Beta Ambiguity and Security Return Characteristics

Zhe Geng and Tan Wang *

Abstract

We develop a model to study the cross-sectional properties of asset returns in the presence of ambiguity in the distribution of asset returns. In our model, the cross-sectional expected returns can be described by a three-factor model, capturing risk, mean ambiguity and variance-covariance ambiguity, respectively. Expected returns include a mean ambiguity premium, a variance-covariance ambiguity premium, as well as the standard risk premium. The expected returns exhibit cross-sectional characteristics consistent with the empirical fact that the overall beta-return relation and IVOL-return relation are both negative, but the beta-return relation is negative and stronger among overpriced stocks while positive and weaker among underpriced stocks, and the IVOL-return relation is negative and stronger among overpriced stocks but positive and weaker among underpriced stocks (Black, Jensen, and Scholes (1972), Ang, Hodrick, Xing, and Zhang (2006), Liu, Stambaugh, and Yuan (2016), and Stambaugh, Yu, and Yuan (2015)).

*Zhe Geng and Tan Wang are with SAIF, Shanghai Jiao Tong University. We thank Yu Yuan for valuable comments.

1 Introduction

In the classical CAPM of Sharpe (1964) and Lintner (1965) theory, stocks with higher betas earn higher premia than stocks with lower betas. Empirical studies document, however, that security market line is too flat relative to the one predicted by the CAPM theory (Black, Jensen, and Scholes (1972) and Fama and MacBeth (1973)). The empirical result has since been referred to as the beta anomaly. Several explanations are provided in the literature. Common to most of the explanations is the assumption of borrowing constraints, be it in the form of constraint on riskless borrowing, margin constraint or constraint on short-sale of stocks (Black (1972), Frazzini and Pedersen (2014), Hong and Sraer (2016), and Liu, Stambaugh, and Yuan (2016)). Recently, Liu, Stambaugh, and Yuan (2016) offer additional evidence on the anomaly. They show that while the security market line for over-priced stocks is flatter than that predicted by the standard CAPM theory as documented in the literature, the security market line for under-priced stocks is not flatter if not steeper, pointing to the dissatisfaction of some of the existing explanations. They also show that the beta anomaly is closely related to the idiosyncratic volatility anomaly (Ang, Hodrick, Xing, and Zhang (2006)). They argue that the beta anomaly is a consequence of the idiosyncratic volatility, as there is strong positive IVOL-beta correlation. The explanation, based on Liu, Stambaugh, and Yuan (2016), relies also on short-sale constraints.

In this paper, we provide an alternative understanding of the beta anomaly that does not require borrowing or short-sale constraint. The financial markets in our model are without any frictions. The key ingredient of our explanation of the beta anomaly is that investors do not have perfect knowledge of the probability distribution of stock returns. As a consequence, investors ask for a premium as the compensation for that ambiguity. That premium is the alpha, or mispricing, relative to the CAPM. Under natural assumptions, the premium exhibits a pattern that is consistent with the beta anomaly.

In our model, agents are homogeneous and are fully aware that there is ambiguity about the probability law of stock returns and the data can only provide a reference distribution. Due to their aversion to ambiguity, they adjust their portfolio computed according to the reference distribution to account for the ambiguity. The adjustment leads to equilibrium returns that deviate from those computed according to the reference distribution. Interestingly, however, that deviation can always be tracked by two tracking portfolios, one for the ambiguity in the expected returns of the stocks and the other for the ambiguity in the covariances of the returns of the stocks. In a way, these two portfolios can be viewed as two factors tracking two sources of ambiguity. As such the premia on those two tracking factors are interpreted as the premia for the two sources of ambiguity. It should, however, be emphasized that those two factors are not factors in the traditional sense. They do not track some fundamental macro or aggregate risks as for example the market portfolio does. They capture the systematic ambiguity for the pricing of stocks.

When agents are ambiguity averse, the two factors earn positive ambiguity premia. Stocks that have higher loading on those factors earn higher premia, while stocks that have lower or negative loading on those factors earn lower or even negative premia. As there is no good reason to believe that the reference distribution is related to the level of ambiguity, beta calculated according to the reference distribution is unrelated to the level of ambiguity. As a consequence, if the stocks are double-sorted on mis-pricing and beta, the alphas can and in fact are likely to exhibit the pattern as shown in Liu, Stambaugh, and Yuan (2016). At the same time, as agents are ambiguity averse, stocks on average earn positive ambiguity premia. As the total premium of a stock has two components, the risk premium and the ambiguity premium, when returns are regressed on risk factor only, the resulting regression line has a slope to flat relative to CAPM.

Through simulation we show that our model produces qualitatively similar patterns of alphas as shown in the literature. The security market line is flatter than predicted by CAPM. The overall IVOL-return relation is negative. However, the beta-return relation is negative and stronger among overpriced stocks while positive and weaker among underpriced stocks, and the IVOL-return relation is negative and stronger among overpriced stocks, but positive and weaker among underpriced stocks.

Our paper is related to two branches of the literature. One branch is that on ambiguity and its implications for asset prices. To model ambiguity averse agents, we follow the multiple-prior approach of Gilboa and Schmeidler (1989). The dynamic version of it is proposed by Epstein and Schneider (2003). In the study of asset pricing implications of ambiguity, similar approach has been taken by Dow and Werlang (1992), Epstein and Wang (1994, 1995), Chen and Epstein (2002) and Epstein and Miao (2003), Kogan and Wang (2003), Easley and O'Hara (2009, 2010), among many others. An alternative approach to modeling ambiguity averse agents is introduced by Hansen and Sargent (2001) and Anderson, Hansen, and Sargent (2003). Maccheroni, Marinacci, and Rustichini (2006a) subsequently axiomatize the robustness preferences and prove it as sub-class of variational preferences. Maccheroni, Marinacci, and Rustichini (2006b) introduce dynamic variational preferences. That approach is taken by Uppal and Wang (2003), Maenhout (2004, 2006), Liu, Pan, and Wang (2005), among others, in their study of portfolio choice, portfolio diversification and option pricing problems. The third, smooth ambiguity preference, approach to modeling ambiguity averse agents is introduced by Klibanoff, Marinacci, and Mukerji (2005). Klibanoff, Marinacci, and Mukerji (2009), Hayashi and Miao (2011) provide a dynamic axiomatization of the smooth ambiguity preference. Following that approach, Chen, Ju, and Miao (2014) study the qualitative and quantitative asset allocation implications of model ambiguity and ambiguity aversion in Bayesian frameworks. Ju and Miao (2012) propose a generalized recursive smooth ambiguity model which permits a three-way separation among risk aversion, ambiguity aversion, and inter-temporal substitution in a consumption-based asset-pricing model. Relative to this branch of the literature, the innovation of our model is that it allows for ambiguity both in the mean and in the covariance matrix, while most of the existing literature assumes away that ambiguity in the covariance matrix. Epstein and Ji (2013) consider ambiguity in the volatility of one asset. Liu and Zeng (2017) study the effect of correlation ambiguity on portfolio under-diversification. Add reference to NYCU.

The second branch of the literature that our paper relates to is the literature on the beta anomaly and the idiosyncratic volatility puzzle. Liu, Stambaugh, and Yuan (2016) provides a excellent summary of that literature. They point out that most of the explanations of the beta anomaly in the literature are beta driven. The common thread is that investors exhibit preference for high beta stocks. When there is borrowing constraint, that preference leads to an equilibrium in which the security market line is latter than predicted by CAPM. In addition to the papers mentioned above, Baker, Bradley, and Wurgler (2011) and Christoffersen and Simutin (2016) share a similar idea. Due to the desire to benchmark their portfolios, fund managers exhibit preference for high beta stocks. For those explanations the challenge is to reconcile with the evidence documented in Liu, Stambaugh, and Yuan (2016), which shows that the price of beta for over-priced and under-priced stocks have opposite signs. We show that over-pricing or under-pricing is driven by ambiguity, which is un-correlated with beta. As argued earlier, on double-sorting by mis-pricing and beta, the alpha can exhibit the pattern shown in Liu, Stambaugh, and Yuan (2016).

The paper that is closely related to ours is Kogan and Wang (2003). We adopt the same mean-variance framework. One difference is our introduction of ambiguity in variance-covariance matrix. The key difference is in our focus on the role of ambiguity for understanding of the beta anomaly and the idiosyncratic volatility puzzle.

The remainder of this paper is organized as follows. Section 2 describes our model. Section 3 presentsits the equilibrium asset pricing implications. Section 4 focus on the role of ambiguity for understanding beta anomaly. Section 5 summarizes the results and concludes.

2 The Model

2.1 The Setting

The setting is similar to that in Kogan and Wang (2003). We consider a frictionless representative agent economy where the agent has constant absolute risk aversion (CARA) utility with risk aversion parameter $\gamma > 0$,

$$U(x) = -\frac{e^{-\gamma x}}{\gamma},$$

The agent is endowed with an initial wealth W_0. Without loss of generality, we assume $W_0 = 1$. Consumption takes place at the end of the period. The agent trades $N + 1$ assets, one riskless asset with instantaneous riskless return r and N risky assets whose returns follow a joint normal distribution. The representative agent knows that the returns are jointly normally distributed. We assume, however, that she is ambiguous about the expected return vector μ and variance-covariance matrix Ω. The setting is otherwise as in a standard mean-variance model. It is this ambiguity about the expected return vector and variance-covariance matrix that sets our setting apart from that of the CAPM theory. We turn now to the description of the ambiguity the agent faces and her aversion to it.

2.2 Ambiguity and Ambiguity Averse Preferences

Due to the lack of perfect knowledge of the probability law of asset returns, the agent's preference can not be represented by the standard expected utility, and instead is represented by a multi-prior max-min utility (Gilboa and Schmeidler (1989)).

$$\min_{\mathcal{Q} \in \mathcal{P}} \{E^Q[u(W)]\}, \tag{1}$$

where $\mathcal{P}$ is a set of probability priors. For our study, the specification of $\mathcal{P}$ is important. We will set that it is determined by the reference model and ambiguity parameters.

