Portfolio Concentration, Portfolio Inertia, and Ambiguous Correlation

Julia Jiang

University of North Carolina at Charlotte

Jun Liu

University of California at San Diego

Weidong Tian

University of North Carolina at Charlotte

Xudong Zeng

Shanghai University of Finance and Economics

Corresponding author: Weidong Tian, University of North Carolina at Charlotte. Email: wtian1@uncc.edu. We thank Nengjiu Ju (discussant), Philipp Karl Illeditsch (discussant), Valery Polkovnichenko (discussant), Alexandre M. Baptista, Pedro Barroso, Michael Brennan, Ethan Chiang, Chris Kirby, Jianjun Miao, Yufeng Han, Hong Yan, Liyan Yang, Tan Wang, Guofu Zhou, and seminar participants at the 6th Conference on Corporate Finance and Capital Markets of SUFE, Ben Graham Centre 6th Symposium on Intelligent Investing, European Financial Management Association 2016 Annual Meetings, 2016 Chinese International Conference in Finance, 2016 SFS Cavalcade Finance Conference, Shanghai Advanced Institute of Finance (SAIF), 2016 Financial Management Association Conference, and Southwestern University of Finance and Economics for helpful comments and suggestions. The authors' email addresses are jjiang6@uncc.edu, junliu@ucsd.edu, wtian1@uncc.edu, and zeng.xudong@mail.shufe.edu.cn. The authors would like to thank the editor, the associate editor and anonymous referees for their constructive comments and insightful suggestions.

Portfolio Concentration, Portfolio Inertia, and Ambiguous Correlation

ABSTRACT

When an investor is ambiguous about the asset returns' correlation and evaluates the portfolio in a multiple-priors framework, we show that the optimal portfolio is either independent of feasible correlation matrices or contains only a fraction of risky assets. In particular, the investor evaluates the risk-return tradeoff of each risky asset, and just one risky asset enters the optimal portfolio if the level of correlation ambiguity is high enough. Moreover, we demonstrate that the optimal portfolio does not change when each asset's Sharpe ratio changes in a range. Ambiguity-aversion on correlation uncertainty explains portfolio concentration and portfolio inertia in household portfolios and retirement accounts, and the model can explain the growth of indexing and ETFs from an optimal portfolio choice perspective. We further show that these properties are not valid in alternative smooth ambiguity models, suggesting that the smooth ambiguity model does depart from the standard model enough to explain portfolio concentration and portfolio inertia.

JEL classification: G11, G12, G13, and D 50

Keywords: Correlation ambiguity, anti-diversification, correlation-invariant, portfolio concentration, portfolio inertia, smooth ambiguity

1 Introduction

Correlation plays a crucial role in portfolio choice since the seminal work in Markowitz (1952). However, as the number of assets increases, Chan, Karceski, and Lakonishok (1999), Jagannathan and Ma (2003), and Kan and Zhou (2007) document that the covariance-variance matrices are often imprecisely estimated and lead to significant estimation errors in constructing the optimal portfolio. In this paper, we study the effect of correlation uncertainty on portfolio choice for any number of risky assets.

In this study, as in Epstein and Halevy (2019),[1] an investor knows perfectly about the marginal distribution of each risky asset; however, the investor's ambiguity on the joint distribution is represented by a set (an ambiguous set) of asset return correlation matrices. The investor evaluates the portfolio based on each feasible correlation matrix and chooses the worst-case one in the sense of value function.[2] We first consider a situation in which the investor knows barely about correlation, and we show that the ambiguity-aversion investor will optimally choose a portfolio with only one risky asset, yielding an *anti-diversified* portfolio.[3]

The intuition of anti-diversification is as follows. Given a known marginal distribution of each risky asset, the investor knows how to construct an optimal portfolio with this risky asset only. Since the investor is ambiguous averse to correlation, the investor prefers a portfolio that is insensitive to correlation. The only portfolios insensitive to all correlation matrices consist of one asset. Therefore, if the ambiguous set is sufficiently large, the optimal portfolio for multiple assets is reduced to consider the portfolio construction for each asset separably. This portfolio construction method without using a correlation matrix is similar

[1] Epstein and Halevy (2019) identifies investors' lack of confidence of joint distribution among multiple random events experimentally. See also Mihm (2016) for ambiguity on the dependent structure.

[2] It is the classical multiple-priors model in Gilboa and Schmeidler (1989) to study the Knightian uncertainty or ambiguity. Since Ellsberg (1961), the investor's uncertainty is consistent with experimental evidences by Dimmock, Kouwenberg, Mitchell, and Peijnenburg (2016), Bianchiand and Tallon (2019) and Epstein and Halevy (2019).

[3] Goldman (1979) coins this term for holding one risky asset in the optimal portfolio. An anti-diversified portfolio is *unique* in that it does not depend on any correlation matrix.

to the separably screening procedure suggested in Carroll (2017) for a principle to screen an agent along several dimensions of private information and the marginal distribution for each component of the agent's type is known. The anti-diversification proposition in the portfolio choice setting shares a similar economic insight of the optimal screening theorem in Carroll (2017).

By the same intuition, the ambiguity-aversion investor prefers a portfolio that is *independent* of the feasible correlation matrices in a generally given ambiguous set. If there exists a portfolio that is *correlation-invariant*, we show that the investor chooses such a portfolio. If there exists no correlation-invariant portfolio for a given ambiguous set, we demonstrate that the optimal portfolio consisting of a few risky assets with significantly different marginal distributions, so the optimal portfolio under correlation uncertainty is highly *concentrated*. The mechanism to derive these features of an optimal portfolio is as follows. Since the benefit of diversification is the essential factor in constructing a portfolio, the ambiguity-aversion investor evaluates the trade-off between diversification benefit and the cost of reducing the portfolio's Sharpe ratio due to the ambiguity-aversion on the correlation uncertainty. The correlation ambiguity affects the Sharpe ratio of the portfolio in general, while each asset Sharpe ratio is not affected. Therefore, the correlation matrix with a small (or even zero) effect on the utility appeals to the investor. In the end, the investor chooses a correlation-invariant portfolio or a concentrated portfolio.

Moreover, we demonstrate a precise *risky asset inertia* property of the optimal portfolio under correlation ambiguity. Portfolio inertia is the concept when the Sharpe ratio of risky assets change, both the list of risky assets and their holdings in the optimal portfolio do not change accordingly. We show that when investors are ambiguity-averse to correlation uncertainty, they stop rebalancing the portfolio. For instance, when the level of correlation ambiguity is high, only the risky asset with the highest Sharpe ratio enters the optimal portfolio. Therefore, this anti-diversified portfolio does not change regardless of changes of all *other* risky assets as long as the other Sharpe ratio is smaller than the highest Sharpe

ratio. The intuition of portfolio inertia for a general situation is also straightforward from the diversification perspective. The investor considers the diversification benefits in rebalancing the portfolio. Even though each risky asset becomes more attractive, if the diversification benefit is not significant enough to dominate the correlation ambiguity concern, the investor would prefer to stay with the original portfolio. This inertia property for risky assets is shown to be robust for any number of risky assets and many ambiguous sets. Both the portfolio concentration and portfolio inertia are consistent with empirical studies in household portfolios and retirement accounts (See, for instance, Agnew, Balduzzi, and Sunden (2003), Bilias, Geogarakos, and Haliassos (2007), Campbell (2006), and Ivković, Sialm, and Weisbenner (2008).).

For a comparison purpose, we further study a smooth ambiguity model of Klibanoff, Marinacci, and Mukerji (2005) in which the investor still has perfect knowledge about each marginal distribution but a second-order distribution on a set of feasible correlation matrices. In contrast with the multiple-priors model, we show that the optimal portfolio is *not* correlation-invariant, i,e., it must depend on the second-order distribution on the ambiguous set. In particular, it is impossible to treat each asset separately to obtain the optimal anti-diversified portfolio regardless of the ambiguous set. Moreover, we show that the investor must rebalance his portfolio when each asset's investment opportunity changes slightly. Therefore, there is no risky asset inertia property when the investor has a smooth preference for the correlation ambiguity on correlation.[4] Our results suggest that the smooth ambiguity model is not significantly different from the standard model enough to explain portfolio inertia and under-diversification.

[4] Maccheroni, Marinacci, and Ruffino (2013) solve an optimal portfolio problem in an analog of the smooth ambiguity model. In a financial market with a risky asset and an ambiguous asset, Maccheroni, Marinacci, and Ruffino (2013) show that the portfolio rebalancing in response to the ambiguity aversion depends on the ambiguous asset's alpha. These authors do not consider the general property of the optimal portfolio in this robust mean-variance framework.

1.1 Related Literature Review

In addition to the papers already mentioned, many economic models of ambiguity are studied in the literature. These models are different in identifying attitudes toward risk and ambiguity. In these previous studies, only a small number of the risky asset is considered. For instance, Dow and Werlang (1992) show that the holding of a risky position is zero in a range of asset price by a model of Choquet expected utility. Easley and O'Hara (2010) introduce Bewley's incomplete preference to demonstrate a region of a price (for one risky asset) over there is no trading. Epstein and Schneider (2008) and Illeditsch (2011) consider an economy with only one risky asset, but the investor has ambiguity aversion to the information quality of the future cash flow of this risky asset. Illeditsch (2011) shows the price inertia that the optimal demand on the risky asset does not change in a range of risky asset prices. Epstein and Schneider (2010) shows the inertia property for risk-free asset and a certain risky portfolio when an investor is ambiguous about both the expected return and volatility of one risky asset. Bossaerts (2010) demonstrate the no participation and portfolio inertia under $\alpha - maximin$ model of Ghirardato, Maccheroni and Marinacci (2004).[5] In general, when the risky asset's payoff is ambiguous, the ambiguity averse investor holds zero position on this risky asset (Murkerji and Tallon (2003)). Nevertheless, this paper shows that the trade-off between the benefit of diversification and ambiguity-aversion to the correlation uncertainty is another channel to lead inertia for any number of risky assets. Moreover, the portfolio inertia property in this paper is different from those studies on the marginal distribution uncertainty in which the correlated structure is known.

Many recent studies document the different implications among different approaches to ambiguity aversion. Gollier (2011) and Caskey (2009) document the second-order effect in risk aversion and implications to portfolio choice and asset prices. By contrast, the multiple

[5] Cao, Han, Hirshleifer, and Zhang (2011) introduce status quo deviation aversion and demonstrate no-trade under familiarity bias for two risky assets. Other models to explain the inertia or "no-trade" include information costs (Abel, Eberly, and Panageas (2013), Van Nieuwerburgh and Veldkamp (2010)), and transaction costs in Liu (2004).

priors model of the expected return and volatility (marginal distribution) ambiguity leads to the first-order effect in risk aversion. See Illeditsch (2011), Condie, Ganhuli, and Illeditsch (2021) and Epstein and Schneider (2010). In this paper, we also demonstrate different implications of the multiple priors and smooth ambiguity model. This paper contributes to the literature by showing that the correlation ambiguity under the multiple priors preferences leads to the first-order effect, similar to the marginal distribution ambiguity. In contrast, the smooth ambiguity preference leads to a second-order effect (see also Seo (2009) for a theoretical discussion of the smooth ambiguity model). Epstein (2010) argued that the smooth ambiguity model might not be appropriate to separate the ambiguity and ambiguity aversion. Our results show that the smooth ambiguity model is not strong enough to distinguish it from the standard expected utility model. On the other hand, Gajods, Hayanshi, Tallon and Vergnaud (2008) suggest that the feasible set in the multiple priors model is different from the logical possible set, so the worst-case scenario under the multiple priors belief is often the "corner" solution, whereas the smooth ambiguity model implies an "interior" solution.

These concentration or correlation-invariant features of the optimal portfolio paired with correlation ambiguity is different from under-diversification in previous ambiguity aversion literature on the expected return or volatility.[6] Previously, under-diversification refers to a bias in individual assets or non-participation in risky assets. For example, Cao, Wang, and Zhang (2005), Easley and O'Hara (2009), and Garlappi, Uppal, and Wang (2007) demonstrate that some (or all) risky assets do not enter the optimal portfolio if these assets are not attractive and investor has ambiguous about them. Using the robust control model of Anderson, Hansen and Sargent (2003) and Strzalecki (2011), Uppal and Wang (2003) also explains expected return ambiguity can cause under-diversification in the sense that the optimal positions are biased relative to the standard mean-variance portfolio. In contrast, we assume that each asset is attractive in our model. Because the investor's ambiguity-

[6]Among the ambiguity literature, Boyle, Garlappi, Uppal and Wang (2012), Cao, Wang, and Zhang (2005), Easley and O'Hara (2009), and Garlappi, Uppal, and Wang (2007) investigate expected return parameter uncertainty. Easley and O'Hara (2009), and Epstein and Ji (2013) discuss volatility parameter uncertainty. In these studies, asset returns are independent, so the correlation is not a concern.

aversion reduces the benefit of diversification (the portfolio's Sharpe ratio), the portfolio construction does not take full advantage of asset correlation. Therefore, the optimal portfolio under correlation uncertainty is under-diversified in the form of portfolio concentration or correlation-invariant.

There are other explanations for under-diversification in literature from different mechanisms. For example, Roche, Tompaidis, and Yang (2013) suggest that financial constraints can lead to under-diversification. Van Nieuwerburgh and Veldkamp (2010) propose an explanation based on information costs. Polkovnichenko (2005) demonstrates household portfolio under-diversification by using a rank-dependent preference. Mitton and Vorkink (2007) explain under-diversification because of investors' heterogeneous beliefs and preferences to skewness. Ang, Bekaert, and Liu (2005) demonstrate portfolio under-diversification when investor has disappointment aversion. By extending the Choquet expected utility model for a large number of assets, Murkerji and Tallon (2001) show that the traditional diversification role fails due to the ambiguity-aversion; thus, the effect of ambiguity-aversion on the financial market is to make the risk-sharing opportunities is less complete than it would be. Guidolin and Liu (2016) show the under-diversification in a smooth ambiguity model. In this group of under-diversification literature, correlation attracts little direct attention. Our study focuses on the effect of ambiguity-aversion on the benefit of diversification.[7]

In a new approach to multiple sources of information to form predictions, Levy and Razin (2020) (see also Levy and Razin (2015)) introduce two factors, correlation ignorance and a bound on pointwise mutual information. Correlation ignorance is a naive interpretation of forecasts to ignore correlation by simply assuming independent marginal distributions. As demonstrated in this paper, the ambiguity-aversion investor does not like the correlation-ignorance portfolio. For the same reason, the optimal portfolio for a naive Bayer investor with correlation ignorance is neither concentrated nor correlation-invariant. Our characterization of the optimal portfolio under correlation uncertainty is significantly different from the one

[7]We also derive a general portfolio choice rule under joint distribution uncertainty in this paper for this comparison exercise.

characterized in Levy and Razin (2020).

Finally, Huang, Zhang, and Zhu (2017) derive the same portfolio choice role for two risky assets in a mean-variance framework. By deriving essentially the same optimal portfolio for two risky opportunities, Easley, O'Hara, and Yang (2015) study the asset price implications for hedge fund regulations. Condie, Ganhuli, and Illeditsch (2021) consider the correlation uncertainty between one firm's future cash flow and its signal (which is not tradable) and obtain short-term momentum due to this ambiguity-aversion. Here, we consider any number of risky assets and study the features of the optimal portfolios. Moreover, we demonstrate that these features disappear in smooth ambiguity models.

The paper is organized as follows. In Section 2, we present a portfolio choice problem with an aversion to correlation ambiguity and solve the problem explicitly. In Section 3, we demonstrate several unique features of the optimal portfolio under correlation uncertainty. Section 4 demonstrates the portfolio inertia property of the optimal portfolio for any number of risky assets. In Section 5, we first explain the empirical evidence and implications to our results and then compare the effect of the correlation ambiguity with the Sharpe ratio (marginal distribution) ambiguity. In Section 6, we study alternative smooth ambiguity model or Bayesian model uncertainty and demonstrate different features of the optimal portfolio. Our conclusions are presented in Section 7. Proofs and other general technical results are given in Appendix.

2 Optimal Mean-Variance Portfolio under Ambiguous Correlation

This section first explains how a set of feasible correlation matrices captures investor's correlation ambiguity when the marginal distribution is known and presents an optimal portfolio choice problem under correlation ambiguity in a mean-variance setting. Then, we explicitly characterize the optimal strategies for an arbitrary set of feasible correlation matrices and

any number of risky assets.

In a universe of a riskless asset and N risky assets, r_f denotes the rate of return on the riskless asset and the rate of return of N risky assets are $r_1, ..., r_N$, respectively. Let $\mu = (\mu_1, ..., \mu_N)^\top$ denote the expected excess return vector of risky assets, where the convention $^\top$ denotes the transpose. Let σ_n denote the standard deviation of excess return n, $n = 1, ..., N$ and let σ be the diagonal matrix with diagonal entries $\sigma_1, ..., \sigma_N$ in order. Let $\rho = (\rho_{ij})$ be the correlation matrix of the excess returns, where $\rho_{ij} = 1$ if $i = j$, and ρ_{ij} is the coefficient of correlation between risky asset i and risky asset j for $i \neq j$. Then the covariance-variance matrix Σ of excess return is $\sigma\rho\sigma$. Define $s = (s_1, ..., s_N)^\top$, where $s_n = \mu_n/\sigma_n$ is the Sharpe ratio of risky asset n. Without loss of generality, we assume that Σ is non-singular.

In the standard mean-variance model of portfolio choice, both the expected excess return vector μ and the covariance-variance matrix Σ are known in precise. Let ϕ_n, $n = 1, ..., N$, denote the dollar amount that is invested in risky asset n and $\phi = (\phi_1, ..., \phi_N)^\top$. The unique optimal strategy to the standard mean-variance portfolio choice problem

$$\max_{\phi} \mu^\top \phi - \frac{A}{2}\phi^\top \Sigma \phi, \tag{1}$$

is given by

$$\phi_{MV}(\rho) = \frac{1}{A}\Sigma^{-1}\mu = \frac{1}{A}\sigma^{-1}\rho^{-1}s, \tag{2}$$

where A is the risk aversion coefficient. The optimal value is obtained by using the strategy $\phi_{MV}(\rho)$ and given by

$$V = \frac{1}{2A}s^\top \rho^{-1} s. \tag{3}$$

Similar to Carroll (2017), the investor in our model has a perfect knowledge about the marginal distribution of each r_n, but does not know the joint distribution of $(r_1, \cdots, r_N)$.

