Expected Utility Theory and Prospect Theory in Skewness Preference

Jiawei Hu*

October 1, 2023

Abstract

Based on Barberis et al. (2021) prospect theory asset pricing model, we build a model that incorporates both expected utility theory term and prospect theory term to study skewness preference. The model is used to explain coskewness premium and idiosyncratic skewness. As expected, prospect theory term mainly responsible for idiosyncratic skewness premium. However, we find that although both terms play a role in explaining coskewness premium, the two terms have competing effects.

Keywords: Skewness preference, expected utility theory, cumulative prospect theory, coskewness premium

*The University of Texas at Dallas, Naveen Jindal School of Management, 800 W Campbell Rd, Richardson, TX 75080, email: Jiawei.Hu@utdallas.edu.

1 Introduction

The effect of skewness preference has been studied for decades. Based on expect utility theory, early papers, such as Rubinstein (1973), Kraus and Litzenberger (1976), and Harvey and Siddique (2000), mainly focus on the systematic skewness and imply that only the coskewness should be priced. However, since diversification is not necessarily desirable for investors who have skewness preference (see Simkowitz and Beedles, 1978; Conine Jr and Tamarkin, 1981), under different frameworks, more recent papers point out that individual stock's total skewness, which includes both systematic skewness and idiosyncratic skewness, is priced (see Brunnermeier et al., 2007; Mitton and Vorkink, 2007; Barberis and Huang, 2008b). Moreover, Bakshi et al. (2003) propose skew laws to link the price of security's total skewness to the price of market skewness and the price of idiosyncratic skewness. Similarly, by empirically decomposing total skewness directly into coskewness and idiosyncratic skewness, Conrad et al. (2013) find that idiosyncratic component matters after controlling for the differences of co-moments. Nonetheless, these papers do not explain the existence of coskewness premium and idiosyncratic skewness simultaneously under one framework.

In this paper, we propose a framework that can generate both coskewness premium and idiosyncratic skewness premium at the same time and study the interaction between the terms that corresponding to the two premiums. While both coskewness premium and idiosyncratic skewness premium are studied separately under different frameworks in the literature, there is rare paper looks at the two premiums using the same preference structure. In theoretical papers, on one hand, the literature that explains coskewness premium is based on the expected utility theory on total wealth with assumption of well-diversification, but well-diversification erases the effect of individual skewness. On the other hand, the literature that generates idiosyncratic skewness premium usually has the assumption of zero co-moments among stocks, but such simplification makes the portfolio skewness determined only through individual skewness, which as a result, fail to generate coskewness premium. To solve the issue, we incorporate both expected utility theory (skewness preference term) and prospect theory (prospect theory term) into one model so that it can generate both premiums simultaneously. By doing so, the degree of diversification is affected by the weight on prospect theory, as illustrated in Barberis and Huang (2008b), so that idiosyncratic skewness premium can survive, while the co-moments are not necessarily be zero so that the coskewness premium is generated from co-moments among stocks under expected utility theory.

By assuming homogeneous preference, our model closely follows the settings of Barberis et al. (2021) but also extend their model by adding skewness preference under expected utility theory on total wealth in the utility function. The advantage of such a model is that it can nest some of the previous model. By assigning zero weights on both skewness preference term and prospect theory term, the model defaults to the seminal Markowitz (1959) mean-variance model. With skewness preference term but without prospect theory term, the model generates the equilibrium described by Kraus and Litzenberger (1976) for coskewness premium. Without skewness preference term but with prospect theory term, the model becomes the Barberis et al. (2021) model, which can explain the idiosyncratic skewness premium, as well as many existing anomalies.

We use the model to study the coskewness premium and idiosyncratic skewness premium. The model indicates that the coskewness premium increases with the weight on skewness preference term but shrinks with weight on prospect theory term. On the contrary, the idiosyncratic premium increases mainly with weight on prospect theory term and increases slightly with weight on skewness preference term. To study why the two terms show an competing relationship in explaining coskewness premium, we decompose the total utility from objective function for coskewness premium. We analyze the effect of each component on the expected return and the interaction between the components. We find that, on one hand, if individual stocks are same among assets, the existence of prospect theory term does not affect the utility from skewness preference term directly but change its influence by introducing underdiversification. On the other hand, the existence of skewness preference term affects the utility from prospect theory term directly if the co-moments of assets are different. The channel is to shift the return distribution parallel so that the utility from prospect theory term on the distribution changes.

Our paper contributes the literature in two ways. First, we study the relationship between expected utility theory term and prospect theory term in one model, linking the two separately explored areas and directly showing that coskewness premium and idiosyncratic skewness premium can exist simultaneously in an equilibrium. Second, by adding skewness preference term into Barberis and Huang (2008b) and Barberis et al. (2021) framework, we show that prospect theory, as a behavioral term, interacts with traditional rational term in equilibrium and provide additional insights in explaining existing risk premiums.

