INNOCENZO GASPARINI INSTITUTE FOR ECONOMIC RESEARCH

Institutional Members: CEPR, NBER and Università Bocconi

WORKING PAPER SERIES

Alpha as Ambiguity: Robust Mean-Variance Portfolio Analysis

Fabio Maccheroni, Massimo Marinacci and Doriana Ruffino

Working Paper n. 373

This Version: October 2010

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

Alpha as Ambiguity: Robust Mean-Variance Portfolio Analysis*

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

[a]Università Bocconi
[b]University of Minnesota

October 2010

Abstract

We derive the analogue of the classic Arrow-Pratt approximation of the certainty equivalent under model uncertainty as defined by the smooth model of decision making under ambiguity of Klibanoff, Marinacci and Mukerji (2005). We study its scope via a portfolio allocation exercise that delivers a tractable mean-variance model adjusted for model uncertainty. In a problem with a risk-free asset, a risky asset, and an ambiguous asset, we find that portfolio rebalancing in response to higher model uncertainty *only* depends on the ambiguous asset's *alpha*, setting the performance of the risky asset as benchmark. In addition, the portfolios recommended by our model are not systematically conservative on the share held in the ambiguous asset: indeed, in general, it is not true that greater ambiguity reduces the optimal demand for the ambiguous asset. The analytical tractability of the enhanced Arrow-Pratt approximation renders our model especially well suited for calibration exercises aimed at exploring the consequences of ambiguity aversion 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 an expected utility maximizer with utility u and wealth w considers a (self-financing) investment h, the Arrow-Pratt approximation of his certainty equivalent for the resulting uncertain prospect $w+h$ is

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

where P is the probabilistic model the agent uses to describe 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 $\sigma_P^2\left(h\right)$ of h with respect to P. This relation between risk and variance is a central pillar of risk management. In particular, the coefficient $\lambda_u\left(w\right)=-u''\left(w\right)/u'\left(w\right)$ that links volatility and risk premium is determined by the agent's risk aversion at w. The practical merit

*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, Giulia Brancaccio, Simone Cerreia-Vioglio, Sujoy Mukerji, and Alessandro Sbuelz 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.

of (1) is in providing the foundation for the mean-variance preference model, where a prospect f is evaluated through

$$U(f) = E_P(f) - \frac{1}{2}\lambda\sigma_P^2(f), \tag{2}$$

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 *ambiguity* (or *Knightian uncertainty*), that is, to the case when the agent is uncertain about the true probabilistic model P that governs the occurrence of the different states. If only risk is present and so the agent fully relies on a single probabilistic model P, then the certainty equivalent $c(w+h,P)$ of $w+h$, that is, 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

$$\begin{aligned} C(w+h) &= v^{-1}(E_\mu(v(c(w+h)))) & (4)\\ &= v^{-1}\left(E_\mu\left(v\left(u^{-1}(E(u(w+h)))\right)\right)\right), & (5)\end{aligned}$$

where $c(w+h)$ is the random variable that associates $c(w+h,Q)$ to each model Q in Δ.

If $u = v$, it can be shown that

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

where $\bar{Q}$ is the reduced probability $\int Qd\mu(Q)$ induced by the prior μ. The certainty equivalent (4) thus reduces to (3) where the reduced probability $P = \bar{Q}$ is considered. A similar reduction holds also when the support of the prior μ is a singleton, that is, μ concentrates on a single probabilistic model. However, if u and v differ (a taste feature) and if the support of μ is nonsingleton (an information feature), this reduction no longer holds – the analysis of ambiguity cannot be reduced to risk only – 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 defined by the smooth ambiguity certainty equivalent of Klibanoff, Marinacci and Mukerji (2005), henceforth abbreviated KMM. Specifically, Theorem 4 shows that:

$$C(w+h) \approx w + E_{\bar{Q}}(h) - \frac{1}{2}\lambda_u(w)\sigma_{\bar{Q}}^2(h) - \frac{1}{2}(\lambda_v(w) - \lambda_u(w))\sigma_\mu^2(E(h)), \tag{6}$$

where $\bar{Q} = \int Qd\mu(Q)$ is the reduced probability induced by the prior μ, and $E(h) : \Delta \to \mathbb{R}$ is a 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 (6) – that is novel relative to (1). In other words, model uncertainty renders volatile the return $E(h)$ of h, thereby affecting agent's certainty equivalent.

In Section 4 we show that unambiguous prospects are characterized by the condition $\sigma_\mu^2(E(h)) = 0$, that is, they are unaffected by model uncertainty. For this special class of prospects, approximation (6) reduces to its classic counterpart (1).

The quadratic approximation (6) 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 by imposing $\bar{Q} = P$, we obtain the following natural and parsimonious extension

$$U(f) = E_P(f) - \frac{\lambda}{2}\sigma_P^2(f) - \frac{\theta}{2}\sigma_\mu^2(E(f)) \tag{7}$$

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 model. The taste parameters λ and θ represent negative attitudes toward risk and ambiguity, respectively. Higher values of these parameters correspond to stronger negative attitudes. The information parameter μ determines the variances $\sigma_{\bar{Q}}^2(h)$ and $\sigma_\mu^2(E(h))$ that measure the risk and model uncertainty perceived in the valuation of prospect h. 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 (7) via a portfolio allocation exercise. In particular, we study a tripartite portfolio problem with a risk-free asset, a risky but unambiguous 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 the optimal portfolio rebalancing in response to changes in ambiguity depends *only* on the risk adjusted performance α – the *alpha*, see eq. (32) – of the ambiguous asset with respect to the risky one. In particular, if the ambiguous asset has a *positive* α, then the proportional holdings of the ambiguous asset decrease as ambiguity rises. Conversely, if the ambiguous asset has a *negative* α, then the proportional holdings of the ambiguous asset increase as ambiguity rises.

As a result, the portfolios recommended by our model are not systematically conservative with respect to the ambiguous asset holdings. The intuition is simple. Following the standard practice of considering the risky asset as a benchmark, α captures the residual performance of the ambiguous asset that cannot be explained in terms of risk only. Thus, an ambiguity averse agent that observes a positive α attributes the, otherwise unexplained, augmented return to an unmeasurable increase in uncertainty (ambiguity) that drives him away from the ambiguous asset. Analogously, a negative α is associated with a diminution of uncertainty that, in turn, makes the ambiguous asset more desirable.

Some fundamental asset allocation problems feature a natural tripartite structure. This is the case for international portfolio allocation problems with domestic "riskless" 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) that affect the assets. We expect this to be often the case.[1] In the last subsection of the paper we construct a concrete example of our tripartite problem. We choose 3-month Treasury bills, the S&P500 index and the MSCI World ex-U.S. index to represent the risk-free asset, the risky asset and the ambiguous asset, respectively. We compute the alpha gains of the ambiguous asset (with respect to the risky asset) before and concurrently with the six recessions ensued between 1971 and today. During expansions, α is positive and our model's implications are consistent with the evidence of home bias. In two instances, during the 1981-1982 and 2007-2009 recessions, α is negative. In this case, our analysis would have advised to tilt the portfolio toward the foreign (ambiguous) index.

Related Works Our work is related to recent papers by Nau (2006), Izhakian and Benninga (2008), and Skiadas (2009) that construct an approximated ambiguity premium to detect and measure the agent's ambiguity aversion in a two-stage smooth model. Their analyses, however, are very different from ours and, more important, they do not study portfolio decisions.

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

Our formulation of the portfolio selection problem shares some features with the models of Epstein and Miao (2003), Taboga (2005), Gollier (2009) and Boyle, Garlappi, Uppal and Wang (2010). In particular, Taboga (2005) proposes a model of portfolio selection based on a two-stage valuation procedure to disentangle ambiguity and ambiguity aversion. Gollier (2009) investigates the comparative statics of more ambiguity aversion in a static two-asset portfolio problem. He exhibits sufficient conditions to guarantee that, ceteris paribus, an increase in ambiguity aversion reduces the optimal exposure to uncertainty. Epstein and Miao (2003) use a recursive multiple priors models to study the home bias, while Boyle, Garlappi, Uppal and Wang (2010) 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 (6) 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 (2009) study the effects of "ambiguous" information on investors' market trades and valuations. Easley and O'Hara (2009) and (2010) 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, Collard, Mukerji, Sheppard and Tallon (2009) match the historical equity premium, while Ju and Miao (2010) generate a variety of dynamic asset pricing phenomena observed in the data.

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 $\sigma_P(X,Y)$ and $\rho_P(X,Y)$ the covariance and correlation coefficients

$$\sigma_P(X,Y) = E_P[(X - E_P(X))(Y - E_P(Y))] \quad \text{and} \quad \rho_P(X,Y) = \frac{\sigma_P(X,Y)}{\sigma_P(X)\sigma_P(Y)}$$

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 a general theorem of Bonnice and Klee (1963) (see their Theorem 4.3), we have the following existence result.

Proposition 1 *Given a Borel probability measure μ on Δ with bounded support,*[2] *there exists a unique $\bar{q} \in \Delta$ such that*

$$\int_{\Omega} X(\omega)\,\bar{q}(\omega)\,dP(\omega) = \int_{\Delta}\left(\int_{\Omega} X(\omega)\,q(\omega)\,dP(\omega)\right)d\mu(q), \qquad \forall X \in L^2. \tag{8}$$

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 $\mathcal{F}$, (8) delivers

$$\bar{Q}(A) = \int_{\Delta} Q(A)\,d\mu(Q), \qquad \forall A \in \mathcal{F}, \tag{9}$$

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, (9) suggests a natural interpretation of $\bar{Q}$ in terms of reduction of compound lotteries. For example, if $\mathrm{supp}\mu = \{Q_1, ..., Q_n\}$ is finite and $\mu(Q_i) = \mu_i$ for $i = 1, ..., n$, then (9) 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.