Since the investor constructs portfolio in a mean-variance setting, it is equivalent to assume that the investor knows both the expected excess return and standard deviation, but not know the correlation matrix ρ.[8] We assume that the set of feasible correlation matrices is denoted by an ambiguous set $\mathcal{C}$. Here, for its general purpose, the ambiguous set $\mathcal{C}$ of correlation matrix is an arbtrary closed convex subset of $\mathcal{B}$, and $\mathcal{B}$ is a set of positive definite symmetric matrices $\rho = (\rho_{ij})$ such that $\rho_{ii} = 1, \forall i, \rho_{ij} \in [\underline{\rho}_{ij}, \overline{\rho}_{ij}], \forall i \neq j$. [9] The size of the ambiguous set reflects the investor's lack of confidence when thinking about returns' correlation structure.

For a pair (i, j) of asset returns we use confidence interval as ambiguity set of coefficients of correlation by a standard method in statistics. Let $R_p = \frac{\sum_{i=1}^{n}(X_i-\bar{X})(Y_i-\bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i-\bar{X})^2(Y_i-\bar{Y})^2}}$ for a paired sample $(X_1, Y_1), ..., (X_n, Y_n)$, with sample mean $(\bar{X}, \bar{Y})$. The Fisher transformation $F(R_p) = \frac{1}{2}\ln(\frac{1+R_p}{1-R_p})$ is approximately normally distributed with mean $\frac{1}{2}\ln(\frac{1+p}{1-p})$ and variance $\frac{1}{n-3}$, where p is the population correlation. The confidence bounds are based on the asymptotic normal distribution. If variables have a multivariate normal distribution, these bounds are accurate for large samples. For this reason, we let $[\underline{\rho}_{ij}, \overline{\rho}_{ij}]$ denote the confidence level of the coefficient of correlation between asset i and asset j. We assume that asset i and asset j are not perfectly correlated, that is, $-1 < \underline{\rho}_{ij} < \overline{\rho}_{ij} < 1$, but we do not impose any restriction on the value of $\overline{\rho}_{ij}$ and $\underline{\rho}_{ij}$ for generality. Since the covariance-variance is positive definite and symmetric, the correlation matrix must be an element of $\mathcal{B}$.

Following Gilboa and Schmeidler (1989), the optimal portfolio choice problem for an ambiguity-averse investor is,

$$J = \max_{\phi} \min_{\rho \in \mathcal{C}} \mu^{\top}\phi - \frac{A}{2}\phi^{\top}\Sigma\phi, \tag{4}$$

where the minimization reflects the agent's aversion to correlation ambiguity. The investor

[8] A general solution of the optimal portfolio choice problem under joint distribution ambiguity is discussed in Section 5 and Appendix (Proposition G.1).

[9] For a matrix $B = (b_{ij})$, $||B|| = \sum_{i=1}^{N}\sum_{j=1}^{N}|b_{ij}|$ yields a norm topology. By a closed set we mean it is closed under this norm topology. A set of matrix is bounded if there exists a positive number $M > 0$ such that $||B|| \leq M$ for all matrix B in this set.

chooses the optimal portfolio that maximizes the worst value functions over all feasible correlation matrices in $\mathcal{C}$. The next result characterizes the unique optimal portfolio under any ambiguous set of correlation matrices.

PROPOSITION 1. *There exists a solution ρ^* of the following optimization problem*

$$\min_{\rho \in \mathcal{C}} \left(s^\top \rho^{-1} s \right). \tag{5}$$

Then $(\phi_{MV}(\rho^*), \rho^*)$ satisfies the saddle-point property

$$f(\phi, \rho^*) \leq f(\phi_{MV}(\rho^*), \rho^*) \leq f(\phi_{MV}(\rho^*), \rho), \forall \rho \in \mathcal{C}, \phi \in \mathbb{R}^N. \tag{6}$$

where $f(\phi, \rho) = \mu^\top \phi - \frac{A}{2} \phi^\top \Sigma \phi$. Moreover, the max-min problem (4) has a unique optimal solution $\phi_{MV}(\rho^*)$.

By Proposition 1, we can change the order of maximization and minimization and obtain

$$J = \max_{\phi} \min_{\rho} \mu^\top \phi - \frac{A}{2} \phi^\top \Sigma \phi = \min_{\rho} \max_{\phi} \mu^\top \phi - \frac{A}{2} \phi^\top \Sigma \phi = \min_{\rho} \frac{1}{2A} s^\top \rho^{-1} s. \tag{7}$$

The existence of solution of the inner maximization problem, $\rho^* = \text{argmin}_{\rho \in \mathcal{C}} \left(s^\top \rho^{-1} s \right)$, is standard due to compactness of $\mathcal{C}$. Proposition 1 states that the optimal portfolio is unique and explicitly given by $\phi_{MV}(\rho^*)$. Then we refer a solution ρ^* as *the worst-case correlation* (which might be not unique), and $\phi_{MV}(\rho^*)$ as the unique optimal portfolio of the max-min problem (4).

To describe correlation ambiguity among asset returns, we provide three examples of the ambiguous set $\mathcal{C}$ for multiple risky assets.

Example 2.1. *The ambiguous set is* $\mathcal{C}[\underline{a}, \overline{a}] = \{ \rho = R(a) \equiv (\rho_{ij}) \in \mathcal{B} : \rho_{ij} = a \in [\underline{a}, \overline{a}], \forall i \neq j, \rho_{ii} = 1, \forall i \ \}$.

The class of equicorrelation matrix is studied in Engel and Kelly (2012). The confidence level of a common pairwise correlation coefficient in this example is $[\underline{a}, \overline{a}]$.

Example 2.2. *Given a positive number $\epsilon > 0$, the ambiguous set is*[10] $\mathcal{C}(\epsilon) = \{\rho \in \mathcal{B} : \lambda_{min}(\rho) \geq \epsilon\}$. *Here $\lambda_{min}(\rho)$ denotes the minimal eigenvalue of ρ.*

If the positive number is sufficiently small, the ambiguous set in this example denotes an investor barely knowing anything about the correlation matrix among asset returns.

Example 2.3. *Fix two correlation matrices ρ_1 and ρ_2 of size $n \times n$ and $k \times k$, respectively The ambiguous set is*

$$\mathcal{C}(\rho_1, \rho_2) = \left\{ \begin{pmatrix} \rho_1 & \zeta \\ \zeta^\top & \rho_2 \end{pmatrix} \in \mathcal{B} : \zeta \in M_{n\times k}(\mathbb{R}) \right\}.$$

Given this ambiguous set, the investor has a perfect knowledge about the joint distribution of $\{r_1, \cdots, r_n\}$ and $\{r_{n+1}, \cdots, r_{n+k}\}$, but the correlated structure between assets in the first class $\{1, \cdots, n\}$ with the asset in the second class $\{n+1, \cdots, n+k\}$ is unknown. Uppal and Wang (2003) and Garlappi, Uppal, and Wang (2007) consider a similar class of ambiguous set. A general product structure of the ambiguous set is considered in Section 3.3.

It is helpful to compare the correlation ambiguity with the marginal distribution ambiguity at this point. Experimentally, Epstein and Halevy (2019) identify investor's lack of confidence of joint distribution among multiple random events. For a mean-variance investor, why do we need to address the correlation ambiguity among asset returns? For an individual asset, it is well known that estimating the expected excess return is a challenge, whereas the standard error of variance estimator decreases with the frequency of data observations (Merton (1980)). Therefore, we need to compare which is simpler to estimate the covariance-variance matrix or the expected return from an econometrician perspective.

[10] It is standard (Serre (2002)) to show that the set $\mathcal{C}(\epsilon)$ is convex and closed as follows. First, since all eigenvalues are continuous functions of all entries of a matrix, $\mathcal{C}(\epsilon)$ is closed. Second, we make use of the classical Weyl's inequality in matrix theory: Given three symmetric matrices A, H, P and $A = H + P$. Assume that the eigenvalues of A, H, P are $\mu_1 \geq \cdots \geq \mu_N, \nu_1 \geq \cdots \geq \nu_N$ and $\rho_1 \geq \cdots \geq \rho_N$, respectively, then $\mu_i \geq \nu_i + \rho_N$. Then for any $H, P \in \mathcal{C}(\epsilon)$, all eigenvalues of $\frac{H+P}{2}$ are greater than ϵ. Hence, $\frac{H+P}{2} \in \mathcal{C}(\epsilon)$ and $\mathcal{C}(\epsilon)$ is convex. Furthermore, a symmetric matrix is positive definite if and only if all eigenvalues are positive numbers.

For this problem, Kan and Zhou (2007) analytically show that the estimation error in the covariance-variance matrix is more costly than estimation errors in the expected return. Chan, Karceski, and Lakonishok (1999) also demonstrate the difficulty of forecasting future covariance-variance matrix in portfolio choice with a large number of risky assets. Moreover, the joint distribution imposes far more risk management issues than the marginal distribution (Levy and Razin (2020)). Hence, it is vital to study the portfolio implications of the correlation ambiguity among many assets. We will present a comparison in detail between the correlation ambiguity and the marginal distribution ambiguity in Section 5.3.

3 Features of the Optimal Portfolios

This section first demonstrates that when the ambiguous set contains a particular correlation matrix, the optimal portfolio includes only one risky asset with the highest Sharpe ratio (anti-diversified). This optimal portfolio rule is robust in the sense that the ambiguity-averse investor always likes the unambiguous portfolio, which is independent of the correlation matrix in the ambiguous set (correlation-invariant). Moreover, the optimal portfolio under correlation ambiguity is likely concentrated, containing only a few risky assets for a general ambiguous set.

3.1 Anti-diversified portfolio

Since the investor is ambiguity-averse to correlation uncertainty, we first study under what condition the optimal portfolio does not depend on any knowledge about the correlation matrix. That is, it includes only one risky asset. The following result provides a sufficient and necessary condition for such an optimal portfolio.

PROPOSITION 2. *Suppose* $|s_1| > \max\{|s_2|, ..., |s_N|\}$. *Then the optimal portfolio contains only the first risky asset if and only if there exists one correlation matrix* $\rho \in \mathcal{C}$ *such that* $\rho_{1i} = \frac{s_i}{s_1}$ *for all* $i = 2, \cdots, N$. *Moreover, if the optimal portfolio under correlation ambiguity*

contains only risky asset 1, then $s_1^2 > s_i^2, i = 2, \cdots, N$.

The intuition in Proposition 2 is straightforward. For each risky asset, the investor can construct an optimal portfolio with a risk-free asset and this risky asset by using its known marginal distribution. Since the investor is ambiguous about the joint distribution, the investor intends to choose the portfolio without using any correlation matrix (i.e., portfolio with only one risky asset). If the investor has very limited knowledge about the joint distribution, or equivalently a large ambiguous set, it turns out that the optimal portfolio for multiple assets under correlation ambiguity is one of these portfolios with a single risky asset constructed by the marginal distribution only.

Proposition 2 is similar to the optimal screening problem studied in Carroll (2017) when a principle screens an agent along several dimensions of private information. If the principle knows the marginal distribution of each component of the agent's type but does not know the joint distribution, and any mechanism is evaluated by its worst-case expected profit over all joint distributions consistent with the known marginals, Carroll (2017) shows that the optimal for the principle is to screen along with each component separately using the known marginal distribution. Proposition 2 demonstrates a similar insight in the optimal portfolio choice setting.

In portfolio choice literature, a portfolio is anti-diversification in the sense of Goldman (1979) if only one risky asset enters the optimal portfolio. Goldman (1979) shows that the buy-and-hold strategy will result in anti-diversification in an infinite time horizon. In his paper, only the asset with the highest risk aversion adjusted expected return will be held. For the mean-variance investor, it is natural to use Sharpe ratio to represent each asset's risk-return characteristics, so the anti-diversified portfolio includes only the asset with the highest Sharpe ratio.

The following examples illustrate anti-diversified portfolio under explicit condition of the ambiguous set.

Example 3.1. *Suppose* $N = 2, |s_1| > |s_2|$, *and the ambiguous set of the correlation coefficient*

ρ_{12} is $[\underline{\rho}, \overline{\rho}]$. Then, the optimal portfolio contains only the first risky asset if and only if $\underline{\rho} \leq \frac{s_2}{s_1} \leq \overline{\rho}$.

Huang, Zhang, and Zhu (2017) solve the same optimal portfolio choice problem with two risky assets, and derive the same result as in Example 3.1. In a different context, Easley, O'Hara, and Yang (2015) also derive the optimal portfolio under correlation ambiguity with two risky assets. The next two examples illustrate anti-diversified optimal portfolio for many ambiguous sets with $N \geq 3$.

Example 3.2. *Suppose $N = 3, |s_1| > \max(|s_2|, |s_3|)$, and the ambiguous set of the correlation matrix is $\mathcal{C}$. Then, the optimal portfolio contains only the first risky asset if and only if*

$$\begin{pmatrix} 1 & \frac{s_2}{s_1} & \frac{s_3}{s_1} \\ \frac{s_2}{s_1} & 1 & \zeta \\ \frac{s_3}{s_1} & \zeta & 1 \end{pmatrix} \in \mathcal{C}$$

for one $\zeta \in (-1, 1)$.

Example 3.3. *Assuming $|s_1| > \max\{|s_i|, i = 2, \cdots, N\}$, and $\underline{\rho}_{ij} \leq \frac{s_i s_j}{s_1^2} \leq \overline{\rho}_{ij}, \forall i \neq j$, then the optimal portfolio for the ambiguous set $\mathcal{C}(\epsilon)$ contains only the first risky asset for sufficiently small positive number ϵ.*

We should point out that the known marginal distribution assumption is crucial in Proposition 2 and in Carroll (2017). Given N risky assets, $X_1, \cdots, X_N$, we consider N portfolios, $Y_1, \cdots, Y_N$ such that each asset X_i can be also spanned by these N portfolios. For instance, $Y_i = \sum_{j=1}^N a_{i,j} X_j$, and the matrix (a_{ij}) is a non-singular matrix with coefficients $a_{ij} \in \mathbb{R}$. Intuitively, if the investor is sufficiently ambiguous among these assets X_i, the investor should also be sufficiently ambiguous among the N portfolios $Y_1, \cdots, Y_N$. Hence, the investor would choose one of the N portfolios $\{Y_1, \cdots, Y_N\}$, which seems contradicts to the anti-diversified portfolio in Proposition 2. We can solve this "inconsistency" issue by noticing *unknown*

marginal distribution of Y_i under correlation ambiguity assumption, and the optimal portfolio based on assets Y_i can be found by a general result in Appendix (Proposition G.1). Indeed, Proposition 2 cannot be applied for these N portfolios Y_i directly. By using Proposition G.1 for $\{Y_1, \cdots, Y_N\}$ in a joint distribution uncertainty setting, the optimal portfolio under the assumption in Proposition 2 is one of $X_1, \cdots, X_N$, being one specific portfolio of these assets $Y_1, \cdots, Y_N$.

Remark 3.1. *In practice, it is possible that the investor is offered by some basic portfolios for which the marginal distributions are statistically known, for instance, ETF portfolios or emerging market portfolios. But the correlation ambiguity is still a concern for these portfolios. Proposition 2 can be applied is this situation as well. When the set of correlation matrix between these portfolios is given in Proposition 2 or Example 3.3, the investor holds only the portfolio with the highest Sharpe ratio even though each basic portfolio is attractive. As an illustrative example, the ambiguous-averse investor might not enter the emerging market if the correlated structure between te U.S market and emerging market is too complicated to be analyzed.*

3.2 Correlation-invariant portfolio

In this subsection, we extend Proposition 2 for a general ambiguous set.

Definition 3.1. *An investment portfolio ϕ is* correlation-invariant *with respect to $\mathcal{C}$ if $f(\phi, \rho)$ is the same for any $\rho \in \mathcal{C}$.*

An anti-diversified portfolio is clearly correlation-invariant with respect to any ambiguous set $\mathcal{C}$ since no correlation matrix is used in constructing the optimal portfolio. Like an anti-diversified portfolio, since its mean-variance utility $f(\phi, \rho)$ of a correlation-invariant portfolio is independent of the correlation matrix, there is *no ambiguity* in the correlation-invariant portfolio when the investor evaluates each possible correlation matrix. Naturally, the ambiguity-averse investor chooses this portfolio if possible.

PROPOSITION 3. *The optimal portfolio under correlation ambiguity is correlation-variant if and only if there exists $\rho_1 \in \mathcal{C}$ such that the vector $\phi_{MV}(\rho_1)$ is correlation-invariant with respect to $\mathcal{C}$.*

According to Proposition 1, the optimal portfolio under correlation ambiguity is $\phi_{MV}(\rho)$ for certain $\rho \in \mathcal{C}$. On one hand, if $\phi_{MV}(\rho)$ is correlation-invariant for one $\rho \in \mathcal{C}$, then the ambiguous-averse investor chooses $\phi_{MV}(\rho)$ optimally. On the other hand, if the portfolio $\phi_{MV}(\rho)$ is *not* correlation-invariant for each feasible correlation matrix $\rho \in \mathcal{C}$, then the optimal portfolio is not correlation-invariant anymore. The existence of a correlation-invariant optimal portfolio relies on the nature of the ambiguous set, as shown by the following two examples.