Specifically, the set of priors, $\mathcal{P}$, is assumed to take the form

$$\mathcal{P} = \left\{ Q : \, v_{J_k}^{\top} \Omega_{J_k}^{-1} v_{J_k} \leq 2\eta_{1,k}, \; tr(\Omega_{J_k}^{-1} U_{J_k}) - \ln |I_{J_k} + \Omega_{J_k}^{-1} U_{J_k}| \leq 2\eta_{2,k}, \; k = 1, ..., K \right\} \tag{2}$$

where Q are probability measures under which the returns of the assets are jointly normally distributed with density function given by

$$f_Q(R) = (2\pi)^{-N/2}|\hat{\Omega}|^{-1/2}e^{-\frac{1}{2}(R-\hat{\mu})^\top\hat{\Omega}^{-1}(R-\hat{\mu})},$$

J_k is a subset of $\{1, 2, \ldots, N\}$, $v = (\mu - \hat{\mu})$, $U = (\hat{\Omega} - \Omega)$, and v_{J_k} denotes the sub-vector consisting of the elements of v in the subset J_k. All the other notations with subscript J_k have similar meaning. In the set $\mathcal{P}$, the probability measure for which $v = 0$ and $U = 0$ is called the reference model and is denoted by P. The density function of the return distribution under P is given by

$$f(R) = (2\pi)^{-N/2}|\Omega|^{-1/2}e^{-\frac{1}{2}(R-\mu)^\top\Omega^{-1}(R-\mu)}.$$

The motivation of the specific form of $\mathcal{P}$ is the same as in Kogan and Wang (2003) and Uppal and Wang (2003). It is a set of priors defined by log likelihood ratio or relative entropy (Anderson, Hansen, and Sargent (2003) and Uppal and Wang (2003)). The set $\mathcal{P}$ is essentially a confidence region. In Kogan and Wang (2003), as there is no ambiguity about the variance-covariance matrix, $\eta_{2,k} = 0$, for $k = 1, \ldots, K$. The $\mathcal{P}$ in (2) is more general and can accommodate ambiguity in variance-covariance matrix as well. We now provide the detailed explanation of what set $\mathcal{P}$ captures. We do so with two elaborated examples.

2.2.1 A Single Source of Information

As the true probability law of asset returns is unknown, an econometrician has to estimate a model of asset returns based on the data available and assumption on the data generating model. Suppose that there is only a single data source of the stock returns. The result of the estimation is the reference model P. However, the econometrician is not completely sure that his reference model is indeed the true model. So he provides, along with the reference model P, a measure of his confidence that the true model is not far from the reference model, say, a 95% confidence region. Let Q be a probability measure that is potentially the true model. As the representative agent knows that the returns follow a joint normal distribution, the return under Q has density given by

$$f_Q(R) = (2\pi)^{-N/2}|\hat{\Omega}|^{-1/2}e^{-\frac{1}{2}(R-\hat{\mu})^\top\hat{\Omega}^{-1}(R-\hat{\mu})},$$

Under this measure, the expected return vector is $\hat{\mu}$ and the variance-covariance matrix is $\hat{\Omega}$. One measure of confidence the econometrician can use is the log likelihood ratio, $E^Q[\ln \xi]$, where $\xi = dQ/dP$ is the density of Q with respect to P. In terms of the reference probability, the likelihood ratio is the relative entropy, $E[\xi \ln(\xi)]$, of Q with respect to P. As argued in Kogan and Wang (2003) and Uppal and Wang (2003), $E[\xi \ln(\xi)]$ is a good approximation of the empirical log-likelihood when the number of observations is large.

It is readily verified that

$$\frac{dQ}{dP} = \xi(R) = \frac{|\Omega|^{\frac{1}{2}}}{|\hat{\Omega}|^{\frac{1}{2}}} e^{-\frac{1}{2}(R-\hat{\mu})^\top \hat{\Omega}^{-1}(R-\hat{\mu}) + \frac{1}{2}(R-\mu)^\top \Omega^{-1}(R-\mu)},$$

A bit of algebra in the appendix shows

$$E[\xi \ln(\xi)] = \frac{1}{2}\left[tr(\Omega^{-1}(\hat{\Omega} - \Omega)) - \ln|I + \Omega^{-1}(\hat{\Omega} - \Omega)| + (\mu - \hat{\mu})^\top \Omega^{-1}(\mu - \hat{\mu})\right] \tag{3}$$

Using relative entropy, the confidence region is of the form

$$\{Q : \, E[\xi \ln(\xi)] \leq \eta\},$$

The level of confidence is determined by η.

For normal distributions, the relative entropy has a very useful property. Fix a Q and suppose that it is the true model. Suppose first that there is no ambiguity about the true variance-covariance matrix. Then the relative entropy, denoted by L_{mean}, is given by

$$L^Q_{\text{mean}} = \frac{1}{2}(\mu - \hat{\mu})^\top \Omega^{-1}(\mu - \hat{\mu})$$

Suppose next that there is no ambiguity about the true mean return vector. Then the relative entropy, denoted L_{cov}, is given by

$$L^Q_{\text{cov}} = \frac{1}{2}\left(tr(\Omega^{-1}(\hat{\Omega} - \Omega)) - \ln|I + \Omega^{-1}(\hat{\Omega} - \Omega)|\right)$$

Comparing with the general expression for relative entropy ratio when there is ambiguity both in the mean return vector and in the variance covariance matrix, we have

$$E[\xi \ln \xi] = L^Q_{\text{mean}} + L^Q_{\text{cov}}$$

In specifying $\mathcal{P}$ in (2), we have taken advantage of this property of relative entropy. For Q to be in $\mathcal{P}$, what we require is that $L^Q_{\text{mean}} \leq \eta_1$ and $L^Q_{\text{cov}} \leq \eta_2$, which then implies that $E[\xi \ln \xi] \leq \eta_1 + \eta_2 = \eta$. Thus, in the case of one source of information, $\mathcal{P}$ in (2) is a confidence region.

2.2.2 Multiple Sources of Information

More realistically, the investors can obtain multiple data sources on the returns and each data source pertains to only a subset of the risky assets. To model multiple sources of information, let $J_k = \{j_1, j_2, ..., j_{N_k}\}$, $k = 1, 2, \ldots, K$, be subsets of $\{1, 2, \ldots, N\}$, and $\cup_k J_k = \{1, 2, ..., N\}$. So overall the agent has some information about each asset. The distribution of asset returns for any source of information J_k is $R_{J_k} = (R_{j_1}, R_{j_2}, ..., R_{j_{N_k}})$. We assume the reference probability law implied by the various sources of information coincides with the marginal distributions of the reference model P (denoted as P_{J_k}). The density function of R_{J_k} under the true model Q is

$$f(R_{J_k}) = (2\pi)^{-1}|\hat{\Omega}_{J_k}|^{-1/2}e^{-\frac{1}{2}(R_{J_k}-\hat{\mu}_{J_k})^\top \hat{\Omega}_{J_k}^{-1}(R_{J_k}-\hat{\mu}_{J_k})},$$

which is the marginal distribution of Q (denoted as Q_{J_k}), where $\hat{\mu}_{J_k}$ and $\hat{\Omega}_{J_k}$ are the mean return vector and variance-covariance return matrix of R_{J_k}. Thus, the likelihood ratio of the marginal distribution Q_{J_k} with respect to P_{J_k} is

$$\xi(R_{J_k}) = \frac{|\Omega_{J_k}|^{\frac{1}{2}}}{|\hat{\Omega}_{J_k}|^{\frac{1}{2}}}e^{-\frac{1}{2}(R_{J_k}-\hat{\mu}_{J_k})^\top \hat{\Omega}_{J_k}(R_{J_k}-\hat{\mu}_{J_k})+\frac{1}{2}(R_{J_k}-\mu_{J_k})^\top \Omega_{J_k}(R_{J_k}-\mu_{J_k})},$$

For convenience, we use the same notation $\hat{\Omega}_{J_k}^{-1}$ ($\Omega_{J_k}^{-1}$) to denote the $N \times N$-matrix whose elements in the j_m-th row and j_n-th column, for j_m and j_n in J_k, is the same as the elements in the m-th row and n-th column of the matrix $\hat{\Omega}_{J_k}^{-1}$ ($\Omega_{J_k}^{-1}$), otherwise it is zero. Then the relative entropy is

$$E[\xi_{J_k}\ln(\xi_{J_k})] = \frac{1}{2}\left[tr(\Omega_{J_k}^{-1}(\hat{\Omega}_{J_k} - \Omega_{J_k})) - \ln|I + \Omega_{J_k}^{-1}(\hat{\Omega}_{J_k} - \Omega_{J_k})| + (\mu - \hat{\mu})^\top \Omega_{J_k}^{-1}(\mu - \hat{\mu})\right] \quad (4)$$

The confidence region resulting from multiple sources of information is then of the form,

$$\{Q:\ E[\xi_{J_k}\ln(\xi_{J_k})] \leq \eta_k,\ k = 1, 2, ..., K\}. \quad (5)$$

By a similar argument given in the case of single source of information, the set $\mathcal{P}$ in (2) is again a confidence region in the more general case of multiple sources of information.

3 Portfolio Choice

In this section, we examine the representative agent's portfolio choices. To understand how the agent trades off ambiguity, risk and return, it is useful to introduce a metric for ambiguity. As argued earlier, a particularly useful property of the relative entropy for normal distributions allows us to decompose the total likelihood into L_{mean} and L_{cov}, which makes it possible for us to measure mean return and variance-covariance ambiguity separately. In the next two subsections, we will introduce our metric for mean return and variance-covariance ambiguity, respectively.

3.1 Measure of Mean Ambiguity

Suppose first that there is no variance-covariance ambiguity. In this case, the relative entropy including mean ambiguity only becomes

$$E[\xi \ln(\xi)] = \frac{1}{2}(\mu - \hat{\mu})^\top \Omega^{-1}(\mu - \hat{\mu}),$$

Let θ denote the portfolio of the agent and $\theta^\top R$ the portfolio return. The metric we use to measure the ambiguity in the mean return of the portfolio is given as

$$\Delta_1(\theta) = \sup_{Q \in \mathcal{P}} \{\theta^\top (\mu - \hat{\mu})\}, \tag{6}$$

where

$$\mathcal{P}_1 = \{Q : \ E[\xi_{J_k} \ln(\xi_{J_k})] = (\mu - \hat{\mu})^\top \Omega_{J_k}^{-1}(\mu - \hat{\mu}) \leq 2\eta_{1,k}, \ k = 1, 2, ..., K\}.$$

By definition of the metric, the difference between the expected return of the portfolio under the reference model P and the true expected return of the portfolio, $\theta^\top(\mu - \hat{\mu})$, falls into the interval $[-\Delta_1(\theta), \Delta_1(\theta)]$. Thus $\Delta_1(\theta)$ is the maximum possible error in using the reference model P to gauge the true expected return of the portfolio, given the confidence region described by $\mathcal{P}_1$. Clearly, the smaller the $\Delta_1(\theta)$, the less ambiguity there is about the expected return of the portfolio. Lemma 1 provides more on the metric.

