left[\frac{g(k+t_nh) - g(k)}{t_n} - \tilde{\ell}(h)\right] \\
&= \lim_{n\to\infty} \left[\frac{g(k+t_nh) - k}{t_n} - \tilde{\ell}(h)\right] \le \lim_{n\to\infty} \left[\frac{\tilde{\ell}(k+t_nh) - k}{t_n} - \tilde{\ell}(h)\right] \\
&= \lim_{n\to\infty} \left[\frac{k + t_n\tilde{\ell}(h) - k}{t_n} - \tilde{\ell}(h)\right] = \ell(h) - \tilde{\ell}(h)
\end{aligned}$$

Thus, $\tilde{\ell}(h) \le \ell(h)$. Since h was arbitrarily chosen, we have $\tilde{\ell} \le \ell$. Since both functionals are linear, we then have $\tilde{\ell} = \ell$, i.e., $\langle \nabla g(k), \cdot \rangle = \ell(\cdot)$. Since k was arbitrarily chosen, the statement follows.

(ii) Let $k, k', h \in O$. For $m > 0$ large enough, we have $k + 1/m \in O$. Let $\{t_n\}$ be such that $t_n = 1/n\bar{m}$ for all n. Clearly, $k + t_n \in O$ for all n. It follows that

$$\begin{aligned}
0 &= \lim_{n\to\infty} \frac{|\langle \nabla g(k + t_n 1_S), h\rangle - \langle \nabla g(k), h\rangle - \langle \nabla^2 g(k)(h), t_n 1_S\rangle|}{t_n} \\
&= \lim_{n\to\infty} \frac{\left|\langle \nabla g\left(k + \frac{1}{n\bar{m}} 1_S\right), h\rangle - \langle \nabla g(k), h\rangle - \left\langle \nabla^2 g(k)(h), \frac{1}{n\bar{m}} 1_S\right\rangle\right|}{\frac{1}{n\bar{m}}} \\
&= \lim_{n\to\infty} \frac{\left|\langle \nabla g(k), h\rangle - \langle \nabla g(k), h\rangle - \left\langle \nabla^2 g(k)(h), \frac{1}{n\bar{m}} 1_S\right\rangle\right|}{\frac{1}{n\bar{m}}} \\
&= \lim_{n\to\infty} \frac{\left|\left\langle \nabla^2 g(k)(h), \frac{1}{n\bar{m}} 1_S\right\rangle\right|}{\frac{1}{n\bar{m}}} = \lim_{n\to\infty} \frac{\frac{1}{n\bar{m}} |\langle \nabla^2 g(k)(h), 1_S\rangle|}{\frac{1}{n\bar{m}}} = \left|\left\langle \nabla^2 g(k)(h), 1_S\right\rangle\right|
\end{aligned}$$

Thus, $\langle \nabla^2 g(k)(h), 1_S\rangle = 0$. In turn, this implies

$$d^2 g(k)(h, k') = \left\langle \nabla^2 g(k)(h), k'\right\rangle = k' \left\langle \nabla^2 g(k)(h), 1_S\right\rangle = 0$$

[19] That is, at all $k \in O$.

Since $d^2 g(k)$ is a symmetric bilinear map, we then have

$$\begin{aligned}\Gamma_k(h+k') &= d^2 g(k)(h+k', h+k') = d^2 g(k)(h,h) + 2d^2 g(k)(h,k') + d^2 g(k)(k',k') \\ &= d^2 g(k)(h,h) = \Gamma_k(h)\end{aligned}$$

as desired. ■

Corollary 1 *A superlinear and constant additive $g : O \to \mathbb{R}$, with $1 \in O$, is Gateaux differentiable on the diagonal if and only if it is linear.*

Proof We prove the "only if" since the converse trivially holds. As g can be uniquely extended on L^∞, we directly consider $g : L^\infty \to \mathbb{R}$. Its restriction on the collection B_0 of Σ-measurable simple functions can be represented as $g(\xi) = \min_{Q \in \mathcal{C}} \mathbb{E}_Q(\xi)$, where $\mathcal{C}$ is a set of probability measures (see Gilboa and Schmeidler, 1989). Clearly, $\mathcal{C} \subseteq \operatorname{core} I$. By Proposition 12, if I is Gateaux differentiable on the diagonal, then $\operatorname{core} I$, so $\mathcal{C}$, is a singleton $\{\ell\}$. Thus, $g(\xi) = \mathbb{E}_\ell(\xi)$ for all $\xi \in B_0$. Since g is Lipschitz continuous and B_0 is dense in L^∞, we conclude that $g(\xi) = \mathbb{E}_\ell(\xi)$ for all $\xi \in L^\infty$, as desired. ■

Sharper differential properties hold in the constant additive case.

Proposition 13 *Let $g : O \to \mathbb{R}$ be constant additive.*

(i) If g is Gateaux differentiable, then $dg(k) = dg(k')$ for all $k, k' \in O$.

(ii) If g is twice continuously Gateaux differentiable, then

(a) $d^2 g(k) = d^2 g(k')$ for all $k, k' \in O$.

(b) $\Gamma_k(h) = \Gamma_{k+k''}(h+k')$ for all $k, k', k'', h \in O$.

Proof Fix $\tilde{k} \in C$ and set $\tilde{O} = O - \tilde{k} = \left\{\xi - \tilde{k} : \xi \in O\right\}$.[20] Define $\tilde{g} : \tilde{O} \to \mathbb{R}$ by $\tilde{g}(\xi) = g(\xi + \tilde{k})$ for all $\xi \in \tilde{O}$. Clearly, $\tilde{g}$ is constant additive and inherits the differentiability properties of g.

(i) Let $k \in O$. Since $k - \tilde{k} \in \tilde{O}$, we have:

$$\begin{aligned} dg(k)(h) &= \lim_{t \to 0} \frac{g(k+th) - g(k)}{t} = \lim_{t \to 0} \frac{\tilde{g}(k+th-\tilde{k}) - \tilde{g}(k-\tilde{k})}{t} \\ &= \lim_{t \to 0} \frac{\tilde{g}(th) + k - \tilde{k} - (k - \tilde{k})}{t} = \lim_{t \to 0} \frac{\tilde{g}(th)}{t} = d\tilde{g}(0)(h) \end{aligned}$$

We conclude that $dg(k) = d\tilde{g}(0)$ for all $k \in O$.

[20] If $0 \in O$, we can just take $\tilde{k} = 0$.

(ii) It holds (see Hamilton, 1982, p. 81),

$$
\begin{aligned}
d^2 g(k)(h,\eta) &= \lim_{t,\tau\to 0} \frac{g(k+th+\tau\eta) - g(k+th) - g(k+\tau\eta) + g(k)}{t\tau} \\
&= \lim_{t,\tau\to 0} \frac{\tilde{g}(k+th+\tau\eta-\tilde{k}) - \tilde{g}(k+th-\tilde{k}) - \tilde{g}(k+\tau\eta-\tilde{k}) + \tilde{g}(k-\tilde{k})}{t\tau} \\
&= \lim_{t,\tau\to 0} \frac{\tilde{g}(th+\tau\eta) + (k-\tilde{k}) - \tilde{g}(th) - (k-\tilde{k}) - \tilde{g}(\tau\eta) - (k-\tilde{k}) + (k-\tilde{k})}{t\tau} \\
&= \lim_{t,\tau\to 0} \frac{\tilde{g}(th+\tau\eta) - \tilde{g}(th) - \tilde{g}(\tau\eta)}{t\tau} = d^2\tilde{g}(0)(h)(\eta)
\end{aligned}
$$

We conclude that $d^2 g(k) = d^2 \tilde{g}(0)$ for all $k \in O$. Moreover, for each $k, k', k'', h \in L^\infty$,

$$
\begin{aligned}
d^2 g(k+k'')(h+k', \eta+k') &= d^2\tilde{g}(0)(h+k', \eta+k') \\
&= \lim_{t,\tau\to 0} \frac{\tilde{g}(t(h+k') + \tau(\eta+k')) - \tilde{g}(t(h+k')) - \tilde{g}(\tau(\eta+k'))}{t\tau} \\
&= \lim_{t,\tau\to 0} \frac{\tilde{g}(th+\tau\eta) + (t+\tau)k' - \tilde{g}(th) - \tilde{g}(\tau\eta) - (t+\tau)k'}{t\tau} \\
&= \lim_{t,\tau\to 0} \frac{\tilde{g}(th+\tau\eta) - \tilde{g}(th) - \tilde{g}(\tau\nu)}{t\tau} = d^2\tilde{g}(0)(h,\eta)
\end{aligned}
$$

We conclude that $\Gamma_k(h) = \Gamma_{k+k''}(h+k')$ for all $k, k', k'', h \in O$. ■

Next we turn to comparative results.

