Alpha as Ambiguity: Robust Mean-Variance Portfolio Analysis*

Fabio Maccheroni[a] Massimo Marinacci[a] Doriana Ruffino[b]

[a]Università Bocconi
[b]University of Minnesota

Revision of March 2012

Abstract

We derive the analogue of the classic Arrow-Pratt approximation of the certainty equivalent under model uncertainty as described by the smooth model of decision making under ambiguity of Klibanoff, Marinacci and Mukerji (2005). We study its scope by deriving a tractable mean-variance model adjusted for ambiguity and solving the corresponding portfolio allocation problem. In the problem with a risk-free asset, a risky asset, and an ambiguous asset, we find that portfolio rebalancing in response to higher ambiguity aversion *only* depends on the ambiguous asset's *alpha*, setting the performance of the risky asset as benchmark. In particular, a positive *alpha* corresponds to a long position in the ambiguous asset, a negative *alpha* corresponds to a short position in the ambiguous asset, and greater ambiguity aversion reduces optimal exposure to ambiguity. The analytical tractability of the enhanced Arrow-Pratt approximation renders our model especially well suited for calibration exercises aimed at exploring the consequences of model uncertainty on equilibrium asset prices.

"*Crises feed uncertainty. And uncertainty affects behaviour, which feeds the crisis.*"

Olivier Blanchard, *The Economist*, January 29, 2009

1 Introduction

When a von Neumann-Morgenstern expected utility maximizer with utility u and wealth w considers an investment h, the Arrow-Pratt approximation of his certainty equivalent for the resulting uncertain prospect $w+h$ is

$$c(w+h,P) \approx w + E_P(h) - \frac{1}{2}\lambda_u(w)\,\sigma_P^2(h), \tag{1}$$

where P is the probabilistic model that describes the stochastic nature of the problem.

This classic approximation has two main merits, a theoretical and a practical one. Its theoretical merit is to show that, for an expected utility agent, the premium associated with facing risk h is proportional to the variance $\sigma_P^2(h)$ of h with respect to P. This relation between risk and variance

*Maccheroni and Marinacci are at the Department of Decision Sciences, Dondena, and Igier, Università Bocconi, Italy; Ruffino is at the Department of Finance, University of Minnesota, USA. We thank David Bates, Emanuele Borgonovo, John Cochrane, Marco Giacoletti, Massimo Guidolin, Lars Hansen, Sujoy Mukerji, Fulvio Ortu, Andrea Resti, Alessandro Sbuelz, Claudio Tebaldi, the co-editor Wolfgang Pesendorfer, three anonymous referees, and especially Giulia Brancaccio and Simone Cerreia-Vioglio for helpful comments. The financial support of ERC (advanced grant, BRSCDP-TEA) and the Carlson School of Management at the University of Minnesota (Dean's Research Grant) is gratefully acknowledged.

is a central pillar of risk management. In particular, the coefficient $\lambda_u(w) = -u''(w)/u'(w)$ that links risk premium and volatility is determined by the agent's risk aversion at w. The practical merit of (1) is in providing the foundation for the mean-variance preference model, where a prospect f is evaluated through

$$U(f) = E_P(f) - \frac{\lambda}{2}\sigma_P^2(f), \tag{2}$$

obtained from (1) by setting $w + h = f$ and $\lambda_u(w) = \lambda$. This model is the workhorse of asset management in the finance industry.

The purpose of this paper is to extend the classic Arrow-Pratt analysis to account for *model uncertainty*: the situation in which the agent is uncertain about the true probabilistic model P that governs the occurrence of different states. If only risk is present, that is, the agent fully relies on a single probabilistic model P, then the certainty equivalent $c(w+h, P)$ of $w+h$, the sure amount of money that he considers equivalent to the uncertain prospect $w+h$, is given by

$$c(w+h, P) = u^{-1}(E_P(u(w+h))). \tag{3}$$

Here u represents the agent's attitude toward risk. If, in contrast, the agent is not able to identify a single probabilistic model P, but he also considers alternative models Q, then $c(w+h, Q)$ becomes a variable amount of money that depends on Q. Suppose μ is the agent's prior probability on the space Δ of possible models and v is his attitude toward model uncertainty (*stricto sensu*; see Section 2.2). The rationale used to obtain the certainty equivalent (3) leads to a (second-order) certainty equivalent

$$\begin{aligned} C(w+h) &= v^{-1}(E_\mu(v(c(w+h)))) \\ &= v^{-1}\left(E_\mu\left(v\left(u^{-1}(E(u(w+h)))\right)\right)\right), \end{aligned} \tag{4}$$

where $c(w+h)$ is the random variable that associates $c(w+h, Q)$ to each model Q in Δ. This is the smooth ambiguity certainty equivalent of Klibanoff, Marinacci and Mukerji (2005), henceforth abbreviated KMM.

The case in which the support of the prior μ is a singleton P corresponds to the absence of model uncertainty. In fact, the agent is fully confident about P and (4) coincides with (3). Analogously, if $v = u$ it can be shown that

$$C(w+h) = c\left(w+h, \bar{Q}\right)$$

where $\bar{Q}$ is the reduced probability $\int Q d\mu(Q)$ induced by the prior μ. In this case, the certainty equivalent (4) reduces to (3) where the probabilistic model P is replaced by $\bar{Q}$; in the jargon of decision theory, the agent is *ambiguity neutral* and the reduced distribution $\bar{Q}$ represents all the uncertainty he is facing (see Ellsberg, 1961, p. 661). However, if the support of μ is nonsingleton (there is model uncertainty, an information feature) and v differs from u (the reactions to model uncertainty and to risk differ, a taste feature) the identification of (3) and (4) no longer holds – model uncertainty cannot be reduced to risk – and the Arrow-Pratt analysis needs to be extended.

The first step in our extension of the Arrow-Pratt analysis is to derive in Section 3 the analogue of approximation (1) under ambiguity, as captured by the KMM certainty equivalent (4). Specifically, Proposition 3 shows that:

$$C(w+h) \approx w + E_{\bar{Q}}(h) - \frac{1}{2}\lambda_u(w)\sigma_{\bar{Q}}^2(h) - \frac{1}{2}(\lambda_v(w) - \lambda_u(w))\sigma_\mu^2(E(h)), \tag{5}$$

where $\bar{Q} = \int Q d\mu(Q)$ is the reduced probability induced by the prior μ, and $E(h) : \Delta \to \mathbb{R}$ is the random variable

$$Q \mapsto E_Q(h)$$

that associates the expected value $E_Q(h)$ to each possible model Q. Its variance $\sigma_\mu^2(E(h))$, along with the difference $\lambda_v(w) - \lambda_u(w)$ in uncertainty attitudes, determines an *ambiguity premium* – the last term in (5) – that is novel relative to (1). In other words, model uncertainty renders volatile the return $E(h)$ of h, thereby affecting the agent's certainty equivalent. Indeed, (5) shows that $\lambda_v(w)$ captures model uncertainty aversion also in "the small," *à la* Pratt (1964), in fact, *ceteris paribus*, the higher $\lambda_v(w)$ the greater the ambiguity premium. In turn, this completes the KMM analysis of ambiguity aversion, as discussed at the end of Section 3.

In Section 4, we study those prospects that are unaffected by model uncertainty, that is, the special class of prospects for which approximation (5) reduces to its classic counterpart (1).

The quadratic approximation (5) allows us to extend, in Section 5, the mean-variance model (2). Specifically, by setting $w + h = f$, $\lambda_u(w) = \lambda$, $\lambda_v(w) - \lambda_u(w) = \theta$, and $\bar{Q} = P$, we obtain the following natural and parsimonious extension

$$U(f) = E_P(f) - \frac{\lambda}{2}\sigma_P^2(f) - \frac{\theta}{2}\sigma_\mu^2(E(f)) \tag{6}$$

of the mean-variance model (2) that is able to deal with ambiguity. This augmented mean-variance model is determined by the three parameters λ, θ, and μ, as opposed to the two parameters λ and P of the classic mean-variance model. The taste parameters λ and θ represent attitudes toward risk and ambiguity, respectively. Higher values of these parameters correspond to stronger negative attitudes. The information parameter μ determines the variances $\sigma_P^2(f)$ and $\sigma_\mu^2(E(f))$ that measure the risk and model uncertainty perceived in the evaluation of prospect f. Higher values of these variances correspond to poorer information on prospect's outcomes and on models.

In Section 6, we study the scope of the augmented mean-variance model (6) via a portfolio allocation exercise. In particular, we study a tripartite portfolio problem with a risk-free asset, a purely risky asset, and an ambiguous one. Relative to more traditional portfolio analyses with a risk-free and a risky asset only, the addition of an ambiguous asset allows for the study of model uncertainty. Our portfolio analysis shows that optimal portfolio rebalancing in response to higher ambiguity aversion *only* depends on the ambiguous asset's *alpha*, setting the performance of the risky asset as benchmark. An asset's *alpha*, it is found, is the component of the expected excess return of the ambiguous asset which is ambiguity specific, that is, uncorrelated with pure risk. More precisely, it is the (expected) return of the ambiguous asset in excess of the return on the risk free asset less the amount that can be explained as the excess return due to the pure risk embedded in the asset.[1] When *alpha* is positive, the asset return offers compensation in excess of its risky component, a compensation that an ambiguity averse agent would need to hold a long position in the asset. Indeed, a positive *alpha* corresponds to a long position in the ambiguous asset, a negative *alpha* corresponds to a short position in the ambiguous asset, and greater ambiguity aversion reduces optimal exposure to ambiguity.

Some fundamental asset allocation problems feature a natural tripartite structure. This is the case for international portfolio allocation problems with domestic bonds, domestic stocks, and foreign stocks. Our analysis is relevant for these problems when the information available to investors is such that the tripartite structure may be interpreted as reflecting different types of uncertainty (i.e., risk and ambiguity) about the assets. We expect this to be often the case.[2]

Related Works Our work is related to recent papers by Nau (2006), Skiadas (2009), Izhakian and Benninga (2011), and Jewitt and Mukerji (2011), that *inter alia* also obtain approximations for the ambiguity premium in the smooth ambiguity model on the basis of special assumptions. More

[1] See (29) and its discussion for details.

[2] See, e.g., French and Poterba (1991), Canner, Mankiw and Weil (1997), and Huberman (2001) for evidence on these and related allocation problems that is inconsistent with existing static choice models.

importantly, these papers do not use the approximation to extend the mean-variance approach and study portfolio decisions.

Our findings on the portfolio selection problem share some features with the ones of Epstein and Miao (2003), Taboga (2005), Boyle, Garlappi, Uppal and Wang (2012), and Gollier (2011). In particular, Taboga (2005) proposes a model of portfolio selection based on a two-stage evaluation procedure to disentangle ambiguity and ambiguity aversion. Gollier (2011) investigates the comparative statics of more ambiguity aversion in a static two-asset portfolio problem. He shows that ambiguity aversion may not reinforce risk aversion and exhibits sufficient conditions to guarantee that, ceteris paribus, an increase in ambiguity aversion reduces the optimal exposure to ambiguity. Gollier's insight has been confirmed in terms of ambiguity premia by Izhakian and Benninga (2011), who show, for CRRA and CARA specifications, that such premium may differ qualitatively from the risk premium. Epstein and Miao (2003) use a recursive multiple priors model to study the home bias, while Boyle, Garlappi, Uppal and Wang (2012) employ the concepts of ambiguity and ambiguity aversion in a multiple priors framework to formalize the idea of investor's "familiarity" toward assets.

In addition, the analytical tractability of the enhanced Arrow-Pratt approximation (5) favors empirical tests of our model's implications to several observationally puzzling (and economically interesting) investment behaviors. These include the home bias puzzle, the equity premium puzzle, as well as the employer-stock ownership puzzle. For this reason, our paper is also related to several papers in the literature that explore the consequences of ambiguity aversion on equilibrium prices. Among others, Chen and Epstein (2002) identify separate excess return premia for risk and ambiguity within a representative agent asset market setting, while Garlappi, Uppal, and Wang (2007) extend a traditional portfolio problem to a multiple priors setting. Caskey (2009) and Illeditsch (2011) study the effects of "ambiguous" information on investors' market trades and valuations. Easley and O'Hara (2010a,b) explain how low trading volumes during part of the recent financial crisis may have resulted from investors' perceived uncertainty and how designing markets to reduce ambiguity may induce participation by both investors and issuers.

By use of recursive versions of the smooth ambiguity model, Ju and Miao (2010) calibrate a representative agent consumption based asset pricing model and generate a variety of dynamic asset pricing phenomena that are observed in the data; Chen, Ju, and Miao (2011) study an investor's optimal consumption and portfolio choice problem;[3] while Collard, Mukerji, Sheppard and Tallon (2011) show the importance of model uncertainty for the analysis of long run risk (LLR) by matching the historical equity premium with a LLR model that features endogenously time-varying ambiguity (e.g., increasing during recessions) based on publicly available data on aggregate consumption and dividend. Hansen and Sargent (2010) consider two risk-sensitivity operators, one of them being a recursive version of the smooth ambiguity model, in a LLR setup and show how sensitive beliefs become to information under model uncertainty; they show how the resulting "model uncertainty premia" affects the price of macroeconomic risk. Finally, Weitzman (2007) shows how model uncertainty can be important in dynamic asset pricing already under standard expected utility (and so, without taking into account agents' specific reaction to model uncertainty) when "persistent" uncertainty may prevent full learning of the true data generating process, so that subjective beliefs on models keep being relevant even with large amounts of observations. Though these dynamic papers are different from our static analysis (where learning issues are absent), they share with this paper the insight that a proper account of model uncertainty is required to better understand the quantitative puzzles of asset markets.