Lemma 1 *Let θ be a portfolio of the agent. A solution to (6) exists. If the portfolio θ is such that $\theta_i \neq 0$ for all $i = 1, \ldots N$, then the solution $v(\theta)$ is unique and is given by,*

$$v(\theta) = \Omega_\mu(\theta)\theta, \tag{7}$$

where $\Omega_\mu(\theta)$

$$\Omega_\mu(\theta) = \left(\sum_{k=1}^{K} \lambda_{1,k}(\theta) \Omega_{J_k}^{-1} \right)^{-1}.$$

and $\lambda_{1,k}$, $k = 1, \ldots, K$, are Lagrangian multipliers for the K constraints in the definition of $\mathcal{P}_1$.

Obviously, $\Delta_1(\theta) = \theta^\top v(\theta)$ depends on the set $\mathcal{P}_1$ and the portfolio θ. The Lagrangian multipliers $\lambda_{1,k}(\theta)$, $k = 1, \ldots, K$, measure how much each source of information contributes to the ambiguity of the portfolio. If $\lambda_{1,k}(\theta) = 0$, for example, then the kth source of information does not help to reduce the ambiguity, at least for the portfolio θ.

3.2 The Measure of Variance-Covariance Ambiguity

Now suppose that there is no ambiguity in the mean return vector. In this case,

$$E\xi \ln \xi] = \frac{1}{2}\left(tr(\Omega^{-1}(\hat{\Omega} - \Omega)) - \ln |I + \Omega^{-1}(\hat{\Omega} - \Omega)|\right)$$

We define the measure of the ambiguity in variance-covariance by

$$\Delta_2(\theta) = \sup_{Q_2 \in \mathcal{P}_2} \{\theta^\top U \theta\}, \tag{8}$$

where $U = (\hat{\Omega} - \Omega)$ and

$$\mathcal{P}_2 = \left\{Q : \frac{1}{2}\left[tr(\Omega_{J_k}^{-1}(\hat{\Omega}_{J_k} - \Omega_{J_k})) - \ln |I_{J_k} + \Omega_{J_k}^{-1}(\hat{\Omega}_{J_k} - \Omega_{J_k})|\right] \leq \eta_{2,k},\ k = 1, 2, ..., K\right\}.$$

If $\hat{\Omega}$ is the true variance-covariance matrix, then the true variance of the portfolio return is $\theta^\top \hat{\Omega} \theta$. However, under the reference model P, the variance is $\theta^\top \Omega \theta$. Thus, by using the reference model, given the confidence region described by $\mathcal{P}_2$, the maximum error in the variance of the return of the portfolio is given by $\Delta_2(\theta)$.

Lemma 2 *For non-zero portfolio θ, the solution of (8) exits and is unique.*

3.3 Portfolio Choice

Having defined the preference of the investor and the measure of ambiguity, we now turn to the portfolio choice problem of the agent. Using the utility function from (1), the representative agents utility maximization problem is

$$\sup_\theta \min_{Q \in \mathcal{P}} \{E^Q[-\gamma^{-1} e^{-\gamma W}]\},$$

where the set $\mathcal{P}$ is as given in (2), subject to the agent's wealth constraint

$$W = W_0[\theta(R - r\mathbf{1}) + 1 + r].$$

where $\mathbf{1}$ is the N-vector $(1, 1, \ldots, 1)^\top$.

Proposition 3 *The agents utility maximization problem has a solution θ given by,*

$$\theta = \gamma^{-1}(\Omega + U(\theta))^{-1}(\mu - r\mathbf{1} - v(\theta)), \tag{9}$$

where $v(\theta)$ and $U(\theta)$ are the solutions of (6) and (8), respectively, given the portfolio θ.

When there is no ambiguity, $v(\theta) = 0$ and $U(\theta) = 0$. (9) reduces to the standard mean-variance optimal portfolio. When there is only ambiguity in the expected returns, (9) reduces to the formula given in Kogan and Wang (2003). More generally, (9) says that in the presence of ambiguity, the agent behaves as if the true expected return vector of the assets is given by $\mu - r\mathbf{1} - v(\theta)$ and the variance-covariance matrix is given by $\Omega + U(\theta)$. The expected portfolio return is then $\theta(\mu - r\mathbf{1}) + (1 + r) - \Delta_1(\theta)$ and the variance of the portfolio return is $\theta^\top \Omega \theta + \Delta_2(\theta)$. That is, the agent behaves as if the expected portfolio return is that under the reference model provided by the econometrician adjusted downward by $\Delta_1(\theta)$, which is the ambiguity in the mean, and the variance is that under the reference model adjusted upward by $\Delta_2(\theta)$, which is the ambiguity in the variance.

4 Equilibrium Expected Returns

To derive the equilibrium, let θ_m denote the market portfolio. In equilbrium, the representative agent holds the market portfolio. By Proposition 3, the expected return on the individual stocks and on the market are given by

$$\begin{aligned} \mu - r1 &= \gamma\Omega\theta_m + \gamma U(\theta_m)\theta_m + v(\theta_m) && (10) \\ \mu_m - r &= \gamma\theta_m^\top\Omega\theta_m + \gamma\theta_m^\top U(\theta_m)\theta_m + \Delta_1(\theta_m), && (11) \end{aligned}$$

The following theorem follows from readily.

Theorem 4 *The equilibrium vector of expected excess returns is given by*

$$\mu - r1 = \lambda\beta + \lambda_\mu\beta_\mu + \lambda_\Omega\beta_\Omega, \tag{12}$$

where

$$\begin{aligned} \beta &= \frac{\Omega\theta_m}{\theta_m^\top\Omega\theta_m}, \quad \lambda = \gamma\theta_m^\top\Omega\theta_m = \gamma\sigma_m^2 \\ \beta_\mu &= \frac{\Omega_\mu(\theta_m)\theta_m}{\theta_m^\top\Omega_\mu(\theta_m)\theta_m}, \quad \lambda_\mu = \Delta_1(\theta_m) = \theta_m^\top\Omega_\mu\theta_m \\ \beta_\Omega &= \frac{U(\theta_m)\theta_m}{\theta_m^\top U(\theta_m)\theta_m}, \quad \lambda_\Omega = \gamma\Delta_2(\theta_m) = \gamma\theta_m^\top U(\theta_m)\theta_m. \end{aligned}$$

where $\Omega_\mu(\theta_m)$ and $U(\Omega)$ are solutions of (6) and (8), respectively.

Equation (12) has the obvious interpretation that $\lambda\beta$ is the risk premium, $\lambda_\Omega\beta_\Omega$ is the variance-covariance ambiguity premium, and $\lambda_\mu\beta_\mu$ is the expected return ambiguity premium. Clearly, when there is no ambiguity, (12) reduces to the standard CAPM. The second and third terms on the right hand side of (12) are equal to zero and the β is the standard CAPM beta. Just as the interpretation for the risk premium where λ is the price of risk and β is the risk, λ_μ and λ_Ω are the prices of ambiguity in the expected return and variance-covariance matrix, and β_μ and β_Ω are the ambiguities in the expected return and varianc-covariance matrix, respectively.

More importantly, Theorem 4 says that the expected returns of the assets have a three-factor structure. The first factor is the CAPM risk factor β. The second factor is the mean ambiguity factor β_μ, which is the covariance of individual asset return with the market return, taking $\Omega_\mu(\theta_m)$ as the variance-covariance matrix. The third factor is the variance-covariance ambiguity factor β_Ω, which is the covariance of individual asset return with the market return, taking $U(\theta_m)$ as the variance-covariance matrix. When the agent invests in asset j, she bears the uncertainty associated with the asset. If there is no ambiguity, the reference probability measure is the true probability measure and the uncertainty is pure risk. The expected return and the beta are completely determined by the reference model P. The systematic risk is given by β_j as in CAPM model. When there is ambiguity, the total uncertainty of investing in asset j comes not only from the risk β_j, but also from the ambiguity in the true mean return and variance-covariance matrix. In this case, the total premium of the asset is not determined by the risk premium only. For example, when there is no ambiguity, a zero-beta portfolio θ that neutralizes the standard market risk ($\theta^\top\beta = 0$) delivers approximately riskless return when the portfolio is large and the idiosyncratic risks are diversified away. When there is ambiguity, however, the return on that portfolio may no longer be approximately riskfree. The three factor structure described in Theorem 4 suggests that a portfolio θ that also neutralize ambiguity, that is, the portfolio such that $\theta^\top\beta = 0$, $\theta^\top\beta_\mu$ and $\theta^\top\beta_\Omega = 0$, is more likely to be riskless. It should be noted, however, even that portfolio may not produce true riskless rate of return as $\mathcal{P}$ is in reality only a confidence region. In other words, that portfolio θ will produce riskless rate of return with the confidence determined by $\mathcal{P}$.

While Theorem 4 provides a three-factor structure for the expected returns of the asset, a fundamental question is whether such a prediction of Theorem 4 is empirically distinguishable from that of the CAPM theory. We provide two examples to elaborate on that.

One Source of Information

When there is only one source of information ($K = 1$), it can be shown that

$$\mu - r\mathbf{1} \quad = \quad (\gamma + \gamma\delta_1 + \delta_2)\Omega\theta_m$$

where $\delta_1 > 0$ and $\delta_2 > 0$ are two positive numbers. Thus it is as if the representative agent lives in a world with risk only and she has a higher level of risk aversion. The standard CAPM holds. This is reminiscient of the result in Anderson, Hansen, and Sargent (2003). This example shows that the presence of ambiguity does not necessarily leads to violation of CAPM. In this case, the standard zero-beta portfolio will neutralize with the confidence determined by $\mathcal{P}$, the uncertainty from both risk and ambiguity.

Multiple Non-overlapping Sources of Information

Another interesting case is one of non-overlapping sources of information. Suppose that there are K sources of information and they are non-overlapping in the sense that each source of information is about a subset of the N assets and the subsets do not overlap. In this case we can divide N assets into K non-overlapping groups and solve (6) and (8) to get explicit expressions for $\Delta_1(\theta)$ and $\Delta_2(\theta)$.

