KIER DISCUSSION PAPER SERIES

KYOTO INSTITUTE
OF
ECONOMIC RESEARCH

Discussion Paper No.943

"Mutual Fund Theorem for Ambiguity-Averse Investors
and the Optimality of the Market Portfolio"

Chiaki Hara and Toshiki Honda

June 2016 (Revised: July 2016)

KYOTO UNIVERSITY

KYOTO, JAPAN

Mutual Fund Theorem for Ambiguity-Averse Investors and the Optimality of the Market Portfolio

Chiaki Hara[*] Toshiki Honda[†]

July 29, 2016

Abstract

We study the optimal portfolio choice problem for an ambiguity-averse investor having a utility function of the form of Klibanoff, Marinacci, and Mukerji (2005) and Maccheroni, Marinacci, and Rufino (2013) in an ambiguity-inclusive CARA-normal setup. We extend the mutual fund theorem to accommodate ambiguity, identify a necessary and sufficient condition for a given portfolio to be optimal for some ambiguity-averse investor, characterize all the ambiguity structure under which the given portfolio is optimal, and find the minimal ones in two senses to be made precise. We also calculate the minimal ambiguity structures based on the U.S. equity market data and find the smallest coefficient of ambiguity aversion with which the market portfolio is optimal is equal to 9.31.

JEL Classification Codes: C38, D81, G11.

Keywords: Ambiguity aversion, optimal portfolio, mutual fund theorem, FF6 portfolios, market portfolio.

[*]The paper has previously been entitled "Asset Demand and Ambiguity Aversion" and circulated since 2013. The paper is part of the academic projects on Economic Analysis of Intergenerational Issues: Searching for Further Development, funded by the Grant-in-Aid for Specially Promoted Research (grant number 22000001), and on Equilibrium Analysis on Financial Markets with Transaction costs, funded by the Grant-in-Aid (A) (grant number 25245046), both from Japan's Ministry of Education, Culture, Sports, Science and Technology. Hara's email address is `hara.chiaki.7x@kyoto-u.ac.jp`.

[†]The paper is part of the academic project funded by the Grant-in-Aid (grant number 25285097) from Japan's Ministry of Education, Culture, Sports, Science and Technology. Honda's email address is `thonda@ics.hit-u.ac.jp`. We are both grateful to Robert Anderson, Peter Bossaerts, Darrell Duffie, Simon Grant, Youichiro Higashi, Mayumi Horie, Peter Klibanoff, Christoph Kuzmics, Sujoy Mukerji, Frank Riedel, Jean-Marc Tallon, Takashi Ui, Katsuhito Wakai, William Zame, and participants of our presentations at Kyoto University, Hitotsubashi University, Meiji University, Yokohama National University, the 2013 Annual Meeting of the Nippon Finance Association, the Swiss-Kyoto Symposium, the 2013 Asian Meeting of the Econometric Society, the 2014 Annual Meeting of the Korean Association of Financial Engineering, the ZiF Research Group Workshop on Knightian Uncertainty in Strategic Interactions and Markets, the 2016 Winter Workshop on Operations Research, Finance and Economics, and the Fourth Asian Quantitative Finance Conference.

1 Introduction

1.1 Motivation

There are many empirical findings in financial economics that cannot be explained by standard theoretical models. A prominent one is that, contrary to the prediction by the Capital Asset Pricing Model, the market portfolio is not mean-variance-efficient. This anomaly is often true even in a stronger sense that a typical mean-variance-efficient portfolio, based on the empirical distribution of asset returns, involves very large long positions and very large (in absolute value) short positions.

Tables 1 and 2 present such an instance, based on the data of the so called FF6 portfolios in the U.S. equity markets, obtained from Ken French's website. The FF6 portfolios are formed by sorting out traded stocks in terms of the market equity (ME), which is either Small or Big, and ratio of the book equity to the market equity (B/M), which is either Low, Neutral, or High, and named the SL, SN, SH, BL, BN, and BH portfolios. They make up a partially aggregated description of the stock market performance that allows us to derive the the mean-variance-efficient portfolio and the market portfolio without suffering from the curse of dimensionality.

Table 1: Sample Means and the Covariance Matrix of the FF6 Portfolios: July 1926-December 2014.

	Mean (%)	SL	SN	SH	BL	BN	BH
R	0.28						
SL	0.98	57.36	50.77	55.76	34.61	35.62	43.74
SN	1.28	50.77	49.64	55.73	31.89	35.74	45.07
SH	1.48	55.76	55.73	67.64	34.64	41.20	53.89
BL	0.91	34.61	31.89	34.64	28.62	27.43	31.63
BN	0.97	35.62	35.74	41.20	27.43	32.89	38.32
BH	1.19	43.74	45.07	53.89	31.63	38.32	50.95

Table 1 reports the sample mean of the risk-free rate (R) and the sample means and covariances of the returns of the FF6 portfolios. The FF6 portfolios are formed in the following manner. First, sort out the stocks traded on NYSE, AMEX, and NASDAQ in terms of the market equity (also known as the market value and the market capitalization), abbreviated as ME, and the ratio of the book equity (also known as the book value) to the market equity, abbreviated as B/M. Second, partition the stocks, with positive book equity, into six groups, according to whether the ME belongs to the top 50% or the bottom 50% (referred to as being Big or Small), and whether the B/M belongs to the top 30%, the bottom 30%, or neither (referred to as being High, Low, or Neutral). Third, form the ME-weighted portfolio of the stocks in each of the six portfolios. The six portfolios thus formed are named SL, SN, SH, BL, BN, and BH in the obvious manner.

The sample means, variances, and covariances of the FF6 portfolios for the period of August 1926 to December 2014 are reported in Table 1. Within each group of a common B/M, the returns of the Small ME portfolios (SH, SN, and SL) have larger sample means and variances than the Big ME portfolios (BH, BN, and BL). Within each group of a

Table 2: The mean-variance efficient portfolio and the market portfolio

	Mean-variance-efficient portfolio	Market portfolio
SL	-3.36	0.02
SN	3.45	0.03
SH	1.12	0.02
BL	1.84	0.51
BN	-1.08	0.31
BH	-0.97	0.11

Table 2 reports the mean-variance efficient portfolio and the market portfolio. Both portfolio vectors are normalized so that the sum of the all elements is one. The market portfolio is the time-series average of market capitalization weight of the FF6 portfolios.

common ME, a higher B/M leads to larger sample means and variances, except that the SL portfolio has a larger variance than the SN portfolio. Table 2 reports the mean-variance-efficient portfolio, based on the data of Table 1, and the market portfolio, defined as the time-series average of ME weights. Both are normalized so that the coordinates add up to one. The mean-variance-efficient portfolio has extremely large long positions of the SN, SH, and BL portfolios and extremely large short positions of the SL, BN, and BH portfolios. In contrast, the market portfolio, by its definition, holds long all FF6 portfolios, and invests less than 3% in any Small ME portfolio.

To reconcile theory with the fact that the market portfolio is far from mean-variance-efficient, there are many possibilities, such as introducing more general utility functions, incomplete markets and market frictions, and heterogeneity in investors' utility functions. Among these possibilities, we opt for the first one by incorporating ambiguity aversion, while sticking to the traditional representative-agent paradigm, because there has been a significant development in the analysis of ambiguity-averse preferences, first in decision theory and subsequently in macroeconomics and asset pricing. Also, since our purpose is to reconcile theory with data, we should aim at the minimal deviation from expected utility functions. In some class of ambiguity-averse utility functions, we can give a precise meaning of minimality.

Among many classes of ambiguity-averse utility functions, we use the one initiated by Klibanoff, Marinacci, and Mukerji (2005) (abbreviated hereafter as KMM) and further explored by Maccheroni, Marinacci, and Ruffino (2013) (hereafter MMR). There is a probability distribution in investor's mind, which KMM called the second-order belief, on a set of distributions of asset returns or other payoff-relevant parameters. The coefficient of ambiguity aversion introduced by KMM is an indispensable building block in our subsequent analysis. Moreover, as explored by MMR, there is a natural ambiguity-inclusive extension of the CARA-normal setup. In the extension, the asset returns are normally distributed conditional on their mean returns, but the mean returns are ambiguous and normally distributed according to the second order belief. The return covariance matrix is, on the other hand, unambiguous. The assumption of ambiguous means and unambiguous variances is motivated by the fact that means are notoriously difficult to estimate but variances are less so.

To grasp the idea on how the extension is done, suppose that there are N risky asset, of which the returns are jointly normally distributed with mean vector μ_X and the covariance matrix Σ_X. Suppose also that the risk-free asset is traded, of which the return is denoted by R. Then a portfolio a of risky assets is optimal for an ambiguity-neutral (expected-utility-maximizing) investor with CARA coefficient θ if and only if

$$a = \frac{1}{\theta}\Sigma_X^{-1}\left(\mu_X - R\mathbf{1}\right), \tag{1}$$

where $\mathbf{1}$ is the vector of $\boldsymbol{R}^N$ of which every coordinate is equal to one. Then a is mean-variance-efficient, and, conversely, every mean-variance-efficient portfolio can be written in this form, for some θ. Denote the n-the coordinate of μ_X (the expected return of asset n) by μ_n and the n-coordinate of $\Sigma_X a$ (the covariance between the return of portfolio a and the return of asset n) by c_n. Then, by multiplying Σ_X to (1), we can rewrite the condition as

$$\frac{1}{\theta}\left(\mu_n - R\right) - c_n = 0 \tag{2}$$

for every n. In words, a portfolio a is mean-variance efficient if and only if its covariances with the risky assets are proportional to their expected excess returns.

When we observe that the investor chooses a portfolio that is not mean-variance-efficient, so that (2) fails to hold for the chosen portfolio a, we attribute it to some kind of ambiguity in his mind regarding the expected returns of the risky assets. We assume that there is a univariate normally distributed random variable $\hat{M}$ and that, for each asset n, the conditional expected return is equal to $s_n\hat{M} + k_n$, where s_n and k_n are constants and the s_n are proportional to the $\theta^{-1}(\mu_n - R) - c_n$. This assumption implies that the conditional expected returns of all assets are positively or negatively perfectly correlated, the absolute value $|s_n|$ of the coefficient determines the sizes of ambiguity that lies in the returns of asset n, and the sign of s_n determines the positivity or negativity of (perfect) correlation among all assets. Once the values of the s_n and the k_n are properly specificed, the portfolio a turns out to be optimal for an ambiguity-averse investor.

1.2 Overview of our results

The argument towards the end of the previous subsection on how to construct ambiguity to make a given portfolio a is optimal is somewhat informal. To give a formal proof, we need to establish three results.

First, we generalize the mutual fund theorem (Theorem 1 in Section 3). In the CARA-normal setup with ambiguity-neutral investors, it is well known that the mutual fund theorem holds, that is, there is a single portfolio of risky assets such that every CARA investor's optimal portfolio of risky assets is a positive multiple of this portfolio. We show that the mutual fund theorem no longer holds for ambiguity-averse investors, but provide a generalized mutual fund theorem that identifies a collection of mutual funds of which each ambiguity-averse investor's optimal portfolio is a linear combination. We also show how the demands for these mutual fund vary as the investor becomes more ambiguity-averse. This theorem is different from our argument in the previous subsection on how to construct ambiguity to make a given portfolio optimal, in that the ambiguity in investor's mind is given rather than constructed. Yet, the theorem is full of other implications, because it clarifies the way in which the ambiguity affects optimal portfolios and is general enough to accommodate multivariate representation of ambiguity.

Second, we provide a necessary and sufficient condition for a given portfolio to be optimal for some ambiguity-averse investors (Theorem 2 in Section 4). The existence of such investors should, of course, not be taken for granted. Indeed, the necessary and sufficient condition shows that there is an ambiguity-averse investor for whom a given portfolio is optimal if and only if the expected excess return of the portfolio is positive. It also identifies the range of the risk aversion coefficients such an ambiguity-averse investor may have, and the class of all the pairs of second-order beliefs (the distribution of the expected returns of the risky assets) and ambiguity aversion coefficients with which a optimal.

The third result, then, identifies the pair that minimizes the ambiguity and the pair that minimizes the ambiguity aversion coefficient (Theorem 3 in Section 4). It, thus, shows the minimal deviation from expected utility functions that is necessary to rationalize the choice of an mean-variance inefficient portfolio by ambiguity aversion.

Once these and other theoretical results have all been established, we apply them to the U.S. equity market data to infer the ambiguity and ambiguity aversion that are embedded in the representative investor, who chooses, by definition, the market portfolio (Section 5). We show that the SH portfolio, the portfolio having a small market equity and a high ratio of the book equity to the market equity, has the most ambiguous return among the FF6 portfolios. The ambiguity aversion coefficient can be defined analogously to the (constant) coefficient of relative risk aversion, in that both measure the elasticity of marginal utility. We find that the smallest ambiguity aversion coefficient the representative investor may have is equal to 9.31.

1.3 Related literature

As we have mentioned, we use the ambiguity-averse utility functions of the form axiomatized by KMM. They also introduced the notion of coefficients of ambiguity aversion, which is similar to coefficients of absolute risk aversion of Arrow and Pratt. The more famous (and earlier) form for ambiguity-averse utility functions was axiomatized by Gilboa and Schmeidler (1989). KMM claimed that their functional form is more tractable and allows them to disentangle ambiguity from ambiguity aversion. Epstein (2010) presented Ellsberg-like thought experiments to argue, among other things, that it is in fact impossible to disentangle ambiguity from ambiguity aversion. Subsequently, Klibanoff, Marinacci, and Mukerji (2012) gave a counterargument. In this paper, we define the minimality in terms of both ambiguity and ambiguity aversion and, in our numerical analysis, we find the ambiguity and the ambiguity aversion coefficient that are minimal according to each of the two definitions.

MMR extended the notion of certainty equivalents from the expected utility functions to the ambiguity-averse utility functions of the KMM type. They also introduced an extended notion of mean-variance utility functions exhibiting ambiguity aversion of the KMM type. In our CARA-normal framework with ambiguity aversion, the utility functions are defined by two negative exponential functions with differing coefficients. The distribution of the expected returns (the second-order belief) and the conditional distribution of returns given the expected returns are both normal. Then, the KMM utility functions coincide with the mean-variance utility functions of MMR. But none of our three main results was obtained by MMR.

Epstein and Miao (2003) considered a continuous-time general equilibrium model of two investors (countries) and two assets with ambiguous returns to address the home bias puzzle. Based on the axiomatization by Hayashi and Miao (2011), Ju and Miao (2012) introduced a generalized recursive smooth ambiguity model, by which they extended the KMM utility functions to the discrete-time setup while allowing for separation between relative risk aversion and elasticity of intertemporal substitution in the manner of Epstein and Zin (1989). While they assumed that there is only one asset of which returns are ambiguous, we focus on the optimal portfolio choice problem by allowing for an arbitrary number of such assets. As such, the setups are quite different, but we share the view that ambiguity aversion is the key to solve asset pricing puzzles and the KMM utility functions are tractable. In addition, we both believe that the traditional Bayesian portfolio choice, in which the investor is assumed to be equally averse to risks in asset returns and uncertainty (ambiguity) in parameters, is insufficient to solve those puzzles.

Collard, Mukerji, Sheppard, and Tallon (2015) found the KMM coefficients of ambiguity aversion, and the countercyclical nature thereof, using the U.S. equity market data. Our framework is static while theirs is dynamic, but we deal with multiple risky assets while they deal only with one risky asset. In particular, we introduce ambiguity aversion in order to rationalize the representative investor's choice of the market portfolio, which, given the sample means, variances, and covariances of asset returns, cannot be optimal for an ambiguity-neutral investor whatever large or small his risk aversion coefficient is. On the other hand, they are more interested in to what extent the equity premium can be attributed to ambiguity aversion.

Ahn, Choi, Gale, and Kariv (2014) and Attanasi, Gollier, Montesano, and Pace (2014) inferred ambiguity aversion from laboratory experiments on portfolio selection. Bossaerts, Ghirardato, Guarnaschelli, and Zame (2010) did the same by letting nearly thirty subjects to trade state-contingent consumptions (or Arrow assets) without having told them the probabilities for some (but not all) states to occur. Ahn, Choi, Gale, and Kariv (2014) and Bossaerts, Ghirardato, Guarnaschelli, and Zame (2010) concluded that many subjects are ambiguity-averse, and that they appear to have utility functions of Gilboa and Schmeidler (1989) but not of KMM, because they tend to choose unambiguous (purely risky) consumption plans, a tendency that is consistent with utility functions of Gilboa and Schmeidler (1989) but not with those of KMM. Note, however, that the KMM utility functions represent approximately the same preference relations as the utility functions of Gilboa and Schmeidler (1989) as the ambiguity aversion coefficient becomes large without bounds, according to Proposition 3 of KMM. Hence, these experimental results could be interpreted as saying that the subjects have KMM utility functions with extremely high ambiguity aversion coefficients. Indeed, in this paper, we deduce from the generalized mutual fund theorem (Theorem 1) that the demand for portfolios with ambiguous returns diminishes to zero, while the demand for a portfolio with unambiguous returns remains constant, as the ambiguity aversion coefficient diverges to infinity. Moreover, using the U.S. equity market data we identify a portfolio with unambiguous returns, as well as the portfolio with the most ambiguous returns.

1.4 Organization of the paper

The rest of this paper is organized as follows. Section 2 sets up the model. Section 3 presents the generalized mutual fund theorem and some of its implications. Section 4 shows how to find an ambiguity-averse investor for whom a given portfolio is optimal. Section 5 calculates the ambiguity covariance matrices and ambiguity aversion coefficients that are implied by the U.S. equity market data. Section 6 concludes and suggests directions of future research. All proofs and most lemmas are given in the appendix.