Example 3.4. *Let the ambiguous set $\mathcal{C} = \mathcal{C}[\underline{a}, \overline{a}]$, a common pairwise correlation coefficient $a \in [\underline{a}, \overline{a}]$. A portfolio vector ϕ is correlation-invariant with respect to $\mathcal{C}[\underline{a}, \overline{a}]$ if and only if $\sum_{i \neq j}(\sigma_i \phi_i)(\sigma_j \phi_j) = 0$. It can be shown that (see the proof of Example 4.3 in Appendix) if $\frac{1-\Omega(s)}{1+(N-1)\Omega(s)} \in [\underline{a}, \overline{a}]$, where*

$$\Omega(s) = \sqrt{\frac{1}{N-1}\left(N\frac{\sum_{n=1}^{N} s_n^2}{(\sum_{n=1}^{N} s_n)^2} - 1\right)},$$

then the optimal portfolio under $\mathcal{C}[\underline{a}, \overline{a}]$ is correlation-invariant. Otherwise, the optimal portfolio under $\mathcal{C}[\underline{a}, \overline{a}]$ is either $\phi_{MV}(\underline{a})$ or $\phi_{MV}(\overline{a})$.

There are two remarkable implications in this example. First, even though a correlation-invariant portfolio is always appealing, an optimal portfolio under ambiguous set could be not correlation-invariant for a given ambiguous set. For example, if $\frac{1-\Omega(s)}{1+(N-1)\Omega(s)} < \underline{a}$ or $\frac{1-\Omega(s)}{1+(N-1)\Omega(s)} > \overline{a}$, the worst-case correlation is associated with the smallest or largest possible pairwise correlation coefficient number, respectively, leading a "corner solution". Second, when an optimal portfolio is indeed correlation-invariant, the optimal portfolio is not necessarily anti-diversified. It is possible that all risky assets enter the correlation-invariant

optimal portfolio.

In Example 3.4, asset returns have a constant pairwise ambiguous correlation coefficient. The next example demonstrates similar features of the correlation-invariant portfolio when some assets are independent.

Example 3.5. *Let*

$$\mathcal{C} = \left\{ T(a) \equiv \begin{pmatrix} 1 & 0 & a \\ 0 & 1 & a \\ a & a & 1 \end{pmatrix} : a \in [\underline{a}, \overline{a}] \right\},$$

and $\overline{a}^2 < \frac{1}{2}$. *A portfolio* ϕ *is correlation-invariant with respect to* $\mathcal{C}$ *if and only if* $(\phi_1+\phi_2)\phi_3 = 0$. *If* $\frac{s_3}{s_1+s_2} \in [\underline{a}, \overline{a}]$, *then* $\phi_{MV}(\frac{s_3}{s_1+s_2})$ *is an optimal portfolio that is correlation-invariant.*

It is interesting to compare the correlation-invariant portfolio with the prediction in the presence of ambiguity over correlation structure in recent literature. In the framework of Levy and Razin (2020), decision makers combine multiple sources of information to form predictions. Similar to our setting and Carroll (2017), the decision makers understand each information source in isolation but uncertain about the correlation between the sources; thus, the decision makers face ambiguity in relation to the set of predictions. Levy and Razin (2020) characterize the set of rational prediction with a set of joint information structures.

Our characterization of the optimal portfolio under correlation ambiguity is related to but different from the approach in Levy and Razin (2020) in several aspects. *First*, Levy and Razin (2020) identifies one important factor, a naive interpretation of forecasts to ignore correlation (Naive-Bayes belief), in forming rational predictions. Moreover, if this Naive-Bayes belief is relatively precise, the investor behaves as if she completely ignore the correlation issue. [11] In the portfolio choice setting, it is well known that the (identically) independent asset return assumption yields a $\frac{1}{N}$ diversification rule (Samuelson (1967)).

[11] The correlation ignorance (Levy and Razin (2015)) suggests that naive investors make a decision by assuming independent marginal distribution. Levy and Razin (2020) demonstrate correlation-ignorance endogenously. See also Ellis snd Piccione (2017).

Therefore, correlation-ignorance leads to different optimal portfolio from the correlation-invariant approach. *Second*, when the Naive-Bayes belief is not precise, Levy and Razin (2020) demonstrates that the correlation ambiguity yields cautious behavior in portfolio construction, and specifically, the level of correlation bound on pointwise mutual information (PMI) is help to characterize the rational prediction. In contrast, we construct the optimal portfolio in which the ambiguous set of correlation matrix in our framework plays a similar role as the PMI parameter. For instance, in both Example 3.4 - 3.5, $\underline{a}$ and $\overline{a}$ provide the lower and upper bound for the pointwise correlation coefficient, Example 3.4 - 3.5 explicitly demonstrate the effect of the PMI parameter to the ambiguity-inverse investor's optimal portfolio.

3.3 Concentrated portfolio

If an investor knows very little about the joint distribution, Proposition 2 states that the investor chooses a highly concentrated, anti-diversified portfolio optimally. In this section, we further extend Proposition 2 by showing that the investor's optimal portfolio is often concentrated, containing only a few assets, although it might not be anti-diversified in a general situation.

Given an ambiguous set $\mathcal{C}$ of correlation matrix of asset $\{1, \cdots, N\}$, for any subset $\{i_1, \cdots, i_M\}$ of these N assets, define

$$\mathcal{C}(i_1, \cdots, i_M) = \left\{\rho_0 \in \mathbb{S}^M_{++} : \rho_0 \text{ is the } (i_1, \cdots, i_M) \text{ minor of an element } \rho \in \mathcal{C}\right\}.$$

It is easy to check that $\mathcal{C}(i_1, \cdots, i_M)$ is convex and compact. Then we view $\mathcal{C}(i_1, \cdots, i_M)$ as an induced ambiguous set among assets $\{i_1, \cdots, i_M\}$ by $\mathcal{C}$.

To simplify notations, let $\mu_0 = (\mu_{i_1}, \cdots, \mu_{i_M}), \sigma_0 = diag(\sigma_{i_1}, \cdots, \sigma_{i_M})$, and

$$J(i_1, \cdots, i_M; \mathcal{C}) = \max_{\phi_0 \in \mathbb{R}^M} \min_{\rho_0 \in \mathcal{C}(i_1, \cdots, i_M)} \mu_0 \top \phi - \frac{A}{2} \phi^\top (\sigma_0 \rho_0 \sigma) \phi. \tag{8}$$

$J(i_1, \cdots, i_M; \mathcal{C})$ is the mean-variance utility of the optimal portfolio with assets $i_1, \cdots, i_M$ and the corresponding ambiguous set $\mathcal{C}(i_1, \cdots, i_M)$ of the correlation matrix. Hence, Proposition 1 can be applied to the max-min problem $J(i_1, \cdots, i_M; \mathcal{C})$ for any subset of risky assets.

PROPOSITION 4. *The optimal portfolio contains assets in* $\{1, \cdots, M\}$ *if and only if there exists a worts-case correlation matrix* ρ_0^* *for the max-min problem* $J(1, \cdots, M; \mathcal{C})$ *and a matrix* ζ *of size* $M \times (N - M)$ *such that*

$$(s_{M+1}, \cdots, s_N) = (s_1, \cdots, s_M)\rho_0^{*-1}\zeta$$

such that

$$\begin{pmatrix} \rho_0^* & \zeta \\ \zeta^\top & \rho_1 \end{pmatrix} \in \mathcal{C}$$

for one $\rho_1 \in \mathcal{S}_{++}^{N-M}$.

Proposition 2 is a special case of Proposition 4 as follows. Let $\zeta = (s_2/s_1, \cdots, s_N/s_1)$. Then the optimal portfolio only holds asset 1 if and only if there exists one feasible matrix $\rho \in \mathcal{C}$ such that $\zeta = (\rho_{1,2}, \cdots, \rho_{1,N})$. Proposition 4 characterizes the general situation in which assets $M+1, \cdots, N$ do not enter the optimal portfolio if their investment opportunities can be generated by assets $1, \cdots, M$.

As a consequence, the next result presents the condition under which one particular asset is not included in the optimal portfolio.

Corollary 3.1. *If there exists a worst-case correlation matrix* ρ_0 *for the max-min problem* $J(1, \cdots, N-1; \mathcal{C}\}$, *and one element* $\rho \in C$ *such that* $\rho_{ij} = (\rho_0)_{ij}$ *for* $1 \leq i, j \leq N-1$, *and*

$$s_N = (s_1, \cdots, s_{N-1})(\rho_0)^{-1}(\rho_{1,N}, \cdots, \rho_{N-1N})^\top, \tag{9}$$

then the asset N is not included in the optimal portfolio.

Corollary 3.1 can be understood as follows. Assume first the asset N is uncorrelated with all other assets, then, this asset is not included in the optimal portfolio if and only if $\mu_N = 0$.[12] In general, let $\tilde{\beta} = (\rho_{1,N}\frac{\sigma_1}{\sigma_N}, ..., \rho_{N-1,N}\frac{\sigma_{N-1}}{\sigma_N})^\top$. $\tilde{\beta}$ is the population regression coefficient of r_N on $(r_1, ..., r_{N-1})$. Then the space of all return $\{r_1, r_2, \cdots, r_N\}$ can be written as $X \oplus Y$ where X consists of $r_1, \cdots, r_{N-1}$, and Y is generated by one asset return with zero expected value if and only if Equation (9) holds. Since the asset in Y is uncorrelated with assets in X and the expected return is zero by Equation (9), the diversification benefit is zero by adding asset in Y; thus, asset N is not included in the optimal portfolio. The intuition of Proposition 4 is similar from the diversification perspective.

Proposition 4 implies that those risky assets in the optimal portfolio under correlation ambiguity have either small correlations with each other or very different risk-return characteristics (Sharpe ratios). If asset N is highly correlated with other assets or the investment opportunity of asset N is generated by other assets (Equation (9) likely holds), the investor intends to exclude it because the benefit of diversification by adding this asset is zero. By using Corollary 3.1 repeatedly, we often come up with an optimal portfolio that is concentrated.

Similar to Carroll (2017), Proposition 4 can be also used when investor has some knowledge about asset correlations. Let us divide all risky assets into several smaller class of the risky asset, $\mathcal{N}_1, \cdots, \mathcal{N}_k$, and the investor knows the joint distribution of assets in each class $\mathcal{N}_i$. The investor can construct an optimal portfolio with assets in $\mathcal{N}_i$ by using the known joint distribution for this smaller class of risky assets. We also assume that $\mathcal{N}_1 = \{1, \cdots, M\}$, and $J(1, \cdots, M; \mathcal{C})$ is larger than the value obtained from other class $\mathcal{N}_j, j > 1$. It means that the optimal portfolio in the class $\mathcal{N}_1$ dominates the optimal portfolios in all other classes $\mathcal{N}_i$ separably. Extending Proposition 2, Proposition 4 characterizes the condition under which

[12] Similarly, for any increasing and concave function $U(\cdot)$, if asset N is independent of all other assets, then asset N is not included in the optimal portfolio if and only if its excess expected return is zero (Samuelson (1967), Theorem III).

the investor can treat each class separably and choose $\mathcal{N}_1$ to construct the optimal portfolio for all assets.

The next example shows the importance of the independent assumption between some asset returns, when some $\mathcal{N}_i$ contain multiple assets.

Example 3.6. *Assume $N = 3$, and ρ_{12} is known, $\mathcal{N}_1 = \{1, 2\}$ and $\mathcal{N}_2 = \{3\}$. We assume that $J(3) < J(1, 2)$. On one hand, if $\rho_{12} = 0$, then the optimal portfolio contains assets in $\{1, 2\}$ if and only if there is one correlation matrix $\rho \in \mathcal{C}$ such that $s_3 = s_1\rho_{1,2} + s_2\rho_{2,3}$. On the other hand, if $\rho_{12} \neq 0$, it is plausible that asset 3 enters the optimal portfolio for any ambiguous set $\mathcal{C}$.*

Consider a situation with three risky assets and order them by the Sharpe ratio in decreasing order, $i-1, i, i+1$. If asset $i-1$ is independent of the asset $i+1$. Then, the asset i does not enter the optimal portfolio if there exists one feasible correlation matrix such that $s_i = s_{i-1}\rho_{i.i-1} + s_{i+1}\rho_{i,i+1}$, whereas asset $i+1$ might enter the optimal portfolio because it adds the benefit of diversification. It is more interesting though to notice that asset $i-1$ might not enter the asset when asset i is independent of asset $i+1$, if $s_{i-1}^2 < s_i^2 + s_{i+1}^2$ holds. However, if asset i and asset $i+1$ are correlated and the investor perfectly knows the correlation coefficient $\rho_{i,i+1}$, although $J(i, i+1) > J(i-1)$ in certain situations in Example 3.6, asset $i-1$ still enters the optimal portfolio for any ambiguous set $\mathcal{C}$. Therefore, the correlation structure in each $\mathcal{N}_i$, and the correlation structure between any two classes matter in constructing the optimal portfolio.

4 Portfolio Inertia

According to our analysis of Section 3, when the ambiguous set of the correlation matrix changes, the optimal portfolio is not necessarily changed. For instance, the optimal portfolio under correlation ambiguity does not smoothly depend on the correlation matrix in the anti-diversification situation. This property is often called portfolio inertia in literature.

This section demonstrates a new portfolio inertia property of the optimal portfolio under correlation ambiguity for multiple assets. Namely, when the ambiguous set is given and fixed, the optimal portfolio under correlation ambiguity does not change if the Sharpe ratio vector (the marginal distribution) changes in a non-trivial region. Since the Sharpe ratio measures the investment opportunity of each asset, this portfolio inertia property states that the investor chooses the *same* optimal portfolio even though the investment opportunity of each asset changes. [13]

4.1 General portfolio inertia property

For simplicity, we first fix the volatility vector σ, and consider all possible Sharpe ratio vector $s \in \mathbb{R}^N_{++}$, the set of vector with all positive components. Later we explain the portfolio inertia property is robust when the volatility changes.

We use $\mathbb{R}^N_{++}$ to denote the set of all Sharpe ratios because of positive expected excess returns of risky assets. Given any ambiguous set $\mathcal{C}$ of the correlation matrix, Proposition 1 shows a unique optimal portfolio, $\frac{1}{A}\sigma^{-1}\rho^{*-1}(s)s$, for a worst-case correlation matrix $\rho^*(s) = argmin_{\rho\in\mathcal{C}} s^\top \rho^{-1} s$. Therefore, there is a well-defined map

$$F_{\mathcal{C}} : \mathbb{R}^N_{++} \to \mathbb{R}^N, F_{\mathcal{C}}(s) = \frac{1}{A}\sigma^{-1}\rho^{*-1}(s)s.$$

If there is no correlation ambiguity, $\mathcal{C} = \{\rho\}$, the map $F_{\mathcal{C}}$ is clearly injective: $s_1 \neq s_2$ implies $\sigma\rho^{-1}s_1 \neq \sigma\rho^{-1}s_2$. By portfolio inertia in out setting we mean that the map $F_{\mathcal{C}}$ is *not* injective. Since the *level set* $F_{\mathcal{C}}^{-1}(\phi)$ consists of all Sharpe ratios with the same optimal portfolio ϕ, our approach is to investigate the structure of the level set and show that the level set is fairly large.

[13] A range of Sharpe ratio can be easily transformed to a range of asset price given a known marginal distribution of asset's future cash flow. Specifically, we write the asset excess return as $\tilde{R} = \frac{\tilde{d}}{p} - r_f$ for a future cash flow $\tilde{d}$ and the asset price p. Then the Sharpe ratio $s = \frac{\mu_d - pr_f}{\sigma_d}$, where μ_d is the expected value of the future cash flow $\tilde{d}$ and σ_d is the standard deviation of $\tilde{d}$. (μ_d, σ_d) is determined by the marginal distribution of the future cash flow $\tilde{d}$. Therefore, a region of the Sharpe ratio vectors is mapped to a region of asset prices.

Definition 4.1. *For an ambiguous set $\mathcal{C}$, a vector $\phi \in \mathbb{R}^N$ has a portfolio inertia property if $dim F_{\mathcal{C}}^{-1}(\phi) \geq 1$.*

In this definition, we use a dimension concept, $dim(X)$, in differential topology to count the points and distinguish subsets in $\mathbb{R}^N$.[14] Briefly speaking, a differential manifold with dimension m is locally likes $\mathbb{R}^m$, and we can do multivariable calculus on it. For a (C^∞) smooth map between two differential manifolds, $F : X \to Y$, we can use the Jacobian matrix DF_p at a point $p \in X$ to analyze the property of F locally.

PROPOSITION 5. *Assume the map $F_{\mathcal{C}}$ is smooth and the dimension of the image set of $F_{\mathcal{C}}$ is strictly smaller than N, then almost all portfolio in the image set of $F_{\mathcal{C}}$ has a portfolio inertia property. In general, if X is a submanifold of $\mathbb{R}^N_+$, and $dim(F_{\mathcal{C}}(X)) < dim(X)$, then almost all portfolio $\phi \in F_{\mathcal{C}}(X)$, the level set $F_{\mathcal{C}}^{-1}(\phi) \cap X$ is a submainfold of X with dimension $dim(X) - dim(F_{\mathcal{C}}(X))$.*

By Sard's theorem in differential topology, the level set of a smooth map is closely related to its image set. Technically speaking, the portfolio inertia property on the level set $F^{-1}(\phi)$ follows from the characterization of the optimal portfolio in Section 3 and the Sard's theorem.