Lemma 5 *Let θ be a portfolio weight vector and θ_{J_k} be the sub-vector of portfolio weights on assets in group k for $k = 1, \ldots, K$. If the K sources of information are non-overlapping, then the solutions of (6) and (8) are given by,*

$$v(\theta) = \begin{bmatrix} \frac{\sqrt{2\eta_{1,1}}}{\sigma_{J_1}}\Omega_{J_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \frac{\sqrt{2\eta_{1,K}}}{\sigma_{J_K}}\Omega_{J_K} \end{bmatrix} \begin{bmatrix} \theta_{J_1} \\ \vdots \\ \theta_{J_K} \end{bmatrix} \tag{13}$$

and

$$U(\theta) \quad = \quad \begin{bmatrix} \frac{2\cdot\Omega_{J_1}\theta_{J_1}\theta_{J_1}^\top\Omega_{J_1}}{(\lambda_{2,1}-2)\cdot\theta_{J_1}^\top\Omega_{J_1}\theta_{J_1}} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \frac{2\cdot\Omega_{J_K}\theta_{J_K}\theta_{J_K}^\top\Omega_{J_K}}{(\lambda_{2,K}-2)\cdot\theta_{J_K}^\top\Omega_{J_K}\theta_{J_K}} \end{bmatrix} \tag{14}$$

where $\sigma^2_{J_k}(\theta_m) = \theta_{J_k}^\top\Omega_{J_k}\theta_{J_k}$, $\lambda_{1,k}$ and $\lambda_{2,k}$ are given as the solutions of

$$\lambda_{1,k} = \sqrt{\frac{\theta_{J_k}^\top\Omega_{J_k}\theta_{J_k}}{2\eta_{1,k}}}, \quad 2\eta_{2,k} = -\ln\left(1 + \frac{2}{\lambda_{2,k}-2}\right) + \frac{2}{\lambda_{2,k}-2}. \tag{15}$$

Given the explicit solutions, it follows from Theorem (4) that, for the asset j in group k, the mean ambiguity beta is, for $j \in J_k$,

$$\beta_{\mu,j} = \frac{1}{\Delta_1(\theta_m)}v_j(\theta_m) = \frac{\sqrt{2\eta_{1,k}}\sigma_{J_k}}{\sum_{k=1}^K\sqrt{2\eta_{1,k}}\sigma_{J_k}}\beta_{J_k,j},$$

where $\beta_{J_k,j} = \text{cov}(r_j, \theta_{J_k}^\top R_{J_k})/\sigma^2(\theta_{J_k})$. Interestingly, the mean ambiguity beta of the market portfolio is the risk beta of portfolio θ_{J_k} scaled down by a weight, with the weight being determined by the ambiguity.

For the variance-covariance ambiguity beta,

$$\beta_{\Omega,j} = \frac{[U\theta_m]_j}{\theta_m^\top U\theta_m} = \frac{\theta_{J_k}^\top U_{J_k}(\theta_m)\theta_{J_k}}{\Delta_2(\theta_m)} \frac{[U_{J_k}(\theta_m)\theta_{J_k}]_j}{\theta_{J_k}^\top U_{J_k}(\theta_m)\theta_{J_k}} = \frac{\frac{2}{\lambda_{2,k}-2}\sigma_{J_k}^2}{\sum_{k=1}^K \frac{2}{\lambda_{2,k}-2}\sigma_{J_k}^2}\beta_{J_k,j}.$$

Putting things together, we have the following corollary,

Corollary 6 *If the K sources of information are non-overlapping, then the expected return on the individual asset j in the group k is given by*

$$\mu_j - r = \gamma\sigma_m^2\beta_j + \left(\sqrt{2\eta_{1,k}}\sigma_{J_k} + \frac{2\gamma\sigma_{J_k}^2}{\lambda_{2,k}-2}\right)\beta_{J_k,j}, \tag{16}$$

Corollary 6 shows that when there are more than one sources of information on the probability distribution of the returns, the equilibrium expected returns in our model differs from those in the CAPM theory.

5 Equilibrium Asset Prices

In this section, to prepare for the analysis in Section 6, we rewrite the equilibrium returns in Theorem 4 in terms of exogenous dividend process and calculate the equilibrium price. Suppose that the exogenous dividend D follows a normal distribution. The reference distribution is one with mean vector d and variance-covariance matrix Σ. Let θ_m denote the market portfolio in terms of portfolio weight and $\bar{\theta}_m$ denotes the market portfolio in terms of amount (total supply). Let P denote the equilibrium price vector P. Then

$$R_j = \frac{D_j}{P_j} - 1, \quad \mu_j = \frac{d_j}{P_j} - 1, \quad \Omega = \text{diag}(1/P)\Sigma\text{diag}(1/P),$$
$$R_m = \frac{\bar{\theta}_m^\top D}{\bar{\theta}_m^\top P} - 1, \quad \theta_m = \text{diag}(P)\bar{\theta}_m, \quad (\Omega\theta_m) = \text{diag}(1/P)\Sigma\bar{\theta}_m.$$

where diag(x) is the diagonal matrix whose diagonal elements are given by the elements of vector x. Note that $\theta_m^\top \mathbf{1}$ is not necessarily equal to one as the riskless rate r is exogenously given.

When there is no ambiguity, the equilibrium price vector is given by,

$$P = \frac{1}{1+r}(d - \gamma\Sigma\bar{\theta}_m),$$

and the beta is given by

$$\beta = \frac{1}{\bar{\theta}_m^\top \Sigma \bar{\theta}_m} \mathrm{diag}(1/P)\Sigma\bar{\theta}_m,$$

The expected excess return of individual asset and market portfolio are respectively,

$$\mu - r\mathbf{1} = \gamma \mathrm{diag}(1/P)\Sigma\bar{\theta}_m, \quad \mu_m - r = \gamma \bar{\theta}_m^\top \Sigma \bar{\theta}_m,$$

The CAPM holds,

$$\mu_j - r = \frac{1}{P_j} \frac{(\Sigma\bar{\theta}_m)_j}{\bar{\theta}_m^\top \Sigma \bar{\theta}_m} (\mu_m - r).$$

When there is mean ambiguity and variance-covariance ambiguity under independent source of information, Corollary 6 in Section 4 shows that the equilibrium price for the asset j in group k is

$$P_j = \frac{1}{1+r}\left(d_j - \gamma(\Sigma\bar{\theta}_m)_j - \left(\sqrt{\frac{2\eta_{1,k}}{\bar{\theta}_{J_k}^\top \Sigma_{J_k} \bar{\theta}_{J_k}}} + \frac{2\gamma}{(\lambda_{2,J} - 2)}\right)(\Sigma_{J_k}\bar{\theta}_{J_k})_j\right). \tag{17}$$

6 Understanding Anomalies

In this section, we show that the theory developed in the preceding sections can be applied to provide an alternative understanding of the beta and IVOL anomalies.

6.1 Over-Pricing and Under-Pricing

An analysis of anomaly typically starts with the mis-pricing of assets according to a benchmark asset pricing theory. To provide our analysis of the beta anomaly and the idiosyncratic volatility anomaly, we first define over-pricing and under-pricing in our model.

The setting of our model is that of CAPM, except that the representative agent has max-min utility instead of the expected utility. Thus the benchmark theory for over-pricing and under-pricing is CAPM. That is,

$$\mu_j - r = \alpha_j + (\mu_m - r)\beta_j,$$

and a non-zero α_j implies mis-pricing. Asset j is under-priced if $\alpha_j > 0$. It is over-priced if $\alpha_j < 0$. It then follows from Theorem 4 that

$$\alpha_j = [\lambda - (\mu_m - r)]\beta_j + \lambda_\mu \beta_{\mu,j} + \lambda_\Omega \beta_{\Omega,j}.$$

Since $\mu_m - r = \lambda + \lambda_\mu + \lambda_\Omega$,

$$\alpha_j = \left(\lambda_\mu \left[\frac{\beta_{\mu,j}}{\beta_j} - 1 \right] + \lambda_\Omega \left[\frac{\beta_{\Omega,j}}{\beta_j} - 1 \right] \right) \beta_j. \tag{18}$$

Equation (18) is the basis on which we provide our analysis of the beta and idiosyncratic volatility anomalies.

6.2 Beta Anomaly

In the classical CAPM of Sharpe (1964) and Lintner (1965) theory, stocks with higher betas should earn higher premia than stocks with lower betas. However, the empirical evidence shows that high-beta stocks earn too little compared to low-beta stocks (Black, Jensen, and Scholes (1972) and Fama and MacBeth (1973)). As noted in the introduciton, there are several explanations in the literature. Here in this section, based on the theory developed earlier, we provide an alternative explanation of the beta anomaly.

To do so, we calibrate our model to the data and double sort α_j by mispricing and beta. The basic idea is that when there is ambiguity, equilibrium return should compensate investors for bearing both risk and ambiguity. However, if the econometrician takes CAPM as the true model and uses realized returns to estimate the expected return and beta for each asset, he will see violation of CAPM and may misunderstand it as irrational mispricing.

We focus on the special case of (18) where there are non-overlapping sources of information about the mean of the liquidating dividends. We assume there is no ambiguity about the variance-covariance matrix here. The model is calibrated as follows.

1. Set the number of stocks n to be 1000. We make 1000 draws from the normal distribution $N(200, 5)$ as the mean vector d of the 1000 liquidating dividends. We use US stocks monthly price and return data to estimate the monthly variance-covariance matrix of the liquidating dividends Σ as follows. We randomly choose 1000 stocks (we require that each stock should have over 20 years' monthly data) and calculate the correlation matrix. We then draw 1000 times from $N(0.22, 0.08)$ and take the absolute values of the 1000 draws as the elements of the diagonal of Σ. The supply of each asset equals to 1. The risk aversion coefficient is 2. The risk-free rate is set to be $r = 3\%$, annualized.

2. Assume that there are non-overlapping sources of information about mean ambiguity of the liquidating dividends. We divide the 1000 stocks into two groups of 500 each. Draw 600 times

from the joint dividends distribution $N(d, \Sigma)$ and take those samples as realized dividends for the assets (dividend data for 50 years). Calculate the equilibrium return based on the simulated dividends, $r_{j,t} = D_{j,t}/P_j - 1$. The mean ambiguity confidence level of the first group and the second group are $\eta_1 = 200$ and $\eta_2 = 250$ respectively.

3. The econometrician uses those realized returns to run regressions to estimate CAPM beta and to calculate the variance of the residuals as Ivol for each asset. Calculate the average of excess returns of each asset as the true return and the average of market excess returns as the true market excess return. Then define the alpha as the difference between the true return and the product term of CAPM betas multiplying average market excess return. We also use αs to proxy for mispricing.

4. We double-sort the stocks by mis-pricing (α) and beta into 5 quintiles each and obtain 5×5 cells. For each cell, we compute the average of the αs of the stocks in that cell. We also compute the t-statistics of the average.

The result of the simulation is reported in the Table (1). Panel A reports the averages of the αs and Panel B reports the t-statistics of the averages. The middle of Panel A are the 5×5 cells of double-sort. The last column of Panel A shows the difference, $H - L$, between the average α of the stocks with the highest β and that with the lowest β. The last row reports the average of αs of stocks sorted by beta. The second last row are the difference in α between the most over-priced and the most under-priced stocks.

