

Beta and Coskewness Pricing: Perspective from Probability Weighting

YUN SHI, XIANGYU CUI, XUN YU ZHOU*

ABSTRACT

The security market line is often flat or downward-sloping. We hypothesize that probability weighting plays a role and that one ought to differentiate between periods in which agents overweight extreme events and those in which they underweight them. Overweighting inflates the probability of extremely bad events and demands greater compensation for beta risk. Underweighting has the opposite effect. Overall, these two effects offset each other, resulting in a flat or slightly negative return–beta relationship. Similarly, overweighting the tails enhances the negative relationship between return and coskewness, whereas underweighting reduces it. We support our theory through an extensive empirical study.

*Shi is at Academy of Statistics and Interdisciplinary Sciences, Faculty of Economics and Management, East China Normal University, Shanghai, China. Email: yshi@fem.ecnu.edu.cn. Cui is at School of Statistics and Management, Shanghai Institute of International Finance and Economics, Shanghai University of Finance and Economics, Shanghai, China. Email: cui.xiangyu@mail.shufe.edu.cn. Zhou is at Department of Industrial Engineering and Operations Research and The Data Science Institute, Columbia University, New York, NY 10027, USA. Email: xz2574@columbia.edu. Shi acknowledges financial support from the National Natural Science Foundation of China under grants 71971083, 71601107, and from the State Key Program of National Natural Science Foundation of China under grant 71931004. Cui acknowledges financial support from the National Natural Science Foundation of China under grant 71601107. Zhou acknowledges financial support through a start-up grant at Columbia University and through the Nie Center for Intelligent Asset Management.

The classical CAPM theory of Sharpe (1964), Lintner (1965) and Mossin (1966) asserts that expected returns increase with beta, leading to an upward-sloping security market line. However, empirical studies have shown that the return–beta slope is flat or even downward-sloping; see e.g., Black, Jensen, and Scholes (1972), Fama and French (1992), and Baker, Bradley, and Wurgler (2011). Various explanations have been offered for why beta is not or is negatively priced, ranging from misspecification of risk (Jagannathan and Wang, 1996) to investor sentiment (Antoniou, Doukas, and Subrahmanyam, 2015) to aggregate disagreement (Hong and Sraer, 2016) to positive correlation with idiosyncratic volatility (Liu, Stambaugh, and Yuan, 2018).

We approach the beta anomaly from a different perspective, one that involves probability weighting. The central task of asset pricing is to characterize how expected returns are related to risk and to investors’ *perceptions* of risk. Probability weighting affects the perception of risk, especially in the two tails of the market returns; therefore, it is posited to play a role in beta pricing. On the other hand, a positive/negative skewness measures the risk of large positive/negative realizations and can be viewed as a part of tail risk. As a result, probability weighting is also relevant to skewness/coskewness pricing because substantially right-/left-skewed events, such as winning a lottery (Barberis and Huang, 2008; Bordalo, Gennaioli, and Shleifer, 2012) or encountering a catastrophe (Kelly and Jiang, 2014; Bollerslev, Todorov, and Xu, 2015; Kozhan, Neuberger, and Schneider, 2013), are more attractive/undesirable to an investor with probability weighting.

In this paper, we provide a theory for beta and coskewness pricing using rank-dependent utilities (Quiggin, 1982; Quiggin, 2012; Schmeidler, 1989; Abdellaoui, 2002). The key difference between the rank-dependent utility theory (RDUT) and the classical expected utility theory (EUT) is that in the former the utility is weighted by a probability weighting function assigned to ranked outcomes. So RDUT nests EUT but also captures risk attitude toward probabilities, especially those represented in the two tails of the return distributions. We first derive an equilibrium asset pricing formula for a representative RDUT economy, and we

then examine the signs and magnitudes of the risk premiums in covariance and coskewness. Our results show that the pricing kernel in this economy is the product of the marginal utility and the derivative of probability weighting, which implies that the *shape* of the weighting function matters for pricing.

The dominant view in the behavioral finance literature is that individuals overweight probabilities of both extremely good and extremely bad events, rendering an inverse S-shaped probability weighting function (Tversky and Kahneman, 1992; Wu and Gonzalez, 1996; Prelec, 1998; Hsu, Krajbich, Zhao, and Camerer, 2009; Tanaka, Camerer, and Nguyen, 2010).¹ Some authors, however, have documented empirical evidence that individuals sometimes *underweight* tail events (Hertwig, Barron, Weber, and Erev, 2004; Humphrey and Verschoor, 2004; Harrison, Humphrey, and Verschoor, 2009; Henrich, Heine, and Norenzayan, 2010), leading to S-shaped weighting functions. Accordingly, Barberis (2013) comments that the coexistence of overweighting and underweighting poses a challenge in the study of extreme events. Employing the method proposed in Polkovnichenko and Zhao (2013), we use S&P 500 index option data and S&P 500 index data from January 4, 1996 to April 29, 2016 to derive option-implied probability weighting functions, updated monthly. Among the total 244 months tested in our study, about 89.3% of the time the implied weighting functions are either S-shaped (about 37.7%) or inverse S-shaped (about 51.6%). This observation suggests that people tend to overweight or underweight the two tails *simultaneously* (Tversky and Kahneman, 1992; Gonzalez and Wu, 1999; Humphrey and Verschoor, 2004). We further use the implied weighting functions to construct a monthly probability weighting index that reflects investors' perceptions regarding tail events, based on which we divide the market into two regimes: overweighting and underweighting. The ratio between the number of months during which the market is in the overweighting regime and during which it is in the underweighting one is around 57% to 43%. Given this empirical finding, we can sepa-

¹Through functional magnetic resonance imaging (fMRI) data, Hsu et al. (2009) find that activities in the striatum during valuation of monetary gambles are nonlinear in probabilities in a manner that is consistent with overweighting tail events.

rate the overweighting and underweighting regimes and analyze the corresponding beta and coskewness pricing *separately and respectively*.

Our main theoretical result is a closed-form expression of the risk premiums of the covariance and coskewness of any given asset with respect to the market. These premiums depend on the utility function, as well as on the weighting function *evaluated* at the probability value of having a positive market return. During an overweighting month (in which the weighting function is inverse S-shaped), the agent overweights both tails. Overweighting the left tail makes the RDUT agent more risk averse toward bad states (related to the convex part of the weighting function) than her EUT counterpart, whereas overweighting the right tail makes her less risk averse (the concave part) toward good states. While these two effects seem to offset each other, the empirical fact that the average probability of having a positive market return is close to 1 for any reasonably long sample period indicates that only the *convex* part of the weighting function is relevant. This results in an enhanced, positive risk premium with respect to the covariance with the market. Symmetrically, during an underweighting month in which the agent underweights both tails of the market return distribution, she is less risk averse toward bad states and more risk averse toward good ones. However, only the former behavior is relevant because the probability of having a positive market return now falls in the *concave* domain of the weighting function. In this case, the agent demands less compensation for the covariance with the market compared to an EUT maximizer, and potentially may even be willing to pay for the risk if and when she sufficiently underweights the bad states. Finally, for a sufficiently long sample period covering comparable runs of overweighting and underweighting months, the elevated beta during the former periods and the reduced beta during the latter periods may cancel each other, leading to a flat or possibly downward-sloping security market line.

Kraus and Litzenberger (1976) and Harvey and Siddique (2000) show that a typical EUT agent is willing to pay for the coskewness, implying a negative premium in coskewness. Probability weighting is likely to strengthen this skewness-inclined preference during an over-

weighting period. This can be explained as follows. When the asset has positive coskewness with the market, the asset return has a fat right tail with respect to the market portfolio. The hope of a larger gain relative to the market due to overweighting further increases the attraction of the risky asset and requires even less compensation when compared to the case of no probability weighting. Symmetrically, when the coskewness is negative, the fear of a larger loss, which is also amplified by the overweighting, reduces the attraction of the risky asset and hence demands greater compensation. Combined, overweighting the tails inflates the negative relationship between expected return and coskewness, generating a significantly negative premium of the coskewness. When the agent underweights both tails of the market return, the probability weighting introduces a skewness-averse preference. Underweighting the tails deflates the negative relationship between expected return and coskewness, resulting in an insignificant (or even potentially positive) premium.

Our empirical study confirms the above key theoretical predictions. To carry out this study, the first important task is to identify the characteristics and level of probability weighting for any given period. We develop several indices for this purpose, using S&P 500 option data for the period January 4, 1996 – April 29, 2016. After deriving the non-parametric, option-implied probability weighting functions as described earlier, we propose an “implied fear index (IFI)” and an “implied hope index (IHI)” as the average values of the derivatives of the weighting function around 1 and 0 respectively. The two indices have highly similar dynamics (see Figure 6), with a correlation coefficient of 0.8. Hence we use the average of the two to define one of our probability weighting indices, PWI_1 , to reflect the level of probability weighting. Moreover, to cross examine how robust PWI_1 is, we introduce a second probability weighting index, PWI_2 , by fitting the parametric specification of Prelec’s weighting function.² The summary statistics of IFI, IHI, PWI_1 and PWI_2 (Table II) show that the percentages of time during which these indices indicate the presence of overweighting are remarkably close to each other. Moreover, PWI_1 and PWI_2 have a correlation coefficient

²Prelec’s weighting function generates inverse S-shaped and S-shaped functions, determined by a single parameter.

of 0.949 (see also Figure 7 for a comparison of their dynamics). Therefore, we use PWI_1 alone as the probability weighting index in our empirical study, and we use its median to divide between the overweighting regime and the underweighting one.

We then perform portfolio analysis to study the impact of probability weighting on the mean-covariance and mean-coskewness relationships, respectively. We use all the common stocks listed on NYSE, AMEX and NASDAQ for the period January 1991 – April 2016. The one-way sort portfolio analysis (Table III) shows that the spread covariance-sorted portfolio has a significantly positive return during overweighting months, a significantly negative return during underweighting months, and a slightly positive (but insignificant) average return for the entire period. The spread coskewness-sorted portfolio, on the other hand, has a significantly negative return during overweighting months, an insignificantly negative return during underweighting months, and a notably negative average return for the entire period. Moreover, the two-way sort portfolio analysis (used to separate the effects of covariance and coskewness) yields similar results (Table IV and Table V). Finally, a Fama–MacBeth cross-sectional regression analysis (Fama and MacBeth, 1973) further confirms all the results (Table VI).

The paper proceeds as follows. We present and solve the portfolio selection problem under RDUT in Section I. We derive the equilibrium asset pricing formula and the three-moment CAPM model in Section II. We present the empirical analysis in Section III. We conclude in Section IV. All proofs are presented in the Appendix.

I. Portfolio Selection under Rank-dependent Utility

A. Basic Setting

Consider a one-period-two-date market. The set of possible states of nature at date 1 is Ω and the set of events at date 1 is a σ -algebra \mathcal{F} of subsets of Ω . There are n assets (one risk-free asset and $n - 1$ risky assets), whose prices at time t are $P_{t1}, P_{t2}, \dots, P_{tn}$, $t = 0, 1$.

These assets are priced by

$$P_{0i} = \mathbb{E}[\tilde{m}P_{1i}], \quad i = 1, 2, \dots, n,$$

where \tilde{m} is the pricing kernel, which is an \mathcal{F} -measurable random variable such that $\mathbb{P}(\tilde{m} > 0) = 1$, $\mathbb{E}[\tilde{m}] < \infty$ and $\mathbb{E}[\tilde{m}P_{1i}] < \infty$ ($i = 1, 2, \dots, n$).

The representative agent in the market has an RDUT preference given by

$$U(\tilde{X}) = \int u(x)w'(1 - F_{\tilde{X}}(x))dF_{\tilde{X}}(x),$$

where $u(\cdot)$ is an (outcome) utility function, $w(\cdot)$ a probability weighting function, and $F_{\tilde{X}}(\cdot)$ the cumulative distribution function (CDF) of the random payoff \tilde{X} . Specific parametric classes of probability weighting functions proposed in the literature include those by Tversky and Kahneman (1992)

$$w(p) = \frac{p^\alpha}{(p^\alpha + (1 - p)^\alpha)^{1/\alpha}},$$

and by Prelec (1998)

$$w(p) = \exp(-(-\log(p))^\alpha),$$

where, in both cases, $0 < \alpha < 1$ corresponds to inverse S-shaped weighting functions (i.e., first concave and then convex; hence overweighting both tails) and $\alpha > 1$ corresponds to S-shaped one (i.e., first concex and then concave; hence underweighting both tails). In this paper, we will use Prelec's weighting functions, which are plotted in Figure 1 with different α 's. The smaller the $\alpha (< 1)$, the higher degree of overweighting, and the larger the $\alpha (> 1)$, the higher degree of underweighting.

[Insert Figure 1 here]

We make the following assumptions on the market.

ASSUMPTION 1: (i) The probability space $(\Omega, \mathcal{F}, \mathbb{P})$ admits no atom.

(ii) u is strictly increasing, strictly concave, continuously differentiable on $(0, \infty)$, and satisfies the Inada condition: $u'(0+) = \infty$, $u'(\infty) = 0$. Moreover, without loss of generality, $u(\infty) > 0$. The asymptotic elasticity of u is strictly less than one: $\lim_{x \rightarrow \infty} \frac{xu'(x)}{u(x)} < 1$.