Remark 4.1. *Since the volatility vector σ is given and fixed, the Sharpe ratio changes equivalents to the expected return changes. To highlight the effect of the volatility vector, we now use $F_{\mathcal{C}}^{\sigma}$ to replace $F_{\mathcal{C}}$. Given a submanifold $X \subseteq \mathbb{R}^N_{++}$, the image set $F_{\mathcal{C}}^{\sigma}$ is $\left\{\frac{1}{A}\sigma^{-1}\rho^{*-1}(s)s : s \in \mathbb{R}^N_{++}\right\}$, where $\rho^*(s)$ depends only on s. Notice the map, $(x_1, \cdots, x_N) \to (\frac{x_1}{\sigma_1}, \cdots, \frac{x_N}{\sigma_N})$, is diffeomorphism. Then, the dimension of the image set $F_{\mathcal{C}}^{\sigma}(X)$ is independent of the choice of volatility vector σ. If $dim F_{\mathcal{C}}^{\sigma_0}(X) < \dim(X)$ for one volatility vector σ_0, then for any other volatility vector σ, we have $dim F_{\mathcal{C}}^{\sigma}(X) = dim F_{\mathcal{C}}^{\sigma_0}(X) < \dim(X)$; hence, by Proposition 5, almost all vectors in $F_{\mathcal{C}}^{\sigma}(X)$ has the portfolio inertia property.*

[14] See Lee (2013) and Milnor (1997) for basic concepts in differential topology. The dimension concept is diffeomorphism invariant in a sense that a nonempty C^∞ smooth manifold of dimension m cannot be diffeomorphic to an n-dimensional manifold unless $m = n$. See Lee (2013), Theorem 2.17. Since the maps we use are C^∞ smooth, we do not distinguish the terminology between differential and (C^∞) smooth manifold, and we sometimes simply refer to manifold. Similarly, a smooth map means a C^∞ smooth map. Notice that only a set with finitely many points has dimension zero. Therefore, a set X with $dim(X) \geq 1$ contains at least infinitely many points.

Next, we fix the expected return, so a range of Sharpe ratio is equivalent to a region of the volatility. To investigate the effect of the changing volatility to the optimal portfolio uncer correlation ambiguity, we fix an expected return vector $\mu \in \mathbb{R}^N_{++}$, and define

$$G_{\mathcal{C}} : (s_1, \cdots, s_N) \in \mathbb{R}^N_{++} \to \frac{1}{A} diag(\frac{s_1}{\mu_1}, \cdots, \frac{s_N}{\mu_N}) \rho^{*-1}(s) s \tag{10}$$

This map $G_{\mathcal{C}}$ is also well-defined by Proposition 1. Similar to Proposition 5, we have the following portfolio inertia property.[15]

PROPOSITION 6. *Assume $G_{\mathcal{C}}$ is smooth and $X \subseteq \mathbb{R}^N_{++}$ is a submanifold with $dim(G_{\mathcal{C}}(X)) < dim(X)$, then almost all portfolios in $G_{\mathcal{C}}(X)$ have a portfolio inertia property. That is, $dim G^{-1}_{\mathcal{C}}(\phi) \geq 1$, and when the Sharpe ratio vector moves in the level set $G^{-1}_{\mathcal{C}}(\phi)$, the optimal portfolio is the same portfolio vector ϕ. Moreover, this portfolio inertia is robust with the changes of the expected return vector $\mu \in \mathbb{R}^N_{++}$.*

In Proposition 5 - Proposition 6, the smooth condition of the map $F_{\mathcal{C}}$ and $G_{\mathcal{C}}$ is technical, which mean the smoothness effect of the Sharpe ratio to the optimal portfolio. It is intuitive since the investor has no ambiguity about the marginal distribution. For a given correlation ambiguous set, this smooth condition can be checked directly. The existence of a submanifold X such that $dim(F_{\mathcal{C}}(X)) < dim(X)$ or $dim(G_{\mathcal{C}}(X)) < dim(X)$ follow from the properties of the optimal portfolio under correlation ambiguity in Section 3.

Similar to the correlation-invariant or concentrated feature of the optimal portfolio, this portfolio inertia property for an optimal portfolio also follows from the portfolio diversification and ambiguity aversion effect. Its mechanism is as follows. The investor considers each asset based on its marginal distribution or the Sharpe ratio. If there is only one asset, investors buy it if it becomes more attractive (a higher Sharpe ratio) and sell it short if the Sharpe ratio decreases (less attractive). For a portfolio with multiple assets, the investor

[15] We can apply the same approach to the situation in which both the expected return vector and the volatility vector change, that is, the marginal distribution change in a non-trivial region. See Example 4.4 for an illustration.

considers the diversification benefits in addition to the attractiveness of each asset (marginal distribution). Because of ambiguity-aversion to the correlation ambiguity, the investor optimally chooses a correlation-invariant or concentrated portfolio with zero or relatively small effect of asset return correlation. Since the diversification benefit is small, the investor does not want to trade away from these optimal portfolios, even though the Sharpe ratio changes inside a reasonably large region.

4.2 Examples of portfolio inertia

In this section we present several examples to illustrate the portfolio inertia property under correlation ambiguity.

We start with a situation in which the optimal portfolio is anti-diversified.

Example 4.1. *Let* $X = \left\{(s_1, \cdots, s_N)^\top \in \mathbb{R}^N_{++}, s_i = s_1\rho_{1i}, i = 2, \cdots, N, \rho \in \mathcal{C}\right\}$. *Then by Proposition 2,* $F_{\mathcal{C}}(X) = \left\{\phi = (\phi_1, 0, \cdots, 0)^\top, \phi_1 > 0\right\}$, *so the dimension of* $F_{\mathcal{C}}^{-1}(\phi) \bigcap X$ *is at least* $dim(X) - 1$. *If the dimension of* X *is at least two, then each portfolio in* $F_{\mathcal{C}}(X)$ *has a portfolio inertia property. For instance, for* $N = 3$, *and* $\mathcal{C} = \left\{\rho \in \mathcal{B} : \rho_{12} \in [\underline{\rho}_{12}, \overline{\rho}_{12}], \rho_{13} \in [\underline{\rho}_{13}, \overline{\rho}_{13}]\right\}$, *we have* $dim(X) = 2$ *and* $dim(F_{\mathcal{C}}^{-1}(\phi) \bigcap X) = 1, \forall \phi = (\phi_1, 0, 0)^\top, \phi_1 > 0$.

The next result follows the characterization of the concentrated portfolio in Proposition 4.

Corollary 4.1. *Assuming the investor knows the correlated structure among assets* $\{1, \cdots, M\}$ *perfectly, and fhe correlation matrix is given by* ρ_0. *Let*

$$X = \left\{ (s_1, \cdots, s_N)^\top \in \mathbb{R}^N_{++} : (s_{M+1}, \cdots, s_N) = (s_1, \cdots, s_M)\rho_0^{-1}\zeta, \begin{pmatrix} \rho_0 & \zeta \\ \zeta^\top & \rho_1 \end{pmatrix} \in \mathcal{C} \right\},$$

and $dim(X) \geq M + 1$. *Then for any* $\phi = (\phi_1, \cdots, \phi_M, 0, \cdots, 0)$ *such that* $F_{\mathcal{C}}^{-1}(\phi) \bigcap X$ *is non-empty,* $F_{\mathcal{C}}^{-1}(\phi) \bigcap X$ *is a submanifold of dimension* $dim(X) - M$.

The next example illustrates Corollary 4.1.

Example 4.2. *Assume $N = 3$, and for simplicity, each $\sigma_i = 1, 1 \leq i \leq 3$. The ambiguous set is*

$$\mathcal{C} = \left\{ \begin{pmatrix} 1 & 0 & a \\ 0 & 1 & a \\ a & a & 1 \end{pmatrix} : a \in [0.2, 0.7] \right\},$$

There is portfolio inertia in this situation.

To see it, for any $s \in X$, $s_3 = as_1 + as_2, a \in [0.2, 0.7]$. Then by Proposition 4 and Corollary 4.1, asset 3 does not enter in the portfolio in $F_{\mathcal{C}}(X)$. Hence, each portfolio in $F_{\mathcal{C}}(X)$ has a portfolio inertia property.

We next show the portfolio inertia property for optimal correlation-invariant portfolio. Let $\mathcal{S} = \left\{ s \in \mathbb{R}^N_{++} : F_{\mathcal{C}}(s) \text{ is correlation-invariant with respect to } \mathcal{C} \right\}$, and $\mathcal{Z} = \{F_{\mathcal{C}}(s) : s \in \mathcal{S}\}$, the set of optimal portfolio that is correlation-invariant.

The next result follows from Proposition 5 easily.

Corollary 4.2. *If both $\mathcal{S}$ and $\mathcal{Z}$ have smooth manifold structures and $dim(\mathcal{Z}) < dim(\mathcal{S})$, then for* almost all *$\phi \in \mathcal{Z}$*[16]*, the dimension of $F^{-1}(\phi)$ is at least $dim(\mathcal{S}) - dim(\mathcal{Z})$. In particular, if $\mathcal{S}$ is a manifold of dimension N, then for almost all portfolio ϕ in $\mathcal{Z}$, the dimension of $F_{\mathcal{C}}^{-1}(\phi)$ is at least $N - dim(\mathcal{Z}) \geq 1$, and the portfolio inertia occurs for the ambiguous set $\mathcal{C}$.*

We use two examples to demonstrate the portfolio inertia that is generated by the correlation-invariant portfolio in Corollary 4.2. The first example is the Engle-Kelly's block dynamic equicorrelation (DECO) model.

Example 4.3. *Let $\mathcal{C} = \{[R(a_1)_{N_1 \times N_1}; \cdots; R(a_k)_{N_k \times N_k}] : \underline{a}_i \leq a_i \leq \overline{a}_i, i = 1, \cdots, k\}$, a block DECO model in Engel and Kelly (2012), then almost all portfolio in $\mathcal{Z}$ has a portfolio*

[16] By almost all we mean the points in $\mathcal{Z}$ which satisfies that the dimension of $F_{\mathcal{C}}^{-1}(\phi)$ is smaller than $dim(\mathcal{S}) - dim(\mathcal{Z})$ has a measure zero.

inertia property. Moreover, for almost all portfolio $\phi \in G_{\mathcal{C}}(\mathcal{S})$, its level set under the smooth map $G_{\mathcal{C}}$ is of dimension greater than one.

As the second example, we consider Example 4.2 again. By Example 3.5, any portfolio in $F_{\mathcal{C}}(X)$ is correlation-invariant. Since X is a smooth manifold and $dim(X) = 3$, by Corollary 4.2, almost all portfolio in $F_{\mathcal{C}}(X)$ have a portfolio inertia property.

Finally, we illustrate that the portfolio inertia property under correlation ambiguity when both the expected return and volatility changes in a non-trivial region. The following example is motivated by Illeditsch (2011) and Epstein and Schneider (2010).

Example 4.4. *Consider two risky assets and the ambiguous set of the correlation matrix is $\mathcal{C}[\underline{a}, \overline{a}]$. Let $(\mu_i^e, \overline{\sigma}_i^2)$ be a benchmark expected return and variance of asset $i = 1, 2$. The region of the expected return and variance of asset i is given by $\mu_i = \mu_i^e + x_i, \sigma_i^2 = \overline{\sigma}_i^2 + \frac{2x_i}{A}$ and $0 \leq x_i \leq \overline{x}_i$. Here $\overline{x}_1$ and $\overline{x}_2$ are two given positive number. There exists portfolio inertia in this situation.*

To see it, we consider the set $\{(\mu_1, \sigma_1, \mu_2, \sigma_2), 0 \leq x_1 \leq \overline{x}_1, 0 \leq x_2 \leq \overline{x}_2\}$. Let $S_i(x_i) = \frac{\mu_i^e + x_i}{\overline{\sigma}_i^2 + \frac{2x_i}{A}}$ be the Sharpe ratio of asset i for the parameter $x_i \in [0, \overline{x}_i], i = 1, 2$. Let $X = \left\{(x_1, x_2) \in [0, \overline{x}_1] \times [0, \overline{x}_2] : \underline{a} \leq \min\left(\frac{S_1(x_1)}{S_2(x_2)}, \frac{S_2(x_2)}{S_1(x_1)}\right) \leq \overline{a}\right\}$. We can show that $dim(X) = 2$. By Proposition 2, the ambiguous-averse investor only invests on one risky asset with the highest Sharpe ratio when $(x_1, x_2) \in X$, that is, the corresponding optimal portfolio is contained in $Y = \{(\phi_1, 0), (0, \phi_2), \phi_1, \phi_2 > 0\}$. Since $dim(Y) = 1$, then for each portfolio $\theta \in Y$, there exists a subminifold X_0 of X of dimension 1 such that the optimal portfolio is the portfolio ϕ, as the expected return and volatility (marginal distribution) of each asset i changes in the region $(x_1, x_2) \in X_0$. By the same idea, we can demonstrate similar portfolio inertia property when both the marginal distributions change in a non-trivial region in other situations.

5 Discussions

This section first provides empirical evidence and implications of our results in Section 3 - Section 4. Then we compare the optimal portfolio under correlation ambiguity and the Sharpe ratio ambiguity.

5.1 Empirical evidence

Our theoretical results are consistent with numbers of empirical studies. First of all, extant empirical studies document that investors usually hold much less risky assets than they could. For example, Campbell (2006) finds that the financial portfolios of households contain only a few risky assets. Goetzmann and Kumar (2008) report that the majority of individual investors hold a single-digit number of assets in a sample data set from 1991-1996. Among many empirical findings regarding under-diversification from various data sets, we refer to Mitton and Vorkink (2007), Calvet, Campbell, and Sodini (2008), and Ivković, Sialm, and Weisbenner (2008). Proposition 4 (and Proposition G.1) suggest that correlation ambiguity explains those empirical findings on concentrated optimal portfolios or portfolio concentration.

Second, many empirical studies have documented risky asset inertia in household portfolios or pension funds. For instance, Agnew, Balduzzi, and Sunden (2003) study nearly 7,000 retirement accounts during the April 1994-August 1998 period, and they find that most asset allocations are extreme (either 100 percent or zero percent in equities) and there is inertia in asset allocations. Bilias, Geogarakos, and Haliassos (2007) uses data representative of the population to document the extent of household portfolio inertia and to link it to household characteristics and to stock market movements. They document considerable portfolio inertia, as regards both changing stockholding participation status and trading stocks, and find that specific household characteristics contribute to the tendency to exhibit such stock inertia. By using data from the Panel Study of Income Dynamics, Brunnermeier and Nagel

(2008) find out that households rebalance only very slowly following inflows and outflows or capital gains and losses. Our ambiguity-aversion model (Section 4) shows that investors' aversion to the correlation uncertainty could be one reason for the risky assets inertia inside their portfolios.

5.2 Implications

Our partial equilibrium model shows that the correlation ambiguity-aversion investor tends to hold a part of stocks and holds these stocks passively. The properties of the optimal portfolio under correlation ambiguity might explain the increasing trading of the index.

Over 40 years, the amount of capital devoted to index investing has grown by more than 4 trillion dollars by 2016 (Bogle (2016)). Some researchers contribute to a low fee of passive investors, and active managers do not outperform the market after fees. Therefore, by creating some indexes (ETFs), it might be optimal to hold these indexes passively. Bond and Garcia (2020) develop a rational expectation equilibrium model of the index and demonstrate the indexing improves price efficiency. Hirshleifer, Huang, and Tech (2019) also develop a rational expectation equilibrium of information asymmetry and ambiguity aversion. With a well-designed risk-adjusted market portfolio (RAMP), Hirshleifer, Huang, and Tech (2019) show that each investor holds the RAMP in equilibrium while the standard weighted-weighted market portfolio (VWMP) does not help the ambiguity-averse investor to participate in the market index.

Our model provides an alternative way to look at the construction of indexes and ETFs. Specifically, an index is appealing if it maximizes the diversification benefit and reduces inefficiency, such as transaction costs. In this regard, the attractiveness (marginal distribution) of each risky asset is known. The issue is to select a few risky assets to span sufficient investment opportunities, and the structure of the portfolio is not significantly sensitive to the market movement of these assets to reduce the rebalancing cost. Therefore, some well-designed optimal portfolios under correlation ambiguity might serve the role of ETFs. For instance, by

choosing several risky assets in Proposition 4 and Corollary 4.1, we have shown that such a portfolio is optimal and does not change with some market movement of risky assets. For the same reason, when a portfolio is invariant to a set of asset correlation structures, it is also optimal to keep this portfolio even when these assets' investment opportunities change in the market (Proposition 3 and Corollary 4.2). Since the ambiguity-aversion effect of the correlation structure makes the portfolio *less sensitive* (portfolio inertia) and *less diversified* (concentrated or anti-diversified) than it would be otherwise, the optimal portfolio under correlation uncertainty might be an attractive candidate for passive investment tool in the market.

We can justify the indexing alternatively. In our discussion of Proposition 2 about the inconsistency issue, let us start with a set of assets $Y_1, \cdots, Y_N$, and consider N portfolios $X_1, \cdots, X_N$ of the underlying assets $Y_1, \cdots, Y_N$. If the marginal distribution of these N basic portfolios $X_1, \cdots, X_N$ are precise, but the induced ambiguous set of these basic portfolios contains a particular correlation matrix, then Proposition 2 implies that one index from $\{X_1, \cdots, X_N\}$ is optimal. Moreover, the investor does not rebalance this index with some market movements according to Corollary 4.1 (see also Proposition 5 - Proposition 6). Our results suggest that it is feasible to index the financial market (with any N assets) such that this index is not necessarily rebalanced continually from an optimal portfolio choice perspective.

5.3 Comparison with the Sharpe ratio uncertainty

In this subsection, we present a comparison between correlation ambiguity and the expected return or variance ambiguity.

Before performing this comparison, we extend Proposition 1 to include both the marginal distribution ambiguity and the correlation ambiguity. Consider a subset $\mathcal{U} \subseteq \mathbb{R}^N \times \mathbb{S}^N_{++}$. Here $\mathbb{S}^N_{++}$ denotes the set of a positive definite and symmetric matrix. We assume that $\mathcal{U}$ is a compact and convex subset of $\mathbb{R}^N \times \mathbb{S}^N_{++}$, representing an ambiguous set of the expected

returns vector μ and the covariance-variance matrix Σ. The optimal portfolio choice problem under joint distribution ambiguity is the following max-min problem

$$\max_{\phi} \min_{(\mu,\Sigma)\in\mathcal{U}} \mu^\top \phi - \frac{A}{2}\phi^\top \Sigma \phi, \tag{11}$$

Proposition G.1 in Appendix A solves this max-min problem under a general ambiguous set $\mathcal{U}$ as follows. Let $(\mu^*, \Sigma^*) = \text{argmin}_{(\mu,\Sigma)\in\mathcal{U}} \left(\mu^\top \Sigma^{-1} \mu\right)$, and $\phi^* = \frac{1}{A}\Sigma^{*-1}\mu^*$. Then, ϕ^* is a unique optimal solution of this general max-min problem.[17] Moreover, if there exists $(\mu^*, \Sigma^*) \in \mathcal{U}$ such that the mean-variance utility $\mu^\top \phi^* - \frac{A}{2}\phi^{*\top}\Sigma\phi^*$ is independent of $(\mu, \Sigma) \in \mathcal{U}$, that is, the portfolio ϕ^* is *unambiguous* with respect to the ambiguous set $\mathcal{U}$, then ϕ^* is the unique optimal portfolio.