[3] The recursive multiple priors case appears in Miao (2009).

2 Preliminaries

2.1 Mathematical Setup

Given a probability space $(\Omega, \mathcal{F}, P)$, let $L^2 = L^2(\Omega, \mathcal{F}, P)$ be the Hilbert space of square integrable random variables on Ω and $L^\infty = L^\infty(\Omega, \mathcal{F}, P)$ be the subset of L^2 consisting of its almost surely bounded elements. Given an interval $I \subseteq \mathbb{R}$, we set

$$L^\infty(I) = \{f \in L^\infty : \operatorname{essinf} f, \operatorname{esssup} f \in I\}.$$

Throughout the paper $\|\cdot\|$ denotes the L^2 norm. The space L^2 is the natural setting for this paper because of our interest in quadratic approximations.

We indicate by $E_P(X)$ and $\sigma_P^2(X)$ the expectation and variance of a random variable $X \in L^2$, respectively. Moreover, we indicate by $\sigma_P(X, Y)$ the covariance

$$\sigma_P(X, Y) = E_P[(X - E_P(X))(Y - E_P(Y))]$$

between two random variables $X, Y \in L^2$.

The set of probability measures Q on $\mathcal{F}$ that have square integrable density $q = dQ/dP$ with respect to P can be identified, via Radon-Nikodym derivation, with the closed and convex subset of L^2 given by

$$\Delta = \left\{ q \in L_+^2 : \int_\Omega q(\omega)\, dP(\omega) = 1 \right\}.$$

By Bonnice and Klee (1963, Th. 4.3), we have the following existence result.

Lemma 1 *Given a Borel probability measure μ on Δ with bounded support,*[4] *there exists a unique $\bar{q} \in \Delta$ such that*

$$\int_\Omega X(\omega)\, \bar{q}(\omega)\, dP(\omega) = \int_\Delta \left(\int_\Omega X(\omega)\, q(\omega)\, dP(\omega) \right) d\mu(q), \qquad \forall X \in L^2. \tag{7}$$

The density $\bar{q}$ is denoted by $\int_\Delta q d\mu(q)$ and is called *barycenter* of μ. Notice that, when restricted to indicator functions 1_A of elements of $\mathcal{F}$, (7) delivers

$$\bar{Q}(A) = \int_\Delta Q(A)\, d\mu(Q), \qquad \forall A \in \mathcal{F}, \tag{8}$$

where the identification of each probability measure Q with its density q allows to write $d\mu(Q)$ instead of $d\mu(q)$. The probability measure $\bar{Q}$ is called *reduction* of μ on Ω. In fact, (8) suggests a natural interpretation of $\bar{Q}$ in terms of reduction of compound lotteries. For example, if $\operatorname{supp}\mu = \{Q_1, ..., Q_n\}$ is finite and $\mu(Q_i) = \mu_i$ for $i = 1, ..., n$, then (8) becomes

$$\bar{Q}(A) = \mu_1 Q_1(A) + ... + \mu_n Q_n(A), \qquad \forall A \in \mathcal{F}.$$

Hence, μ can be seen as a lottery whose outcomes are all possible models, which in turn can be seen as lotteries that determine the state.

[4] A *carrier* of μ is any Borel subset of Δ having full measure. If the intersection of all closed carriers is a carrier, it is called *support* of μ and denoted by $\operatorname{supp}\mu$. Since Δ inherits the L^2 distance, $\operatorname{supp}\mu$ is *bounded* when $\{\|q\| : q \in \operatorname{supp}\mu\}$ is a bounded set of real numbers. If $\mathcal{F}$ is finite the support always exists and it is bounded.

2.2 Decision Theoretic Setup

Given any nonsingleton interval $I \subseteq \mathbb{R}$ of monetary outcomes, we consider decision makers (DMs) who behave according to the smooth model of decision making under ambiguity of KMM.[5] That is, DMs who rank prospects through the functional $V : L^{\infty}(I) \to \mathbb{R}$ defined by

$$V(f) = \int_{\Delta} \phi\left(\int_{\Omega} u(f(\omega))\, q(\omega)\, dP(\omega)\right) d\mu(q), \qquad \forall f \in L^{\infty}(I), \tag{9}$$

where μ is a Borel probability measure on Δ with bounded support, and $u : I \to \mathbb{R}$ and $\phi : u(I) \to \mathbb{R}$ are smooth and strictly increasing functions.

Lemma 2 *The functional $V : L^{\infty}(I) \to \mathbb{R}$ is well defined, with $V(L^{\infty}(I)) = \phi(u(I))$.*

The certainty equivalent function $C : L^{\infty}(I) \to I$ induced by V is defined by $V(C(f)) = V(f)$ for all prospects f, that is,

$$C(f) = u^{-1}\left(\phi^{-1}\left(\int_{\Delta} \phi\left(\int_{\Omega} u(f(\omega))\, q(\omega)\, dP(\omega)\right) d\mu(q)\right)\right), \qquad \forall f \in L^{\infty}(I). \tag{10}$$

In the monetary setting of the present paper, where outcomes are amounts of money and prospects are financial assets, it is natural to consider monetary certainty equivalents. To this end, set $v = \phi \circ u : I \to \mathbb{R}$ (see KMM p. 1859). It is then possible to rewrite (9) as

$$V(f) = \int_{\Delta} \left(v \circ u^{-1}\right)\left(\int_{\Omega} u(f(\omega))\, q(\omega)\, dP(\omega)\right) d\mu(q), \qquad \forall f \in L^{\infty}(I), \tag{11}$$

and so (10) as

$$C(f) = v^{-1}\left(\int_{\Delta} v\left(u^{-1}\left(\int_{\Omega} u(f(\omega))\, q(\omega)\, dP(\omega)\right)\right) d\mu(q)\right), \qquad \forall f \in L^{\infty}(I). \tag{12}$$

Here the certainty equivalent $C(f)$ is viewed as the composition of two monetary certainty equivalents,

$$c(f,q) = u^{-1}\left(\int_{\Omega} u(f(\omega))\, q(\omega)\, dP(\omega)\right) \quad \text{and} \quad v^{-1}\left(\int_{\Delta} v(c(f,q))\, d\mu(q)\right).$$

This is the approach we sketched in the introduction, motivated by the paper monetary setting.

In KMM the function v represents attitudes toward *stricto sensu* model uncertainty, that is, the uncertainty that agents face when dealing with alternative possible probabilistic models. The function v is characterized in KMM along with the prior μ through prospects whose outcomes depend only on models and, as such, are only affected by model uncertainty.

Model uncertainty cumulates with the state uncertainty that any nontrivial probabilistic model features. The combination of these two sources of uncertainty determines in the KMM model the ambiguity that DMs face in ranking prospects $f : \Omega \to \mathbb{R}$. KMM show that overall attitudes toward ambiguity are captured by the function ϕ. In particular, its concavity characterizes ambiguity aversion, which therefore implies positive Arrow-Pratt coefficients $\lambda_{\phi} = -\phi''/\phi'$. Since

$$\lambda_{\phi}(u(w)) = \frac{1}{u'(w)}\left(\lambda_v(w) - \lambda_u(w)\right) \tag{13}$$

we conclude that ambiguity aversion amounts to $\lambda_v - \lambda_u \geq 0$, a key condition for the paper.[6]

Ambiguity neutrality is modelled by $\phi(x) = x$, that is, $v = u$, while absence of model uncertainty is modelled by a trivial μ with singleton support (i.e., a Dirac measure). In both cases criterion (9) reduces to expected utility, though in one case the reduction originates from a taste component – a neutral attitude, under which the two sources of uncertainty "linearly" combine via reduction (8) – while in the other case it originates from an information component (absence of a source of uncertainty, i.e., model uncertainty).

[5] We use the terms decision maker and agent interchangeably in the paper.
[6] See Lemma 12 in the appendix.

3 Quadratic Approximation

Let $w \in \operatorname{int} I$ be a scalar interpreted as current wealth. To ease notation, we also denote by w the degenerate random variable $w1_\Omega$. Given any prospect $h \in L^\infty$ such that $w + h \in L^\infty(I)$, we are interested in the certainty equivalent $C(w+h)$ of $w+h$, that is,

$$C(w+h) = v^{-1}\left(\int_\Delta v\left(u^{-1}\left(\int_\Omega u(w+h)\,qdP\right)\right)d\mu(q)\right). \tag{14}$$

For all $h \in L^\infty$, the functions

$$E(h) : q \mapsto \int_\Omega hqdP \quad \text{and} \quad \sigma^2(h) : q \mapsto \int_\Omega h^2 qdP - \left(\int_\Omega hqdP\right)^2$$

are continuous and bounded on Δ, and so belong to $L^\infty(\Delta, \mathcal{B}, \mu)$. The variance of $E(h)$ with respect to μ

$$\int_\Delta \left(\int_\Omega h(\omega)\,q(\omega)\,dP(\omega)\right)^2 d\mu(q) - \left(\int_\Delta \left(\int_\Omega h(\omega)\,q(\omega)\,dP(\omega)\right)d\mu(q)\right)^2$$

is denoted by $\sigma_\mu^2(E(h))$. This variance reflects the uncertainty on the expectation $E(h)$ which, in turn, is implied by model uncertainty. Thus, higher values of $\sigma_\mu^2(E(h))$ correspond to a higher incidence of model uncertainty on the expectation of h.

We can now state the second order approximation of the certainty equivalent (14).

Proposition 3 *Let μ be a Borel probability measure with bounded support on Δ and $u, v : I \to \mathbb{R}$ be twice continuously differentiable with $u', v' > 0$. Then,*

$$C(w+h) = w + E_{\bar{Q}}(h) - \frac{1}{2}\lambda_u(w)\,\sigma_{\bar{Q}}^2(h) - \frac{1}{2}\left(\lambda_v(w) - \lambda_u(w)\right)\sigma_\mu^2(E(h)) + R_2(h) \tag{15}$$

for all $h \in L^\infty$ such that $w + h \in L^\infty(I)$, where

$$\lim_{t \to 0} \frac{R_2(th)}{t^2} = 0. \tag{16}$$

Moreover, if $\mathcal{F}$ is finite, then $R_2(h) = o\left(\|h\|^2\right)$ as $h \to 0$ in L^2.

Notice that the first three components on the right hand side of (15) correspond to the Arrow-Pratt approximation of the ambiguity neutral certainty equivalent $u^{-1}\left(\int u(w+h)\,d\bar{Q}\right)$ of $w+h$. To the contrary, the fourth component represents the effects of ambiguity: the sign and magnitude of these effects on the certainty equivalent depend on the difference $\lambda_v(w) - \lambda_u(w)$. In particular, provided model uncertainty affects the expectation of h, that is $\sigma_\mu^2(E(h)) \neq 0$, ambiguity has no effects when the DM is neutral to it at w, i.e., $\lambda_v(w) = \lambda_u(w)$.[7] Finally the fifth component "the error" tends to zero faster than the square of the size of the uncertainty exposure.

The variance $\sigma_{\bar{Q}}^2(h)$ can be decomposed along the two sources of uncertainty as

$$\sigma_{\bar{Q}}^2(h) = E_\mu\left(\sigma^2(h)\right) + \sigma_\mu^2(E(h))$$

State uncertainty, which exists within each model, underlies the average variance $E_\mu\left(\sigma^2(h)\right)$. Model uncertainty, instead, determines the variance of averages $\sigma_\mu^2(E(h))$. Approximation (15) can thus be rearranged according to the Arrow-Pratt coefficients of u and v as follows:

$$C(w+h) = w + E_{\bar{Q}}(h) - \frac{\lambda_u(w)}{2}E_\mu\left(\sigma^2(h)\right) - \frac{\lambda_v(w)}{2}\sigma_\mu^2(E(h)) + R_2(h). \tag{17}$$

[7] Notice that this may be the case even if $v \neq u$, that is, the requirement $\lambda_v(w) = \lambda_u(w)$ is a requirement of ambiguity neutrality in "the small" rather than in "the large."

This formulation shows that, when the indexes u and v are sufficiently smooth, both state and model uncertainty play at most a second order effect in the evaluation.[8] Specifically, risk aversion determines the DM's reaction to the average variance $E_\mu\left(\sigma^2(h)\right)$ and model uncertainty aversion determines his reaction to the variance of averages $\sigma^2_\mu(E(h))$.

We conclude by observing that formulas (13) and (15) allow to interpret ϕ as capturing ambiguity aversion both in the large and in the small *à la* Pratt (1964). Specifically, given $u, v_1, v_2 : I \to \mathbb{R}$ twice continuously differentiable with $u', v_1', v_2' > 0$, then, by (13),

$$\lambda_{\phi_1}(u(w)) \geq \lambda_{\phi_2}(u(w)) \Leftrightarrow \frac{1}{u'(w)}(\lambda_{v_1}(w) - \lambda_u(w)) \geq \frac{1}{u'(w)}(\lambda_{v_2}(w) - \lambda_u(w))$$
$$\Leftrightarrow \lambda_{v_1}(w) \geq \lambda_{v_2}(w).$$

In "the large," by Corollary 3 of KMM, this implies that agent 1 is more ambiguity averse than agent 2 if and only if he is more model uncertainty averse than agent 2. In "the small," by (15), this is equivalent to say that the ambiguity premium for agent 1 is greater than the ambiguity premium for agent 2 (while risk premia coincide).