2 Setup

2.1 Formulation

The setup of this paper is essentially a special case of those of KMM and MMR and especially close to that of Section 6 of MMR, but we lay it out in a manner that is more suitable to accommodate the additional parametric assumptions we will impose on ambiguity in the expected asset returns. Let $(\Omega, \mathscr{F}, P)$ be a probability space. Let M be a random vector defined on Ω. We will later see that M describes the ambiguity in the mind of investors.

For each $\theta > 0$, define $u_\theta : \boldsymbol{R} \to \boldsymbol{R}$ by letting $u_\theta(x) = -\exp(-\theta x)$ for every $x \in \boldsymbol{R}$. This felicity function exhibits constant absolute risk aversion (CARA) and its coefficient is equal to θ. For each $\gamma > 0$ and each $\theta > 0$, define a utility function $U_{\gamma,\theta}$ over a set of random variables Z on Ω by

$$U_{\gamma,\theta}(Z) = E\left[u_\gamma\left(u_\theta^{-1}\left(E\left[u_\theta(Z) \mid M\right]\right)\right)\right]. \tag{3}$$

If we write $\varphi_{\gamma,\theta} = u_\gamma \circ u_\theta^{-1}$ and $\eta = \gamma/\theta - 1$, then

$$\varphi_{\gamma,\theta}(z) = -(-z)^{\gamma/\theta - 1} = -(-z)^{1+\eta} \tag{4}$$

for every $z < 0$, and

$$U_{\gamma,\theta}(Z) = E\left[\varphi_{\gamma,\theta}\left(E\left[u_\theta(Z) \mid M\right]\right)\right].$$

Since

$$-\frac{\varphi''_{\gamma,\theta}(z)(-z)}{\varphi'_{\gamma,\theta}(z)} = \eta \tag{5}$$

for every $z < 0$, η is equal to the elasticity of the marginal utility from conditional expectations given M. We call η the coefficient of ambiguity aversion, because, according to Theorem 2 of KMM, the more concave the function $\varphi_{\gamma,\theta}$ is, the more ambiguity-averse the investor is, and the larger the value of η, the function $\varphi_{\gamma,\theta}$ is more concave. Our usage of the ambiguity aversion coefficient differs from that of KMM in that they defined the ambiguity aversion coefficient as $-\varphi''_{\gamma,\theta}(z)/\varphi'_{\gamma,\theta}(z)$. We opt for the left-hand side of

(5) because it is constant (independent of z) and still represents the same ranking of concavity as KMM's definition.

If $\gamma = \theta$, then $\eta = 0$ and $\varphi_{\gamma,\theta}$ is the identity map. Thus, by the law of iterated expectation, $U_{\gamma,\theta}(Z) = E\left[u_\theta(Z)\right]$. In this case, therefore, $U_{\gamma,\theta}$ is an expected utility function with CARA coefficient θ. We then say that the investor is ambiguity-neutral. We say that an investor who has the utility function $U_{\gamma,\theta}$ with $\gamma > \theta$, or $\eta > 0$, is ambiguity-averse. If $\gamma < \theta$, or $\eta < 0$, then the investor is ambiguity-loving, though we will not pay much attention to this case.

Assume that two types of assets are traded. The first one is composed of N assets whose gross returns are represented by an N-variate random vector X defined on Ω. The second one is the bond whose gross return is deterministic and equal to $R \in \boldsymbol{R}$. We assume also that M (as well as X) is an N-variate random vector, and M and X are jointly normally distributed. We further assume that $E[M] = E[X]$ and $\mathrm{Cov}[M, X] = \mathrm{Cov}[M, M]$. We can thus write

$$\begin{pmatrix} M \\ X \end{pmatrix} \sim \mathscr{N}\left(\begin{pmatrix} \mu_M \\ \mu_M \end{pmatrix}, \begin{pmatrix} \Sigma_M & \Sigma_M \\ \Sigma_M & \Sigma_X \end{pmatrix}\right). \tag{6}$$

This assumption involves no loss of generality. It can indeed be shown that even if M did not satisfy this assumption (possibly with a dimension greater or smaller than N), some linear transformation of M added by some vector of $\boldsymbol{R}^N$ would satisfy this assumption.[1]

Then the conditional return of X given M is normally distributed:

$$X|M \sim \mathscr{N}\left(M, \Sigma_{X|M}\right),$$

where $\Sigma_{X|M} = \Sigma_X - \Sigma_M$. Our utility function, thus, embodies the idea that the investor believes that the expected returns of the risky assets are ambiguous; when the expected excess return vector is equal to $m \in \boldsymbol{R}^N$, the asset returns are distributed according to $\mathscr{N}(m, \Sigma_{X|M})$; and these models of return distributions are distributed according to $\mathscr{N}(\mu_M, \Sigma_M)$.

[1]This point can be articulated as follows. Suppose that F is a K-dimensional random vector and

$$\begin{pmatrix} F \\ X \end{pmatrix} \sim \mathscr{N}\left(\begin{pmatrix} \mu_F \\ \mu_X \end{pmatrix}, \begin{pmatrix} \Sigma_F & \Sigma_{FX} \\ \Sigma_{XF} & \Sigma_X \end{pmatrix}\right)$$

Then there is a $D \in \boldsymbol{R}^{K \times N}$ such that $\Sigma_{XF} = D^\top \Sigma_F$. Define $M = D^\top F + (\mu_X - D^\top \mu_F)$. Then (M, X) has the same distribution as assumed in (6).

2.2 Some decision-theoretic facts on our utility functions

Though not explicitly used in the subsequent analysis, it will be hepful to give some facts on our specification (3) of utility functions from a decision-theoretic perspective. First, while the KMM utility functions, in general, may not be concave, $U_{\gamma,\theta}$ is concave for every $\gamma > 0$ and every $\theta > 0$. The reason is that since the CARA function u_θ satisfies the condition in Section 3.16 of Hardy, Littlewood, and Polya (1952),[2] the mapping $Z \mapsto u_\theta^{-1}\left(E\left[u_\theta(Z)|M\right]\right)$ is a concave function at every realization of M. In particular, $U_{\gamma,\theta}$ is concave even when the investor is ambiguity-loving. This should be contrasted with a fact on the α-MEU utility functions of Marinacci (2002), that if the decision maker is ambiguity-loving, then his α-MEU utility function is never concave. Second, we could have defined $U_{\gamma,\theta}$ by $U_{\gamma,\theta}(Z) = E\left[u_\gamma\left(E\left[u_\theta(Z) \mid M\right]\right)\right]$, that is, we could have used the negative exponential function $-\exp(-\gamma z)$ in place of the negative power function $-(-z)^{\gamma/\theta-1}$. We do not take the negative exponential function for two reasons. The first reason is that, while it represents constant ambiguity aversion and is tractable in general,[3] it does not do so much as the negative power function when combined with the CARA utility function u_θ. The second reason is that, since we aim to extend the CARA-normal model to a model of ambiguity-averse utility functions and since CARA is defined in terms of consumption levels, we believe that if constant ambiguity aversion is to be imposed, then it should be done so on conditional certainty equivalents $u_\theta^{-1}\left(E\left[u_\theta(Z) \mid M\right]\right)$ rather than conditional expected utility levels $E\left[u_\theta(Z) \mid M\right]$. Then the resultant function $u_\gamma \circ u_\theta^{-1}$ must necessarily be a negative power function.

While the setup of this paper may appear to be quite different from those of KMM and MMR, our utility function $U_{\gamma,\theta}$ in (3) can indeed be obtained in the setup of MMR and, to a lesser extent, that of KMM as well. We now show how this can be done, taking μ_M, Σ_X, and Σ_M with the properties assumed in our setup as given. The difference between KMM and MMR setups lies in whether we take the space of objectively defined probability distributions inside or outside the state space. A consequence of this difference is that in the KMM setup, there are so-called roulette lotteries, that is, acts that give rise to single probability distributions of consequences regardless of the choice of second-order beliefs, but there is no such acts in the MMR setup.

First, we show that our utility function can be obtained in the setup of MMR. As in our model, let Ω be a measurable space, representing physical uncertainty. Denote by Δ the

[2]Strictly speaking, we need to reverse the sign condition on the second derivative in paragraph 106 in that section and check that this leads us to the (reversed) necessary and sufficient condition in our casecase, because they dealt with convex, not concave, functions.

[3]For example, Ahn, Choi, Gale, and Kariv (2015) and Collard, Mukerji, Sheppard, and Tallon (2015) assumed constant ambiguity aversion.

set of all probability distributions on Ω, representing the set of probability distributions that the investor perceive as possible models of the physical uncertainty. Let μ be a probability distribution on Δ, which KMM call the second order belief, representing the subjective likelihood of these models. Let $X : \Omega \to \boldsymbol{R}^N$ be an N-dimensional random vector, representing the returns of N assets. Now, suppose that $\Omega = \boldsymbol{R}^N$. For each measurable subset Ψ of Ω, write $\Delta_\Psi = \{\mathscr{N}(\omega, \Sigma_X - \Sigma_M) \in \Delta \mid \omega \in \Psi\}$. Suppose that for every measurable subset Ψ of Ω, $\mu(\Delta_\Psi)$ is equal to the measure that $\mathscr{N}(\mu_M, \Sigma_M)$ gives to Ψ. Suppose that X is the identity map on Ω. Write, for each $\omega \in \Omega$, an integral of a random variable with respect to $\mathscr{N}(\omega, \Sigma_X - \Sigma_M)$ using its Radon-Nikodym derivative with respect to, say, $\mathscr{N}(0, \Sigma_X)$ (of which the support coincides with $\boldsymbol{R}^N$). Then the utility function derived from μ, u_θ, and $\varphi_{\gamma,\theta}$ in the manner of equality (9) of MMR coincides with our utility function $U_{\gamma,\theta}$ in (3). The second order belief in this setup coincides with the (marginal) distribution of M in our setup, and, for each $\omega \in \boldsymbol{R}^N$, the distribution of X with respect to the probability distribution $\mathscr{N}(\omega, \Sigma_X - \Sigma_M)$ in this setup coincides with the conditional distribution of X given $M = \omega$ in our setup. Hence, the ambiguity-averse mean-variance utility function derived from this specification coincides with the ambiguity-averse mean-variance utility function $V_{\gamma,\theta}$, which will be defined in (8).

Next, we show that our utility function can be obtained in the setup that is identical, in spirit, to the setup of KMM. Let Ω be a measurable space, representing physical uncertainty to which the investor is unsure of the probability distribution to attach. Let Υ be another measurable space, representing physical uncertainty to which the investor is sure of the probability distribution to attach. Denote the probability distribution on Υ by λ. We deviate from the setup of KMM in that Υ and λ are different from the unit interval $[0, 1]$ and the Lebesgue measure on it, though λ is still interpreted as an objective probability distribution. Write $S = \Omega \times \Upsilon$. Denote by Δ the set of all product (joint and yet independent) distributions of a probability distribution on Ω and λ. Then Δ is a subset of the set of all probability distributions on S, representing the set of probability distributions that the investor perceive as possible models of the physical uncertainty. Let μ be a probability distribution on Δ, which KMM call the second order belief, representing the subjective likelihood of these models. Let $X : S \to \boldsymbol{R}^N$ be an N-dimensional random vector, representing the returns of N assets. Now, suppose that $\Omega = \boldsymbol{R}^N$, $\Upsilon = \boldsymbol{R}^N$, and $\lambda = \mathscr{N}(0, \Sigma_X - \Sigma_M)$. For each $\omega \in \Omega$, denote by δ_ω the (degenerate) probability distribution on Ω such that $\delta_\omega(\{\omega\}) = 1$. Then $\delta_\omega \otimes \lambda \in \Delta$. For each measurable subset Ψ of Ω, write $\Delta_\Psi = \{\delta_\omega \otimes \lambda \in \Delta \mid \omega \in \Psi\}$. Recalling that $\Omega = \boldsymbol{R}^N$, suppose that for every measurable subset Ψ of Ω, $\mu(\Delta_\Psi)$ is equal to the measure that $\mathscr{N}(\mu_M, \Sigma_M)$ gives to Ψ. Denote by M the projection of S onto Ω and by L the projection of S onto Υ. Recalling that $\Omega = \Upsilon = \boldsymbol{R}^N$, suppose that $X = M + L$. Then the utility function derived from μ, u_θ, and $\varphi_{\gamma,\theta}$ in the manner of equality (2) of KMM coincides with our utility function $U_{\gamma,\theta}$ in (3). The second order belief in this setup coincides, via the isomorphism $\omega \mapsto \delta_\omega \otimes \lambda$, with the (marginal) distribution of M in our setup, and, for each $\omega \in \boldsymbol{R}$, the distribution of X with respect to the probability distribution $\delta_\omega \otimes \lambda$ in this setup coincides with the conditional distribution of X given $M = \omega$ in our setup. Hence, the ambiguity-averse mean-variance utility function derived from this specification coincides with the ambiguity-averse mean-variance utility function $V_{\gamma,\theta}$, which will be defined in (8).

Even on the interpretative side, our setup is not different from that of KMM. Indeed, when justifying the use of second order acts (acts that are defined on Δ) in their axioms, they claimed that in a portfolio choice problem, second order acts could be bets about parameter values that characterizes the asset returns and be determined, among others, by events that take place inside the firm and in the wider market. In our notation, the randomness of such events is captured by the distribution $\mathscr{N}(\mu_M, \Sigma_M)$, while the remaining randomness is captured by the conditional distribution $\mathscr{N}(m, \Sigma_X - \Sigma_M)$ given $M = m$, which is understood to be unambiguous, because it is presumably quantifiable by a sufficiently large set of historical return data with stationarity.

2.3 Preliminary analysis

Denoted by $\mathscr{S}^N$ the set of all $N \times N$ symmetric matrices. Denote by $\mathscr{S}^N_{++}$ the set of all symmetric and positive definite $N \times N$ matrix, and by $\mathscr{S}^N_{+}$ the set of all symmetric positive semidefinite $N \times N$ matrix. Then $\mathscr{S}^N_{++} \subset \mathscr{S}^N_{+} \subset \mathscr{S}^N$. We assume that the total covariance matrix $\Sigma_X \in \mathscr{S}^N_{++}$ but allow for the ambiguity covariance matrix $\Sigma_M \in \mathscr{S}^N_{+} \backslash \mathscr{S}^N_{++}$ and $\Sigma_{X|M} \in \mathscr{S}^N_{+} \backslash \mathscr{S}^N_{++}$. Note that for every $\Sigma \in \mathscr{S}^N$, $\operatorname{Row}\Sigma = \operatorname{Col}\Sigma$ and $\operatorname{Ker}\Sigma = (\operatorname{Row}\Sigma)^\perp = (\operatorname{Col}\Sigma)^\perp$ and that, for every $\Sigma \in \mathscr{S}^N_{+}$ and every $v \in \boldsymbol{R}^N$, $v \in \operatorname{Ker}\Sigma$ if and only if $v^\top \Sigma v = 0$.

While our assumption excludes perfect correlation between any pair of linear combinations of the returns of the N assets with respect to the total covariance matrix Σ_X, it allows for perfect correlation with respect to the ambiguity covariance matrix Σ_M and the pure-risk covariance matrix $\Sigma_X - \Sigma_M$. In particular, each coordinate (a random variable) of the N-dimensional random vector M is a linear combination of some collection of $\operatorname{rank}\Sigma_M$ coordinates (random variables) of M. From the investor's viewpoint, therefore, $\operatorname{rank}\Sigma_M$ is the essential number of the sources of ambiguity in the N-asset markets. We will see in Section 4 that many seemingly suboptimal portfolios are in fact optimal for some ambiguity-averse investor even when there is essentially only one source of ambiguity, no matter how large the number N of assets may be.

Denote by $(a, b) \in \boldsymbol{R}^N \times \boldsymbol{R}$ a portfolio of these $N+1$ assets, representing the monetary amounts invested in each of these assets. Once the state is realized, the portfolio pays out $a^\top X + bR$. Denote the initial wealth by $W \in \boldsymbol{R}$. Let $\mathbf{1}$ be the vector in $\boldsymbol{R}^N$ of which the N coordinates are all equal to one. Then the budget constraint on the portfolio $(a, b) \in \boldsymbol{R}^N \times \boldsymbol{R}$ is $\mathbf{1}^\top a + b \leqslant W$. The decision maker's utility maximization problem is given by

$$\begin{aligned} \max_{(a,b)\in \boldsymbol{R}^N \times \boldsymbol{R}} \quad & U_{\gamma,\theta}(a^\top X + bR) \\ \text{subject to} \quad & \mathbf{1}^\top a + b \leqslant W. \end{aligned} \tag{7}$$

Define $V_{\gamma,\theta} : \boldsymbol{R}^N \times \boldsymbol{R} \to \boldsymbol{R}$ by letting

$$V_{\gamma,\theta}(a, b) = \mu_M^\top a + Rb - \frac{1}{2} a^\top \left(\gamma \Sigma_M + \theta \Sigma_{X|M}\right) a \tag{8}$$

for every $(a, b) \in \boldsymbol{R}^N \times \boldsymbol{R}$. Since $\Sigma_{X|M} = \Sigma_X - \Sigma_M$, this can be rewritten as

$$V_{\gamma,\theta}(a, b) = \mu_M^\top a + Rb - \frac{\theta}{2} a^\top \Sigma_X a - \frac{\gamma - \theta}{2} a^\top \Sigma_M a.$$

Thus, it is what MMR calls a robust mean-variance utility function. The following lemma shows that $V_{\gamma,\theta}$ represents the same preference ordering over the portfolios as $U_{\gamma,\theta}$.