The rest of the paper is organized as follows. In Section 2, we review the literature about skewness preference and prospect theory. Section 3 present the model and the equilibrium structure. In Section 4, we discuss the implication of the model and its mechanism.

2 Literature Review

In this section, we briefly review the literature in skewness preference, including both systematic skewness and idiosyncratic skewness. We also provide introduction to prospect theory and narrow framing and show how literature uses them to explain phenomenon in finance area.

2.1 Systematic Skewness

With the findings that the empirical results are usually inconsistent with the traditional form of the Sharpe-Lintner mode[1], researchers have their eye on higher order moments. By taking the expectation of the exact Taylor series expansion, Rubinstein (1973) extends the mean-variance security valuation model to a general parameter-preference model. Ignoring the

[1] See, for example, Friend and Blume (1970), Black et al. (1972), and Blume and Friend (1973).

fourth and higher orders, Kraus and Litzenberger (1976) define the investor's expected utility over the first three central moments and propose a unconditional three-moment CAPM. Through nonlinear pricing kernel, Harvey and Siddique (2000) and Dittmar (2002) develop the conditional version of three-moment CAPM.

While the theoretical model is well developed, there are several issues regarding the skewness preference. First, literature has a discussion about the direction of the skewness preference or the existence of premium based on the expected utility theory. Along with the sense that investors prefer positive skewness, most papers argue that the stock with higher skewness should has lower expected return. By studying the two-moment investing strategy in a three-moment world, Beedles (1979) shows that the investors must sacrifice some return skewness to gain the mean-variance benefits. Scott and Horvath (1980) prove that, for risk-averse investors who are consistent in direction of preference of moments, the preference direction is positive for positive skewness. Applying polynomial goal programming in international stock markets, Chunhachinda et al. (1997) and Prakash et al. (2003) find that investors who have preference for skewness do trade expected return of the portfolio for skewness. Using different measures of coskewness, Harvey and Siddique (2000) and Langlois (2020) find a significant positive premium on systematic skewness in US stock market. Using option data, Bali and Murray (2013) and Christoffersen et al. (2021) also find empirical supports that are consistent with positive skewness preference. Nonetheless, other papers find that the sign is not conclusive. Friend and Westerfield (1980) employ a more comprehensive testing of the Kraus-Litzenberger thesis but find that the evidence of premium on positive skewness is not conclusive[2]. Moreover, based on the theory of Tchebychev systems of functions, Brockett and Kahane (1992) show that the assumed positive skewness preference is theoretically unsound when the choice set of decision makers includes arbitrary distribution. Similarly, by including higher order derivatives, Simaan (1993) show that, for risk-averse investors, the sign of the skewness premium is not determined beforehand. Moreover, Kozhan

[2]Barone-Adesi (1985) attributes the contrasting results between K-L and F-W to the serious econometric problems

et al. (2013) find that skewness risk is tightly related to variance risk, so the risk premium on skewness become insignificant after hedging out the variance risk. Although these papers argue from different perspectives, they all imply that the skewness preference may be influenced by other factors as investors are complicated. They inspire that additional terms in utility function may have opposite effect on skewness preference based on expected utility theory, as we find in later section.

Second, the existence of skewness preference may make investors hold under-diversified portfolio. Simkowitz and Beedles (1978) show that portfolio skewness decreases with diversification, so for investors who have preference on skewness, diversification is not necessarily desirable. Conine Jr and Tamarkin (1981) and Kane (1982) theoretically link the portfolio skewness preference to the empirical findings of underdiversification. It is such underdiversification leaves room for idiosyncratic skewness premium. Using different behavioral model, Brunnermeier et al. (2007), Mitton and Vorkink (2007), and Barberis and Huang (2008b) generate equilibrium with idiosyncratice skewness premium and underdiversification.

In conclusion, skewness preference shows additional power beyond traditional mean-variance framework, but something beyond systemtic skewness may also affect the role of skewness preference in asset pricing.

2.2 Idiosyncratic Skewness

According to the rational model based on expected utility theory, only systematic skewness should be priced. However, empirical findings, such as Boyer et al. (2010), show that idiosyncratic skewness is also priced. This implies that investors may also care about idiosyncratic component of returns beyond the systematic component.

Several papers propose different theories that can help explain why investors prefer individual stocks with positive skewness. First, following the optimal belief framework proposed by Brunnermeier and Parker (2005), where investors have subjective probability and persistently err in attaining their goal, Brunnermeier et al. (2007) show that investors have idiosyncratic skewness preference beyond systematic skewness preference and the relative importance of idiosyncratic skewness and systematic skewness in asset pricing may vary with the subjective belief. Their study also fit the fact that investors hold heterogeneous and imperfect-diversified portfolios.