Lemma 1 *Let $g_1, g_2 : O \to \mathbb{R}$ be monotone, normalized and Gateaux differentiable. If $g_1 \geq g_2$, then $\nabla g_1(k) = \nabla g_2(k)$ for all $k \in O$.*

Proof Let $k \in O$. By Proposition 11, $\nabla g_1(k)$ and $\nabla g_2(k)$ are two probability measures. Then, for each $h \in L^\infty$,

$$
\langle \nabla g_1(k), h\rangle = \lim_{t\to 0} \frac{g_1(k+th) - g_1(k)}{t} \leq \lim_{t\to 0} \frac{g_2(k+th) - g_2(k)}{t} = \langle \nabla g_2(k), h\rangle
$$

because $g_1(k) = g_2(k) = k$. By the Fundamental Theorem of Duality (see, e.g., Aliprantis and Border, 2006, p. 212), it follows that the two probability measures $\nabla g_1(k)$ and $\nabla g_2(k)$ are equal. ■

Proposition 14 *Let $g_1, g_2 : O \to \mathbb{R}$ be monotone, normalized and twice continuously Gateaux differentiable. If $g_1 \geq g_2$ then $\Gamma_k^1 \leq \Gamma_k^2$ for all $k \in O$.*

Proof Let $k \in O$. Since $g_1 \geq g_2$, by Lemma 1 we have $dg_1(k) = dg_2(k)$. Denote by ℓ this common value. By (14), for $i = 1, 2$ we have

$$
g_i(k+th) = k + \ell(h)t + \frac{1}{2}\left\langle \nabla^2 g_i(k)(h), h\right\rangle t^2 + o\left(t^2\right)
$$

for all $h \in L^{\infty}$ and t small enough. Thus,

$$\left\langle \nabla^2 g_1(k)(h), h \right\rangle + \frac{o(t^2)}{t^2} \geq \left\langle \nabla^2 g_2(k)(h), h \right\rangle + \frac{o(t^2)}{t^2}$$

for all $h \in L^{\infty}$ and t small enough. As $t \to 0$, we conclude that

$$\Gamma_k^1(h) = \left\langle \nabla^2 g_1(k)(h), h \right\rangle \geq \left\langle \nabla^2 g_2(k)(h), h \right\rangle = \Gamma_k^2(h)$$

for all $h \in L^{\infty}$, as desired. ■

A useful balanced corollary follows.

Corollary 2 *Let $g : O \to \mathbb{R}$ be normalized and balanced. If g is twice continuously Gateaux differentiable, then $\Gamma_k \leq 0$ for all $k \in O$.*

Proof By Proposition 12, $\operatorname{core} g = \{\ell\}$ where $\ell = dg(k)$ for all $k \in O$. By Proposition 11, $\ell(1) = 1$. So, by taking $g_1 = \ell$ and $g_2 = g$ in Proposition 14, we have $d^2 g(k)(h, h) \leq d^2 \ell(k)(h, h) = 0$ for all $h \in L^{\infty}$. ■

Next we consider the law invariant case.

Lemma 2 *Let $g : O \to \mathbb{R}$ be law invariant.*

(i) If g is Gateaux differentiable at $k \in O$, then $dg(k)$ is law invariant.

(ii) If $g : O \to \mathbb{R}$ is twice continuously Gateaux differentiable at $k \in O$, then Γ_k is law invariant.

Proof Let g be Gateaux differentiable at $k \in O$. Let $h, h' \in L^{\infty}$ with $h \overset{d}{\sim} h'$. Assume that h and h' are both nonzero, otherwise the result is trivially true. Clearly, $k + th \overset{d}{\sim} k + th'$ for each $t \in \mathbb{R}$. As g is law invariant, we have $g(k + th') = g(k + th')$ for each $t \in \mathbb{R}$. In turn, this implies

$$h \overset{d}{\sim} h' \implies dg(k)(h) = dg(k)(h') \tag{16}$$

for all $h, h' \in L^{\infty}$. This proves that $dg(k)$ is law invariant. Similarly, when $g : O \to \mathbb{R}$ is twice continuously Gateaux differentiable at $k \in O$, it holds (see Hamilton, 1982, p. 81),

$$\begin{aligned} d^2 g(k)(h, h) &= \lim_{t,\tau \to 0} \frac{g(k + (t + \tau) h) - g(k + th) - g(k + \tau h) + g(k)}{t\tau} \\ &= \lim_{t,\tau \to 0} \frac{g(k + (t + \tau) h') - g(k + th') - g(k + \tau h') + g(k)}{t\tau} = d^2 g(k)(h', h') \end{aligned}$$

thus proving that

$$h \overset{d}{\sim} h' \implies d^2 g(k)(h, h) = d^2 g(k)(h', h') \tag{17}$$

for all $h, h' \in L^{\infty}$. Thus, Γ_k is law invariant. ■

Proposition 15 *Let P be nonatomic. If a normalized, monotone and law invariant $g : O \to \mathbb{R}$ is Gateaux differentiable at $k \in O$, then $\nabla g(k) = P$.*

Proof By Proposition 11, $\nabla g(k)$ is a probability measure. By (16), for all $E, F \in \Sigma$ it holds

$$P(E) = P(F) \implies \nabla g(k)(E) = dg(k)(F)$$

Since P is nonatomic, it holds $P = \nabla g(k)$.[21] ■

We close with a continuity result.

Proposition 16 *Let $g : O \to \mathbb{R}$ be a Lebesgue continuous and twice continuously Gateaux differentiable and $\xi \in O$. Then, $\Gamma_\xi : L^\infty \to \mathbb{R}$ is continuous if and only if it is Lebesgue continuous.*

Proof A Lebesgue continuous Γ_ξ is easily seen to be continuous. Conversely, let Γ_ξ be continuous. There exists $\varepsilon > 0$ such that $B_\varepsilon(\xi) \subseteq O$. Set $\psi_t(h) = g(\xi + 2th) - 2g(\xi + th) + g(\xi)$ for all $(t, h) \in \mathbb{R} \times L^\infty$ such that $|t| \|h\|_\infty < \varepsilon$. As twice continuous Gateaux differentiability implies twice Frechet differentiability, there exists[22] $0 < \delta_\varepsilon \leq \varepsilon$ such that, for each $h \in L^\infty$ and $t \in \mathbb{R}$,

$$|\Gamma_\xi(h) - \psi_t(h)| \leq 3\varepsilon t^2 \|h\|_\infty^2 \qquad \forall\, |t| \|h\|_\infty < \delta_\varepsilon \tag{18}$$

Let $\{h_n\} \subseteq L^\infty$ be a uniformly bounded sequence that converges P-a.e. to $h \in L^\infty$. Then, there exists $M > 0$ such that $\|h_n\|_\infty \leq M$ for all n. Thus, $\|h\|_\infty \leq M$. By hypothesis, $\lim_n g_t(h_n) = g_t(h)$. Thus, $\lim_n \psi_t(h_n) = \psi_t(h)$. Take

$$|t| < \frac{\delta_\varepsilon}{M}$$

so that $|t| \|h_n\|_\infty < \delta_\varepsilon$ for all $n \geq 0$. By (18), we thus have:

$$\begin{aligned}
|\Gamma_\xi(h) - \Gamma_\xi(h_n)| &= |\Gamma_\xi(h) - \psi_t(h) + \psi_t(h) - \psi_t(h_n) + \psi_t(h_n) - \Gamma_\xi(h_n)| \\
&\leq |\Gamma_\xi(h) - \psi_t(h)| + |\psi_t(h) - \psi_t(h_n)| + |\psi_t(h_n) - \Gamma_\xi(h_n)| \\
&\leq 3\varepsilon t^2 \|h\|_\infty^2 + |\psi_t(h) - \psi_t(h_n)| + 3\varepsilon t^2 \|h_n\|_\infty^2 \\
&\leq 6\varepsilon t^2 M^2 + |\psi_t(h) - \psi_t(h_n)|
\end{aligned}$$

So,

$$\lim_n |\Gamma_\xi(h) - \Gamma_\xi(h_n)| \leq 6\varepsilon t^2 M^2 \qquad \forall\, |t| < \frac{\delta_\varepsilon}{M}$$

We conclude that $\lim_n |\Gamma_\xi(h) - \Gamma_\xi(h_n)| = 0$, as desired. ■

[21] See, e.g., Marinacci (2000).

[22] See, e.g., eq. (3.3) of Ambrosetti and Prodi (1993).

A.3 Adding a utility function

Throughout this section, we consider a continuous utility function $u : C \to \mathbb{R}$ defined on open interval of the real line.

Lemma 3 *The operator $v : L^{\infty}(C) \to L^{\infty}$ defined by $v(\xi) = u \circ \xi$ for all $\xi \in L^{\infty}(C)$ is well defined and continuous.*

(i) If u is continuously differentiable, then v is continuously Gateaux differentiable with

$$dv(\xi)(h) = u'(\xi)h$$

at each $\xi \in L^{\infty}(C)$.

(ii) If u is twice continuously differentiable, then v is twice continuously Gateaux differentiable with

$$d^2v(\xi)(h)(\eta) = u''(\xi)h\eta$$

at each $\xi \in L^{\infty}(C)$.