Lemma 1 *For every $(a, b) \in \boldsymbol{R}^N \times \boldsymbol{R}$, $U_{\gamma,\theta}(a^\top X + bR) = u_\gamma\left(V_{\gamma,\theta}(a, b)\right)$.*

If (a, b) is a solution to the utility maximization problem (7), then $\mathbf{1}^\top a + b = W$. Hence, by Lemma 1, for every $(a, b) \in \boldsymbol{R}^N \times \boldsymbol{R}$, (a, b) is a solution to (7) if a is a solution to

$$\max_{a \in \boldsymbol{R}^N} V_{\gamma,\theta}(a, W - \mathbf{1}^\top a) \tag{9}$$

and $b = W - \mathbf{1}^\top a$. Since $\gamma \Sigma_M + \theta \Sigma_{X|M} \in \mathscr{S}^N_{++}$, the first-order condition gives the solution to the problem (7):

$$\begin{aligned} a &= (\gamma \Sigma_M + \theta \Sigma_{X|M})^{-1}(\mu_M - R\mathbf{1}) \qquad (10) \\ &= \left(\theta \Sigma_X + (\gamma - \theta)\Sigma_M\right)^{-1}(\mu_M - R\mathbf{1}) \\ &= \frac{1}{\theta}\left(\Sigma_X + \eta \Sigma_M\right)^{-1}(\mu_M - R\mathbf{1}). \qquad (11) \end{aligned}$$

The equality (24) of MMR is an equivalent characterization of the optimal portfolio. (11) shows that for any two pairs (Σ_M, η) and (Σ'_M, η') of an ambiguity covariance matrix and an ambiguity aversion coefficient, they share the same optimal portfolio if $\eta \Sigma_M = \eta' \Sigma'_M$.

If the investor is ambiguity-neutral, that is, $\gamma = \theta$, then

$$a = \frac{1}{\theta}\Sigma_X^{-1}(\mu_M - R\mathbf{1}),$$

and a is mean-variance efficient, that is, it attains the maximum of the expected excess return $(\mu_M - R\mathbf{1})^\top a$ of a portfolio a subject to the variance constraint of the form $a^\top \Sigma_X a \leqslant \sigma$ for any $\sigma > 0$. However, if the investor is ambiguity-averse, that is, $\gamma > \theta$, then a need not be mean-variance efficient. In Section 4, we examine to what extent ambiguity aversion can justify the investor's choice of a portfolio that is not mean-variance efficient.

We write $Q \equiv \Sigma_X^{-1}\Sigma_M$ and define $\zeta : (-1, \infty) \to \boldsymbol{R}^N$ by letting

$$\zeta(\eta) = (I + \eta Q)^{-1}\, \Sigma_X^{-1}(\mu_M - R\mathbf{1}) \tag{12}$$

for every $\eta \in (-1, \infty)$. Then the solution (10) to the problem (7) satisfies $a = \theta^{-1}\zeta(\eta)$. In other words, the function ζ tells us how the investor's portfolio depends on η. In particular, $\zeta(0) = \Sigma_X^{-1}(\mu_M - R\mathbf{1})$, and $\theta^{-1}\zeta(0)$ coincides with the optimal portfolio for the ambiguity-neutral investor.

The following characterization of eigenvectors of Q will be illustrative. Take any portfolio $v \in \boldsymbol{R}^N$. Then $\left(v^\top \Sigma_M v\right) / \left(v^\top \Sigma_X v\right)$ is the fraction of the variance of the return $v^\top X$ of the portfolio v that can be accounted for by the ambiguity M. This fraction lies between zero and one. Consider, then, the problem of finding a portfolio that minimizes this fraction:

$$\min_{v \in \boldsymbol{R}^N \setminus \{0\}} \quad \frac{v^\top \Sigma_M v}{v^\top \Sigma_X v}. \tag{13}$$

Denote a solution by v_1. Next, let $n \geqslant 2$ and $v_1, v_2, \ldots v_{n-1}$ be portfolios, and consider the problem of finding a portfolio that minimizes the fraction, subject to the constraint that the returns of the portfolios must be uncorrelated with every preceding portfolio v_m with $m \leqslant n-1$:

$$\begin{aligned} &\min_{v \in \boldsymbol{R}^N \setminus \{0\}} && \frac{v^\top \Sigma_M v}{v^\top \Sigma_X v} \\ &\text{s. t.} && v_m^\top \Sigma_X v = 0 \text{ for every } m \leqslant n-1. \end{aligned} \tag{14}$$

We say that a sequence $(v_1, v_2, \ldots, v_N)$ of portfolios is a sequence of solutions to the sequence of problems (14) if $v_1, v_2, \ldots, v_N$ are obtained iteratively by solving (13) and (14). There is indeed a sequence $(v_1, v_2, \ldots, v_N)$ of solutions to the sequence of problems (14), because the objective functions are continuous and the domains can further be restricted to $\{v \in \boldsymbol{R}^N \mid v^\top \Sigma_X v = 1\}$. Moreover, for every such sequence $(v_1, v_2, \ldots, v_N)$, the returns of the portfolios in the sequence are uncorrelated with each other and satisfy

$$\frac{v_1^\top \Sigma_M v_1}{v_1^\top \Sigma_X v_1} \leqslant \frac{v_2^\top \Sigma_M v_2}{v_2^\top \Sigma_X v_2} \leqslant \cdots \leqslant \frac{v_N^\top \Sigma_M v_N}{v_N^\top \Sigma_X v_N}.$$

The following proposition characterizes the eigenvalues and eigenvectors of Q as the sequences of solutions to the sequence of problems (14).

Proposition 1 *1. For every sequence $(v_1, v_2, \ldots, v_N)$ of solutions to the sequence of problems* (14) *and for every n, v_n is an eigenvector of Q and its corresponding eigenvalue is equal to $\left(v_n^\top \Sigma_M v_n\right) / \left(v_n^\top \Sigma_X v_n\right)$.*

2. For every sequence $(v_1, v_2, \ldots, v_N)$ of eigenvectors of Q, if their returns are uncorrelated with each other and the sequence of the corresponding eigenvalues is non-decreasing, then it is a sequence of solutions to the sequence of problems (14)*.*

3. The returns of the eigenvectors of Q that correspond to distinct eigenvalues are uncorrelated with each other.

4. All eigenvalues of Q belong to the closed unit interval $[0, 1]$*.*

This proposition states that the eigenvectors of Q can be obtained with the non-decreasing order of the corresponding eigenvalues by iteratively minimizing the fraction of variance of portfolio returns that can be accounted for by the ambiguity. It also implies that Q is, though not necessarily symmetric, diagonalizable.

3 Generalized mutual fund theorem

3.1 Theorem

The following theorem is a generalized version of the mutual fund theorem, which is applicable to the utility functions $U_{\gamma,\theta}$ with $\gamma \neq \theta$ and clarifies how the original version of the mutual fund theorem fails for investors who are not ambiguity-neutral. Mathematically, it expresses how the value $\zeta(\eta)$ of the function ζ defined by (12) depend on η. Since the optimal portfolio for the investor with the risk aversion coefficient θ and the ambiguity aversion coefficient η (whence $\gamma = (1+\theta)\eta$) coincides with $\theta^{-1}\zeta(\eta)$, it shows how the optimal portfolio changes as the ambiguity aversion coefficient changes, while the risk aversion coefficient is fixed. The total covariance matrix Σ_X and the ambiguity covariance matrix Σ_M are also fixed as they are embedded in the definition of ζ. In the language of KMM on page 1869, therefore, the theorem measures the pure effect of introducing greater ambiguity aversion into a given economic situation. We eliminate the case where $\mu_M - R\mathbf{1} = 0$, because the portfolio demand is, then, equal to zero for all values of γ and θ.

Theorem 1 (Generalized Mutual Fund Theorem) *Suppose that $\mu_M - R\mathbf{1} \neq 0$. Then there are a positive integer K and K eigenvectors $v_1, v_2, \dots, v_K$ of Q with distinct eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_K$ such that for every $\eta > -1$,*

$$\zeta(\eta) = \sum_{k=1}^{K} \frac{1}{1 + \lambda_k \eta} v_k. \tag{15}$$

Since the eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_K$ are distinct, it follows from part 3 of Proposition 1 that the returns of the portfolios $v_1, v_2, \dots, v_K$ are uncorrelated with each other, that is, $v_k^\top \Sigma_X v_\ell = 0$ whenever $k \neq \ell$. Thus, the number K of mutual funds cannot be greater than the number of distinct eigenvalues of Q, which cannot be greater than N. The theorem, thus, represents the optimal portfolios for all ambiguity-averse investors in terms of K mutual funds of which the returns are uncorrelated with each other.

This theorem is rich in interpretation. First, it is a generalized mutual fund theorem: there are K mutual funds, or portfolios of the N risky assets, $v_1, v_2, \dots, v_K$, that cater for all investors who exhibit any degrees of ambiguity aversion. Second, if $K = 1$, that is, $\zeta(0)$ is an eigenvector of Q, then the original mutual fund theorem holds: a single mutual fund v_1 is sufficient to satisfy all investors' portfolio demands, regardless of whether they are ambiguity-neutral or not. Third, if $\lambda_k > 0$, then the demand for the k-th mutual fund v_k decreases and converges to zero as the coefficient η of ambiguity aversion diverges to the infinity; but if $\lambda_k = 0$, then the demand for the k-th mutual fund does not depend on η. This should come as no surprise because, then, $v_k \in \operatorname{Ker} \Sigma_M$ and the return of v_k involves no ambiguity. Finally, since

$$\frac{(1 + \lambda_k \eta)^{-1}}{(1 + \lambda_\ell \eta)^{-1}} = \frac{1 + \lambda_\ell \eta}{1 + \lambda_k \eta} = \frac{\lambda_\ell}{\lambda_k} + \left(1 - \frac{\lambda_\ell}{\lambda_k}\right)\left(\frac{1}{1 + \lambda_k \eta}\right),$$

if $\lambda_k > \lambda_\ell$, then $(1 + \lambda_k \eta)^{-1}/(1 + \lambda_\ell \eta)^{-1}$ is a strictly decreasing function of η and converges to λ_ℓ/λ_k. Therefore, as $\eta \to \infty$, $\zeta(\eta)$ converges to v_k if $\lambda_k = 0$, and $\zeta(\eta)$ converges to 0 but tends to be proportional to $(\lambda_1^{-1}, \lambda_2^{-1}, \dots, \lambda_K^{-1})$ if $\lambda_k > 0$ for every k.

The theorem also allows us to decompose the expected excess return $\mu_N - R\mathbf{1}$ into two parts, one due to pure risks and the other due to ambiguity. Indeed, it follows from the definition of $\zeta(\eta)$ and $\theta a = \zeta(\eta)$ that

$$\mu_M - R\mathbf{1} = (\Sigma_X - \Sigma_M)\zeta(\eta) + (1 + \eta)\Sigma_M \zeta(\eta). \tag{16}$$

The first term in the right-hand side of (16) is the expected excess return that would induce the (ambiguity-neutral or not) investor to hold $\zeta(\eta)$ if the ambiguity were completely removed,[4] and the second term in the right-hand side of (16) is the expected excess return that would induce the ambiguity-averse investor to hold $\zeta(\eta)$ if the pure risks were completely removed.[5] Our contention here is that of the expected excess returns, the first term should be attributed to pure risks and the second to ambiguity, and each should be quantified based on empirical data, as we shall do in Section 5. By substituting the decomposition (15) into (16), we can rewrite the decomposition (16) as

$$\mu_M - R\mathbf{1} = \left(\sum_{k=1}^{K} \frac{1}{1+\lambda_k \eta} (\Sigma_X - \Sigma_M) v_k \right) + \left(\sum_{k=1}^{K} \frac{1+\eta}{1+\lambda_k \eta} \Sigma_M v_k \right),$$

which we can further rewrite as

$$\mu_M - R\mathbf{1} = \left(\sum_{k=1}^{K} \frac{1-\lambda_k}{1+\lambda_k \eta} \Sigma_X v_k \right) + \left(\sum_{k=1}^{K} \frac{\lambda_k + \lambda_k \eta}{1+\lambda_k \eta} \Sigma_X v_k \right),$$

because $\Sigma_M v_k = \lambda_k \Sigma_X v_k$. This decomposition implies that if $\lambda_k > 0$ for every k, then as $\eta \to \infty$, the right-hand side converges to

$$0 + \left(\sum_{k=1}^{K} \Sigma_X v_k \right) = 0 + (\mu_M - R\mathbf{1}).$$

That is, when the investor is extremely ambiguity-averse, the expected excess returns are entirely attributed to ambiguity. If $\lambda_k = 0$ for some k, then let, say, $\lambda_1 = 0$. Then, as $\eta \to \infty$, the right-hand side converges to

$$\Sigma_X v_1 + \Sigma_X \left(\sum_{k=2}^{K} v_k \right) = (\Sigma_X - \Sigma_M) v_1 + \Sigma_M \left(\sum_{k=2}^{K} \frac{1}{\lambda_k} v_k \right).$$

That is, if there is a portfolio that involves no ambiguity, then there is a nonzero part of the expected excess returns that can be attributed to pure risk, even when the investor is extremely ambiguity-averse.

[4] In this hypothetical situation, the ambiguous mean vector is deterministic and the covariance matrix of asset returns, as well as its conditional covariance matrix given the ambiguous mean vector, coincides with $\Sigma_X - \Sigma_M$.

[5] In this hypothetical situation, the covariance matrix of the ambiguous mean vector coincides with Σ_M and the asset returns, once conditioned on the ambiguous mean vector, are deterministic.

3.2 Special cases of one or two mutual funds

As we noted right after Theorem 1, the number K of mutual funds is bounded from above by the number of distinct eigenvalues of Q. In this subsection, we give sufficient conditions, in terms of Σ_X and Σ_M, for there to exist at most two distinct eigenvalues.

Proposition 2 *1. Let K be a positive integer and suppose that there are K distinct nonnegative numbers $\lambda_1, \lambda_2, \ldots, \lambda_K$ such that $\operatorname{rank}(\lambda_k \Sigma_X - \Sigma_M) < N$ for every k and*

$$\sum_{k=1}^{K} \operatorname{rank}(\lambda_k \Sigma_X - \Sigma_M) = (K-1)N.$$

Then, for every k, there exists a $v_k \in \operatorname{Ker}(\lambda_k \Sigma_X - \Sigma_M)$ such that, for every $\eta > -1$,

$$\zeta(\eta) = \sum_{k=1}^{K} \frac{1}{1+\lambda_k \eta} v_k.$$

2. If there is a $\lambda \geqslant 0$ such that $\lambda \Sigma_X = \Sigma_M$, then there is a $v \in \boldsymbol{R}^N$ such that, for every $\eta > -1$,

$$\zeta(\eta) = \frac{1}{1+\lambda\eta} v.$$

3. If $0 < \operatorname{rank}\Sigma_M < N$ and there is a $\lambda > 0$ such that $\operatorname{rank}\Sigma_M + \operatorname{rank}(\lambda\Sigma_X - \Sigma_M) = N$, then there are a $v_{\mathrm{R}} \in \operatorname{Ker}\Sigma_M$ and a $v_{\mathrm{A}} \in \operatorname{Ker}(\Sigma_X - \Sigma_M)$ such that, for every $\eta > -1$,

$$\zeta(\eta) = v_{\mathrm{R}} + \frac{1}{1+\lambda\eta} v_{\mathrm{A}}.$$

4. If $1 = \operatorname{rank}\Sigma_M < N$, then there are a $\lambda > 0$, a $v_{\mathrm{R}} \in \operatorname{Ker}\Sigma_M$, and a $v_{\mathrm{A}} \in \operatorname{Ker}(\Sigma_X - \Sigma_M)$ such that, for every $\eta > -1$,

$$\zeta(\eta) = v_{\mathrm{R}} + \frac{1}{1+\lambda\eta} v_{\mathrm{A}}.$$

Part 1 of Proposition 2 is similar to, but different from, Theorem 1 in that it is stated in terms of Σ_X and Σ_M and that Q is assumed to be diagonalizable, while proved to be so in Theorem 1 via Proposition 1. It serves as a general fact from which the mutual funds in the special cases of parts 2 and 3 are derived. Part 2 shows that if Σ_M is a scaled down version of Σ_X, then a single mutual fund is sufficient. Part 3 gives a sufficient condition under which two mutual funds are sufficient. The subscripts R and A of v_{R} and v_{A} stand for risk and ambiguity. They do indeed make sense, because the demand for v_{R} does not depend on the coefficient η of ambiguity aversion, while the demand for v_{A} vanishes as η

becomes large without bounds. Part 4 is a special case of part 3, which shows that the condition of part 3 is met whenever there is essentially only one source of ambiguity.

It is worthwhile to dwell on the special case of parts 3 and 4 of Proposition 2 in which $\lambda = 1$. The result can then be simplified to

$$\zeta(\eta) = v_{\mathrm{R}} + \frac{1}{1+\eta} v_{\mathrm{A}}.$$

Moreover, the decomposition (16) can be rewritten as

$$\mu_M - R\mathbf{1} = \Sigma_X v_{\mathrm{R}} + \Sigma_X v_{\mathrm{A}} = (\Sigma_X - \Sigma_M) v_{\mathrm{R}} + \Sigma_M v_{\mathrm{A}},$$

which has the virtue of being independent of η. This special case with $\lambda = 1$ can be used in factor models, in which the factors are traded assets and purely ambiguous (that is, ambiguous but not risky) and the idiosyncratic shocks are purely risky (that is, risky but not ambiguous).[6] The single-factor model of this type is a special case of part 4 with $\lambda = 1$.

The following corollary of Proposition 2 shows how an increase in the coefficient η of ambiguity aversion affect optimal portfolios. Denote by $\zeta_n(\eta)$ the n-the coordinate of $\zeta(\eta)$.