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. That is, DMs who rank prospects through the functional $V : L^\infty(I) \to \mathbb{R}$ defined by

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

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,

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

In the monetary setting of the present paper, where outcomes are amounts of money and acts 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 (10) as

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

and so (11) as

$$C(f) = v^{-1}\left(\int_{\Delta} v\left(u^{-1}\left(\int_{\Omega} u(f(\omega))\,q(\omega)\,dP(\omega)\right)\right)d\mu(q)\right), \qquad \forall f \in L^\infty(I). \tag{13}$$

[2] 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 $\mathrm{supp}\mu$.

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 the prior μ through second order acts, whose outcomes depend 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 monetary acts $f : \Omega \to \mathbb{R}$. 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

$$\lambda_\phi(u(w)) = \frac{1}{u'(w)}\left(\lambda_v(w) - \lambda_u(w)\right) \tag{14}$$

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

Ambiguity neutrality is modelled by $\phi(x) = x$, that is, $\lambda_v = \lambda_u$, while absence of ambiguity is modelled by a trivial μ with singleton support (i.e., a Dirac measure). In both cases the criterion (10) reduces to subjective expected utility, though in one case the reduction originates in a taste component – a neutral attitude, under which the two sources of uncertainty "linearly" combine via the reduction (9) – while in the other case it originates in an information component (absence of a source of uncertainty, i.e., model uncertainty).

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,

$$C(w+h) = v^{-1}\left(\int_\Delta v\left(u^{-1}\left(\int_\Omega u(w+h)\, q dP\right)\right) d\mu(q)\right). \tag{15}$$

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

$$E(h) : q \mapsto \int_\Omega hq dP$$

is continuous and bounded on Δ, and hence belongs to $L^\infty(\Delta, \mathcal{B}, \mu)$. Its variance 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 $E(h)$, due to the model uncertainty that the DM perceives. Thus, higher values of $\sigma_\mu^2(E(h))$ correspond to a higher incidence of model uncertainty in the valuation of $E(h)$.

We can now state the second order approximation of the certainty equivalent (15). We start by considering the special case when $\mathcal{F}$ is finite (e.g., because Ω is finite).

Proposition 3 *Let μ be a Borel probability measure on Δ and $u, v : I \to \mathbb{R}$ be twice continuously differentiable with $u', v' > 0$. If $\mathcal{F}$ is finite, then*

$$C(w+h) = w + E_{\bar{Q}}(h) - \frac{1}{2}\lambda_u(w)\, \sigma_{\bar{Q}}^2(h) - \frac{1}{2}\left(\lambda_v(w) - \lambda_u(w)\right) \sigma_\mu^2(E(h)) + o\left(\|h\|^2\right) \tag{16}$$

as $h \to 0$ in L^2.

Thus, the sign and magnitude of the effect of perceived ambiguity on the certainty equivalent depend on the difference $\lambda_v(w) - \lambda_u(w)$. In the rest of the paper we will focus on the ambiguity averse case, that is, $\lambda_v(w) - \lambda_u(w) \geq 0$. Similarly, we will consider the risk averse case $\lambda_u(w) \geq 0$.

Notice that the first three components of the right hand side of (16) correspond to the Arrow-Pratt approximation of $u^{-1}\left(\int u(w+h)\,d\bar{Q}\right)$. To the contrary, the fourth component represents model uncertainty. In particular, (16) reduces to (1) under either the taste condition $\lambda_v(w) = \lambda_u(w)$ – the DM is neutral to ambiguity – or the information condition $\sigma^2_\mu(E(h)) = 0$ – model uncertainty does not affect the expectation of h.

We now extend the approximation to a general state space Ω. In this case, the Peano remainder is in the sense of Gateaux, as clarified by (18). This approximation will suffice for our purposes.

Theorem 4 *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,*

$$C(w+h) = w + E_{\bar{Q}}(h) - \frac{1}{2}\lambda_u(w)\,\sigma^2_{\bar{Q}}(h) - \frac{1}{2}(\lambda_v(w) - \lambda_u(w))\,\sigma^2_\mu(E(h)) + R_2(h), \tag{17}$$

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

$$\lim_{t \to 0} \frac{R_2(th)}{t^2} = 0. \tag{18}$$

4 Approximately Unambiguous Prospects

Since our setting encounters both risk and ambiguity, it is important to identify unambiguous prospects, i.e., prospects that are not affected by the presence of ambiguity. These are the prospects considered in classic risk theory.

Definition 5 *A prospect $h \in L^2$ is* unambiguous *if $F_{h,Q} = F_{h,Q'}$ for all $Q, Q' \in \mathrm{supp}\mu$.*

In words, a prospect is unambiguous if its distribution F_h is invariant across all models in the support of μ, that is, according to all models that a DM with prior μ deems possible.[3] Equivalently, $h \in L^2$ is unambiguous if and only if $Q(A) = Q'(A)$ for all events A that belong to the σ-algebra generated by h. Hence, the notion of unambiguous prospects that we use here is a special case, for the present setup, of the notion proposed by Ghirardato, Maccheroni, and Marinacci (2004).[4]

When restricted to the family of unambiguous prospects, the certainty equivalent (13) coincides with $u^{-1}\left(\int u(f)\,d\bar{Q}\right)$. Classic risk theory can thus be viewed as the special case in which all prospects are unambiguous. In particular, for unambiguous prospects our quadratic approximation (17) coincides with the Arrow-Pratt one. The class of prospects for which this is the case is, however, larger than that of unambiguous prospects. Given our focus on the quadratic approximation, it is important to identify this larger class of prospects.

Definition 6 *A prospect $h \in L^2$ is* (first order) approximately unambiguous *if $E_Q(h) = E_{Q'}(h)$ for all $Q, Q' \in \mathrm{supp}\mu$.*

In words, a prospect is approximately unambiguous if it has the same *first moment* according to all models that the DM deems possible. For an unambiguous prospect, *all moments* coincide and, hence, unambiguous prospects are approximately unambiguous.[5] The converse is false.

[3] Recall that $F_{h,Q}(x) = Q(h \leq x)$ for all $x \in \mathbb{R}$.

[4] 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 (2009) (see their Theorem 21).

[5] It is easy to see that a bounded h is unambiguous if and only if $E_Q(h^n) = E_{Q'}(h^n)$ for all $Q, Q' \in \mathrm{supp}\mu$ and for all $n \geq 1$, that is, all of its moments coincide on $\mathrm{supp}\mu$.

Proposition 7 *For a prospect $h \in L^2$, the following properties are equivalent:*

(i) *h is approximately unambiguous;*

(ii) $\sigma_\mu^2 (E(h)) = R_2(h)$;[6]

(iii) $\sigma_\mu^2 (E(h)) = 0$.

In words, a prospect h is approximately unambiguous if and only if its variance σ_μ^2 is zero. When $w + h \in L^\infty (I)$ this amounts to say that, in evaluating such prospect, the DM is indistinguishable in the second order approximation

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

from a subjective expected utility DM with utility u and subjective beliefs given by the reduced probability measure $\bar{Q}$ induced by μ on Ω. This second equivalence motivates the "approximately unambiguous" terminology and shows that approximately unambiguous prospects form the larger class of prospects for which our quadratic approximation (17) coincides with that of Arrow-Pratt.

Approximately unambiguous prospects are thus equivalent to risky ones in our approximation. The imprecision that is peculiar to any approximation – the cost of its higher analytical tractability – results here in an approximate notion of unambiguous prospects.

That said, next we show that the approximate and the "full-fledged" notions of unambiguous prospects become equivalent when they apply to all prospects. This is an important consistency check of the approximate notion.

Proposition 8 *The following facts are equivalent:*

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

(ii) *all prospects in L^2 are unambiguous;*

(iii) $\mu = \delta_{\bar{Q}}$;

(iv) $\sigma_\mu^2 (E(1_A)) = 0$ *for all* $A \in \mathcal{F}$.

The equivalence of (i) and (ii) is the announced equivalence among approximate and full-fledged unambiguity when they apply to all prospects. In turn, they are equivalent to (iii), that is, to a trivial μ. In this case the DM is subjective expected utility with utility u and subjective beliefs given by the reduced probability measure $\bar{Q}$ induced by μ on Ω. The equivalence with (iv) shows that for this to happen is actually enough that the bets 1_A be approximately unambiguous.

Next we give a further characterization of the absence of ambiguity.

Proposition 9 *Let $\lambda_v(w) - \lambda_u(w) \neq 0$. There is an absorbing*[7] *subset B of L^∞ such that $w + h \in L^\infty(I)$ and*

$$C(w+h) = w + E_{\bar{Q}}(h) - \frac{1}{2}\lambda_u(w)\,\sigma_{\bar{Q}}^2(h) + R_2(h) \qquad \forall h \in B \tag{19}$$

if and only if the prior μ is trivial, i.e., $\mu = \delta_{\bar{Q}}$.[8]

[6] Here $R_2(h)$ is as in (18).

[7] A subset B of L^∞ 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 open ball that contains the origin is absorbing.

[8] In turn, this is equivalent to conditions (i), (ii) and (iv) of the previous result.

This result shows that the *only* case in which our approximation (17) coincides with that of Arrow-Pratt for all "small" prospects (i.e., the prospects in B) is when there was no ambiguity to start with, that is, when μ is trivial. 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, ambiguity may keep having second order effects in the quadratic approximation of arbitrarily small prospects. Ambiguity may never fade away, even approximately, for arbitrarily "small" prospects (the counterparts in our more general setting of what are sometimes called "small risks" in risk theory).