Second, assuming heterogeneous preference rather than belief, Mitton and Vorkink (2007) generate equilibrium with underdiversification. The two different types of investors are "Traditional Investor" with mean-variance utility function and "Lotto Investor" with mean-variance-skewness utility function. By assuming zero co-moments among securities, they show that idiosyncratic skewness have impact on equilibrium prices beyond coskewness. In addition, underdiversification is associated with skewness not by coincident.

Third, within homogeneous preference and belief framework, Barberis and Huang (2008b) show that stock's own skewness, not just coskewness, can be priced. In their model, the investors have heterogeneous holdings based on the prospect theory framework, in which probability weighting and some degree of loss aversion contribute to the result, and the pricing of skewed stock is generated because of the trade-off between diversification and portfolio-level skewness preference. They also point out that the pricing of idiosyncratic skewness can be derived by applying prospect theory under narrow framing. Barberis et al. (2021) extend this idea and use the model with narrow framing to quantitatively predict the expected idiosyncratic skewness anomaly, as well as many other empirical anomalies.

In general, idiosyncratic skewness seems to have additional effect on asset pricing beyond systematic skewness, and this effect is associated with underdiversification.

2.3 Prospect Theory and Narrow Framing

Since we employ prospect theory with narrow framing as the method to generate idiosyncratic skewness premium, we briefly introduce prospect theory and narrow framing in this section.

To solve some conflicts between laboratory findings and the expected utility theory, Kahneman and Tversky (1979) propose the first version of prospect theory to serve as an alternative way in explaining people's decision under risk. It is extended by Tversky and Kahneman (1992) to incorporate cumulative functional and risky prospects with more than two outcomes. The advanced version is cumulative prospect theory, which is applied in many financial papers and in this paper.

To see how cumulative prospect theory works, consider a gamble with $m+n+1$ states, where m states have negative payoffs, 1 state has zero payoff, and n states have positive payoff. In an ascending order by payoffs, it can be expressed as

$$(x_{-m}, p_{-m}; ...; x_{-1}, p_{-1}; x_0, p_0; x_1, p_1; ...; x_n, p_n). \tag{1}$$

In equation 1, x_i is the payoff and p_i is its corresponding probability. A cumulative prospect theory individual assigns the value to this gamble as

$$\sum_{i=-m}^{n} \pi_i v(x_i), \tag{2}$$

where

$$\pi_i = \begin{cases} w(p_i + ... + p_n) - w(p_{i+1} + ... + p_n), & 0 \leq i \leq n \\ w(p_{-m} + ... + p_i) - w(p_{-m} + ... + p_{i-1}), & -m \leq i < 0 \end{cases} \tag{3}$$

and where $v(.)$ and $w(.)$ are value function and probability weighting function, respectively. They have the functional forms

$$v(x) = \begin{cases} x^\alpha, & x \geq 0 \\ -\lambda(-x)^\alpha, & x < 0 \end{cases} \tag{4}$$

and

$$w(P) = \frac{P^\delta}{(P^\delta + (1-P)^\delta)^{1/\delta}}. \tag{5}$$

where $\alpha, \delta \in (0,1)$ and $\lambda > 1$.

Distinct from the expected utility theory, the cumulative prospect theory have the following features on value function and probability distribution. First, the value function is defined over gains and losses, concave over gains but convex over losses (diminishing sensitivity), and kinked at the origin with steeper slope for losses than for gains (loss aversion). Second, the objective probability distribution is transformed so that people overweight small probabilities but underweight moderate and high probabilities (probability weighting). Figure 2 shows the pattern of the value function and the probability function under cumulative prospect theory.

[insert Figure 2 here]

Narrow framing, first demonstrated by Tversky and Kahneman (1981), means that when making several decisions, people are usually focusing on the outcome of each decision separately instead of focusing on the combined outcome of decisions[3]. In finance field, it means that investors maximize their utility function, to some extent, over individual security gains and losses rather than solely over final wealth. Barberis et al. (2006) also argue that the commonly observed rejection of small, independent, but favorable gamble is evidence not only for loss aversion but also for narrow framing.

The cumulative prospect theory and narrow framing are used to explain several phenomena in finance field[4]. The loss aversion, combined with narrow framing in some cases, catches researchers' eye first. Using myopic loss aversion, a combination of loss aversion and a short evaluation period, Benartzi and Thaler (1995) propose a static model to explain the equity premium. The investors who calculate utility based on prospect theory have extreme discomfort with the return variability even if the returns in short-run have no effect on consumption. Allowing the representative investor to get utility from both the fluctuation of the financial wealth and the consumption, Barberis et al. (2001) apply prospect theory to a dynamic equilibrium model that can generate equity premium. In addition, they point out

[3] See Barberis and Huang (2008a) for more detailed interpretation and examples for narrow framing, as well as loss aversion.