Proof Let $\xi \in L^{\infty}(C)$. Since $[\operatorname{ess\,inf}\xi, \operatorname{ess\,sup}\xi] \subseteq C$ and u is continuous, it is easy to see that $\|u(\xi)\|_{\infty} < \infty$, so that $u \circ \xi \in L^{\infty}$. It follows that v is well defined. Clearly, v is continuous.[23]

(i) Assume that u is continuously differentiable. Define the identity operator $\iota : L^{\infty}(C) \to L^{\infty}(C)$ by $\iota(\xi) = \xi$ for all $\xi \in L^{\infty}(C)$. Clearly, ι is continuously Gateaux differentiable, with $d\iota(\xi)(h) = h$ for all $\xi \in L^{\infty}(C)$. By the chain rule (see, e.g., Hamilton, 1982, p. 78), the composition $v = u \circ \iota$ is continuously Gateaux differentiable, with, at each $\xi \in L^{\infty}(C)$,

$$dv(\xi)(h) = du(\iota(\xi))(d\iota(\xi)(h)) = u'(\xi)h \qquad \forall h \in L^{\infty}$$

as desired.

(ii) Assume that u twice continuously differentiable. The identity operator ι is also twice continuously Gateaux differentiable with $d^2\iota(\xi) = 0$ at each $\xi \in L^{\infty}(C)$. By the chain rule (see, e.g., Hamilton, 1982, p. 81), the composition $v = u \circ \iota$ is twice continuously Gateaux differentiable, with, at each $\xi \in L^{\infty}(C)$,

$$\begin{aligned} d^2v(\xi)(h,\eta) &= d^2u(\iota(\xi))(d\iota(\xi)h, d\iota(\xi)\eta) + du(\iota(\xi))\left(d^2\iota(\xi)(h,\eta)\right) \\ &= u''(\xi)h\eta + du(\iota(\xi))(0) = u''(\xi)h\eta + u'(\xi)0 = u''(\xi)h\eta \end{aligned}$$

[23]The map v associates to each $\xi \in L^{\infty}(C)$ the equivalence class in L^{∞} obtained by composing u with a representative element of ξ whose range is fully contained in C. Since u is continuous, the choice of the representative element in ξ is irrelevant.

as desired. ■

A strictly increasing $u : C \to \mathbb{R}$ has a nonempty open interval as its image $\operatorname{Im} u$, so the set $L^\infty(\operatorname{Im} u)$ is open and convex.

Proposition 17 *Let $u : C \to \mathbb{R}$ be strictly increasing.*

(i) If $g : L^\infty(\operatorname{Im} u) \to \mathbb{R}$ is continuously Gateaux differentiable and u is continuously differentiable, then $\psi = g \circ v$, then ψ is continuously Gateaux differentiable, with

$$d\psi(\xi)(h) = \langle \nabla g(u(\xi)), u'(\xi) h \rangle$$

at each $\xi \in L^\infty(C)$.

(ii) If $g : L^\infty(\operatorname{Im} u) \to \mathbb{R}$ is twice continuously Gateaux differentiable and u is twice continuously differentiable, then $\psi = g \circ v$ is twice continuously Gateaux differentiable, with

$$d^2\psi(\xi)(h, \eta) = \left\langle \nabla^2 g(v(\xi))(u'(\xi) h), u'(\xi)\eta \right\rangle + \langle \nabla g(v(\xi)), u''(\xi) h\eta \rangle$$

at each $\xi \in L^\infty(C)$.

Proof (i) By Lemma 3-(i), v is continuously Gateaux differentiable. By the chain rule (see, e.g., Hamilton, 1982, p. 78), the composition $\psi = g \circ v$ is continuously Gateaux differentiable, with

$$d\psi(\xi)(h) = dg(v(\xi))(dv(\xi)(h)) = dg(v(\xi))(u'(\xi) h) = \langle \nabla g(v(\xi)), u'(\xi) h \rangle$$

as desired.

(ii) By Lemma 3-(ii), v is twice continuously Gateaux differentiable. By the chain rule (see, e.g., Hamilton, 1982, p. 81), the composition $\psi = g \circ v$ is twice continuously Gateaux differentiable, with

$$\begin{aligned} d^2\psi(\xi)(h, \eta) &= d^2 g(v(\xi))(dv(\xi)(h), dv(\xi)(\eta)) + dg(v(\xi))\left(d^2 v(\xi)(h, \eta)\right) \\ &= d^2 g(v(\xi))(u'(\xi) h, u'(\xi)\eta) + dg(v(\xi))(u''(\xi) h\eta) \\ &= \left\langle \nabla^2 g(v(\xi))(u'(\xi) h), u'(\xi)\eta \right\rangle + \langle \nabla g(v(\xi)), u''(\xi) h\eta \rangle \end{aligned}$$

as desired. ■

We say that a functional $g : L^\infty(\operatorname{Im} u) \to \mathbb{R}$ is *internal* when $\operatorname{Im} g$ is an open interval contained in $\operatorname{Im} u$. For a such functional we can define a certainty equivalent $c : L^\infty(C) \to C$ by

$$c(\xi) = u^{-1}(g(u(\xi)))$$

Note that a normalized and monotone g is internal because $\operatorname{Im} g = \operatorname{Im} u$.

Proposition 18 *Assume that:*

(i) $g : L^{\infty}(\operatorname{Im} u) \to \mathbb{R}$ *is twice continuously Gateaux differentiable;*

(ii) $u : C \to \mathbb{R}$ *is twice continuously differentiable, with* $u' > 0$*;*

(iii) g *is normalized and internal.*

Then, the certainty equivalent $c : L^{\infty}(C) \to C$ *is twice continuously Gateaux differentiable, with, at each* $k \in C$*,*

$$dc(k)(h) = \langle \nabla g(u(k)), h \rangle$$

and

$$\begin{aligned} d^2 c(\xi)(h)(\eta) &= u'(k) \left\langle \nabla^2 g(u(k))(h), \eta \right\rangle \\ &\quad + \frac{u''(k)}{u'(k)} \left(\langle \nabla g(u(k)), h\eta \rangle - \langle \nabla g(u(k)), h \rangle \langle \nabla g(u(k)), \eta \rangle \right) \end{aligned}$$

Proof Set $\varphi = u^{-1}$, so that $u' = 1/\varphi' \circ u$ and $u'' = -\varphi'' \circ u / (\varphi' \circ u)^3$. By the chain rule (see, e.g., Hamilton, 1982, p. 78 and p. 81), the composition $\varphi \circ \psi$ is twice continuously Gateaux differentiable, with

$$\begin{aligned} dc(\xi)(h) &= d(\varphi \circ \psi)(\xi)(h) = d\varphi(\psi(\xi))(d\psi(\xi)(h)) = \varphi'(\psi(\xi)) \langle \nabla g(v(\xi)), u'(\xi) h \rangle \\ &= \varphi'(u(\varphi(\psi(\xi)))) \langle \nabla g(v(\xi)), u'(\xi) h \rangle = \frac{1}{u'(u^{-1}(\psi(\xi)))} \langle \nabla g(v(\xi)), u'(\xi) h \rangle \end{aligned}$$

and

$$\begin{aligned} d^2 c(\xi)(h)(\eta) &= d^2(\varphi \circ \psi)(\xi)(h, \eta) = d^2 \varphi(\psi(\xi))(d\psi(\xi)(h), d\psi(\xi)(\eta)) \\ &\quad + d\varphi(\psi(\xi)) d^2 \psi(\xi)(h, \eta) \end{aligned}$$

For $\xi = k$, we have $v(k) = u(k)$ as well as $\psi(k) = g(v(k)) = u(k)$ because g is normalized. Thus,

$$\begin{aligned} dc(k)(h) &= \frac{1}{u'(u^{-1}(\psi(k)))} \langle \nabla g(u(k)), u'(k) h \rangle \\ &= \frac{u'(k)}{u'(k)} \langle \nabla g(u(k)), h \rangle = \langle \nabla g(u(k)), h \rangle \end{aligned}$$