We close this section with a useful decomposition. In view of Proposition 7, the collection M of all approximately unambiguous prospects is easily seen to be a closed linear subspace of L^2 given by

$$M = \left\{ h \in L^2 : \sigma_\mu^2 \left(E \left(h \right) \right) = 0 \right\}.$$

As a consequence, M contains all *risk-free* (constant) prospects, and its orthogonal complement $M^\perp$ is a closed subspace of $\left\{ h \in L^2 : E_P \left(h \right) = 0 \right\}$. The classic Projection Theorem then implies the following decomposition of each prospect.

Proposition 10 *For each prospect $h \in L^2$ there exist unique $h_c \in \mathbb{R}$, $h_g \in M$ with $E_P \left(h_g \right) = 0$, and $h_a \in M^\perp$ such that*

$$h = h_c + h_g + h_a. \tag{20}$$

Moreover,

$$\sigma_P^2 \left(h \right) = \sigma_P^2 \left(h_g \right) + \sigma_P^2 \left(h_a \right) \tag{21}$$

and

$$\sigma_\mu^2 \left(E \left(h \right) \right) = \sigma_\mu^2 \left(E \left(h_a \right) \right). \tag{22}$$

In particular, h is approximately unambiguous if and only if $h_a = 0$, and it is risk-free if and only if $h_g = h_a = 0$.

In view of decomposition (20), the constant h_c – which is equal to $E_P \left(h \right)$ – can be interpreted as the risk-free component of h. Indeed, $h = h_c$ if and only if $\sigma_P^2 \left(h \right) = 0$. The next component, h_g, can be viewed as a fair gamble because $h_g \in M$ and $E_P \left(h_g \right) = 0$. The sum $h_c + h_g$ of the first two components is approximately unambiguous. In contrast, (21) and (22) 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).

Summing up, h_c is the risk-free component of h, h_g is the (fair) gamble component of h, and h_a is the residual ambiguous component of h.

5 Robust Mean-Variance Preferences

In the next section we will apply the quadratic approximation (17) to a portfolio allocation problem. To this end, we first generalize standard mean-variance preferences to account for model uncertainty. Specifically, we consider a DM who ranks prospects h in L^2 through the robust mean-variance functional $C : L^2 \to \mathbb{R}$ given by

$$C \left(h \right) = E_{\bar{Q}} \left(h \right) - \frac{\lambda}{2} \sigma_{\bar{Q}}^2 \left(h \right) - \frac{\theta}{2} \sigma_\mu^2 \left(E \left(h \right) \right), \qquad \forall h \in L^2, \tag{23}$$

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 (9).

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 (17), which shows that (23) can be viewed as a local approximation of a KMM preference functional (13) at a constant w such that $\lambda = \lambda_u \left(w \right)$ and $\theta = \lambda_v \left(w \right) - \lambda_u \left(w \right)$. 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 $\sigma^2_{\bar{Q}}(h)$ and $\sigma^2_{\mu}(E(h))$ that measure the risk and model uncertainty that the DM perceives in the valuation of prospect h. 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, (23) takes the form

$$C(h) = E_P(h) - \frac{\lambda}{2}\sigma^2_P(h) - \frac{\theta}{2}\sigma^2_{\mu}(E(h)), \qquad \forall h \in L^2, \tag{24}$$

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

$$C(h) = E_P(h) - \frac{\lambda}{2}\sigma^2_P(h) \tag{25}$$

for prospect h. 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. Similarly, when the taste condition $\theta = 0$ holds, the standard mean-variance evaluation (25) holds for all prospects h.[9]

Like standard mean-variance preferences, our robust mean-variance preferences (24) nicely separate taste parameters, λ and θ, and uncertainty measures, $\sigma^2_P(h)$ and $\sigma^2_{\mu}(E(h))$. 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 (24) 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.

Couple of final remarks. First, thanks to the decomposition (20) the robust mean-variance preference functional (24) can be written as

$$C(h) = C(h_c + h_g) - \frac{\lambda}{2}\sigma^2_P(h_a) - \frac{\theta}{2}\sigma^2_{\mu}(E(h_a)), \qquad \forall h \in L^2$$

where $C(h_c + h_g)$ is the standard Arrow-Pratt functional (25) of the approximately unambiguous prospect $h_c + h_g$. The decomposition (20) thus provides a further view of the robust mean-variance functional as an "augmented" mean-variance functional.

Second, 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.

[9] Notice that condition $\sigma^2_{\mu}(E(h)) = 0$ is about prospect h and only affects it, while condition $\theta = 0$ is a general taste condition that affects all prospects.

6 The Portfolio Allocation Problem

In this section we apply the newly obtained robust mean-variance preferences to a portfolio allocation problem. In particular, we consider a portfolio of three assets: a risk-free asset, a 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 straightforward application of our setting with domestic Treasury bonds viewed as the risk-free asset, other domestic assets viewed as the risky assets, and foreign assets viewed as the ambiguous assets. This will be our motivating example in this section.

6.1 The General Setting

We consider the one-period allocation 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 $r_i \in L^2$. The $(n \times 1)$ vector of returns on the first n assets is then 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

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

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, $\mathbf{w}$ can be optimally chosen in $\mathbb{R}^n$ and the portfolio problem can be written as

$$\max_{\mathbf{w} \in \mathbb{R}^n} C(r_{\mathbf{w}}) = \max_{\mathbf{w} \in \mathbb{R}^n} \left(E_P(r_{\mathbf{w}}) - \frac{\lambda}{2} \sigma_P^2(r_{\mathbf{w}}) - \frac{\theta}{2} \sigma_\mu^2(E(r_{\mathbf{w}})) \right). \tag{26}$$

Straightforward computation delivers the optimality condition

$$[\lambda \mathrm{Var}_P[\mathbf{r}] + \theta \mathrm{Var}_\mu[\mathrm{E}[\mathbf{r}]]] \, \widehat{\mathbf{w}} = \mathrm{E}_P[\mathbf{r} - \mathbf{1} r_f], \tag{27}$$

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,
- $\mathrm{E}_P[\mathbf{r} - \mathbf{1} r_f] = [E_P(r_i - r_f)]_{i=1}^n$ is the vector of expected excess returns under P.

A key feature of condition (27) is that it allows us to make use of the vast body of research on mean-variance preferences developed for problems involving risk to analyze problems involving ambiguity. On the other hand, the ability to take advantage of conventional risk theory is a feature of the KMM model.

6.2 The Case of One Ambiguous Asset

If $n = 1$, then there is only one uncertain asset and (27) delivers

$$\widehat{w} = \frac{E_P(r) - r_f}{\lambda \sigma_P^2(r) + \theta \sigma_\mu^2(E(r))}. \tag{28}$$

If either r is approximately unambiguous – i.e., $\sigma_\mu^2(E(r)) = 0$ – or the DM is ambiguity neutral – i.e., $\theta = 0$ – then (30) reduces to the standard mean-variance Markowitz (1952) solution

$$\widehat{w} = \frac{E_P(r) - r_f}{\lambda \sigma_P^2(r)}. \tag{29}$$

Ambiguity does not affect excess returns and the difference between (28) and (29) lies in their denominators only. Specifically, an increase in $\theta\sigma_\mu^2(E(r))$ – that is, an increase in either perceived ambiguity $\sigma_\mu^2(E(r))$ or ambiguity aversion θ – leads to an increase in the DM's demand for the risk-free asset.

By decomposing r orthogonally, as in (20), we obtain $r = r_c + r_g + r_a$. This allows us to rewrite (28) as

$$\widehat{w} = \frac{E_P(r) - r_f}{\lambda \sigma_P^2(r) + \theta\sigma_\mu^2(E(r_a))}. \tag{30}$$

where in the argument of $E(\cdot)$ in place of r appears, more accurately, its ambiguity component r_a. In this sharper version the approximate unambiguous nature of r is modelled by $r_a = 0$, while $\sigma_\mu^2(E(r_a))$ models changes in perceived ambiguity.

6.3 The Case of One Risky and One Ambiguous Assets

We now turn to the case of two assets with uncertain returns r_m and r_e in L^2, interpreted as representing a home security index and a foreign security index respectively. For this reason, we choose r_m to be approximately unambiguous (i.e., purely risky for robust mean-variance preferences) and r_e to be approximately ambiguous, according to the definitions of Section 4. The DM can now invest in a risk-free asset r_f, in a purely risky one r_m, and in an uncertain one r_e.

Here condition (27) 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

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

Set

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

The optimal portfolio weights associated with the risky and the ambiguous assets are

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

respectively, provided $CD \neq H^2$. We are interested in how changes in the preference parameters affect the optimal amounts $\widehat{w}_m$ and $\widehat{w}_e$, as well as their ratio

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

This ratio varies with μ, the prior probability over models, with θ, the ambiguity aversion parameter, as well as with λ, the risk aversion parameter. Before stating the results, we collect all the assumptions that we will need (most of them are non-triviality assumptions).

Omnibus Condition Suppose $CD \neq H^2$, i.e., the portfolio problem admits a unique solution; $CA \neq HB$, i.e., the ratio of optimal portfolio weights is well-defined; $A > 0$ and $B > 0$, i.e., excess returns on uncertain assets are both positive.

The next simple lemma simplifies the analysis of variations in model uncertainty.

Lemma 11 *Suppose the Omnibus Condition holds and set $\sigma_\mu^2 = \sigma_\mu^2 (E(r_e))$. Then,*

$$\frac{\partial (\widehat{w}_m / \widehat{w}_e)}{\partial \sigma_\mu^2} = \frac{\theta}{\sigma_\mu^2} \frac{\partial (\widehat{w}_m / \widehat{w}_e)}{\partial \theta}, \quad \frac{\partial \widehat{w}_m}{\partial \sigma_\mu^2} = \frac{\theta}{\sigma_\mu^2} \frac{\partial \widehat{w}_m}{\partial \theta} \quad \text{and} \quad \frac{\partial \widehat{w}_e}{\partial \sigma_\mu^2} = \frac{\theta}{\sigma_\mu^2} \frac{\partial \widehat{w}_e}{\partial \theta}. \tag{31}$$

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 will only consider variations in θ and we will generally refer to them as variations in ambiguity.