The simulation results in Table 1 are consistent with what have been documented in the literature. First, the last row of Panel A shows that overall there is a negative relation between beta and α. The $H - L$ of that row shows that the difference is -0.24% per month. Panel B shows that the difference is statistically significantly different from zero. This is consistent with teh beta anomaly documented in Black, Jensen, and Scholes (1972) and Fama and MacBeth (1973). Next Panel A shows that there is negative relation between beta and α among over-priced stocks. The H-L for over-priced stocks is -0.15% per month, which is statistically significantly different from zero as shown in Panel B of the table. At the same time, Panel A also shows that the relation between beta and α among under-priced stocks is positive, which is in contrast with that for over-priced stocks. The H-L for under-priced stocks is 0.07% per month. Panel B shows that the difference is statistically insignificantly different from zero. Overall, the pattern is qualitatively similar to that shown in Liu, Stambaugh, and Yuan (2016).

Table 1: Alphas for Portfolios Sorted on Beta and Mispricing

The table reports the alpha for portfolios formed by an independent 5 × 5 sort on Beta and Mispricing.

Mispricing	Beta Quintile					
Quintile	Lowest	2	3	4	Highest	H-L
			A. Alpha (%)			
Overpriced	-0.66	-0.63	-0.59	-0.69	-0.80	-0.15
2	-0.24	-0.24	-0.26	-0.25	-0.24	-0.01
3	0.02	-0.01	0.01	0.01	0.03	0.01
4	0.27	0.25	0.25	0.24	0.26	-0.01
Underpriced	0.66	0.65	0.67	0.71	0.72	0.07
Over-Under	-1.32	-1.28	-1.26	-1.40	-1.53	
All stocks	0.10	0.05	0.04	-0.07	-0.13	-0.24
			B. T statistics			
Overpriced	-14.99	-16.63	-15.55	-17.44	-27.39	-2.59
2	-15.54	-17.74	-20.74	-19.17	-15.80	-0.33
3	1.97	-0.65	1.22	1.17	1.87	0.65
4	21.32	24.28	22.70	19.04	13.05	-0.40
Underpriced	20.30	16.24	15.29	14.41	19.77	1.35
Over-Under	-24.17	-22.26	-20.36	-21.91	-32.73	
All stocks	3.29	1.80	1.35	-2.14	-2.85	-4.21

The explanations of the beta anomaly provided in the literature are mostly based on short-selling or borrowing constraints. One argument is that when short-selling constraint is binding, investors behave as if they are holding the market portfolio and a zero-beta portfolio (Black (1972), Frazzini and Pedersen (2014)). The expected return on the zero-beta portfolio is higher than that of the riskless rate. Thus it appears that the security market line is flatter than the one predicted by the CAPM theory. Another argument is that heterogeneous expectations and short-sale constraints tend to lead to over-pricing of high beta stocks. Thus the security market line is flatter or even downward sloping in time of higher disagreement (Hong and Sraer (2016)). The third and more recent argument is that the beta anomaly maybe the consequence of the idiosyncratic volatility anomaly (Liu, Stambaugh, and Yuan (2016)).

Our explanation of the beta anomaly, as illustrated by the result of the simulation, is different from all those explanations. The key to understanding our explanation is equation (18). Consider the case where there is ambiguity in the mean only, which is the case for the simulation. In that case, (18) reduces to

$$\alpha_j = \lambda_\mu \left[\frac{\beta_{\mu,j}}{\beta_j} - 1\right] \beta_j. \tag{19}$$

This equation shows that for assets with positive betas, which is the case for most assets in the real data and in our simulation, there is over-pricing if and only if $\beta_{\mu,j} < \beta_j$. If $\beta_{\mu,j}$ and β_j are un-correlated, then, for over-priced (under-priced) stocks with positive betas, α_j is increasing (decreasing) β_j and hence H-L is negative (positive) for over-priced (under-priced) stocks. Thus, double sorting by mis-pricing and beta is likely to lead to what is seen in the middle part of Table 1. When the ambiguity is not too large, $\beta_{\mu,j}$ is relatively small and $\frac{\beta_{\mu,j}}{\beta_j} < 1$ for most stocks. Consequently, over-pricing occurs more often than under-pricing and the result on under-pricing stocks is less likely to be statistically significant.

Next, let I_k denotes the number of stocks in column k. Summing over each column of Table 1, we have

$$\sum_{j \in I_k} \alpha_j / I_k = \lambda_\mu \left[\sum_{j \in I_k} \beta_{\mu,j} / I_k - \sum_{j \in I_k} \beta_j / I_k\right]. \tag{20}$$

Over the k = low, 2, 3, 4, high, columns, $\sum_{j \in I_k} \beta_j / I_k$ is increasing. When $\beta_{\mu,j}$ and β_j are un-correlated, $\sum_{j \in I_k} \beta_{\mu,j} / I_k$ is likely constant across the columns. Thus $\sum_{j \in I_k} \beta_{\mu,j} / I_k - \sum_{j \in I_k} \beta_j / I_k$ is decreasing across the columns. As a consequence, $\sum_{j \in I_k} \alpha_j / I_k$ is decreasing across the column.

Figure 1 illustrates the explanation described above for the case where $n = 100$. It plots the average excess returns of the assets against their CAPM betas. The red line is the security market

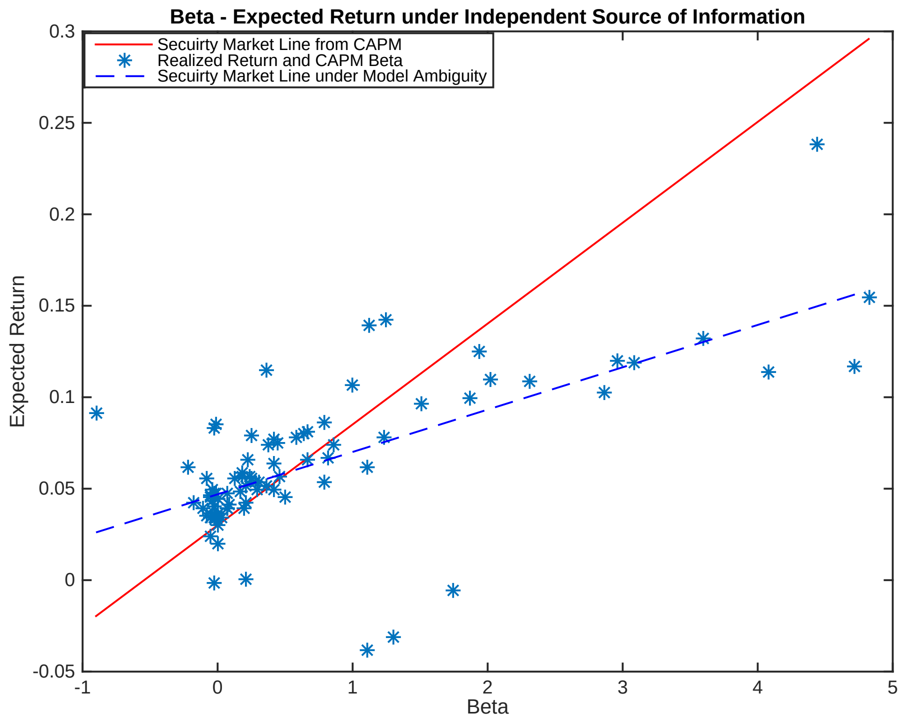

Figure 1: This figure plots the relation between betas and average returns of stocks.

line implied by CAPM and the blue line is the regression line. Assets that lie below the security line implied by CAPM have negative alphas and hence are over-priced. Those that lie above are under-priced. Overall, to the right, there is more over-pricing than under-pricing and vice versa, as described above. This is the beta anomaly, which is the last row of Panel A of Table 1.

It should be clear from the figure that if we regress the average returns of those assets that lie below the red CAPM line against their betas, the regression line will likely be flatter than the red CAPM line. If we regress the average returns of those assets that lie above the red CAPM line, the regression line will likely be steeper than the red CAPM line. If we do the regressions for the most over-price stocks and most under-priced stocks, the results will likely be statistically significant, which correspond to the first and fifth rows of Panel A of Table 1.

6.3 Idiosyncratic Volatility Anomaly

Ang, Hodrick, Xing, and Zhang (2006) find a puzzling empirical pattern that stocks with higher idiosyncratic volatility have subsequent lower returns, which is difficult to understand, because traditional theories predict either no relation between idiosyncratic volatility and expected returns (CAPM theory) or a positive relation due to market incompleteness and frictions (Merton (1987), Hirshleifer (1988)). Several potential explanations have been provided in the literature, such as preferences for expected idiosyncratic positive skewness (Barberis and Huang (2008), Boyer, Mitton, and Vorkink (2010)), coskewness (Chabi-Yo and Yang (2009)), lottery-like payoffs proxied by maximum daily return (Bali, Cakici, and Whitelaw (2011)), earnings surprises (Jiang, Xu, and Yao (2009), Wong (2011)), one-month return reversal (Fu (2009), Huang, Liu, Rhee, and Zhang (2010)), and illiquidity (Bali and Cakici (2008), Han and Lesmond (2011)). Chen and Petkova (2012) argue that IVOL proxies for sensitivity to a negative priced average volatility factor. Portfolios with high (low) idiosyncratic volatility relative to the Fama and French (1993) model have positive (negative) exposures to innovations in average stock variance and therefore lower (higher) expected returns. Recently, Stambaugh, Yu, and Yuan (2015) find a interesting empirical pattern that IVOL-return relation is negative among overpriced stocks but positive among underpriced stocks. Furthermore, the negative relation among overpriced stocks is stronger, especially for stocks less easily shorted, and consequently the overall IVOL-return relation is negative, which poses a challenge to all the explanations above.

In this paper, we provide an alternative explanation of the IVOL anomaly, similar in spirit to our explanation for the beta anomaly given in Section 6.2. We also show that our explanation based on ambiguity can be potentially consistent with empirical regularity documented in Stambaugh, Yu, and Yuan (2015).

First, following a similar setting as in Section 6.2 except the last step, we simulate the stock returns and assign stocks to mis-pricing (α) quintiles and IVOL quintiles and obtain 5×5 intersecting cells. We report the results in the Table 2. We see in the last row of Panel A of Table 2 that overall there is significantly negative relation between IVOL and the expected return of stocks, which is consistent with the negative relation reported in Ang, Hodrick, Xing, and Zhang (2006). Moreover, there is a significantly negative (positive) relation between expected return and IVOL of over-priced (under-priced) stocks, which is consistent with the IVOL-return empirical pattern documented in Stambaugh, Yu, and Yuan (2015). In terms of magnitude, the negative relation -0.29 is much stronger than the positive relation 0.18.

Next we provide a potential explanation of why our model can generate a similar IVOL-return empirical pattern documented in Stambaugh, Yu, and Yuan (2015). The basis is again equation

Table 2: Alphas for Portfolios Sorted on IVOL and Mispricing

The table reports the alpha for portfolios formed by an independent 5 × 5 sort on IVOL and Mispricing.