Moreover,

$$\begin{aligned} d^2 \psi(k)(h, \eta) &= \left\langle \nabla^2 g(u(k))(u'(k) h), u'(k) \eta \right\rangle + \langle \nabla g(u(k)), u''(k) h\eta \rangle \\ &= u'(k)^2 \left\langle \nabla^2 g(u(k))(h), \eta \right\rangle + u''(k) \langle \nabla g(u(k)), h\eta \rangle \end{aligned}$$

and so

$$
\begin{aligned}
d^2c(\xi)(h)(\eta) &= d^2\varphi(\psi(k))(d\psi(k)(h), d\psi(k)(\eta)) + d\varphi(\psi(k))\, d^2\psi(k)(h,\eta) \\
&= \varphi''(u(k)) \langle \nabla g(u(k)), u'(k)h\rangle \langle \nabla g(u(k)), u'(k)\eta\rangle \\
&+ \varphi'(u(k)) \left(u'(k)^2 \left\langle \nabla^2 g(u(k))(h), \eta\right\rangle + u''(k)\langle \nabla g(u(k)), h\eta\rangle\right) \\
&= -u''(k)\varphi'(u(k))^3 u'(k) \langle \nabla g(u(k)), h\rangle \langle \nabla g(u(k)), \eta\rangle \\
&+ \frac{1}{u'(k)} \left(u'(k)^2 \left\langle \nabla^2 g(u(k))(h), \eta\right\rangle + u''(k)\langle \nabla g(u(k)), h\eta\rangle\right) \\
&= -\frac{u''(k)}{u'(k)} \langle \nabla g(u(k)), h\rangle \langle \nabla g(u(k)), \eta\rangle \\
&+ u'(k)\left\langle \nabla^2 g(u(k))(h), \eta\right\rangle + \frac{u''(k)}{u'(k)} \langle \nabla g(u(k)), h\eta\rangle \\
&= u'(k)\left\langle \nabla^2 g(u(k))(h), \eta\right\rangle \\
&+ \frac{u''(k)}{u'(k)} \left(\langle \nabla g(u(k)), h\eta\rangle - \langle \nabla g(u(k)), h\rangle \langle \nabla g(u(k)), \eta\rangle\right)
\end{aligned}
$$

as desired. ■

Define $\Phi_\xi : L^\infty \to \mathbb{R}$ by $\Phi_\xi(h) = -d^2g(\xi)(h,h)$ for all $h \in L^\infty$.

Corollary 3 *Let $k \in O$. Under the hypotheses of Proposition 18, it holds, for each $h \in L^\infty$ with $h + k \in L^\infty(C)$,*

$$
c(k+h) = k + \mathbb{E}_Q(h) - \frac{1}{2}\lambda_u(k)\,\mathbb{V}_Q(h) - \frac{1}{2}u'(k)\,\Phi_k(h) + R(h)
$$

where $Q = \nabla g(u(k)) \in ba$ and $R(th) = o(t^2)$.

Proof It holds $v(k) = u(k)$. Thus, in view of Proposition 18, by (15) we have:

$$
\begin{aligned}
c(k+h) &= c(k) + \langle \nabla g(u(k)), h\rangle \\
&+ \frac{1}{2}\left[u'(k)\left\langle \nabla^2 g(u(k))(h), h\right\rangle + \frac{u''(k)}{u'(k)}\left(\left\langle \nabla g(u(k)), h^2\right\rangle - (\langle \nabla g(u(k)), h\rangle)^2\right)\right] \\
&= k + \mathbb{E}_Q(h) - \frac{1}{2}\lambda_u(k)\left(\mathbb{E}_Q\left(h^2\right) - \mathbb{E}_Q^2(h)\right) + \frac{1}{2}u'(k)\left\langle \nabla^2 g(u(k))(h), h\right\rangle \\
&= k + \mathbb{E}_Q(h) - \frac{1}{2}\lambda_u(k)\,\mathbb{V}_Q(h) - \frac{1}{2}u'(k)\,\Phi_k(h)
\end{aligned}
$$

as desired. ■

Denote by $\Xi : L^\infty \to L_M^\infty$ the map

$$
h \longmapsto \mathbb{E}_{(\cdot)}h
$$

By what observed at the end of Section 2, Ξ is a bounded linear map with $\Xi(1) = 1_M$. Moreover, $\Xi(L^\infty(C)) \subseteq L_M^\infty(C)$. As before, for a given $u : C \to \mathbb{R}$, define $v : L^\infty(C) \to L^\infty$ by $v(\xi) = u \circ \xi$. Given a function $g : L_M^\infty(\operatorname{Im} u) \to \mathbb{R}$, we define $\hat{g} : L^\infty(C) \to \mathbb{R}$ by $\hat{g} = g \circ \Xi \circ v$.

Lemma 4 *Assume that:*

(i) $u : C \to \mathbb{R}$ *is twice continuously differentiable, with* $u' > 0$*;*

(ii) $g : L_M^\infty(\operatorname{Im} u) \to \mathbb{R}$ *is twice continuously Gateaux differentiable;*

(iii) g *is normalized, monotone and law invariant under* μ*.*

Then, the map $\hat{c} : L^\infty(C) \to C$ *defined by* $\hat{c} = u^{-1} \circ \hat{g}$ *is twice continuously Gateaux differentiable at* $k \in C$*, with*

$$\langle \nabla \hat{c}(k), h \rangle = \mathbb{E}_{\bar{\mu}}(h)$$

and

$$\left\langle \nabla^2 \hat{c}(k)(h), h \right\rangle = u'(k) \left\langle \nabla^2 \hat{g}(k)(\Xi(h)), \Xi(h) \right\rangle_M - \lambda_u(k) \mathbb{V}_{\bar{\mu}}(h)$$

Proof As u is strictly increasing and continuous, $\operatorname{Im} u$ is an open and convex set, and so are $L^\infty(\operatorname{Im} u)$ and $L_M^\infty(\operatorname{Im} u)$. Since g is normalized and monotone and u is strictly increasing, $\operatorname{Im} \hat{g}$ is an open interval with $\operatorname{Im} \hat{g} \subseteq \operatorname{Im} u$.

The Gateaux differential of Ξ at each $\xi \in L^\infty$ is Ξ itself. In view of Lemma 3, $\Xi \circ v$ is Gateaux continuously differentiable on $L^\infty(C)$ with

$$d(\Xi \circ v)(\xi)(h) = \langle \cdot, u'(\xi) h \rangle \tag{19}$$

at each $\xi \in L^\infty(C)$. By assumption, g is twice continuously Gateaux differentiable on $L_M^\infty(\operatorname{Im} u)$. By Lemma 3, v is twice continuously Gateaux differentiable. By the chain rule (see, e.g., Hamilton, 1982, p. 78 and p. 81), in view of (19) it follows that $\hat{g}$ is twice continuously Gateaux differentiable, with, at each $\xi \in L^\infty(C)$,

$$\langle \nabla \hat{g}(\xi), h \rangle = \langle \nabla \hat{g}(\xi), \langle \cdot, u'(\xi) h \rangle \rangle_M$$

and

$$\left\langle \nabla^2 \hat{g}(k)(h), h \right\rangle = u'(k)^2 \left\langle \nabla^2 \hat{g}(k)(\Xi(h)), \Xi(h) \right\rangle_M + u''(k) \left\langle \nabla \hat{g}(k), \left\langle \cdot, h^2 \right\rangle \right\rangle_M$$

Define $\varphi = u^{-1}$. The function φ is strictly increasing and twice continuously differentiable, with $u' = 1/\varphi' \circ u$ and $u'' = -\varphi'' \circ u / (\varphi' \circ u)^3$. So, $\hat{c}$ is twice continuously Gateaux differentiable, with

$$\langle \nabla \hat{c}(\xi), h \rangle = \langle \varphi'(\hat{g}(\xi)) \nabla \hat{g}(\xi), h \rangle = \varphi'(\hat{g}(\xi)) \langle \nabla \hat{g}(\xi), \langle \cdot, u'(\xi) h \rangle \rangle_M$$

For $\xi = k$ we have $\hat{g}(k) = u(k)$. In view of Maccheroni et al. (2013), we then have:

$$\begin{aligned}
\langle \nabla \hat{c}(k), h \rangle &= \varphi'(\hat{g}(k)) \langle \nabla \hat{g}(k), \langle \cdot, u'(k) h \rangle \rangle_M = \varphi'(u(k)) u'(k) \langle \nabla \hat{g}(k), \langle \cdot, h \rangle \rangle_M \\
&= \frac{1}{u'(k)} u'(k) \langle \nabla \hat{g}(k), \langle \cdot, h \rangle \rangle_M = \langle \nabla \hat{g}(k), \langle \cdot, h \rangle \rangle_M \overset{*}{=} \langle \nabla g(\Xi(u(k))), \langle \cdot, h \rangle \rangle_M \\
&= \mathbb{E}_\mu \left(\mathbb{E}_{(\cdot)}(h) \right) = \mathbb{E}_{\bar{\mu}}(h)
\end{aligned}$$