The orthogonal decompositions of r_m and r_e are $r_m = r_{m,c} + r_{m,g}$ and $r_e = r_{e,c} + r_{e,g} + r_{e,a}$. In particular, *inter alia* we have

$$\sigma_P^2 (r_m) = \sigma_P^2 (r_{m,g}), \quad \sigma_\mu^2 (E(r_e)) = \sigma_\mu^2 (E(r_{e,a})), \quad \sigma_P (r_m, r_e) = \sigma_P (r_{m,g}, r_{e,g}),$$

and so we can write more accurately

$$C = \lambda \sigma_P^2 (r_{m,g}), \quad D = \lambda \sigma_P^2 (r_e) + \theta \sigma_\mu^2 (E(r_{e,a})), \quad H = \lambda \sigma_P (r_{m,g}, r_{e,g}).$$

Here the variance $\sigma_\mu^2 (E(r_{e,a}))$ only depends on the ambiguous component of r_e and the covariance $\sigma_P (r_{m,g}, r_{e,g})$ reflects the correlation among the risk components of r_m and r_e. Though to ease notation we will not mention explicitly these orthogonal components, the variances and covariances that will appear in the portfolio analysis can be read in their terms.

6.3.1 Uncorrelated Asset Returns

We first consider the case in which r_m and r_e are uncorrelated, that is, $\sigma_P (r_m, r_e) = 0$ (which implies $H = 0$). In this case, the optimal sums $\widehat{w}_m$ and $\widehat{w}_e$ are similar to those derived in Section 6.2 with a single uncertain asset. In fact, we have

$$\widehat{w}_m = \frac{E_P (r_m) - r_f}{\lambda \sigma_P^2 (r_m)} \quad \text{and} \quad \widehat{w}_e = \frac{E_P (r_e) - r_f}{\lambda \sigma_P^2 (r_e) + \theta \sigma_\mu^2 (E(r_e))}.$$

The value of $\widehat{w}_m$ is the same as (29) and it is not affected by changes in model uncertainty. Such changes, in contrast, determine the optimal $\widehat{w}_e$. In particular, an increase in $\theta \sigma_\mu^2 (E(r_e))$ determines a decrease in $\widehat{w}_e$ – i.e., a decrease in the DM's demand for the uncertain asset. On the other hand, an increase in λ decreases the values of both $\widehat{w}_m$ and $\widehat{w}_e$. Additionally, when asset returns are uncorrelated, the optimal ratio of risky to ambiguous assets is given by

$$\frac{\widehat{w}_m}{\widehat{w}_e} = \frac{BD}{CA},$$

so that

$$\frac{\partial}{\partial \theta} \left(\frac{\widehat{w}_m}{\widehat{w}_e} \right) = \sigma_\mu^2 \frac{B}{CA} > 0 \quad \text{and} \quad \frac{\partial}{\partial \lambda} \left(\frac{\widehat{w}_m}{\widehat{w}_e} \right) = -\theta \sigma_\mu^2 \frac{B}{AC^2} \sigma_P^2 (r_m) < 0.$$

Thus, an increase in ambiguity aversion rises the relative share of risky asset holdings, while an increase in risk aversion produces the opposite effect.

6.3.2 Correlated Asset Returns I: Changes in θ

Let us now turn to the case of non-zero correlation between returns, that is, $\sigma_P(r_m, r_e) \neq 0$ (which implies $H \neq 0$). Here we consider the effects of changes in θ, in the next section we study the effects of changes in λ.

Following common practice, we take the index r_m as a benchmark and we measure the performance of r_e through the *alpha-beta decomposition*, that is,

$$E_P(r_e) - r_f = \alpha_P(r_m, r_e) + \underbrace{\frac{\sigma_P(r_m, r_e)}{\sigma_P^2(r_m)}}_{=\beta_P(r_m, r_e)} (E_P(r_m) - r_f). \tag{32}$$

Here β represents the excess return's sensitivity of r_e to systematic risk as represented by r_m, and α captures the residual uncertainty that cannot be explained in terms of systematic risk only.

The next stark result shows that the sign of α alone determines the effects on the ratio of the optimal amounts $\widehat{w}_m$ and $\widehat{w}_e$ as ambiguity aversion varies.

Proposition 12 *Suppose the Omnibus Condition holds. Then,*

$$\operatorname{sgn} \frac{\partial}{\partial \theta} \left(\frac{\widehat{w}_m}{\widehat{w}_e} \right) = \operatorname{sgn} \alpha_P(r_m, r_e). \tag{33}$$

As anticipated, the interpretation is natural: an ambiguity averse agent interprets a positive α and the corresponding extra return as the premium for an unmeasurable increase in uncertainty. This drives him away from the ambiguous asset as ambiguity aversion increases. Analogously, a negative α is associated with an unmeasurable diminution of uncertainty that, in turn, makes the ambiguous asset more desirable.

As the above intuition suggests, the sign of α also governs the absolute (rather than relative) variations of the optimal $\widehat{w}_e$ when model uncertainty varies. On the other hand, also the sign of β becomes relevant to describe the absolute variation of $\widehat{w}_m$.

Proposition 13 *Suppose the Omnibus Condition holds. Then,*

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

and

$$\frac{\partial \widehat{w}_m}{\partial \theta} = -\beta_P(r_m, r_e) \frac{\partial \widehat{w}_e}{\partial \theta}. \tag{35}$$

Thanks to Propositions 12 and 13 we can show how variations in θ affect the optimal portfolio composition, both in relative and in absolute terms. Two possible cases arise.

Case 1 If α is *positive*, then

$$\triangle\theta > 0 \implies \triangle \left(\frac{\widehat{w}_m}{\widehat{w}_e} \right) > 0 \text{ and } \triangle \widehat{w}_e < 0, \tag{36}$$

where $\triangle$ denotes a small variation.[10] Here, an increase in ambiguity aversion determines a higher ratio $\widehat{w}_m / \widehat{w}_e$ and a lower optimal $\widehat{w}_e$. The sign of the variation in $\widehat{w}_m$ coincides with the sign of β, that is, of the covariance $\sigma_P(r_m, r_e)$,

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

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

Case 2 If α is *negative*, then

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

That is, an increase in θ determines a lower ratio $\widehat{w}_m/\widehat{w}_e$, due to a higher optimal $\widehat{w}_e$ and a lower optimal $\widehat{w}_m$ (it is easy to check that $\alpha < 0$ implies $\beta > 0$).

In sum, depending on the values of the technical risk measures α and β, we have altogether different effects of variations in θ on the composition of the optimal portfolio. For example, even though the valuation of the risky asset is not directly affected by changes in θ, only under positivity of both α and β higher ambiguity aversion results in a higher value of $\widehat{w}_m$ and a lower value of $\widehat{w}_e$, that is, in an increase in the optimal amount of the risky asset and a decrease in that of the uncertain asset. If β is negative, the ratio $\widehat{w}_m/\widehat{w}_e$ still increases, but only because the optimal amount $\widehat{w}_m$ decreases less than $\widehat{w}_e$. Under negativity of α, everything is reversed: now, higher ambiguity aversion results in a higher value of $\widehat{w}_e$ and a lower value of $\widehat{w}_m$.

When the ambiguous asset underperforms the risky one (Case 2), agents hold relatively more of it as uncertainty rises, formally violating the risk-return trade-off rule (and behaving much alike consumers of Giffen goods do when prices rise).[11] To the contrary, when the ambiguous asset outperforms the risky one (Case 1), higher uncertainty drives agents away from it.

The analysis of the above cases fully characterizes the effect of higher ambiguity aversion on optimal portfolios. Next we provide a numerical illustration of our comparative statics.

Example 1 Consider a DM allocating \$1 among a safe asset, an approximately unambiguous (risky) asset and an approximately ambiguous asset. Suppose that $\alpha > 0$ and $\beta > 0$. Then, as θ increases, holding risk attitude fixed, uncertainty in the payoff of the ambiguous asset drives the agent away from it. Hence, the ratio of risky to ambiguous assets holdings also increases. Figure I gives a graphical representation of this effect.

Insert Figure I Here

Now suppose that $\beta < 0$. Then, as θ increases, the agent will want to diversify away ambiguity in payoffs and hold a larger amount in the safe asset and a relatively larger amount in the risky asset. Notice that it is the sign of the correlation between the return of the risky asset and the return of the ambiguous asset that determines the variations of risky asset holdings in Case 1. That is, since the risky asset and the ambiguous asset are not perfectly correlated, the risky asset provides a hedge to fluctuations in model uncertainty. Figure II depicts these variations in portfolio shares.

Insert Figure II Here

Finally, if the risky asset has a higher expected return per unit of risk than the ambiguous asset – i.e., $\alpha < 0$ – then, the agent trades off return against risk. In particular, higher model uncertainty results in a smaller value of $\widehat{w}_m$ and a higher value of of $\widehat{w}_e$, which combined decrease the ratio $\widehat{w}_m/\widehat{w}_e$. These changes are exhibited in Figure III.

Insert Figure III Here

[11] The Giffen analogy for uncertain assets was inspired by Gollier (2009).