Corollary 1 1. *If there is a $\lambda \geqslant 0$ such that $\lambda \Sigma_X = \Sigma_M$, then, for every n, the sign of $\zeta_n(\eta)$ does not depend on $\eta > -1$. If, in addition, $\lambda \neq 0$, then $|\zeta_n(\eta)|$ converges strictly decreasingly to 0 as $\eta \to \infty$.*

2. *Let L be a positive integer smaller than N. Write $X = (\check{X}, \widehat{X})$ with $\check{X}$ being L-dimensional and $\widehat{X}$ being $(N-L)$-dimensional. Write*

$$\Sigma_X = \begin{pmatrix} \Sigma_{\check{X}} & \Sigma_{\check{X}\widehat{X}} \\ \Sigma_{\widehat{X}\check{X}} & \Sigma_{\widehat{X}} \end{pmatrix},$$

and similarly for M. Suppose that $\Sigma_{\check{M}} = 0$ and there is a $\lambda \geqslant 0$ such that $\Sigma_{\widehat{M}} = \lambda \left(\Sigma_{\widehat{X}} - \Sigma_{\widehat{X}\check{X}} \Sigma_{\check{X}}^{-1} \Sigma_{\check{X}\widehat{X}} \right)$. Then, for every $n > L$, the sign of $\zeta_n(\eta)$ does not depend

[6] To see this point more formally, let $X \sim \mathscr{N}(\mu_X, \Sigma_X)$ with $\Sigma_X \in \mathscr{S}^N_{++}$, and $A \in \boldsymbol{R}^{N \times L}$ with rank $A = L$. We take the returns $A^\top X$ of L portfolios as factor returns. Define $B = \left(A^\top \Sigma_X A\right)^{-1} A^\top \Sigma_X \in \boldsymbol{R}^{L \times N}$. Then B represents the factor loadings of the N assets. Define $M = (AB)^\top X + (I_N - (AB))^\top \mu_X$. Then M represents the parts of the asset returns that can be accounted for by the factor returns. Let $\mu_M = \mu_X$ and $\Sigma_M = \Sigma_X A \left(A^\top \Sigma_X A\right)^{-1} A^\top \Sigma_X$. Then (6) is met. Moreover, the factor returns are purely ambiguous and the idiosyncratic shocks, which are the parts of asset returns that cannot be attributed to the factor returns, are purely risky. Furthermore, since $\Sigma_M A = \Sigma_X A$, rank Σ_M + rank $(\Sigma_X - \Sigma_M) = N$ and part 3 is applicable.

on $\eta > -1$. If, in addition, $\lambda \neq 0$, then $|\zeta_n(\eta)|$ converges strictly decreasingly to 0 *as $\eta \to \infty$.*

Part 1 of Corollary 1 states that when a single mutual fund is sufficient, the sign of the holding in any asset does not depend on the ambiguity aversion coefficients and, unless it is equal to zero, its absolute value converges to zero as the coefficient becomes large without bounds. Part 2 of Corollary 1 obtains the same result in a more complicated case, in which two mutual funds are sufficient, as can be seen in the proof. To see what the case is like, note that the assumption that $\Sigma_{\widecheck{M}} = 0$ means that the returns of the first L assets are unambiguous and that $\Sigma_{\widehat{X}} - \Sigma_{\widehat{X}\widecheck{X}}\Sigma_{\widecheck{X}}^{-1}\Sigma_{\widecheck{X}\widehat{X}}$ is the conditional (total) covariance matrix of the returns of the last $N - L$ assets given the returns of the first L assets. Thus, $\Sigma_{\widehat{X}} - \Sigma_{\widehat{X}\widecheck{X}}\Sigma_{\widecheck{X}}^{-1}\Sigma_{\widecheck{X}\widehat{X}}$ is the conditional (total) covariance matrix of the returns of the ambiguous assets given the returns of the unambiguous assets. The existence of a $\lambda > 0$ such that $\Sigma_{\widehat{M}} = \lambda\left(\Sigma_{\widehat{X}} - \Sigma_{\widehat{X}\widecheck{X}}\Sigma_{\widecheck{X}}^{-1}\Sigma_{\widecheck{X}\widehat{X}}\right)$ means that the parts of the returns of the ambiguous assets that cannot be accounted for by the returns of the unambiguous assets is, up to scalar multiplication, due solely to the ambiguity of the expected returns of the ambiguous assets. This part is applicable whenever $L = N - 1$, because, then, $\Sigma_{\widehat{M}}$ and $\lambda\left(\Sigma_{\widehat{X}} - \Sigma_{\widehat{X}\widecheck{X}}\Sigma_{\widecheck{X}}^{-1}\Sigma_{\widecheck{X}\widehat{X}}\right)$ are scalars. This part, therefore, generalizes Proposition 8 of MMR, which concentrates on the case where $N = 2$ and $L = 1$, in two respects. First, the number of assets (and those of purely risky assets and ambiguous assets) is arbitrary. Second, the assumption is given only in terms of the sings of optimal holdings for the ambiguous assets.

4 Implied ambiguity covariance matrix and ambiguity aversion

4.1 Paradox on the market portfolio

As we noted in the introduction, many empirical studies have found that, unlike the market portfolio, the mean-variance efficient portfolios typically involve large short positions. On a more analytical note, Brennan and Lo (2010) proved that if the covariance matrix is randomly chosen from some parameterized family of positive definite matrices according to some probability distribution, then the probability that every portfolio on the mean-variance efficiency frontier involves at least one short position converges to one as the number of assets increases without bounds.[7] Since an ambiguity-neutral investor would

[7]Note that the random choice of covariance matrices is devoid of equilibrium consideration. This was a point of the subsequent discussions by Levy and Roll (2014), Ingersoll (2014), and Brennan and Lo

optimally choose a mean-variance efficient portfolio, these findings are not consistent with the hypothesis that the representative investor, being the holder of the market portfolio, is ambiguity-neutral. On the other hand, Theorem 1 shows that the optimal portfolio for an ambiguity-averse investor may not be mean-variance-efficient. We are thus led to ask whether and to what extent ambiguity aversion solves the apparent paradox that the market portfolio is not mean-variance-efficient.

We give two answers in this section. The first one is a simple equivalent condition for a given portfolio, of which the market portfolio is our prime example, to be optimal for some pair of an ambiguity covariance matrix and an ambiguity aversion coefficient. As it turns out, whenever there is one such pair, there are more than one. The second answer is, in senses to be made precise later, the minimal ambiguity covariance matrix and the minimal ambiguity aversion coefficient with which the given portfolio is optimal. Finding these minimal ones is important as they measure the minimal deviation from ambiguity neutrality that is necessary to justify the optimality of the market portfolio.

In the subsequent analysis, we first take the total covariance matrix Σ_X, the mean vector μ_M, the risk-free rate R, and a particular portfolio a as given. Then we infer a coefficient θ of absolute risk aversion, a coefficient η of ambiguity aversion, and an ambiguity covariance matrix Σ_M that together make a given portfolio a optimal for the ambiguity-averse investor, so that

$$a = \frac{1}{\theta}(\Sigma_X + \eta\Sigma_M)^{-1}(\mu_M - R\mathbf{1}). \tag{17}$$

From a mathematical viewpoint, we solve equation (17) for θ, Σ_M, and η, with a given portfolio a and given parameter values R, μ_M, and Σ_X. From a decision-theoretic viewpoint, we show how much of the total covariance matrix Σ_X the investor attributes to ambiguity. We further show what extent the investor is averse to risk and ambiguity if portfolio a is optimal for him.

As we already noted after (11), we cannot disentangle ambiguity from ambiguity aversion, as Epstein (2010) claimed. In fact, if a pair (Σ_M, η) is a solution to (17), then another pair (Σ'_M, η') is also a solution to (17) whenever $\eta\Sigma_M = \eta'\Sigma'_M$. However, we cannot treat Σ_M and η completely symmetrically, in the following sense: if a is optimal with the ambiguity covariance matrix Σ_M and the ambiguity aversion coefficient η, then, for every $\tau \in (0, 1)$, a is also optimal with the ambiguity covariance matrix $\tau\Sigma_M$ and the ambiguity aversion coefficient $\tau^{-1}\eta$, but a need not be optimal with the ambiguity covariance matrix $\tau^{-1}\Sigma_M$ and the ambiguity aversion coefficient $\tau\eta$, because $\Sigma_X - \tau^{-1}\Sigma_M$ need not be positive semidefinite. For this reason, the smallest ambiguity

(2014).

aversion coefficient η with which a is optimal is well defined, but the smallest (or minimal) ambiguity covariance matrix Σ_M with which a is optimal is not well defined.

We could instead take, say, the covariance matrix $\Sigma_{X|M} = \Sigma_X - \Sigma_M$ of unambiguous (purely risky) asset returns as given. But we opt for taking Σ_X as given, because if we fixed $\Sigma_{X|M}$ and chose Σ_M, then the total covariance matrix $\Sigma_X = \Sigma_{X|M} + \Sigma_M$ would be changed according to our choice of Σ_M. This would undermine our argument that ambiguity aversion solves the paradox of the inefficient market portfolio, because our choice of Σ_M changes the set of mean-variance efficient portfolios through changes in Σ_X. A consequence of fixing Σ_X but varying Σ_M is that the covariance matrix $\Sigma_{X|M}$ of unambiguous asset returns is also changed. In the setup of KMM, to which our setup is shown in Subsection 2.1 to be equivalent, this would mean that a change in the subjective probability distribution on the domain of second-order acts leads to a change in the conditional distributions over consequences given the realizations in the domain of second-order acts. In search for an implied covariance matrix Σ_M, therefore, we need to take this secondary effect into consideration and to see the quantitative implications on implied ambiguity and ambiguity aversion.[8]

4.2 When can a given portfolio be optimal?

The following theorem establishes a necessary and sufficient condition for a given portfolio vector $a \in \boldsymbol{R}^N$ to be optimal for some ambiguity-averse investor.

Theorem 2 *Let $\Sigma_X \in \mathscr{S}^N_{++}$, $\mu_M \in \boldsymbol{R}^N$, $R \in \boldsymbol{R}$, and $a \in \boldsymbol{R}^N$. For each $\theta > 0$, define*

$$v^\theta = \frac{1}{\theta}\Sigma_X^{-1}(\mu_M - R\mathbf{1}) - a \in \boldsymbol{R}^N \tag{19}$$

[8] To see the difference in quantitative implications, imagine that we wish to estimate Σ_M based on sample means and covariances of asset returns. In our approach, we take Σ_X as the sample covariance matrix and use (17) to infer Σ_M. In the alternative approach, we would take $\Sigma_{X|M}$ as the sample covariance matrix to infer Σ_M. Note that (17) is equivalent to

$$a = \frac{1}{\theta}\left(\Sigma_{X|M} + (\eta+1)\Sigma_M\right)^{-1}(\mu_M - R\mathbf{1}). \tag{18}$$

Comparing the two expressions (17) and (18) of the optimal portfolio, we can see that if we multiply $1 + \eta^{-1}$ to any ambiguity covariance matrix inferred in the alternative approach, then we can obtain an ambiguity covariance matrix that may be inferred in our approach. Conversely, if we divide by $1 + \eta^{-1}$ any ambiguity covariance matrix inferred in our approach, then we can obtain an ambiguity covariance matrix that may be inferred in the alternative approach. The factor $1 + \eta^{-1}$, therefore, quantifies the secondary effect of taking Σ_X as fixed.

and suppose that there is no $\theta > 0$ for which $v^\theta = 0$. Write

$$\bar{\theta} = \frac{a \cdot (\mu_M - R\mathbf{1})}{a^\top \Sigma_X a} \in \boldsymbol{R}. \tag{20}$$

If $\bar{\theta} > 0$, define, for each $\theta \in \left(0, \bar{\theta}\right)$,

$$\Sigma_M^\theta = \frac{1}{\left(v^\theta\right)^\top \Sigma_X v^\theta} \left(\Sigma_X v^\theta\right) \left(\Sigma_X v^\theta\right)^\top \in \mathscr{S}_+^N \quad \textit{and} \quad \eta^\theta = \frac{\left(v^\theta\right)^\top \Sigma_X v^\theta}{a^\top \Sigma_X v^\theta} \in \boldsymbol{R}_{++}. \tag{21}$$

1. *Assume that $a \cdot (\mu_M - R\mathbf{1}) > 0$. For every $(\Sigma_M, \eta, \theta) \in \mathscr{S}_+^N \times \boldsymbol{R}_{++} \times \boldsymbol{R}_{++}$, if $\theta < \bar{\theta}$, $v^\top(\eta\Sigma_M)v \geqslant v^\top(\eta^\theta \Sigma_M^\theta)v$ for every $v \in \boldsymbol{R}^N$, and $a^\top(\eta\Sigma_M)a = a^\top(\eta^\theta\Sigma_M^\theta)a$, then (Σ_M, η, θ) satisfies (17).*

2. *Assume that $\mu_M - R\mathbf{1} \neq 0$. For every $(\Sigma_M, \eta, \theta) \in \mathscr{S}_+^N \times \boldsymbol{R}_{++} \times \boldsymbol{R}_{++}$, if (Σ_M, η, θ) satisfies (17) and $v^\top \Sigma_X v \geqslant v^\top \Sigma_M v$ for every $v \in \boldsymbol{R}^N$, then $a \cdot (\mu_M - R\mathbf{1}) > 0$, $\theta < \bar{\theta}$, $\eta \geqslant \eta^\theta$, $v^\top (\eta\Sigma_M) v \geqslant v^\top \left(\eta^\theta \Sigma_M^\theta\right) v$ for every $v \in \boldsymbol{R}^N$, and $a^\top(\eta\Sigma_M)a = a^\top(\eta^\theta\Sigma_M^\theta)a$.*

Part 1 of Theorem 2 shows that a is the optimal portfolio of some ambiguity-averse investor whenever $a \cdot (\mu_M - R\mathbf{1}) > 0$, that is, the expected excess return of the portfolio a is positive. It implies, in particular, that a is optimal for an investor with the ambiguity covariance matrix Σ_M^θ, ambiguity aversion coefficient η^θ, and risk aversion coefficient θ. Part 2 shows that the sufficient conditions identified in Part 1 for the optimality of a for some ambiguity-averse investor are also necessary. Together, they provide the range $\left(0, \bar{\theta}\right)$ of risk aversion coefficients with which such an ambiguity-averse investor exists.

Theorem 2 characterizes the ambiguity-covariance matrix Σ_M and ambiguity aversion coefficient η with which the given portfolio a is optimal, for each fixed risk aversion coefficient $\theta \in (0, \bar{\theta})$. Among these candidates, we shall focus, in the subsequent analysis, on the ambiguity-covariance matrix Σ_M^θ and ambiguity aversion coefficient η^θ that are defined in (21). The main reason is that the aversion-adjusted ambiguity matrix $\eta^\theta\Sigma_M^\theta$ is the smallest one in the sense that $\eta\Sigma_M - \eta^\theta\Sigma_M^\theta$ is positive semidefinite, and the ambiguity aversion coefficient η^θ is the smallest one, for every pair of Σ_M and η that satisfies the condition of the theorem,[9] and we are interested in minimal deviations from the expected utility functions that makes the portfolio a optimal.

We now justify our informal arugment, towards the end of Subsection 1.1, on how to construct ambiguity, by relating the elements of the ambiguity covariance matrix Σ_M^θ, defined by (21), to the means, variances, and covariances of asset returns. Moreover, we characterize the decomposition of the optimal portfolio and the expected excess returns

[9] Thus, every Schatten norm (of which the spectral norm and the Frobenius norm are special cases) of $\eta^\theta\Sigma_M^\theta$ is smaller than or equal to that of $\eta\Sigma_M$.

in the forms we represented after the mutual fund theorem (Theorem 1). Furthermore, when the portfolio a is the market portfolio, we restate these results in terms of the alphas and betas of the N assets.

In order to identify the elements of the ambiguity matrix Σ_M^θ, it is sufficient to know how large each diagonal element is and whether each off-diagonal elements is positive or negative, because Σ_M^θ is of rank one. Indeed, denote the (n, m)-element of Σ_M^θ by σ_{nm}^θ. Then $\sigma_{nm}^\theta = \left(\sigma_{nn}^\theta \sigma_{mm}^\theta\right)^{1/2}$ if it is positive, and $\sigma_{nm}^\theta = -\left(\sigma_{nn}^\theta \sigma_{mm}^\theta\right)^{1/2}$ if it is negative. Note also that the sign of σ_{nm}^θ is equal to that of $\sigma_{nk}^\theta \sigma_{km}^\theta$ for every k. To identify the signs of the off-diagonal elements, therefore, it is sufficient to partition the set $\{1, 2, \dots, N\}$ of all assets into two subsets, within each of which the assets are perfectly positively correlated with each other.

Denote the n-th coordinates of μ_M, a, and $\Sigma_X a$ by μ_n, a_n, and c_n. These are the expected return of asset n, the position of asset n in the portfolio a, and the covariance between the return of portfolio a and the return of asset n. Define two sets of assets, S_+^θ and S_-^θ, by

$$S_+^\theta = \left\{ n \;\middle|\; \frac{1}{\theta}(\mu_n - R) > c_n \right\} \quad \text{and} \quad S_-^\theta = \left\{ n \;\middle|\; \frac{1}{\theta}(\mu_n - R) < c_n \right\}.$$

Since

$$\Sigma_M^\theta v^\theta = \frac{1}{\theta}(\mu_M - R\mathbf{1}) - \Sigma_X a,$$

the following proposition can be easily derived from the definition (21) of Σ_M^θ.

Proposition 3 *Define $\bar{\theta}$ by* (20) *and, for each $\theta \in (0, \bar{\theta})$, define v^θ, Σ_M^θ and η^θ by* (19) *and* (21).