as well as

$$
\begin{aligned}
&\left\langle \nabla^2 \hat{c}(k)(h), h \right\rangle \\
&= \varphi''(\hat{g}(k)) \langle \nabla \hat{g}(k), h \rangle^2 + \varphi'(\hat{g}(k)) \left\langle \nabla^2 \hat{g}(k)(h), h \right\rangle \\
&= \varphi''(u(k)) \langle \nabla \hat{g}(k), h \rangle^2 + \varphi'(u(k)) \left\langle \nabla^2 \hat{g}(k)(h), h \right\rangle \\
&= \varphi''(u(k)) u'(k)^2 \langle \nabla \hat{g}(k), \langle \cdot, h \rangle \rangle_M^2 + \varphi'(u(k)) [u'(k)^2 \left\langle \nabla^2 \hat{g}(k)(\Xi(h)), \Xi(h) \right\rangle_M \\
&+ u''(k) \left\langle \nabla \hat{g}(k), \left\langle \cdot, h^2 \right\rangle \right\rangle_M] \\
&= -\frac{u''(k)}{u'(k)^3} u'(k)^2 \langle \nabla \hat{g}(k), \langle \cdot, h \rangle \rangle_M^2 + \frac{1}{u'(k)} u'(k)^2 \left\langle \nabla^2 \hat{g}(k)(\Xi(h)), \Xi(h) \right\rangle_M \\
&+ \frac{u''(k)}{u'(k)} \left\langle \nabla \hat{g}(k), \left\langle \cdot, h^2 \right\rangle \right\rangle_M \\
&= u'(k) \left\langle \nabla^2 \hat{g}(k)(\Xi(h)), \Xi(h) \right\rangle_M + \frac{u''(k)}{u'(k)} (\left\langle \nabla \hat{g}(k), \left\langle \cdot, h^2 \right\rangle \right\rangle_M - \langle \nabla \hat{g}(k), \langle \cdot, h \rangle \rangle_M^2) \\
&\overset{*}{=} u'(k) \left\langle \nabla^2 \hat{g}(k)(\Xi(h)), \Xi(h) \right\rangle_M + \frac{u''(k)}{u'(k)} (\mathbb{E}_\mu \left(\left\langle \cdot, h^2 \right\rangle \right) - (\mathbb{E}_\mu (\langle \cdot, h \rangle))^2) \\
&= u'(k) \left\langle \nabla^2 \hat{g}(k)(\Xi(h)), \Xi(h) \right\rangle_M + \frac{u''(k)}{u'(k)} (\mathbb{E}_{\bar{\mu}} \left(h^2 \right) - \mathbb{E}_{\bar{\mu}}(h)^2) \\
&= u'(k) \left\langle \nabla^2 \hat{g}(k)(\Xi(h)), \Xi(h) \right\rangle_M - \lambda_u(k) \mathbb{V}_{\bar{\mu}}(h)
\end{aligned}
$$

as desired. The equalities $\overset{*}{=}$ hold, in view of Proposition 15, because of the hypothesis (iii). ■

B Proofs in the main text and related material

Proof of Proposition 1 (i) Continuity is in Theorem 2 in Kurepa (1959). As to homogeneity,[24] by (3), we get $2\Phi(0) = 4\Phi(0)$, so $\Phi(0) = 0$. By (3), we also get $\Phi(\xi') + \Phi(-\xi') = 2\Phi(0) + 2\Phi(\xi')$, so $\Phi(\xi') = \Phi(-\xi')$. The functional Φ is thus even. Next, again by (3) we get $\Phi(2\xi) + \Phi(0) = 4\Phi(\xi)$, so $\Phi(2\xi) = 2^2\Phi(\xi)$. By induction, we then have

$$
\Phi(n\xi) = n^2 \Phi(\xi) \qquad \forall \xi \in L^\infty \tag{20}
$$

for all $n \geq 0$. Indeed, assume that this equality holds for all natural numbers $< n$. Then,

$$
\begin{aligned}
\Phi(n\xi) &= \Phi((n-1)\xi + \xi) = 2\Phi((n-1)\xi) + 2\Phi(\xi) - \Phi((n-2)\xi) \\
&= [2(n-1)^2 + 2 - (n-2)^2] \Phi(\xi) = n^2 \Phi(\xi)
\end{aligned}
$$

We conclude that (20) holds. In turn, this implies that $\Phi(\xi) = \Phi(n(\xi/n)) = n^2\Phi(\xi/n)$, so

$$
\Phi\left(\frac{\xi}{n}\right) = \frac{1}{n^2} \Phi(\xi) \qquad \forall \xi \in L^\infty \tag{21}
$$

[24] Here we extend an argument used by Kurepa (1959) p. 58 for functions defined on the real line.

for all $n \geq 1$. By (20) and (21), we then have

$$\Phi\left(\frac{n}{m}\xi\right) = \Phi\left(n\frac{\xi}{m}\right) = n^2\Phi\left(\frac{\xi}{m}\right) = \frac{n^2}{m^2}\Phi(\xi)$$

for all $n \geq 0$ and $m \geq 1$. As Φ is even, we conclude that, for each $\xi \in L^\infty$, it holds $\Phi(q\xi) = q^2\Phi(\xi)$ for all $q \in \mathbb{Q}$. The continuity of Φ then implies the result.

(iii) Let $\Phi \geq 0$. For each $\xi, \xi' \in L^\infty$ it holds:

$$\begin{aligned}\Phi\left(\frac{\xi+\xi'}{2}\right) &= \frac{1}{4}\Phi(\xi+\xi') = \frac{2\Phi(\xi)+2\Phi(\xi')-\Phi(\xi-\xi')}{4} \\ &= \frac{\Phi(\xi)+\Phi(\xi')}{2} - \frac{\Phi(\xi-\xi')}{4}\end{aligned}$$

Since $\Phi \geq 0$, this proves that Φ is midpoint convex. As Φ is continuous by (i), we conclude that Φ is convex. Conversely, let Φ be convex. It holds:

$$\frac{1}{4}\Phi(\xi) = \frac{1}{4}\Phi(2\xi-\xi) = \frac{\Phi(2\xi)+\Phi(\xi)}{2} - \Phi\left(\frac{2\xi+\xi}{2}\right) \geq 0$$

Thus, $\Phi \geq 0$. ■

Note that the arguments used in this last proof hold, more generally, for a continuous Φ defined on a topological vector space. That said, next we present a general version of Proposition 2.

Proposition 19 *Let (S, Σ, P) be adequate.*[25] *The following conditions are equivalent:*

(i) $\Phi : L^\infty \to \mathbb{R}$ *is a Lebesgue continuous and law invariant quadratic functional;*

(ii) there exist $a, b \in \mathbb{R}$ *such that*

$$\Phi(\xi) = a\mathbb{V}_P(\xi) + b\mathbb{E}_P^2(\xi)$$

for all $\xi \in L^\infty$.

In this case a and b are uniquely determined. Moreover,

1. Φ *is positive (i.e., convex) if and only if* $a, b \geq 0$*;*
2. $\Phi(1) = 0$ *if and only if* $b = 0$.

[25] That is, either *nonatomic* or such that Σ is generated by a finite partition of equiprobable events (we call this case *discrete-uniform*).

In reading point 1, recall that Lebesgue continuity implies supnorm continuity. Moreover, observe that this result holds when L^∞ is replaced by L^2 and Lebesgue continuity with norm continuity, providing an axiomatic characterization of variance.

Proof As it is easy to check that (ii) implies (i), we only show that (i) implies (ii). We first prove the discrete-uniform case and then the non-atomic one.

Discrete-uniform case. Let $\Pi = \{E_1, ..., E_k\}$ be a generating partition of Σ such that $P(E_i) = 1/k$ for all $i = 1, ..., k$. Each $\xi \in L^\infty$ can be uniquely written as

$$\xi = \sum_{i=1}^{k} x_i^\xi 1_{E_i} \tag{22}$$

and the map

$$\begin{array}{rccl} \iota & : & L^\infty & \to & \mathbb{R}^k \\ & & \xi & \mapsto & x^\xi \end{array}$$

is an isometric isomorphism of Banach spaces (when $\mathbb{R}^k$ is endowed with the supnorm too). So, its inverse

$$\begin{array}{rcccc} \iota^{-1} & : & \mathbb{R}^k & \to & L^\infty \\ & & x & \mapsto & \sum_{i=1}^k x_i 1_{E_i} \end{array}$$

is an isometric isomorphism too. Set $\eta = \iota^{-1}$. We can thus define $\varphi = \Phi \circ \eta : \mathbb{R}^k \to \mathbb{R}$ and observe that

$$\begin{aligned} \varphi(x + x') + \varphi(x - x') &= \Phi(\eta(x + x')) + \Phi(\eta(x - x')) = \Phi(\eta(x) + \eta(x')) + \Phi(\eta(x) - \eta(x')) \\ &= 2\Phi(\eta(x)) + 2\Phi(\eta(x')) = 2\varphi(x) + 2\varphi(x') \end{aligned}$$

for all $x, x' \in \mathbb{R}^k$. Continuity of φ descends from continuity of Φ and η. So, we have the representation $\varphi(x) = x^\top Ax$ where A is a symmetric matrix of order k (cf. Kurepa, 1959, p. 63). As Φ is law invariant, we have that, for each $x \in \mathbb{R}^k$ and each permutation σ of $\{1, ..., k\}$,

$$\varphi(x_\sigma) = \varphi\left(\left(x_{\sigma(1)}, ..., x_{\sigma(n)}\right)^\top\right) = \Phi\left(\sum_{i=1}^k x_{\sigma(i)} 1_{E_i}\right) = \Phi\left(\sum_{i=1}^k x_i 1_{E_i}\right) = \varphi(x)$$