6.3.3 Correlated Asset Returns II: Changes in λ

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

Proposition 14 *Suppose the Omnibus Condition holds. Then,*

$$\operatorname{sgn} \frac{\partial \widehat{w}_e}{\partial \lambda} = \operatorname{sgn} \frac{\partial}{\partial \lambda} \left(\frac{\widehat{w}_m}{\widehat{w}_e} \right) = - \operatorname{sgn} \alpha_P \left(r_m, r_e \right). \tag{37}$$

Interestingly, variations in the ratio $\widehat{w}_m / \widehat{w}_e$ due to changes in λ have opposite sign relative to changes in θ. That is, changes in risk attitudes vary the portfolio composition among uncertain assets in the opposite direction to changes in ambiguity attitudes (and in ambiguity itself). 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 α. As to $\partial \widehat{w}_m / \partial \lambda$, we have:

$$\frac{\partial \widehat{w}_m}{\partial \lambda} = \widehat{w}_e \frac{\partial \left(\widehat{w}_m / \widehat{w}_e \right)}{\partial \lambda} + \frac{\widehat{w}_m}{\widehat{w}_e} \frac{\partial \widehat{w}_e}{\partial \lambda}. \tag{38}$$

If $\widehat{w}_m$ and $\widehat{w}_e$ are positive, then variations in $\widehat{w}_m$ have the same sign as those in $\widehat{w}_m / \widehat{w}_e$ and $\widehat{w}_e$. That is, increases in $\widehat{w}_m / \widehat{w}_e$ correspond to higher values of both $\widehat{w}_m$ 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 (38).

Example 2 Consider the DM of Example 1 and assume $\widehat{w}_m$ and $\widehat{w}_e$ are positive. If also α is positive, then, irrespective of the sign of the correlation between asset returns, an increase in λ results in systematically lower $\widehat{w}_m$, $\widehat{w}_e$ and $\widehat{w}_m / \widehat{w}_e$. For this reason, Figure IV and V present similar trends in comparative statics.

Insert Figure IV and Figure V Here

Assume now negativity of α. In this case, the risky asset has a higher expected return per unit of risk than the ambiguous asset. Then, as λ increases, the agent trades off return against risk but the ratio $\widehat{w}_m / \widehat{w}_e$ still increases. This is because the optimal amount $\widehat{w}_e$ increases less than $\widehat{w}_m$. Moreover, in all three cases, changes in risk attitudes vary the portfolio composition in the opposite direction to changes in attitudes toward model uncertainty.

Insert Figure VI Here

6.4 U.S. Business Cycles

In what follows we employ historical data from the past four decades to calculate the alpha gains of an ambiguous asset relative to a purely risky asset to determine an investor's optimal portfolio rebalancing in response to increasing ambiguity aversion. We choose the S&P500 index to represent the purely risky asset and the MSCI World ex-U.S. index to represent the ambiguous asset. The return on 3-month Treasury bills proxies the return of the risk-free asset.[12]

According to the National Bureau of Economic Research six recessions ensued between 1971 and today: 1973(IV) - 1975(I), 1980(I) - 1980(III), 1981(III) - 1982(IV), 1990(III) - 1991(I), 2001(I) - 2001(IV) and 2007(IV) - 2009(III).[13] During all six episodes the correlation between the two indexes spiked. However, in only two instances, 1981-1982 and 2007-2009, the ambiguous asset's alpha became negative, which suggested rebalancing toward the foreign index. Table I contains our results.

[12] All asset prices are in real U.S. dollars.

[13] Quarterly dates are in parentheses. The National Bureau of Economic Research does not define a recession in terms of two consecutive quarters of decline in real GDP. Rather, a recession is a significant decline in economic activity spread across the economy, lasting more than a few months, normally visible in real GDP, real income, employment, industrial production, and wholesale-retail sales.

Table I: U.S. Business Cycle Contractions

Reference Dates		Duration in Months	Sgn($\alpha_P(r_m, r_e)$)	
Peak	**Trough**		**Pre-crisis**	**Crisis**
November 1973 (IV)	March 1975 (I)	16	+	+
January 1980 (I)	July 1980 (III)	6	+	+
July 1981 (III)	November 1982 (IV)	16	+	−
July 1990 (III)	March 1991(I)	8	+	+
March 2001 (I)	November 2001 (IV)	8	+	+
December 2007 (IV)	July 2009 (III)	19	+	−

Despite the fact that Sharpe Ratio calculations may be blunted by noise over short periods of time, it is indicative that a "switch" occurred over recessions of comparable depth. Indeed, the Bureau of Economic Analysis reported a percentage change of -6.4 in real Gross Domestic Product (quarterly data seasonally adjusted at annual rates) in the first quarter of both 1982 and 2009 (U.S. National Income and Product Accounts). On the other hand, while considering the S&P500 as approximately unambiguous during expansions seems reasonable, the same intuition is less clear during recessions when information quality might well become poorer.

7 Conclusions

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

Then, we derive the solution to the static portfolio problem when the agent perceives ambiguity with respect to the true probability of an asset's expected return. The comparative statics of more ambiguity aversion engender two noteworthy results: (i) portfolio rebalancing in response to higher uncertainty depends *solely* on the return generated by the uncertain asset in excess of the risky asset return after correcting for risk (the *alpha*), and (ii) ambiguity aversion does not generally reinforce risk aversion and, indeed, an increase in ambiguity aversion may increase the optimal demand for 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 equity premium puzzle, the asset allocation puzzle, the home bias puzzle, and the employer-stock ownership puzzle.

A Proofs and Related Analysis

To prove the quadratic approximation (17) we need the following version of standard results on differentiation under the integral sign.

Lemma 15 *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

$$\partial_j G(\mathbf{x}) = \int \partial_j g(\mathbf{x},\omega)\, d\pi(\omega) \tag{39a}$$

$$\partial_{jk} G(\mathbf{x}) = \int \partial_{jk} g(\mathbf{x},\omega)\, d\pi(\omega) \tag{39b}$$

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

A.1 Quadratic Approximation

Here we prove Proposition 3 and Theorem 4. 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 16 *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 Theorem yields Proposition 3 and Theorem 4 as corollaries.

Theorem 17 *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 $w + \mathbf{x}\cdot\mathbf{h} \in L^\infty(I)$,*

$$C(w+\mathbf{x}\cdot\mathbf{h}) = w + E_{\bar{Q}}(\mathbf{x}\cdot\mathbf{h}) - \frac{1}{2}\lambda_u(w)\,\sigma^2_{\bar{Q}}(\mathbf{x}\cdot\mathbf{h}) - \frac{1}{2}(\lambda_v(w)-\lambda_u(w))\,\sigma^2_\mu(E(\mathbf{x}\cdot\mathbf{h})) + o\left(|\mathbf{x}|^2\right) \tag{40}$$

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 as in Lemma 16) for all $\mathbf{x} \in (-\delta,\delta)^N$.[14] In particular, for all $\mathbf{x} \in (-\delta,\delta)^N$ and all $\omega \in \Omega$,

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

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

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

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 $g:O\times\Omega\to\mathbb{R}$ as $(\mathbf{x},\omega)\mapsto u(w+\mathbf{x}\cdot\mathbf{h}(\omega))$. Next we show that g satisfies assumptions (a), (b), and (c) of Lemma 15.

(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.

(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 16)

$$|g(\mathbf{x},\omega)|=|u(w+\mathbf{x}\cdot\mathbf{h}(\omega))|\le M\quad;\quad|\partial_j g(\mathbf{x},\omega)|=|u'(w+\mathbf{x}\cdot\mathbf{h}(\omega))|\,|h_j(\omega)|\le M\|h_j\|_{\sup}$$
$$|\partial_{jk}g(\mathbf{x},\omega)|=|u''(w+\mathbf{x}\cdot\mathbf{h}(\omega))|\,|h_j(\omega)|\,|h_k(\omega)|\le M\|h_j\|_{\sup}\|h_k\|_{\sup}$$

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

By Lemma 15, for each $q\in\Delta$, the function defined on O by

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

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

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

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 $(\mathbf{x},q)\in O\times\Delta$,

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

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 15.

(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.[15]

(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 $j,k\in\{1,2,...,N\}$

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

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

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

(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 16 and K as above)

$$|f(\mathbf{x},q)| = |\phi(G(\mathbf{x},q))| \le M \quad ; \quad |\partial_j f(\mathbf{x},q)| = |\phi'(G(\mathbf{x},q))|\,|\partial_j G(\mathbf{x},q)| \le MK$$
$$|\partial_{jk} f(\mathbf{x},q)| \le |\phi''(G(\mathbf{x},q))|\,|\partial_k G(\mathbf{x},q)|\,|\partial_j G(\mathbf{x},q)| + |\phi'(G(\mathbf{x},q))|\,|\partial_{jk} G(\mathbf{x},q)| \le MK^2 + MK$$

and the latter majorization holds term by term.