1. *For every $\theta \in (0, \bar{\theta})$, there is a $\kappa^\theta > 0$ such that $\sigma_{nn}^\theta = \kappa^\theta (\theta^{-1}(\mu_n - R) - c_n)^2$ for every n.*

2. *If $n \in S_+^\theta$ and $m \in S_+^\theta$, then $\sigma_{nm}^\theta > 0$. If $n \in S_-^\theta$ and $m \in S_-^\theta$, then $\sigma_{nm}^\theta > 0$. If $n \in S_+^\theta$ and $m \in S_-^\theta$, then $\sigma_{nm}^\theta < 0$.*

Part 1 of Proposition 3 shows that the ambiguity of asset n is related to the deviation of the covariance of the return of the given portfolio a with the return of asset n from the expected excess return of asset n, divided by the risk aversion coefficient θ. Part 2 shows that for any pair of assets, if they both belong to either S_+^θ or S_-^θ, then their returns, as measured by the ambiguity covariance matrix, are perfectly positively correlated, but if one belongs to S_+^θ and the other to S_-^θ, then their returns are perfectly negatively correlated.

The results of Proposition 3 can be better interpreted when a is the normalized market portfolio, so that $\mathbf{1}^\top a = 1$, that is, each coordinate a_n represents the proportion of the value (market capitalization) of asset n in the market portfolio. Then the beta β_n of asset n is defined by c_n divided by the variance of the market portfolio, and the so called alpha α_n of the asset n is defined as the part of the expected excess return that cannot be explained by its beta:

$$\beta_n = \frac{c_n}{a^\top \Sigma_X a} \quad \text{and} \quad \alpha_n = (\mu_n - R) - (a \cdot (\mu_M - R\mathbf{1}))\beta_n. \tag{22}$$

There are an asset n such that $\alpha_n > 0$ and another asset m such that $\alpha_m < 0$ unless a is mean-variance-efficient. Define two sets of assets, T_+^θ and T_-^θ, by

$$T_+^\theta = \left\{ n \mid \alpha_n + (a^\top \Sigma_X a)(\bar{\theta} - \theta)\beta_n > 0 \right\},$$
$$T_-^\theta = \left\{ n \mid \alpha_n + (a^\top \Sigma_X a)(\bar{\theta} - \theta)\beta_n < 0 \right\}.$$

The following corollary follows from Proposition 3 and the equality

$$(\mu_n - R) - \theta c_n = \alpha_n + (a^\top \Sigma_X a)(\bar{\theta} - \theta)\beta_n.$$

Corollary 2 *Define $\bar{\theta}$ by* (20) *and, for each $\theta \in (0, \bar{\theta})$, define v^θ, Σ_M^θ and η^θ by* (19) *and* (21)*.*

1. *For every $\theta \in (0, \bar{\theta})$, there is a $\kappa^\theta > 0$ such that $\sigma_{nn}^\theta = \kappa^\theta(\alpha_n + (a^\top \Sigma_X a)(\bar{\theta} - \theta)\beta_n)^2$ for every n.*

2. *If $n \in T_+^\theta$ and $m \in T_+^\theta$, then $\sigma_{nm}^\theta > 0$. If $n \in T_-^\theta$ and $m \in T_-^\theta$, then $\sigma_{nm}^\theta > 0$. If $n \in T_+^\theta$ and $m \in T_-^\theta$, then $\sigma_{nm}^\theta < 0$.*

Part 1 of Corollary 2 presents the diagonal elements of the ambiguity matrix Σ_M^θ in terms of the alphas and betas. It shows that the size of ambiguity of asset n depends not only on its alpha α_n but also its beta β_n. Part 2 identifies the two groups of assets such that the ambiguity of two assets is perfectly positively correlated if they belong to the same group, but it is perfectly negatively correlated if they belong to the different groups. As can be seen from the definitions of T_+^θ and T_-^θ, it is the ratio α_n/β_n, if well defined, that determines to which group the asset belongs.

Having characterized the ambiguity covariance matrix Σ_M^θ, we now characterize the optimal portfolio a in the form (15) of the generalized mutual fund theorem (Theorem 1). Define $Q^\theta = \Sigma_X^{-1}\Sigma_M^\theta$ and define $\zeta^\theta : (-1, \infty) \to \boldsymbol{R}^N$ by letting

$$\zeta^\theta(\eta) = \left(I_N + \eta Q^\theta\right)^{-1} \Sigma_X^{-1} (\mu_M - R\mathbf{1})$$

for each $\eta > -1$. As we already see in parts 3 and 4 of Proposition 2, the portfolio a is decomposed into two mutual funds, which are defined by

$$v_{\mathrm{A}}^{\theta} = \frac{v^{\theta} \cdot (\mu_M - R\mathbf{1})}{(v^{\theta})^{\top} \Sigma_X v^{\theta}} v^{\theta} \qquad \text{and} \qquad v_{\mathrm{R}}^{\theta} = \Sigma_X^{-1} (\mu_M - R\mathbf{1}) - v_{\mathrm{A}}^{\theta}, \tag{23}$$

where subscripts R and A stand for risk and ambiguity. Denote the n-th coordinates of v_{A}^{θ}, v_{R}^{θ}, and $\Sigma_X^{-1}(\mu_M - R\mathbf{1})$ by $v_{\mathrm{A}n}^{\theta}$, $v_{\mathrm{R}n}^{\theta}$, and d_n.

Proposition 4 *Define $\bar{\theta}$ by* (20) *and, for each $\theta \in (0, \bar{\theta})$, define v^{θ}, Σ_M^{θ} and η^{θ} by* (19) *and* (21).

1. *For every $\eta > -1$,*
$$\zeta^{\theta}(\eta) = v_{\mathrm{R}}^{\theta} + \frac{1}{1+\eta} v_{\mathrm{A}}^{\theta}. \tag{24}$$

2. *For every $\theta \in (0, \bar{\theta})$, there is a $\kappa^{\theta} > 0$ such that $v_{\mathrm{A}n}^{\theta} = \kappa^{\theta}(\theta^{-1} d_n - a_n)$ for every n.*

Part 1 of Proposition 4 gives the decomposition of the optimal portfolio into two mutual funds in the form (15) of the generalized mutual fund theorem (Theorem 1). Part 2 shows that the position of asset n in the ambiguity-driven portfolio v_{A}^{θ} is determined by how much the position n is different from the position d_n in the mean-variance-efficient portfolio. Since $v^{\theta} \cdot (\mu_M - R\mathbf{1}) > 0$, v_{A}^{θ} is a positive multiple of v^{θ}. This part, therefore, follows from the definition (19) of v^{θ}.

The final result of this subsection is concerned with the decomposition of the expected excess returns $\mu_M - R\mathbf{1}$, stated after Theorem 1:

$$\mu_M - R\mathbf{1} = \left(\Sigma_X - \Sigma_M^{\theta}\right) v_{\mathrm{R}}^{\theta} + \Sigma_M^{\theta} v_{\mathrm{A}}^{\theta} \tag{25}$$

The first term on the right-hand side of (25) is the expected excess return that would induce the (ambiguity-neutral or ambiguity-averse) investor to hold the portfolio a if the ambiguity were completely removed. The second term is the expected excess return that would induce the ambiguity-averse investor to hold the portfolio a if the risk were completely removed. The following proposition gives some important properties of these terms. We denote the n-th coordinates of $\Sigma_M^{\theta} v_{\mathrm{A}}^{\theta}$ by $\left(\Sigma_M^{\theta} v_{\mathrm{A}}^{\theta}\right)_n$.

Proposition 5 *Define $\bar{\theta}$ by* (20) *and, for each $\theta \in (0, \bar{\theta})$, define v^{θ}, Σ_M^{θ} and η^{θ} by* (19) *and* (21). *Then, for every $\theta \in (0, \bar{\theta})$, there is a $\kappa^{\theta} > 0$ such that*

$$\left(\Sigma_M^{\theta} v_{\mathrm{A}}^{\theta}\right)_n = \kappa^{\theta}((\mu_n - R) - \theta c_n) = \kappa^{\theta}(\alpha_n + (a^{\top} \Sigma_X a)(\bar{\theta} - \theta)\beta_n)$$

for every n.

Proposition 5 shows the ambiguity-related expected return of asset n is proportional to the deviation of the covariance of the portfolio a with the return of asset n from the expected return of asset n. We omit the proof of this proposition.

4.3 Minimal ambiguity and minimal ambiguity aversion that are independent of risk aversion coefficients

According to Part 2 of Theorem 2, the ambiguity aversion coefficient η^θ and the aversion-adjusted ambiguity covariance matrix $\eta^\theta \Sigma_M^\theta$ are the smallest ones that make a optimal. Note, however, they are smallest within the class of the pairs of Σ_M and η given a risk aversion coefficient θ. In this subsection, we consider the problems of minimizing the ambiguity aversion coefficient η^θ and the largest eigenvalue of the aversion-adjusted ambiguity covariance matrix $\eta^\theta \Sigma_M^\theta$ by varying θ, thereby finding the smallest of the ambiguity aversion coefficients and the smallest of the largest eigenvalues of the aversion-adjusted ambiguity covariance matrices that are independent of the choice of $\theta \in (0, \bar{\theta})$. We solve not just one but the two minimization problems because there are two conflicting interpretations of ambiguity covariance matrices and ambiguity aversion coefficients. KMM and Klibanoff, Marinacci, and Mukerji (2012) argued that these two are separable, while Epstein (2010) cast some doubts on the argument of KMM. The minimization of the ambiguity aversion coefficient is motivated by KMM's argument, while the minimization of the largest eigenvalue of the aversion-adjusted ambiguity covariance matrix is motivated by Epstein's argument.

We should add a caveat to this exercise. As we mentioned in Subsection 2.1, Theorem 2 of KMM showed that an investor is more ambiguity-averse than another if and only if the function $\varphi_{\gamma,\theta}$ is more concave for the first one than for the second. As KMM repeatedly emphasized, however, the theorem is valid only when the two investors are assumed to have the same risk attitude. Thus, although their Theorem 2 is valid to Theorem 2 of this paper, it is not applicable to our Theorem 3 below. In particular, even when $\eta^\theta > \eta^{\theta'}$ for two risk aversion coefficients θ and θ', we cannot say that the investor with the utility function $U_{(1+\eta^\theta)\theta,\theta}$ is more ambiguity-averse than the investor with the utility function $U_{(1+\eta^{\theta'})\theta',\theta'}$. Nonetheless, we shall consider the problem of minimizing η^θ (and the largest eigenvalue of $\eta^\theta \Sigma_M^\theta$) by varying θ for two reasons. First, as our generalized mutual fund theorem (Theorem 1) shows that the composition of investment among the N assets with random returns, which is of our prime concern, is determined by the ambiguity aversion coefficient η and independent of the risk aversion coefficient θ. Second, since the ambiguity aversion coefficient η can be interpreted as the elasticity of marginal utilities, we can, as we will do in Subsection 5.1, assess whether a particular value of η is reasonable based on this interpretation.

Theorem 3 *Let Σ_X, μ_M, a, v^θ, $\bar{\theta}$, Σ_M^θ, and η^θ be as in Theorem 2. Assume that $a \cdot (\mu_M - R\mathbf{1}) > 0$.*

1. *As a function of $\theta \in (0, \bar{\theta})$, the largest eigenvalue of $\eta^\theta \Sigma_M^\theta$ is minimized at*

$$\theta^* = \left(\bar{\theta}^{-1} + \frac{\|\bar{\theta}^{-1}(\mu_M - R\mathbf{1}) - \Sigma_X a\|}{\|\mu_M - R\mathbf{1}\|} \right)^{-1},$$

and, at $\theta = \theta^*$, it is equal to

$$\frac{2}{a \cdot (\mu_M - R\mathbf{1})} \left(\|\bar{\theta}^{-1}(\mu_M - R\mathbf{1}) - \Sigma_X a\| \, \|\mu_M - R\mathbf{1}\| - \left(\bar{\theta}^{-1}(\mu_M - R\mathbf{1}) - \Sigma_X a\right) \cdot (\mu_M - R\mathbf{1}) \right). \tag{26}$$

2. *As a function of $\theta \in (0, \bar{\theta})$, η^θ is minimized at*

$$\theta^{**} = \left(\bar{\theta}^{-1} + \left(\frac{\left(\bar{\theta}^{-1}(\mu_M - R\mathbf{1}) - \Sigma_X a\right)^\top \Sigma_X^{-1} \left(\bar{\theta}^{-1}(\mu_M - R\mathbf{1}) - \Sigma_X a\right)}{(\mu_M - R\mathbf{1})^\top \Sigma_X^{-1} (\mu_M - R\mathbf{1})} \right)^{1/2} \right)^{-1},$$

and $\eta^{\theta^{**}}$ is equal to

$$\begin{aligned} &\frac{2}{a \cdot (\mu_M - R\mathbf{1})} \\ &\times \left[\left(\left(\bar{\theta}^{-1}(\mu_M - R\mathbf{1}) - \Sigma_X a\right)^\top \Sigma_X^{-1} \left(\bar{\theta}^{-1}(\mu_M - R\mathbf{1}) - \Sigma_X a\right) \right)^{1/2} \left((\mu_M - R\mathbf{1})^\top \Sigma_X^{-1} (\mu_M - R\mathbf{1}) \right)^{1/2} \right. \\ &\qquad \left. - \left(\bar{\theta}^{-1}(\mu_M - R\mathbf{1}) - \Sigma_X a\right)^\top \Sigma_X^{-1} (\mu_M - R\mathbf{1}) \right]. \end{aligned} \tag{27}$$

Theorem 3 gives the minimum of the largest eigenvalue of the averion-adjusted ambiguity matrix $\eta \Sigma_M$ and the minimum of the ambiguity aversion η. Since $\eta^\theta \Sigma_M^\theta$ is of rank one, the largest eigenvalue is the only strictly positive eigenvalue (and zero is the other eigenvalue, with multiplicity $N - 1$), the spectral norm of $\eta^{\theta^*} \Sigma_M^{\theta^*}$ is smaller than or equal to that of $\eta \Sigma_M$, even when (17) holds for (Σ_M, η, θ) with $\theta \neq \theta^*$. However, for such a (Σ_M, η, θ), $\eta \Sigma_M - \eta^{\theta^*} \Sigma_M^{\theta^*}$ need not be positive semidefinite.

It is easy to check that $\bar{\theta}$ in Theorem 2 is homogeneous of degree -1 in a, that is, if a is multiplied by factor $\tau > 0$, then it is multiplied by factor τ^{-1}. Thus, by the definition (19), $v^{\bar{\theta}}$ is homogeneous of degree 1 in a. By Theorem 3, both θ^* and θ^{**} are homogeneous of degree -1 in a. On the other hand, none of the ambiguity matrices $\Sigma_M^{\theta^*}$ and $\Sigma_M^{\theta^{**}}$ and the ambiguity aversion coefficients η^{θ^*} and $\eta^{\theta^{**}}$ is affected by any scalar multiplication to

a. In particular, any normalization of portfolio a, such as $\mathbf{1}^\top a = 1$, would not affect any of these ambiguity parameters, although it does change risk aversion parameters θ^* and θ^{**}. Our numerical results in Section 5, in particular, do not depend on the normalization of the market portfolio a.

5 Examples based on the U.S. equity market data

5.1 Data

Tables 1 shows the sample means, variances, and covariances of the FF6 portfolios in the U.S. equity markets, obtained from Ken French's website. Table 2 reports the mean-variance-efficient portfolio, based on the data of Table 1, and the market portfolio, defined as the time-series average of ME weights. Throughout this section, we take the risk-free rate R, the expected return vector μ_M, and the total covariance matrix Σ_X to be the the sample means and covariances in Table 1. We then find the ambiguity covariance matrix Σ_M^θ and the ambiguity aversion coefficient η^θ, defined in (21), with which the market portfolio weight in Table 2 is optimal.[10]

Theorem 2 shows the necessary and sufficient condition for a portfolio to be optimal for some ambiguity-averse investor. With our data, $\mu_M - R\mathbf{1} \in \boldsymbol{R}_{++}^N$. Moreover, if a is the market portfolio in Table 2, then $a \in \boldsymbol{R}_{++}^N$, $a \cdot (\mu_M - R\mathbf{1}) > 0$, and $\Sigma_X a \in \boldsymbol{R}_{++}^N$ because all elements of Σ_X are strictly positive. Then the market portfolio a satisfies the necessary and sufficient condition $a \cdot (\mu_M - R\mathbf{1}) > 0$.

Table 3 reports the CAPM alphas and betas. The distinguish property is that the SN and SH portfolios have significantly positive alphas and the SL portfolio has a significantly negative alpha. Nonzero alphas have been reported in many empirical studies, such as Black, Jensen, and Scholes (1972) and Fama and French (1992).

5.2 Implied coefficients of risk and ambiguity aversion

Theorem 2 finds the upper the upper bound $\bar{\theta}$, defined in (20), on the coefficients of constant absolute risk aversion with which the market portfolio a is optimal. With our data, the upper bond $\bar{\theta}$ is equal to 0.0226.

Theorem 3 finds the coefficients θ^* and θ^{**} of constant absolute risk aversion that minimize the largest eigenvalue of the aversion-adjusted ambiguity covariance matrix $\eta^\theta \Sigma_M^\theta$ and the coefficient η^θ of ambiguity aversion. With our sample data, $\theta^* = 0.0192$

[10] We concentrate on these $\left(\Sigma_M^\theta, \eta^\theta\right)$'s, rather than arbitrary (Σ_M, η)'s, because the former represents minimal deviations from the ambiguity-neutral investors, as explained in Section 4.