In the rest of the paper we will focus on the case in which the DM is ambiguity averse at w, i.e., $\lambda_v(w) - \lambda_u(w) \geq 0$. Similarly, we will assume risk aversion at w, i.e., $\lambda_u(w) \geq 0$.

4 Approximately Unambiguous Prospects

As previously observed, the ambiguity premium

$$\frac{1}{2}(\lambda_v(w) - \lambda_u(w))\,\sigma^2_\mu(E(h)) \tag{18}$$

may vanish for two reasons: the ambiguity neutrality condition $\lambda_v(w) = \lambda_u(w)$ or the information condition $\sigma^2_\mu(E(h)) = 0$. In this section we consider the latter case since it directly refers to model uncertainty (as captured by the parameter μ) and our main focus is on ambiguity non-neutral behavior.

Definition 4 *A prospect $h \in L^2$ is* approximately unambiguous *if $\sigma^2_\mu(E(h)) = 0$.*

Clearly, if h is approximately unambiguous, then approximation (15) collapses to

$$C(w+h) = w + E_{\bar{Q}}(h) - \frac{1}{2}\lambda_u(w)\,\sigma^2_{\bar{Q}}(h) + R_2(h),$$

that is, the DM's evaluation of h is indistinguishable (in the second order approximation) from the certainty equivalent of an (ambiguity neutral) expected utility maximizer with utility u and beliefs given by the reduced probability measure $\bar{Q}$ induced by μ on Ω. The equivalence of (i) and (ii) in the next result shows that also the converse is true, thus motivating the term "approximately unambiguous" for such a prospect.

Proposition 5 *For a prospect $h \in L^2$, the following properties are equivalent:*

(i) h is approximately unambiguous;

(ii) $\sigma^2_\mu(E(h)) = R_2(h)$;

(iii) $E_Q(h) = E_{Q'}(h)$ for all $Q, Q' \in \mathrm{supp}\mu$.

[8] For standard expected utility this follows from the Arrow-Pratt approximation. Segal and Spivak (1990) is a classic study of orders of risk aversion. In Maccheroni, Marinacci, and Ruffino (2011) we study in detail orders of risk aversion and of model uncertainty aversion in the smooth ambiguity model.

The equivalence between (i) and (iii) is also noteworthy. It has two different implications. First, it says that h is approximately unambiguous if and only if the first moment of h is invariant across all models in the support of μ, that is, according to all models that a DM with prior μ deems plausible.[9] Second, it shows that the set of all approximately unambiguous prospects forms a closed linear subspace of L^2, and hence any prospect can be decomposed into an approximately unambiguous part and a residual ambiguous one (see Appendix A.3).

We conclude by showing that under model uncertainty there always exist arbitrarily small "ambiguous" (that is, not approximately unambiguous) prospects.

Proposition 6 *If $\lambda_v(w) \neq \lambda_u(w)$, then the following properties are equivalent:*

(i) all prospects in L^2 are approximately unambiguous;

(ii) there is an absorbing[10] subset B of L^∞ such that $w + h \in L^\infty(I)$ and

$$C(w+h) = w + E_{\bar{Q}}(h) - \frac{1}{2}\lambda_u(w)\,\sigma^2_{\bar{Q}}(h) + R_2(h) \tag{19}$$

for all $h \in B$;

(iii) μ is a Dirac measure.

This result shows that the *only* case in which our approximation (15) coincides with that of Arrow-Pratt for all "small prospects" (i.e., the prospects in B) is the one in which there is no model uncertainty to start with, that is, the prior μ is degenerate. Otherwise, for each $\varepsilon > 0$, however small, there is a prospect h with $\|h\| < \varepsilon$ for which the two approximations differ. In other words, as long as μ is not trivial and the agent is not ambiguity neutral, ambiguity effects may never fade away, even approximately for arbitrarily "small prospects" (the counterparts of "small risks" in risk theory).

This result may seem at odds with the findings of Skiadas (2009) who shows that ambiguity effects may fade in the small for specific vanishing nets of prospects. But there are two nontrivial differences that explain the presumed contradiction; for simplicity we elucidate them in the simplest possible case in which $\Omega = \{1, -1\}$, thus L^2 can be identified with $\mathbb{R}^2$.

- First, condition (ii) of Proposition 6 refers to an absorbing subset. For example, $\left\{(h,k) \in \mathbb{R}^2 : h^2 + k^2 < \varepsilon\right\}$. While Skiadas (2009) considers vanishing nets which are just paths closing to the origin. For example, $\left\{(t,-t) \in \mathbb{R}^2 : t \in (0,\varepsilon)\right\}$.
- Second, and conceptually more important, here μ is fixed. While in Skiadas (2009) it depends on t. For example, identifying Δ with the interval $[0,1]$,[11] fix $p, p' \in (0,1)$ and consider as μ_t the uniform distribution on $[p-t, p'+t]$ for $t \in (0,\varepsilon)$.

Clearly, the second point is very important because it allows model uncertainty to vanish when μ_t converges to a Dirac measure.[12] A complete analysis can be found in Maccheroni, Marinacci, and Ruffino (2011). In a nutshell, there we show that ambiguity effects fade if and only if model uncertainty vanishes, which is the perturbed version of Proposition 6 above.

[9] The interpretation of $\mathrm{supp}\mu$ as the set of plausible models, *à la* Ghirardato, Maccheroni, and Marinacci (2004), is formally discussed by Cerreia-Vioglio, Maccheroni, Marinacci, and Montrucchio (2011, Th. 21). Also notice that a simple prospect h is unambiguous in the sense of Ghirardato, Maccheroni, and Marinacci (2004) if and only if *all* of its moments (not only the *first*) coincide on the support of μ.

[10] A subset B of a vector space is *absorbing* if for any point of the space there exists a (strictly) positive multiple of B that contains the segment joining the point and zero. For example, any ball that contains the origin is absorbing.

[11] $q \in [0,1]$ being the probability of state 1.

[12] In the example, this happens if and only if $p = p'$, in which case μ_t weakly converges to δ_p.

5 Robust Mean-Variance Preferences

Inspired by the quadratic approximation (15), in this section we generalize standard mean-variance preferences to account for model uncertainty. Specifically, we consider a DM who ranks prospects f in L^2 through the robust mean-variance functional $C : L^2 \to \mathbb{R} \cup \{-\infty\}$ given by

$$C(f) = E_{\bar{Q}}(f) - \frac{\lambda}{2}\sigma^2_{\bar{Q}}(f) - \frac{\theta}{2}\sigma^2_{\mu}(E(f)), \qquad \forall f \in L^2, \tag{20}$$

where λ and θ are (strictly) positive coefficients, and μ is a Borel probability measure on Δ with bounded support and barycenter $\bar{Q}$ given by (8).

As mentioned in the introduction, this preference functional is fully determined by three parameters: λ, θ, and μ. Its theoretical foundation is given by the quadratic approximation (15), which shows that (20) can be viewed as a local approximation of a KMM preference functional (12) at a constant w such that $\lambda = \lambda_u(w)$ and $\theta = \lambda_v(w) - \lambda_u(w)$. Thus, the taste parameters λ and θ model the DM's negative attitudes toward risk and ambiguity, respectively. In particular, higher values of these parameters correspond to stronger negative attitudes.

In turn, the information parameter μ determines the variances $\sigma^2_{\bar{Q}}(f)$ and $\sigma^2_{\mu}(E(f))$ that measure the risk and model uncertainty that the DM perceives in the evaluation of prospect f. Higher values of these variances correspond to a DM's poorer information on prospect's outcomes and on models.

Since the probability measure $\bar{Q}$ is the reference model for a DM with prior μ, in order to facilitate comparison with the classic case we identify the barycenter $\bar{Q}$ of Δ with the baseline probability P. That is, in the rest of the paper we maintain the following assumption:

Assumption 1 *$\bar{Q} = P$.*

Under this assumption, $C(f)$ is always finite and (20) takes the form

$$C(f) = E_P(f) - \frac{\lambda}{2}\sigma^2_P(f) - \frac{\theta}{2}\sigma^2_{\mu}(E(f)), \qquad \forall f \in L^2, \tag{21}$$

which we will consider hereafter. When the information condition $\sigma^2_{\mu}(E(f)) = 0$ holds, we obtain the standard mean-variance evaluation

$$C(f) = E_P(f) - \frac{\lambda}{2}\sigma^2_P(f) \tag{22}$$

for prospect f. Approximately unambiguous prospects are thus regarded as *purely risky* by robust mean-variance preferences, that is, they form the class of prospects on which robust and conventional mean-variance preferences agree.

Like standard mean-variance preferences, our robust mean-variance preferences (21) separate taste parameters, λ and θ, and uncertainty measures, $\sigma^2_P(f)$ and $\sigma^2_{\mu}(E(f))$. This sharp separation gives standard mean-variance preferences an unsurpassed tractability and is the main reason for their success and widespread use. These key features fully extend to robust mean-variance preferences, as (21) shows. As a result, they are well-suited for finance and macroeconomics applications and can improve calibration and other quantitative exercises. Their scope will be illustrated in detail in the portfolio problem of next section.

Finally, we expect that a monotonic version of robust mean-variance preferences can be derived by suitably generalizing what Maccheroni, Marinacci, Rustichini, and Taboga (2009) established for conventional mean-variance preferences.

6 The Portfolio Allocation Problem

In this section we apply the newly obtained robust mean-variance preferences to a portfolio allocation problem. The theoretical difference between these preferences and the smooth ambiguity preferences of KMM is analogous to the one that separates the expected utility approach and the mean-variance utility approach in the case of portfolio selection under risk only.[13]

In the final part of the section we focus on a portfolio of three assets: a risk-free asset, a purely risky asset, and an ambiguous asset. This problem is the natural extension of the standard portfolio problem (with a risk-free and a risky asset) to our setting with model uncertainty. As mentioned in the introduction, international portfolio allocation problems provide a natural application of our setting with domestic Treasury bonds viewed as risk-free assets, other domestic assets viewed as purely risky assets, and foreign assets viewed as ambiguous assets. This will be our motivating example.

6.1 The General Setting

We consider the one-period optimization problem of an agent who has to decide how to allocate a unit of wealth among $n+1$ assets at time 0. The gross return on asset i after one period, $i = 1, ..., n$, is denoted by $r_i \in L^2$. Then, the $(n \times 1)$ vector of returns on the first n assets is denoted by $\mathbf{r}$ and the $(n \times 1)$ vector of portfolio weights (indicating the fraction of wealth invested in each asset) is denoted by $\mathbf{w}$. The return on the $(n+1)$-th asset is risk-free and it is equal to a constant r_f.

The end-of-period wealth $r_{\mathbf{w}}$ induced by a choice $\mathbf{w}$ is given by

$$r_{\mathbf{w}} = r_f + \mathbf{w} \cdot (\mathbf{r} - r_f \mathbf{1}),$$

where $\mathbf{1}$ is the n-dimensional unit vector. We assume frictionless financial markets in which assets are traded in the absence of transaction costs and both borrowing and short-selling are allowed without restrictions. Then, the portfolio problem can be written as

$$\max_{\mathbf{w} \in \mathbb{R}^n} C(r_{\mathbf{w}}) = \max_{\mathbf{w} \in \mathbb{R}^n} \left(E_P(r_{\mathbf{w}}) - \frac{\lambda}{2} \sigma_P^2(r_{\mathbf{w}}) - \frac{\theta}{2} \sigma_\mu^2(E(r_{\mathbf{w}})) \right). \tag{23}$$

A simple argument, in appendix, delivers the optimality condition

$$[\lambda \mathrm{Var}_P[\mathbf{r}] + \theta \mathrm{Var}_\mu[\mathrm{E}[\mathbf{r}]]] \,\widehat{\mathbf{w}} = \mathrm{E}_P[\mathbf{r} - r_f \mathbf{1}], \tag{24}$$

where:

- $\mathrm{Var}_P[\mathbf{r}] = [\sigma_P(r_i, r_j)]_{i,j=1}^n$ is the variance-covariance matrix of returns under P,
- $\mathrm{Var}_\mu[\mathrm{E}[\mathbf{r}]] = [\sigma_\mu(E(r_i), E(r_j))]_{i,j=1}^n$ is the variance-covariance matrix of expected returns under μ,
- $\mathrm{E}_P[\mathbf{r} - r_f \mathbf{1}] = [E_P(r_i - r_f)]_{i=1}^n$ is the vector of expected excess returns under P.

Thus, the solution to (23) is an ambiguity-adjusted mean-variance portfolio whose weights reflect uncertainty about expected returns as captured by $\mathrm{Var}_\mu[\mathrm{E}[\mathbf{r}]]$. A key feature of condition (24) is that it allows to make use of the vast body of research on mean-variance preferences developed for problems involving risk to analyze problems involving ambiguity. For example, also here the optimal investments are smooth functions of the taste parameters λ and θ and of the information parameters $\mathrm{E}_P[\mathbf{r} - \mathbf{1} r_f]$, $\mathrm{Var}_P[\mathbf{r}]$, and $\mathrm{Var}_\mu[\mathrm{E}[\mathbf{r}]]$. This is crucial for the comparative statics analysis. In fact, given the available information on assets' returns,[14] condition (24) allows to explicitly write the

[13] An empirical comparison of the two (a la Levy and Markowitz, 1979) goes beyond the scope of this paper and it is left for future research.

[14] That is, $\mathrm{E}_P[\mathbf{r} - \mathbf{1} r_f]$, $\mathrm{Var}_P[\mathbf{r}]$, and $\mathrm{Var}_\mu[\mathrm{E}[\mathbf{r}]]$.

optimal portfolios as functions of the uncertainty attitudes,[15] and so to study how different attitudes influence optimal holdings. Conversely, given the uncertainty attitudes of an agent, condition (24) allows to explicitly write the optimal portfolios as functions of the available information on assets' returns, and so to study how information and its quality affect optimal holdings.