Then,

$$x^\top Ax = x_\sigma^\top A x_\sigma$$

In particular, denoting by $\{e_1, ..., e_k\}$ the canonical basis of $\mathbb{R}^k$, we have

$$a_{ij} = e_i^\top A e_j = \left((e_i)_\sigma\right)^\top A\,(e_j)_\sigma = \left(e_{\sigma^{-1}(i)}\right)^\top A\left(e_{\sigma^{-1}(j)}\right) = a_{\sigma^{-1}(i)\sigma^{-1}(j)}$$

for all $i, j \in \{1, ..., k\}$.[26] We conclude that $a_{ij} = a_{\tau(i)\tau(j)}$ for all permutations τ of $\{1, ..., k\}$. Consequently, there exist constants c and d such that $a_{ii} = a_{11} = c$ for all i, and $a_{ij} = a_{12} = d$ for all $i \neq j$. We have shown that

$$A = (c - d) I + d\mathbf{1}\mathbf{1}^\top = \frac{a}{k} I + \frac{k^2 d}{k^2} \mathbf{1}\mathbf{1}^\top$$

where $a = k(c - d)$. Thus, for each $\xi \in L^\infty$,

$$\begin{aligned}
\Phi(\xi) &= \Phi\left(\eta\left(\eta^{-1}(\xi)\right)\right) = \Phi(\eta(\iota(\xi))) = \varphi\left(x^\xi\right) = \frac{a}{k}\left[\left(x^\xi\right)^\top I x^\xi\right] + \frac{k^2 d}{k^2}\left[\left(x^\xi\right)^\top \mathbf{1}\mathbf{1}^\top x^\xi\right] \\
&= a\frac{\left(x_1^\xi\right)^2 + ... + \left(x_k^\xi\right)^2}{k} + k^2 d \frac{\left(x_1^\xi + ... + x_k^\xi\right)^2}{k^2} = a\mathbb{E}_P\left(\xi^2\right) + k^2 d \mathbb{E}_P^2(\xi) \\
&= a\mathbb{E}_P\left(\xi^2\right) - a\mathbb{E}_P^2(\xi) + \left(k^2 d + a\right)\mathbb{E}_P^2(\xi) = a\mathbb{V}_P(\xi) + b\mathbb{E}_P^2(\xi)
\end{aligned}$$

where $b = k^2 d + a$.

Take any ξ which is not a.s. constant, but $\mathbb{E}_P(\xi) = 0$. For instance, $1_{B_1} - 1_{B_2}$ with $B_1, B_2 \in \Sigma$ disjoint and $P(B_2) = P(B_1)$. Then, $\mathbb{V}_P(\xi) > 0$ and $\mathbb{E}_P^2(\xi) = 0$. So,

$$a = \frac{\Phi(\xi)}{\mathbb{V}_P(\xi)}$$

while

$$b = \Phi(1_S)$$

These two relations prove the rest of the statement (in both the discrete-uniform case and the uniform one).

Nonatomic case. Recall that for each $\xi \in L^0$, $F_\xi(t) = P(\xi \leq t)$ for all $t \in \mathbb{R}$, is the distribution function of ξ, with pseudo-inverse $F_\xi^{-1}(p) = \inf\{x \in \mathbb{R} : F_\xi(x) \geq p\}$ for all $p \in (0, 1)$. F_ξ is increasing and right-continuous, with $\lim_{t\to-\infty} F_\xi(t) = 0$ and $\lim_{t\to+\infty} F_\xi(t) = 1$. F_ξ^{-1} is increasing (hence Borel measurable) and left-continuous. It is bounded if $\xi \in L^\infty$.

[26] When a vector x is seen as the map

$$\begin{array}{cccc} x & : & \{1, ..., k\} & \to & \mathbb{R} \\ & & j & \mapsto & x(j) = x_j \end{array}$$

then

$$\begin{array}{cccc} x_\sigma & : & \{1, ..., k\} & \to & \mathbb{R} \\ & & j & \mapsto & x_{\sigma(j)} \end{array}$$

is the map $x \circ \sigma$. With this e_i is the map $j \mapsto \delta_i(j)$ and $(e_i)_\sigma = \delta_i \circ \sigma$. For $j = \sigma^{-1}(i)$, $(\delta_i \circ \sigma)(j) = \delta_i(\sigma(j)) = \delta_i(i) = 1$, for $j \neq \sigma^{-1}(i)$, $\sigma(j) \neq i$, and $(\delta_i \circ \sigma)(j) = \delta_i(\sigma(j)) = 0$. That is $(e_i)_\sigma = e_{\sigma^{-1}(i)}$.

Also recall that nonatomicity guarantees that there exists $\upsilon : S \to (0,1)$ in L^∞ with uniform distribution, i.e., $F_\upsilon(t) = t$ for all $t \in (0,1)$. For each $\xi \in L^0$, by defining $\xi_\upsilon = F_\xi^{-1} \circ \upsilon$ we have $\xi \overset{d}{\sim} \xi_\upsilon$.

We start the proof with a key observation. If we show that there exist $a, b \in \mathbb{R}$ such that

$$\Phi(\xi_\upsilon) = a\mathbb{V}_P(\xi_\upsilon) + b\mathbb{E}_P^2(\xi_\upsilon) \tag{23}$$

for all $\xi \in L^\infty$, then law invariance yields

$$\Phi(\xi) = \Phi(\xi_\upsilon) = a\mathbb{V}_P(\xi_\upsilon) + b\mathbb{E}_P^2(\xi_\upsilon) = a\mathbb{V}_P(\xi) + b\mathbb{E}_P^2(\xi)$$

With this, we next proceed to prove (23). To this end, for each $n \in \mathbb{N}$ denote by

$$\Psi_n = \left\{ \left(0, \frac{1}{2^n}\right], \left(\frac{1}{2^n}, \frac{2}{2^n}\right], \ldots, \left(\frac{2^n-1}{2^n}, \frac{2^n}{2^n}\right) \right\}$$

the partition of $(0,1)$ into segments of length 2^{-n}. Note that

$$\Pi_n = \Pi_n^\upsilon = \left\{ \upsilon^{-1}(D_n) : D_n \in \Psi_n \right\}$$

is a partition of S in Σ such that $P(E) = 1/2^n$ for all $E \in \Pi_n^\upsilon$. Set $\Sigma_n = \sigma(\Pi_n) = \upsilon^{-1}(\sigma(\Psi_n))$ for all $n \in \mathbb{N}$. $\{\Sigma_n\}_{n\in\mathbb{N}}$ is a filtration in Σ. As usual, $\Sigma_\infty = \sigma\left(\bigcup_{n\in\mathbb{N}} \Sigma_n\right)$.

It can be shown (see, e.g., Maccheroni et al., 2023) that $\Sigma_\infty = \sigma(\upsilon)$. By the Martingale Convergence Theorem,

$$\mathbb{E}(\xi_\upsilon \mid \Sigma_n) \to \xi_\upsilon \qquad \forall \xi \in L^1$$

where convergence is both almost everywhere and in L^1.

Now consider, for each $n \in N$ the restriction of Φ to $L_n^\infty = L^\infty(S, \Sigma_n, P)$. Note that there exist $a_n, b_n \in \mathbb{R}$ such that

$$\Phi(\zeta) = a_n\mathbb{V}_P(\zeta) + b_n\mathbb{E}_P^2(\zeta) \tag{24}$$

for all $\zeta \in L_n^\infty$. Since both $1_{\upsilon^{-1}((0,1/2])} - 1_{\upsilon^{-1}((1/2,1))}$ and 1_S belong to L_n^∞ for all $n \in \mathbb{N}$, we then have

$$a_n = \frac{\Phi\left(1_{\upsilon^{-1}((0,1/2])} - 1_{\upsilon^{-1}((1/2,1))}\right)}{\mathbb{V}_P\left(1_{\upsilon^{-1}((0,1/2])} - 1_{\upsilon^{-1}((1/2,1))}\right)}$$

and $b_n = \Phi(1_S)$, irrespective of n. Summing up,

$$\Phi(\zeta) = a\mathbb{V}_P(\zeta) + b\mathbb{E}_P^2(\zeta) \tag{25}$$

for all $\zeta \in \bigcup_{n\in\mathbb{N}} L_n^\infty$.