Table 3: Characteristics of the FF6 portfolios

n	θ_n	$\hat{\theta}_n$	α_n	β_n	α_n/β_n	a_n
SL	0.019	−3.45	−0.15	1.20	−0.12	0.025
SN	0.028	2.95	0.18	1.16	0.16	0.030
SH	0.030	1.36	0.29	1.30	0.22	0.021
BL	0.022	0.09	−0.03	0.93	−0.03	0.507
BN	0.022	−0.09	−0.01	1.00	−0.01	0.312
BH	0.025	−0.23	0.08	1.19	0.07	0.106

Table 3 reports the θ_n and the $\hat{\theta}_n$ defined in (28), as well as the CAPM alphas and betas, defined by (22), their ratios, and the weights in the market portfolio, based on the μ_M, Σ_X, and R.

and $\theta^{**} = 0.0124$. Theorem 3 also shows the corresponding values of ambiguity aversion coefficients η^{θ^*} and $\eta^{\theta^{**}}$. With our data, $\eta^{\theta^*} = 16.69$ and $\eta^{\theta^{**}} = 9.31$.

The latter ambiguity aversion coefficient, $\eta^{\theta^{**}}$, is the smallest coefficient of ambiguity aversion with which the market portfolio is optimal. Since it is the elasticity of marginal utilities from conditional expected utilities given M, whether the value is plausible may be judged based on the literature on the equity premium puzzle of Mehra and Prescott (1985). As put forward by Kocherlakota (1996), the consensus seems that a theoretical model is considered as being consistent with empirical data only if the prediction of the model coincides with the data when involving the coefficient of relative risk aversion, which is nothing but the elasticity of marginal utilities from consumption, is no higher than ten. We can therefore conclude that our model of an ambiguity-averse representative investor is consistent with the U.S. equity market data to the extent that the elasticity of marginal utilities from conditional expected utilities is comparable to that from consumption.

In the next two subsections, we quantitatively determine the ambiguity variances and covariances, as well as the decomposition of the optimal portfolio and the expected excess returns. Before doing so, we relate the obtained risk aversion coefficients to some parameters that characterize FF6 portfolios.

As we did in Section 4, denote the n-th coordinates of μ_M, a, $\Sigma_X a$, and $\Sigma_X^{-1}(\mu_M - R\mathbf{1})$ by μ_n, a_n, c_n, and d_n. Then $\mu_n - R > 0$, $a_n > 0$, and $c_n > 0$ for every n. We can thus define θ_n as the ratio of the expected excess return of asset n to its covariance with the market portfolio, and $\hat{\theta}_n$ as the ratio of the position of asset n in the mean-variance-efficient portfolio to the position of asset n in the optimal portfolio a:

$$\theta_n = \frac{\mu_n - R}{c_n} \in \boldsymbol{R}_{++} \quad \text{and} \quad \hat{\theta}_n = \frac{d_n}{a_n} \in \boldsymbol{R}. \tag{28}$$

The θ_n are not all equal, because a is not mean-variance-efficient. Moreover, there are an asset n for which $\theta_n < \bar{\theta}$ and, at the same time, another asset n for which $\theta_n > \bar{\theta}$, because and the weighted averages of the θ_n with weights $(a_n c_n)\left(a^\top \Sigma_X a\right)^{-1}$ is equal to $\bar{\theta}$. The same can be said of for the $\hat{\theta}_n$. Table 3 reports the values of θ_n, $\hat{\theta}_n$, and a_n, from which we can rank the values:

$$0 < \theta^{**} < \theta_{\mathrm{SL}} < \theta^{*} < \theta_{\mathrm{BL}} < \theta_{\mathrm{BN}} < \bar{\theta} < \theta_{\mathrm{BH}} < \theta_{\mathrm{SN}} < \theta_{\mathrm{SH}} \tag{29}$$

$$\hat{\theta}_{\mathrm{SL}} < \hat{\theta}_{\mathrm{BH}} < \hat{\theta}_{\mathrm{BN}} < 0 < \theta^{**} < \theta^{*} < \bar{\theta} < \hat{\theta}_{\mathrm{BL}} < \hat{\theta}_{\mathrm{SH}} < \hat{\theta}_{\mathrm{SN}}. \tag{30}$$

5.3 Implied ambiguity covariance matrices

Part 1 of Proposition 3 characterizes the ambiguity of each asset, given a particular risk aversion coefficient $\theta \in (0, \bar{\theta})$. For $\theta = \theta^*$ and $\theta = \theta^{**}$, Table 4 shows the values of the diagonal elements σ^{θ}_{nn} of Σ^{θ}_M, which represents the ambiguity of the returns of the FF6 portfolios. According to part 1 of Proposition 3, the diagonal elements σ^{θ}_{nn} are proportional to the $c_n^2(\theta_n - \theta)^2$. However, the orders of diagonal elements σ^{θ}_{nn} are the same as the order of the values of θ_n in (29) for both $\theta = \theta^*$ and $\theta = \theta^{**}$. The result shows, therefore, that the covariances c_n of the return of the market portfolio a with the return of each FF6 portfolio just happen to be irrelevant to determine the order of ambiguity.

Table 4: Diagonal elements of the ambiguity covariance matrix

n	$\sigma^{\theta^*}_{nn}$	$\sigma^{\theta^{**}}_{nn}$
SL	0.01	1.49
SN	2.73	8.23
SH	5.29	13.34
BL	0.15	1.96
BN	0.28	2.61
BH	1.24	5.70

Table 4 reports the the values of the diagonal elements of the aversion-adjusted ambiguity covariance matrix $\eta^{\theta}\Sigma^{\theta}_M$ when θ is set equal to θ^* and θ^{**} to minimize the largest eigenvalue of $\eta^{\theta}\Sigma^{\theta}_M$ and η^{θ}.

Part 2 of Proposition 3 characterizes the correlations of ambiguity. According to the order in (29), if $\theta = \theta^{**}$, then the ambiguity of all FF6 portfolios are perfectly positively correlated with each other. That is, $\sigma^{\theta}_{nm} > 0$ for all $n, m \in \{\mathrm{SL}, \mathrm{SN}, \mathrm{SH}, \mathrm{BL}, \mathrm{BN}, \mathrm{BH}\}$. If $\theta = \theta^*$, then the ambiguity of the SL portfolio is perfectly negatively correlated with the ambiguity of any other portfolio, and the ambiguity of the SN, SH, BL, BN, and BH portfolios is perfectly positively correlated with each other. That is, $\sigma^{\theta}_{\mathrm{SL},n} < 0$ for all $n \in \{\mathrm{SN}, \mathrm{SH}, \mathrm{BL}, \mathrm{BN}, \mathrm{BH}\}$.

Although the values of the diagonal elements of Σ^{θ}_M are reported in Table 4 only for

$\theta = \theta^*$ and $\theta = \theta^{**}$, it is illustrative to see how they change as θ varies on $(0, \bar{\theta})$. We claim that

$$\sigma^{\theta}_{\mathrm{BL,BL}} < \sigma^{\theta}_{\mathrm{BN,BN}} < \sigma^{\theta}_{\mathrm{SL,SL}} < \sigma^{\theta}_{\mathrm{BH,BH}} < \sigma^{\theta}_{\mathrm{SN,SN}} < \sigma^{\theta}_{\mathrm{SH,SH}}$$

for every θ sufficiently close to 0, and

$$\sigma^{\theta}_{\mathrm{BN,BN}} < \sigma^{\theta}_{\mathrm{BL,BL}} < \sigma^{\theta}_{\mathrm{BH,BH}} < \sigma^{\theta}_{\mathrm{SL,SL}} < \sigma^{\theta}_{\mathrm{SN,SN}} < \sigma^{\theta}_{\mathrm{SH,SH}}$$

for every θ sufficiently close to $\bar{\theta}$. This is because

$$\mu_{\mathrm{BL}} < \mu_{\mathrm{BN}} < \mu_{\mathrm{SL}} < \mu_{\mathrm{BH}} < \mu_{\mathrm{SN}} < \mu_{\mathrm{SH}},$$
$$|\alpha_{\mathrm{BN}}| < |\alpha_{\mathrm{BL}}| < |\alpha_{\mathrm{BH}}| < |\alpha_{\mathrm{SL}}| < |\alpha_{\mathrm{SN}}| < |\alpha_{\mathrm{SH}}|,$$

and

$$(\mu_n - R) - \theta a_n = \alpha_n + (a^{\top}\Sigma_X a)(\bar{\theta} - \theta)\beta_n = \begin{cases} \mu_n - R & \text{if } \theta = 0, \\ \alpha_n & \text{if } \theta = \bar{\theta}. \end{cases} \tag{31}$$

Moreover, since the SH and SN portfolios attain the highest expected returns and alphas, $\sigma^{\theta}_{nn} < \sigma^{\theta}_{\mathrm{SN,SN}} < \sigma^{\theta}_{\mathrm{SH,SH}}$ for every $\theta \in (0, \bar{\theta})$ and every $n \in \{\mathrm{BL, BN, BH, SL}\}$. That is, the SH portfolio has the largest ambiguity and the SN portfolio has the second largest ambiguity regardless of the values of θ. Although this is mainly due to the fact that these two portfolios have the highest expected returns and alphas, the role of the beta in affecting the order of the ambiguity variance as θ varies should not be missed. In addition, the ambiguity tend to be larger for the portfolios with high B/M's.

This last result should be contrasted with the assumption that Bossaerts, Ghirardato, Guarnaschelli, and Zame (2010) used when they explained the so called value effect, the observation that value stocks (stocks with high B/M's) tend to have higher returns than growth stocks (stocks with low B/M's). They argued that since firms of growth stocks have more growth potential, it seems natural to assume that their returns are more ambiguous than those of value stocks. They observed, in the experimental markets with heterogenous investors, that more risk-averse investors tend to be more ambiguity-averse as well. They deduced, from this observation, that growth stocks are held and priced primarily by investors who are less averse to both risk and ambiguity, while value stocks are held and priced more evenly by all investors in the market. They, then, concluded that value stocks have higher returns than growth stocks at equilibrium of heterogeneous investors. In contrast, we found, in the analysis of an ambiguity-averse representative agent based on the U.S. equity market data, that the value stocks tend to be more, rather than less, ambiguous.[11]

[11]It may appear that the difference in the findings is due to the difference in the models, in that

5.4 Optimal portfolios and expected excess returns

Proposition 4 decomposes the optimal portfolio into two mutual funds in the form the generalized mutual fund theorem (Theorem 1). In particular, the portfolio is decomposed into the ambiguity-driven portfolio and the risk-driven portfolio. The decomposition of the optimal portfolio is numerically reported in Table 5.

Table 5: Decomposition of the market portfolio

n	$(1/\theta^*)v_{\mathrm{R}}^{\theta^*}$	$(1/\theta^*)v_{\mathrm{A}}^{\theta^*}$	$(1/\theta^{**})v_{\mathrm{R}}^{\theta^{**}}$	$(1/\theta^{**})v_{\mathrm{A}}^{\theta^{**}}$
SL	0.29	−4.70	0.76	−7.60
SN	−0.24	4.77	−0.72	7.74
SH	−0.07	1.54	−0.22	2.50
BL	0.39	2.02	0.16	3.58
BN	0.42	−1.83	0.58	−2.78
BH	0.19	−1.46	0.33	−2.30

Table 5 lists the pure-risk portfolios v_{R}^{θ} and the ambiguity portfolios v_{A}^{θ} for $\theta^* = 0.0192$ and $\theta^{**} = 0.0124$.

With our sample data, by (28), the ambiguity-driven portfolio $v_{\mathrm{A}n}^{\theta}$ is positive for every $\theta \in (0, \bar{\theta})$ and $n \in \{\mathrm{SN}, \mathrm{SH}, \mathrm{BL}\}$, and is negative for every $\theta \in (0, \bar{\theta})$ and $n \in \{\mathrm{SL}, \mathrm{BN}, \mathrm{BH}\}$. In other words, the sign pattern of the ambiguity-driven portfolio $v_{\mathrm{A}n}^{\theta}$ is independent of θ.

Proposition 5 shows the ambiguity-related expected return of asset n. The decomposition is numerically reported in Table 6. Because of the ranking (29), the ambiguity-induced expected excess returns of the SL, BL, and BN portfolios are strictly negative, and those of the BH, SN, and SH portfolios are strictly positive.

As regards to how these positions change as θ varies over $(0, \bar{\theta})$, we make three claims. First, for every θ sufficiently close to 0,

$$\left(\Sigma_M^{\theta} v_{\mathrm{A}}^{\theta}\right)_{\mathrm{BL}} < \left(\Sigma_M^{\theta} v_{\mathrm{A}}^{\theta}\right)_{\mathrm{BN}} < \left(\Sigma_M^{\theta} v_{\mathrm{A}}^{\theta}\right)_{\mathrm{SL}} < 0 < \left(\Sigma_M^{\theta} v_{\mathrm{A}}^{\theta}\right)_{\mathrm{BH}} < \left(\Sigma_M^{\theta} v_{\mathrm{A}}^{\theta}\right)_{\mathrm{SN}} < \left(\Sigma_M^{\theta} v_{\mathrm{A}}^{\theta}\right)_{\mathrm{SH}}.$$

Second, for every θ sufficiently close to $\bar{\theta}$,

$$\left(\Sigma_M^{\theta} v_{\mathrm{A}}^{\theta}\right)_{\mathrm{SL}} < \left(\Sigma_M^{\theta} v_{\mathrm{A}}^{\theta}\right)_{\mathrm{BL}} < \left(\Sigma_M^{\theta} v_{\mathrm{A}}^{\theta}\right)_{\mathrm{BN}} < 0 < \left(\Sigma_M^{\theta} v_{\mathrm{A}}^{\theta}\right)_{\mathrm{BH}} < \left(\Sigma_M^{\theta} v_{\mathrm{A}}^{\theta}\right)_{\mathrm{SN}} < \left(\Sigma_M^{\theta} v_{\mathrm{A}}^{\theta}\right)_{\mathrm{SH}}.$$

they used a model of heterogeneous investors while we use a representative-investor model. However, the difference in the models should not really account for the the difference in findings, because the representative investor's ambiguity-averse utility function would reflect (the heterogeneity of) individual investors' ambiguity aversion as long as markets were complete. The problem of characterizing the representative investor's ambiguity aversion by maximizing a weighted sum of individual investors' utilities seems, though, largely unexplored in the literature.

Table 6: Decomposition of the expected excess returns

n	$\left(\Sigma_X - \Sigma_M^{\theta*}\right) v_{\mathrm{R}}^{\theta*}$	$\Sigma_M^{\theta*} v_{\mathrm{A}}^{\theta*}$	$\left(\Sigma_X - \Sigma_M^{\theta**}\right) v_{\mathrm{R}}^{\theta**}$	$\Sigma_M^{\theta**} v_{\mathrm{A}}^{\theta**}$
SL	0.72	−0.02	0.44	0.26
SN	0.67	0.32	0.39	0.61
SH	0.75	0.45	0.43	0.77
BL	0.55	0.07	0.33	0.30
BN	0.59	0.10	0.35	0.34
BH	0.70	0.22	0.41	0.51

Table 6 lists the pure-risk-induced expected excess returns $\left(\Sigma_X - \Sigma_M^{\theta}\right) v_{\mathrm{R}}^{\theta}$ and the ambiguity-induced expected excess returns $\Sigma_M^{\theta} v_{\mathrm{A}}^{\theta}$ for for $\theta* = 0.0192$ and $\theta** = 0.0124$.

Third, for every $\theta \in (0, \bar{\theta})$ and every $n \in \{\mathrm{BL}, \mathrm{BN}, \mathrm{SL}\}$,

$$(\Sigma_M^{\theta} v_A^{\theta})_n < 0 < (\Sigma_M^{\theta} v_A^{\theta})_{\mathrm{BH}} < (\Sigma_M^{\theta} v_A^{\theta})_{\mathrm{SN}} < (\Sigma_M^{\theta} v_A^{\theta})_{\mathrm{SH}}.$$

That is, the SH portfolio has the largest ambiguity-induced expected excess return, the SN portfolio has the second largest one, and the BH has the third largest one, regardless of the values of θ. All these results follow from (31) and

$$\mu_{\mathrm{BL}} < \mu_{\mathrm{BN}} < \mu_{\mathrm{SL}} < \mu_{\mathrm{BH}} < \mu_{\mathrm{SN}} < \mu_{\mathrm{SH}},$$
$$\alpha_{\mathrm{SL}} < \alpha_{\mathrm{BL}} < \alpha_{\mathrm{BN}} < \alpha_{\mathrm{BH}} < \alpha_{\mathrm{SN}} < \alpha_{\mathrm{SH}}.$$

This follows from the fact that the top-three rankings in terms of the alphas and expected returns are the same. The bottom-three rankings depends on the value of θ because the betas of the BL, BN, and SL portfolios are different.

5.5 Summary of our numerical results

In this section, we used the data on the FF6 portfolios of the U.S. equity markets to calculate the ambiguity aversion coefficients and the ambiguity covariance matrix with which the market portfolio is optimal. There are two main results obtained from our numerical analysis, one with regards to the minimum ambiguity aversion coefficient and the other with regards to the distributions of ambiguity in expected returns across the FF6 portfolios. First, the minimum ambiguity aversion coefficient (with which the market portfolio is optimal for some specification of the ambiguity covariance matrix) is within the range of values deemed plausible in view of the contributions of the equity premium puzzle. Second, the returns of the portfolios with small market equities, especially those having high and neutral ratios of the book equity to the market equity, tend to be more ambiguous and positively correlated with each other, while they may be negatively correlated with the returns of those with big market equities and neutral or low ratios of book to market equities.

6 Conclusion

Using the ambiguity-inclusive CARA-normal setup of an investor who has an ambiguity-averse utility function of the form of KMM and MMR, we have generalized the mutual fund theorem to accommodate ambiguity; identified necessary and sufficient conditions for a given, not mean-variance efficient, portfolio to be optimal for some ambiguity-averse investor; characterized all the pairs of ambiguity covariance matrices and ambiguity aversion coefficients with which the give portfolio is optimal; and found the minimal ones. We have numerically derived the minimal ambiguity aversion coefficient and the ambiguity covariance matrix based on the U.S. equity market data.