Next, we study condition (24) for the case of a single ambiguous asset and for the case of one purely risky asset and one ambiguous asset. These two cases are important because they immediately contrast the optimal portfolio solution with and without ambiguity.

6.2 One Ambiguous Asset

If $n = 1$, then there is only one uncertain asset and (24) delivers

$$\widehat{w} = \frac{E_P(r) - r_f}{\lambda \sigma_P^2(r) + \theta \sigma_\mu^2(E(r))}. \tag{25}$$

If r is purely risky, i.e., $\sigma_\mu^2(E(r)) = 0$, then (25) reduces to the standard mean-variance Markowitz (1952) solution

$$\widehat{w} = \frac{E_P(r) - r_f}{\lambda \sigma_P^2(r)}. \tag{26}$$

Notice that ambiguity does not affect excess returns so that the difference between (25) and (26) lies in their denominators only. Specifically, an increase in $\theta \sigma_\mu^2(E(r))$ – that is, an increase in either ambiguity aversion θ or ambiguity in expectations $\sigma_\mu^2(E(r))$ – makes the ambiguous asset less desirable and increases the DM's demand for the risk-free asset (a flight-to-quality effect). As shown by Gollier (2011), this very intuitive result is not in general true for the smooth ambiguity preferences of KMM. Specifically, Gollier (2011) gives conditions under which more ambiguity aversion reduces the optimal level of exposure to uncertainty for KMM preferences; in our simplified setting no additional condition is needed.

6.3 One Purely Risky and One Ambiguous Assets

We now turn to the case of two uncertain assets with returns r_m and r_e in L^2, which we interpret as a domestic and a foreign security index, respectively. For this reason, we choose r_m to be purely risky for robust mean-variance preferences, i.e., $\sigma_\mu^2(E(r_m)) = 0$, and r_e to be ambiguous, i.e., $\sigma_\mu^2(E(r_e)) > 0$. The DM can now invest in a risk-free asset, with rate of return r_f, in a purely risky asset, with rate of return r_m, and in an ambiguous one, with rate of return r_e.

Here, condition (24) becomes

$$\lambda \begin{bmatrix} \sigma_P^2(r_m) & \sigma_P(r_m, r_e) \\ \sigma_P(r_m, r_e) & \sigma_P^2(r_e) \end{bmatrix} \begin{bmatrix} \widehat{w}_m \\ \widehat{w}_e \end{bmatrix} + \theta \begin{bmatrix} 0 & 0 \\ 0 & \sigma_\mu^2(E(r_e)) \end{bmatrix} \begin{bmatrix} \widehat{w}_m \\ \widehat{w}_e \end{bmatrix} = \begin{bmatrix} E_P(r_m) - r_f \\ E_P(r_e) - r_f \end{bmatrix},$$

that is,

$$E_P(r_m) - r_f = \widehat{w}_m \lambda \sigma_P^2(r_m) + \widehat{w}_e \lambda \sigma_P(r_m, r_e)$$

and

$$E_P(r_e) - r_f = \widehat{w}_m \lambda \sigma_P(r_m, r_e) + \widehat{w}_e \left[\lambda \sigma_P^2(r_e) + \theta \sigma_\mu^2(E(r_e))\right].$$

[15] That is, λ and θ.

For convenience, we set

$$\begin{aligned}
A &= E_P(r_e) - r_f, \\
B &= E_P(r_m) - r_f, \\
C &= \lambda \sigma_P^2(r_m), \\
D &= \lambda \sigma_P^2(r_e) + \theta \sigma_\mu^2(E(r_e)), \\
H &= \lambda \sigma_P(r_m, r_e).
\end{aligned}$$

Then, the optimal portfolio weights associated with the purely risky and the ambiguous assets are

$$\widehat{w}_m = \frac{BD - HA}{CD - H^2} \quad \text{and} \quad \widehat{w}_e = \frac{CA - HB}{CD - H^2}$$

and $CD - H^2 > 0$.[16] We are interested in determining how changes in the preference parameters affect the optimal amounts $\widehat{w}_m$ and $\widehat{w}_e$, as well as their ratio

$$\frac{\widehat{w}_m}{\widehat{w}_e} = \frac{BD - HA}{CA - HB}.$$

These quantities vary with μ, the prior probability over models, with θ, the parameter of ambiguity aversion, and with λ, the parameter of risk aversion. Before stating the results, we collect all the assumptions we make, most of which are non-triviality assumptions.

Omnibus Condition Suppose that $A > 0$ and $B > 0$, i.e., excess returns on uncertain assets are both positive, and $CA \neq HB$, i.e., the optimal investment in the ambiguous asset is nonzero.

The next lemma simplifies our analysis of the effects of varying ambiguity on the optimal portfolio.

Lemma 7 *Suppose that the Omnibus Condition holds and set $\sigma_\mu^2 = \sigma_\mu^2(E(r_e))$. Then,*

$$\frac{\partial(\widehat{w}_m/\widehat{w}_e)}{\partial \sigma_\mu^2} = \frac{\theta}{\sigma_\mu^2} \frac{\partial(\widehat{w}_m/\widehat{w}_e)}{\partial \theta}, \quad \frac{\partial \widehat{w}_m}{\partial \sigma_\mu^2} = \frac{\theta}{\sigma_\mu^2} \frac{\partial \widehat{w}_m}{\partial \theta}, \quad \text{and} \quad \frac{\partial \widehat{w}_e}{\partial \sigma_\mu^2} = \frac{\theta}{\sigma_\mu^2} \frac{\partial \widehat{w}_e}{\partial \theta}. \tag{27}$$

Notice that, since the optimal portfolio allocation varies with μ (the prior over models) only through σ_μ^2 (the variance of expected returns), changes in μ can be measured by taking derivatives with respect to σ_μ^2. Moreover, variations in $\widehat{w}_m$, $\widehat{w}_e$ and $\widehat{w}_m/\widehat{w}_e$ due to changes in σ_μ^2 and θ share the same sign. In view of this result, hereafter we only consider variations in θ and we generally refer to them as variations in ambiguity.

6.3.1 Changes in θ

We begin our comparative statics analysis by studying the effects of changes in θ; later, we study the effects of changes in λ.

For the comparative analysis it is convenient to decompose the excess return of the ambiguous asset by means of the ordinary least square coefficients. To this end, we project the excess return of the ambiguous asset over the excess return of the purely risky one, that is, we consider the solutions of

$$\min_{\alpha,\beta \in \mathbb{R}} \|(r_e - r_f) - (\alpha + \beta(r_m - r_f))\|.$$

As well known, they are given by

$$\beta_P(r_m, r_e) = \frac{\sigma_P(r_m, r_e)}{\sigma_P^2(r_m)} \tag{28}$$

[16] In fact, $CD - H^2 = \lambda^2 \left[\sigma_P^2(r_m)\sigma_P^2(r_e) - \sigma_P(r_m, r_e)^2\right] + \lambda\theta\sigma_P^2(r_m)\sigma_\mu^2(E(r_e))$ with the first summand nonnegative by the Cauchy-Schwartz inequality and the second one positive since all of its factors are positive.

and

$$\alpha_P(r_m, r_e) = E_P(r_e - r_f) - \beta_P(r_m, r_e) E_P(r_m - r_f). \tag{29}$$

The beta coefficient $\beta_P(r_m, r_e)$ measures the level of pure risk – as embodied in r_m – of asset r_e; in the finance jargon, $\beta_P(r_m, r_e)$ is a pure risk adjustment.[17] As a result, $\beta_P(r_m, r_e) E_P(r_m - r_f)$ is what index r_e is expected to earn/lose, net of r_f, given its level of pure risk sensitivity. The residual component $\alpha_P(r_m, r_e)$ of the expected excess return $E_P(r_e - r_f)$ is then what r_e is expected to earn/lose, net of r_f, given its level of uncertainty uncorrelated with pure risk.[18] Such uncertainty is specific to the ambiguous asset.

The next result shows that the sign of $\alpha_P(r_m, r_e)$ alone determines the effect of changes in ambiguity on the optimal proportion between risky and ambiguous holdings $\widehat{w}_m$ and $\widehat{w}_e$, as well as on the variation of $\widehat{w}_e$. On the other hand, also the sign of $\beta_P(r_m, r_e)$ becomes relevant to describe the variation of $\widehat{w}_m$.

Proposition 8 *Suppose the Omnibus Condition holds. Then,*

$$\operatorname{sgn} \frac{\partial}{\partial \theta}\left(\frac{\widehat{w}_m}{\widehat{w}_e}\right) = \operatorname{sgn} \alpha_P(r_m, r_e). \tag{30}$$

Moreover,

$$\operatorname{sgn} \frac{\partial \widehat{w}_e}{\partial \theta} = -\operatorname{sgn} \alpha_P(r_m, r_e) \quad \textit{and} \quad \operatorname{sgn} \frac{\partial \widehat{w}_m}{\partial \theta} = \operatorname{sgn} \alpha_P(r_m, r_e) \beta_P(r_m, r_e). \tag{31}$$

The sign of $\alpha_P(r_m, r_e)$ also governs the sign of optimal ambiguous holdings $\widehat{w}_e$.

Proposition 9 *Suppose the Omnibus Condition holds. Then*

$$\operatorname{sgn} \widehat{w}_e = \operatorname{sgn} \alpha_P(r_m, r_e). \tag{32}$$

In other words, the agent uses $\alpha_P(r_m, r_e)$ as a criterion to decide whether to take a long or short position in the ambiguous asset, that is, to decide in which side of the market of asset r_e to be. As anticipated in the introduction, when this term is positive, the ambiguous asset return offers compensation in excess of its risky component, which induces the agent to hold a long position on the asset (the symmetric argument holds in case of negativity).

In the jargon of investment practitioners, our agent "seeks the *alpha*" (buys/sells the ambiguous fund if $\alpha_P(r_m, r_e)$ is positive/negative) and he is aware that this extra return comes from the ambiguous nature of the investment, "therefore" he reduces exposure to ambiguity as ambiguity aversion rises.

Thanks to Propositions 8 and 9 we can show how variations in θ affect the optimal portfolio composition, both in relative and in absolute terms. Two possible cases arise.[19]

Case 1 If α is *positive*, then $\widehat{w}_e > 0$ and

$$\triangle\theta > 0 \implies \triangle\left(\frac{\widehat{w}_m}{\widehat{w}_e}\right) > 0 \text{ and } \triangle\widehat{w}_e < 0, \tag{33}$$

[17] Equation (50) in the appendix further illustrates this interpretation by showing that $\sigma_P(r_m, r_e) = \sigma_P(r_m, r_e - r_e^a)$. That is, only the purely risky projection $r_e - r_e^a$ of the ambiguous r_e enters its covariance with the purely risky r_m.

[18] To see why it is uncorrelated, notice that, by the Hilbert Decomposition Theorem,

$$r_e - r_f = A_P(r_m, r_e) + \beta_P(r_m, r_e)(r_m - r_f)$$

where $A_P(r_m, r_e) \in L^2$ is the part of the excess return $r_e - r_f$ that is uncorrelated with $r_m - r_f$ and $\alpha_P(r_m, r_e)$ is the expected value of the uncorrelated part $A_P(r_m, r_e)$, that is, $\alpha_P(r_m, r_e) = E_P(A_P(r_m, r_e))$.

[19] To ease notation, in what follows we just write α and β without the subscript $_P(r_m, r_e)$.

where $\triangle$ denotes a small variation.[20] Here, an increase in ambiguity aversion determines a higher ratio $\widehat{w}_m/\widehat{w}_e$ and a lower optimal $\widehat{w}_e$. Since $\widehat{w}_e$ is positive, a lower optimal $\widehat{w}_e$ corresponds to a lower exposure to ambiguity. Finally, the sign of the variation in $\widehat{w}_m$ coincides with the sign of β, that is, of the covariance $\sigma_P(r_m, r_e)$,

$$\triangle\theta > 0 \implies \triangle\widehat{w}_m \gtreqless 0 \text{ if and only if } \beta_P(r_m, r_e) \gtreqless 0.$$

When β is positive, higher ambiguity aversion thus results in a higher value of $\widehat{w}_m$ and a lower value of $\widehat{w}_e$. Otherwise, the ratio $\widehat{w}_m/\widehat{w}_e$ still increases, but only because the optimal amount $\widehat{w}_m$ decreases less than $\widehat{w}_e$.

Case 2 If α is *negative*, then $\widehat{w}_e < 0$ and

$$\triangle\theta > 0 \implies \triangle\left(\frac{\widehat{w}_m}{\widehat{w}_e}\right) < 0, \text{ with } \triangle\widehat{w}_e > 0 \text{ and } \triangle\widehat{w}_m < 0.$$

Here, an increase in ambiguity aversion determines a lower ratio $\widehat{w}_m/\widehat{w}_e$, a higher optimal $\widehat{w}_e$, and a lower optimal $\widehat{w}_m$ (it is easy to check that $\alpha < 0$ implies $\beta > 0$). Since $\widehat{w}_e$ is negative, a higher optimal $\widehat{w}_e$ corresponds again to lower exposure to ambiguity.

In sum, depending on the values of the technical risk measures α and β, we have different effects of variations in θ on the composition of the optimal portfolio. But, in any case our DM:

- goes long on r_e when α is positive and short otherwise (Proposition 9);
- reduces exposure to r_e as ambiguity increases (Proposition 8).