Assuming the correlation matrix is known and asset returns are independent, then, the worst-case (μ^*, σ^*) in Proposition G.1 is given by

$$(\mu^*, \sigma^*) = \text{argmin}_{(\mu,\sigma)} s^\top s, \phi^* = \frac{1}{A}\sigma^{*-1}\sigma^{*-1}\mu^*. \tag{12}$$

Previous results in Cao, Wang, and Zhang (2005), Epstein and Schneider (2010), and Easley and O'Hara (2009) can be derived easily from Equation (12). For instance, a sufficiently large of ambiguity on the expected return of each asset implies $\mu_i = 0, i = 1, \cdots, N$ for an element (μ, σ) in the ambiguous set, then the investor is away from the risky assets (no-participation).

Indeed, the expected return uncertainty could also lead to under-diversification or a concentrated portfolio. For instance, if asset 1, $\cdots$, asset M are not attractive in the sense that their expected returns could be zero, Equation (12) implies that these assets do not

[17] Epstein and Schneider (2010) formulate the same problem and demonstrate that it can be also applied in a continuous-time setting. When the covariance-variance matrix is known, Garlappi, Uppal, and Wang (2007) solve the primary problem in Proposition G.1 for several examples of ambiguous set of the expected return vector explicitly by a topology method. Kim and Boyd (2008) solve a relevant max-min problem in which the objective function is the Sharpe ratio of a portfolio. As shown in Kan and Zhou (2007), the problem (11) is appealing to finance since the robust Sharpe ratio measure in Kim and Boyd (2008) is independent of the leverage of the portfolio.

enter the optimal portfolio.[18] However, the reason to choose such a concentrated portfolio is different from the correlation ambiguity. Under expected return ambiguity and asset returns are independent, the ambiguity-aversion investors take zero positions in the unattractive assets they are ambiguous about. In contrast, the concentrated portfolio under correlation ambiguity follows from the consideration of the diversification benefit. As shown in Proposition 4 and Corollary 3.1, even though each asset is attractive, the investor excludes asset i when this asset is highly correlated with other assets and its investment opportunity can be spanned by other assets. Therefore, the correlation ambiguity derives a concentrated portfolio from a different channel of the expected return ambiguity.

More importantly, the portfolio inertia property differs significantly between correlation uncertainty and Sharpe ratio uncertainty in literature. For instance, in the expected return ambiguity, there is portfolio inertia only when some assets do not enter the optimal portfolio. That is, there is zero holding when one asset is not attractive regardless of other assets. As another example, Illeditsch (2011) considers a signal quality ambiguity in which an ambiguous signal affects both the expected return and the volatility of a risky asset, say, $\mu = \mu^e + x, \sigma^2 = \overline{\sigma}^2 + \frac{2x}{A}$, where A is the risk aversion parameter, and $0 \leq x \leq \overline{x}$ for a fixed positive number $\overline{x}$. If $\overline{x} \geq \mu^e - A\overline{\sigma}^2 > 0$, then there is a portfolio inertia for the non-zero portfolio $\theta^* = 1$.[19] In this group of literature about the Sharpe ratio or marginal distribution uncertainty, the asset return correlation matrix is *known*, the investor's ambiguity aversion leads to the same optimal portfolio when the asset's marginal distribution changes in a range. By contrast, we demonstrate the portfolio inertia property when the correlation matrix is *unknown* and the investor is ambiguous averse to a given ambiguous set $\mathcal{C}$ of correlation matrix. We demonstrate that the optimal portfolio is still the same, although each asset's Sharpe ratio (marginal distribution) changes in a non-trivial region. This general portfolio

[18] See Easley and O'Hara (2009), and Epstein and Ji (2013). As explained in Epstein and Schneider (2010), the volatility ambiguity changes the positions on the risky assets but does not affect the participation feature.

[19] It is straightforward to derive non-zero portfolio which has the portfolio inertia property for multiple assets.

inertia property is consistent with asset allocation inertia in pension portfolios for multiple assets.

A natural question is which correlation ambiguity and Sharpe ratio ambiguity has a first-order effect on portfolio diversification. We first consider the expected return ambiguity. For simplicity, we assume asset returns are independent. For a dependent case, the economic insight is similar. Assuming $\overline{\mu} = (\overline{\mu}_1, \cdots, \overline{\mu}_N)$ is a benchmark expected return and σ is the volatility vector, The expected return changes in the region $[\overline{\mu}_1 - \epsilon, \overline{\mu}_1 + \epsilon] \times \cdots \times [\overline{\mu}_N - \epsilon, \overline{\mu}_N + \epsilon_N]$. Let $\phi(\epsilon)$ the corresponding optimal portfolio. Then

$$\frac{\partial \phi(\epsilon)_i}{\partial \epsilon} = \frac{1}{A\sigma_i^2} \neq 0.$$

It means that the expected return uncertainty introduces the first order effect of risk version in the sense of Segal and Spival (1990). Moreover, $\frac{\partial J(\epsilon)}{\partial \epsilon} \neq 0$, the first-order effect to the expected utility, where $J(\epsilon)$ denotes the expected utility under expected return uncertainty.

We now apply the same analysis to the correlation uncertainty. Recall the expected utility is $\mu^\top \phi - \frac{A}{2}\phi^\top \sigma\rho\sigma\phi$. Even though the correlation matrix ρ only appears in the second term $\phi^\top \sigma\rho\sigma\phi$, it actually influences the first term $\mu^\top\phi$ through its affect to the optimal portfolio ϕ. In other words, a first-order effect to the optimal portfolio ϕ also implies the first-order effect of the expected utility. When the investor is ambiguous averse to the correlation uncertainty and the investor's preference is represented in a multiple prior framework, the correlation uncertainty also implies a first-order effect in risk aversion, in a generic case.

We first illustrate this point by two following examples.

Example 5.1. *Assuming $\rho_0 = R(a_0)$ is a benchmark correlation matrix, $a_0 \in (-1, 1)$, and the ambiguous set of the correlation matrix is $\mathcal{C}[a_0 - \epsilon, a_0 + \epsilon]$ for sufficiently small $\epsilon > 0$. Let $\phi(\epsilon)$ denotes the optimal portfolio for this ambiguous set. Then, the correlation ambiguity leads to the first-order effect if and only if $a_0 \neq \frac{1-\Omega(s)}{1+(N-1)\Omega(s)}$.*

Example 5.2. *Consider a benchmark correlation matrix $\rho_0 = T(a_0)$, and the correlation*

matrix moves in the region of the ambiguous set in Example 3.4 with $a_0 - \epsilon \leq a \leq a_0 + \epsilon$. *Then the correlation ambiguity implies a first-order effect if and only of* a_0 *is* not *a local minimal point of the function* $f(a) = s^\top T(a)^{-1} s$.

The above two examples show that the correlation ambiguity leads to the first-order effect in risk aversion and expected utility in a generic case. It is also true in general as follows. Let ρ_0 be a benchmark correlation matrix, and the ambiguous set $\mathcal{C} = \rho_0 + \epsilon\mathcal{A}$ for a suitable set $\mathcal{A}$. The optimal portfolio is written as $\phi(\epsilon)$ to highlight the role of ϵ, and $\phi(\epsilon) = argmin_{\rho \in \rho_0 + \epsilon A} s^\top \rho^{-1} s$. Notice that for each $A \in \mathcal{U}, (\rho_0 + \epsilon A)^{-1} = \rho_0 - \epsilon \rho_0^{-1} A \rho_0 + \frac{1}{2}\epsilon^2 (\rho_0^{-1} A)^2 \rho_0^{-1} + \cdots$. In a generic case, the worst-case correlation matrix is the "corner solution" of the multiple prior preference model (as explained with many examples in Epstein and Schneider (2010)), ϵ has a first-order effect to $(\rho_0 + \epsilon A)^{-1}$ and the optimal portfolio $\phi(\epsilon)$. In the next section, we will show that the smooth ambiguity model leads to an "interior solution" of the worst-case correlation matrix, implying a second-order effect in risk aversion. Nevertheless, in a non-generic case that $\epsilon = 0$ is a locally minimal point of the function $s^\top (\rho_0 + \epsilon A)^{-1} s$, because of the portfolio inertia feature, the correlation ambiguity has no effect to the optimal portfolio and the expected utility.

6 Smooth Ambiguity Models

In this section we demonstrate that both correlation-invariant and portfolio inertia do **not** satisfy in an alternative ambiguity model.

We consider an investor who has marginal distribution on each return, and his ambiguous about the joint distribution is denoted by an ambiguous set $\mathcal{C}$ of correlation matrix. To compare with the max-min mean-variance problem (4), we assume that the joint distribution of asset returns belongs to a class of multivariate normal distribution, and the investor's preference is $u(w) = -e^{-Aw}$, where A is the constant absolute risk aversion parameter. Note that, for any portfolio vector ϕ, W is the portfolio wealth following this strategy, $\mathbb{E}[u(W)]$

is $u\left(f(\phi,\rho)\right)$ where $f(\phi,\rho)=\mu^{\top}\phi-\frac{A}{2}\phi^{\top}\Sigma\phi$. Therefore, the max-min problem

$$\max_{\phi}\min_{\rho\in\mathcal{C}}\mathbb{E}[u(W)] \tag{13}$$

is the same as the max-min problem (4).

As a comparison to the optimization problem (13), we study the optimal portfolio in a smooth ambiguity model of Klibanoff, Marinacci, and Mukerji (2005). Let $\mathcal{P}$ be a prior distribution of the correlation matrix $\rho\in\mathcal{C}$, and the investor's correlation ambiguity is expressed by this second-order distribution, and a concave function $h(x)$ measures the investor's attitude towards ambiguity aversion to the correlation matrix. Specifically, the investor prefers one portfolio ϕ to another portfolio $\tilde{\phi}$ in this smooth ambiguity model if and only if $\mathbb{E}_{\mathcal{P}}\left[h\left(\mathbb{E}_{\rho}[u(W)]\right)\right]\geq\mathbb{E}_{\mathcal{P}}\left[h\left(\mathbb{E}_{\rho}[u(\tilde{W})]\right)\right]$, where $\mathbb{E}_{\rho}[\cdot]$ denotes the expectation operator for which the asset return correlation is ρ. We do not specify the prior distribution $\mathcal{P}$ of the correlation matrix, and refer to Murihead (1982) for the discussion of the distribution of correlation matrix for multivariate normal distribution. In its general version, the investor's optimal portfolio choice problem is

$$\max_{\phi}\mathbb{E}_{\mathcal{P}}\left[h\left(\mathbb{E}_{\rho}[u(W)]\right)\right]. \tag{14}$$

Following Caskey (2009), Gollier (2011), Condie, Ganhuli, and Illeditsch (2021), we choose a power-specification of the function $h(x)$ to exhibit constant relative ambiguity aversion:

$$h(x)=-\frac{1}{1+\alpha}(-x)^{1+\alpha},\alpha\geq 0.$$

If the constant $\alpha=0$, $h(x)=x$, then $\mathbb{E}_{\mathcal{P}}\left[h\left(\mathbb{E}_{\rho}[u(W)]\right)\right]$ is reduced to an expectation under a product probability space over the state space and the correlation matrix state space, which is referred to a Bayesian model uncertainty. Hence, we assume the constant

$\alpha \geq 0$ to include both Bayesian model uncertainty and smooth ambiguity model.

We show that the optimal portfolio is *not* correlation-invariant always, and the portfolio-inertia property fails to be satisfied under the following Assumption A.

Assumption A. For the Sharpe ratio vector s, and any $\tau \in \mathcal{C}$, the map: $\rho \in \mathcal{C} \to s^\top \tau^{-1} \rho \tau^{-1} s$ is not a constant map.

PROPOSITION 7. *For any Sharpe ratio vector $s \in \mathbb{R}^N_{++}$, there exists a unique optimal portfolio ϕ^* of the optimization problem (14). If Assumption A holds, the optimal portfolio ϕ^* is* **not** *correlation-invariant. Moreover, for any $s \in \mathbb{R}^N_{++}$, there is an open region U_s of s such that the map $s \to \phi^*$ is* **one-to-one** *over the region U_s and the dimension of U_s is N.*

Proposition 7 demonstrates significant difference between the smooth ambiguity model and the correlation ambiguity in the multiple-priors framework in several aspects. First, the optimal portfolio is not correlation-invariant. In other words, the optimal portfolio ϕ^* in the smooth ambiguity model must depend on the correlation matrix in the ambiguity set. Second, when the investment opportunities change slightly, the optimal portfolio changes accordingly. Therefore, there is no portfolio inertia property in the smooth ambiguity model and the situation with Bayesian model uncertainty.

Assumption A is a rather weak condition about the Sharpe ratio s and the ambiguous set. We can show that for any ambiguous set $\mathcal{C}$ with more than one element, Assumption A holds for almost all $s \in \mathbb{R}^N_{++}$.[20] The following example shows that it can be verified directly.

Example 6.1. *Let $\mathcal{C} = \mathcal{C}[\underline{a}, \overline{a}]$. Then, Assumption A holds for any $s \in \mathbb{R}^N_{++}$.*

Remark 6.1. *The difference between the smooth ambiguity preference and the multiple prior preference is well studied in the literature. As Epstein and Schneider (2010) documented,*

[20] Given an ambiguous set C with more than one point, then the set of s that Assumption A fails must be included in the set $\bigcup_{\rho_1} \bigcap_{\rho} \{s : s^\top(\rho_1^{-1}\rho\rho_1^{-1} - \rho_1^{-1})s = 0\}$. By a standard continuity argument we can replace this set by a countably union $\bigcup_{\rho_1 \in \mathcal{C}(Q)} \bigcap_{\rho} \{s : s^\top(\rho_1^{-1}\rho\rho_1^{-1} - \rho_1^{-1})s = 0\}$, here $\rho_1 \in \mathcal{C}(Q)$ means that all entries are rational numbers. Therefore, the Lebesgue measure the last set is zero. Hence, almost all s satisfy Assumption A.

"the smooth ambiguity model is more similar to standard expected utility model than the multiple priors model in the sens tea locally risk neutral, portfolio reacts smoothly to changes in return expectation and diversification is beneficial". Gollier (2011) also demonstrates the second-order effect in risk aversion in a portfolio choice setting. Epstein (2010) questioned whether the smooth ambiguity model separates the ambiguity and the attitude towards ambiguity appropriately. In a portfolio choice with correlation ambiguity setting, we show that the smoothness feature of the preference plays an important role in distinguishing the optimal portfolio property such as correlation-invariant and portfolio inertia.

7 Conclusions

In this paper, we solve an optimal portfolio choice problem under correlation ambiguity for any number of risky assets and derive new properties of the optimal portfolio. We show that the optimal portfolio consists of only one risky asset when correlations are sufficiently ambiguous. In general, only part of risky assets enters the optimal portfolio (concentrated portfolio), or the optimal portfolio is correlation-invariant for the ambiguous set because of the ambiguity-aversion effect. Moreover, we demonstrate a general risky asset inertia property of the optimal portfolio under correlation ambiguity. We demonstrate significantly different implications of correlation uncertainty from the expected return or volatility uncertainty.

This paper explains portfolio concentration and portfolio inertia concurrently in one ambiguity-aversion model. We also use the model to explain the growth of indexing and ETFs from an optimal portfolio choice perspective. Furthermore, we demonstrate that both the correlation-invariant feature and the portfolio inertia property are not valid anymore in an alternative smooth ambiguity model of correlation uncertainty. Overall, our results suggest that correlation ambiguity in a multiple-priors framework has important implications for portfolio choice.

Appendix A: Proofs

Proof of Proposition 1:

We first prove the uniqueness property of the optimal portfolio (if there exists) in the max-min problem (4). Assuming there are two portfolios $\phi_1, \phi_2 \in \mathbb{R}^n$ such that $J = \min_{\rho\in\mathcal{C}} f(\phi_i, \rho), i = 1, 2$. Then the set $X_i = \left\{(\mu^\top\phi_i, \phi_i^\top\sigma\rho\sigma\phi_i) : \rho \in \mathcal{C}\right\}$ lies above the line l on the plane $(x, y) \in \mathbb{R}^2$:

$$l : x - \frac{A}{2}y = J,$$

and the set X_i does not entirely lie above this line (X_i meets the line on at one point). Let $\phi_3 = \frac{\phi_1+\phi_2}{2}$ and assume that $\phi_1 \neq \phi_2$, then $(\phi_1 - \phi_2)\top\sigma\rho\sigma(\phi_1 - \phi_2) > 0, \forall\rho \in \mathcal{C}$, implying $\phi_3^\top\sigma\rho\sigma\phi_3 < \frac{\phi_1^\top\sigma\rho\sigma\phi_1+\phi_2^\top\sigma\rho\sigma^\top\phi_2}{2}$. Hence, $f(\phi_3, \rho) = \mu^\top\phi_3 - \frac{A}{2}\phi_3^\top\sigma\rho\sigma\phi_3 > J, \forall\rho \in \mathcal{C}$. Since $\mathcal{C}$ is compact, we have $\min_{\rho\in\mathcal{C}} f(\phi_3, \rho) > J$, which contradicts to the definition of J.