There are a couple of directions of future research. The first one is to rationalize the so-called $1/N$ portfolio, the portfolio in which the investment is split equally among all the assets by ambiguity aversion. This is important because the $1/N$ portfolio is often regarded as a rule of thumb, rather than derived from the optimization behavior, but its empirical performance is no worse than the sample-based mean-variance-efficient portfolio, as shown, for example, by DeMiguel, Garlappi, and Uppal (2009). The second one is to develop a factor model that explicitly takes ambiguity of factors and idiosyncratic shocks into account. Incorporating ambiguity into a factor model allows us to analyze the optimal portfolio for an investor who believe the validity of the factor model probabilistically, rather than deterministically. We should aim at generalizing such factor models as those of Pástor (2000), Pástor and Stambaugh (2000), and Wang (2005).

A Lemmas and Proofs for Section 2

Proof of Lemma 1 By the properties of the moment generating function,

$$\begin{aligned}
&E\left[u_\theta\left(a^\top X + bR\right)|M\right] \\
&= -\exp(-\theta bR) E\left[\exp\left((-\theta a)^\top X\right)|M\right] \\
&= -\exp(-\theta bR)\exp\left((-\theta a)^\top M + \frac{1}{2}(\theta a)^\top \Sigma_{X|M}(\theta a)\right) \\
&= -\exp\left(-\theta\left(a^\top M + Rb - \frac{\theta}{2}a^\top \Sigma_{X|M} a\right)\right).
\end{aligned}$$

Then it follows from (4) that

$$\varphi_{\gamma,\theta}\left(E\left[u_\theta\left(a^\top X + bR\right)|M\right]\right) = -\exp\left(-\gamma\left(a^\top M + Rb - \frac{\theta}{2}a^\top \Sigma_{X|M} a\right)\right).$$

Thus, again by the properties of the moment generating function,

$$\begin{aligned}
&U_{\gamma,\theta}(a^\top X + bR)\\
=&E\left[-\exp\left(-\gamma\left(a^\top M + Rb - \frac{\theta}{2}a^\top \Sigma_{X|M} a\right)\right)\right]\\
=&-\exp\left(-\gamma\left(Rb - \frac{\theta}{2}a^\top \Sigma_{X|M} a\right)\right)E\left[\exp\left((-\gamma a)^\top M\right)\right]\\
=&-\exp\left(-\gamma\left(Rb - \frac{\theta}{2}a^\top \Sigma_{X|M} a\right)\right)\exp\left(\mu_M^\top(-\gamma a) + \frac{1}{2}(-\gamma a)^\top \Sigma_M(-\gamma a)\right)\\
=&u_\gamma\left(V_{\gamma,\theta}(a,b)\right).
\end{aligned}$$

///

To prove Proposition 1, we need a lemma. To state it, let $P \in \mathscr{S}^N$ and consider the problem of minimizing the quadratic form defined by P:

$$\min_{w \in \boldsymbol{R}^N \setminus \{0\}} \frac{w^\top P w}{\|w\|^2}. \tag{32}$$

Denote a solution by w_1. Next, let $n \geqslant 2$ and $w_1, w_2, \dots w_{n-1}$ belong to $\boldsymbol{R}^N$, and consider the problem of minimizing the quadratic form subject to the constraint that the solution must be orthogonal to $w_1, w_2, \dots, w_{n-1}$:

$$\begin{aligned}
\min_{w \in \boldsymbol{R}^N \setminus \{0\}} \quad & \frac{w^\top P w}{\|w\|^2}\\
\text{s. t.} \quad & w_m \cdot w = 0 \text{ for every } m \leqslant n-1.
\end{aligned} \tag{33}$$

We say that a sequence $(w_1, w_2, \dots, w_N)$ of vectors in $\boldsymbol{R}^N$ is a sequence of solutions (33) if $w_1, w_2, \dots, w_N$ are obtained iteratively by solving (32) and (33). There is a sequence of solutions to the problems (33), because the objective functions are continuous and the domains can be restricted to $\{w \in \boldsymbol{R}^N \mid \|w\| = 1\}$. Moreover, for every sequence $(w_1, w_2, \dots, w_N)$ of solutions, $(w_1, w_2, \dots, w_N)$ is orthogonal and

$$\frac{w_1^\top P w_1}{\|w_1\|^2} \leqslant \frac{w_2^\top P w_2}{\|w_2\|^2} \leqslant \cdots \leqslant \frac{w_N^\top P w_N}{\|w_N\|^2}$$

for every n. Since the following lemma, which characterizes the eigenvectors and eigenvalues of P, is well known,[12] we omit the proof.

Lemma 2 *1. For every sequence $(w_1, w_2, \ldots, w_N)$ of solutions to the sequence of problems (33) and for every n, w_n is an eigenvector of P and its corresponding eigenvalue is equal to $w_n^\top P w_n / \|w_n\|^2$.*

2. For every sequence $(w_1, w_2, \ldots, w_N)$ of eigenvectors of P that is orthogonal and of which the sequence of the corresponding eigenvalues is non-decreasing is a sequence of solutions to the sequence of problems (33).

3. If $w_1, w_2, \ldots, w_K$ are eigenvectors of P that correspond to distinct eigenvectors, then $(w_1, w_2, \ldots, w_K)$ is orthogonal.

Proposition 1 can be proved using this lemma as follows.

Proof of Proposition 1 Define $P = \Sigma_X^{-1/2} \Sigma_M \Sigma_X^{-1/2}$. Then, $P \in \mathscr{S}^N$. Moreover, assume that $w \in \boldsymbol{R}^N$, $v \in \boldsymbol{R}^N$, $\lambda \in \boldsymbol{R}$ and $v = \Sigma_X^{-1/2} w$. Then

$$\frac{v^\top \Sigma_M v}{v^\top \Sigma_X v} = \frac{w^\top P w}{\|w\|^2},$$

and $(\Sigma_X^{-1} \Sigma_M) v = \lambda v$ if and only if $Pw = \lambda w$. Thus, parts 1 and 2 of this proposition follow from parts 1 and 2 of Lemma 2. If, in addition, $w' \in \boldsymbol{R}^N$, $v' \in \boldsymbol{R}^N$, and $v' = \Sigma_X^{-1/2} w'$, then $v^\top \Sigma_X v = w \cdot w'$. Thus, part 3 of this proposition follows from part 3 of Lemma 2. It remains to prove part 4. Let λ be an eigenvalue of Q and v be a corresponding eigenvector. Then $\Sigma_M v = \lambda \Sigma_X v$. Thus $v^\top \Sigma_M v = \lambda v^\top \Sigma_X v$, that is,

$$\lambda = \frac{v^\top \Sigma_M v}{v^\top \Sigma_M v + v^\top \Sigma_{X|M} v}.$$

Since $v^\top \Sigma_{X|M} v \geqslant 0$, $0 \leqslant \lambda \leqslant 1$. ///

B Proofs for Section 3

Proof of Theorem 1 Let Λ be the set of all eigenvalues of Q. It follows from Proposition 1 that $\Lambda \subset [0, 1]$. For each $\lambda \in \Lambda$, denote by V_λ the eigenspace that correspond to λ. It also follows that V_λ is a linear subspace of $\boldsymbol{R}^N$, $(V_\lambda)_{\lambda \in \Lambda}$ is linearly independent (that

[12] It is used, for example, in Campbell, Lo, and MacKinlay (1997, Section 6.4), in which the minimum is replaced by the maximum.

is, if $v_\lambda \in V_\lambda$ for every $\lambda \in \Lambda$ and $\sum_{\lambda\in\Lambda} v_\lambda = 0$, then $v_\lambda = 0$ for every $\lambda \in \Lambda$), and $\sum_{\lambda\in\Lambda} V_\lambda = \boldsymbol{R}^N$.

Then, for each $\lambda \in \Lambda$, there is a $v_\lambda \in V_\lambda$ such that $\zeta(0) = \sum_{\lambda\in\Lambda} v_\lambda$. Since $\zeta(0) \neq 0$, there is a $\lambda \in \Lambda$ such that $v_\lambda \neq 0$. Let $\{\lambda_1, \lambda_2, \ldots, \lambda_K\}$ be the set of all such λ's. For each k, write $v_k = v_{\lambda_k}$, then $\zeta(0) = \sum_{k=1}^K v_k$.

Since $(I + \eta Q)v_k = (1 + \eta\lambda_k)v_k$, $(I + \eta Q)^{-1}v_k = (1 + \eta\lambda_k)^{-1}v_k$. Thus,

$$\zeta(\eta) = (I + \eta Q)^{-1}\zeta(0) = \sum_{k=1}^K (I + \eta Q)^{-1} v_k = \sum_{k=1}^K \frac{1}{1 + \lambda_k \eta} v_k.$$

///

Proof of Proposition 2 1. By the definition of Q, $\operatorname{Ker}(\lambda_k\Sigma_X - \Sigma_M) = \operatorname{Ker}(\lambda_k I_N - Q)$ and

$$\dim \operatorname{Ker}(\lambda_k I_N - Q) = N - \operatorname{rank}(\lambda_k\Sigma_X - \Sigma_M) > 0.$$

Thus $\operatorname{Ker}(\lambda_k\Sigma_X - \Sigma_M)$ is the eigenspace of Q that corresponds to eigenvalue λ_k. Moreover,

$$\begin{aligned}
&\sum_{k=1}^K \dim \operatorname{Ker}(\lambda_k\Sigma_X - \Sigma_M) \\
&= \sum_{k=1}^K (N - \operatorname{rank}(\lambda_k\Sigma_X - \Sigma_M)) \\
&= KN - \sum_{k=1}^K \operatorname{rank}(\lambda_k\Sigma_X - \Sigma_M) \\
&= KN - (K-1)N = N.
\end{aligned}$$

Thus $\boldsymbol{R}^N$ coincides with the direct sum of the K eigenspaces $\operatorname{Ker}(\lambda_k\Sigma_X - \Sigma_M)$. Thus $\lambda_1, \lambda_2, \ldots, \lambda_K$ are the eigenvalues of Q. Part 1 therefore follows from Theorem 1.

2. This is a special case of part 2 with $K = 1$.

3. Since $\operatorname{rank}\Sigma_M < N$ and $\operatorname{rank}(\Sigma_X - \Sigma_M) < N$, this is a special case of part 1 with $K = 2$ and $\{\lambda_1, \lambda_2\} = \{0, \lambda\}$.

4. Since $\operatorname{rank} Q = \operatorname{rank}\Sigma_M = 1$, by Proposition 1, there are a $\lambda > 0$ and $v \in \boldsymbol{R}^N\backslash\{0\}$ such that $Qv = \lambda v$, that is, $(\lambda\Sigma_X - \Sigma_M)v = 0$. Thus $\operatorname{rank}(\lambda\Sigma_X - \Sigma_M) \leqslant N$. Hence $\operatorname{rank}(\lambda\Sigma_X - \Sigma_M) + \operatorname{rank}\Sigma_M \leqslant N$. On the other hand, $\operatorname{rank}(\lambda\Sigma_X - \Sigma_M) + \operatorname{rank}\Sigma_M \geqslant \operatorname{rank}\lambda\Sigma_X = N$. Thus $\operatorname{rank}(\lambda\Sigma_X - \Sigma_M) + \operatorname{rank}\Sigma_M = N$. The conclusion follows then from part 3. ///

Proof of Corollary 1 1. This follows from part 2 of Proposition 2.

2. By assumption, $\operatorname{rank}\Sigma_M = N - L$. As for $\operatorname{rank}(\lambda\Sigma_X - \Sigma_M)$, recall from linear algebra that the row space of any matrix will not be changed by subtracting from row vectors any linear combinations of the other vectors. Thus

$$\begin{aligned}
&\operatorname{Row}(\lambda\Sigma_X - \Sigma_M) \\
&= \operatorname{Row}\begin{pmatrix} \Sigma_{\check{X}} & \Sigma_{\check{X}\hat{X}} \\ \Sigma_{\hat{X}\check{X}} & \Sigma_{\hat{X}} - \lambda^{-1}\Sigma_{\widehat{M}} \end{pmatrix} \\
&= \operatorname{Row}\begin{pmatrix} \Sigma_{\check{X}} & \Sigma_{\check{X}\hat{X}} \\ \Sigma_{\hat{X}\check{X}} - \Sigma_{\hat{X}\check{X}}\Sigma_{\check{X}}^{-1}\Sigma_{\check{X}} & \Sigma_{\hat{X}} - \lambda^{-1}\Sigma_{\widehat{M}} - \Sigma_{\hat{X}\check{X}}\Sigma_{\check{X}}^{-1}\Sigma_{\check{X}\hat{X}} \end{pmatrix} \\
&= \operatorname{Row}\begin{pmatrix} \Sigma_{\check{X}} & \Sigma_{\check{X}\hat{X}} \\ 0 & 0 \end{pmatrix}.
\end{aligned}$$

Thus $\operatorname{rank}(\lambda\Sigma_X - \Sigma_M) = L$. By part 2 of Proposition 2, there are a $v_{\mathrm{R}} \in \operatorname{Ker}\Sigma_M$ and a $v_{\mathrm{A}} \in \operatorname{Ker}(\Sigma_X - \Sigma_M)$ such that for every $\eta > -1$,

$$\zeta(\eta) = v_{\mathrm{R}} + \frac{1}{1+\lambda\eta} v_{\mathrm{A}}.$$

Since $\operatorname{Ker}\Sigma_M = \boldsymbol{R}^L \times \{0\}$, the n-th coordinate of v_{R} is equal to zero for every $n > L$. This part then follows from the above equality. ///

C Lemmas and Proofs for Section 4

Theorems 2 is proved via the following Lemma 3 and 4.

Lemma 3 *Let $a \in \boldsymbol{R}^N$ and $c \in \boldsymbol{R}^N$.*

1. *Assume that $a \cdot c > 0$. Define*

$$\Xi^* = (a \cdot c)^{-1} cc^\top \in \mathscr{S}_+^N, \tag{34}$$

then $\|c\|^2 (a \cdot c)^{-1}$ is an eigenvalue of Ξ^* of multiplicity 1 with a corresponding eigenvector c, and 0 is an eigenvalue of Ξ^* of multiplicity $N-1$.

2. *Assume that $a \cdot c > 0$. For every $\Xi \in \mathscr{S}_+^N$, if $v^\top \Xi v \geqslant v^\top \Xi^* v$ for every $v \in \boldsymbol{R}^N$ and $a^\top \Xi a = a^\top \Xi^* a$, then $\Xi a = c$, where Ξ^* is defined by (34).*

3. *Assume that $c \neq 0$. For every $\Xi \in \mathscr{S}_+^N$, if $\Xi a = c$, then $a \cdot c > 0$, $v^\top \Xi v \geqslant v^\top \Xi^* v$ for every $v \in \boldsymbol{R}^N$, and $a^\top \Xi a = a^\top \Xi^* a$, where Ξ^* is defined by (34).*

The first part of Lemma 3 defines a positive semidefinite matrix Ξ^* based on two vectors. Using it, the last two parts jointly gives a necessary and sufficient condition for a positive semidefinite matrix to map one of the two vectors to the other.

Proof of Lemma 3 1. This is straightforward, possibly except for the statement on multiplicity, but it follows directly from the fact that the column space of Ξ^* coincides with the line spanned by c.

2. Since $\Xi - \Xi^* \in \mathscr{S}_+^N$, $(\Xi - \Xi^*)^{1/2}$ is well defined. Moreover, $0 = a^\top (\Xi - \Xi^*)\, a = \left\|(\Xi - \Xi^*)^{1/2} a\right\|^2$. Thus $(\Xi - \Xi^*)^{1/2} a = 0$. Hence $\Xi a = \Xi^* a = c$.

3. Since $\Xi a = c \neq 0$, $a \cdot c = a^\top \Xi a = \|\Xi^{1/2} a\|^2 \geqslant 0$ and $\Xi^{1/2} a \neq 0$. Thus $\|\Xi^{1/2} a\| > 0$ and, hence, $a \cdot c > 0$. Since $\Xi^* a = c$, $a^\top \Xi^* a = a \cdot c = a^\top \Xi a$. It remains to prove that $v^\top \Xi v \geqslant v^\top \Xi^* v$ for every $v \in \boldsymbol{R}^N$. Note that $v^\top \Xi v \geqslant v^\top \Xi^* v$ if and only if $(v^\top \Xi v)(a \cdot c) \geqslant (c \cdot v)^2$. Since $\Xi a = c$, this is equivalent to $(v^\top \Xi v)(a^\top \Xi a) \geqslant (a^\top \Xi v)^2$, which can further be rewritten as $\left\|\Xi^{1/2} v\right\|^2 \left\|\Xi^{1/2} a\right\|^2 \geqslant \left(\left(\Xi^{1/2} v\right) \cdot \left(\Xi^{1/2} a\right)\right)^2$. This last inequality follows from the Cauchy-Schwarz inequality. ///

Lemma 4 *Let $a \in \boldsymbol{R}^N$, $c \in \boldsymbol{R}^N$, and $\Sigma \in \mathscr{S}_{++}^N$. For each $\theta > 0$, define*

$$v^\theta = \frac{1}{\theta} \Sigma^{-1} c - a$$

and assume that there is no $\theta > 0$ for which $v^\theta = 0$.

1. *Assume that $a \cdot c > 0$. Write $\bar{\theta} = (a^\top \Sigma a)^{-1} (a \cdot c)$. For each $\theta \in \left(0, \bar{\theta}\right)$, define*

$$\Gamma^\theta = \frac{1}{\left(v^\theta\right)^\top \Sigma v^\theta} \left(\Sigma v^\theta\right) \left(\Sigma v^\theta\right)^\top \in \mathscr{S}_+^N,$$
$$\eta^\theta = \frac{\left(v^\theta\right)^\top \Sigma \left(v^\theta\right)}{a^\top \Sigma v^\theta} \in \boldsymbol{R}_{++},$$

   then $\|\Sigma v^\theta\|^2 \left(a^\top \Sigma v^\theta\right)^{-1}$ is an eigenvalue of $\eta^\theta \Gamma^\theta$ of multiplicity 1 with a corresponding eigenvector Σv^θ, and 0 is an eigenvalue of $\eta^\theta \Gamma^\theta$ of multiplicity $N-1$. Moreover, $v^\top \Sigma v \geqslant v^\top \Gamma^\theta v$ for every $v \in \boldsymbol{R}^N$.