For example, in an international portfolio interpretation of our tripartite analysis, this means that higher ambiguity results in higher home bias.

6.3.2 Changes in λ

We now study the effects of changes in risk attitudes on the agent's assets holdings.

Proposition 10 *Suppose the Omnibus Condition holds. Then,*

$$\operatorname{sgn}\frac{\partial}{\partial\lambda}\left(\frac{\widehat{w}_m}{\widehat{w}_e}\right) = \operatorname{sgn}\frac{\partial\widehat{w}_e}{\partial\lambda} = -\operatorname{sgn}\alpha_P(r_m, r_e). \tag{34}$$

Variations in the ratio $\widehat{w}_m/\widehat{w}_e$ due to changes in λ thus have opposite sign relative to changes in θ. That is, changes in risk attitudes vary the portfolio proportional composition in the opposite direction to changes in ambiguity attitudes. This confirms the numerical findings of KMM p. 1878, in the general theoretical setting of this paper.

Moreover, variations in $\widehat{w}_m/\widehat{w}_e$ and $\widehat{w}_e$ share the same sign, which again is determined by $\alpha_P(r_m, r_e)$. As to $\partial\widehat{w}_m/\partial\lambda$, we have:

$$\frac{\partial\widehat{w}_m}{\partial\lambda} = \widehat{w}_e\frac{\partial\left(\widehat{w}_m/\widehat{w}_e\right)}{\partial\lambda} + \frac{\widehat{w}_m}{\widehat{w}_e}\frac{\partial\widehat{w}_e}{\partial\lambda}. \tag{35}$$

If $\widehat{w}_m$ and $\widehat{w}_e$ are both positive, then variations in $\widehat{w}_m$ have the same sign as those in $\widehat{w}_m/\widehat{w}_e$ and $\widehat{w}_e$. If $\widehat{w}_m$ and $\widehat{w}_e$ are not both positive, then the relations among variations in $\widehat{w}_m/\widehat{w}_e$ and $\widehat{w}_e$ and variations in $\widehat{w}_m$ are more complicated, but can be determined through (35).

[20] Not to be confounded with the set of probability measures Δ.

7 Conclusions

In this paper, we study how the classic Arrow-Pratt approximation of the certainty equivalent is altered by model uncertainty. Under the smooth ambiguity model of Klibanoff, Marinacci and Mukerji (2005), we find that the adjusted approximation contains an additional ambiguity premium that depends both on the degree of ambiguity aversion displayed by the DM and on the incidence of model uncertainty on the expectation of the prospect he is evaluating.

Then, we introduce robust mean-variance preferences, which are the counterpart for the smooth ambiguity model of standard mean-variance preferences for the expected utility model. We illustrate their scope by studying a static portfolio problem under model uncertainty. In the special case of a risk-free asset, a purely risky one, and an ambiguous one, the implied comparative statics of ambiguity aversion carries two noteworthy consequences: (i) portfolio rebalancing in response to ambiguity depends solely on the return *alpha* generated by the ambiguous asset in excess of the purely risky asset after correcting for *beta* (e.g., the DM takes a long/short position on the ambiguous asset if and only if *alpha* is positive/negative) and (ii) an increase in ambiguity aversion decreases the optimal exposure in the ambiguous asset.

Finally, we note that the analytical tractability of the enhanced approximation renders our model particularly fit for the study of puzzling investment behaviors including the home bias puzzle, the asset allocation puzzle, the equity premium puzzle, and the employer-stock ownership puzzle. The natural direction of development of this project is therefore the derivation of a robust CAPM, corresponding to robust Mean-Variance preferences, and its calibration.

A Proofs and Related Analysis

To prove the quadratic approximation (15) we need the following version of standard results on differentiation under the integral sign.

Lemma 11 *Let O be an open subset of $\mathbb{R}^N$, $(\Omega, \mathcal{F})$ be a measurable space, and $g : O \times \Omega \to \mathbb{R}$ be a function with the following properties:*

(a) for each $\mathbf{x} \in O$, $\omega \mapsto g(\mathbf{x}, \omega)$ is $\mathcal{F}$-measurable;

(b) for each $\omega \in \Omega$, $\mathbf{x} \mapsto g(\mathbf{x}, \omega)$ is twice continuously differentiable on O;

(c) the functions g, $\partial_j g$, and $\partial_{jk} g$ are bounded on $O \times \Omega$ for all $j, k \in \{1, 2, ..., N\}$.

Then,

(i) for each probability measure π on $\mathcal{F}$, the function defined on O by $G(\mathbf{x}) = \int g(\mathbf{x}, \omega)\, d\pi(\omega)$ is twice continuously differentiable;

(ii) the functions $\omega \mapsto \partial_j g(\mathbf{x}, \omega)$ and $\omega \mapsto \partial_{jk} g(\mathbf{x}, \omega)$ are measurable for all $\mathbf{x} \in O$, with

$$\partial_j G(\mathbf{x}) = \int \partial_j g(\mathbf{x}, \omega)\, d\pi(\omega) \tag{36a}$$

$$\partial_{jk} G(\mathbf{x}) = \int \partial_{jk} g(\mathbf{x}, \omega)\, d\pi(\omega) \tag{36b}$$

for all $\mathbf{x} \in O$ and $j, k \in \{1, 2, ..., N\}$.

A.1 Quadratic Approximation

Here we prove Proposition 3. We assume throughout the section that μ is a Borel probability measure with bounded support on Δ and the functions $u : I \to \mathbb{R}$ and $v : I \to \mathbb{R}$ are twice continuously differentiable, with $u', v' > 0$. We start with a simple lemma.

Lemma 12 *Let $\phi = v \circ u^{-1} : u(I) \to v(I)$ and $\psi = v^{-1} : v(I) \to I$. The functions ϕ and ψ are twice continuously differentiable on $u(\operatorname{int} I)$ and $v(\operatorname{int} I)$, respectively. In particular, there exist $\varepsilon > 0$ such that $[w-\varepsilon, w+\varepsilon] \subseteq \operatorname{int} I$ and $M > 1$ such that the absolute values of u, v, ϕ, and ψ – as well as their first and second derivatives – are bounded by M on $[w-\varepsilon, w+\varepsilon]$, $[w-\varepsilon, w+\varepsilon]$, $u([w-\varepsilon, w+\varepsilon])$, and $v([w-\varepsilon, w+\varepsilon])$, respectively. Finally, for all $x \in \operatorname{int} I$:*

$$\phi'(u(x)) = \frac{v'(x)}{u'(x)}, \quad \phi''(u(x)) = \frac{v''(x)}{u'(x)^2} - v'(x)\frac{u''(x)}{u'(x)^3},$$
$$\psi'(\phi(u(x))) = \frac{1}{v'(x)}, \quad \psi''(\phi(u(x))) = -\frac{v''(x)}{v'(x)^3}.$$

If $\mathbf{h} \in L^\infty(\Omega, \mathcal{F}, P)^N$ and $\mathbf{x} \in \mathbb{R}^N$, set $\mathbf{x} \cdot \mathbf{h} = \sum_{i=1}^N x_i h_i \in L^\infty$. Denote by $|\cdot|$ the Euclidean norm of $\mathbb{R}^N$. The next result yields Proposition 3 as a corollary.

Proposition 13 *Let μ be a Borel probability measure with bounded support on Δ and $u, v : I \to \mathbb{R}$ be twice continuously differentiable, with $u', v' > 0$. Then, for each $\mathbf{h} \in L^\infty(\Omega, \mathcal{F}, P)^N$ and all $\mathbf{x} \in \mathbb{R}^N$ such that $w + \mathbf{x} \cdot \mathbf{h} \in L^\infty(I)$,*

$$C(w + \mathbf{x} \cdot \mathbf{h}) = w + E_{\bar{Q}}(\mathbf{x} \cdot \mathbf{h}) - \frac{1}{2}\lambda_u(w)\sigma^2_{\bar{Q}}(\mathbf{x} \cdot \mathbf{h}) - \frac{1}{2}(\lambda_v(w) - \lambda_u(w))\sigma^2_\mu(E(\mathbf{x} \cdot \mathbf{h})) + o\left(|\mathbf{x}|^2\right) \quad (37)$$

as $\mathbf{x} \to \mathbf{0}$.

Proof Let $\mathbf{h} = (h_1, ..., h_N)$ and (wlog) assume that all the h_is are bounded.

Clearly, $\|\mathbf{x} \cdot \mathbf{h}\|_{\sup} \le \sum_{i=1}^N |x_i| \|h_i\|_{\sup}$ therefore there exists $\delta > 0$ such that $\|\mathbf{x} \cdot \mathbf{h}\|_{\sup} < \varepsilon$ (ε is the one obtained in Lemma 12) for all $\mathbf{x} \in (-\delta, \delta)^N$.[21] In particular, for all $\mathbf{x} \in (-\delta, \delta)^N$ and all $\omega \in \Omega$,

$$w - \varepsilon < w - \|\mathbf{x} \cdot \mathbf{h}\|_{\sup} \le w + \mathbf{x} \cdot \mathbf{h}(\omega) \le w + \|\mathbf{x} \cdot \mathbf{h}\|_{\sup} < w + \varepsilon$$

that is, $w + \mathbf{x} \cdot \mathbf{h}(\omega) \in (w-\varepsilon, w+\varepsilon)$, and so $w + \mathbf{x} \cdot \mathbf{h} \in L^\infty([w-\varepsilon, w+\varepsilon]) \subseteq L^\infty(I)$. Set $O = (-\delta, \delta)^N$.

Define

$$\begin{array}{rccc} g : & O \times \Omega & \to & \mathbb{R} \\ & (\mathbf{x}, \omega) & \mapsto & u(w + \mathbf{x} \cdot \mathbf{h}(\omega)) \end{array}.$$

Next we show that g satisfies assumptions (a), (b), (c) of Lemma 11.

(a) For each $\mathbf{x} \in O$, $\omega \mapsto g(\mathbf{x}, \omega)$ is $\mathcal{F}$-measurable; in fact, $\omega \mapsto w + \mathbf{x} \cdot \mathbf{h}(\omega) \in (w-\varepsilon, w+\varepsilon)$ is measurable and $u : (w-\varepsilon, w+\varepsilon) \to \mathbb{R}$ is continuous.

[21] Take for example $\delta = \varepsilon\left(\sum_{i=1}^N \|h_i\|_{\sup} + 1\right)^{-1}$. Then,

$$\|\mathbf{x} \cdot \mathbf{h}\|_{\sup} \le \sum_{i=1}^N |x_i| \|h_i\|_{\sup} \le \delta \sum_{i=1}^N \|h_i\|_{\sup} = \varepsilon\left(\sum_{i=1}^N \|h_i\|_{\sup}\right) / \left(\sum_{i=1}^N \|h_i\|_{\sup} + 1\right) < \varepsilon.$$

(b) For each $\omega \in \Omega$, $\mathbf{x} \mapsto g(\mathbf{x}, \omega)$ is twice continuously differentiable on O; in fact, given $\omega \in \Omega$, for all $\mathbf{x} \in O$ and all $j, k \in \{1, 2, ..., N\}$

$$\partial_j g(\mathbf{x}, \omega) = u'(w + \mathbf{x} \cdot \mathbf{h}(\omega)) h_j(\omega) \quad \text{and} \quad \partial_{jk} g(\mathbf{x}, \omega) = u''(w + \mathbf{x} \cdot \mathbf{h}(\omega)) h_j(\omega) h_k(\omega)$$

and the latter equation defines (for fixed ω, j, k) a continuous function on O.

(c) The functions g, $\partial_j g$, and $\partial_{jk} g$ are bounded on $O \times \Omega$ for all $j, k \in \{1, 2, ..., N\}$; in fact, given $j, k \in \{1, 2, ..., N\}$, for all $(\mathbf{x}, \omega) \in O \times \Omega$ (choosing M like in Lemma 12)

$$\begin{aligned} |g(\mathbf{x}, \omega)| &= |u(w + \mathbf{x} \cdot \mathbf{h}(\omega))| \leq M \\ |\partial_j g(\mathbf{x}, \omega)| &= |u'(w + \mathbf{x} \cdot \mathbf{h}(\omega))| |h_j(\omega)| \leq M \|h_j\|_{\sup} \\ |\partial_{jk} g(\mathbf{x}, \omega)| &= |u''(w + \mathbf{x} \cdot \mathbf{h}(\omega))| |h_j(\omega)| |h_k(\omega)| \leq M \|h_j\|_{\sup} \|h_k\|_{\sup} \end{aligned}$$

and indeed a uniform bound K for the supnorms on $O \times \Omega$ of all these functions can be chosen.

Then (by Lemma 11) for each $q \in \Delta$, the function defined on O by

$$G(\mathbf{x}, q) = \int g(\mathbf{x}, \omega) \, dQ(\omega) \quad \left(= \int_\Omega u(w + \mathbf{x} \cdot \mathbf{h}) \, q dP\right)$$

is twice continuously differentiable, the functions $\omega \mapsto \partial_j g(\mathbf{x}, \omega)$ and $\omega \mapsto \partial_{jk} g(\mathbf{x}, \omega)$ are measurable for all $\mathbf{x} \in O$, and

$$\begin{aligned} \partial_j G(\mathbf{x}, q) &= \int \partial_j g(\mathbf{x}, \omega) \, dQ(\omega) \quad \left(= \int_\Omega u'(w + \mathbf{x} \cdot \mathbf{h}) h_j q dP\right) \\ \partial_{jk} G(\mathbf{x}, q) &= \int \partial_{jk} g(\mathbf{x}, \omega) \, dQ(\omega) \quad \left(= \int_\Omega u''(w + \mathbf{x} \cdot \mathbf{h}) h_j h_k q dP\right) \end{aligned}$$

for all $\mathbf{x} \in O$ and $j, k \in \{1, 2, ..., N\}$.