By Lemma 15, the function defined on O by

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

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

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

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

$$c(\mathbf{x}) = c(\mathbf{0}) + \nabla c(\mathbf{0})\,\mathbf{x} + \frac{1}{2}\mathbf{x}^\intercal\nabla^2 c(\mathbf{0})\,\mathbf{x} + o\left(|\mathbf{x}|^2\right). \tag{43}$$

Next we explicitly compute it using repeatedly the relations obtained above as well as those provided by Lemma 16. 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 $j,k\in\{1,2,...,N\}$

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

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.[16] Now

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

[16] 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

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

Finally

$$c\left(\mathbf{0}\right)=w \tag{44}$$

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

$$\nabla c\left(\mathbf{0}\right)\mathbf{x}=E_{\bar{Q}}\left(\mathbf{x}\cdot\mathbf{h}\right)\qquad\forall\mathbf{x}\in O \tag{45}$$

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

$$\begin{aligned}\partial_{jk}c\left(\mathbf{0}\right)&=\psi''\left(F\left(\mathbf{0}\right)\right)\partial_{k}F\left(\mathbf{0}\right)\partial_{j}F\left(\mathbf{0}\right)+\psi'\left(F\left(\mathbf{0}\right)\right)\partial_{jk}F\left(\mathbf{0}\right)=\psi''\left(\phi\left(u\left(w\right)\right)\right)v'\left(w\right)^{2}E_{\bar{Q}}\left(h_{k}\right)E_{\bar{Q}}\left(h_{j}\right)\\&\quad+\psi'\left(\phi\left(u\left(w\right)\right)\right)\left(\left(v''\left(w\right)-v'\left(w\right)\frac{u''\left(w\right)}{u'\left(w\right)}\right)E_{\mu}\left(\left\langle h_{k},\cdot\right\rangle\left\langle h_{j},\cdot\right\rangle\right)+\frac{v'\left(w\right)}{u'\left(w\right)}u''\left(w\right)E_{\bar{Q}}\left(h_{j}h_{k}\right)\right)\\&=-\frac{v''\left(w\right)}{v'\left(w\right)}E_{\bar{Q}}\left(h_{k}\right)E_{\bar{Q}}\left(h_{j}\right)+\left(\frac{v''\left(w\right)}{v'\left(w\right)}-\frac{u''\left(w\right)}{u'\left(w\right)}\right)E_{\mu}\left(\left\langle h_{k},\cdot\right\rangle\left\langle h_{j},\cdot\right\rangle\right)+\frac{u''\left(w\right)}{u'\left(w\right)}E_{\bar{Q}}\left(h_{j}h_{k}\right)\\&=\lambda_{v}\left(w\right)E_{\bar{Q}}\left(h_{j}\right)E_{\bar{Q}}\left(h_{k}\right)+\left(\lambda_{u}\left(w\right)-\lambda_{v}\left(w\right)\right)E_{\mu}\left(\left\langle h_{j},\cdot\right\rangle\left\langle h_{k},\cdot\right\rangle\right)-\lambda_{u}\left(w\right)E_{\bar{Q}}\left(h_{j}h_{k}\right)\\&=-\left[\lambda_{u}\left(w\right)\sigma_{\bar{Q}}\left(h_{j},h_{k}\right)+\left(\lambda_{v}\left(w\right)-\lambda_{u}\left(w\right)\right)\sigma_{\mu}\left(\left\langle h_{j},\cdot\right\rangle,\left\langle h_{k},\cdot\right\rangle\right)\right].\end{aligned}$$

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 $\left[\sigma_{\mu}\left(\left\langle h_{j},\cdot\right\rangle,\left\langle h_{k},\cdot\right\rangle\right)\right]_{j,k=1}^{N}$

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

and

$$\frac{1}{2}\mathbf{x}^{\intercal}\nabla^{2}c\left(\mathbf{0}\right)\mathbf{x}=-\frac{1}{2}\lambda_{u}\left(w\right)\sigma_{\bar{Q}}^{2}\left(\mathbf{x}\cdot\mathbf{h}\right)-\frac{1}{2}\left(\lambda_{v}\left(w\right)-\lambda_{u}\left(w\right)\right)\sigma_{\mu}^{2}\left(E\left(\mathbf{x}\cdot\mathbf{h}\right)\right)\qquad\forall\mathbf{x}\in O. \tag{46}$$

This concludes the proof since replacement of (44), (45), and (46) into (43) delivers

$$c\left(\mathbf{x}\right)=w+E_{\bar{Q}}\left(\mathbf{x}\cdot\mathbf{h}\right)-\frac{1}{2}\lambda_{u}\left(w\right)\sigma_{\bar{Q}}^{2}\left(\mathbf{x}\cdot\mathbf{h}\right)-\frac{1}{2}\left(\lambda_{v}\left(w\right)-\lambda_{u}\left(w\right)\right)\sigma_{\mu}^{2}\left(E\left(\mathbf{x}\cdot\mathbf{h}\right)\right)+o\left(\left|\mathbf{x}\right|^{2}\right)$$

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

$$o\left(\left|\mathbf{x}\right|^{2}\right)=C\left(w+\mathbf{x}\cdot\mathbf{h}\right)-\left[w+E_{\bar{Q}}\left(\mathbf{x}\cdot\mathbf{h}\right)-\frac{1}{2}\lambda_{u}\left(w\right)\sigma_{\bar{Q}}^{2}\left(\mathbf{x}\cdot\mathbf{h}\right)-\frac{1}{2}\left(\lambda_{v}\left(w\right)-\lambda_{u}\left(w\right)\right)\sigma_{\mu}^{2}\left(E\left(\mathbf{x}\cdot\mathbf{h}\right)\right)\right]$$

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

Proof of Proposition 3. 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 $\gamma:\mathbf{x}\mapsto\sum_{i=1}^{N}x_{i}1_{A_{i}}=\mathbf{x}\cdot\mathbf{h}$ is a norm isomorphism between $\mathbb{R}^{N}$ and L^{2}.[17] In particular, choosing $\delta>0$ as in the proof of Theorem 17, for all $x\in\gamma\left((-\delta,\delta)^{N}\right)=\left\{\mathbf{x}\cdot\mathbf{h}:\mathbf{x}\in(-\delta,\delta)^{N}\right\}$

$$C\left(w+x\right)=w+E_{\bar{Q}}\left(x\right)-\frac{1}{2}\lambda_{u}\left(w\right)\sigma_{\bar{Q}}^{2}\left(x\right)-\frac{1}{2}\left(\lambda_{v}\left(w\right)-\lambda_{u}\left(w\right)\right)\sigma_{\mu}^{2}\left(E\left(x\right)\right)+o\left(\left|\mathbf{x}\right|^{2}\right) \tag{47}$$

[17] Finite dimensionality guarantees that Δ is bounded and *a fortiori* the support of μ is bounded too.

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 $x \in \gamma\left((-\delta,\delta)^N\right)$, $\underline{p} = \min_{i=1,\ldots,N} P(A_i)$, and $\overline{p} = \max_{i=1,\ldots,N} P(A_i)$. 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 (non-zero) 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 (47) 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. ■

Proof of Theorem 4. Let $h \in L^\infty(\Omega, \mathcal{F}, P)$. By Theorem 17, 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 $w + th \in L^\infty(I)$,

$$R_2(th) = C(w+th) - \left[w + E_{\bar{Q}}(th) - \frac{1}{2}\lambda_u(w)\sigma^2_{\bar{Q}}(th) - \frac{1}{2}(\lambda_v(w) - \lambda_u(w))\sigma^2_\mu(E(th))\right] \quad (48)$$

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

$$\sigma^2_\mu(E(th)) = \sigma^2_\mu(tE(h)) = t^2\sigma^2_\mu(E(h)). \quad (49)$$

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

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

It remains to show that $\sigma^2_\mu(E(h)) = 0$ implies that h is approximately unambiguous. If $h \in L^2$ and $\sigma^2_\mu(E(h)) = 0$, 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 $q^* \in \text{supp}\mu$ 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 h is approximately unambiguous. ■

Proof of Proposition 8. Clearly, (iii) $\Rightarrow$ (ii) $\Rightarrow$ (i) $\Rightarrow$ (iv). Next we show that (iv) $\Rightarrow$ (iii). Assume per contra suppμ is not a singleton. Then, there are $Q' \neq Q''$ in suppμ. In particular, $Q'(A) \neq Q''(A)$ for some $A \in \mathcal{F}$, and this contradicts $\sigma_\mu^2(E(1_A)) = 0$. In fact, $\sigma_\mu^2(E(1_A)) = 0$ and Proposition 7 imply that 1_A is approximately unambiguous, and hence $Q'(A) = E_{Q'}(1_A) = E_{Q''}(1_A) = Q''(A)$. ∎

Proof of Proposition 9 We prove the "only if," the converse being trivial. 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 Theorem 4, $\lim_{t\to 0} F(th)/t^2 = \lim_{t\to 0} G(th)/t^2 = 0$. Therefore, for all $h \in B$, setting $k = 2(\lambda_v(w) - \lambda_u(w))^{-1}$,

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

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 (iv) of Proposition 8 holds. In turn, this implies $\mu = \delta_{\bar{Q}}$. ∎

Proof of Proposition 10. Let $h \in L^2$. By definition of $M^\perp$ and by the Hilbert Decomposition Theorem, decomposition (20) 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 (21).

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. ∎

A.3 Portfolio

First order conditions for (26). Setting

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

(26) 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 11 Since $\partial D/\partial\sigma_\mu^2=\theta$ and $\partial D/\partial\theta=\sigma_\mu^2$, simple algebra shows that

$$\frac{\partial\left(\widehat{w}_m/\widehat{w}_e\right)}{\partial\sigma_\mu^2}=\theta\frac{B}{CA-HB}\quad\text{and}\quad\frac{\partial\left(\widehat{w}_m/\widehat{w}_e\right)}{\partial\theta}=\sigma_\mu^2\frac{B}{CA-HB}\tag{50}$$

and this implies the first equality in (31). Analogously,

$$\frac{\partial\widehat{w}_m}{\partial\sigma_\mu^2}=\theta\frac{(AC-BH)\,H}{\left(CD-H^2\right)^2}\quad\text{and}\quad\frac{\partial\widehat{w}_e}{\partial\sigma_\mu^2}=-\theta\frac{(CA-HB)\,C}{\left(CD-H^2\right)^2},\tag{51}$$