(iii) w is strictly increasing and continuously differentiable on $[0, 1]$ and satisfies $w(0) = 0$, $w(1) = 1$.

B. Portfolio Selection

Use x_i to denote the portfolio weight in asset i and let shorting be disallowed. Then, the representative investor faces a one-period portfolio selection problem:

$$\begin{aligned}
 (P) \quad & \max_{\{x_i\}} \int u(x)w'(1 - F_{\tilde{R}_P}(x))dF_{\tilde{R}_P}(x), \\
 & \text{s.t.} \quad \tilde{R}_P = \sum_{i=1}^n x_i \tilde{R}_i, \\
 & \quad \quad \sum_{i=1}^n x_i = 1, \\
 & \quad \quad x_i \geq 0, \quad i = 1, 2, \dots, n,
 \end{aligned} \tag{1}$$

where \tilde{R}_P is the total return of the portfolio and \tilde{R}_i is the total return of asset i .

Noting $P_{0i} = \mathbb{E}[\tilde{m}P_{1i}]$, we have $\mathbb{E}[\tilde{m}\tilde{R}_i] = 1$. Thus, problem (P) is, equivalently,

$$\begin{aligned}
 (P') \quad & \max_{\tilde{R}_P \in \mathcal{F}} \int u(x)w'(1 - F_{\tilde{R}_P}(x))dF_{\tilde{R}_P}(x), \\
 & \text{s.t.} \quad \mathbb{E}[\tilde{m}\tilde{R}_P] = 1, \\
 & \quad \quad \mathbb{P}(\tilde{R}_P \geq 0) = 1.
 \end{aligned} \tag{2}$$

This problem specializes to the one solved by Xia and Zhou (2016).³ Define a function N by

$$N(q) = - \int_{\bar{w}^{-1}(q)}^{1-} Q_{\tilde{m}}^{-}(1-p) dp, \quad q \in [0, 1],$$

where $Q_{\tilde{m}}^{-}(p) \triangleq \sup\{x \in \mathbb{R} \mid F_{\tilde{m}}(x) < p\}$, $p \in (0, 1]$, is the lower quantile function of m , and \bar{w} is the dual of w : $\bar{w}(p) = 1 - w(1-p)$.

Applying Theorem 3.3 of Xia and Zhou (2016), we have the following result.

THEOREM 1: *Assume \tilde{m} has a continuous CDF $F_{\tilde{m}}$, and*

$$\int_0^{1-} (u')^{-1}(\mu \hat{N}'(\bar{w}(p))) Q_{\tilde{m}}^{-}(1-p) dp < \infty$$

for all $\mu > 0$, where \hat{N} is the concave envelope of N and \hat{N}' the right derivative of \hat{N} . Then the total return of the optimal portfolio is

$$\tilde{R}_P^* = (u')^{-1} \left(\lambda^* \hat{N}'(1 - w(F_{\tilde{m}}(\tilde{m}))) \right),$$

where the Lagrange multiplier λ^* is determined by

$$\mathbb{E} \left[\tilde{m} (u')^{-1} \left(\lambda^* \hat{N}'(1 - w(F_{\tilde{m}}(\tilde{m}))) \right) \right] = 1.$$

³In Problem (2.2) of Xia and Zhou (2016), setting $I = 1$, $u_{0i}(\cdot) = 0$, $u_{1i}(\cdot) = u(\cdot)$, $w_i(\cdot) = w(\cdot)$, $\beta_i = 1$, $\mathbb{E}[\tilde{\rho} \tilde{e}_{1i}] = 1$, $e_{0i} = 0$, and $\tilde{c}_{1i} = \tilde{R}_P$, we recover our Problem (P').

II. Equilibrium Asset Pricing and Three-Moment CAPM

A. Equilibrium Asset Pricing Formula

Denote by

$$\tilde{R}_M = \sum_{i=1}^n x_i^M \tilde{R}_i$$

the total return of the market portfolio where x_i^M is the market capitalization weight of asset i with $\sum_{i=1}^n x_i^M = 1$, and by \tilde{r}_M , \tilde{r}_i and r_f the return rates of the market portfolio, the asset i , and the risk-free asset, respectively. Define a function m by

$$m(x) = \frac{w'(1 - F_{\tilde{R}_M}(x))u'(x)}{(1 + r_f)\mathbb{E}\left[w'(1 - F_{\tilde{R}_M}(\tilde{R}_M))u'(\tilde{R}_M)\right]}, \quad x > 0. \quad (3)$$

THEOREM 2: *When the market is in equilibrium with \tilde{R}_M having a continuous CDF $F_{\tilde{R}_M}$, the pricing kernel is given by*

$$\tilde{m} = m(\tilde{R}_M) \equiv \frac{w'(1 - F_{\tilde{R}_M}(\tilde{R}_M))u'(\tilde{R}_M)}{(1 + r_f)\mathbb{E}\left[w'(1 - F_{\tilde{R}_M}(\tilde{R}_M))u'(\tilde{R}_M)\right]}. \quad (4)$$

Moreover, if \tilde{R}_M has a continuously differentiable density function $f_{\tilde{R}_M}$, then, for any risky asset i , we have the three-moment CAPM:

$$\mathbb{E}[\tilde{r}_i] = r_f + ACov(\tilde{r}_i, \tilde{r}_M) + \frac{1}{2}BCov(\tilde{r}_i, \tilde{r}_M^2) + o(\tilde{r}_M^2), \quad (5)$$

where the risk premiums A and B are given as

$$\begin{aligned}
A &= -(1+r_f)m'(1) \\
&= -\mathbb{E}^{-1}[w'(1-F_{\tilde{R}_M}(\tilde{R}_M))u'(\tilde{R}_M)] \\
&\quad \cdot \left[w'(1-F_{\tilde{R}_M}(1))u''(1) - w''(1-F_{\tilde{R}_M}(1))u'(1)f_{\tilde{R}_M}(1) \right], \tag{6}
\end{aligned}$$

$$\begin{aligned}
B &= -(1+r_f)m''(1) \\
&= -\mathbb{E}^{-1}[w'(1-F_{\tilde{R}_M}(\tilde{R}_M))u'(\tilde{R}_M)] \\
&\quad \cdot \left[w'(1-F_{\tilde{R}_M}(1))u'''(1) + w'''(1-F_{\tilde{R}_M}(1))u'(1)f_{\tilde{R}_M}^2(1) \right. \\
&\quad \left. - w''(1-F_{\tilde{R}_M}(1))[2u''(1)f_{\tilde{R}_M}(1) + u'(1)f'_{\tilde{R}_M}(1)] \right]. \tag{7}
\end{aligned}$$

The expression (4) indicates that the equilibrium pricing kernel \tilde{m} is not only influenced by the utility function u , but also by the probability weighting function w evaluated at $1 - F_{\tilde{R}_M}(\tilde{R}_M)$ and, *consequently*, by the market return distribution. To illustrate, we use Figure 2. There, we consider Prelec's probability weighting function, the CRRA utility function and two types of market scenarios:

- Market I: $1 - F_{\tilde{R}_M}(1) > e^{-1}$, i.e., $1 - F_{\tilde{R}_M}(1)$ lies in the convex domain (concave domain) of $w(\cdot)$ when $\alpha < 1$ ($\alpha > 1$).⁴
- Market II: $1 - F_{\tilde{R}_M}(1) < e^{-1}$, i.e., $1 - F_{\tilde{R}_M}(1)$ lies in the concave domain (convex domain) of $w(\cdot)$ when $\alpha < 1$ ($\alpha > 1$).

In Figure 2, both market scenarios are considered. Take \tilde{r}_M to be skew-normal with mean 7.6%, standard deviation 15.8% and skewness -0.339, which can be easily verified to be a case of Market I. For Market II we take \tilde{r}_M to be skew-normal with mean -7.6%, standard deviation 15.8% and skewness -0.339. Figures 2(a) and 2(d) show the plots of the pricing kernel against the market return. They show that, in both market scenarios, the RDUT agent overweighting both tails pays higher prices in the extreme states, resulting in a U-shaped

⁴ $p^* = e^{-1}$ is the solution of the equation $\exp(-(-\log(p))^\alpha) = p$ for any α ; hence it is the reflection point separating the convex and concave domains of Prelec's weighting function.

pricing kernel.⁵ The case of underweighting both tails has the opposite characteristics, with a bell-shaped pricing kernel.

Figures 2(b) and 2(e) show the implied relative risk aversion, $-m'(\tilde{R}_M) \cdot \tilde{R}_M/m(\tilde{R}_M)$, as a function of the market return \tilde{R}_M . These plots exhibit consistent qualitative behaviors in the two market scenarios. Overweighting (underweighting) both tails decreases (increases) risk aversion as market return improves; so the agent takes more (less) risk when the return is higher. The classical CRRA case has a flat relative risk aversion, as expected.

Figures 2(c) and 2(f) depict the implied relative risk aversion as a function of α when the market return is fixed at $\tilde{R}_M = 1$, which has the same sign of the parameter A (see (6)). In Market I, it is a decreasing function of α ; so a higher level of overweighting the tails leads to a larger relative risk aversion and hence demands a higher risk compensation. In particular, the risk aversion is negative when one sufficiently underweights, yielding a risk-seeking behavior. In Market II, the implied relative risk aversion is increasing in α , which implies that the more (less) overweighting (underweighting) the tails the less risk aversion is entailed.

[Insert Figure 2 here]

B. Risk Premiums

In the three-moment CAPM model, the terms A and B are the risk premiums of the covariance and the coskewness, respectively. The signs of the two parameters determine the risk preference with respect to covariance and coskewness. As A and B depend on the utility function, the weighting function, and the distribution of the market return in a complex way, it is difficult to analytically study their signs in general. However, if we choose

⁵U-shaped pricing kernels have been presented in literature under various settings. For example, Shefrin (2008) finds that the pricing kernel has a U-shape when agents have heterogeneous risk tolerance. Baele, Driessen, Ebert, Londono, and Spalt (2018) observe that pricing kernels under the cumulative prospect theory (CPT) are U-shaped.

a specific weighting function - in this case, Prelec's - then an analytical examination of the signs becomes possible.

As noted earlier, the market return distribution also impacts pricing and risk premiums. For Prelec's weighting function w , we have introduced two scenarios, Market I and Market II, depending on whether $1 - F_{\tilde{R}_M}(1)$ lies in the convex or concave domain of w . Notice that $1 - F_{\tilde{R}_M}(1) = \mathbb{P}(\tilde{R}_M > 1)$, i.e., the probability that the market has a positive return rate. So this quantity reflects the trend of the overall economy/market. Let us first examine which type of Market is more plausible for the S&P 500. Following De Giorgi and Legg (2012), we assume that the S&P 500 annual return rate, \tilde{r}_M , follows either a normal distribution, a skew-normal distribution, or a log-normal distribution. Based on the annual historical data of the S&P 500 from January 1946 to January 2009, the annual return rate, \tilde{r}_M , has a mean of 7.6% and a standard deviation of 15.8%. Moreover, in the case of skew-normal, its skewness is estimated to be -0.339. Based on these values, assuming \tilde{r}_M is respectively normal, skew-normal, and log-normal, we can estimate the corresponding value of $1 - F_{\tilde{R}_M}(1)$ to be respectively 0.6847, 0.6980, and 0.6543, all greater than e^{-1} . Next, we consider the S&P 500 monthly return. Following Polkovnichenko and Zhao (2013), we use the daily historical data of the S&P 500 from January 4, 1996 to April 29, 2016 and estimate an EGARCH model for the S&P 500 with 1250 daily index return data up to the end of each month. Then, we simulate 100,000 samples based on the EGARCH model and compute the nonparametric probability density function for the monthly return of the S&P 500. We find that $1 - F_{\tilde{R}_M}(1) > e^{-1}$ holds for all the months tested. Therefore, we conclude that Market I is a more likely scenario for the S&P 500, and hence we focus on this case in the following discussion.⁶

Let us now examine first the sign of A . From its expression on the right hand side of (6), it follows that the first term (including the factor in front of the brackets) is positive, representing the part of the risk aversion arising from the outcome utility function u . The

⁶Essentially, Market I is the case when the economy grows most of the time, which has historically been the case for US markets.

second term is more complicated, representing the part of the risk aversion arising from the probability weighting, whose sign depends on whether the quantity $1 - F_{\tilde{R}_M}(1)$ lies in the convex or concave domain of the weighting function w .

PROPOSITION 1: *Assume that the market is in equilibrium with \tilde{R}_M having a density function $f_{\tilde{R}_M}$, and that $w(p) = \exp(-(-\log(p))^\alpha)$ and $1 - F_{\tilde{R}_M}(1) > e^{-1}$. Then,*

- (i) $A > 0$ if $0 < \alpha < 1$.
- (ii) $A < 0$ if $\alpha > \alpha^*$ for some $\alpha^* > 1$.

To interpret Proposition 1, consider first the case $0 < \alpha < 1$ when w is inverse S-shaped and the agent overweights both extremely good and extremely bad states. Overweighting the left tail makes the RDUT agent more risk averse toward bad states than her CRRA counterpart, and overweighting the right tail makes her less risk averse toward good states. While these two effects seem to offset each other, the condition $1 - F_{\tilde{R}_M}(1) > e^{-1}$, which is consistent with the empirical observation for the S&P 500, as discussed earlier, dictates that *only* the former effect is relevant. This results in a positive risk premium with respect to the covariance with the market; see the solid red line in Figure 3.