Notice that, by point (c) above, for all $j, k \in \{1, 2, ..., N\}$ and all $(\mathbf{x}, q) \in O \times \Delta$,

$$|G(\mathbf{x}, q)| \leq K, \ |\partial_j G(\mathbf{x}, q)| \leq K, \text{ and } |\partial_{jk} G(\mathbf{x}, q)| \leq K$$

and that, by definition, $G(\mathbf{x}, q) \in u([w - \varepsilon, w - \varepsilon])$ where ϕ is twice continuously differentiable.

Set $f = \phi \circ G$. Next we show that the function $f : O \times \Delta \to \mathbb{R}$, with $(\mathbf{x}, q) \mapsto \phi(G(\mathbf{x}, q))$, satisfies assumptions (a), (b), and (c) of Lemma 11.

(a) For each $\mathbf{x} \in O$, $q \mapsto f(\mathbf{x}, q)$ is Borel measurable; in fact, given $\mathbf{x} \in O$, the function $f(\mathbf{x}, \cdot) = \phi(G(\mathbf{x}, \cdot)) = \phi(\langle u(w + \mathbf{x} \cdot \mathbf{h}), \cdot \rangle)$, being a composition of continuous functions, is continuous.[22]

(b) For each $q \in \Delta$, $\mathbf{x} \mapsto f(\mathbf{x}, q)$ is twice continuously differentiable on O; this follows from the fact that it is a composition of twice continuously differentiable functions, specifically, given $q \in \Delta$, for all $\mathbf{x} \in O$ and all $j, k \in \{1, 2, ..., N\}$

$$\begin{aligned} \partial_j f(\mathbf{x}, q) &= \phi'(G(\mathbf{x}, q)) \partial_j G(\mathbf{x}, q) \\ \partial_{jk} f(\mathbf{x}, q) &= \phi''(G(\mathbf{x}, q)) \partial_k G(\mathbf{x}, q) \partial_j G(\mathbf{x}, q) + \phi'(G(\mathbf{x}, q)) \partial_{jk} G(\mathbf{x}, q) \end{aligned}$$

and the latter equation defines (for fixed q, j, k) a continuous function on O.

(c) the functions f, $\partial_j f$, and $\partial_{jk} f$ are bounded on $O \times \Delta$ for all $j, k \in \{1, 2, ..., N\}$; in fact, given $j, k \in \{1, 2, ..., N\}$, for all $(\mathbf{x}, q) \in O \times \Delta$ (choosing M like in Lemma 12 and K as above)

$$\begin{aligned} |f(\mathbf{x}, q)| &= |\phi(G(\mathbf{x}, q))| \leq M \\ |\partial_j f(\mathbf{x}, q)| &= |\phi'(G(\mathbf{x}, q))| |\partial_j G(\mathbf{x}, q)| \leq MK \\ |\partial_{jk} f(\mathbf{x}, q)| &\leq |\phi''(G(\mathbf{x}, q))| |\partial_k G(\mathbf{x}, q)| |\partial_j G(\mathbf{x}, q)| + |\phi'(G(\mathbf{x}, q))| |\partial_{jk} G(\mathbf{x}, q)| \leq MK^2 + MK \end{aligned}$$

and the latter majorization holds term by term.

[22] The duality pairing $E_P(XY)$ in L^2 is denoted, as usual, by $\langle X, Y \rangle$ for all $X, Y \in L^2$.

By Lemma 11, the function defined on O by

$$F(\mathbf{x}) = \int f(\mathbf{x}, q)\, d\mu(q) \quad \left(= \int_\Delta \phi\left(\int_\Omega u(w + \mathbf{x}\cdot\mathbf{h})\, qdP\right) d\mu(q)\right)$$

is twice continuously differentiable, the functions $q \mapsto \partial_j f(\mathbf{x}, q)$ and $q \mapsto \partial_{jk} f(\mathbf{x}, q)$ are measurable for all $\mathbf{x} \in O$, and

$$\partial_j F(\mathbf{x}) = \int \partial_j f(\mathbf{x}, q)\, d\mu(q) \quad \text{and} \quad \partial_{jk} F(\mathbf{x}) = \int \partial_{jk} f(\mathbf{x}, q)\, d\mu(q)$$

for all $\mathbf{x} \in O$ and $j, k \in \{1, 2, ..., N\}$.

Finally, for all $\mathbf{x} \in O$ and all $q \in \Delta$, $G(\mathbf{x}, q) \in u([w - \varepsilon, w - \varepsilon])$ implies $f(\mathbf{x}, q) = \phi(G(\mathbf{x}, q)) \in v\left(u^{-1}(u([w - \varepsilon, w - \varepsilon]))\right) = v([w - \varepsilon, w - \varepsilon])$ and $F(\mathbf{x}) \in v([w - \varepsilon, w - \varepsilon])$. Thus

$$c(\mathbf{x}) = \psi \circ F(\mathbf{x}) \qquad \forall \mathbf{x} \in O$$

is well defined and twice continuously differentiable on $O = (-\delta, \delta)^N$. Its second order McLaurin expansion is

$$c(\mathbf{x}) = c(\mathbf{0}) + \nabla c(\mathbf{0})\, \mathbf{x} + \frac{1}{2} \mathbf{x}^\intercal \nabla^2 c(\mathbf{0})\, \mathbf{x} + o\left(|\mathbf{x}|^2\right). \tag{40}$$

Next we explicitly compute it using repeatedly the relations obtained above as well as those provided by Lemma 12. For all $\mathbf{x} \in O$,

$$\partial_j c(\mathbf{x}) = \psi'(F(\mathbf{x}))\, \partial_j F(\mathbf{x}) \quad \text{and} \quad \partial_{jk} c(\mathbf{x}) = \psi''(F(\mathbf{x}))\, \partial_k F(\mathbf{x})\, \partial_j F(\mathbf{x}) + \psi'(F(\mathbf{x}))\, \partial_{jk} F(\mathbf{x})$$

in particular for $\mathbf{x} = \mathbf{0}$,

$$\partial_j c(\mathbf{0}) = \psi'(F(\mathbf{0}))\, \partial_j F(\mathbf{0}) \quad \text{and} \quad \partial_{jk} c(\mathbf{0}) = \psi''(F(\mathbf{0}))\, \partial_k F(\mathbf{0})\, \partial_j F(\mathbf{0}) + \psi'(F(\mathbf{0}))\, \partial_{jk} F(\mathbf{0})$$

but $F(\mathbf{0}) = \phi(u(w))$ for all $j, k \in \{1, 2, ..., N\}$

$$\begin{aligned}
\partial_j F(\mathbf{0}) &= \int_\Delta \partial_j f(\mathbf{0}, q)\, d\mu(q) = \int_\Delta \phi'(G(\mathbf{0}, q))\, \partial_j G(\mathbf{0}, q)\, d\mu(q) \\
&= \int_\Delta \phi'(u(w)) \left(\int_\Omega u'(w)\, h_j qdP\right) d\mu(q) = \phi'(u(w))\, u'(w) \int_\Delta \left(\int_\Omega h_j qdP\right) d\mu(q) \\
&= v'(w) \int_\Delta \left(\int_\Omega h_j qdP\right) d\mu(q) = v'(w)\, E_{\bar{Q}}(h_j)
\end{aligned}$$

and

$$\begin{aligned}
\partial_{jk} F(\mathbf{0}) &= \int_\Delta \partial_{jk} f(\mathbf{0}, q)\, d\mu(q) = \int_\Delta \phi''(G(\mathbf{0}, q))\, \partial_k G(\mathbf{0}, q)\, \partial_j G(\mathbf{0}, q) + \phi'(G(\mathbf{0}, q))\, \partial_{jk} G(\mathbf{0}, q)\, d\mu(q) \\
&= \int_\Delta \phi''(G(\mathbf{0}, q))\, \partial_k G(\mathbf{0}, q)\, \partial_j G(\mathbf{0}, q)\, d\mu(q) + \int_\Delta \phi'(G(\mathbf{0}, q))\, \partial_{jk} G(\mathbf{0}, q)\, d\mu(q)
\end{aligned}$$

where the last equality is justified by the fact that both summands are continuous and bounded in q.[23] Now

$$\begin{aligned}
&\int_\Delta \phi''(G(\mathbf{0}, q))\, \partial_k G(\mathbf{0}, q)\, \partial_j G(\mathbf{0}, q)\, d\mu(q) = \int_\Delta \phi''(u(w)) \left(\int_\Omega u'(w)\, h_k qdP\right) \left(\int_\Omega u'(w)\, h_j qdP\right) d\mu(q) \\
&= \phi''(u(w))\, u'(w)^2 \int_\Delta \langle h_k, q\rangle \langle h_j, q\rangle\, d\mu(q) = \left(v''(w) - v'(w) \frac{u''(w)}{u'(w)}\right) E_\mu(\langle h_k, \cdot\rangle \langle h_j, \cdot\rangle)
\end{aligned}$$

[23] Boundedness was already observed. Continuity in q descends from $G(\mathbf{0}, q) = u(w)$, $\partial_i G(\mathbf{0}, q) = \langle u'(w)\, h_i, q\rangle$ for $i = j, k$, and $\partial_{jk} G(\mathbf{0}, q) = \langle u''(w)\, h_j h_k, q\rangle$.

and

$$\int_{\Delta}\phi'\left(G\left(\mathbf{0},q\right)\right)\partial_{jk}G\left(\mathbf{0},q\right)d\mu\left(q\right)=\int_{\Delta}\phi'\left(u\left(w\right)\right)\left(\int_{\Omega}u''\left(w\right)h_jh_kqdP\right)d\mu\left(q\right)=\frac{v'\left(w\right)}{u'\left(w\right)}u''\left(w\right)E_{\bar{Q}}\left(h_jh_k\right).$$

Finally

$$c\left(\mathbf{0}\right)=w \tag{41}$$

for all $j\in\{1,2,...,N\}$

$$\partial_j c\left(\mathbf{0}\right)=\psi'\left(F\left(\mathbf{0}\right)\right)\partial_j F\left(\mathbf{0}\right)=\psi'\left(\phi\left(u\left(w\right)\right)\right)v'\left(w\right)E_{\bar{Q}}\left(h_j\right)=\frac{1}{v'\left(w\right)}v'\left(w\right)E_{\bar{Q}}\left(h_j\right)=E_{\bar{Q}}\left(h_j\right)$$

so that

$$\nabla c\left(\mathbf{0}\right)\mathbf{x}=E_{\bar{Q}}\left(\mathbf{x}\cdot\mathbf{h}\right)\qquad\forall\mathbf{x}\in O \tag{42}$$

and, for all $j,k\in\{1,2,...,N\}$,

$$\begin{aligned}\partial_{jk}c\left(\mathbf{0}\right)&=\psi''\left(F\left(\mathbf{0}\right)\right)\partial_k F\left(\mathbf{0}\right)\partial_j F\left(\mathbf{0}\right)+\psi'\left(F\left(\mathbf{0}\right)\right)\partial_{jk}F\left(\mathbf{0}\right)=\psi''\left(\phi\left(u\left(w\right)\right)\right)v'\left(w\right)^2E_{\bar{Q}}\left(h_k\right)E_{\bar{Q}}\left(h_j\right)\\&\quad+\psi'\left(\phi\left(u\left(w\right)\right)\right)\left(\left(v''\left(w\right)-v'\left(w\right)\frac{u''\left(w\right)}{u'\left(w\right)}\right)E_{\mu}\left(\langle h_k,\cdot\rangle\langle h_j,\cdot\rangle\right)+\frac{v'\left(w\right)}{u'\left(w\right)}u''\left(w\right)E_{\bar{Q}}\left(h_jh_k\right)\right)\\&=-\frac{v''\left(w\right)}{v'\left(w\right)}E_{\bar{Q}}\left(h_k\right)E_{\bar{Q}}\left(h_j\right)+\left(\frac{v''\left(w\right)}{v'\left(w\right)}-\frac{u''\left(w\right)}{u'\left(w\right)}\right)E_{\mu}\left(\langle h_k,\cdot\rangle\langle h_j,\cdot\rangle\right)+\frac{u''\left(w\right)}{u'\left(w\right)}E_{\bar{Q}}\left(h_jh_k\right)\\&=\lambda_v\left(w\right)E_{\bar{Q}}\left(h_j\right)E_{\bar{Q}}\left(h_k\right)+\left(\lambda_u\left(w\right)-\lambda_v\left(w\right)\right)E_{\mu}\left(\langle h_j,\cdot\rangle\langle h_k,\cdot\rangle\right)-\lambda_u\left(w\right)E_{\bar{Q}}\left(h_jh_k\right)\\&=-\left[\lambda_u\left(w\right)\sigma_{\bar{Q}}\left(h_j,h_k\right)+\left(\lambda_v\left(w\right)-\lambda_u\left(w\right)\right)\sigma_{\mu}\left(\langle h_j,\cdot\rangle,\langle h_k,\cdot\rangle\right)\right].\end{aligned}$$

denoting by $\Sigma_{\bar{Q}}$ and Σ_μ the variance-covariance matrixes $\left[\sigma_{\bar{Q}}\left(h_j,h_k\right)\right]_{j,k=1}^N$ and $\left[\sigma_\mu\left(\langle h_j,\cdot\rangle,\langle h_k,\cdot\rangle\right)\right]_{j,k=1}^N$

$$\nabla^2 c\left(\mathbf{0}\right)=-\left[\lambda_u\left(w\right)\Sigma_{\bar{Q}}+\left(\lambda_v\left(w\right)-\lambda_u\left(w\right)\right)\Sigma_\mu\right]$$

and

$$\frac{1}{2}\mathbf{x}^\intercal\nabla^2 c\left(\mathbf{0}\right)\mathbf{x}=-\frac{1}{2}\lambda_u\left(w\right)\sigma^2_{\bar{Q}}\left(\mathbf{x}\cdot\mathbf{h}\right)-\frac{1}{2}\left(\lambda_v\left(w\right)-\lambda_u\left(w\right)\right)\sigma^2_\mu\left(E\left(\mathbf{x}\cdot\mathbf{h}\right)\right)\qquad\forall\mathbf{x}\in O. \tag{43}$$