[4] See Barberis (2013) for other applications of prospect theory in economy.

that in order to generate equity premium, loss aversion should be combined with the effect of prior outcomes. To further address the time-series and cross-sectional behavior of individual stocks, Barberis and Huang (2001) consider economies with two kinds of narrow framing, which are stock-level and portfolio-level. They find that while both kinds of narrow framing can explain the features of the data with loss aversion, individual stock mental accounting is more successful. Barberis and Huang (2009) also propose new preference specification compared to Barberis et al. (2001) to overcome some limitations, such as no explicit value function and intractable preference in partial equilibrium settings.

More recently, the role of probability weighting is also explored. Barberis and Huang (2008b) build a model to study stock price in economies where investors evaluate risk using prospect theory. Their results show that investors pay very high price for stocks with positive skewness, and can be used to explain idiosyncratic skewness pricing when it is applied with narrow framing. In the equilibrium structure, it is probability weighting with some degree of loss aversion playing the crucial role to generate such outcomes. Probability weighting is also used to explain casino gambling by Barberis (2012). He shows that a prospect theory agent would be willing to take the gamble even if it offers bets with no skewness and with zero or negative expected value. He also shows that probability weighting could make the agent have time-inconsistent strategy. To study the effect of prospect theory in cross-sectional stock return behavior, Barberis et al. (2016) propose a framework where investors apply prospect theory to past stock-level gains and losses. They find that, on average, a stock with high prospect theory value on its past return distribution earns a low subsequent return, and it is probability weighting primarily responsible for this predictive power.

The general implication of prospect theory as a whole is studied recently. Baele et al. (2019) develop a equilibrium model with cumulative prospect theory preferences and apply it to U.S. equity index option returns using GMM. They show that it can simultaneously explain the low returns on out-of-the-money (OTM) equity index put options, the low returns on OTM equity index call options, and the large variance premium. Barberis et al. (2021)

propose a model incorporating all the elements in prospect theory, as well as narrow framing and capital gain overhang, to explain many common anomalies and make quantitative prediction on asset's average return.

From the literature summarized above, we know that the combination of prospect theory and narrow framing can help explain many phenomena in finance field. In addition, this combination can be used to generate underdiversification and idiosyncratic skewness premium in our model. The probability weighting emphasizes on the tails and the loss aversion brings asymmetric attitude towards gains and losses, so prospect theory, as a whole, generates skewness preference. Restricting it to individual stock as described by narrow framing, this skewness preference is set at the stock-level so that the idiosyncratic skewness premium can be explained beyond the power of expected utility theory. Moreover, narrow framing also isolate the contribution of a stock's co-moments to the portfolio skewness as prospect theory do not provide direct channel to do so. However, since the structure of prospect theory cannot clearly differentiate the contribution between coskewness and individual skewness in portfolio-level gains and losses, we don't apply it to portfolio-level but rather let expected utility theory to do the job for generating coskewness premium.

3 Model and Equilibrium Structure

In the literature review, we show that the inclusion of skewness preference implies a premium on systematic skewness and the inclusion of prospect theory helps to explain the premium on idiosyncratic skewness. In this section, we include both terms into the traditional mean-variance framework and discuss the equilibrium structure given this homogeneous preference. Our model closely follows the settings of Barberis et al. (2021).

3.1 Model Setup

Consider an economy with three dates, t = -1, 0, and 1, and investors make decision at date 0. There is a risk-free asset with gross per-period return R_f and N risky assets with gross per-period return $\tilde{R}_i$ for risky asset i. Let $\tilde{R} = (\tilde{R}_1, ..., \tilde{R}_N)'$ be the return vector and has cumulative distribution function $P(\tilde{R})$. The vector of expected returns on the risky assets is $\bar{R} = (\bar{R_1}, ..., \bar{R_N})$ and the covariance matrix of returns is $\Sigma = \{\sigma_{ij}\}$.

Assume investors in the economy are identical in their preferences, wealth at time -1 (W_{-1}), and wealth at time 0 (W_0). Let the fraction of wealth at time 0 that an investor allocates to risky asset i is Θ_i, and the allocation vector is $\Theta = (\Theta_1, ..., \Theta_N)'$. Thus, wealth at time 1 is

$$\tilde{W_1} = W_0((1 - 1'\Theta)R_f + \Theta'\tilde{R}). \tag{6}$$

At date 0, each investor solves the following objective function to determine the allocation:

$$\max_{\Theta_1,...,\Theta_N} E(\tilde{W_1}) - \tfrac{\gamma}{2}Var(\tilde{W_1}) + \tfrac{\phi}{6}Sk(\tilde{W_1}) + b\textstyle\sum_{i=1}^{N} V(\tilde{G_i}), \tag{7}$$

where

$$\begin{aligned} Var(\tilde{W_1}) &= E\left[(\tilde{W_1} - \bar{W_1})^2\right] \\ &= W_0^2\Theta'\Sigma\Theta, \end{aligned} \tag{8}$$

$$\begin{aligned} Sk(\tilde{W_1}) &= E\left[(\tilde{W_1} - \bar{W_1})^3\right] \\ &= W_0^3 \textstyle\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{k=1}^{N} \Theta_i\Theta_j\Theta_k E[(\tilde{R_i} - \bar{R_i})(\tilde{R_j} - \bar{R_j})(\tilde{R_k} - \bar{R_k})] \end{aligned} \tag{9}$$

$$\tilde{G_i} = W_0\Theta_i(\tilde{R_i} - R_f) + W_{-1}\Theta_{i,-1}g_i. \tag{10}$$