Mispricing	IVOL Quintile					
Quintile	Lowest	2	3	4	Highest	H-L
			A. Alpha (%)			
Overpriced	-0.62	-0.60	-0.60	-0.62	-0.91	-0.29
2	-0.22	-0.25	-0.27	-0.21	-0.26	-0.04
3	0.00	0.01	0.00	0.00	0.01	0.01
4	0.26	0.26	0.28	0.26	0.28	0.01
Underpriced	0.60	0.60	0.61	0.67	0.79	0.18
Over-Under	-1.22	-1.20	-1.21	-1.30	-1.69	
All stocks	0.03	0.02	0.02	0.02	-0.09	-0.12
			B. T statistics			
Overpriced	-12.76	-21.44	-20.34	-17.27	-22.08	-3.46
2	-18.87	-19.26	-21.06	-17.59	-15.30	-1.88
3	0.43	1.11	-0.37	0.04	0.90	0.54
4	24.07	24.51	21.63	17.44	15.80	0.62
Underpriced	19.11	18.88	21.45	19.67	15.85	2.48
Over-Under	-22.09	-28.42	-29.46	-26.08	-26.50	
All stocks	1.19	0.64	0.59	0.48	-1.79	-2.13

(19). It follows from Theorem 4,

$$\beta_j = \frac{(\Omega\theta_m)_j}{\theta_m^\top \Omega \theta_m}, \quad \text{and} \quad \beta_{\mu,j} = \frac{(\Omega_\mu \theta_m)_j}{\theta_m^\top \Omega_\mu \theta_m}$$

Let θ_μ be the portfolio defined by $\theta_\mu = \Omega^{-1}\Omega_\mu\theta_m$. Then the return of the portfolio is given by $R_\mu = \theta_\mu R$. By construction, $\Omega_\mu\theta_m = \Omega\theta_\mu$ and $\theta_m^\top\Omega_\mu\theta_m = \theta_\mu\Omega\Omega_\mu^{-1}\Omega\theta_\mu$. Let $\rho_{j,\mu}$ denote the correlation between asset j and the portfolio θ_μ and σ_μ^2 the variance of the return of the portfolio θ_μ. Similarly, let $\rho_{j,m}$ denote the correlation between asset j and the market portfolio θ_m and σ_m^2 the variance of the return of the market portfolio. Then

$$\alpha_j = \lambda_\mu(\beta_{\mu,j} - \beta_j) = \left[\frac{\sigma_\mu\sigma_m}{\theta_m^\top\Omega_\mu\theta_m}\rho_{j,\mu} - \rho_{j,m}\right]\frac{\sigma_j}{\sigma_m} \tag{21}$$

Clearly, over-pricing (under-pricing) occurs if and only if $\frac{\sigma_\mu\sigma_m}{\theta_m^\top\Omega_\mu\theta_m}\rho_{j,\mu} - \rho_{j,m}$ is less (greater) than zero. As long as $\frac{\sigma_\mu\sigma_m}{\theta_m^\top\Omega_\mu\theta_m}\rho_{j,\mu} - \rho_{j,m}$ is un-correlated with σ_j across assets, double-sorting by mis-pricing and σ_j will lead the empirical pattern exhibited in Table 2. One may argue that σ_j is the total volatility, not the idiosyncratic volatility. What equation (21) shows is the relationship between expected returns and total volatility. Empirically, however, over 95% of the total volatility of an asset is idiosyncratic volatility (Ang, Hodrick, Xing, and Zhang (2009)). Thus, relation exhibited in equation (21) between expected returns and total volatilities is likely true between expected returns and idiosyncratic volatilities, which is in fact what is found in Table 2.

If, instead of double-sorting by mis-pricing and σ_j, we sort the returns only by σ_j, then

$$\sum_{j\in I_k}\alpha_j/I_k = \lambda_\mu\left[\frac{\sigma_\mu\sigma_m}{\theta_m^\top\Omega_\mu\theta_m}\sum_{j\in I_k}\rho_{j,\mu}/I_k - \sum_{j\in I_k}\rho_{j,m}/I_k\right]\frac{\sigma_j}{\sigma_m} \tag{22}$$

where, for simplicity of illustration, we have assumed that across each column, assets have the same σ_j. As long as over the columns, $k =$ low, 2, 3, 4, high, $\frac{\sigma_\mu\sigma_m}{\theta_m^\top\Omega_\mu\theta_m}\sum_{j\in I_k}\rho_{j,\mu}/I_k - \sum_{j\in I_k}\rho_{j,m}/I_k$ is negative, to the first order approximation, $\sum_{j\in I_k}\alpha_j/I_k$ is decreasing across the column, which is what is seen in the last row of Panel A of Table 2.

7 Conclusion

We develop a model that is useful for understanding the cross-sectional characteristics of asset returns. The model is otherwise standard. The additional ingredient is that the agent is ambiguous about the probability distribution of the returns of the assets and he is ambiguity averse. The ambiguity can be about the mean as well as the variance-covariance matrix of the returns. The equilibrium cross-sectional expected returns can be described by a three-factor model, capturing risk, mean ambiguity and variance-covariance ambiguity respectively. Expected returns include a mean ambiguity premium, a variance-covariance ambiguity premium, as well as the standard risk premium. Our model helps explain a number of cross-sectional asset return behavior that is silent in standard models. Beta and idiosyncratic volatility are positively correlated. Overall the alpha in our model decreases with beta. However, when sorted by mis-pricing, alpha of over-priced assets decreases with beta, while alpha of under-priced assets increases with beta. The alphas' exhibit similar characteristics when sorted by total or idiosyncratic volatility. Alpha of over-priced assets decreases with total or idiosyncratic volatility, while alpha of under-priced assets increases with total or idiosyncratic volatility. Overall alpha decreases with beta total or idiosyncratic volatility. As argued by Liu, Stambaugh, and Yuan (2016), these cross-sectional characteristics of asset returns help explain the beta anomaly (Black, Jensen, and Scholes (1972) and Fama and MacBeth (1973)) and the idiosyncratic volatility anomaly (Ang, Hodrick, Xing, and Zhang (2006)).

A Appendix

A.1 Relative Entropy

Suppose that $R \sim N(\mu, \Omega)$ under P and $R \sim N(\hat{\mu}, \hat{\Omega})$ under Q. Then

$$\begin{aligned}
E[\xi \ln(\xi)] &= E^Q[\ln \xi] = \frac{1}{2} E^Q \left[\ln(\frac{|\Omega|}{|\hat{\Omega}|}) - (R-\hat{\mu})^\top \hat{\Omega}^{-1}(R-\hat{\mu}) + (R-\mu)^\top \Omega^{-1}(R-\mu) \right] \\
&= \frac{1}{2} \ln(\frac{|\Omega|}{|\hat{\Omega}|}) + \frac{1}{2} E_Q[-tr(\hat{\Omega}^{-1}(R-\hat{\mu})(R-\hat{\mu})^\top) + tr(\Omega^{-1}(R-\mu)(R-\mu)^\top] \\
&= \frac{1}{2}[\ln(\frac{|\Omega|}{|\hat{\Omega}|}) - N + tr(\Omega^{-1}\hat{\Omega}) + (\mu-\hat{\mu})^\top \Omega^{-1}(\mu-\hat{\mu})] \\
&= \frac{1}{2}[tr(\Omega^{-1}(\hat{\Omega}-\Omega)) - \ln|\Omega^{-1}\hat{\Omega}| + (\mu-\hat{\mu})^\top \Omega^{-1}(\mu-\hat{\mu})],
\end{aligned}$$

as is to be shown. ∎

A.2 Proof of Lemma 1

The first statement of the lemma is the same as that in the lemma 1 in Kogan and Wang (2003). The second statement of the lemma is a straightforward application of the Lagrangian duality approach. ∎

A.3 Proof of Lemma 2

Uniqueness: Note that the objective function is a linear function of $\hat{\Omega}$. In order to prove the uniqueness of the solution, we first prove the convexity of the constraints function. For any $k \in K$, denote

$$\begin{aligned}
g(\hat{\Omega}_{J_k}) &= \frac{1}{2}[\ln(\frac{|\Omega_{J_k}|}{|\hat{\Omega}_{J_k}|}) - N_{J_k} + tr(\Omega_{J_k}^{-1}\hat{\Omega}_{J_k})] - \phi^2 \eta_{2,k}, \\
&= \frac{1}{2}[-\ln(|\hat{\Omega}_{J_k}|) + tr(\Omega_{J_k}^{-1}\hat{\Omega}_{J_k})] + \frac{1}{2}[\ln(|\Omega_{J_k}|) - N_{J_k}] - \phi^2 \eta_{2,k}, \\
&= \frac{1}{2}[-\ln(|\hat{\Omega}_{J_k}|) + tr(\Omega_{J_k}^{-1}\hat{\Omega}_{J_k})] + C_k,
\end{aligned}$$

where $C_k = \frac{1}{2}[\ln(|\Omega_{J_k}|) - N_{J_k}] - \phi^2 \eta_{2,k}$ is a constant. Next we need to show, for any $\hat{\Omega}^1_{J_k}$, $\hat{\Omega}^2_{J_k}$ and $a \in (0,1)$,

$$g(a\hat{\Omega}^1_{J_k} + (1-a)\hat{\Omega}^2_{J_k}) \le a g(\hat{\Omega}^1_{J_k}) + (1-a) g(\hat{\Omega}^2_{J_k}),$$

Need to show,

$$
\begin{gathered}
-\ln(|a\hat{\Omega}^1_{J_k} + (1-a)\hat{\Omega}^2_{J_k}|) + tr[\Omega^{-1}_{J_k}(a\hat{\Omega}^1_{J_k} + (1-a)\hat{\Omega}^2_{J_k})] \leq \\
a[-\ln(|\hat{\Omega}^1_{J_k}|) + tr(\Omega^{-1}_{J_k}\hat{\Omega}^1_{J_k})] + (1-a)[-\ln(|\hat{\Omega}^2_{J_k}|) + tr(\Omega^{-1}_{J_k}a\hat{\Omega}^2_{J_k})],
\end{gathered}
$$

which is,

$$
\ln(|a\hat{\Omega}^1_{J_k} + (1-a)\hat{\Omega}^2_{J_k}|) \geq a\ln(|\hat{\Omega}^1_{J_k}|) + (1-a)\ln(|\hat{\Omega}^2_{J_k}|),
$$

From the simple version of Minkowski Inequality, if A and B are positive semidefinite Hermite Matrices, we can have,

$$
|A+B| \geq |A| + |B|,
$$

Therefore, only need to show,

$$
\ln(a|\hat{\Omega}^1_{J_k}| + (1-a)|\hat{\Omega}^2_{J_k}|) \geq a\ln(|\hat{\Omega}^1_{J_k}|) + (1-a)\ln(|\hat{\Omega}^2_{J_k}|),
$$

which is obvious because of the concavity of the log function.