This concludes the proof since replacement of (41), (42), and (43) into (40) delivers

$$c\left(\mathbf{x}\right)=w+E_{\bar{Q}}\left(\mathbf{x}\cdot\mathbf{h}\right)-\frac{1}{2}\lambda_u\left(w\right)\sigma^2_{\bar{Q}}\left(\mathbf{x}\cdot\mathbf{h}\right)-\frac{1}{2}\left(\lambda_v\left(w\right)-\lambda_u\left(w\right)\right)\sigma^2_\mu\left(E\left(\mathbf{x}\cdot\mathbf{h}\right)\right)+o\left(|\mathbf{x}|^2\right)$$

as $\mathbf{x}\to\mathbf{0}$, and $c\left(\mathbf{x}\right)=v^{-1}\left(F\left(\mathbf{x}\right)\right)=C\left(w+\mathbf{x}\cdot\mathbf{h}\right)$ for all $\mathbf{x}\in O$. If $\mathbf{x}\in\mathbb{R}^N\setminus O$ and $w+\mathbf{x}\cdot\mathbf{h}\in L^\infty\left(I\right)$ just set

$$o\left(|\mathbf{x}|^2\right)=C\left(w+\mathbf{x}\cdot\mathbf{h}\right)-\left[w+E_{\bar{Q}}\left(\mathbf{x}\cdot\mathbf{h}\right)-\frac{1}{2}\lambda_u\left(w\right)\sigma^2_{\bar{Q}}\left(\mathbf{x}\cdot\mathbf{h}\right)-\frac{1}{2}\left(\lambda_v\left(w\right)-\lambda_u\left(w\right)\right)\sigma^2_\mu\left(E\left(\mathbf{x}\cdot\mathbf{h}\right)\right)\right]$$

the property of vanishing faster than $|\mathbf{x}|^2$ as $\mathbf{x}\to\mathbf{0}$ has no bite there. ■

Proof of Proposition 3 We first consider the case in which $\mathcal{F}$ is finite. Let $\mathcal{A}=\{A_1,...,A_N\}$ be the family of atoms of $\mathcal{F}$ that are assigned a positive probability by P. Then $\{1_{A_1},...,1_{A_N}\}$ is a base for L^2 and, setting $\mathbf{h}=(1_{A_1},...,1_{A_N})$, the map

$$\gamma:\mathbf{x}\mapsto\sum_{i=1}^N x_i1_{A_i}=\mathbf{x}\cdot\mathbf{h}$$

is a norm isomorphism between $\mathbb{R}^N$ and L^2.[24] In particular, choosing $\delta > 0$ as in the proof of Proposition 13, for all $x \in \gamma\left((-\delta,\delta)^N\right) = \left\{\mathbf{x}\cdot\mathbf{h} : \mathbf{x} \in (-\delta,\delta)^N\right\}$

$$C(w+x) = w + E_{\bar{Q}}(x) - \frac{1}{2}\lambda_u(w)\,\sigma^2_{\bar{Q}}(x) - \frac{1}{2}(\lambda_v(w) - \lambda_u(w))\,\sigma^2_\mu(E(x)) + o\left(|\mathbf{x}|^2\right) \tag{44}$$

as $\mathbf{x} \to \mathbf{0}$ in $\mathbb{R}^N$. Set

$$R_2(x) = C(w+x) - \left[w + E_{\bar{Q}}(x) - \frac{1}{2}\lambda_u(w)\,\sigma^2_{\bar{Q}}(x) - \frac{1}{2}(\lambda_v(w) - \lambda_u(w))\,\sigma^2_\mu(E(x))\right]$$

for all $x \in \gamma\left((-\delta,\delta)^N\right)$, $\underline{p} = \min_{i=1,\dots,N} P(A_i)$, and $\overline{p} = \max_{i=1,\dots,N} P(A_i)$, then

$$\underline{p}\sum_{i=1}^N x_i^2 \le \|x\|^2 \le \overline{p}\sum_{i=1}^N x_i^2.$$

Now, if x_n is a (nonzero) vanishing sequence in $\gamma\left((-\delta,\delta)^N\right)$,

$$\frac{|R_2(x_n)|}{\overline{p}\sum_{i=1}^N (x_n)_i^2} \le \frac{|R_2(x_n)|}{\|x_n\|^2} \le \frac{|R_2(x_n)|}{\underline{p}\sum_{i=1}^N (x_n)_i^2}$$

and by (44) the three sequences above vanish as $n \to \infty$. That is, $R_2(x) = o\left(\|x\|^2\right)$ since $\gamma\left((-\delta,\delta)^N\right)$ is a neighborhood of 0 in L^2.

In the general case, let $h \in L^\infty(\Omega,\mathcal{F},P)$. By Proposition 13, for all $t \in \mathbb{R}$ such that $w + th \in L^\infty(I)$,

$$C(w+th) = w + E_{\bar{Q}}(th) - \frac{1}{2}\lambda_u(w)\,\sigma^2_{\bar{Q}}(th) - \frac{1}{2}(\lambda_v(w) - \lambda_u(w))\,\sigma^2_\mu(E(th)) + o\left(t^2\right)$$

as $t \to 0$. That is, setting, for all $t \in \mathbb{R}$ such that $w + th \in L^\infty(I)$,

$$R_2(th) = C(w+th) - \left[w + E_{\bar{Q}}(th) - \frac{1}{2}\lambda_u(w)\,\sigma^2_{\bar{Q}}(th) - \frac{1}{2}(\lambda_v(w) - \lambda_u(w))\,\sigma^2_\mu(E(th))\right] \tag{45}$$

it results $\lim_{t\to 0} R_2(th)/t^2 = 0$. Moreover, the assumption $w + h \in L^\infty(I)$ guarantees that we can consider $t = 1$ in (45), that is

$$C(w+h) = w + E_{\bar{Q}}(h) - \frac{1}{2}\lambda_u(w)\,\sigma^2_{\bar{Q}}(h) - \frac{1}{2}(\lambda_v(w) - \lambda_u(w))\,\sigma^2_\mu(E(h)) + R_2(h)$$

as wanted. ■

A.2 Approximately Unambiguous Prospects

Proof of Proposition 5 (iii) trivially implies (i), which in turn implies (ii). To complete the proof, we show that (ii) implies (iii). First notice that for all $h \in L^2$ and all $t \in \mathbb{R}$

$$\sigma^2_\mu(E(th)) = \sigma^2_\mu(tE(h)) = t^2\sigma^2_\mu(E(h)). \tag{46}$$

[24] Finite dimensionality guarantees that Δ is bounded, therefore – in this case – the assumption that the support of μ is bounded is automatically satisfied.

Therefore, $\sigma_\mu^2(E(h)) = R_2(h)$ implies

$$0 = \lim_{t\to 0} \frac{\sigma_\mu^2(E(th))}{t^2} = \sigma_\mu^2(E(h)),$$

It remains to show that $\sigma_\mu^2(E(h)) = 0$ implies that $\langle h, \cdot\rangle$ is constant on suppμ. If $h \in L^2$ and $\sigma_\mu^2(E(h)) = 0$, then

$$\langle h, q\rangle = E_\mu(\langle h, q\rangle) = E_{\bar{Q}}(h)$$

for μ-almost all $q \in \Delta$. If, per contra, there exists $q^* \in$ suppμ such that $\langle h, q^*\rangle \neq E_{\bar{Q}}(h)$, then the continuity of $\langle h, \cdot\rangle$ on Δ implies the existence of an open subset G of Δ such that $\langle h, q\rangle \neq E_{\bar{Q}}(h)$ for all $q \in G$. But $G \cap \text{supp}\mu \neq \emptyset$, and so $\mu(G) > 0$, a contradiction. We conclude that $\langle h, q\rangle = E_{\bar{Q}}(h)$ for all $q \in$ suppμ. ■

Proof of Proposition 6 (i) trivially implies (ii) and (iii) implies (i). Next we show that (ii) implies (iii). For all $h \in B$, set

$$F(h) = C(w+h) - \left[w + E_{\bar{Q}}(h) - \frac{1}{2}\lambda_u(w)\sigma_{\bar{Q}}^2(h)\right]$$

and

$$G(h) = C(w+h) - \left[w + E_{\bar{Q}}(h) - \frac{1}{2}\lambda_u(w)\sigma_{\bar{Q}}^2(h) - \frac{1}{2}(\lambda_v(w) - \lambda_u(w))\sigma_\mu^2(E(h))\right]$$

by (19) and Proposition 3, $\lim_{t\to 0} F(th)/t^2 = \lim_{t\to 0} G(th)/t^2 = 0$. Therefore, for all $h \in B$, setting $k = 2(\lambda_v(w) - \lambda_u(w))^{-1}$,

$$\sigma_\mu^2(E(h)) = \lim_{t\to 0}\frac{\sigma_\mu^2(E(th))}{t^2} = \lim_{t\to 0} k\frac{G(th) - F(th)}{t^2} = 0.$$

Since B is absorbing in L^∞, for all $A \in \mathcal{F}$ there is $\varepsilon = \varepsilon_A > 0$ such that $\varepsilon 1_A \in B$, thus

$$\sigma_\mu^2(E(1_A)) = \frac{\sigma_\mu^2(E(\varepsilon 1_A))}{\varepsilon^2} = 0$$

and 1_A is approximately unambiguous for all $A \in \mathcal{F}$. Then, by Proposition 5, for all $A \in \mathcal{F}$,

$$Q(A) = E_Q(1_A) = E_{Q'}(1_A) = Q'(A) \qquad \forall Q, Q' \in \text{supp}\mu,$$

that is, $Q = Q'$. ■

A.3 Orthogonal Decomposition

By Proposition 5, the collection M of all approximately unambiguous prospects is easily seen to be a closed linear subspace of L^2. Clearly, M contains all risk-free (constant) prospects, and its orthogonal complement $M^\perp$ is a closed subspace of $\left\{h \in L^2 : E_P(h) = 0\right\}$. The Hilbert Decomposition Theorem then implies the following decomposition of each prospect.

Proposition 14 *For each prospect $h \in L^2$ there exist unique $h^c \in \mathbb{R}$, $h^g \in M$ with $E_P(h^g) = 0$, and $h^a \in M^\perp$ such that*

$$h = h^c + h^g + h^a. \tag{47}$$

Moreover,

$$\sigma_P^2(h) = \sigma_P^2(h^g) + \sigma_P^2(h^a) \tag{48}$$

and

$$\sigma_\mu^2(E(h)) = \sigma_\mu^2(E(h^a)). \tag{49}$$

In particular, h is approximately unambiguous if and only if $h^a = 0$, and it is risk-free if and only if $h^g = h^a = 0$.

Proof Let $h \in L^2$. By definition of $M^\perp$ and by the Hilbert Decomposition Theorem, decomposition (47) and its uniqueness are easily checked. In particular, $E_P(h) = h^c$. Moreover, the maps $h \mapsto h^g \in M$ and $h \mapsto h^a \in M^\perp$ are linear and continuous operators.

Since h^g and h^a are orthogonal and have zero mean, then

$$\begin{aligned}\sigma_P^2(h) &= \|h - E_P(h)\|^2 = \|E_P(h) + h^g + h^a - E_P(h)\|^2 = \|h^g + h^a\|^2 \\ &= \|h^g\|^2 + \|h^a\|^2 = \sigma_P^2(h^g) + \sigma_P^2(h^a),\end{aligned}$$

which proves (48).

Finally, observe that $\langle h, \cdot\rangle = \langle h^c + h^g, \cdot\rangle + \langle h^a, \cdot\rangle$ and $h^c + h^g \in M$ implies that $\langle h^c + h^g, \cdot\rangle$ is μ-almost surely constant, thus $\sigma_\mu^2(\langle h, \cdot\rangle) = \sigma_\mu^2(\langle h^a, \cdot\rangle)$. The rest is trivial. ∎

In view of decomposition (47), the constant h^c – which is equal to $E_P(h)$ – can be interpreted as the risk-free component of h. Indeed, $h = h^c$ if and only if $\sigma_P^2(h) = 0$. The next component, h^g, can be viewed as a fair gamble because $h^g \in M$ and $E_P(h^g) = 0$. The sum $h^c + h^g$ of the first two components is approximately unambiguous. In contrast, (48) and (49) show that the "residual" component h^a reflects both risk and ambiguity in pure variability terms (net of any level effect factored out by the constant h^c).

In our portfolio exercise of Section 6.3.1, this implies the orthogonal decompositions $r_m = E_P(r_m) + r_m^g$ and $r_e = E_P(r_e) + r_e^g + r_e^a$ of the purely risky and the ambiguous assets. In particular, $E_P(r_e) + r_e^g = r_e - r_e^a$ is the purely risky component of the ambiguous asset. Simple algebra shows that

$$\beta_P(r_m, r_e) = \frac{\sigma_P(r_m, r_e - r_e^a)}{\sigma_P^2(r_m)} \tag{50}$$

which confirms the interpretation of β as a measure of the risk sensitivity of r_e relative to r_m.