We next determine the optimal value J. Since $f(\phi, \cdot)$ is quasi-concave and $f(\cdot, \rho)$ is quasi-convex, the Sion's minimax theorem implies that

$$J = \max_{\phi}\min_{\rho} f(\phi, \rho) = \min_{\rho}\max_{\phi} f(\phi, \rho) = \frac{1}{2A}\min_{\rho\in\mathcal{C}} s^\top\rho^{-1}s.$$

We now show that the unique optimal portfolio is $\phi_{MV}(\rho^*)$ for $\rho^* = \text{argmin}_{\rho\in\mathcal{C}} s^\top\rho^{-1}s$ and the saddle-point property. For this purpose, we next prove that the set of vectors $\{\phi_{MV}(\rho) : \rho \in \mathcal{C}\}$ is bounded. First, there exists a positive number c such that $det(\rho) \geq c, \forall\rho \in \mathcal{C}$. In fact, if $\mathcal{B}$ is bounded, then $\mathcal{C}$ is bounded. By assumption, $\mathcal{C}$ is closed, then $\mathcal{C}$ is a compact convex subset of $\mathcal{B}$. Since $det : \mathcal{C} \to \mathbb{R}$ is continuous, then the Weierstrass theorem guarantees a global minimal point of the function $det(\cdot)$ on $\mathcal{C}$. Since each $\rho \in \mathcal{C}$ is positive-definite, there exists $c > 0$ such that $det(\rho) \geq c, \forall\rho \in \mathcal{C}$. Next, the inverse matrix ρ^{-1} is $\frac{1}{det(\rho)}\hat{\rho}$, where the (i, j)-th entry of the matrix $\hat{\rho}$ of cofactors is the minor of order $N-1$ obtained by removing the ith row and the jth column multiplied by $(-1)^{i+j}$. Clearly, each

entry of the matrix $\hat{\rho}$ is uniformly bounded for all $\rho \in \mathcal{C}$ since $|\rho_{ij}| \leq 1$. Since $\frac{1}{det(\rho)} \leq \frac{1}{c}$, then the set $\{\rho^{-1} : \rho \in \mathcal{C}\}$ is uniformly bounded. Therefore, the set of admissible vectors, $\{\phi_{MV}(\rho) : \rho \in \mathcal{C}\}$, is bounded. We now choose a convex and bounded set X such that $\phi_{MV}(\rho) \in X \subseteq \mathbb{R}^N, \forall \rho \in \mathcal{C}$. We apply the minimax result in Corollary 37.6.2 of Rockafellar (1970) to the max-min problem $\max_{\phi \in \mathcal{X}} \min_{\rho \in \mathcal{C}} f(\phi, \rho)$, then there exists a saddle point (ϕ^*, ρ^*) such that

$$f(\phi, \rho^*) \leq f(\phi^*, \rho^*) \leq f(\phi^*, \rho), \forall \rho \in \mathcal{C}, \phi \in \mathcal{X}. \tag{A-1}$$

Equation (A-1) implies that $\rho^* = \text{argmin}_{\rho \in \mathcal{C}} s^\top \rho^{-1} s$, $\phi^* = \phi_{MV}(\rho^*)$, and $J = \frac{1}{2A} \min_{\rho \in C} s^\top \rho^{-1} s = \min_{\rho \in C} f(\phi^*, \rho) = f(\phi^*, \rho^*)$. That is, $\phi_{MV}(\rho^*)$ is the unique optimal portfolio in problem (4). Finally, for any $\phi \notin X$, we have $f(\phi, \rho^*) \leq f(\phi_{MV}(\rho^*), \rho^*) = f(\phi^*, \rho^*)$. Hence $(\phi_{MV}(\rho^*), \rho^*)$ satisfies the required saddle-point property. The proof is finished. □

We first prove a general result, Proposition 3, and Proposition 2 follows from Proposition 3 as a special case.

Proof of Proposition 3:

By Proposition 1, the unique optima portfolio is of the form $\phi_{MV}(\rho^*)$ and ρ^* solves the problem (5), the necessary part is proved. For the sufficient part, let $\phi_0 = \phi_{MV}(\rho_0), \rho_0 \in \mathcal{C}$ and assume ϕ_0 is correlation-invariant. Then $f(\phi_0, \rho) = f(\phi_0, \rho_0), \forall \rho \in \mathcal{C}$. Hence, $\min_\rho f(\phi_0, \rho) = f(\phi_0, \rho_0)$. By Proposition 1, we have

$$f(\phi_0, \rho_0) = \min_\rho f(\phi_0, \rho) \leq J = \min_\rho \max_\phi f(\phi, \rho) \leq \max_\phi f(\phi, \rho_0) = f(\phi_0, \rho_0).$$

Therefore, $J = \min_\rho f(\phi_0, \rho)$, hence ϕ_0 is the unique optimal portfolio. Moreover,

$$f(\phi_0, \rho_0) = \min_\rho \max_\phi f(\phi, \rho) \leq \max_\phi f(\phi, \rho), \forall \rho \in \mathcal{C},$$

yielding $s^\top \rho_0^{-1} s \leq s^\top \rho^{-1} s, \forall \rho \in \mathcal{C}$. Hence ρ_0 is one solution of the minimization problem (5).

□

Proof of Proposition 2:

If $|s_1| > \max\{|s_2|, \cdots, |s|_N\}$, and there exists one $\hat{\rho} \in \mathcal{C}$ such that $\hat{\rho}_{1i} = \frac{s_i}{s_1}, i = 2, \cdots, N$. Then $\hat{\rho}(s_1, \cdots, 0)^\top = (s_1, \cdots, s_N)^\top$. Hence, $\phi_{MV}(\hat{\rho}) = \frac{1}{A}\sigma^{-1}\hat{\rho}^{-1}s = (\frac{1}{A}\frac{s_1}{\sigma_1}, 0, \cdots, 0)$ is anti-diversified and thus correlation-invariant. By Proposition 3, the anti-diversified portfolio $\phi_{MV}(\hat{\rho})$ is the unique optimal portfolio with the ambiguous set $\mathcal{C}$. □

Proof of Example 3.5:

By direct calculation, the inverse matrix of $T(a)$ defined in the example (for $2a^2 \neq 1$) is

$$T(a)^{-1} = \frac{1}{1-2a^2}\begin{pmatrix} 1-a^2 & a^2 & -a \\ a^2 & 1-a^2 & -a \\ -a & -a & 1 \end{pmatrix}.$$

Let $\phi = \frac{1}{A}\sigma^{-1}T(a)^{-1}s^{-1}$. Then ϕ is correlation-invariant with respect to $\mathcal{C}$ if and only if $(\phi_1 + \phi_2)\phi_3 = 0$, and it reduces to either $-a(s_1 + s_2) + s_3 = 0$ or

$$\frac{(1-a^2)s_1 + a^2 s_2 - a s_3}{\sigma_1} + \frac{a^2 s_1 + (1-a^2)s_2 - a s_3}{\sigma_2} = 0.$$

If $\frac{s_3}{s_1+s_2} \in [\underline{a}, \overline{a}]$, then the corresponding vector $\phi_{MV}(\frac{s_3}{s_1+s_2})$ is correlation-invariant. Actually, if there exists a root x of the quadratic equation $(\sigma_2 - \sigma_1)(s_2 - s_1)x^2 - (\sigma_1 + \sigma_2)s_3 x + \sigma_1 s_2 + \sigma_2 s_1 = 0$ over the region $[\underline{a}, \overline{a}]$, then $\phi_{MV}(x)$ is the correlation-invariant portfolio. Otherwise, there is no optimal portfolio that is correlation-invariant. □

Proof of Proposition 4:

We first prove that

$$J(i_1, \cdots, i_M; \mathcal{C}) \leq J, \forall (i_1, \cdots, i_M), \forall \{i_1, \cdots, i_M\} \subseteq \{1, \cdots, N\}. \tag{A-2}$$

By using Proposition 1 for the max-min problem (8), there exists an optimal portfolio $\phi \in \mathbb{R}^M$ for $J(1, \cdots, M; \mathcal{C})$. We let $\tilde{\phi} \in \mathbb{R}^N$ such that $\tilde{\phi}_{i_j} = \phi_j, j = 1, \cdots, M$ and other

entry of $\tilde{\phi}$ is zero. It is straightforward to show that

$$\min_{\rho_0 \in \mathcal{C}(i_1,\cdots,i_M)} \mu_0^\top \phi - \frac{A}{2}\phi^\top \sigma_0 \rho_0 \phi = \min_{\rho \in \mathcal{C}} f(\tilde{\phi}, \rho). \tag{A-3}$$

Therefore,

$$J(i_1, \cdots, i_M; \mathcal{C}) = \min_{\rho_0 \in \mathcal{C}(i_1,\cdots,i_M)} \mu_0^\top \phi - \frac{A}{2}\phi^\top \sigma_0 \rho_0 \phi \leq J. \tag{A-4}$$

By the same proof, we show that

$$J(i_1, \cdots, i_M; \mathcal{C}) \leq J(k_1, \cdots, k_K; \mathcal{C}), \forall \{i_1, \cdots, i_M\} \subseteq \{k_1, \cdots, k_K\} \subseteq \{1, \cdots, N\}. \tag{A-5}$$

Therefore, the optimal portfolio holds asset in $\{1, \cdots, M\}$ if and only if $J = J(1, \cdots, M; \mathcal{C})$. To proceed, we make use of the following observation:

Observation: For any $a_\rho \geq 0, b_\rho \geq 0$, $\min_\rho(a_\rho + b_\rho) = \min_\rho a_\rho$ if and only if there exists $\rho^* = \text{argmin}(a_\rho)$ such that $b_{\rho^*} = 0$.

Proof of Observation: If for any $\rho^* = \text{argmin}(a_\rho)$, $b_{\rho^*} > 0$, then $a_{\rho^*} + b_{\rho *} > a_{\rho^*}$. On the other hand, for any $\rho \neq \text{argmin}(a_\rho)$, $a_\rho + b_\rho \geq a_\rho > a_{\rho^*}$. Therefore, $\min_\rho(a_\rho + b_\rho) > \min_\rho a_\rho$.

For any $\rho \in \mathcal{C}$, we write $\Sigma = \Sigma_\rho = \sigma\rho\sigma = \begin{pmatrix} \Sigma_1 & E^\top \\ E & \Sigma_2 \end{pmatrix}$. Here Σ_1 and Σ_2 is of size $M \times M$ and $(N-M) \times (N-M)$, respectively. For any $\phi = \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} \in \mathbb{R}^N, \phi_1 \in \mathbb{R}^M, \phi_2 \in \mathbb{R}^{N-M}$, let

$$\psi = U\phi = \begin{pmatrix} I_M & \Sigma_1^{-1}E^\top \\ 0 & I_{N-M} \end{pmatrix} \times \phi = \begin{pmatrix} \phi_1 + \Sigma_1^{-1}E^\top\phi_2 \\ \phi_2 \end{pmatrix}.$$

Here 0 represents zero vector with suitable size, I_k is the identity matrix of size $k \times k$. We write $\mu^{(1)} = (\mu_1, \cdots, \mu_M)^\top, \mu^{(2)} = (\mu_{M+1}, \cdots, \mu_N)^\top$ and define $s^{(1)}, s^{(2)}$ similarly. Then, we

have

$$\max_{\phi} f(\phi,\rho)=\max_{\psi_1}\left\{(\mu^{(1)})^\top\psi_1-\frac{A}{2}\phi_1^\top\Sigma_1\psi_1\right\}+\max_{\psi_2}\left\{(\mu^{(2)}-\mu^{(1)}\Sigma_1^{-1}E^\top)\psi_2-\frac{A}{2}\psi_2^\top(\Sigma_2-E\Sigma_1^{-1}E^\top)\psi_2\right\},$$

where both optimal values in the right side depend on $\rho\in\mathcal{C}$. By Proposition 1,

$$J=\min_{\rho\in\mathcal{C}}\left\{\max_{\psi_1}(\mu^{(1)})^\top\psi_1-\frac{A}{2}\phi_1^\top\Sigma_1\psi_1+\max_{\psi_2}(\mu^{(2)}-\mu^{(1)}\Sigma_1^{-1}E^\top)\psi_2-\frac{A}{2}\psi_2^\top(\Sigma_2-E\Sigma_1^{-1}E^\top)\psi_2\right\}.$$

By Equation (A-3), and Proposition 1 for the problem $J(1,\cdots,M;\mathcal{C})$, we have

$$J(1,\cdots,M;\mathcal{C})=\min_{\rho\in\mathcal{C}}\left\{\max_{\psi_1}(\mu^{(1)})^\top\psi_1-\frac{A}{2}\phi_1^\top\Sigma_1\psi_1\right\}.$$

Therefore, by the above "observation", $J=J(1,\cdots,M;\mathcal{C})$ if and only if there exists a worst-case correlation matrix ρ_0^* for the max-min problem $J(1,\cdots,M;\mathcal{C})$ such that

$$\max_{\psi_2}(\mu^{(2)}-\mu^{(1)}\Sigma_1^{-1}E^\top)\psi_2-\frac{A}{2}\psi_2^\top(\Sigma_2-E\Sigma_1^{-1}E^\top)\psi_2=0.$$

The last equation equivalents to $\mu^{(2)}-\mu^{(1)}\Sigma_1^{-1}E^\top=0$. Here, $\Sigma_1=\sigma_0\rho_0^*\sigma_0$. The proof is finished. □

Proof of Example 3.6:

For the first part, we first show that $s_3^2<s_1^2+s_2^2$ if and only if there exists $x,y\in(-1,1)$ such that $s_3=s_1x+s_2y$ and $x^2+y^2<1$. Without loss of generality, we assume that $s_2\neq 0$, then it reduces to the existence of x such that $x^2+\left(\frac{s_3-s_1x}{s_2}\right)^2<1$. By examing this quadratic equation, the existence of such a real number x is the same as $(s_3s_1)^2>(s_1^2+s_2^2)(s_3^2-s_2^2)$, that is, $s_3^2<s_1^2+s_2^2$.

We demonstrate the second part by a numerical example. We choose $s_1=0.3, s_2=1, \rho_{12}=0.8, s_3=\sqrt{2.2}$, and $A=1$. Then $J(3)=1.1<J(1,2)=1.15$. We show that by a

contradiction argument there exists no $x, y \in \mathbb{R}$ such that $s_3 = s_1 x + s_2 y$ and the matrix

$$\begin{pmatrix} 1 & \rho_{12} & x \\ \rho_{12} & 1 & y \\ x & y & 1 \end{pmatrix}$$

is positive definite. Therefore, by Proposition 4, asset 3 must enter the optimal portfolio for any ambiguous set $\mathcal{C}$. If not, by plugging $y = \frac{s_3 - s_1 x}{s_2}$ and using the positive-definite condition, there exists x such that

$$1 - \rho_{12}^2 - x^2 - \left(\frac{s_3 - s_1 x}{s_2}\right)^2 + 2\rho_{12} x \frac{s_3 - s_1 x}{s_2} > 0.$$

Equivalently, there exists one real number x such that

$$(s_1^2 + s_2^2 + 2\rho_{12} s_1 s_2)x^2 - 2(s_1 s_3 + \rho_{12} s_2 s_3)x + s_3^2 - (1 - \rho_{12}^2)s_2^2 < 0. \tag{A-6}$$

We can verify that for given parameters, the following inequality holds: $(s_1 s_3 + \rho_{12} s_2 s_3)^2 < (s_1^2 + s_2^2 + 2\rho_{12} s_1 s_2)\,(s_3^2 - (1 - \rho_{12}^2)s_2^2)$. Hence, there is no such a real number x in (A-6), yielding a contradiction. □

The following results on differential topology can be found in Milnor (1997, p. 17 and p. 11 separably).

Lemma 7.1. *(Sard's theorem) Let $f : U \to \mathbb{R}^p$ be a smooth map, with U open in $\mathbb{R}^n$, and C be the set of critical points, then $f(C) \subseteq \mathbb{R}^p$ has measure zero.*

Lemma 7.2. *Let $f : M \to N$ is a smooth map between manifolds of dimension $m \geq n$, and if $y \in N$ is a regular value, then the set $f^{-1}(y) \subseteq M$ is a smooth manifold of dimension $m - n$.*

Proof of Proposition 5:

We consider the differential map $F : X \to F(X)$. By Lemma 7.1, the set of regular

values of F is everywhere dense in $F(X)$. By a regular value we mean an element in the complementary set $F(\mathcal{X}) - F(X_0)$ where $X_0 \subseteq \mathcal{X}$ is the set of $s \in \mathcal{X}$ such that the Jacobian matrix DF_s has rank less than $dim(F(X))$. X_0 is called the critical set. Sard's theorem shows that the Lebesgue measue of $F(X_0)$ is zero. Moreover, by Lemma 7.2, for all regular set $\phi \in \mathcal{F}(\mathcal{X})$, $F^{-1}(\phi)$ is a manifold of dimension $dim(X) - dim(F(X)) \geq 1$.