2. *Assume that $a \cdot c > 0$. For every $(\Gamma, \eta, \theta) \in \mathscr{S}_+^N \times \boldsymbol{R}_{++} \times \boldsymbol{R}_{++}$, if $\theta < \bar{\theta}$, $v^\top (\eta \Gamma) v \geqslant v^\top (\eta^\theta \Gamma^\theta) v$ for every $v \in \boldsymbol{R}^N$, and $a^\top (\eta \Gamma) a = a^\top (\eta^\theta \Gamma^\theta) a$, then $\theta(\Sigma + \eta^\theta \Gamma^\theta) a = c$, where $\bar{\theta}$, Γ^θ, and η^θ are defined as in part 1.*

3. *Assume that $c \neq 0$. For every $(\Gamma, \eta, \theta) \in \mathscr{S}_+^N \times \boldsymbol{R}_{++} \times \boldsymbol{R}_{++}$, if $\theta(\Sigma + \eta \Gamma) a = c$ and $v^\top \Sigma v \geqslant v^\top \Gamma v$ for every $v \in \boldsymbol{R}^N$, then $a \cdot c > 0$, $\theta < \bar{\theta}$, $\eta \geqslant \eta^\theta$, $v^\top (\eta \Gamma) v \geqslant v^\top \left(\eta^\theta \Gamma^\theta\right) v$*

for every $v \in \boldsymbol{R}^N$, and $a^\top(\eta\Gamma)a = a^\top\left(\eta^\theta\Gamma^\theta\right)a$, where $\bar{\theta}$, Γ^θ, and η^θ are defined as in part 1.

Proof of Lemma 4 1. For every $\theta \in \left(0, \bar{\theta}\right)$, $a \cdot (\Sigma v^\theta) = \theta^{-1} a \cdot c - a^\top \Sigma a > 0$. Moreover,

$$\eta^\theta \Gamma^\theta = \frac{1}{a^\top \Sigma v^\theta}\left(\Sigma v^\theta\right)\left(\Sigma v^\theta\right)^\top \in \mathscr{S}_+^N.$$

This part, therefore, follows from part 1 of Lemma 3, except that $v^\top \Sigma v \geqslant v^\top \Gamma^\theta v$ for every $v \in \boldsymbol{R}^N$. To show this, note first that it is equivalent to the condition that if v^* is a solution to the problem

$$\begin{array}{ll} \max\limits_{v \in \boldsymbol{R}^N} & v^\top \Gamma^\theta v \\ \text{subject to} & v^\top \Sigma v \leqslant 1, \end{array}$$

then $(v^*)^\top \Gamma^\theta v^* \leqslant 1$. Indeed, by the definition of Γ^θ, the objective function can be replaced by $\left|\left(\Sigma v^\theta\right) \cdot v\right|$. Thus, by the first-order condition, either v^* or $-v^*$ coincides with

$$\left(\left(v^\theta\right)^\top \Sigma v^\theta\right)^{-1/2} \Sigma^{-1}\left(\Sigma v^\theta\right) = \left(\left(v^\theta\right)^\top \Sigma v^\theta\right)^{-1/2} v^\theta.$$

It is then easy to show that

$$(v^*)^\top \Gamma^\theta v^* = \frac{\left(\left(v^\theta\right)^\top \Sigma v^\theta\right)^2}{\left(\left(v^\theta\right)^\top \Sigma v^\theta\right)^2} = 1.$$

2. This part follows from part 2 of Lemma 3.

3. This part follows from part 3 of Lemma 3, except that $a \cdot c > 0$, $\theta < \bar{\theta}$, and $\eta \geqslant \eta^\theta$. First, since $a \cdot (\Sigma v^\theta) > 0$, $a \cdot c > \theta a^\top \Sigma a$. Thus $a \cdot c > 0$ and $\theta < \bar{\theta}$. As for $\eta \geqslant \eta^\theta$, since $v^\top \Sigma v \geqslant v^\top \Gamma v$ for every $v \in \boldsymbol{R}^N$ and $(v^\theta)^\top \Sigma v^\theta = (v^\theta)^\top \Gamma^\theta v^\theta$,

$$\begin{aligned} \eta - \eta^\theta &= \left((v^\theta)^\top \Sigma v^\theta\right)^{-1}\left(\eta (v^\theta)^\top \Sigma v^\theta - \eta^\theta (v^\theta)^\top \Gamma^\theta v^\theta\right) \\ &\geqslant \left((v^\theta)^\top \Sigma v^\theta\right)^{-1}\left((v^\theta)^\top\left(\eta\Gamma - \eta^\theta\Gamma^\theta\right) v^\theta\right) \geqslant 0. \end{aligned}$$

///

Proof of Theorem 2 This theorem can be derived from parts 2 and 3 of Lemma 4 for the case where $c = \mu_M - R\mathbf{1}$, $\Sigma = \Sigma_X$ and $\Gamma = \eta\Sigma_M$. ///

As in Lemma 4, let $a \in \boldsymbol{R}^N$, $c \in \boldsymbol{R}^N$, and $\Sigma \in \mathscr{S}_{++}^N$ and assume that there is no $\theta > 0$ such that $\theta \Sigma a = c$. Then, for each $\theta > 0$, there is a $(\Gamma, \eta) \in \mathscr{S}_+^N \times \boldsymbol{R}_{++}$ such that $\theta(\Sigma + \eta\Gamma)a = c$ and $\Sigma - \Gamma \in \mathscr{S}_+^N$ if and only if $\theta < \bar{\theta}$ (where $\bar{\theta} = (a^\top \Sigma a)^{-1}(a \cdot c)$), the largest eigenvalue of $\eta\Gamma$ is larger than or equal to $\|\theta^{-1}c - \Sigma a\|^2(\theta^{-1}(a \cdot c) - a^\top \Sigma a)^{-1}$, and η is larger than or equal to $\left((\theta^{-1}c - \Sigma a)^\top \Sigma^{-1} (\theta^{-1}c - \Sigma a)\right)(a \cdot (\theta^{-1}c - \Sigma a))^{-1}$. We are interested in minimizing these two values by varying θ over the interval $(0, \bar{\theta})$. The following lemma is general enough to deal with both of the two minimization problems.

Lemma 5 *Let $a \in \boldsymbol{R}^N$, $c \in \boldsymbol{R}^N$, $\Sigma \in \mathscr{S}^N_{++}$, and $S \in \mathscr{S}^N_{++}$. For each $\theta > 0$, define*

$$v^\theta = \frac{1}{\theta}\Sigma^{-1}c - a$$

and assume that there is no $\theta > 0$ for which $v^\theta = 0$. Assume also that $a \cdot c > 0$ and write $\bar{\theta} = \left(a^\top \Sigma a\right)^{-1}(a \cdot c)$. Then there is a unique $\theta > 0$ that minimizes the function

$$\begin{array}{ccc} \left(0, \bar{\theta}\right) & \to & \boldsymbol{R}_{++} \\ \theta & \mapsto & \dfrac{\left(v^\theta\right)^\top S v^\theta}{a^\top \Sigma v^\theta}, \end{array} \tag{35}$$

which is given by

$$\theta^{-1} = \bar{\theta}^{-1} + \left(\frac{\left(v^{\bar{\theta}}\right)^\top S v^{\bar{\theta}}}{c^\top \Sigma^{-1} S \Sigma^{-1} c}\right)^{1/2} \tag{36}$$

and the minimized value of the function is equal to

$$\frac{2}{a \cdot c}\left(\left(c^\top \Sigma^{-1} S \Sigma^{-1} c\right)^{1/2}\left(\left(v^{\bar{\theta}}\right)^\top S v^{\bar{\theta}}\right)^{1/2} + \left(v^{\bar{\theta}}\right)^\top S \Sigma^{-1} c\right). \tag{37}$$

To find the minimum, over $\theta \in (0, \bar{\theta})$, of the largest eigenvalue of $\eta^\theta \Gamma^\theta$ in Lemma 4, we take $S = \Sigma^2$. Then (36) can be rewritten as

$$\theta^{-1} = \bar{\theta}^{-1} + \frac{\left\|\bar{\theta}^{-1}c - \Sigma a\right\|}{\|c\|}, \tag{38}$$

and (37) can be rewritten as

$$\frac{2}{a \cdot c}\left(\|c\|\|\bar{\theta}^{-1}c - \Sigma a\| + \left(\bar{\theta}^{-1}c - \Sigma a\right) \cdot c\right). \tag{39}$$

To find the minimum of η^θ over $\theta \in (0, \bar{\theta})$, we take $S = \Sigma$. Then (36) can be rewritten as

$$\theta^{-1} = \bar{\theta}^{-1} + \left(\frac{\left(\bar{\theta}^{-1}c - \Sigma a\right)^\top \Sigma^{-1}\left(\bar{\theta}^{-1}c - \Sigma a\right)}{c^\top \Sigma^{-1} c}\right)^{1/2}, \tag{40}$$

and (37) can be rewritten as

$$\frac{2}{a \cdot c}\left(\left(\left(c^\top \Sigma^{-1} c\right)\left(\bar{\theta}^{-1} c - \Sigma a\right)^\top \Sigma^{-1}\left(\bar{\theta}^{-1} c - \Sigma a\right)\right)^{1/2} + \left(\bar{\theta}^{-1} c - \Sigma a\right)^\top \Sigma^{-1} c\right). \tag{41}$$

Proof of Lemma 5 The function (35) is continuous. Its value diverges to infinity as $\theta \uparrow \bar{\theta}$ (because, then, $a^\top \Sigma v^\theta \to 0$) and as $\theta \downarrow 0$. Thus there is a θ that minimizes the function. To show that (36) gives the unique solution, write $\beta = \theta^{-1}$ and differentiate

$$\ln \frac{\left(\beta \Sigma^{-1} c - a\right)^\top S \left(\beta \Sigma^{-1} c - a\right)}{a^\top \Sigma \left(\beta \Sigma^{-1} c - a\right)}$$

with respect to β to obtain the first-order necessary condition

$$\frac{2(\beta \Sigma^{-1} c - a)^\top S \Sigma^{-1} c}{\left(\beta \Sigma^{-1} c - a\right)^\top S \left(\beta \Sigma^{-1} c - a\right)} - \frac{a \cdot c}{\beta a \cdot c - a^\top \Sigma a} = 0,$$

which is equivalent to

$$(a \cdot c)\left(c^\top \Sigma^{-1} S \Sigma^{-1} c\right)\beta^2 - 2\left(a^\top \Sigma a\right)\left(c^\top \Sigma^{-1} S \Sigma^{-1} c\right)\beta + 2\left(a^\top \Sigma a\right)\left(a^\top S \Sigma^{-1} c\right) - \left(a^\top S a\right)(a \cdot c) = 0,$$

which is, in turn, equivalent to

$$\left(\beta - \bar{\theta}^{-1}\right)^2 - \frac{\left(v^{\bar{\theta}}\right)^\top S v^{\bar{\theta}}}{c^\top \Sigma^{-1} S \Sigma^{-1} c} = 0. \tag{42}$$

Since $v^{\bar{\theta}} \neq 0$, the second term of the left-hand side is strictly positive. Hence (42) has two distinct solutions, one larger and the other smaller than $\bar{\theta}^{-1}$. The former satisfies the first-order condition but the latter does not, because it is necessary that $\beta > \bar{\theta}^{-1}$. Therefore, there is only one solution to the problem of minimizing the function (35), which is given by (36).

As for the minimized value of the function, note that

$$v^\theta = v^{\bar{\theta}} + \left(\frac{\left(v^{\bar{\theta}}\right)^\top S v^{\bar{\theta}}}{c^\top \Sigma^{-1} S \Sigma^{-1} c}\right)^{1/2} \Sigma^{-1} c.$$

Thus

$$a^\top \Sigma v^\theta = (a \cdot c) \left(\frac{\left(v^{\bar{\theta}}\right)^\top S v^{\bar{\theta}}}{c^\top \Sigma^{-1} S \Sigma^{-1} c} \right)^{1/2},$$

$$\begin{aligned}\left(v^\theta\right)^\top S v^\theta &= \left(v^{\bar{\theta}}\right)^\top S v^{\bar{\theta}} + 2 \left(\frac{\left(v^{\bar{\theta}}\right)^\top S v^{\bar{\theta}}}{c^\top \Sigma^{-1} S \Sigma^{-1} c} \right)^{1/2} \left(v^{\bar{\theta}}\right)^\top S \Sigma^{-1} c + \frac{\left(v^{\bar{\theta}}\right)^\top S v^{\bar{\theta}}}{c^\top \Sigma^{-1} S \Sigma^{-1} c} c^\top \Sigma^{-1} S \Sigma^{-1} c \\ &= 2 \left(v^\theta\right)^\top S v^\theta + 2 \left(\frac{\left(v^\theta\right)^\top S v^\theta}{c^\top \Sigma^{-1} S \Sigma^{-1} c} \right)^{1/2} v^{\bar{\theta}} S \Sigma^{-1} c.\end{aligned}$$

Thus (37) is obtained. ///

We can now prove Theorem 3.

Proof of Theorem 3 This theorem can be derived Lemma 5 for the case where $c = \mu_M - R\mathbf{1}$, $\Sigma = \Sigma_X$, $\Gamma = \eta\Sigma_M$, and $S = \Sigma^2$ or $S = \Sigma$, as explained right after the lemma. ///

References

[1] Ahn, D., Choi, S., Gale, D. and Kariv, S. 2014, Estimating ambiguity aversion in a portfolio choice experiment. *Quantitative Economics* 5:195-223. doi: 10.3982/QE243

[2] Attanasi, G., Gollier, C., Montesano, A., and Pace, N. 2014. Eliciting ambiguity aversion in unknown and in compound lotteries: A smooth ambiguity model experimental study. *Theory and decision* 77(4): 485-530.

[3] Black, F., M. C. Jensen, and M. Scholes, 1972, The capital asset pricing model: Some empirical tests, in Jensen, M. (ed.), *Studies in the Theory of Capital Markets*, Praeger, New York

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

[5] Brennan, T.J., and A. W. Lo, 2010, Impossible Frontiers. *Management Science* 56(6):905–923.

[6] Brennan, T.J., and A. W. Lo, 2015, Reply to "(Im)possible Frontiers: A Comment", *Critical Financial Review* 4: 157–171.

[7] Campbell, J. Y., A. Lo, and C. MacKinlay. 1997. *The Econometrics of Financial Markets*. Princeton University Press.

[8] Collard, F., S. Mukerji, K. Sheppard, and J. Tallon. 2011, *Ambiguity and the historical equity premium* Available at SSRN: http://ssrn.com/abstract=1836297 or http://dx.doi.org/10.2139/ssrn.1836297

[9] DeMiguel, V., L. Garlappi, and R. Uppal. 2009. Optimal Versus Naive Diversification: How Inefficient is the $1/N$ Portfolio Strategy? *The Review of Financial Studies* 5:1915–1953.

[10] Epstein, L. G. 2010, A paradox for the "smooth ambiguity" model of preference, Econometrica 78: 2085–2099.

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

[12] Epstein, L. G., and S. E. Zin, 1989, Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework, *Econometrica* 57: 937–969.

[13] Fama, E. F., and K. R. French. 1992. The cross-section of expected stock returns. it Journal of finance 47(2):427-465.

[14] Epstein, L. G., and S. E. Zin, 1989, Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework, *Econometrica* 57, 937–969.

[15] Garlappi, L., R. Uppal, and T. Wang. 2007. Portfolio Selection with Parameter and Model Uncertainty: A Multi-Prior Approach. *Review of Financial Studies* 20:41–81.

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

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

[18] Hayashi, T., and J. Miao, 2011, Intertemporal substitution and recursive smooth ambiguity preferences, *Theoretical Economics* 6, 423–472.

[19] Ju, N., and J. Miao, 2012, Ambiguity, Learning, and Asset Returns. *Econometrica* 80, 559–591.

[20] Klibanoff, P., M. Marinacci, and S. Mukerji. 2005. A Smooth Model of Decision Making under Ambiguity. *Econometrica* 73:1849–1892.

[21] Klibanoff, P., M. Marinacci, and S. Mukerji. 2012. On the smooth ambiguity model: A reply. *Econometrica* 80:1303–1321.

[22] Kocherlakota, N. R., 1996, The equity premium: It's still a puzzle, Journal of Economic Literature 34, 42–71.

[23] Levy, M., and R. Roll, 2010, The Market Portfolio May Be Mean/Variance Efficient After All. *Review of Financial Studies* 23(6): 2464–2491.

[24] Levy, M., and R. Roll, 2015, (Im)Possible Frontiers: A Comment, *Critical Finance Review* 4, 139–148.

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

[26] Marinacci, M., 2002, Probabilistic sophistication and multiple priors, *Econometrica* 70, 755–765.

[27] Mehra R., and E. C. Prescott, 1985, The equity premium: A puzzle, Journal of Monetary Economics 15, 145–161.

[28] Pástor, L. 2000. Portfolio Selection and Asset Pricing Models. *Journal of Finance* 55:179–223.

[29] Pástor, L., and R. F. Stambaugh. 2000. Comparing Asset Pricing Models: An Investment Perspective. *Journal of Financial Economics* 56:335–381.

[30] Wang, Z. 2005. A Shrinkage Approach to Model Uncertainty and Asset Allocation. *Review of Financial Studies* 18:673–705.