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

which imply the other equalities in (31). ■

Proof of Proposition 12 By definition

$$\beta_P\left(r_m,r_e\right)=\frac{H}{C}\quad\text{and}\quad\alpha_P\left(r_m,r_e\right)=E_P\left(r_e\right)-r_f-\beta_P\left(r_m,r_e\right)\left(E_P\left(r_m\right)-r_f\right)=A-\frac{H}{C}B\tag{53}$$

while, by (50),

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

thus

$$\operatorname{sgn}\frac{\partial}{\partial\theta}\left(\frac{\widehat{w}_m}{\widehat{w}_e}\right)=\operatorname{sgn}\left(CA-HB\right)=\operatorname{sgn}\left(\frac{CA-HB}{C^2}\right)=\operatorname{sgn}\alpha_P\left(r_m,r_e\right)$$

as wanted. ■

Proof of Proposition 13 By (52),

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

that is (35) holds. Moreover, (52) again implies

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

as wanted. ■

Proof of Proposition 14. Direct computation delivers

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

which together with Proposition 12 determines the sign of $\partial\left(\widehat{w}_m/\widehat{w}_e\right)/\partial\lambda$. Moreover,

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

hence

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

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

$$r_e-E_P\left(r_e\right)=k\left(r_m-E_P\left(r_m\right)\right)$$

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

A.4 Proofs of Section 2

Proof of Proposition 1. Consider the duality inclusion ι :supp$\mu \to (L^2)^*$ 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

$$\langle X, \bar{q} \rangle = \int_{\mathrm{supp}\mu} \langle X, \iota(q) \rangle \, d\mu(q), \qquad \forall X \in L^2. \tag{54}$$

By (54) 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 $x 1_\Omega \in L^\infty(I)$ and $V(L^\infty(I)) \ni V(x 1_\Omega) = \int_\Delta \phi \circ \langle u(x), \cdot \rangle \, d\mu = \phi(u(x)) = z$. ■

Proof of Eq. (14) Set $\varphi = u^{-1}$. It holds

$$\begin{aligned}
\lambda_\phi(u(w)) &= -\frac{\phi''(u(w))}{\phi'(u(w))} = -\frac{v''(\varphi(u(w)))(\varphi'(u(w)))^2 + v'(\varphi(u(w)))\varphi''(u(w))}{v'(\varphi(u(w)))\varphi'(u(w))} \\
&= -\varphi'(u(w)) \frac{v''(\varphi(u(w)))}{v'(\varphi(u(w)))} - \frac{\varphi''(u(w))}{\varphi'(u(w))} = -\frac{1}{u'(w)} \frac{v''(w)}{v'(w)} + \frac{u''(w)}{u'(w)^2} \\
&= \frac{1}{u'(w)} (\lambda_v(w) - \lambda_u(w)).
\end{aligned}$$

as desired. ■

References

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

[2] H. Bauer, *Measure and Integration Theory*, de Gruyter, Berlin, 2001.

[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, Unpublished manuscript, 2010.

[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, *The Review of Financial Studies*, 22, 3595-3627, 2009

[7] S. Cerreia-Vioglio, F. Maccheroni, M. Marinacci, and L. Montrucchio, Uncertainty Averse Preferences, Unpublished manuscript, 2009.

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

[9] S. Collard, S. Mukerji, K. Sheppard and J-M. Tallon, Ambiguity and the historical equity premium, Unpublished manuscript, 2009.

[10] D. Easley and M O'Hara, Liquidity and valuation in an uncertain World, *Journal of Financial Economics*, forthcoming.

[11] D. Easley and M O'Hara, Microstructure and ambiguity, *Journal of Finance*, forthcoming.

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

[13] G.B. Folland, *Real Analysis*, Wiley, 1999.

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

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

[16] C. Gollier, Portfolio choices and asset prices: The comparative statics of ambiguity aversion, Unpublished manuscript, 2009.

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

[18] P. Illeditsch, Amiguous information, risk aversion and asset pricing, Unpublished manuscript, 2009.

[19] Y. Izhakian and S. Benninga, The uncertainty premium in an ambiguous economy, Unpublished manuscript, 2008.

[20] N. Ju and J. Miao, Ambiguity, learning and asset returns, Unpublished manuscript, 2010.

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

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

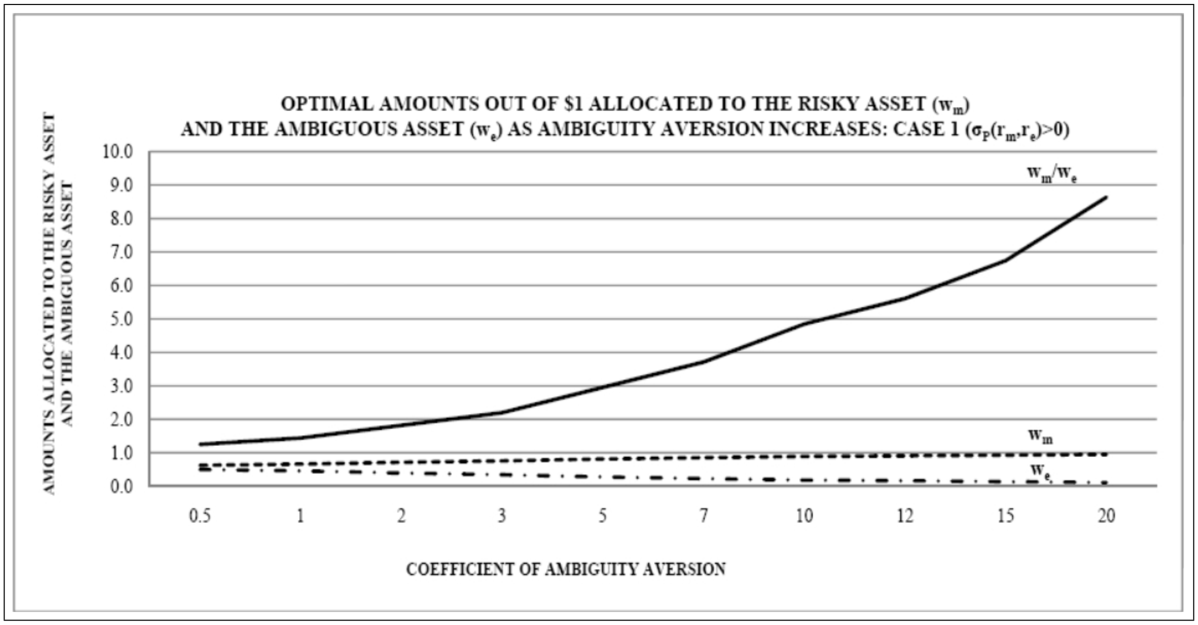
[23] H. M. Markowitz, Portfolio selection, *Journal of Finance*, 7, 77-91, 1952.

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

[25] C. Skiadas, Smooth ambiguity toward small risks and continuous-time recursive utility, Unpublished manuscript, 2009.

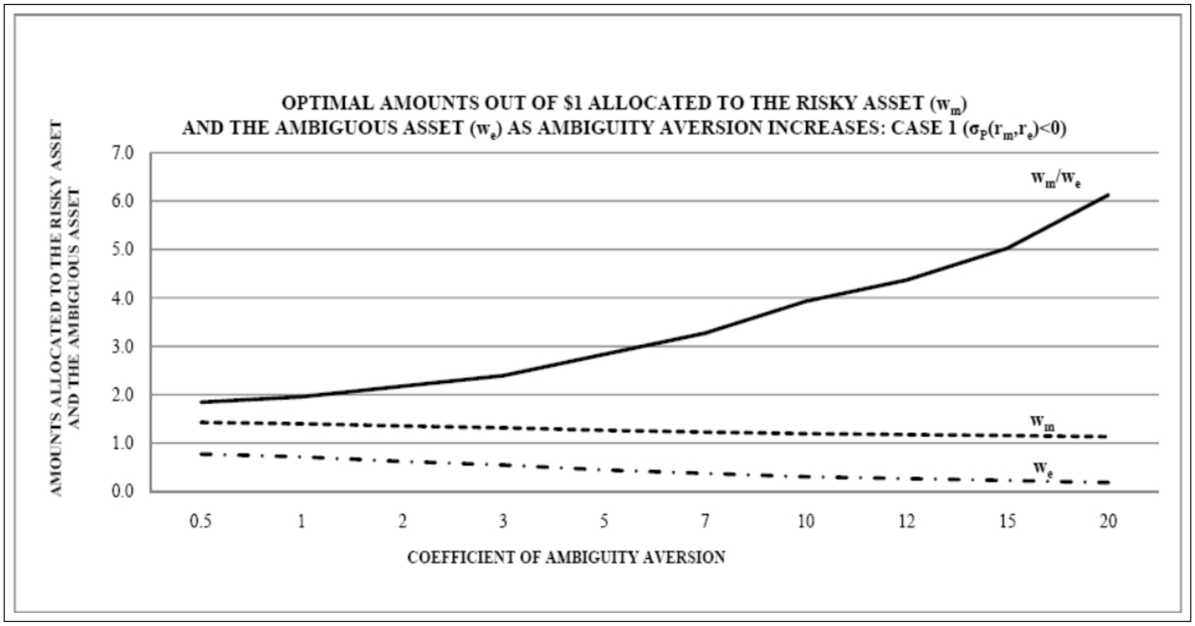
[26] M. Taboga, Portfolio selection with two-stage preferences, *Finance Research Letters*, 2, 152-164, 2005.