For all regular value ϕ (which is almost everywhere in $F(X)$, and $s \in F^{-1}(\phi)$, then, $F(s) = \phi$. The optimal strategy is always ϕ when the Sharpe ratios move in the region $F^{-1}(\phi)$. Since all regular values $\phi \in F(X)$ are almost everywhere, and $dim(F^{-1}(\phi)) \geq 1$, the portfolio inertia is generated almost everywhere. □

Proof of Proposition 6:

Its proof is same as Proposition 5 by replacing the smooth map F by the smooth map G. □

Proof of Corollary 4.1:

The restriction of F on X equals to the map:

$$F_1 : (s_1, \cdots, s_M, s_{M+1}, \cdots, s_N) \to (\phi_1, \cdots, \phi_M)^\top = \sigma_0^{-1}\rho_0^{-1}(s_1, \cdots, s_M)^\top \in \mathbb{R}^M.$$

By calculation, the Jacobian matrx of DF_1 is M. Then F_1 is a smooth map with constant rank M. The proposition follows from the constant-rank level set theorem (Lee (2013), Theorem 5.12). □

Proof of Example 4.3:

We first demonstrate the situation with single block. That is, $k = 1$. By Sherman-Morrison formula, we obtain the inverse matrix of $R(a)$ as follows.

$$R(a)^{-1} = \frac{1}{1-a} I_N - \frac{1}{1+(N-1)a} \frac{a}{1-a} (1, \cdots, 1)^\top (1, \cdots, 1)$$

where I_N is the identity matrix. Then the value function is

$$J = \frac{1}{2A} \min_{a\in[\underline{a},\overline{a}]} s^\top R(a)^{-1} s = \frac{1}{2AN} \min_{a\in[\underline{a},\overline{a}]} \frac{N\sum_{n=1}^N s_n^2 - (\sum_{n=1}^N s_n)^2}{1-a} + \frac{(\sum s_n)^2}{1+(N-1)a}.$$

By direct calculation, the worst-case correlation matrix $a^* = \text{argmin}_{a\in[\underline{a},\overline{a}]} \left(s^\top R(a)^{-1} s\right)$ is obtained by

$$\rho^* = \begin{cases} R(\underline{a}), & \text{if } \underline{a} > \frac{1-\Omega(s)}{1+(N-1)\Omega(s)}, \\ R(\overline{a}), & \text{if } \overline{a} < \frac{1-\Omega(s)}{1+(N-1)\Omega(s)}, \\ R\left(\frac{1-\Omega(s)}{1+(N-1)\Omega(s)}\right), & \text{if } \underline{a} \le \frac{1-\Omega(s)}{1+(N-1)\Omega(s)} \le \overline{a}. \end{cases} \tag{A-7}$$

Moreover, when $\rho^* = R\left(\frac{1-\Omega(s)}{1+(N-1)\Omega(s)}\right)$, the corresponding optimal strategy $\phi_{MV}(\rho^*)$ satisfies that $\sum_{i\neq j}(\sigma_i\phi_i)(\sigma_j\sigma_j) = 0$, a correlation-invariant portfolio with respect to $\mathcal{C}$. Moreover, when $\rho^* = R(a)$, or $\rho^* = R(b)$, the corresponding optimal strategy is not correlation-invariant. Therefore, the set $\mathcal{S}$ is characterized by:

$$\mathcal{S} = \left\{ s : \underline{a} \le \frac{1-\Omega(s)}{1+(N-1)\Omega(s)} \le \overline{a} \right\} = \left\{ s : \frac{1-\overline{a}}{1+(N-1)\overline{a}} \le \Omega(s) \le \frac{1-\underline{a}}{1+(N-1)\underline{a}} \right\}.$$

We show that $\mathcal{S}$ has a smooth manifold structure and $dim(\mathcal{S}) = N$. To the end, for each real number t, the intersection of $\mathcal{S}$ with the hyperplane $s_1 + \cdots + s_N = t$ is the same as the intersection of the following set

$$\left\{ \frac{t^2}{N}\left((N-1)(\frac{1-\overline{a}}{1+(N-1)\overline{a}})^2 + 1\right) \le \sum_{n=1}^N s_n^2 \le \frac{t^2}{N}\left((N-1)(\frac{1-\underline{a}}{1+(N-1)\underline{a}})^2 + 1\right) \right\}$$

with the hyperplane $s_1 + \cdots + s_N = t$, being a submanifold of R^N of dimension $N-1$. Therefore, $\mathcal{S}$ is of dimension N.

In general, for $k > 1$, $\rho^{-1} = [R(a_1)^{-1}; \cdots; R(a_k)^{-1}]$. Therefore,

$$s^\top \rho^{-1} s = s_1^\top R(a_1)^{-1} s_1 + \cdots + s_k^\top R(a_k)^{-1} s_k$$

where $s_1, \cdots, s_k$ are the associated sub-vector of the Sharpe ratio vector s. Then, the optimal correlation coefficient in the value function $s^\top \rho^{-1} s$ is given by $[R(a_1^*); \cdots; R(a_k^*)]$ where $a_i^* = \text{argmin}_{\underline{a}_i \le a \le \overline{a}_i}(s_i^\top R(a)^{-1} s_i)$.

By the proof in the first part $s \in \mathcal{S}$ if and only if each subvector $s_1, \cdots, s_k$ satisfies that

$$\frac{1-\overline{a}_i}{1+(N_i-1)\overline{a}_i} \le \Omega(s_i) \le \frac{1-\underline{a}_i}{1+(N_i-1)\underline{a}_i}, i = 1, \cdots, k.$$

Therefore, the dimension of $\mathcal{S}$ equals to $N_1 + \cdots + N_k = N$. The proof for the map $G_{\mathcal{C}}$ is the same and omitted. □

The next result solves a general optimization problem under both expected mean and covariance-variance uncertainty.

PROPOSITION G.1. *Let $g(\phi, \mu, \Sigma) = \mu^\top \phi - \frac{A}{2}\phi^\top \Sigma \phi, \forall \phi \in \mathbb{R}^N$. There exists a solution of the minimization problem*

$$\min_{(\mu,\Sigma)\in\mathcal{U}} \mu^\top \Sigma^{-1} \mu.$$

Let $(\mu^*, \Sigma^*) = \text{argmin}_{(\mu,\Sigma)\in\mathcal{U}} \left(\mu^\top \Sigma^{-1} \mu\right)$, and $\phi^* = \frac{1}{A}\Sigma^{*-1}\mu^*$, then $(\phi^*, \mu^*, \Sigma^*)$ satisfies the saddle-point property

$$g(\phi, \mu^*, \Sigma^*) \le g(\phi^*, \mu^*, \Sigma^*) \le g(\phi^*, \mu, \Sigma), \forall \phi \in \mathbb{R}^N, (\mu, \Sigma) \in \mathcal{U}.$$

Moreover, ϕ^* is a unique optimal solution of the max-min problem

$$\max_{\phi} \min_{(\mu,\Sigma)\in\mathcal{U}} \mu^\top \phi - \frac{A}{2}\phi^\top \Sigma \phi,$$

Moreover, if there exists $(\mu_0, \Sigma_0) \in \mathcal{U}$ such that $g(\phi_0, \mu, \sigma)$ is independent of any $(\mu, \Sigma) \in$]calU, where $\phi_0 = \frac{1}{A}\Sigma_0^{-1}\mu_0$, then ϕ_0 is the optimal portfolio with the ambiguous set $\mathcal{U}$.

Proof:

We notice that $\left\{\frac{1}{A}\Sigma^{-1}\mu : (\mu,\Sigma)\in\mathcal{U}\right\}$ is included in a bounded convex subset $\mathcal{X}\subseteq\mathbb{R}^N$ because of the compactness assumption of $\mathcal{U}$. The proof for the first part is similar to Proposition G.1 and omitted. For the second part, if $g(\phi_0,\mu,\Sigma)$ is independent of $(\mu,\Sigma)\in\mathcal{U}$, then $g(\mu_0,\Sigma_0)=\min_{(\mu,\Sigma)\in\mathcal{U}} g(\phi_0,\mu,\Sigma)$. Then, by the saddle-point property in the first part, we have

$$g(\phi_0,\mu_0,\Sigma_0) \;\leq\; \max_{\phi}\min_{(\mu,\Sigma)} g(\phi,\mu,\Sigma) = \min_{(\mu,\Sigma)}\max_{\phi} g(\phi,\mu,\Sigma) \leq \max_{\phi} g(\phi,\mu_0,\Sigma_0) = g(\phi_0,\mu_0,\Sigma_0).$$

This implies $J=g(\phi_0,\mu_0,\Sigma_0)=\min_{(\mu,\Sigma)\in\mathcal{U}} g(\phi_0,\mu,\Sigma)$. So ϕ_0 is the unique optimal portfolio uncer the ambiguous set $\mathcal{U}$. □

Proof of Example 5.1:

Assume first $\frac{1-\Omega(s)}{1+(N-1)\Omega(s)} < a_0$, then for sufficiently small position number *epsilon*, our argument in Example 3.4 and Example 4.3 show that the optimal portfolio $\phi(\epsilon)=\frac{1}{A}\sigma^{-1}R(a_0-\epsilon)^{-1}s$. It is straightforward to show that

$$\frac{\partial\phi(\epsilon)_i}{\partial\epsilon} = \frac{1}{A\sigma_i(1-a_0)^2}\left(\frac{S(1+(N-1)a_0^2)}{(1+(N-1)a_0)^2} - s_i\right), \tag{A-8}$$

where $S=s_1+\cdots+s_N$. It is impossible that all first-order derivative $\frac{\partial\phi(\epsilon)_i}{\partial\epsilon}=0,\forall i$. Hence, there exists at least i such that $\frac{\partial\phi(\epsilon)_i}{\partial\epsilon}\neq 0$; thus, there is first-order effect in risk aversion. Moreover, for the expected utility under correlation ambiguity,

$$\frac{J(\epsilon)}{\partial\epsilon} = \frac{1}{2A}\frac{S^2}{(1-a_0)^2(1+(N-1)a_0)^2}\left((1+(N-1)a_0^2) - \frac{s_1^2+\cdots+s_N^2}{S^2}(1+(N-1)a_0)^2\right)\neq 0,$$

yielding the first-order effect to the expected utility.

Second, if $\frac{1-\Omega(s)}{1+(N-1)\Omega(s)} > a_0$, then for sufficiently small number ϵ, the optimal portfolio is $\phi(\epsilon)=\frac{1}{A}\sigma^{-1}R(a_0+\epsilon)^{-1}s$. By the same proof as above, there exists first-order effect in risk aversion and the expected utility.

Third, if $a_0=\frac{1-\Omega(s)}{1+(N-1)\Omega(s)}$, then for any small positive number ϵ, by Example 3.4, the

optimal portfolio is the same (portfolio inertia). Therefore, $\frac{\partial \phi(\epsilon)}{\partial \epsilon} = 0$. □

The proof for Example 5.2:

By Proposition 1, the optimal portfolio $\phi(\epsilon) = \frac{1}{A}\sigma^{-1}\rho^{*-1}s$, where $\rho^* = argmin_{a\in[a_0-\epsilon,a_0+\epsilon]} s^\top T(a)s$. TIf a_0 is not the locally minimal point of the function $s^\top T(a)s$, then for small ϵ, the optimal portfolio is either $frac1A\sigma^{-1}\rho(a_0+\epsilon)-1s$ or $\frac{1}{A}\sigma^{-1}\rho(a_0-\epsilon)-1s$. In either case, we see that $\frac{\phi(\epsilon)}{\partial\epsilon} \neq 0$ by the expression of $T(a)^{-1}$ in Example 3.4. □

Proof of Proposition 7:

The proof is divided into five steps.

Step 1. We demonstrate the existence and uniqueness optimal portfolio and characterize the optimal portfolio.

The objective function in the optimal portfolio choice problem is denoted by $g(\phi) = \mathbb{E}_{\mathcal{P}}\left[h(\mathbb{E}_\rho[u(W)])\right]$, W is the portfolio wealth by using the portfolio investment vector ϕ. Then $g_i(\phi) \equiv \frac{\partial g(\phi)}{\partial \phi_i} = \mathbb{E}_{\mathcal{P}}\left[h'\left(\mathbb{E}_\rho[u(W)]\right)\mathbb{E}_\rho[u'(W)(r_i - r_f)]\right]$, and

$$\begin{aligned} g_{ij}(\phi) \equiv \frac{\partial^2 g(\phi)}{\partial\phi_i\partial\phi_j} &= \mathbb{E}_{\mathcal{P}}\left[h''\left(\mathbb{E}_\rho[u(W)]\right)\mathbb{E}_\rho[u'(W)(r_i-r_f)]\mathbb{E}_\rho[u'(W)(r_j-r_f)]\right] \\ &+\mathbb{E}_{\mathcal{P}}\left[h'\left(\mathbb{E}_\rho[u(W)]\right)\mathbb{E}_\rho[u''(W)(r_i-r_f)(r_j-r_f)]\right]. \end{aligned}$$

For any real numbers $\zeta_1, \cdots, \zeta_N \in \mathbb{R}$, we have

$$\begin{aligned} \sum_{i,j=1}^N g_{ij}\zeta_i\zeta_j &= \mathbb{E}_{\mathcal{P}}\left[h''\left(\mathbb{E}_\rho[u(W)]\right)\left(\sum_{i=1}^N \zeta_i\mathbb{E}_\rho[u'(W)(r_i-r_f)]\right)^2\right] \\ &+\mathbb{E}_{\mathcal{P}}\left[h'\left(\mathbb{E}_\rho[u(W)]\right)\mathbb{E}_\rho\left[u''(W)\left(\sum_i \zeta_i(r_i-r_f)\right)^2\right]\right]. \end{aligned}$$

Since both $u(\cdot)$ and $h(\cdot)$ are strictly concave, the Hessian matrix of the function g is negative-definite and the function $g(\cdot)$ is concave. There exists an optimal portfolio in the problem (14). The result also holds when $h(x)$ is linear and $h'(x) < 0$.

For the uniqueness property of the optimal portfolio, we need the strictly concave or

strictly quasi-concave property of the objective function $g(\phi)$. Given the specification of the function $h(x)$, $g(\phi)$ is $\mathbb{E}_{\mathcal{P}}[k(f(\phi,\rho)]$, where $k(x) = h(u(x)) = -\frac{1}{1+\alpha}e^{-A(1+\alpha)x}$. Notice that $k(x)$ is strictly increasing and strictly concave. Moreover, $f(\phi,\rho) = \mu^\top\phi - \frac{A}{2}\phi^\top(\sigma\rho\sigma)\phi$ is strictly concave with respect to ϕ. For simplicity, we write $f_i(\phi) = \frac{\partial f}{\partial \phi_i}$, $f_{ij}(\phi) = \frac{\partial^2 f}{\partial\phi_i\partial\phi_j}$, then $g_i(\phi) = \mathbb{E}_{\mathcal{P}}[k'(f(\phi,\rho))f_i(\phi)]$. For any non-zero vector $\zeta \in \mathbb{R}^N$, we have

$$\sum_{i=1,j=1}^{N} g_{ij}\zeta_i\zeta_j = \mathbb{E}_{\mathcal{P}}\left[k''(f)(\sum_i \zeta_i f_i)^2\right] + \mathbb{E}_{\mathcal{P}}\left[k'(f)\sum_{i=1,j=1}^{N}\zeta_i\zeta_j f_{ij}\right] < 0.$$

We have thus proved the strictly concave property of the function $g(\phi)$, yielding the existence and the uniqueness of the optimal portfolio ϕ^*. Moreover, ϕ^* is uniquely solved by the following N equations:

$$\mathbb{E}_{\mathcal{P}}\left[k'(f(\phi,\rho))(\mu_i - A\sum_{j=1}^{N}\sigma_i\sigma_j\rho_{ij}\phi_j)\right] = 0, i = 1,\cdots,N. \tag{A-9}$$

Step 2. We show that the optimal portfolio is not correlation-invariant by a contradiction argument.

Assuming not, $f(\phi^*,\rho)$ is independent of the correlation matrix ρ, then Equation (A-9) implies that $s = A\sigma\mathbb{E}_{\mathcal{P}}[\rho]\phi^*, \phi^* = \frac{1}{A}\sigma^{-1}\rho_0^{-1}s$. Here, $\rho_0 = \mathbb{E}_{\mathcal{P}}[\rho] = (\mathbb{E}_{\mathcal{P}}[\rho_{ij}])$ is the expected correlation matrix under the prior distribution. Since $\mathbb{E}[\rho]$, as an expectation of the variable ρ under measure $\mathcal{P}$, is an element of the closure of the convex hull of the set $\mathcal{C}$, and $\mathcal{C}$ is convex and compact (and thus complete), we obtain $\rho_0 \in \mathcal{C}$.

If the vector $\frac{1}{A}\sigma^{-1}\rho_0^{-1}s$ is correlation-invariant, then $f(\frac{1}{A}\sigma^{-1}\rho_0^{-1}s,\rho) = f(\frac{1}{A}\sigma^{-1}\rho_0^{-1}s,\rho_0), \forall\rho \in \mathcal{C}$. Therefore, we have $s^\top(\rho_0^{-1}\rho\rho_0^{-1})s = s^\top\rho_0^{-1}s, \forall\rho\in\mathcal{C}$. However, Assumption A implies that the map $\rho \in \mathcal{C} \to s^\top(\rho_0^{-1}\rho\rho_0^{-1})s$ is not a constant map. This contradiction shows that the optimal portfolio ϕ^* is not correlation-invariant.

Step 3. We show that the map $s \to \phi^*$ is a smooth map.