The first two terms in equation 7 is the traditional mean-variance preference under expected utility theory on total wealth, where the parameter γ controls the degree of aversion to portfolio risk. The third term is portfolio skewness (the third central moment) under expected utility theory, which is new to the Barberis et al. (2021) model. The parameter ϕ measures the degree of preference to the portfolio skewness, and it theoretically can be either positive, negative, or zero[5]. We can also see from this term that individual skewness affect the portfolio skewness through the terms where $i = j = k$ but it possibly does not affect much compared to the greater amount of co-moments. The fourth term captures both prospect theory and narrow framing. It is the sum of N components, where each $V(\tilde{G}_i)$ corresponds to a specific asset i. $V(.)$ is the cumulative prospect theory value of the asset following equation 2, and $\tilde{G}_i$ is the potential gain or loss on asset i. In equation 10, we set $\tilde{G}_i$ as the sum of two terms: the potential future gain or loss on asset i between time 0 and time 1 ($W_0\Theta_i(\tilde{R}_i - R_f)$) and the gain or loss the investor experienced in his holdings of asset i prior to time 0 ($W_{-1}\Theta_{i-,-1}g_i$)[6]. The parameter b measures relative importance of the prospect theory term.

A natural question is why we use the third moments to capture skewness preference but use prospect theory value to capture idiosyncratic component effect and why we don't exchange them. First, the choice of third moments for skewness preference allows us to clearly see the effect of co-moments in portfolio skewness and differentiate it from the effect of individual skewness. In rational model under expected utility theory, only coskewness is priced, so we can assign same value to individual skewness among stocks to achieve this outcome. On the contrary, if we use prospect theory value in portfolio level, we cannot set co-moments different but individual skewness same among stocks to generate skewness preference as in Mitton and Vorkink (2007) and Barberis and Huang (2008b). Second, the choice of prospect theory value in idiosyncratic skewness allows us to generate heterogeneous holdings and underdiversification. In a homogeneous preference model, it is hard to generate heterogeneous holdings and underdiversification using only third moments. Technically, it does not provide enough opposite curvature to generate multiple global optima in the utility curve. Intuitively, individual skewness would not matter much if investors hold

[5]According to the discussion in literature review, we set this parameter to positive value or zero to compare the effect of portfolio skewness preference. Negative value is mathematically feasible but not intuitive in the context.

[6]The parameter g_i in the second term is corresponding to the capital gain overhang on asset i and is treated as part of the benchmark to measure the total gain or loss. It can also be set as zero, risk-free rate, negative market return, or other reasonable values depending on which benchmark we assume investors use.

well-diversified portfolio. That's one reason why we see Mitton and Vorkink (2007) choose heterogeneous preference model while Barberis and Huang (2008b) choose prospect theory in homogeneous preference model to explain skewness preference.

One merit of this utility function is that it allows our model to nest some well-known equilibria in the literature. Specfically, if we set $\phi = 0$ and $b = 0$ in equation 7, the utility function defaults to the traditional mean-variance utility function, which can generate outcomes as Markowitz (1959). If we set $\phi \neq 0$ but $b = 0$, the utility function becomes the mean-variance-skewness utility function used in Kraus and Litzenberger (1976) and can indicate coskewness premium with asymmetric returns distribution. If we set $\phi = 0$ but $b \neq 0$, the utility function turns into the one used in Barberis et al. (2021) model, and it can be used to explain idiosyncratic skewness premium.

For simplicity, we have the following assumptions similar to Barberis et al. (2021) with some adjustments. The effect of these assumptions to the model can be found in their paper. Since our new term, third moments for skewness preference, does not work as a channel for heterogeneous holdings, their findings can be applied to our model. First, we assume that the second term in equation 10 to be identical across investors. In addition, since it is a term not directly related to skewness, we also set same value to all assets. Empirically, we estimate the average gain or loss since purchase across all holders of the asset and average it among all assets[7]. Second, we set $\Theta_{i,-1}$ to a asset i's weight in the market portfolio for all investors. Thus, we can treat the term $W_{-1}\Theta_{i,-1}g_i$ as exogenous, so investors only need to choose the allocation to asset i at time 0, namely Θ_i. Third, we use the approximation $W_{-1} \approx W_0$. More accurate approximation of this relationship may set $W_{-1} \approx W_0/n$, where n is an assumed factor reflecting average return on investor wealth, but it have little impact on our implication, especially when we set g_i to be identical among stocks to control for the effect of capital gain overhang.