There is more to it. Write

$$A = \frac{w'(1 - F_{\tilde{R}_M}(1))u'(1)}{\mathbb{E}[w'(1 - F_{\tilde{R}_M}(\tilde{R}_M))u'(\tilde{R}_M)]} \left[-\frac{u''(1)}{u'(1)} + \frac{w''(1 - F_{\tilde{R}_M}(1))}{w'(1 - F_{\tilde{R}_M}(1))} f_{\tilde{R}_M}(1) \right]. \quad (8)$$

The first term inside the brackets, $-\frac{u''(1)}{u'(1)} > 0$, measures the risk aversion arising from the outcome utility function u in the classical EUT framework. The second term, $\frac{w''(1 - F_{\tilde{R}_M}(1))}{w'(1 - F_{\tilde{R}_M}(1))} f_{\tilde{R}_M}(1)$, which arises from the probability weighting *independent* of the outcome utility function, is also positive when $0 < \alpha < 1$. This second term *elevates* the level of risk aversion. This may shed light on the equity premium puzzle (Mehra and Prescott, 1985): the reason the empirical data of the S&P 500 implies an extremely implausibly high level of risk aversion within the classical EUT framework is that the latter has not accounted for the part of the risk aversion emanating from probability weighting.

When $\alpha > 1$, w is S-shaped and the agent underweights the two tails of the market return distribution. This entails less risk aversion toward bad states and more risk aversion toward good states. However, the condition $1 - F_{\tilde{R}_M}(1) > e^{-1}$ implies that $1 - F_{\tilde{R}_M}(1)$ is now in the *concave* domain of w and hence only the former behavior is relevant. In this case, the agent demands less compensation for the beta risk compared with a classical utility maximizer, and, indeed, even becomes willing to pay for the risk (i.e., $A < 0$) when she sufficiently underweights bad states (i.e., when α is sufficiently large). This case is illustrated by the dashed blue line in Figure 3.⁷

[Insert Figure 3 here]

Next we investigate the sign of B . First, when there is no probability weighting, we have $B = -\mathbb{E}^{-1}[u'(\tilde{R}_M)]u'''(1) < 0$ assuming that $u''' > 0$, which is satisfied by most commonly used utility functions.⁸ This implies that a typical EUT agent is willing to pay for the coskewness; see also Harvey and Siddique (2000). In the presence of probability weighting, the first term of B (see (7)) is negative if, again, $u''' > 0$. The second and third terms are more complicated, in that their signs depend on the utility function, the weighting function, and the market return distribution, all intertwined. However, as in the case of A , whether $1 - F_{\tilde{R}_M}(1)$ lies in the convex or concave domain of w is also critical for the sign of B . We provide several sufficient conditions for B to be positive or negative.

PROPOSITION 2: *Assume that the market is in equilibrium with \tilde{R}_M having a continuously differentiable density function $f_{\tilde{R}_M}$, and that $u(x) = \frac{1}{1-\gamma}x^{1-\gamma}$ where $\gamma > 0$, $w(p) = \exp(-(-\log(p))^\alpha)$ where $\alpha > 0$, and $1 - F_{\tilde{R}_M}(1) > e^{-1}$. Then,*

- (i) $B < 0$ if $0 < \alpha < 1$ and $\gamma \geq \gamma^* := \frac{f'_{\tilde{R}_M}(1)}{2f_{\tilde{R}_M}(1)}$.
- (ii) $B < 0$ if $\alpha > 1$ and $\gamma > \bar{\gamma}(\alpha)$ where $\bar{\gamma}(\alpha)$ is a threshold depending on α .

⁷Figure 3 is completely consistent with Figure 2(c).

⁸Brockett and Golden (1987) refer to the class of increasing utility functions with derivatives that alternate in sign as the class that contains “all commonly used utility functions.” In particular, $u''' > 0$ implies that the utility function has nonincreasing absolute risk aversion, which is one of the essential properties of a risk-averse individual; see Harvey and Siddique (2000).

(iii) $B > 0$ if $\alpha > 1$ and $\gamma^* \leq \gamma < \underline{\gamma}(\alpha)$, where $\underline{\gamma}(\alpha)$ is a threshold depending on α .

We have noted earlier that in the absence of probability weighting, $B < 0$ for “most commonly used utility functions,” suggesting an inherent skewness-inclined preference in the classical EUT framework. Now, when the agent overweights both tails of the market return (the case of $0 < \alpha < 1$), the proof of Proposition 2 (Appendix C) shows that $w'''(1 - F_{\tilde{R}_M}(1)) > 0$. Thus probability weighting is likely to *enhance* the skewness-inclined preference when overweighting occurs. This enhancement can be explained intuitively as follows. When the asset has positive coskewness with the market, the asset return has a fat right tail with respect to the market portfolio. The hope of a larger gain relative to the market, which is strengthened by overweighting, further increases the attraction of the risky asset and requires even less compensation when compared to the case of no probability weighting. Symmetrically, when the coskewness is negative, the fear of a larger loss, which is also amplified by the overweighting, reduces the attraction of the risky asset and hence demands greater compensation. Combined, overweighting the tails inflates the negative relationship between expected return and coskewness, generating a significantly negative B . This enhancement is confirmed by comparing the empirical results across Panels A and B in Table I.⁹

When the agent underweights both tails of the market return (i.e., $\alpha > 1$), $w'''(1 - F_{\tilde{R}_M}(1))$ becomes negative, which implies that the probability weighting introduces a skewness-averse preference. Underweighting the tails deflates the negative relationship between expected return and coskewness. Whether B is eventually positive or negative depends on which one dominates: the skewness-inclination inherent in u or the skewness-aversion in w . Proposition 2-(ii) and -(iii) provide sufficient conditions under which one of these is the case. These results are illustrated by the dashed blue line in Figure 4 and by the values of B in Panel C of Table

⁹The additional condition $\gamma > \gamma^*$ in Proposition 2-(i) is purely technical, as it is used to control the product term $u''w''$ in the proof (Appendix C). Indeed, from that proof it is clear that this condition is sufficient, but by no means necessary. In Panel A of Table I, all the B values are negative even if the condition is violated. Moreover, a numerical plot when $\alpha = 0.5$ shows that $B < 0$ for any $\gamma > 0$; see Figure 4(a).

I.¹⁰

[Insert Table I here]

[Insert Figure 4 here]

III. Empirical Analysis

Our empirical analysis tests our preceding theory regarding how probability weighting impacts the risk premiums on covariance and coskewness. To this end, we first identify the characteristics and level of probability weighting from option data. Then, we provide one-way sort and two-way sort portfolio analyse to study the effect of probability weighting on the mean-covariance and mean-coskewness relationships, respectively. Finally, we run a Fama-MacBeth regression to investigate the impact of probability weighting on cross-sectional asset pricing.

A. *Option-implied Probability Weighting Functions and Indices*

Option data have been used to estimate empirical pricing kernels (see, e.g., Rosenberg and Engle, 2002). As the pricing kernel under RDUT is the product of the marginal utility and the derivative of probability weighting (Theorem 3.1), upon specifying a suitable utility function we can obtain implied probability weighting functions from option price data.

We use S&P 500 index option data and S&P 500 index data to derive the option-implied probability weighting functions and to construct the corresponding probability weighting indices for any given period of time (monthly in this paper). These indices are then used to gauge the overall probability weighting level in the market for the period concerned. The

¹⁰Again, the condition $\gamma \geq \gamma^*$ in Proposition 2-(iii) is only a technical condition used in the proof. In Panel C of Table I we observe positive B when $\gamma < \gamma^*$.

option data and index data are obtained from the OptionMetrics daily files, with a time window from January 4, 1996 to April 29, 2016.¹¹

More specifically, following the procedure of Polkovnichenko and Zhao (2013), with some variations, we derive implied probability weighting functions in three steps. First, at the end of each month during the period January 4, 1996 to April 29, 2016, we use both the S&P 500 index and index option data to construct the implied volatility curves for the options with the nearest maturity and the second nearest maturity. We then compute the implied volatility curve for the options with one-month maturity by spline interpolation. Next, compute the one-month maturity option prices based on the implied volatility curve and extract the nonparametric risk-neutral probability density function for the one-month return of the S&P 500 index, $f_{\tilde{R}_M}^Q$, by taking the second-order derivative of the one-month maturity option prices with respect to the strike prices (see Bliss and Panigirtzoglou, 2002 and Kostakis, Panigirtzoglou, and Skiadopoulos, 2011 for details).¹² Second, at the end of each month, we estimate an EGARCH model for the S&P 500 index with 1250 daily S&P 500 index return data up to the end of the month. Then, simulate 100,000 samples based on the EGARCH model and compute the nonparametric real-world probability density function for the one-month return of the S&P 500 index, $f_{\tilde{R}_M}$. Third, we estimate the probability

¹¹OptionMetrics is a provider of historical options data for empirical and econometric research. We access OptionMetrics via Wharton Research Data Services (WRDS) at <http://www.whartonwrds.com>.

¹²Different from Polkovnichenko and Zhao (2013), who directly select monthly option quotes closest to 28, 45, and 56 days from expiration and compute the risk-neutral probability density function for the 28, 45, and 56-day returns, we focus on the last trading day of each month. As there are no traded options with one-month maturity at the end of each month in our experiments, we need to numerically compute the prices of one-month maturity options based on the options with other maturities. Therefore, we apply the nonparametric method in Bliss and Panigirtzoglou (2002) and Kostakis et al. (2011) to derive the risk-neutral probability density function, which is more suitable for our analysis.

weighting function according to the following equation:

$$\begin{aligned}
w(1-p) &= \int_p^1 w'(1-y)dy = \int_{F_{\tilde{R}_M}^{-1}(p)}^{+\infty} w'(1-F_{\tilde{R}_M}(x))f_{\tilde{R}_M}(x)dx \\
&= \int_{F_{\tilde{R}_M}^{-1}(p)}^{+\infty} \frac{m(x)}{(1+r_f)\mathbb{E}\left[w'(1-F_{\tilde{R}_M}(\tilde{R}_M))u'(\tilde{R}_M)\right]u'(x)}f_{\tilde{R}_M}(x)dx \quad (9) \\
&= c \int_{F_{\tilde{R}_M}^{-1}(p)}^{+\infty} \frac{f_{\tilde{R}_M}^Q(x)}{u'(x)}dx,
\end{aligned}$$

where the function m is defined in (3). The constant c can be determined, by noticing $w(1) = 1$, as

$$c = \left(\int_0^{+\infty} \frac{f_{\tilde{R}_M}^Q(x)}{u'(x)}dx \right)^{-1}.$$

Based on this we derive w numerically by specifying the utility function to be CRRA with $\gamma = 2$.¹³ The solid red lines in Figure 5 show two typical types of implied probability weighting functions derived in our study, an inverse S-shaped one (overweighting both tails) and an S-shaped one (underweighting both tails). Among the total 244 months tested in our study, most of the time (about 89.3%) the implied weighting functions are either S-shaped or inverse S-shaped, an observation that reconciles with the well-documented result that individuals tend to underweight or overweight the two tails *simultaneously*.

[Insert Figure 5 here]

Probability weighting is captured by the derivative of the weighting function w , mainly around the two ends of its domain, $p = 1$ and $p = 0$. We propose two indices, those of “implied fear” and “implied hope”, capture the representative agent’s biases toward extremely

¹³Equation (9) is similar to equation (24) in Polkovnichenko and Zhao (2013), where the same utility function is taken.

good and bad events:

$$\text{Implied fear index (IFI)} = \sum_{i=95}^{99} 0.2w'(0.01i),$$

$$\text{Implied hope index (IHI)} = \sum_{i=1}^5 0.2w'(0.01i),$$

which are the average values of the derivatives of the probability weighting function around 1 and 0, respectively. When $\text{IFI} > 1$ ($\text{IHI} > 1$), the agent overweights extreme bad (good) events. Figure 6 shows the monthly time series of these indices. Evidently, the two indices have very similar dynamics and move together most of the time (their correlation coefficient is 0.8). This reiterates the aforementioned simultaneous underweighting/overweighting of the two tails.

[Insert Figure 6 here]

Given the highly correlated dynamics of IFI and IHI, we use their average as an option-implied probability weighting index to reflect the level of probability weighting:

$$PWI_1 = \frac{\text{IFI} + \text{IHI}}{2}. \quad (10)$$

Figure 7(a) depicts the movement of PWI_1 during the testing period.