A.4 Portfolio

Derivation of (24). Setting

$$\begin{aligned}&\mathbf{m} = [E_P(r_1 - r_f), ..., E_P(r_n - r_f)]^\intercal, \quad \Sigma_P = [\sigma_P(r_i, r_j)]_{i,j=1}^n \\ &\Sigma_\mu = [\sigma_\mu(E(r_i), E(r_j))]_{i,j=1}^n, \quad \Xi = \lambda\Sigma_P + \theta\Sigma_\mu\end{aligned}$$

(23) becomes

$$\max_{\mathbf{w}\in\mathbb{R}^n}\left\{r_f + \mathbf{w}\cdot\mathbf{m} - \frac{\lambda}{2}\mathbf{w}^\intercal\Sigma_P\mathbf{w} - \frac{\theta}{2}\mathbf{w}^\intercal\Sigma_\mu\mathbf{w}\right\}$$

which is equivalent to

$$\max_{\mathbf{w}\in\mathbb{R}^n}\left(\mathbf{w}\cdot\mathbf{m} - \frac{1}{2}\mathbf{w}^\intercal\Xi\mathbf{w}\right)$$

so that the optimal solution $\widehat{\mathbf{w}}$ satisfies $\Xi\widehat{\mathbf{w}} = \mathbf{m}$. ∎

Proof of Lemma 7 Since $\partial D/\partial\sigma_\mu^2 = \theta$ and $\partial D/\partial\theta = \sigma_\mu^2$, simple algebra shows that

$$\frac{\partial(\widehat{w}_m/\widehat{w}_e)}{\partial\sigma_\mu^2} = \theta\frac{B}{CA - HB} \quad \text{and} \quad \frac{\partial(\widehat{w}_m/\widehat{w}_e)}{\partial\theta} = \sigma_\mu^2\frac{B}{CA - HB} \tag{51}$$

and this implies the first equality in (27). Analogously,

$$\frac{\partial\widehat{w}_m}{\partial\sigma_\mu^2} = \theta\frac{(AC - BH)H}{(CD - H^2)^2} \quad \text{and} \quad \frac{\partial\widehat{w}_e}{\partial\sigma_\mu^2} = -\theta\frac{(CA - HB)C}{(CD - H^2)^2}, \tag{52}$$

$$\frac{\partial\widehat{w}_m}{\partial\theta} = \sigma_\mu^2\frac{(AC - BH)H}{(CD - H^2)^2} \quad \text{and} \quad \frac{\partial\widehat{w}_e}{\partial\theta} = -\sigma_\mu^2\frac{(CA - HB)C}{(CD - H^2)^2} \tag{53}$$

which imply the other equalities in (27). ■

Proof of Proposition 8 By definition

$$\beta_P(r_m, r_e) = \frac{H}{C} \quad \text{and} \quad \alpha_P(r_m, r_e) = E_P(r_e) - r_f - \beta_P(r_m, r_e)(E_P(r_m) - r_f) = A - \frac{H}{C}B \quad (54)$$

while, by (51),

$$\frac{\partial(\widehat{w}_m/\widehat{w}_e)}{\partial\theta} = \sigma_\mu^2 \frac{B}{CA - HB}$$

thus

$$\operatorname{sgn} \frac{\partial}{\partial\theta}\left(\frac{\widehat{w}_m}{\widehat{w}_e}\right) = \operatorname{sgn}(CA - HB) = \operatorname{sgn}\left(\frac{CA - HB}{C}\right) = \operatorname{sgn}\alpha_P(r_m, r_e)$$

that is (30) holds.

Moreover, by (53) and (54)

$$\operatorname{sgn} \frac{\partial\widehat{w}_e}{\partial\theta} = \operatorname{sgn}(HB - CA) = -\operatorname{sgn}\alpha_P(r_m, r_e) \quad (55)$$

that is the first part of (31) holds. Moreover, (53) and (54) again deliver

$$\frac{\partial\widehat{w}_m}{\partial\theta} = -\left(-\sigma_\mu^2 \frac{(CA - HB)}{(CD - H^2)^2} C\right)\frac{H}{C} = -\frac{\partial\widehat{w}_e}{\partial\theta}\beta_P(r_m, r_e)$$

which together with (55) delivers the second part of (31). ■

Proof of Proposition 9 By (54)

$$\alpha_P(r_m, r_e) = A - \frac{H}{C}B$$

but

$$\widehat{w}_e = \frac{CA - HB}{CD - H^2}$$

which together with $CD - H^2 > 0$ and $C > 0$ delivers (32). ■

Proof of Proposition 10 Direct computation delivers

$$\lambda \frac{\partial(\widehat{w}_m/\widehat{w}_e)}{\partial\lambda} = -\theta \frac{\partial(\widehat{w}_m/\widehat{w}_e)}{\partial\theta}$$

which together with Proposition 8 determines the sign of $\partial(\widehat{w}_m/\widehat{w}_e)/\partial\lambda$. Moreover,

$$\frac{\partial\widehat{w}_e}{\partial\lambda}\frac{\partial(\widehat{w}_m/\widehat{w}_e)}{\partial\lambda} = -\frac{\theta\sigma_\mu^2 B}{\lambda^2(CD - H^2)^2}\left(H^2 - C\left(D - \theta\sigma_\mu^2\right)\right)$$

hence

$$\frac{\partial\widehat{w}_e}{\partial\lambda}\frac{\partial(\widehat{w}_m/\widehat{w}_e)}{\partial\lambda} > 0 \Leftrightarrow H^2 - C\left(D - \theta\sigma_\mu^2\right) < 0 \Leftrightarrow \sigma_P(r_m, r_e)^2 < \sigma_P^2(r_m)\,\sigma_P^2(r_e).$$

By the Cauchy-Schwartz inequality the above relation fails if and only if

$$r_e - E_P(r_e) = k(r_m - E_P(r_m))$$

for some $k \in \mathbb{R}$, but this would imply that r_e is approximately unambiguous, which is absurd. ■

A.5 Proofs of Section 2

Proof of Lemma 1 Consider the duality inclusion ι :supp$\mu \to \left(L^2\right)^*$ given by $q \mapsto \langle \cdot, q\rangle$. For all $X \in L^2$, the composition $X \circ \iota$:supp$\mu \to \mathbb{R}$ given by $q \mapsto \langle X, \iota(q)\rangle = \langle X, q\rangle$ is (norm) continuous and hence Borel measurable on suppμ, that is, ι is weak* measurable (see, e.g., Aliprantis and Border, 2006, Ch. 11.9). The range of ι is norm bounded since suppμ is norm bounded and ι is an isometry. Therefore, ι is Gelfand integrable over suppμ (*ibidem*, Cor. 11.53). In particular, there exists a unique $\bar{q} \in L^2$ such that

$$\langle X, \bar{q}\rangle = \int_{\mathrm{supp}\mu} \langle X, \iota(q)\rangle \, d\mu(q), \qquad \forall X \in L^2. \tag{56}$$

By (56) it readily follows that $\bar{q} \in \Delta$. ■

Proof of Lemma 2 Let $f \in L^\infty(I)$ and set $a = \operatorname{essinf} f$ and $b = \operatorname{esssup} f$. There exists $A \in \mathcal{F}$ with $P(A) = 1$ such that $f(A) \subseteq [a,b] \subseteq I$. Since u is increasing and continuous $u(f(\omega)) \in [u(a), u(b)] \subseteq u(I)$ for all $\omega \in A$. Moreover, $u \circ f_{|A} : A \to [u(a), u(b)]$ is measurable since $f_{|A}$ is measurable and u is continuous. Therefore, $u(f)$ is defined P-almost surely on Ω, measurable, and $u(a) \leq u(f) \leq u(b)$ P-almost surely. It follows that $\langle u(f), \cdot\rangle : \Delta \to \mathbb{R}$, with $q \mapsto \int_\Omega u(f)\, q dP$, is norm continuous, affine, with range in $[u(a), u(b)] \subseteq u(I)$. Therefore, $\phi \circ \langle u(f), \cdot\rangle : \Delta \to \mathbb{R}$, with $q \mapsto \phi\left(\int_\Omega u(f)\, q dP\right)$, is well defined, norm continuous, with range in $[\phi(u(a)), \phi(u(b))] \subseteq \phi(u(I))$. Therefore, $V(f) = \int_\Delta \phi \circ \langle u(f), \cdot\rangle \, d\mu \in [\phi(u(a)), \phi(u(b))] \subseteq \phi(u(I))$ is well defined.

The first part of this proof yields $V(L^\infty(I)) \subseteq \phi(u(I))$. Conversely, if $z = \phi(u(x))$ for some $x \in I$, then $x1_\Omega \in L^\infty(I)$ and $V(L^\infty(I)) \ni V(x1_\Omega) = \int_\Delta \phi \circ \langle u(x), \cdot\rangle \, d\mu = \phi(u(x)) = z$. ■

References

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

[2] K.J. Arrow, The theory of risk aversion, in *Aspects of the Theory of Risk Bearing*, Yrjo Jahnsson Foundation, Helsinki, 28-44, 1965.

[3] W. Bonnice and V.L. Klee, The generation of convex hulls, *Mathematische Annalen*, 152, 1-29, 1963.

[4] P. Boyle, L. Garlappi, R. Uppal, and T. Wang, Keynes meets Markowitz: The tradeoff between familiarity and diversification, *Management Science*, 58, 253-272, 2012.

[5] N. Canner, N.G. Mankiw, and D. N. Weil, An asset allocation puzzle, *American Economic Review*, 87, 181-191, 1997.

[6] J. Caskey, Information in equity markets with ambiguity-averse investors, *Review of Financial Studies*, 22, 3595-3627, 2009.

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

[8] Z. Chen and L. Epstein, Ambiguity, risk, and asset returns in continuous time, *Econometrica*, 70, 1403-1443, 2002.

[9] H. Chen, N. Ju, and J. Miao, Dynamic asset allocation with ambiguous return predictability, mimeo, 2011.

[10] S. Collard, S. Mukerji, K. Sheppard and J-M. Tallon, Ambiguity and the historical equity premium, Economics Dept. WP 550, Oxford University, 2011.

[11] D. Easley and M. O'Hara, Liquidity and valuation in an uncertain world, *Journal of Financial Economics*, 97, 1-11, 2010a.

[12] D. Easley and M. O'Hara, Microstructure and ambiguity, *Journal of Finance*, 65, 1817-1846, 2010b.

[13] L. Epstein and J. Miao, A two-person dynamic equilibrium under ambiguity, *Journal of Economic Dynamics and Control*, 27, 1253-1288, 2003.

[14] K.R. French and J.M. Poterba, Investor diversification and international equity markets, *American Economic Review P&P*, 81, 222-226, 1991.

[15] L. Garlappi, R. Uppal, and T. Wang, Portfolio selection with parameter and model uncertainty: a multi-prior approach, *Review of Financial Studies*, 20, 41-81, 2007.

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

[17] C. Gollier, Portfolio choices and asset prices: The comparative statics of ambiguity aversion, *Review of Economic Studies*, 78, 1329-1344, 2011.

[18] L. P. Hansen and T. Sargent, Fragile beliefs and the price of uncertainty, *Quantitative Economics*, 1, 129-162, 2010.

[19] G. Huberman, Familiarity breeds investment, *Review of Financial Studies*, 14, 659-680, 2001.

[20] P. Illeditsch, Ambiguous information, portfolio inertia, and excess volatility, *Journal of Finance*, 66, 2213-2247, 2011.

[21] Y. Izhakian and S. Benninga, The uncertainty premium in an ambiguous economy, *Quarterly Journal of Finance*, 1, 323-354, 2011.

[22] I. Jewitt and S. Mukerji, Ordering ambiguous acts, Economics Dept. WP 553, Oxford University, 2011.

[23] N. Ju and J. Miao, Ambiguity, learning and asset returns, *Econometrica*, forthcoming, 2012.

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

[25] H. Levy and H.M. Markowitz, Approximating expected utility by a function of mean and variance, *American Economic Review*, 69, 308-317, 1979.

[26] F. Maccheroni, M. Marinacci and D. Ruffino, Does uncertainty vanish in the small? The smooth ambiguity case, IGIER WP 391, 2011.

[27] F. Maccheroni, M. Marinacci, A. Rustichini, and M. Taboga, Portfolio selection with monotone mean-variance preferences, *Mathematical Finance*, 19, 487-521, 2009.

[28] H.M. Markowitz, Portfolio selection, *Journal of Finance*, 7, 77-91, 1952.

[29] Miao, J., Ambiguity, risk and portfolio choice under incomplete information, *Annals of Economics and Finance*, 10, 257-279, 2009.

[30] R.F. Nau, Uncertainty aversion with second-order utilities and probabilities, *Management Science*, 52, 136-145, 2006.

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

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

[33] C. Skiadas, Smooth ambiguity toward small risks and continuous-time recursive utility, mimeo, 2009.

[34] M. Taboga, Portfolio selection with two-stage preferences, *Finance Research Letters*, 2, 152-164, 2005.

[35] J. von Neumann and O. Morgenstern, *The Theory of Games and Economic Behavior*, 2nd ed., Princeton University Press, Princeton, 1947.

[36] M. Weitzman, Subjective expectations and asset-return puzzles, *American Economic Review*, 97, 1102-1130, 2007.