Finally, choose any $\xi \in L^\infty$. As we discussed above, almost surely

$$\underbrace{\mathbb{E}\left(\xi_v \mid \Sigma_n\right)}_{\xi_n} \to \xi_v \qquad \forall \xi \in L^1$$

Moreover, since F_ξ^{-1} is bounded, then ξ_v is bounded and $\|\xi_n\|_\infty \leq \|\xi_v\|_\infty$ for all $n \in \mathbb{N}$. Thus, by the Lebesgue continuity of Φ we have $\lim_{n\to\infty} \Phi\left(\xi_n\right) = \Phi\left(\xi_v\right)$. By the Dominated Convergence Theorem, also first and second moments are Lebesgue continuous. Thus

$$\Phi\left(\xi_v\right) = \lim_{n\to\infty} \Phi\left(\xi_n\right) = \lim_{n\to\infty}\left(a\mathbb{V}_P(\xi_n) + b\mathbb{E}_P^2\left(\xi_n\right)\right) = a\mathbb{V}_P(\xi_v) + b\mathbb{E}_P^2\left(\xi_v\right)$$

as desired. ∎

Proof of Proposition 3 We need to show that $\sum_{i=1}^n \alpha_i \mathbb{V}_{Q_i}$ is not a variance. Suppose, *per contra*, that there is a probability measure $Q_0 : \Sigma \to [0,1]$ such that $\mathbb{V}_{Q_0} = \sum_{i=1}^n \alpha_i \mathbb{V}_{Q_i}$. For $E \in \Sigma$, let $\beta_E = 1_E - 1_{E^c}$. Then, $\mathbb{V}_{Q_i}\left(2^{-1}\beta_E\right) = Q_i\left(E\right)\left(1 - Q_i\left(E\right)\right)$ for each $i = 0, 1, ..., n$. Hence, for each $\alpha \in [0,1]$,

$$Q_i\left(E\right) = \alpha \quad \forall i \geq 1 \Longrightarrow Q_0\left(E\right) = \{\alpha, 1-\alpha\} \tag{26}$$

By the Lyapunov Convexity Theorem, for each n there is a finite partition Π_n in Σ such that $Q_i\left(E\right) = 2^{-n}$ for each $i \geq 1$ and each $E \in \Pi_n$. By (26), we have $Q_0\left(E\right) \in \{2^{-n}, 1-2^{-n}\}$. Let A be any union of 2^{n-1} elements of Π_n. Since $Q_i\left(A\right) = 1/2$ for each $i \geq 1$, by (26) we have $Q_0\left(A\right) = 1/2$. As each $E \in \Pi_n$ is contained in some such set A, we conclude that $Q_0\left(E\right) = 2^{-n}$ for each $E \in \Pi_n$. In turn, this easily implies that Q_0 is strongly continuous.

Again by the Lyapunov Convexity Theorem, there is $F \in \Sigma$ such that $Q_i\left(F\right) = 1/2$ for each $i \geq 1$. By (26), for each $E \in \Sigma$,

$$Q_i\left(E\right) = Q_i\left(F\right) \quad \forall i \geq 1 \Longrightarrow Q_0\left(E\right) = Q_0\left(F\right)$$

By Theorem 20 in Marinacci and Montrucchio (2003), it holds $Q_0 \in \mathrm{co}\{Q_1, ..., Q_n\}$. Hence, Q_0 is countably additive. So, it is convex-ranged. Thus, there exists $H \in \Sigma$ such that $Q_0\left(H\right) = 1/2$. Then,

$$\frac{1}{4} = \mathbb{V}_{Q_0}\left(\frac{1}{2}\beta_H\right) = \sum_{i=1}^n \alpha_i \mathbb{V}_{Q_i}\left(\frac{1}{2}\beta_H\right) = \sum_{i=1}^n \alpha_i Q_i\left(H\right)\left(1 - Q_i\left(H\right)\right) \tag{27}$$

For each $i \geq 0$, we have $Q_i\left(E\right)\left(1 - Q_i\left(E\right)\right) \leq 1/4$ for all $E \in \Sigma$. Thus, (27) implies

$$Q_i\left(H\right)\left(1 - Q_i\left(H\right)\right) = 1/4 \qquad \forall i \geq 1$$

that is, $Q_i(H) = 1/2$ for each $i \geq 1$. In particular, we have, for each $i \geq 1$,

$$Q_0(E) = Q_0(H) \Longrightarrow Q_i(E) = Q_i(H) \qquad \forall E \in \Sigma$$

By Theorem 1 in Marinacci (2000), it holds $Q_0 = Q_i$ for each $i \geq 1$, which contradicts the hypothesis that the measures $\{Q_i\}_{i=1}^n$ are distinct. We conclude that $\sum_{i=1}^n \alpha_i \mathbb{V}_{Q_i}$ is not a variance. ∎

Next we prove a more general, local version, of Proposition 4. Analyticity at w has the obvious meaning.

Lemma 5 *Let V be a rational decision criterion. If V is analytical at $w \in C$, then there exists a unique reference probability Q_w and a unique quadratic index of variability Φ_w such that, for each $w \in C$,*

$$c(w+h) = w + \mathbb{E}_{Q_w}(h) - \frac{\lambda_u(w)}{2}\mathbb{V}_{Q_w}(h) - \frac{u'(w)}{2}\Phi_w(h) + R(h)$$

where $R(th) = o(t^2)$. If, in addition, P is nonatomic and V is probabilistically sophisticated, then $Q_w = P$.

Proof Let $g = I$ and $k = w$. Recall that $\Phi_w : L^\infty \to \mathbb{R}$ is defined by $\Phi_w(h) = -d^2 g(u(w))(h,h)$ for all $h \in L^\infty$. It is easy to see that Φ_w is continuous, homogeneous of degree 2 and quadratic; thus, it is a quadratic index of variability.

By setting $Q_w = \nabla g(u(w)) \in ba$, in view of Corollary 3 we have, for each $h \in L^\infty$ with $w + h \in L^\infty(C)$,

$$c(w+h) = w + \mathbb{E}_{Q_w}(h) - \frac{1}{2}\lambda_u(w)\mathbb{V}_{Q_w}(h) - \frac{1}{2}u'(w)\Phi_w(h) + R(h) \tag{28}$$

where $R(th) = o(t^2)$.

As to uniqueness, assume that $\tilde{Q}$ and $\tilde{\Phi}$ satisfy this quadratic approximation, i.e.,

$$c(w+h) = w + \mathbb{E}_{\tilde{Q}}(h) - \frac{1}{2}\lambda_u(w)\mathbb{V}_R(h) - \frac{1}{2}u'(w)\tilde{\Phi}(h) + \tilde{R}(h)$$

where $\tilde{R}(th) = o(t^2)$. For each $h \in L^\infty$ with $w + h \in L^\infty(C)$ define

$$\Psi_w(h) = -\frac{1}{2}\lambda_u(w)\mathbb{V}_{Q_w}(h) - \frac{1}{2}u'(w)\Phi_w(h) \quad ; \quad \tilde{\Psi}(h) = -\frac{1}{2}\lambda_u(w)\mathbb{V}_{\tilde{Q}}(h) - \frac{1}{2}u'(w)\tilde{\Phi}(h)$$

Let $0 \neq h \in L^\infty$. Consider $\{t_n\} \subseteq \mathbb{R}\backslash\{0\}$ with $w + t_n h \in L^\infty(C)$ for all n and $t_n \to 0$. For each n, we have:

$$0 = |c(w+t_nh) - c(w+t_nh)| = \Big|\mathbb{E}_{Q_w}(t_nh) - \frac{1}{2}\lambda_u(w)\mathbb{V}_{Q_w}(t_nh) - \frac{1}{2}u'(w)\Phi_w(t_nh) + R(t_nh)$$
$$-\mathbb{E}_R(t_nh) + \frac{1}{2}\lambda_u(w)\mathbb{V}_{\tilde{Q}}(t_nh) + \frac{1}{2}u'(w)\tilde{\Phi}(t_nh) - \tilde{R}(t_nh)\Big|$$

that is, for each n,

$$0 = \left| \mathbb{E}_{Q_w}(t_n h) - \mathbb{E}_{\tilde{Q}}(t_n h) + \Psi_w(t_n h) - \tilde{\Psi}(t_n h) + R(t_n h) - \tilde{R}(t_n h) \right| \tag{29}$$

Thus,

$$\begin{aligned}
\left| \mathbb{E}_{Q_w}(h) - \mathbb{E}_{\tilde{Q}}(h) \right| &= \left| \frac{\mathbb{E}_{Q_w}(t_n h) - \mathbb{E}_{\tilde{Q}}(t_n h)}{t_n} \right| \\
&\leq \left| \frac{\mathbb{E}_{Q_w}(t_n h) - \mathbb{E}_{\tilde{Q}}(t_n h)}{t_n} + \frac{\Psi_w(t_n h) - \tilde{\Psi}(t_n h)}{t_n} + \frac{R(t_n h) - \tilde{R}(t_n h)}{t_n} \right| \\
&+ \left| - \frac{\Psi_w(t_n h) - \tilde{\Psi}(t_n h)}{t_n} - \frac{R(t_n h) - \tilde{R}(t_n h)}{t_n} \right| \\
&= \left| \frac{\Psi_w(t_n h) - \tilde{\Psi}(t_n h)}{t_n} \right| + \left| \frac{R(t_n h) - \tilde{R}(t_n h)}{t_n} \right| \\
&\leq |t_n| \left| \Psi_w(h) - \tilde{\Psi}(h) \right| + \left| \frac{R(t_n h) - \tilde{R}(t_n h)}{t_n^2} \right| |t_n| \to 0
\end{aligned}$$

proving that $\mathbb{E}_{Q_w}(h) = \mathbb{E}_{\tilde{Q}}(h)$. Since $h \neq 0$ was arbitrarily chosen, it follows that $\tilde{Q} = Q_w$. By (29), we then have

$$0 = \left| -\frac{1}{2} u'(w) \Phi_w(h) + \frac{1}{2} u'(w) \tilde{\Phi}(h) + R(t_n h) - \tilde{R}(t_n h) \right| \tag{30}$$