To cross examine how robustly PWI_1 captures the degree of probability weighting in the market over time, we introduce a second probability weighting index by fitting the parametric specification of Prelec's weighting function. Specifically, we calibrate Prelec's parameter α by minimizing the Euclidean distance between Prelec's weighting function and the obtained nonparametric weighting function:

$$\hat{\alpha} = \arg \min_{\alpha} \sum_{i=1}^{99} |w(0.01i) - \exp(-(-\log(0.01i))^{\alpha})|^2,$$

which we call the implied α . Note that for Prelec's weighting function, $0 < \alpha < 1$ corresponds

to inverse S-shaped (overweighting) and $\alpha > 1$ to S-shaped (underweighting), with $\alpha = 1$ dividing the two; see also Figure 1. The greater (smaller) α , the greater the degree of overweighting (underweighting). This suggests that the following index, PWI_2 , also reflects the level of probability weighting:

$$PWI_2 = \frac{1}{\hat{\alpha}}. \quad (11)$$

[Insert Table II here]

Table II reports the summary statistics of IFI, IHI, $\hat{\alpha}$, PWI_1 and PWI_2 . The percentages of time when these indices indicate the occurrence of overweighting are remarkably close to each other: investors overweight extreme events 56%-59% of the time and underweight extreme events 41%-44% of the time. Moreover, Figure 7 shows that PWI_1 and PWI_2 have very similar dynamics. Indeed, they have a correlation coefficient 0.949. Thus, in the following analysis, we focus on PWI_1 , and we use it to separate the whole sample period (from January 1996 to April 2016) into two categories: overweighting periods and underweighting periods. Specifically, we calculate the median level of PWI_1 over the entire period. When the PWI_1 value at the end of month t is above its median level, the next month $t + 1$ is labelled as an overweighting period; otherwise, $t + 1$ is an underweighting period.¹⁴

[Insert Figure 7 here]

¹⁴Given that the percentages of overweighting periods and underweighting periods are 56%-59% and 41%-44% respectively as shown in Table II, it is reasonable to use the median as the separating point. We have also used the value 1 to separate these periods, and we found the main results to be the same.

B. One-way Sort Portfolio Analysis

We now conduct one-way sort portfolio empirical analysis to test the mean-covariance and mean-coskewness relationships during different probability weighting periods. In our study, we use all the common stocks listed on the NYSE, AMEX, and NASDAQ to perform cross-sectional analysis. Data on prices, returns, and shares outstanding are obtained from the Centre for Research in Security Prices (CRSP) monthly files, which range from January 1991 to April 2016.¹⁵

At the end of each month t from January 1991 to April 2016, we use the immediate prior 60-month sample covariance and sample coskewness to estimate the pre-ranking covariance and coskewness in month t ,

$$\begin{aligned}\text{covar}_{i,t} &= \frac{1}{59} \sum_{j=0}^{59} (\tilde{r}_{i,t-j} - \bar{r}_i)(\tilde{r}_{M,t-j} - \bar{r}_M), \\ \text{coskew}_{i,t} &= \frac{1}{59} \sum_{j=0}^{59} (\tilde{r}_{i,t-j} - \bar{r}_i)(\tilde{r}_{M,t-j}^2 - \bar{r}_M^2),\end{aligned}$$

where $\tilde{r}_{i,t-j}$ is the monthly return rate of stock i in month $t-j$, $\tilde{r}_{M,t-j}$ is the monthly return rate of the value-weighted market portfolio in month $t-j$, and \bar{r}_i , \bar{r}_M and \bar{r}_M^2 are the sample means of $\tilde{r}_{i,t}$, $\tilde{r}_{M,t}$, and $\tilde{r}_{M,t}^2$, respectively.

At each t , we classify stocks into 10 covariance-sorted groups based on their pre-ranking covariance. Group 1 (10) contains stocks with the lowest (highest) covariance. For each group, we build an equal-weighted portfolio. Then, we construct a spread portfolio (10-1) that longs Portfolio 10 and shorts Portfolio 1. We then calculate the next month ($t+1$)'s post-ranking portfolio returns for these portfolios and repeat the procedure for the next month. Similarly, we construct the coskewness-sorted portfolios.

Table III reports the average next month's returns of covariance-sorted portfolios and coskewness-sorted portfolios for the entire period, overweighting periods and underweighting

¹⁵CRSP is a provider of historical options data for use in empirical research and econometric studies. We access CRSP via Wharton Research Data Services (WRDS) at <http://www.whartonwrds.com>.

periods, respectively, where the overweighting periods and underweighting periods are determined by the index PWI_1 , as described earlier. We observe that the spread covariance-sorted portfolio (10-1) has a significantly positive monthly return during the overweighting periods and a significantly negative monthly return during the underweighting periods. Meanwhile, the spread coskewness-sorted portfolio (10-1) has a significantly negative monthly return during the overweighting periods and an insignificantly negative monthly return during the underweighting periods. Moreover, the negative return of this spread portfolio nearly doubles its size from an average of -1.01% monthly over the entire period to an average of -1.81% monthly during the overweighting periods.

[Insert Table III here]

Figures 8 (a) and (b) plot the regression lines of the average next month's returns on the average covariance and on the average coskewness, respectively. The return-covariance line has a steep upward (downward) slope during the overweighting (underweighting) periods. The return-coskewness line has a steep downward slope during the overweighting periods and a nearly flat slope during the underweighting periods. These findings are consistent with our theoretical predictions in Section II.B. Finally, for the entire sample period, the return-covariance line is almost flat, reconciling with the beta anomaly, and the return-coskewness line is downward-sloping, conforming to the skewness-inclined preference that is well documented in the literature.

[Insert Figure 8 here]

C. *Two-way Sort Portfolio Analysis*

Next, we conduct two-way sort portfolio analysis to separate the effects of covariance and coskewness. For each month t from January 1996 to April 2016, we classify all the stocks into five coskewness-sorted groups based on their pre-ranking coskewness. Then, in

each coskewness-sorted group, we further group the stocks into five covariance-sorted groups based on their pre-ranking covariance. Thus, there are now a total of 25 groups. For each group, we construct the equal-weighted portfolio, calculate next month $(t+1)$'s post-ranking portfolio returns, and repeat the procedure from the next month onwards. As the stocks have similar values of coskewness within the same coskewness-sorted group, we can thereby control the influence of the coskewness and compare the performance of the five covariance-sorted portfolios.

Table IV reports the average next month's returns for the entire period, overweighting periods and underweighting periods, respectively, where the overweighting periods and underweighting periods are determined by PWI_1 . We can see that with the coskewness controlled, the spread portfolios (5-1) still have significantly positive monthly returns during the overweighting periods and significantly negative monthly returns during the underweighting periods.

[Insert Table IV here]

Similarly, we analyze the 25 covariance-coskewness-sorted portfolios based first on their pre-ranking covariance and then on their pre-ranking coskewness, and report the results in Table V. Again, with covariance controlled, the spread portfolios (5-1) have significantly negative monthly returns during the overweighting periods and insignificantly monthly returns during the underweighting periods.

[Insert Table V here]

Our sort portfolio analysis suggests that one can significantly improve the performance of the long-short covariance-sorted and coskewness-sorted strategies by making use of the additional information from the probability weighting indices proposed in this paper. It confirms that probability weighting is an important factor for pricing securities and building profitable portfolios.

D. Fama-MacBeth Cross-Sectional Regression

We conduct a Fama-MacBeth regression (Fama and MacBeth, 1973) to investigate the impact of probability weighting on cross-sectional asset pricing. Following Bali, Engle, and Murray (2016), we first compute the covariance and coskewness of risky stock i for each month t from January 1996 to April 2016, by means of a 60-month-window regression:

$$\tilde{r}_{i,\tau} = k_{i,t} + \text{covariance}_{i,t} \tilde{r}_{M,\tau} + \text{coskewness}_{i,t} \tilde{r}_{M,\tau}^2 + \epsilon_{i,\tau}, \quad \tau = t - 59, \dots, t,$$

where $k_{i,t}$ is the intercept and $\epsilon_{i,\tau}$ the residual. In contrast to computing the covariance and coskewness from the samples directly, as in Subsection III.B and III.C, the above regression rules out interactions between covariance and coskewness. Next, for each t , we consider the following cross-sectional regression:

$$\tilde{r}_{i,t+1} = k_t + A_t \text{covariance}_{i,t} + B_t \text{coskewness}_{i,t}, \quad i = 1, \dots, n.$$

The significance of the risk premiums A and B is determined by a t -test on the time series $\{A_t\}$ and $\{B_t\}$.¹⁶

Table VI reports the Fama-MacBeth regression results, in which the market overweighting and underweighting periods are divided by the index PWI_1 . If we do not separate the time periods based on probability weighting, the overall risk premium of covariance, A , is not significant and the risk premium of coskewness, B , is modestly negative. These are consistent with the empirical findings in the literature. When we *do* separate the whole period

¹⁶Besides Bali et al. (2016)'s approach, Harvey and Siddique (2000) propose to compute $\text{covariance}_{i,t}$ and $\text{coskewness}_{i,t}$ by the following:

$$\begin{aligned} \tilde{r}_{i,\tau} &= k_{i,t} + \text{covariance}_{i,t} \tilde{r}_{M,\tau} + \epsilon_{i,\tau}, \quad \tau = t - 59, \dots, t, \\ \text{coskewness}_{i,t} &= \frac{\frac{1}{59} \sum_{j=0}^{59} (\epsilon_{i,t-j}) (\tilde{r}_{M,t-j}^2 - \bar{r}_M^2)}{\sqrt{\frac{1}{60} \sum_{j=0}^{59} \epsilon_{i,t-j}^2 \cdot \frac{1}{59} \sum_{j=0}^{59} (\tilde{r}_{M,t-j} - \bar{r}_M)^2}}, \end{aligned}$$

where \bar{r}_M and \bar{r}_M^2 are the sample means of $\tilde{r}_{M,\tau}$ and $\tilde{r}_{M,\tau}^2$, respectively. We obtain the same results based on this method (omitted here due to space).

into two regimes (overweighting and underweighting), however, A becomes significantly positive during overweighting periods and significantly negative during underweighting periods, and B almost doubles its overall value during overweighting periods yet becomes insignificant during underweighting periods. Again, these results reaffirm our theoretical and prior empirical findings.

[Insert Table VI here]

Finally, we note that during overweighting periods, the mean-covariance tradeoff and mean-coskewness tradeoff are statistically significant. Indeed, Table VII shows that the standard deviation of covariance is 0.8586 and that of coskewness is 10.8123. Thus, during an overweighting period, a one-standard-deviation increase in covariance is associated with a roughly 0.76% (the standard deviation of covariance times the coefficient A , i.e., $0.8586 \times 8.888 \times 10^{-3}$) increase in expected monthly return, and a one-standard-deviation decrease in coskewness is associated with a roughly 0.89% (the standard deviation of coskewness times the coefficient B , i.e., $(-10.8123) \times (-0.823) \times 10^{-3}$) increase in expected monthly return. This finding strengthens the notion that probability weighting has important implications in pricing and investment.

[Insert Table VII here]

E. Robustness of Results with Other Control Variables

Finally, we conduct a Fama-MacBeth regression to test the robustness of our main findings after controlling six other variables. These control variables are Size ($Size$), Book-to-market ratio (BM), Momentum (Mom), Short-term reversal (Rev), Idiosyncratic volatility ($IVol$), and Idiosyncratic skewness ($ISkew$). Among them, the first four are well-studied factors long known to be relevant in asset pricing, and the last two are relatively new in the literature. Liu et al. (2018) argue that the beta anomaly arises from beta's positive correla-

tion with idiosyncratic volatility (IVol). Barberis and Huang (2008) find that CPT investors are willing to pay more for lottery-like stocks. This type of stocks has high idiosyncratic skewness, which could be largely independent of the co-skewness with the market. If probability weighting has an impact on the latter, as we have demonstrated, then, very likely, it has an impact on the former as well.

Following Bali et al. (2016), these control variables are computed as follows

$$\begin{aligned}
Size_{i,t} &= \log \left(\frac{PRC_{i,t} \cdot SHROUT_{i,t}}{1000} \right), \\
BM_{i,t} &= \log \left(\frac{1000 \cdot BE_{i,y}}{PRC_{i,y} \cdot SHROUT_{i,y}} \right), \\
Mom_{i,t} &= 100 \left[\prod_{j=1}^{11} (\tilde{r}_{i,t-j} + 1) - 1 \right], \\
Rev_{i,t} &= 100 \cdot \tilde{r}_{i,t}, \\
IVol_{i,t} &= 100 \cdot \sqrt{\frac{1}{56} \sum_{j=0}^{59} \varepsilon_{i,t-j}^2} \cdot \sqrt{12}, \\
ISkew_{i,t} &= \frac{\frac{1}{60} \sum_{j=0}^{59} \varepsilon_{i,t-j}^3}{\left(\frac{1}{60} \sum_{j=0}^{59} \varepsilon_{i,t-j}^2 \right)^{3/2}},
\end{aligned}$$

where y is the end of a fiscal year, t is the end of a month between June of year $y + 1$, and May of year $y + 2$, $PRC_{i,t}$ is the price of stock i at t , $SHROUT_{i,t}$ is the number of shares outstanding of stock i at t , $BE_{i,y}$ is the book value of common equity of stock i at y , $PRC_{i,y}$ is the price of stock i at y , $SHROUT_{i,y}$ is the number of shares outstanding of stock i at y , $\tilde{r}_{i,t-j}$ is the return rate of stock i during month $t - j$, and $\varepsilon_{i,t-j}$, $j = 0, \dots, 59$, are the residuals of the following regression:

$$\tilde{r}_{i,\tau} = k_{i,t} + \delta_{i,t}^1 MKT_{\tau} + \delta_{i,t}^2 SMB_{\tau} + \delta_{i,t}^3 HML_{\tau} + \varepsilon_{i,\tau}, \quad \tau = t - 59, \dots, t,$$

where MKT_{τ} is the return of the market factor during month τ and SMB_{τ} and HML_{τ} are the returns of the size and value factor mimicking portfolios, respectively, during month τ .