[7]As an input calculated from data, the actual difference of the capital gain before average across all deciles is already small (i.e., -1.1 percent vs -3.7 percent for coskewness premium), so even we don't assume it to be identical across deciles, the influence of it is small. For our purpose to concentrate on specific variables, setting it to be same is better for theoretical implication.

For the prospect theory term, $V(\tilde{G}_i)$, we can write it as the following way for $\Theta_i > 0$[8]:

$$\begin{aligned} &-\lambda W_0^\alpha \int_{-\infty}^{R_f-\Theta_{i,-1}g_i/\Theta_i} (\Theta_i(R_f - R_i) - \Theta_{i,-1}g_i)^\alpha dw(P(R_i)) \\ &-W_0^\alpha \int_{R_f-\Theta_{i,-1}g_i/\Theta_i}^{\infty} (\Theta_i(R_i - R_f) + \Theta_{i,-1}g_i)^\alpha dw(1 - P(R_i)). \end{aligned} \tag{11}$$

where $P(R_i)$ is the marginal cumulative distribution function of asset i's returns and $w(.)$ is the probability weighting function in equation 5. Similarly, for $\Theta_i < 0$[9], we can write it as

$$\begin{aligned} &\lambda W_0^\alpha \int_{R_f-\Theta_{i,-1}g_i/\Theta_i}^{\infty} (\Theta_i(R_f - R_i) - \Theta_{i,-1}g_i)^\alpha dw(1 - P(R_i)) \\ &+W_0^\alpha \int_{-\infty}^{R_f-\Theta_{i,-1}g_i/\Theta_i} (\Theta_i(R_i - R_f) + \Theta_{i,-1}g_i)^\alpha dw(P(R_i)). \end{aligned} \tag{12}$$

The top row of equation 11 and 12 corresponds to losses and is therefore multiplied by λ. The bottom row of the equations corresponds to gains.

We also need to specify the probability distribution for stock returns in the model to calculate the prospect theory value. Since skewness is the main point in analysis, the distribution should allow asymmetry and can capture the characteristics of stock return distribution as accurately as possible. We employ the generalized hyperbolic (GH) skewed t distribution here since it is good at modeling skewness and fat tails in asset returns[10]. The density function of a one-dimensional GH skewed t distribution for asset i is

$$\begin{aligned} p(R_i) &= \frac{2^{1-\frac{\nu+1}{2}}}{\Gamma(\frac{\nu}{2})(\pi\nu S_i)^{\frac{1}{2}}} . \frac{K_{\frac{\nu+1}{2}}(\sqrt{(\nu+(R_i-\mu_i)^2/S_i)\zeta_i^2/S_i})exp((R_i-\mu_i)\zeta_i/S_i)}{(\sqrt{(\nu+(R_i-\mu_i)^2/S_i)\zeta_i^2/S_i})^{-\frac{\nu+1}{2}}(1+(R_i-\mu_i)^2\nu^{-1}/S_i)^{\frac{\nu+1}{2}}}, \quad && \zeta_i \neq 0 \\ p(R_i) &= \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})(\pi\nu S_i)^{\frac{1}{2}}} . (1 + (R_i - \mu_i)^2\nu^{-1}/S_i)^{-\frac{\nu+1}{2}}, && \zeta_i = 0 \end{aligned} \tag{13}$$

where $\Gamma(.)$ is the Gamma function and K_l is the modified Bessel function of the second kind with order l.

[8] Since $dw(1 - P(R_i))$ generates a negative sign when transferred to $dw(R_i)$, the bottom row of equation 11 has a negative sign to be consistent with equation 4. Same reason for the positive sign for the top row of equation 12

[9] When $\Theta_i = 0$, the return distribution will not affect the value of this part, so $V(\tilde{G}_i)$ equals $W_0^\alpha(g_i\Theta_{M,i})^\alpha$ if $g_i \geq 0$ or $-\lambda W_0^\alpha(-g_i\Theta_{M,i})^\alpha$ if $g_i < 0$.

[10] See Barberis et al. (2021) for the advantages of adopting this distribution in calculating prospect theory value.