Then we follow the same idea from lemma 1 in Kogan and Wang (2003). Suppose to the contrary that there exist two distinct solution $\hat{\Omega}^1$ and $\hat{\Omega}^2$. The convexity of all the constraints functions implies that for any $a \in (0,1)$, denote $\hat{\Omega}^a = a\hat{\Omega}^1 + (1-a)\hat{\Omega}^2$ and let $\hat{\Omega}^h_{J_k}$, $h = (1,2,a)$ denote the corresponding solution for J_k,

$$
\begin{aligned}
g(\hat{\Omega}^a_{J_k}) &= \frac{1}{2}[\ln(\frac{|\Omega_{J_k}|}{|\hat{\Omega}^a_{J_k}|}) - N_{J_k} + tr(\Omega^{-1}_{J_k}\hat{\Omega}^a_{J_k})] - \phi^2\eta_{2,k}, \\
&\leq ag(\hat{\Omega}^1_{J_k}) + (1-a)g(\hat{\Omega}^2_{J_k}), \\
&\leq 0, \quad k = 1,2,...,K.
\end{aligned}
$$

For k from 1 to K, we want to find all the possible k satisfy the following,

$$
\frac{1}{2}[\ln(\frac{|\Omega_{J_k}|}{|\hat{\Omega}^a_{J_k}|}) - N_{J_k} + tr(\Omega^{-1}_{J_k}\hat{\Omega}^a_{J_k})] - \phi^2\eta_{2,k} = 0, \quad for\, a = 0,1,\bar{a}.
$$

where $\bar{a} \in (0,1)$. Then we can have $\hat{\Omega}^1_{J_k} = \hat{\Omega}^2_{J_k}$ because of the convexity. Denote by A the set of such k. If

$$
J_A = \cup_{k\in A} J_k = 1,2,...,N,
$$

then $\hat{\Omega}^1 = \hat{\Omega}^2$, contradiction. When $J_A \neq 1, 2, ..., N$, WLOG, assume the first element is not in J_A. Thus for all $\hat{\Omega}$ of the following form

$$\hat{\Omega} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1N} \\ \sigma_{21} & \sigma_{12}^1 & \vdots & \sigma_{2N}^1 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{N1} & \sigma_{N2}^1 & \cdots & \sigma_{NN}^1 \end{bmatrix} \tag{23}$$

satisfy

$$\frac{1}{2}[\ln(\frac{|\Omega_{J_k}|}{|\hat{\Omega}_{J_k}|}) - N_{J_k} + tr(\Omega_{J_k}^{-1}\hat{\Omega}_{J_k})] - \phi^2\eta_{2,k} = 0, \quad for\, k \in A.$$

where $\sigma_{1i} \in \mathcal{R},\ i = 1, 2, ..., N$ is the variance (covariance). Note that for $a = \frac{1}{2}$,

$$\hat{\Omega}^a = \begin{bmatrix} \frac{\sigma_{11}^1+\sigma_{11}^2}{2} & \frac{\sigma_{12}^1+\sigma_{12}^2}{2} & \cdots & \frac{\sigma_{1N}^1+\sigma_{1N}^2}{2} \\ \frac{\sigma_{21}^1+\sigma_{21}^2}{2} & \frac{\sigma_{22}^1+\sigma_{22}^2}{2} & \vdots & \frac{\sigma_{2N}^1+\sigma_{2N}^2}{2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\sigma_{N1}^1+\sigma_{N1}^2}{2} & \frac{\sigma_{N2}^1+\sigma_{N2}^2}{2} & \cdots & \frac{\sigma_{NN}^1+\sigma_{NN}^2}{2} \end{bmatrix} = \begin{bmatrix} \frac{\sigma_{11}^1+\sigma_{11}^2}{2} & \frac{\sigma_{12}^1+\sigma_{12}^2}{2} & \cdots & \frac{\sigma_{1N}^1+\sigma_{1N}^2}{2} \\ \frac{\sigma_{21}^1+\sigma_{21}^2}{2} & \sigma_{22}^1 & \vdots & \sigma_{2N}^1 \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\sigma_{N1}^1+\sigma_{N1}^2}{2} & \sigma_{N2}^1 & \cdots & \sigma_{NN}^1 \end{bmatrix}$$

because $J_A = 2, ..., N$. Then we have,

$$\frac{1}{2}[\ln(\frac{|\Omega_{J_k}|}{|\hat{\Omega}_{J_k}|}) - N_{J_k} + tr(\Omega_{J_k}^{-1}\hat{\Omega}_{J_k})] - \phi^2\eta_{2,k} < 0, \quad for\, k \notin A.$$

From continuity, there exists a $\epsilon > 0$ such that for all the $\hat{\Omega}$ in (23) with $\sigma_{1i} \in (\frac{\sigma_{1i}^1+\sigma_{1i}^2}{2} - \epsilon, \frac{\sigma_{1i}^1+\sigma_{1i}^2}{2} + \epsilon),\ i = 1, 2, ..., N$.

$$\frac{1}{2}[\ln(\frac{|\Omega_{J_k}|}{|\hat{\Omega}_{J_k}|}) - N_{J_k} + tr(\Omega_{J_k}^{-1}\hat{\Omega}_{J_k})] - \phi^2\eta_{2,k} \leq 0, \quad for\, k \notin A.$$

Combining with the case $k \in A$, we have,

$$\frac{1}{2}[\ln(\frac{|\Omega_{J_k}|}{|\hat{\Omega}_{J_k}|}) - N_{J_k} + tr(\Omega_{J_k}^{-1}\hat{\Omega}_{J_k})] - \phi^2\eta_{2,k} \leq 0, \quad for\, k = 1, 2, ..., K.$$

As we mentioned before, the objective function is a linear function of $\hat{\Omega}$, so we have,

$$\frac{\theta^\top\hat{\Omega}^1\theta}{\theta^\top\Omega\theta} = \frac{\theta^\top\hat{\Omega}^2\theta}{\theta^\top\Omega\theta} = \frac{\theta^\top\frac{\hat{\Omega}^1+\hat{\Omega}^2}{2}\theta}{\theta^\top\Omega\theta},$$

But now, all the $\hat{\Omega}$ in (23) with $\sigma_{1i} \in (\frac{\sigma_{1i}^1+\sigma_{1i}^2}{2} - \epsilon, \frac{\sigma_{1i}^1+\sigma_{1i}^2}{2} + \epsilon),\ i = 1, 2, ..., N$ are in the choice set, we can choose specific ϵ ($\hat{\Omega}^\epsilon$) to achieve higher value of $\frac{\theta^\top\hat{\Omega}^\epsilon\theta}{\theta^\top\Omega\theta}$, contradiction!

Next, we will apply the standard Lagrangian duality approach to solve the optimal matrix. We first write down the Lagrangian function as follows,

$$\mathcal{L} = \frac{\theta^\top \hat{\Omega}\theta}{\theta^\top \Omega\theta} - \sum_{k=1}^{K} \lambda_{2,k}\{\frac{1}{2}[\ln(\frac{|\Omega_{J_k}|}{|\hat{\Omega}_{J_k}|}) - N + tr(\Omega_{J_k}^{-1}\hat{\Omega}_{J_k})] - \phi^2\eta_{2,k}\},$$

Note that $\partial tr(\Omega^{-1}U)/\partial u_{ij} = tr(\Omega^{-1}U_{ij})$ where U_{ij} is the matrix which has zero everywhere except in the ith row and jth column where it is equal to 1. $\partial \ln|I + \Omega^{-1}U|/\partial u_{ij} = (I + U)_{ij}^{-1}$. $tr(A^\top B) = \sum_i \sum_j A_{ij}B_{ij}$. FOC is

$$\frac{\partial \mathcal{L}}{\partial \hat{\Omega}} = \frac{\theta\theta^\top \circ S}{\theta^\top \Omega\theta} - \sum_{k=1}^{K} \frac{\lambda_{2,k}}{2}(-\hat{\Omega}_{J_k}^{-1} + \Omega_{J_k}^{-1}) = 0,$$

where S is a sign matrix whose elements take 1 if there is variance-covariance ambiguity information about the corresponding elements in Ω and takes 0 otherwise. $\theta\theta^\top \circ S$ is the entry-wise product between two matrices, which produces another matrix where each element ij is the product of elements ij of the original two matrices. So

$$\sum_{k=1}^{K} \lambda_{2,k}(\theta)[\hat{\Omega}^*_{J_k}(\theta)]^{-1} = \sum_{k=1}^{K} \lambda_{2,k}(\theta)\Omega_{J_k}^{-1} - \frac{2\theta\theta^\top \circ S}{\theta^\top \Omega\theta}.$$

Note, similar with Lemma 1, the proof above is also based on the assumption that there are multiple sources of information (K) and the information can cover all the assets in the market. So $\sum_{k=1}^{K} \lambda_{2,k}(\theta)\Omega_{J_k}^{-1}$ should be a full-rank matrix. If there is no variance-covariance ambiguity about some elements in the original Ω, the above equations becomes $0 = 0$ in the corresponding elements, which means those equations are redundant. ∎

A.4 Proof of Proposition 1

The agents utility maximization problem is

$$\sup_{\theta} \min_{Q\in\mathcal{P}} E^Q\left[-\frac{1}{\gamma}e^{-\gamma(\theta^\top(R-r\mathbf{1})+(1+r))}\right] = \sup_{\theta}\left[-\frac{1}{\gamma}e^{-\gamma[\theta^\top(\mu-r\mathbf{1})+1+r-\Delta_1(\theta)]+\frac{1}{2}\gamma^2[\theta^\top\Omega\theta+\Delta_2(\theta)]}\right],$$

The FOC for θ is given by

$$\mu - r\mathbf{1} - \Delta_1'(\theta) - \gamma\Omega\theta - \frac{1}{2}\gamma\Delta_2'(\theta) = 0,$$

So the optimal portfolio choice follows

$$\mu - r\mathbf{1} = \Delta_1'(\theta) + \gamma\Omega\theta + \frac{1}{2}\gamma\Delta_2'(\theta).$$

By envelope theorem, we have, $\Delta_1'(\theta) = v(\theta)$ and $\Delta_2'(\theta) = 2U(\theta)\theta$. Thus

$$\mu - r\mathbf{1} = v(\theta) + \gamma(\Omega + U(\theta))\theta,$$

as is to be shown. ∎

If there is only one source of information, we can write down the optimal solution explicitly. For mean ambiguity,

$$v^*(\theta) = \lambda_1^{-1}(\theta)\Omega\theta,$$

plugging into the constraint, we can solve for $\lambda_1(\theta)$,

$$\lambda_1(\theta) = \sqrt{\frac{\theta^\top \Omega\theta}{2\eta_1}},$$

so

$$v^*(\theta) = \sqrt{\frac{2\eta_1}{\theta^\top \Omega\theta}}\Omega\theta.$$

when ambiguity aversion coefficient ϕ or mean ambiguity level η_1 takes 0, then $\phi v^*(\theta) = 0$. Hence there is no mean-ambiguity effect.