We consider three Fama-MacBeth regressions, $i = 1, \dots, n$,

$$\text{I} : \tilde{r}_{i,t+1} = \delta_{0,t} + A_t \text{covariance}_{i,t} + B_t \text{coskewness}_{i,t} + \delta_{3,t} \text{IVol}_{i,t} + \delta_{4,t} \text{ISkew}_{i,t},$$

$$\begin{aligned} \text{II} : \tilde{r}_{i,t+1} = & \delta_{0,t} + A_t \text{covariance}_{i,t} + B_t \text{coskewness}_{i,t} + \delta_{3,t} \text{IVol}_{i,t} + \delta_{4,t} \text{ISkew}_{i,t} \\ & + \delta_{5,t} \text{Size}_{i,t} + \delta_{6,t} \text{BM}_{i,t}, \end{aligned}$$

$$\begin{aligned} \text{III} : \tilde{r}_{i,t+1} = & \delta_{0,t} + A_t \text{covariance}_{i,t} + B_t \text{coskewness}_{i,t} + \delta_{3,t} \text{IVol}_{i,t} + \delta_{4,t} \text{ISkew}_{i,t} \\ & + \delta_{5,t} \text{Size}_{i,t} + \delta_{6,t} \text{BM}_{i,t} + \delta_{7,t} \text{Mom}_{i,t} + \delta_{8,t} \text{Rev}_{i,t}. \end{aligned}$$

The regression results, including the ones without these control variables, are reported in Table VIII. They show that even with the commonly used variables controlled in a cross-sectional analysis, our main empirical finding that the risk premium of covariance (coskewness) is significantly positive (negative) during overweighting periods and significantly negative (insignificant) during underweighting periods still stands.

It is interesting to note that, in Table VIII, IVol exhibits a significantly negative relationship with expected return only during underweighting periods. In other words, the mispricing (the negative risk premiums for beta and idiosyncratic volatility) only appears when investors underweight the tail risk, which echoes the finding in Liu et al. (2018). On the other hand, similar to coskewness, ISkew shows a significantly negative relation with expected return only during overweighting periods. This implies that overweighting both tails leads to RDUT investors developing a strong taste for both coskewness and idiosyncratic skewness.

[Insert Table VIII here]

IV. Conclusion

We show that probability weighting is a key driver in beta and coskewness pricing. Individuals tend to overweight or underweight the two tails simultaneously. Whether they overweight or underweight changes dynamically over time, but the total length of overweighting periods is comparable to that of underweighting periods. If we analyze a sufficiently long sample period, the aggregate effects of overweighting and underweighting may cancel out, potentially deceiving us into overlooking the significance of probability weighting. It is, therefore, important to *separate out* the overweighting and underweighting periods and to study them individually and respectively. In doing so, as demonstrated in this paper, we are able to understand the significant role that probability weighting plays in pricing and in building profitable portfolios.

We carry out the theoretical analysis using an RDUT model, and confirm its implications with empirical tests. We believe that our model offers a partial explanation of the beta anomaly and hints at a viable way to potentially resolve other puzzles in asset pricing.

Appendix A. Proof of Theorem 2

When the market is in equilibrium, the total return of the optimal portfolio of the representative investor equals the total return of the market portfolio, i.e., $\tilde{R}_P^* = \tilde{R}_M$. Applying Theorem 5.2 of Xia and Zhou (2016), we obtain

$$\tilde{m} = (\lambda^*)^{-1} w'(1 - F_{\tilde{R}_M}(\tilde{R}_M)) u'(\tilde{R}_M).$$

However, the pricing formula of the risk-free asset yields

$$1 = \mathbb{E}[\tilde{m}(1 + r_f)] = \mathbb{E}[\tilde{m}](1 + r_f),$$

which in turn implies

$$\lambda^* = (1 + r_f)\mathbb{E} \left[w'(1 - F_{\tilde{R}_M}(\tilde{R}_M))u'(\tilde{R}_M) \right]. \quad (\text{A1})$$

This establishes (4).

Next we rewrite

$$\tilde{m} = (\lambda^*)^{-1}w'(1 - F_{\tilde{R}_M}(1 + \tilde{r}_M))u'(1 + \tilde{r}_M).$$

Applying the Taylor expansion up to the second order of \tilde{r}_M , we can express the pricing kernel as follows:¹⁷

$$\begin{aligned} \tilde{m} &= (\lambda^*)^{-1}w'(1 - F_{\tilde{R}_M}(1))u'(1) \\ &\quad + (\lambda^*)^{-1} \left[w'(1 - F_{\tilde{R}_M}(1))u''(1) - w''(1 - F_{\tilde{R}_M}(1))u'(1)f_{\tilde{R}_M}(1) \right] \tilde{r}_M \\ &\quad + \frac{1}{2}(\lambda^*)^{-1} \left\{ w'(1 - F_{\tilde{R}_M}(1))u'''(1) - w''(1 - F_{\tilde{R}_M}(1))[2u''(1)f_{\tilde{R}_M}(1) + u'(1)f'_{\tilde{R}_M}(1)] \right. \\ &\quad \left. + w'''(1 - F_{\tilde{R}_M}(1))u'(1)f_{\tilde{R}_M}^2(1) \right\} \tilde{r}_M^2 + o(\tilde{r}_M^2). \end{aligned} \quad (\text{A2})$$

Applying the pricing formula to the risky assets, we have

$$\mathbb{E}[\tilde{m}(1 + \tilde{r}_i)] = 1, \quad i = 2, 3, \dots, n.$$

Hence,

$$\text{Cov}(\tilde{m}, (1 + \tilde{r}_i)) + \mathbb{E}[1 + \tilde{r}_i]\mathbb{E}[\tilde{m}] = 1,$$

¹⁷Here, we Taylor expand $F_{\tilde{R}_M}(1 + \tilde{r}_M)$ around 1, then expand $w'(1 - F_{\tilde{R}_M}(1 + \tilde{r}_M))$ around $1 - F_{\tilde{R}_M}(1)$, and expand $u'(1 + \tilde{r}_M)$ around 1.

or

$$\mathbb{E}[\tilde{r}_i] = \frac{1 - \text{Cov}(\tilde{m}, \tilde{r}_i)}{\mathbb{E}[\tilde{m}]} - 1 = r_f + A\text{Cov}(\tilde{r}_i, \tilde{r}_M) + \frac{1}{2}B\text{Cov}(\tilde{r}_i, \tilde{r}_M^2) + o(\tilde{r}_M^2),$$

thanks to (A1) and (A2).

Appendix B. Proof of Proposition 1

i) When $\alpha < 1$, w is inverse S-shaped. Hence, $1 - F_{\tilde{R}_M}(1)$ lies within the convex domain of w under the assumptions of the proposition. Therefore, the two terms inside the brackets of (6) are both negative. This implies $A > 0$.¹⁸

ii) When $\alpha > 1$, w is S-shaped, and $1 - F_{\tilde{R}_M}(1)$ lies within the concave domain of w . With $w(p) = \exp(-(-\log(p))^\alpha)$, we have

$$\begin{aligned} A &\equiv - \frac{w'(1 - F_{\tilde{R}_M}(1))}{\mathbb{E}[w'(1 - F_{\tilde{R}_M}(\tilde{R}_M))u'(\tilde{R}_M)]} \left[u''(1) - \frac{w''(1 - F_{\tilde{R}_M}(1))}{w'(1 - F_{\tilde{R}_M}(1))} u'(1) f_{\tilde{R}_M}(1) \right] \\ &= - \frac{w'(1 - F_{\tilde{R}_M}(1))}{\mathbb{E}[w'(1 - F_{\tilde{R}_M}(\tilde{R}_M))u'(\tilde{R}_M)]} \left[u''(1) - \frac{\alpha y_0^\alpha - \alpha + 1 - y_0}{(1 - F_{\tilde{R}_M}(1))y_0} u'(1) f_{\tilde{R}_M}(1) \right], \end{aligned}$$

where $y_0 = -\log(1 - F_{\tilde{R}_M}(1))$. It follows from the assumptions of the proposition that $0 \leq y_0 < 1$, which implies $\lim_{\alpha \rightarrow +\infty} \alpha y_0^\alpha - \alpha = -\infty$. Hence,

$$\lim_{\alpha \rightarrow +\infty} - \frac{\alpha y_0^\alpha - \alpha + 1 - y_0}{(1 - F_{\tilde{R}_M}(1))y_0} u'(1) f_{\tilde{R}_M}(1) = +\infty.$$

Thus, there must exist $\alpha^* > 1$ such that when $\alpha > \alpha^*$, $-\frac{\alpha y_0^\alpha - \alpha + 1 - y_0}{(1 - F_{\tilde{R}_M}(1))y_0} u'(1) f_{\tilde{R}_M}(1) > -u''(1)$, implying $A < 0$.

¹⁸This part of the proof does not depend on the specific form of w ; it only requires $1 - F_{\tilde{R}_M}(1)$ to be in the convex domain of w .

Appendix C. Proof of Proposition 2

With $w(p) = \exp(-(-\log(p))^\alpha)$, we have

$$\begin{aligned}
 B &= - \frac{w'(1 - F_{\tilde{R}_M}(1))u'(1)}{\mathbb{E}[w'(1 - F_{\tilde{R}_M}(\tilde{R}_M))u'(\tilde{R}_M)]} \left[\gamma(\gamma + 1) \right. \\
 &\quad - \frac{\alpha y_0^\alpha - \alpha + 1 - y_0}{(1 - F_{\tilde{R}_M}(1))y_0} (-2\gamma f_{\tilde{R}_M}(1) + f'_{\tilde{R}_M}(1)) \\
 &\quad \left. + \frac{(\alpha y_0^\alpha - \alpha + 1 - y_0)^2 + (1 - \alpha) + (1 - \alpha - y_0)(\alpha y_0^\alpha - y_0)}{(1 - F_{\tilde{R}_M}(1))^2 y_0^2} f_{\tilde{R}_M}^2(1) \right] \\
 &:= - \frac{w'(1 - F_{\tilde{R}_M}(1))u'(1)}{\mathbb{E}[w'(1 - F_{\tilde{R}_M}(\tilde{R}_M))u'(\tilde{R}_M)]} [B_1 + B_2 + B_3],
 \end{aligned}$$

where B_1, B_2, B_3 denote the three terms in the brackets, and $y_0 = -\log(1 - F_{\tilde{R}_M}(1))$.

Clearly, $B_1 > 0$.

i) When $0 < \alpha < 1$ and $\gamma \geq \frac{f'_{\tilde{R}_M}(1)}{2f_{\tilde{R}_M}(1)}$, we have $B_2 \geq 0$. We now examine B_3 . It follows from $1 - F_{\tilde{R}_M}(1) > e^{-1}$ that $0 \leq y_0 < 1$. On the other hand, the first-order condition of the function $(\alpha y_0^\alpha - y_0)$ is

$$\frac{\partial}{\partial y_0} (\alpha y_0^\alpha - y_0) = \alpha^2 y_0^{\alpha-1} - 1 = 0,$$

whose solution is $y_0^* = \alpha^{\frac{2}{1-\alpha}} \in (0, 1)$. Then, the extreme values of this function on $y_0 \in [0, 1]$ are

$$\max_{y_0 \in [0, 1]} (\alpha y_0^\alpha - y_0) = \max \{ \alpha 0^\alpha - 0, \alpha 1^\alpha - 1, \alpha^{\frac{1+\alpha}{1-\alpha}} - \alpha^{\frac{2}{1-\alpha}} \} = \alpha^{\frac{1+\alpha}{1-\alpha}} - \alpha^{\frac{2}{1-\alpha}} > 0,$$

$$\min_{y_0 \in [0, 1]} (\alpha y_0^\alpha - y_0) = \min \{ \alpha 0^\alpha - 0, \alpha 1^\alpha - 1, \alpha^{\frac{1+\alpha}{1-\alpha}} - \alpha^{\frac{2}{1-\alpha}} \} = \alpha - 1 < 0.$$

When $1 - \alpha - y_0 > 0$, we have

$$(1 - \alpha) + (1 - \alpha - y_0)(\alpha y_0^\alpha - y_0) \geq (1 - \alpha) + (1 - \alpha - 0)(\alpha - 1) = (1 - \alpha)\alpha > 0.$$

When $1 - \alpha - y_0 < 0$, we have

$$\begin{aligned}
& (1 - \alpha) + (1 - \alpha - y_0)(\alpha y_0^\alpha - y_0) \\
& \geq (1 - \alpha) + (1 - \alpha - 1)(\alpha^{\frac{1+\alpha}{1-\alpha}} - \alpha^{\frac{2}{1-\alpha}}) \\
& = (1 - \alpha^{\frac{2}{1-\alpha}})(1 - \alpha) \\
& > 0.
\end{aligned}$$

When $1 - \alpha - y_0 = 0$, we have

$$(1 - \alpha) + (1 - \alpha - y_0)(\alpha y_0^\alpha - y_0) = (1 - \alpha) > 0.$$

This establishes that $B_3 > 0$, leading to $B < 0$. (Here we have actually proved that, in this case, $w'''(1 - F_{\tilde{R}_M}(1)) > 0$.)