Thus,

$$\begin{aligned}
&\left| \frac{1}{2} u'(w) \tilde{\Phi}(h) - \frac{1}{2} u'(w) \Phi_w(h) \right| \\
&= \left| \frac{\frac{1}{2} u'(w) \tilde{\Phi}(t_n h) - \frac{1}{2} u'(w) \Phi_w(t_n h)}{t_n^2} \right| \\
&= \left| \frac{\frac{1}{2} u'(w) \tilde{\Phi}(t_n h) - \frac{1}{2} u'(w) \Phi_w(t_n h)}{t_n^2} + \frac{R(t_n h) - \tilde{R}(t_n h)}{t_n^2} - \frac{R(t_n h) - \tilde{R}(t_n h)}{t_n^2} \right| \\
&\leq \left| \frac{\frac{1}{2} u'(w) \tilde{\Phi}(t_n h) - \frac{1}{2} u'(w) \Phi_w(t_n h)}{t_n^2} + \frac{R(t_n h) - \tilde{R}(t_n h)}{t_n^2} \right| + \left| \frac{R(t_n h) - \tilde{R}(t_n h)}{t_n^2} \right| \\
&= \left| \frac{R(t_n h) - \tilde{R}(t_n h)}{t_n^2} \right| \to 0
\end{aligned}$$

proving that $\Phi_w(h) = \tilde{\Phi}(h)$ for all $h \neq 0$ since $u' > 0$. As $\Phi_w(0) = \tilde{\Phi}(0) = 0$, we conclude that $\Phi_w = \tilde{\Phi}$. This completes the proof of uniqueness. The probabilistically sophisticated part follows from Proposition 15. ■

Proof of Proposition 5 The equality $Q_w^1 = Q_w^2$ follows from Lemma 1, the inequality $\Phi_w^1 \geq \Phi_w^2$ from Proposition 14. ■

Proof of Proposition 6 The equality $Q_w = Q_{w'}$ follows from Proposition 12. Moreover, by Corollary 2 it holds $\Phi_w \leq 0$. ■

Proof of Proposition 7 By Proposition 4,

$$c(w+h) = w + \mathbb{E}_P(h) - \frac{\lambda_u(w)}{2}\mathbb{V}_P(h) - \frac{u'(w)}{2}\Phi_w(h) + R(h)$$

where $R(th) = o(t^2)$. By Proposition 2, Φ_w is law invariant. By Proposition 16, Φ_w is Lebesgue-continuous, so continuous in law. By Proposition 19, $\Phi_w(h) = b_w \mathbb{V}_P(h)$. ■

Proof of Proposition 8 Since I is constant additive, by Proposition 13-(i) we have $Q_w = \nabla g(u(w)) = \nabla g(u(w')) = Q_{w'}$ for all $w, w' \in C$. By part (a) of Proposition 13-(ii), we also have

$$\Phi_w(h) = -\left\langle \nabla^2 g(u(w))(h), h\right\rangle = -\left\langle \nabla^2 g(u(w'))(h), h\right\rangle = \Phi_{w'}(h)$$

for all $w, w' \in C$. We conclude that Q_w and Φ_w are both independent of w. By part (b) of Proposition 13-(ii), Φ_w is a protovariance. The result then follows from Proposition 4. When V is variational, it can be shown that I is balanced. By Corollary 2, $\Phi_w \geq 0$. ■

C A general two-stage analysis

We close with a general version of the smooth ambiguity criterion. A *two-stage decision criterion* V features a functional I given by

$$I(\xi) = J\left(\mathbb{E}_{(\cdot)}\xi\right)$$

where $J : L_M^\infty(\operatorname{Im} u) \to \mathbb{R}$ is normalized and monotone. So, this criterion has the form

$$V(f) = I(u \circ f) = J\left(\mathbb{E}_{(\cdot)} u \circ f\right) \tag{31}$$

In particular, the smooth case correspond to J of the form $J(\xi) = \phi^{-1}(\mathbb{E}_\mu(\phi(\xi)))$.

For the two-stage criterion we can generalize the mean-variance approximation (8) when the prior is adequate.

Proposition 20 *Let V be an analytical two-stage rational decision criterion. Assume that V is continuously probabilistically sophisticated under an adequate prior $\mu \in \Delta(M)$, that is:*

(i) for all $\xi, \xi' \in L^\infty(C)$,

$$\mathbb{E}_{(\cdot)}\xi \overset{d}{\sim}_\mu \mathbb{E}_{(\cdot)}\xi' \Longrightarrow I(\xi) = I(\xi)$$

(ii) for each uniformly bounded sequence $\{\xi_n\}$ in $L^\infty(C)$,

$$\mathbb{E}_{(\cdot)}\xi_n \rightsquigarrow \mathbb{E}_{(\cdot)}\xi \Longrightarrow I(\xi_n) \to I(\xi)$$

Then, at each $w \in C$ there exists $b_w \in \mathbb{R}$ such that

$$c(w+h) = w + \mathbb{E}_{\bar{\mu}}(h) - \frac{\lambda_u(w)}{2}\mathbb{V}_{\bar{\mu}}(h) - \frac{u'(w)}{2} b_w \mathbb{V}_\mu\left(\mathbb{E}_{(\cdot)}h\right) + R(h) \tag{32}$$

where $R(th) = o(t^2)$.

To put this result in perspective, note that the smooth ambiguity criterion is continuously probabilistically sophisticated under the prior μ (that, however, is not required to be adequate). Approximation (8) is, formally, the special case with $b_w = \lambda_\phi(u(w))$.

Proof In view of (15) and of Lemma 4, we have

$$\begin{aligned}\hat{c}(w+h) &= \hat{c}(w) + \langle \nabla \hat{c}(w), h\rangle + \frac{1}{2}\left\langle \nabla^2 \hat{c}(w)(h), h\right\rangle + R(h) \\ &= w + \mathbb{E}_{\bar{\mu}}(h) - \frac{\lambda_u(w)}{2}\mathbb{V}_{\bar{\mu}}(h) + \frac{u'(w)}{2}\left\langle \nabla^2 \hat{g}(w)(\Xi(h)), \Xi(h)\right\rangle_M\end{aligned}$$

where $R(th) = o(t^2)$. By Propositions 16 and 19,

$$\left\langle \nabla^2 \hat{g}(w)(\Xi(h)), \Xi(h)\right\rangle_M = b_w \mathbb{V}_\mu\left(\mathbb{E}_{(\cdot)}h\right)$$

This completes the proof of (32). ■

References

[1] C. D. Aliprantis and K. C. Border, *Infinite dimensional analysis*, 3rd ed., Springer-Verlag, Berlin, 2006.

[2] A. Ambrosetti and G. Prodi, *A primer of nonlinear analysis*, Cambridge, Cambridge University Press, 1993.

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

[4] K. Arrow, *Essays in the theory of risk-bearing*, Markham, Chicago, 1971.

[5] K. P. S. Bhaskara Rao and M. Bhaskara Rao, *Theory of charges*, Academic Press, New York, 1983.

[6] P. Billingsley, *Probability and measure*, 3rd ed., Wiley, New York, 1995.

[7] S. Cerreia-Vioglio, F. Maccheroni, and M. Marinacci, Absolute and relative ambiguity aversion: A preferential approach, *Journal of Economic Theory*, 199, 2022.

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

[9] S. Cerreia-Vioglio, F. Maccheroni, M. Marinacci, and L. Montrucchio, Risk measures: rationality and diversification, *Mathematical Finance*, 21, 743-774, 2011.

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

[11] P. Ghirardato, F. Maccheroni and M. Marinacci, Differentiating ambiguity and ambiguity attitude, *Journal of Economic Theory*, 118, 133-173, 2004.

[12] L. P. Hansen and T. J. Sargent, Robust control and model uncertainty, *American Economic Review*, 91, 60-66, 2001.

[13] L. P. Hansen and T. J. Sargent, *Robustness*, Princeton University Press, Princeton, 2008.

[14] C. Hara, Measure of ambiguity aversion for smooth utility functions, mimeo, 2022.

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

[16] S. Kurepa, On the quadratic functional, *Publications de l'Institut Mathématique,* 13, 57-72, 1959

[17] P. Jordan and J. von Neumann. On inner products in linear, metric spaces, *Annals of Mathematics*, 36, 719-723, 1935.

[18] E. Jouini, W. Schachermayer and N. Touzi, Law invariant risk measures have the Fatou property, in *Advances in mathematical economics*, 49-71, 2006

[19] F. Maccheroni, M. Marinacci, and D. Ruffino, Alpha as ambiguity: Robust mean-variance portfolio analysis, *Econometrica*, 81, 1075-1113, 2013.

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

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

[22] J. W. Pratt, Risk aversion in the small and in the large, *Econometrica*, 122-136, 1964.