For variance-covariance ambiguity,

$$\lambda_2(\theta)[\hat{\Omega}^*(\theta)]^{-1} = \lambda_2(\theta)\Omega^{-1} - \frac{2\theta\theta^\top}{\theta^\top\Omega\theta}.$$

so

$$\begin{aligned}
\hat{\Omega}^*(\theta) &= (\Omega^{-1} - \frac{2}{\lambda_2(\theta)}\frac{\theta\theta^\top}{\theta^\top\Omega\theta})^{-1}, \\
&= \Omega + \frac{2}{[\lambda_2(\theta)-2]\theta^\top\Omega\theta}\Omega\theta\theta^\top\Omega,
\end{aligned}$$

plugging into the constraint, we can solve for $\lambda_2(\theta)$,

$$\begin{aligned}
2\phi^2\eta_2 &= \ln(\frac{|\Omega|}{|\hat{\Omega}^*(\theta)|}) - n + tr(\Omega^{-1}\hat{\Omega}^*(\theta)) \\
&= \ln(\frac{|\Omega|}{|\Omega||E + \frac{2}{[\lambda_2(\theta)-2]\theta^\top\Omega\theta}\theta\theta^\top\Omega|}) - n + tr(E + \frac{2}{[\lambda_2(\theta)-2]\theta^\top\Omega\theta}\theta\theta^\top\Omega) \\
&= -\ln(|E + \frac{2}{[\lambda_2(\theta)-2]\theta^\top\Omega\theta}\theta\theta^\top\Omega|) + \frac{2}{[\lambda_2(\theta)-2]\theta^\top\Omega\theta}tr(\theta^\top\Omega\theta) \\
&= -\ln(1 + \frac{2}{\lambda_2(\theta)-2}) + \frac{2}{\lambda_2(\theta)-2} = -\ln(1+\frac{2}{\lambda_2-2}) + \frac{2}{\lambda_2-2}.
\end{aligned}$$

when ambiguity aversion coefficient ϕ or variance-covariance ambiguity level η_2 takes 0, then $\frac{2}{\lambda_2-2} = 0$, and $\hat{\Omega}^*(\theta) = \Omega$. Hence there is no variance-covariance-ambiguity effect. ∎

References

Anderson, E. W., L. P. Hansen, and T. J. 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., R. J. Hodrick, Y. Xing, and X. Zhang, 2006, "The Cross-Section of Volatility and Expected Returns," *The Journal of Finance*, 61, 259–299.

Ang, A., R. J. Hodrick, Y. Xing, and X. Zhang, 2009, "High idiosyncraticvolatilityandlowreturns:Internationaland further U.S.evidence," *JournalofFinancialEconomics*, 11, 1–23.

Baker, M., B. Bradley, and J. Wurgler, 2011, "Benchmarks as limits to arbitrage: understanding the low-volatility anomaly," *Financial Analysts Journal*, 67, 4054.

Bali, T. G., and N. Cakici, 2008, "Idiosyncratic Volatility and the Cross Section of Expected Returns," *Journal of Financial and Quantitative Analysis*, 43, 2958.

Bali, T. G., N. Cakici, and R. F. Whitelaw, 2011, "Maxing out: Stocks as lotteries and the cross-section of expected returns," *Journal of Financial Economics*, 99, 427–446.

Barberis, N., and M. Huang, 2008, "Stocks as Lotteries: The Implications of Probability Weighting for Security Prices," *American Economic Review*, 98, 2066–2100.

Black, F., 1972, "Capital Market Equilibrium with Restricted Borrowing," *The Journal of Business*, 45, 444–455.

Black, F., M. C. Jensen, and M. S. Scholes, 1972, "The capital asset pricing model: some empirical tests," *Studies in the Theory of Capital Markets*.

Boyer, B., T. Mitton, and K. Vorkink, 2010, "Expected Idiosyncratic Skewness," *The Review of Financial Studies*, 23, 169–202.

Chabi-Yo, F., and J. Yang, 2009, "Default Risk, Idiosyncratic Coskewness and Equity Returns," *working paper*.

Chen, H., N. Ju, and J. Miao, 2014, "Dynamic Asset Allocation with Ambiguous Return Predictability," *Review of Economic Dynamics*, 17, 799–823.

Chen, Z., and L. Epstein, 2002, "Ambiguity, Risk, and Asset Returns in Continuous Time," *Econometrica*, 70, 1403–1443.

Chen, Z., and R. Petkova, 2012, "Does Idiosyncratic Volatility Proxy for Risk Exposure?," *Review of Financial Studies*, 25, 2745–2787.

Christoffersen, S. E. K., and M. Simutin, 2016, "On the demand for high-beta stocks: evidence from mutual funds," *working paper, University of Toronto*.

Dow, J., and S. R. d. C. Werlang, 1992, "Uncertainty Aversion, Risk Aversion, and the Optimal Choice of Portfolio," *Econometrica*, 60, 197–204.

Easley, D., and M. O'Hara, 2009, "Ambiguity and Non-participation: The Role of Regulation," *Review of Financial Studies*, 22, 1817–1844.

Easley, D., and M. O'Hara, 2010, "Microstructure and Ambiguity," *Journal of Finance*, 65, 1827–1846.

Epstein, L. G., and S. Ji, 2013, "'Ambiguous Volatility and Asset Pricing in Continuous Time," *Review of Financial Studies*, 26, 1740 – 1786.

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

Epstein, L. G., and T. Wang, 1994, "Intertemporal Asset Pricing under Knightian Uncertainty," *Econometrica*, 62, 283–322.

Epstein, L. G., and T. Wang, 1995, "Uncertainty, Risk-Neutral Measures and Security Price Booms and Crashes," *Journal of Economic Theory*, 67, 40 – 82.

Fama, E., and J. MacBeth, 1973, "Risk, Return, and Equilibrium: Empirical Tests," *Journal of Political Economy*, 81, 607–636.

Fama, E. F., and K. R. French, 1993, "Common risk factors in the returns on stocks and bonds," *Journal of Financial Economics*, 33, 3–56.

Frazzini, A., and L. H. Pedersen, 2014, "Betting against beta," *Journal of Financial Economics*, 111, 125.

Fu, F., 2009, "Idiosyncratic risk and the cross-section of expected stock returns," *Journal of Financial Economics*, 91, 24–37.

Gilboa, I., and D. Schmeidler, 1989, "Maxmin expected utility with non-unique prior," *Journal of Mathematical Economics*, 18, 141 – 153.

Han, Y., and D. Lesmond, 2011, "Liquidity Biases and the Pricing of Cross-sectional Idiosyncratic Volatility," *The Review of Financial Studies*, 24, 1590–1629.

Hansen, L. P., and T. J. Sargent, 2001, "Robust Control and Model Uncertainty," *The American Economic Review*, 91, 60–66.

Hayashi, T., and J. Miao, 2011, "Recursive Smooth Ambiguity Preferences,," *Theoretical Economics*, 6, 423–475.

Hirshleifer, D., 1988, "Residual Risk, Trading Costs, and Commodity Futures Risk Premia," *Review of Financial Studies*, 1, 173–193.

Hong, H., and D. Sraer, 2016, "Speculative betas," *The Journal of Finance*, 51, 2095–2144.

Huang, W., Q. Liu, S. G. Rhee, and L. Zhang, 2010, "Return Reversals, Idiosyncratic Risk, and Expected Returns," *The Review of Financial Studies*, 23, 147–168.

Jiang, G. J., D. Xu, and T. Yao, 2009, "The Information Content of Idiosyncratic Volatility," *Journal of Financial and Quantitative Analysis*, 44, 128.

Ju, N., and J. Miao, 2012, “Ambiguity, Learning, and Asset Returns,” *Econometrica*, 80, 559–591.

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

Klibanoff, P., M. Marinacci, and S. Mukerji, 2009, “Recursive Smooth Ambiguity Preferences,,” *Journal of Economic Theory*, 144, 930–976.

Kogan, L., and T. Wang, 2003, “A Simple Theory of Asset Pricing under Model Uncertainty,” Working paper.

Lintner, J., 1965, “The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets,” *The Review of Economics and Statistics*, 47, 13–37.

Liu, J., J. Pan, and T. Wang, 2005, “An Equilibrium Model of Rare-Event Premia and Its Implication for Option Smirks,” *Review of Financial Studies*, 18, 131–164.

Liu, J., R. F. Stambaugh, and Y. Yuan, 2016, “Absolving Beta of Volatilitys Effects,” *Journal of Financial Economics*, forthcoming.

Liu, J., and X. Zeng, 2017, “Correlation Ambiguity,” *working paper*.

Maccheroni, F., M. Marinacci, and A. Rustichini, 2006a, “Ambiguity aversion, robustness, and the variational representation of preferences,” *Econometrica*, 74, 1447–1498.

Maccheroni, F., M. Marinacci, and A. Rustichini, 2006b, “Dynamic Variational Preferences,” *Journal of Economic Theory*, 128, 4–44.

Maenhout, P., 2004, “Robust Portfolio Rules and Asset Pricing,” *Review of Financial Studies*, 17, 951–983.

Maenhout, P., 2006, “Robust portfolio rules and detection-error probabilities for a mean-reverting risk premium,” *Journal of Economic Theory*, 128, 136–163.

Merton, R. C., 1987, “A Simple Model of Capital Market Equilibrium with Incomplete Information,” *The Journal of Finance*, 42, 483–510.

Sharpe, W. F., 1964, “Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk,” *The Journal of Finance*, 19, 425–442.

Stambaugh, R. F., J. Yu, and Y. Yuan, 2015, “Arbitrage Asymmetry and the Idiosyncratic Volatility Puzzle,” *The Journal of Finance*, 70, 1903–1948.

Uppal, R., and T. Wang, 2003, “Model Misspecification and Underdiversification,” *The Journal of Finance*, 58, 2465–2486.

Wong, P., 2011, “Earnings shocks and the idiosyncratic volatility discount in the cross-section of expected returns,” *working paper*.