ii) When $\alpha > 1$, we consider $B_1 + B_2 + B_3$ as a function of γ and α :

$$B_1 + B_2 + B_3 = \gamma^2 + (1 + 2g_1(\alpha)f_{\tilde{R}_M}(1))\gamma - g_1(\alpha)f'_{\tilde{R}_M}(1) + g_2(\alpha)f_{\tilde{R}_M}^2(1),$$

where $g_1(\alpha)$ and $g_2(\alpha)$ are continuous functions of α :

$$\begin{aligned}
g_1(\alpha) & := \frac{\alpha y_0^\alpha - \alpha + 1 - y_0}{(1 - F_{\tilde{R}_M}(1))y_0}, \\
g_2(\alpha) & := \frac{(\alpha y_0^\alpha - \alpha + 1 - y_0)^2 + (1 - \alpha) + (1 - \alpha - y_0)(\alpha y_0^\alpha - y_0)}{(1 - F_{\tilde{R}_M}(1))^2 y_0^2}.
\end{aligned}$$

It is easy to show that, for any given finite $\alpha > 1$, there exists a threshold

$$\bar{\gamma}(\alpha) := \begin{cases} \frac{-(1+2g_1(\alpha)f_{\tilde{R}_M}(1)) + \sqrt{(1+2g_1(\alpha)f_{\tilde{R}_M}(1))^2 - 4(-g_1(\alpha)f'_{\tilde{R}_M}(1) + g_2(\alpha)f_{\tilde{R}_M}^2(1))}}{2}, & \text{if } (1 + 2g_1(\alpha)f_{\tilde{R}_M}(1))^2 - 4(-g_1(\alpha)f'_{\tilde{R}_M}(1) + g_2(\alpha)f_{\tilde{R}_M}^2(1)) \geq 0, \\ 0, & \text{if } (1 + 2g_1(\alpha)f_{\tilde{R}_M}(1))^2 - 4(-g_1(\alpha)f'_{\tilde{R}_M}(1) + g_2(\alpha)f_{\tilde{R}_M}^2(1)) < 0, \end{cases}$$

such that $B_1 + B_2 + B_3 > 0$ (and, hence, $B < 0$) whenever $\gamma > \bar{\gamma}(\alpha)$.

iii) As $0 \leq y_0 < 1$ and $1 - F_{\tilde{R}_M}(1)$ lies within the concave domain of w , we have

$$\alpha y_0^\alpha - \alpha + 1 - y_0 \leq 0.$$

Thus, $B_2 \leq 0$ when $\gamma \geq \gamma^*$. Now,

$$B_1 + B_3 = \gamma(\gamma + 1) + g_2(\alpha) f_{\tilde{R}_M}^2(1).$$

It is easy to prove that, for any given finite $\alpha > 1$, there must exist a threshold

$$\underline{\gamma}(\alpha) := \begin{cases} \frac{\sqrt{1 - g_2(\alpha) f_{\tilde{R}_M}^2(1)} - 1}{2}, & \text{if } g_2(\alpha) \leq 0, \\ 0, & \text{if } g_2(\alpha) > 0, \end{cases}$$

such that $B_1 + B_3 < 0$ when $0 \leq \gamma < \underline{\gamma}(\alpha)$. This completes the proof.

REFERENCES

- Abdellaoui, Mohammed, 2002, A genuine rank-dependent generalization of the von neumann-morgenstern expected utility theorem, *Econometrica* 70, 717–736.
- Antoniou, Constantinos, John A Doukas, and Avanidhar Subrahmanyam, 2015, Investor sentiment, beta, and the cost of equity capital, *Management Science* 62, 347–367.
- Baele, Lieven, Joost Driessen, Sebastian Ebert, Juan M Londono, and Oliver G Spalt, 2018, Cumulative prospect theory, option returns, and the variance premium, *The Review of Financial Studies* 32, 3667–3723.
- Baker, Malcolm, Brendan Bradley, and Jeffrey Wurgler, 2011, Benchmarks as limits to arbitrage: Understanding the low-volatility anomaly, *Financial Analysts Journal* 67, 40–54.
- Bali, Turan G, Robert F Engle, and Scott Murray, 2016, *Empirical asset pricing: the cross section of stock returns* (John Wiley & Sons).
- Barberis, Nicholas, 2013, The psychology of tail events: Progress and challenges, *American Economic Review* 103, 611–16.
- Barberis, Nicholas, and Ming Huang, 2008, Stocks as lotteries: The implications of probability weighting for security prices, *The American Economic Review* 98, 2066–2100.
- Black, Fischer, Michael C Jensen, and Myron Scholes, 1972, *The capital asset pricing model: Some empirical tests*, 9–124 (Praeger Publishers Inc.).
- Bliss, Robert R, and Nikolaos Panigirtzoglou, 2002, Testing the stability of implied probability density functions, *Journal of Banking & Finance* 26, 381–422.
- Bollerslev, Tim, Viktor Todorov, and Lai Xu, 2015, Tail risk premia and return predictability, *Journal of Financial Economics* 118, 113–134.

- Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer, 2012, Saliency theory of choice under risk, *The Quarterly Journal of Economics* 127, 1243–1285.
- Brockett, Patrick L, and Linda L Golden, 1987, A class of utility functions containing all the common utility functions, *Management Science* 33, 955–964.
- De Giorgi, Enrico G, and Shane Legg, 2012, Dynamic portfolio choice and asset pricing with narrow framing and probability weighting, *Journal of Economic Dynamics and Control* 36, 951–972.
- Fama, Eugene F, and Kenneth R French, 1992, The cross-section of expected stock returns, *The Journal of Finance* 47, 427–465.
- Fama, Eugene F, and James D MacBeth, 1973, Risk, return, and equilibrium: Empirical tests, *Journal of Political Economy* 81, 607–636.
- Gonzalez, Richard, and George Wu, 1999, On the shape of the probability weighting function, *Cognitive psychology* 38, 129–166.
- Harrison, Glenn W, Steven J Humphrey, and Arjan Verschoor, 2009, Choice under uncertainty: evidence from ethiopia, india and uganda, *The Economic Journal* 120, 80–104.
- Harvey, Campbell R, and Akhtar Siddique, 2000, Conditional skewness in asset pricing tests, *The Journal of Finance* 55, 1263–1295.
- Henrich, Joseph, Steven J Heine, and Ara Norenzayan, 2010, The weirdest people in the world?, *Behavioral and brain sciences* 33, 61–83.
- Hertwig, Ralph, Greg Barron, Elke U Weber, and Ido Erev, 2004, Decisions from experience and the effect of rare events in risky choice, *Psychological Science* 15, 534–539.
- Hong, Harrison, and David A Sraer, 2016, Speculative betas, *The Journal of Finance* 71, 2095–2144.

- Hsu, Ming, Ian Krajbich, Chen Zhao, and Colin F Camerer, 2009, Neural response to reward anticipation under risk is nonlinear in probabilities, *Journal of Neuroscience* 29, 2231–2237.
- Humphrey, Steven J, and Arjan Verschoor, 2004, The probability weighting function: experimental evidence from uganda, india and ethiopia, *Economics Letters* 84, 419–425.
- Jagannathan, Ravi, and Zhenyu Wang, 1996, The conditional capm and the cross-section of expected returns, *The Journal of Finance* 51, 3–53.
- Kelly, Bryan, and Hao Jiang, 2014, Tail risk and asset prices, *The Review of Financial Studies* 27, 2841–2871.
- Kostakis, Alexandros, Nikolaos Panigirtzoglou, and George Skiadopoulos, 2011, Market timing with option-implied distributions: A forward-looking approach, *Management Science* 57, 1231–1249.
- Kozhan, Roman, Anthony Neuberger, and Paul Schneider, 2013, The skew risk premium in the equity index market, *The Review of Financial Studies* 26, 2174–2203.
- Kraus, Alan, and Robert H Litzenberger, 1976, Skewness preference and the valuation of risk assets, *The Journal of Finance* 31, 1085–1100.
- Lintner, John, 1965, Security prices, risk, and maximal gains from diversification, *The Journal of Finance* 20, 587–615.
- Liu, Jianan, Robert F Stambaugh, and Yu Yuan, 2018, Absolving beta of volatility’s effects, *Journal of Financial Economics* 128, 1–15.
- Mehra, Rajnish, and Edward C Prescott, 1985, The equity premium: A puzzle, *Journal of Monetary Economics* 15, 145–161.
- Mossin, Jan, 1966, Equilibrium in a capital asset market, *Econometrica* 34, 768–783.

- Polkovnichenko, Valery, and Feng Zhao, 2013, Probability weighting functions implied in options prices, *Journal of Financial Economics* 107, 580–609.
- Prelec, Drazen, 1998, The probability weighting function, *Econometrica* 66, 497–527.
- Quiggin, John, 1982, A theory of anticipated utility, *Journal of Economic Behavior & Organization* 3, 323–343.
- Quiggin, John, 2012, *Generalized expected utility theory: The rank-dependent model* (Springer Science & Business Media).
- Rosenberg, Joshua V, and Robert F Engle, 2002, Empirical pricing kernels, *Journal of Financial Economics* 64, 341–372.
- Schmeidler, David, 1989, Subjective probability and expected utility without additivity, *Econometrica* 57, 571–587.
- Sharpe, William F, 1964, Capital asset prices: A theory of market equilibrium under conditions of risk, *The Journal of Finance* 19, 425–442.
- Shefrin, Hersh, 2008, *A Behavioral Approach to Asset Pricing* (New York: Elsevier.).
- Tanaka, Tomomi, Colin F Camerer, and Quang Nguyen, 2010, Risk and time preferences: Linking experimental and household survey data from vietnam, *American Economic Review* 100, 557–71.
- Tversky, Amos, and Daniel Kahneman, 1992, Advances in prospect theory: Cumulative representation of uncertainty, *Journal of Risk and Uncertainty* 5, 297–323.
- Wu, George, and Richard Gonzalez, 1996, Curvature of the probability weighting function, *Management science* 42, 1676–1690.
- Xia, Jianming, and Xun Yu Zhou, 2016, Arrow–debreu equilibria for rank-dependent utilities, *Mathematical Finance* 26, 558–588.

Table I The values of B under different settings. The probability weighting function is that of Prelec with $\alpha = 0.5$, 1 and 1.5. The utility function is CRRA with various values of relative risk aversion parameter γ . The distribution of the market portfolio's total return is normal ($\mu = 1.076$, $\sigma = 0.158$), skew-normal ($\mu = 1.076$, $\sigma = 0.158$, $\text{skewness} = -0.339$), or log-normal ($\mu = 1.076$, $\sigma = 0.158$). The parameters γ^* , $\underline{\gamma}$ and $\bar{\gamma}$ are defined in Proposition 2 and its proof (Appendix C).

Panel A: Overweighting case ($\alpha = 0.5$)

Distribution of \tilde{R}_M	$\gamma = 1$	$\gamma = 2$	$\gamma = 3$	$\gamma = 4$	$\gamma = 5$
Normal $\gamma^* = 1.5222$	-34.6634	-39.9503	-43.4616	-44.6584	-43.4897
Skew-normal $\gamma^* = 2.3799$	-30.2979	-36.4819	-41.0653	-43.2547	-42.8122
Log-normal $\gamma^* = 0.7816$	-32.5639	-34.8329	-35.4924	-16.2108	-12.7188

Panel B: no probability weighting case ($\alpha = 1$)

Distribution of \tilde{R}_M	$\gamma = 1$	$\gamma = 2$	$\gamma = 3$	$\gamma = 4$	$\gamma = 5$
Normal $\gamma^* = 1.5222$	-2.1102	-6.5066	-13.0431	-21.2156	-30.1970
Skew-normal $\gamma^* = 2.3799$	-2.0897	-6.3884	-12.7085	-20.5003	-28.8738
Log-normal $\gamma^* = 0.7816$	-2.1385	-6.3540	-12.2682	-19.2142	-26.2404

Panel C: Underweighting case ($\alpha = 1.5$)

Distribution of \tilde{R}_M	$\gamma = 1$	$\gamma = 2$	$\gamma = 3$	$\gamma = 9$	$\gamma = 10$
Normal $\gamma^* = 1.5222$, $\gamma(\alpha) = 1.4765$, $\bar{\gamma}(\alpha) = 8.2577$	8.9976	15.8929	20.7699	-8.4330	-20.4957
Skew-normal $\gamma^* = 2.3799$, $\underline{\gamma}(\alpha) = 1.4129$, $\bar{\gamma}(\alpha) = 7.1000$	-1.3199	4.8151	9.0720	-16.7676	-26.2754
Log-normal $\gamma^* = 0.7816$, $\underline{\gamma}(\alpha) = 1.6202$, $\bar{\gamma}(\alpha) = 8.8817$	21.4745	28.4943	32.8405	-1.3287	-12.4790

Table II Summary statistics of IFI, IHI, implied α , PWI_1 , and PWI_2 . For implied α , the percentage of overweighting is that of the index being less than 1. For all the other indices, the percentage of overweighting is that of the corresponding index being greater than 1.