Figure I



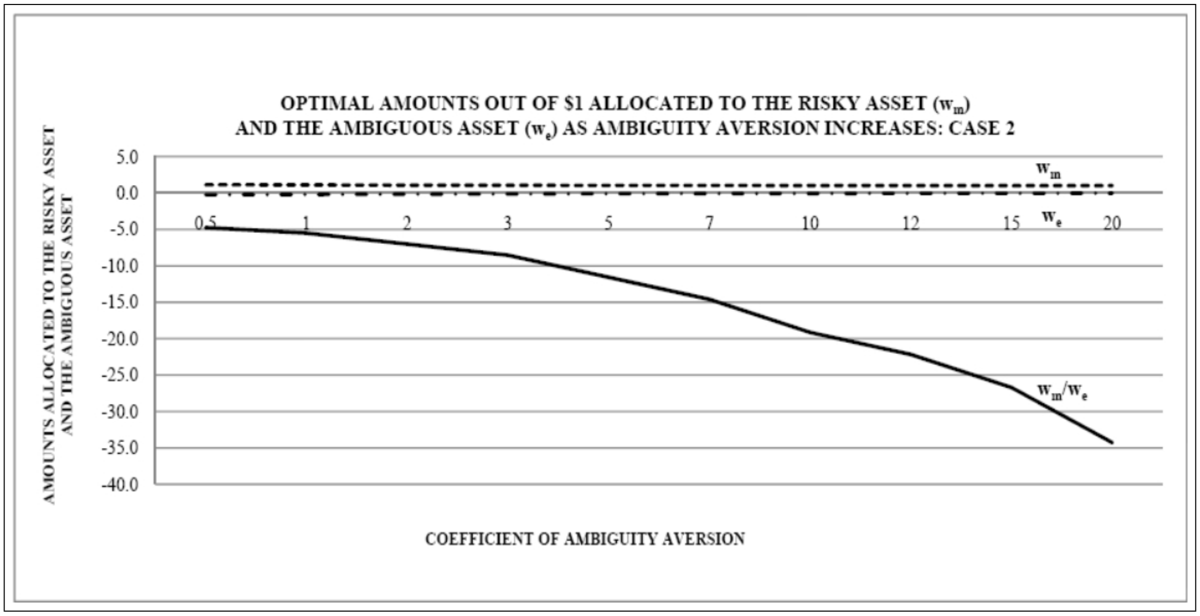
Notes: If the sum of the optimal amounts allocated to the risky and to the ambiguous assets exceeds \$1, then the safe asset is sold short.

Figure II



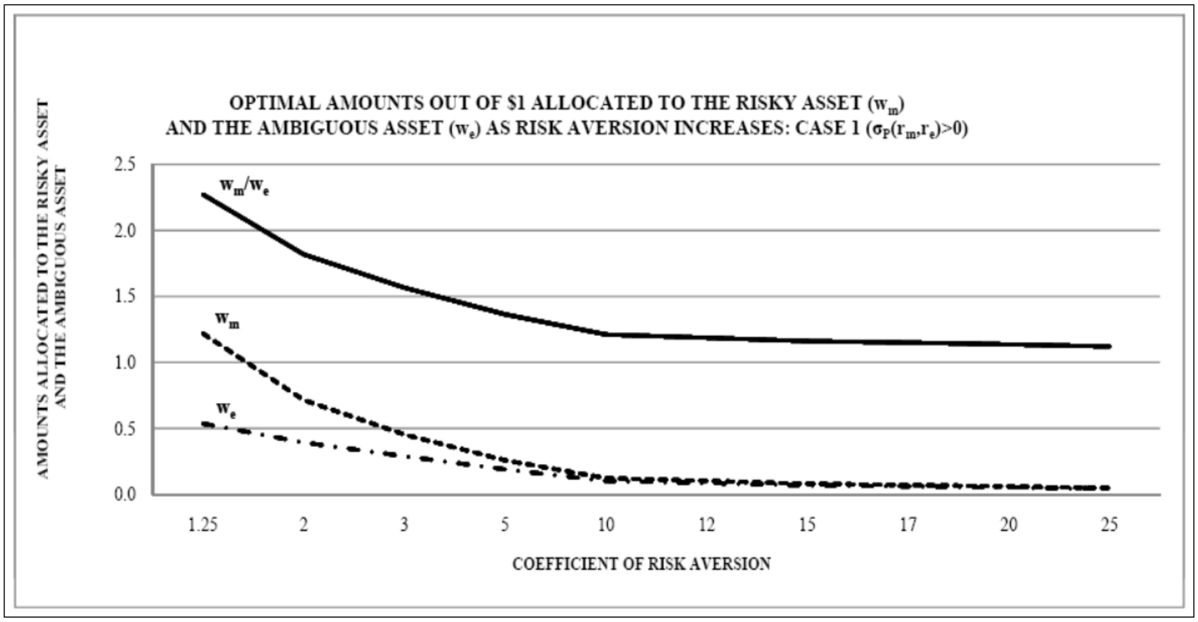
Notes: If the sum of the optimal amounts allocated to the risky and to the ambiguous assets exceeds $1, then the safe asset is sold short.

Figure III



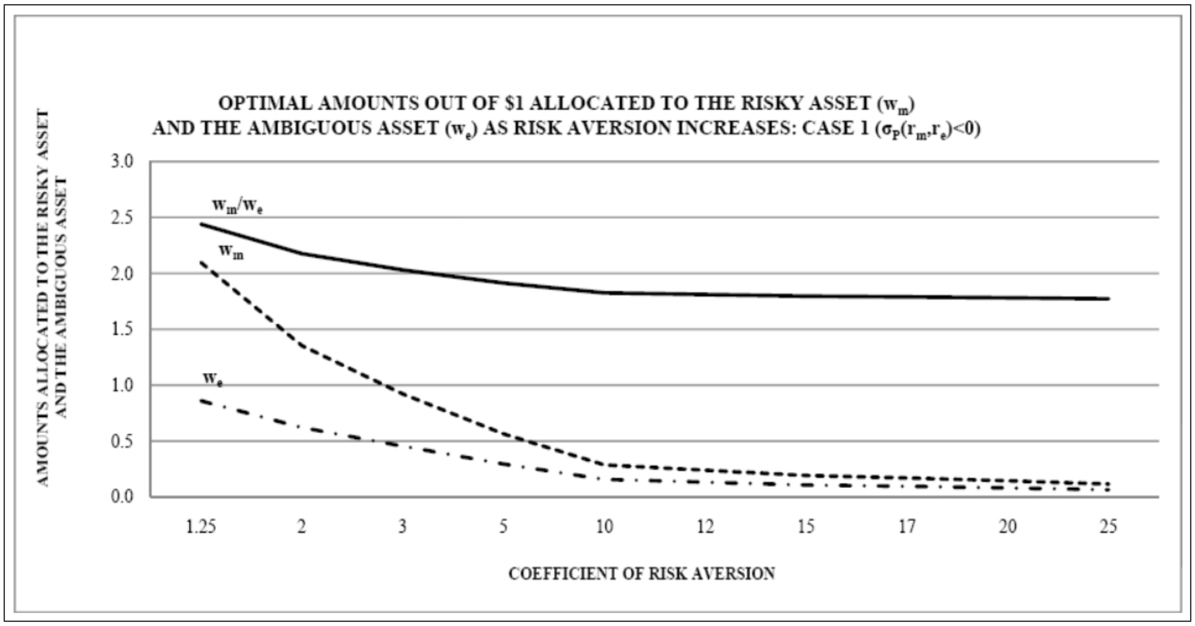
Notes: If the sum of the optimal amounts allocated to the risky and to the ambiguous assets exceeds $1, then the safe asset is sold short.

Figure IV



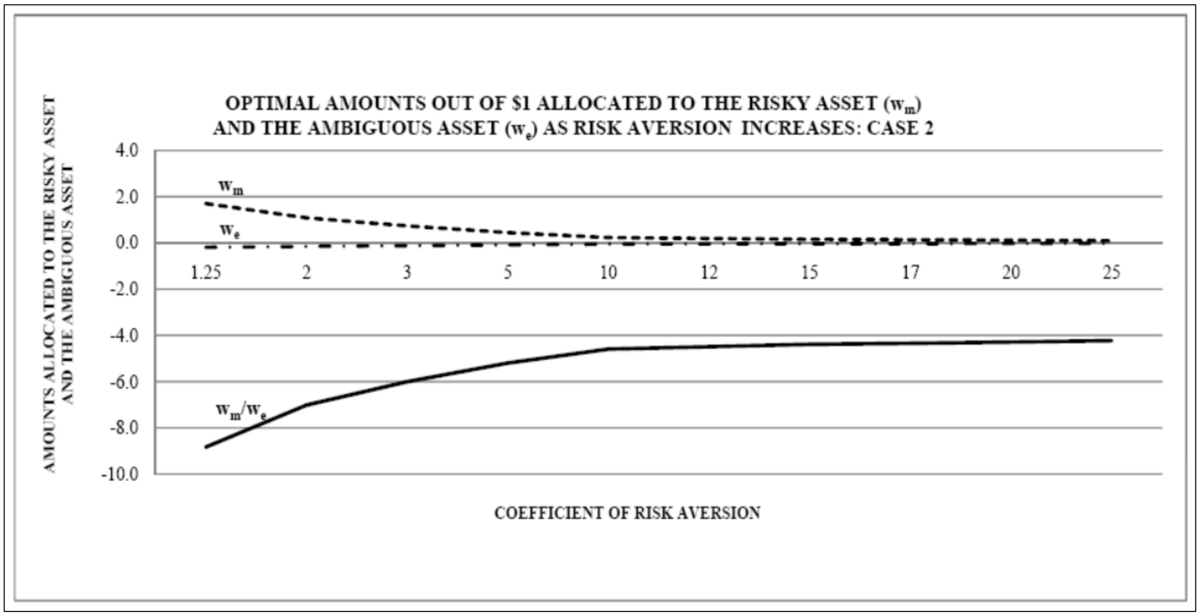
Notes: If the sum of the optimal amounts allocated to the risky and to the ambiguous assets exceeds \$1, then the safe asset is sold short.

Figure V



Notes: If the sum of the optimal amounts allocated to the risky and to the ambiguous assets exceeds $1, then the safe asset is sold short.

Figure VI



Notes: If the sum of the optimal amounts allocated to the risky and to the ambiguous assets exceeds $1, then the safe asset is sold short.