There are four parameters in equation 11 and equation 12: μ_i, S_i, ζ_i, and ν. The parameter μ_i is the location parameter that controls for the mean of the distribution; the parameter S_i is the dispersion parameter that determines the dispersion in returns; the parameter ζ_i is the asymmetry parameter that helps to generate the skewness of returns; and the parameter ν_i is the degree of freedom that affects the heaviness of the tails of the distribution. For the distribution of an asset, the mean, variance, and skewness can be calculated as

$$E(\tilde{R}_i) = \bar{R}_i = \mu_i + \tfrac{\nu}{\nu-2}\zeta_i, \tag{14}$$

$$Var(\tilde{R}_i) = \tfrac{\nu}{\nu-2}S_i + \tfrac{2\nu^2}{(\nu-2)^2(\nu-4)}\zeta_i^2, \tag{15}$$

$$Skew(\tilde{R}_i) = \tfrac{2\zeta_i\sqrt{\nu(\nu-4)}}{\sqrt{S_i}(2\nu\zeta_i^2/S_i+(\nu-2)(\nu-4))^{\frac{3}{2}}}[3(\nu-2) + \tfrac{8\nu\zeta_i^2}{S_i(\nu-6)}], \tag{16}$$

respectively.

3.2 Equilibrium Structure

Similar to Barberis et al. (2021), there could be three possible types of equilibrium in our economy. We now discuss about when will they show up and which one we study. We will also explain the one we use in detail.

The first one is full rationality with homogeneous holdings. Since the investors in our economy are identical in their preferences, with same inputs as assumed in each asset, they may naturally choose same portfolio holdings $\{\Theta_i\}_{i=1}^N$ as the market supply of each asset. Technically, investors choose a vector of $\mu = (\mu_1, ..., \mu_N)'$ such that the objective function in 7 has a unique global maximum $\Theta^* = (\Theta_1^*, ..., \Theta_N^*)$. Moreover, this maximum should satisfy that $\Theta_i^* = \Theta_{M,i}$ for all i, where where $\Theta_{M,i}$ is the market value of asset i divided by the total market value of all traded assets.

This equilibrium structure happens when we set the parameter b, the weight of prospect theory term, equals 0. For our purpose, it reflects the choice of identical investors with skewness preference under expected utility theory only. In this case, as in Kraus and Litzenberger (1976), only coskewness is priced (as long as ϕ is not 0) while idiosyncratic skewness premium is zero. It may also happens when the calculated utility from skewness preference term is relatively small compared to the utility from prospect theory term[11]. The intuition is that the idiosyncratic skewness is so small that even an extreme large holdings could not compensate on the loss from underdiversification caused on other preference terms. Technically, the prospect theory value could not provide enough opposite curvature compared to variance aversion and skewness preference under expected utility theory to generate multiple global maxima in the objective function 7. In other words, even the objective function is a function of Θ_i, Θ_i^2, Θ_i^3, and Θ_i^{α}[12] where $\alpha \in (0, 1)$, it does not have multiple global maxima. However, this second case does not happen in our trails, but it has the same reason as in the third equilibrium structure about why we have some assets have homogeneous holdings.

The second one is full rationality with heterogeneous holdings. When the objective function has multiple global maximum, it means that investors have heterogeneous portfolio holdings so that we can clear the market in each stock. Technically, there exists a vector of $\mu = (\mu_1, ..., \mu_N)'$ such that the objective function in 7 has multiple global maxima. Moreover, we can clear the market by allocating appropriate fraction of investors to each maximum. However, this equilibrium structure is computationally infeasible to determine whether it exists or not[13]. For example, suppose we have N assets and K candidate values for each element of the location vector μ. Then, it implies that we have to try K^N times to find the possible location vectors. For each trial, we also need to find whether it has multiple global maxima and whether we can clear the market by allocating investors to the various maxima.

[11] It depends on both the inputs for calculating the utility from each part of the objective function and the relative weights of each part (i.e., γ, ϕ, and b).

[12] A combination of Θ_i, Θ_i^2, and Θ_i^3 could not generate multiple global maximum. It is Θ_i^{α} bring the possibility to generate multiple global maxima.

[13] See Barberis et al. (2021) for a detailed example

Even if we have further assumptions to reduce both N and K in calculation[14], K^N trials are too many to calculate since each trial take the computer to calculate for non-trivial time.

The third one is bounded rationality with heterogeneous holdings. To overcome the calculation difficulty in full rationality with heterogeneous holdings, we follow Barberis et al. (2021) method by assuming a mild bounded rationality. Specifically, we assume that the investors treat their holdings of the other $N-1$ risky assets equal the market supply of those assets ($\Theta_j = \Theta_{M,j}$ for all $j \neq i$) when they determine the allocation Θ_i to asset i to maximize the objective function 7. This will not be exactly true since investors usually have portfolios that are less diversified than the market portfolios, but it is likely to has a negligible impact on the model's implication because the mechanism of relationship between portfolio-level skewness preference and stock-level skewness preference is not affected by the holdings itself in most cases[15].