	Minimum	25th percentile	Median	75th percentile	Maximum	Percentage of overweighting
IFI	0.000	0.388	1.319	2.600	5.865	56.6
IHI	0.000	0.312	1.233	2.256	4.703	57.4
Implied α	0.375	0.698	0.895	1.203	2.470	58.6
PWI_1	0.000	0.401	1.512	2.447	5.103	56.5
PWI_2	0.405	0.831	1.166	1.434	2.667	58.6

Table III Average next month's returns of covariance-sorted portfolios and coskewness-sorted portfolios. The overweighting periods and underweighting periods are separated by the index PWI_1 . All periods (All), overweighting periods (Over), and underweighting periods (Under) include 244 months, 122 months, and 122 months, respectively. The spread portfolio (10-1) is constructed by longing the 10th portfolio and shorting the 1st portfolio. The t -statistics of the returns of the spread portfolios are reported. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Panel A: covariance-sorted portfolios												
	1	2	3	4	5	6	7	8	9	10	(10-1)	t -statistic of (10-1)
All	1.05%	1.16%	1.18%	1.21%	1.34%	1.27%	1.21%	1.33%	1.18%	1.18%	0.13%	(0.23)
Over	1.09%	1.16%	1.21%	1.25%	1.47%	1.63%	1.63%	2.08%	2.24%	2.75%	1.65%*	(1.77)
Under	1.01%	1.16%	1.16%	1.16%	1.22%	0.91%	0.80%	0.58%	0.12%	-0.40%	-1.40%**	(-2.48)

Panel B: coskewness-sorted portfolios												
	1	2	3	4	5	6	7	8	9	10	(10-1)	t -statistic of (10-1)
All	1.73%	1.46%	1.42%	1.40%	1.20%	1.14%	1.20%	0.93%	0.91%	0.72%	-1.01%***	(-3.04)
Over	2.85%	2.24%	1.97%	1.79%	1.47%	1.49%	1.53%	1.07%	1.05%	1.04%	-1.81%***	(-3.01)
Under	0.61%	0.67%	0.87%	1.01%	0.93%	0.79%	0.87%	0.80%	0.78%	0.39%	-0.21%	(-0.79)

Table IV Average next month's returns of coskewness-covariance-sorted portfolios. The overweighting periods (Over) and underweighting periods (Under) are separated by PWI_1 . The spread portfolio (5-1) is constructed by longing the 5th portfolio and shorting the 1st portfolio. The t -statistics of the returns of the spread portfolios are reported. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Periods	coskewness	covariance						t -statistic of (5-1)
		1	2	3	4	5	(5-1)	
All	1	1.12%	1.27%	1.33%	1.55%	1.43%	0.32%	(0.76)
	2	1.17%	1.27%	1.26%	1.35%	1.19%	0.01%	(0.03)
	3	1.30%	1.33%	1.37%	1.09%	1.03%	-0.27%	(-0.57)
	4	1.22%	1.18%	1.28%	1.19%	1.38%	0.17%	(0.32)
	5	1.00%	0.77%	1.10%	1.07%	1.02%	0.02%	(0.03)
Over	1	1.28%	1.38%	1.72%	2.19%	2.63%	1.34%*	(1.87)
	2	1.23%	1.18%	1.45%	2.16%	2.67%	1.44%*	(1.87)
	3	1.26%	1.43%	1.57%	1.61%	2.28%	1.02%	(1.32)
	4	1.21%	1.20%	1.46%	1.91%	2.79%	1.58%*	(1.83)
	5	0.87%	0.71%	1.34%	1.49%	2.23%	1.36%*	(1.81)
Under	1	0.95%	1.16%	0.95%	0.91%	0.24%	-0.71%*	(-1.77)
	2	1.12%	1.35%	1.07%	0.53%	-0.29%	-1.41%***	(-3.05)
	3	1.33%	1.23%	1.17%	0.57%	-0.23%	-1.56%***	(-2.91)
	4	1.22%	1.16%	1.11%	0.47%	-0.03%	-1.25%**	(-2.43)
	5	1.14%	0.83%	0.86%	0.65%	-0.20%	-1.33%**	(-2.58)

Table V Average next month's returns of covariance-coskewness-sorted portfolios. The overweighting periods (Over) and underweighting periods (Under) are separated by PWI_1 . The spread portfolio (5-1) is constructed by longing the 5th portfolio and shorting the 1st portfolio. The t -statistics of the returns of the spread portfolios are reported. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Periods	covariance	coskewness						t -statistic of (5-1)
		1	2	3	4	5	(5-1)	
All	1	1.30%	1.16%	1.14%	1.05%	0.87%	-0.42%	(-1.44)
	2	1.46%	1.25%	1.23%	1.04%	1.00%	-0.46%**	(-2.32)
	3	1.61%	1.32%	1.29%	1.21%	1.09%	-0.52%**	(-2.45)
	4	1.61%	1.19%	1.32%	1.13%	1.10%	-0.51%**	(-2.31)
	5	1.41%	1.14%	1.25%	1.07%	1.02%	-0.39%	(-1.33)
Over	1	1.80%	1.17%	1.12%	0.85%	0.68%	-1.12%**	(-2.10)
	2	1.61%	1.30%	1.16%	1.08%	0.99%	-0.62%*	(-1.83)
	3	2.13%	1.53%	1.37%	1.37%	1.33%	-0.80%**	(-2.33)
	4	2.48%	1.88%	1.81%	1.57%	1.52%	-0.96%***	(-2.70)
	5	2.93%	2.35%	2.54%	2.43%	2.21%	-0.72%	(-1.36)
Under	1	0.80%	1.16%	1.15%	1.25%	1.07%	0.27%	(1.16)
	2	1.31%	1.20%	1.31%	1.00%	1.00%	-0.31%	(-1.46)
	3	1.09%	1.11%	1.21%	1.04%	0.86%	-0.23%	(-0.95)
	4	0.75%	0.49%	0.83%	0.68%	0.68%	-0.07%	(-0.25)
	5	-0.11%	-0.07%	-0.04%	-0.29%	-0.18%	-0.07%	(-0.25)

Table VI Fama-MacBeth regressions. All period (All), underweighting periods (Under), and overweighting periods (Over) are identified by PWI_1 index. The data cover January 1996 to April 2016. The average values of the coefficients A and B , which are multiplied by 1000, are reported above, and the corresponding t -statistics are reported below, in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

	A	B	A	B	A	B
All	1.507 (0.65)	-0.438*** (-2.87)				
Over			8.888** (2.20)	-0.823*** (-2.89)		
Under					-5.873*** (-2.74)	-0.053 (-0.53)

Table VII Summary statistics of the covariance and coskewness in Fama-MacBeth regression. The covariance and coskewness of risky stock i for each month t from Jan. 1996 to Apr. 2016, by the 60-month-window regression, $\tilde{r}_{i,\tau} = k_{i,t} + \text{covariance}_{i,t} \tilde{r}_{M,\tau} + \text{coskewness}_{i,t} \tilde{r}_{M,\tau}^2 + \epsilon_{i,\tau}$, $\tau = t - 59, \dots, t$.

	Minimum	25th percentile	Median	75th percentile	Maximum	Standard Deviation
covariance	-7.9545	0.5770	1.0239	1.5692	15.2319	0.8586
coskewness	-141.0023	-4.9601	-0.6000	3.3440	320.0555	10.8123

Table VIII Fama-MacBeth regression results with other control variables. “Whole period” is from January 1996 to April 2016, while “underweighting periods” and “overweighting periods” are determined by PWI_1 index. The first column in each panel is the result without control variables. The second, third, and fourth columns in each panel are the results with two, four, and six control variables. The average values of the coefficients, which are multiplied by 1000, are reported above and the corresponding t-statistics are reported below, in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively. The row labelled “Adj. R^2 ” reports the average adjusted r-square of each regression. The row labelled “ n ” reports the average number of stocks in each regression.

	Panel I: Whole period				Panel II: Overweighting periods				Panel III: Underweighting periods			
<i>A</i>	1.51 (0.65)	1.24 (0.65)	2.30 (1.20)	1.72 (1.03)	8.89** (2.20)	7.47** (2.28)	8.74*** (2.65)	7.23** (2.59)	-5.87*** (-2.74)	-4.99*** (-2.73)	-4.14** (-2.35)	-3.79** (-2.25)
<i>B</i>	-0.44*** (-2.87)	-0.38*** (-2.94)	-0.33*** (-2.63)	-0.24** (-2.47)	-0.82*** (-2.89)	-0.66*** (-2.78)	-0.60** (-2.56)	-0.43** (-2.51)	-0.05 (-0.53)	-0.09 (-1.05)	-0.06 (-0.72)	-0.05 (-0.53)
<i>IVol</i>	0.01 (0.15)	0.01 (0.15)	-0.06 (-1.34)	-0.05 (-1.35)	0.10 (1.02)	0.10 (1.02)	0.01 (0.07)	0.01 (0.13)		-0.09 * (-1.69)	-0.12 *** (-2.76)	-0.12*** (-2.69)
<i>ISkew</i>	-0.72 (-1.17)	-0.72 (-1.17)	-1.03* (-1.68)	-0.74 (-1.39)	-2.46** (-2.32)	-2.60** (-2.48)	-2.60** (-2.48)	-1.92** (-2.18)	1.01 (1.64)	1.01 (1.64)	0.54 (0.89)	0.44 (0.74)
<i>Size</i>			-1.66*** (-3.27)	-1.43*** (-3.23)			-2.23** (-2.45)	-1.66** (-2.13)			-1.09** (-2.43)	-1.21*** (-2.79)
<i>BM</i>			1.68** (2.35)	1.93*** (2.71)			1.30 (1.11)	1.64 (1.39)			2.06** (2.51)	2.22*** (2.76)
<i>Mom</i>				0.01 (0.45)				-0.02 (-0.59)				0.04** (2.25)
<i>Rev</i>				-0.34*** (-6.38)				-0.54*** (-5.94)				-0.15*** (-2.81)
Adj. R^2	0.02	0.03	0.04	0.05	0.02	0.03	0.04	0.06	0.02	0.02	0.03	0.04
<i>n</i>	2695	2693	2691	2689	2819	2817	2815	2813	2571	2569	2567	2565

Figure 1. Prelec’s probability weighting functions with different α ’s. When $\alpha < 1$, it has an inverse S-shape (i.e., overweighting both tails). When $\alpha > 1$, it has an S-shape (i.e., underweighting both tails). The point $p^* = e^{-1}$ separates the convex and concave domains of the probability weighting functions.

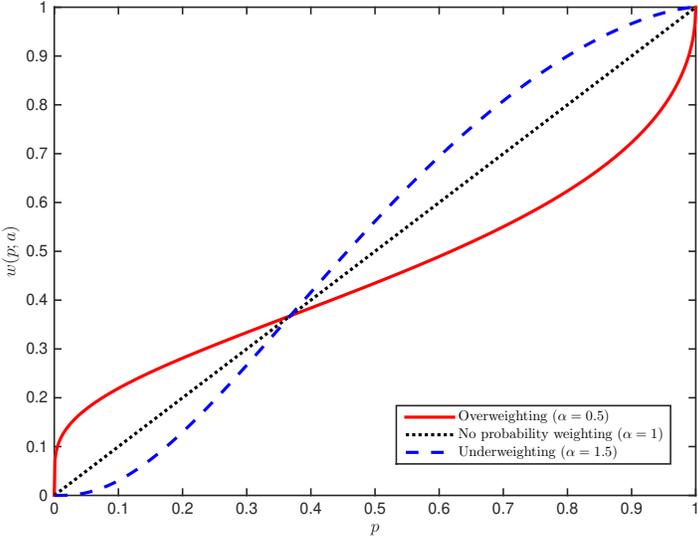
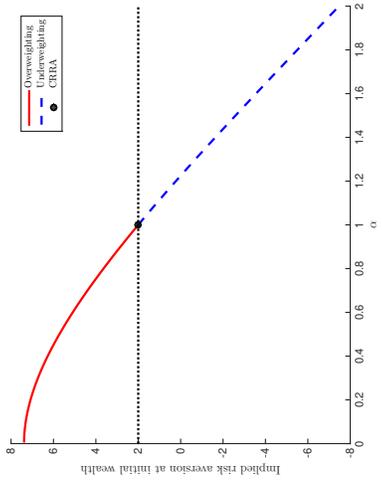
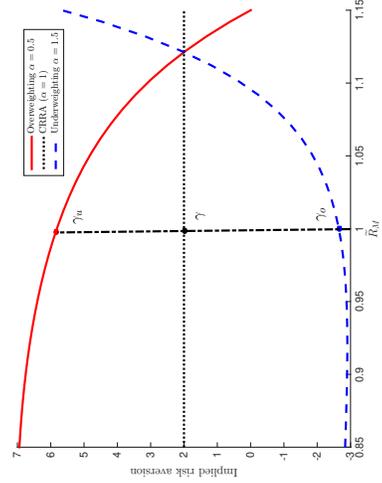


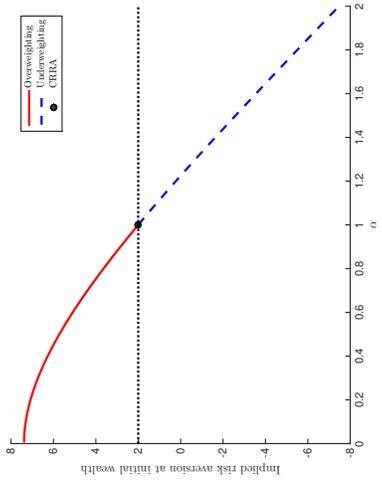
Figure 2. The pricing kernels and implied relative risk aversions for CRRR and RDUT, under both market scenarios. The CRRR has a relative risk aversion $\gamma = 2$ and the RDUT has Prelec's probability weighting functions. The dot black lines are those of the CRRR model, the solid red lines represent the RDUT model with overweighing the tails, and the dashed blue lines show the RDUT model with underweighing the tails. For Market I, the market return follows a skew-normal distribution with mean 7.6%, standard deviation 15.8%, and skewness -0.339. For Market II, the market return follows a skew-normal distribution with mean -15.2%, standard deviation 15.8%, and skewness -0.339. In (a), (b), (d), (e), $\alpha = 0.5$ for overweighing and $\alpha = 1.5$ for underweighing.