Define $Z = \left\{\phi \in \mathbb{R}^N : \phi \text{ satisfies Equation (A-9)}\right\}$. We characterize the optimal portfolio

$\phi^* \in Z$ in terms of smooth functions of s. Let

$$h_i(s,\phi) = \mu_i - A\sum_{j=1}^{N}\sigma_i\sigma_i\sigma_j\rho_{ij}\phi_j = \sigma_i(s_i - A\sum_{j=1}^{N}\sigma_j\rho_{ij}\phi_j), i = 1,\cdots,N \tag{A-10}$$

and define $G : (s,\phi) \to (G_1,\cdots,G_N) \in \mathbb{R}^N$, where

$$G_i(s,\phi) = \mathbb{E}_{\mathcal{P}}\left[k'(f(\phi,\rho))h_i(s,\phi)\right], i = 1,\cdots,N.$$

We demonstrate the rank of the matrix $\left(\frac{\partial G_i}{\partial \phi_j}\right), i,j = 1,\cdots,N$ is N. By calculation,

$$\begin{aligned} a_{ij} &\equiv \frac{\partial G_i}{\partial \phi_j} \\ &= \mathbb{E}_{\mathcal{P}}\left[k''(f(\phi,\rho))h_j h_i\right] + \mathbb{E}_{\mathcal{P}}\left[k'(f(\phi,\rho))(-A\sigma_i\sigma_j\rho_{ij})\right] \\ &= -A\mathbb{E}_{\mathcal{P}}\left[k'(f(\phi,\rho)\left(h_i h_j(1+\alpha) + \sigma_i\sigma_j\rho_{ij}\right)\right] \end{aligned}$$

here we use the fact that $k''(x) = -A(1+\alpha)k'(x)$ in the last equation. If the rank of the matrix (a_{ij}) is less than N, then there exists no-zero vector $\zeta = (\zeta_1,\cdots,\zeta_N)^\top$ such that $\zeta^\top(a_{ij})\zeta = 0$. That is

$$\mathbb{E}_{\mathcal{P}}\left[k'(f(\phi,\rho))\left((1+\alpha)\sum_{i,j}h_i h_j\zeta_i\zeta_j + \sum_{i,j=1}^{N}\sigma_i\sigma_j\rho_{ij}\zeta_i\zeta_j\right)\right] = 0. \tag{A-11}$$

However, $\sum_{i,j} g_i g_j\zeta_i\zeta_j \geq 0$ implies $\mathbb{E}_{\mathcal{P}}\left[k'(f(\phi,\rho))\sum_{i,j}h_i h_j\zeta_i\zeta_j\right] \geq 0$, and $\sum_{i,j=1}^{N}\sigma_i\sigma_j\rho_{ij}\zeta_i\zeta_j > 0, \forall\rho \in \mathcal{C}$ implies $\mathbb{E}_{\mathcal{P}}\left[k'(f(\phi,\rho))\left(\sum_{i,j=1}^{N}\sigma_i\sigma_j\rho_{ij}\zeta_i\zeta_j\right)\right] > 0$. Therefore, Equation (A-11) is impossible. Then, the rank of the matrix (a_{ij}) is N. By the implicit function theorem (Lee (2013) Theorem C. 40), there is a well-defined unique map $H : s \to H(s) = \phi^* \in \mathbb{R}^N$. Moreover, $H(s)$ is smooth.

Step 4. We show that H is locally one-to-one, thus H^{-1} is defined locally and smoothly.

For this purpose, we show the full rank of the matrix $(\frac{\partial G_i}{\partial s_j})$ at the point (s,ϕ) satisfying

the Equation (A-9). Notice that $\frac{\partial f}{\partial s_j} = \sigma_j \phi_j$ by writing $\mu_j = s_j \sigma_j$. Then,

$$\frac{\partial G_i}{\partial s_j} = \begin{cases} \mathbb{E}_{\mathcal{P}}[k''(f)\sigma_j \phi_j h_i], & j \neq i, \\ \mathbb{E}_{\mathcal{P}}[k''(f)\sigma_i \phi_i h_i] + \mathbb{E}_{\mathcal{P}}[k'(f)\sigma_i], & j = i. \end{cases}$$

Since $k''(x) = -A(1+\alpha)k'(x)$, we can write the $N \times N$ matrix $(\frac{\partial G_i}{\partial s_j})$ as follows.

$$(\frac{\partial G_i}{\partial s_j}) = diag(\mathbb{E}_{\mathcal{P}}[k'(f)]\sigma_i) - A(1+\alpha)(\mathbb{E}_{\mathcal{P}}[k''(f)\sigma_j \phi_j h_i]).$$

Here the first term on the right side is a diagonal matrix with i'th component $\mathbb{E}_{\mathcal{P}}[k'(f)]\sigma_i], i = 1, \cdots, N$, and the second term can be written as $UV^\top, U = (\mathbb{E}_{\mathcal{P}}[k'(f)h_1], \cdots, \mathbb{E}_{\mathcal{P}}[k'(f)h_N])^\top, V = (-A(1+\alpha)\sigma_1 \phi_1, \cdots, -A(1+\alpha)\sigma_N \phi_N)^\top \in \mathbb{R}^N$. By using Sherman-Morrison formula, to show the matrix $(\frac{\partial G_i}{\partial s_j})$ is invertible, it suffices to show that

$$1 + V^\top diag(\frac{1}{\mathbb{E}_{\mathcal{P}}[k'(f)]\sigma_i])U \neq 0.$$

To the end, we notice that $\mathbb{E}_{\mathcal{P}}[k'(f)h_i\phi_i] = \mathbb{E}_{\mathcal{P}}[k'(f)h_i]\phi_i = 0$ by the Equation (A-9), then

$$1 + V^\top diag(\frac{1}{\mathbb{E}_{\mathcal{P}}[k'(f)]\sigma_i])U = 1 - A(1+\alpha)\sum_i \frac{\mathbb{E}_{\mathcal{P}}[k'(f)h_i\phi_i]}{\mathbb{E}_{\mathcal{P}}[k'(f)]} = 1 \neq 0.$$

Since the rank of the matrix $(\frac{\partial G_i}{\partial s_j})$ is N, by using the implicit function theorem again (Lee (2013) Theorem C. 40), we can write $s = K(\phi)$ for a well defined map K to solves Equation (A-9) in a small region of ϕ. Therefore, $K = H^{-1}$ in a small region of s, by standard Calculus argument.

Step 5. For any s, we show that the corresponding optimal portfolio ϕ^* does not satisfy the inertia property.

By Step 3 - Step 4, Z is a smooth manifold with dimension N. Therefore, for any s there exists an open region U_s, U_s is a smooth manifold of dimension N, and $H : s \to \phi^*$ is

a one-to-one smooth map. It means that $H(s)$ changes accordingly if s changes in an open region U_s. Hence, there is no portfolio inertia property. The proof is completed. □

Proof of Example 6.1:

For any $s \in \mathbb{R}^N$ such that $s^\top R(a)^{-1}s$ is a constant c, we show that each $s_i = 0, i = 1, \cdots, N$. By a calculation in Corollary 5.1, we have

$$\frac{N\sum s_n^2 - (\sum_n s_n)^2}{1-a} + \frac{(\sum s_n)^2}{1+(N-1)a} = c, \forall \underline{a} \leq a \leq \overline{a}.$$

Then $(N\sum s_n^2 - (\sum_n s_n)^2)(1+(N-1)a) + \sum_n s_n^2(1-a) = c(1-a)(1+(N-1)a), \forall \underline{a} \leq a \leq \overline{a}$. By comparing the coefficients on both side, we derive $c = 0$, $\sum_n s_n = 0$, and $N\sum s_n^2 - (\sum_n s_n)^2 = 0$, yielding $\sum s_n^2 = 0$. Therefore, $s_i = 0, \forall i = 1, \cdots, N$. □

References

Abel, A., J. Eberly., and S. Panageas. 2013. Optimal Inattention to the Stock Market with Information Costs and Transaction Costs. *Econometrica* 81 (4), 1455-1481.

Agnew, J., P. Balduzzi., and A. Sunden. 2003. Portfolio Choice and Trading in a Large 401(k) Plan. *American Economic Review* 93 (1), 193-215.

Anderson, E., PL. Hansen., and T. Sargent. 2003. A Quartet of Semigroups for Model Specification, Robustness, Prices of Risk and Model Detection. *Journal of the European Economic Association* 1, 68-123.

Ang, A., G. Bekaert., and J. Liu. 2005. Why Stocks may Disappoint. *Journal of Financial Economics* 76, 471-508.

Bianchi, M., and J.-M. Tallon. 2019. Ambiguity Preferences and Portfolio Evidence from the Field. *Management Science* 65, 1486-1501.

Bilias, Y., D. Georgarakos., and M. Haliassos. 2007. Portfolio Inertia and Stock Market Fluctuations. *Journal of Money, Credit and Banking* 42 (4), 715-742.

Bogle, J. C. 2016. The Index Mutual Fund: 40 Years of Growth, Change and Challenge. *Financial Analyst Journal*, January/February, 9-13.

Bond, P., D. Garcia. 2020. The Equilibrium Consequence of Indexing. Working paper, University of Washington.

Bossaerts, P., P. Ghirardato, S. Guarnaschelli, and W. Zame. 2010. Ambiguity in Asset Markets: Theory and Experiment. *Review of Financial Studies* 23, 1325-1359.

Boyle, P., L. Garlappi., R. Uppal, R., and T. Wang. 2012. Keynes meets Markowitz: The Trade-Off Between Familiarity and Diversification. *Management Science* 58 (2), 253-272.

Brunnermeier, M., S. Nagel. 2008. Do Wealth Fluctuations Generate Time-Varying Risk Aversion? Micro-evidence on Individual. *American Economic Review* 98 (3), 713-36.

Calvet, L.E., Y. Y. Campbell., and P. Sodini. 2008. Down or Out: Assessing the Welfare Costs of Household Investment Mistakes. *Journal of Political Economy* 115, 707-747.

Campbell, J.Y., 2006. Household Finance. *Journal of Finance* 61(4), 1553-1604.

Carroll, G. 2017. Robustness and Separation in Multidimensional Screening. *Econometrica* 85, 453-488.

Cao, H., T. Wang., and H. Zhang. 2005. Model Uncertainty, Limited Market Participation, and Asset Prices. *Review of Financial Studies* 18, 1219-1251.

Cao, H., H. Han., D. Hirsheleifer., and H. Zhang. 2011. Fear of the Unknown: Familiarity and Economic Decisions. *Review of Finance* 15, 173-206.

Caskey, J. 2009. Information in Equity Markets with Ambiguity-Averse Investors. *Review of Financial Studies* 22, 3595-3627.

Chan, K.C., J. Karceski., and J. Lakonishok. 1999. On Portfolio Optimization: Forecasting Covariances and Choosing the Risk Model. *Review of Financial Studies* 12, 937-974.

Condie, S., J. Ganhuli., and P. K.Illeditsch. 2021. Information Inertia. *Journal of Finance* 76, 443-479.

DeMiguel, V., L. Garlappi, R. Uppal, 2007. Optimal versus Naive Diversification: How Inefficient is the 1/N Portfolio Strategy? *Review of Financial Studies* 22 (5), 1915-1953.

Dimmock, S. G., R. Kouwenberg., O. S. Mitchell, and K. Peijnenburg. 2016. Ambiguity Aversion and Household Portfolio Choice Puzzles: Empirical Evidence, *Journal of Financial Economics* 119, 559-577.

Dow, J., and S. Werlang. 1992. Uncertainty Aversion, Risk Aversion, and the Optimal Portfolio of Choice. *Econometrica* 60 (1), 197-204.

Easley, D., and M. O'Hara. 2009. Ambiguity and Nonparticipation : The Role of Regulation. *Review of Finance Studies* 22 (5), 1817-1846.

Easley, D., and M. O'Hara. 2010. Liquidity and Valuation in an Uncertain World. *Journal of Financial Economics* 97, 1-11.

Easley, D., M. O'Hara., and L. Yang. 2015. Opaque Trading, Disclosure and Asset Prices: Implications for Hedge Fund Regulations. *Review of Financial Studies* 27, 1190-1237.

Ellis, A., and M. Piccione. 2017. Correlation Misperception in Choice. *American Economic Review* 107, 1264-1292.

Ellsberg, D. 1961. Risk, Ambiguity, and the Savage Axioms. *Quarterly Journal of Economics* 75, 643-669.

Engle, R., and B. Kelly. 2012. Dynamic Equicorrelation. *Journal of Business and Economic Statistic* 30 (2), 212-228.

Epstein, L., and Y. Halevy. 2019. Ambiguous Correlation. *Review of Economic Studies* 86 (2), 668-693.

Epstein, L., and S. Ji. 2013. Ambiguity Volatility and Asset Pricing in Continuous Time. *Review of Financial Studies* 26 (7), 1740-1786.

Epstein, L., and M. Schneider. 2008. Ambiguity, Information Quality, and Asset Pricing. *Journal of Finance* 63, 197-228.

Epstein, L., and M. Schneider. 2010. Ambiguity and Asset Markets. *Annual Review of Financial Economics* 2, 315-346.

Epstein, L. 2010. A Paradox for the "Smooth Ambiguity" Model of Preference. *Econometrica* 78, 2085-2099.

Gajdos, T., T. Hayashi., J.-M. Tallon., and J.-C. Vergnaud. 2008. Attitude Toward Imprecise Information. *Journal of Economic Theory* 140, 27-65.

Garlappi, L., R. Uppal., and T. Wang. 2007. Portfolio Selection with Parameter and Model Uncertainty: A Multi-Prior Approach. *Review of Financial Studies* 20 (1), 41-81.

Ghirardato, P., F. Maccheroni., and M. Marinacci. 2004. Differentiating Ambiguity and Ambiguity Attitude. *Journal of Economic Theory* 118, 133-173.

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

Goetzmann, W., and A. Kumar. 2008. Equity Portfolio Diversification. *Review of Finance* 12, 433-463.

Goldman, M. 1979. Antidiversification or Optimal Programmes for Infrequently Revised Portfolios. *Journal of Finance* 34 (2), 505-516.

Gollier, C. 2011. Portfolio Choices and Ssset Prices: The Comparative Statics of Ambiguity Aversion. *Review of Economic Studies* 78, 1329-1344.

Guidolin, M., and Liu, H. L., 2016. Ambiguity Aversion and Under-Diversification. *Journal of Financial and Quantitative Analysis* 51 (4), 1297-1323.

Hirshleifer, D., C. Huang., and S. H. Tech. 2019. Index Investing and Asset Pricing under Information Asymmetry and Ambiguity Aversion. Available at SSRN-ID2898992.

Huang, H., S. Zhang., and W. Zhu. 2017. Limited Participation under Ambiguity of Correlation. *Journal of Financial Markets* 32, 97-143.

Ivković, Z., S. Sialm., and S. Weisbenner. 2008. Portfolio Concentration and the Performance of Individual Investors. *Journal of Financial and Quantitative Analysis* 43, 613-656.

Jagannathan, R., and T. Ma. 2003. Risk Reduction in Large Portfolios: Why Imposing the Wrong Constraints Helps. *Journal of Finance* 58, 1651-1684.

Kan, R. and G. Zhou, 2007. Optimal Portfolio Choice with Parameter Uncertainty. *Journal of Financial and Quantitative Analysis* 42 (3), 621-656.

Kacperczyk, M., C. Sialm., and L. Zheng. 2005. On the Industry Concentration of Actively Managed Equity Mutual Funds. *Journal of Finance* 60, 1983-2011.

Kim, S. J. and Boyd, S., 2008. A Minimax Theorem with Applications to Machine Learning, Signal Processing, and Finance. *SIAM Journal of Optimization* 19, No. 3, 1344-1367.

Klibanoff, P., M. Marinacci. and S. Mukerji. 2005. A Smooth Model of Decision Making under Ambiguity, *Econometrica* 73, 1849-1892.

Lee, John. 2013. *Introduction to Smooth Manifolds.* Second Edition. Graduate Texts in Mathematics, Vol 218, Springer.

Levy, G., and R. Razin. 2015. Correlation Neglect, Voting Behavior, and Information Aggregation. *American Economic Review* 105 (4), 1634-1645.

Levy, G., and R. Razin. 2020. Combining Forecasts in the Presence of Ambiguity over Correlation Structures. *Journal of Economic Theory*, 105075.

Liu, H. 2004. Optimal Consumption and Investment with Transaction Costs and Multiple Risky Assets. *Journal of Finance* 59, 289-338.

Illeditsch, P. K. 2011. Ambiguous Information, Portfolio Inertia, and Excess Volatility. *Journal of Finance* 66, 2213-2247.

Maccheroni, F., M. Marinacci., and D. Ruffino. 2013. Alpha as Ambiguity: Robust Mean-Variance Portfolio Analysis. *Econometrica* 81 (3), 1075-1113.

Markowitz, H. M. 1952. Portfolio Selection. *Journal of Finance* 7, 77-91.

Merton, R. 1980. On Estimating the Expected Return on the Market: An Exploratory Investigation. *Journal of Financial Economics* 8(4), 323-361.

Mihm, M. 2016. Reference Dependent Ambiguity. *Journal of Economic Theory* 163, 495-524.

Milnor, John W., 1997. *Topology from the Differential Viewpoint,* (Revised Edition), Princeton University Press.

Mitton, T., and K. Vorkink. 2007. Equilibrium Underdiversification and the Preference for Skewness. *Review of Financial Studies* 20, 1255-1288.

Murihead, R. 1982. *Aspects of Multivariate Statistical Theory.* Wiley.

Mukerji, S., and Jean-Marc Tallon. 2001. Ambiguity Aversion and Incompleteness of Financial Markets. *Review of Economic Studies* 68, 883-904.

Mukerji, S., and Jean-Marc Tallon. 2003. Ellsberg's Two-Color Experiment, Portfolio Inertia and Ambiguity. *Journal of Mathematical Economics* 39, 299-315.

Polkovnichenko, V., 2005. Household Portfolio Diversification: a Case for Rank-Dependent Preference. *Review of Financial Studies* 18 (4), 1467-1502.

Roche, R., S. Tompaidis., and C. Yang. 2013. Why Does Junior Put All His Eggs in One Basket? A Potential Rational Explanation for Holding Concentrated Portfolios. *Journal of Financial Economics* 109, 775-796.

Rockafellar, R. T. 1970. *Convex Analysis.* Princeton University Press, Princeton, New Jersey, 1970.

Samuelson, P. A. 1967. General Proof that Diversification Pays. *Journal of Financial and Quantitative Analysis* 2, 1-13.

Segal, U., and A. Spival. 1990. First Order versus Second Order Risk Aversion. *Journal pf Economic Theory* 51, 111-125.

Seo, K. 2009. Ambiguity and Second-Order belief. *Econometrica* 77, 1575-605.

Serre, D. 2002. *Matrices: Theory and Applications*. Graduate Texts in Mathematics, 216, Springer.

Strzalecki, T. 2011. Axiomatic Foundations of Multiplier Preferences. *Econometrica* 79, 47-73.

Uppal, R., and T. Wang. 2003. Model Misspecification and Under-Diversification. *Journal of Finance* 58, No. 6, 2465-2486.

Van Nieuwerburgh, S. V., and L. Veldkamp. 2010. Information Acquisition and Portfolio Under-Diversification. *Review of Economic Studies* 77, 779-805.