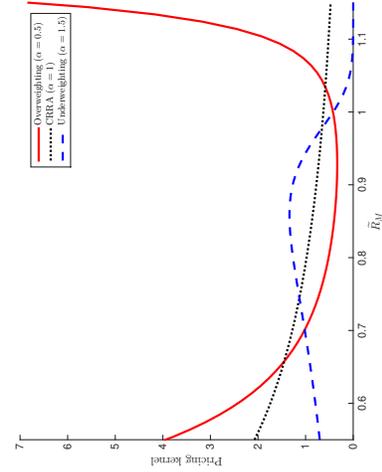
(a) Pricing kernel (Market I)



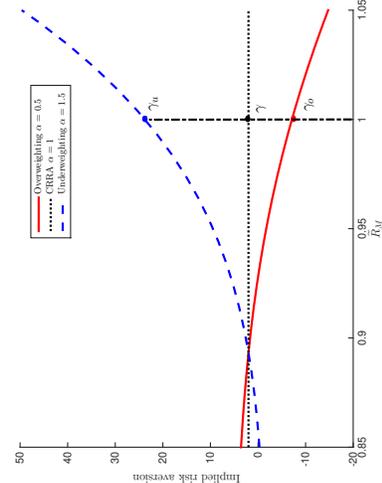
(b) Implied risk aversion (Market I)



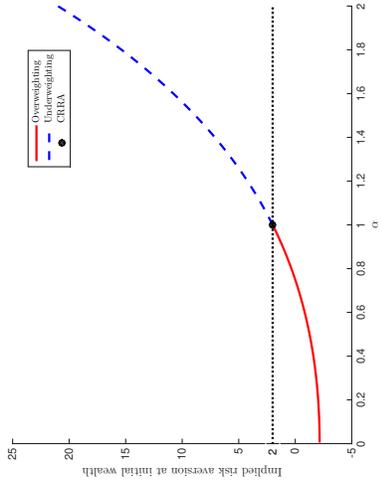
(c) Implied risk aversion at $\tilde{R}_M = 1$ (Market I)



(d) Pricing kernel (Market II)



(e) Implied risk aversion (Market II)



(f) Implied risk aversion at $\tilde{R}_M = 1$ (Market II)

Figure 3. Relationship between A and α . The probability weighting function is that of Prelec with parameter α and the utility is CRRA with relative risk aversion $\gamma = 2$. The market return follows a skew-normal distribution with mean 7.6%, standard deviation 15.8%, and skewness -0.339.

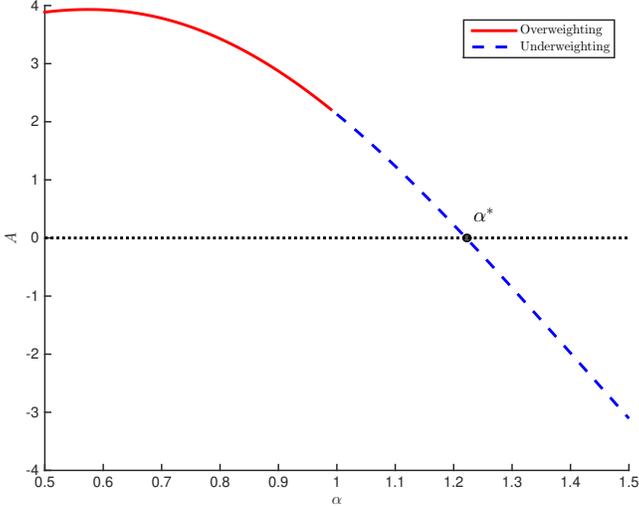
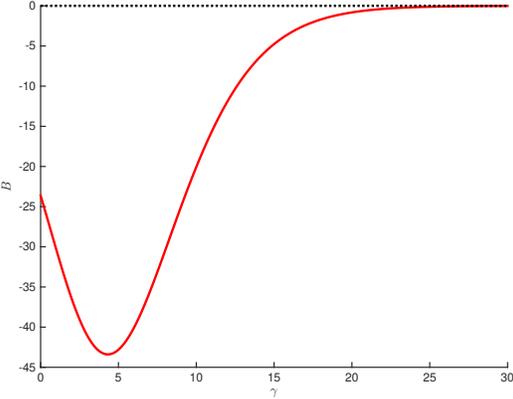
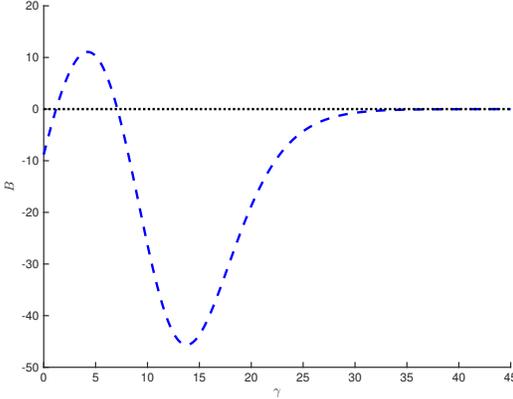


Figure 4. Relationship between B and γ . The probability weighting function is that of Prelec with $\alpha = 0.5$ or $\alpha = 1.5$ and the utility is CRRA with relative risk aversion γ . The market return follows a skew-normal distribution with mean 7.6%, standard deviation 15.8%, and skewness -0.339.

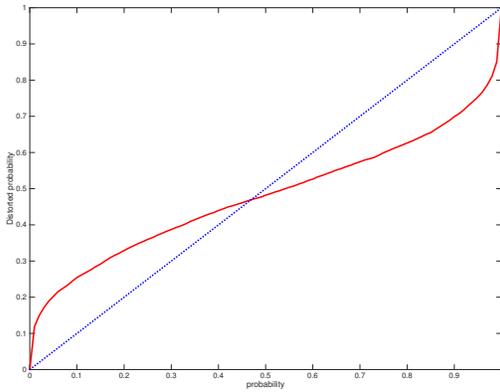


(a) Overweighting Case $\alpha = 0.5$

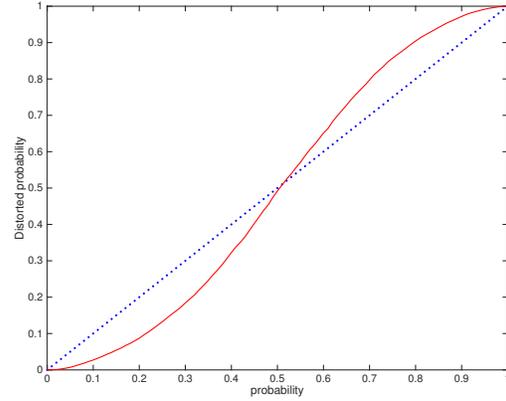


(b) Underweighting Case $\alpha = 1.5$

Figure 5. Two typical option-implied probability weighting functions from our empirical study. The weighting function on September 30, 2008 is inverse S-shaped, implying overweighting of both tails. Among the 244 months tested, there are 126 months (about 51.6%) in which the implied weighting functions are inverse S-shaped. The weighting function on December 30, 2006 is S-shaped, yielding underweighting of both tails. Among the 244 months, there are 92 months (about 37.7%) in which the implied weighting functions are S-shaped.



(a) 2008-09-30



(b) 2006-12-30

Figure 6. Implied fear index (the solid blue line) and implied hope index (the dotted red line). The dashed black line separates overweighting and underweighting.

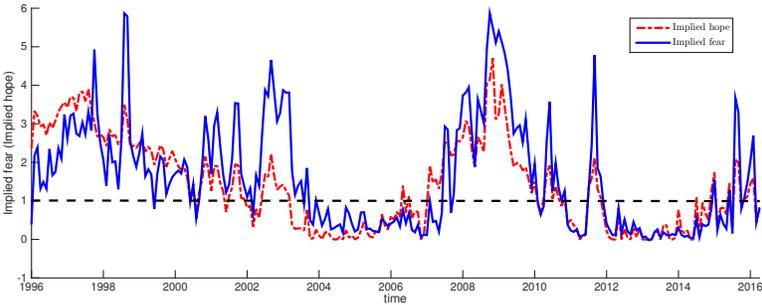
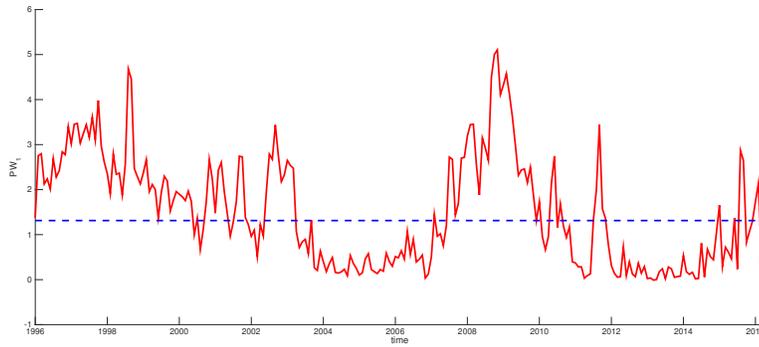
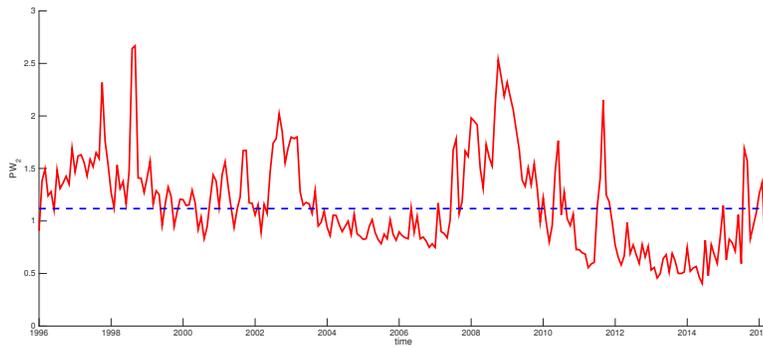


Figure 7. Option implied probability weighting indices PWI_1 and PWI_2 (the solid red lines). The dashed blue line is the median of each index.

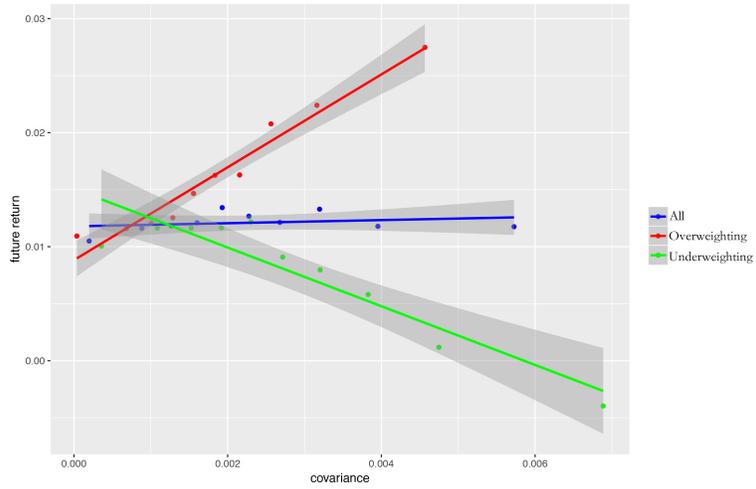


(a) Time series of PWI_1

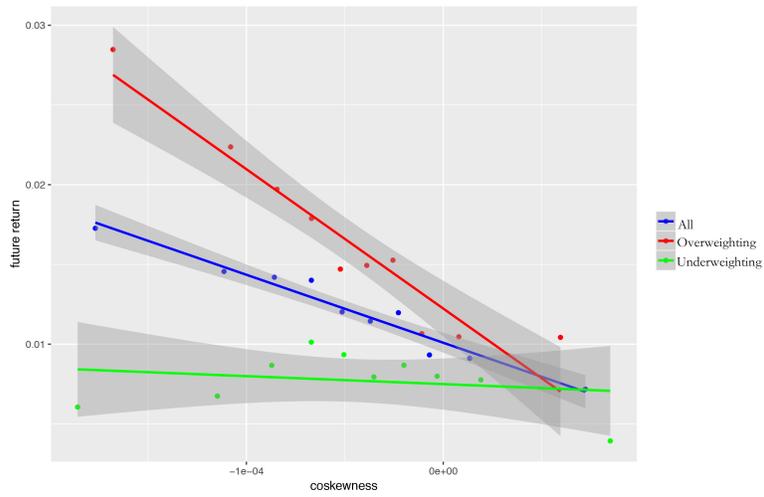


(b) Time series of PWI_2

Figure 8. The relationship between the average next month's returns and the average covariance (coskewness) of the ten covariance-sorted portfolios (ten coskewness-sorted portfolios).



(a) Covariance-sorted portfolios



(b) Coskewness-sorted portfolios