In the equilibrium of bounded rationality with heterogeneous holdings, we search for a location vector μ such that, for each asset, the solution to the objective function 7 involves one or multiple global maxima that can clear the market by allocating each investor to one of the maxima. Specifically, for each risky asset i, we view the objective function as a function of Θ_i. The bounded rationality assumption is reflected by setting $\Theta_j = \Theta_{M,j}$ for all $j \neq i$. Then, by linear transformation, for $\Theta_i \geq 0$, the function can be written as[16]

[14] In individual stock-level, N is over 1000, and even with assumption of identical characteristics in each decile of stocks by sorting on a specific characteristic, N is 10. To have a 0.001 increment each time for return location between 0.8 to 1.2, K is 400. In total, K^N is 400^{10} in a very rough searching way.

[15] As explained in section 4.3, the implication exists because of the influence of co-moments on expected return, so as long as the co-moments between stock i and other stocks in the portfolio do not add up to zero by specific holdings, the implication will hold.

[16] The function for $\Theta_i < 0$ is similar to equation 17 with adjustment consistent with 12.

$$
\begin{aligned}
&\Theta_i\left(\mu_i + \tfrac{\nu\zeta_i}{\nu-2} - R_f\right) - \tfrac{\hat{\gamma}}{2}\left(\Theta_i^2 E[(R_i-\bar{R}_i)^2] + 2\Theta_i \sum_{j\neq i}\Theta_{M,j}E[(R_i-\bar{R}_i)(R_j-\bar{R}_j)]\right)\\
&\quad +\tfrac{\hat{\phi}}{6}\Bigg(\Theta_i^3 E[(R_i-\bar{R}_i)^3] + 3\Theta_i^2\sum_{j\neq i}\Theta_{M,j}E[(R_i-\bar{R}_i)^2(R_j-\bar{R}_j)]\\
&\qquad +3\Theta_i\sum_{j\neq i}\Theta_{M,j}^2 E[(R_i-\bar{R}_i)(R_j-\bar{R}_j)^2]\\
&\qquad +3\Theta_i\sum_{j\neq i}\sum_{k\neq\{j,i\}}\Theta_{M,j}\Theta_{M,k}E[(R_i-\bar{R}_i)(R_j-\bar{R}_j)(R_k-\bar{R}_k)]\Bigg)\\
&\quad -\lambda\hat{b}\int_{-\infty}^{R_f-\Theta_{i,-1}g_i/\Theta_i}(\Theta_i(R_f-R_i)-\Theta_{i,-1}g_i)^{\alpha}dw(P(R_i))\\
&\quad -\hat{b}\int_{R_f-\Theta_{i,-1}g_i/\Theta_i}^{\infty}(\Theta_i(R_i-R_f)+\Theta_{i,-1}g_i)^{\alpha}dw(1-P(R_i)),
\end{aligned}
\tag{17}
$$

where

$$
\hat{\gamma} = \gamma W_0, \quad \hat{\phi} = \phi W_0^2, \quad \hat{b} = bW_0^{\alpha-1}. \tag{18}
$$

For each asset i, if the function in equation 7 has a unique global maximum, then $\Theta_i = \Theta_{M,i}$. This means that the investors do not need to have heterogeneous holdings for all risky assets. When we implement the equilibrium, we indeed find that investors have homogeneous holdings for some risky assets. If the function in equation 7 has multiple global maxima, then some of which are below $\Theta_{M,j}$ and some of which are above $\Theta_{M,j}$ so that we can clear the market in asset i by allocating some investors to the lower optima and others to the upper optima. In fact, we find that, based on our inputs and parameter values, all asset i for which investors have heterogeneous holdings, the function in equation 17 has just two global maxima Θ_i^* and Θ_i^{**}. Moreover, they satisfy that $0 < \Theta_i^* < \Theta_{M,i} < \Theta_i^{**}$ and Θ_i^* are always much closer to $\Theta_{M,i}$. It implies that, first, no investors choose to short stock, and, second, most investors are allocated to Θ_i^*[17].

To make the model inputs easier to compute, we rescale the holdings from the proportion to all traded assets to the proportion to all risky assets. Specifically, let $\Theta_{M,R} = \sum_{i=1}^{N}\Theta_{M,i}$,

[17]The reason for no short-selling investors and two global maxima are similar to Barberis et al. (2021) because based on our inputs and parameter values, the portfolio-level skewness preference does not provide much curvature when Θ_i is close to 0. See their paper for detailed discussion.

where $\Theta_{M,i}$ is the market value of asset i divided by the total market value of all traded assets. Thus, $\Theta_{M,R}$ is the market value of all risky assets relative to the market value of all assets. Define

$$
\begin{aligned}
&\theta_i = \Theta_i/\Theta_{M,R} \\
&\theta_{M,i} = \Theta_{M,i}/\Theta_{M,R} \\
&\theta_{i,-1} = \Theta_{i,-1}/\Theta_{M,R}.
\end{aligned}
\tag{19}
$$

Then, with some simplification, the equation 17 to 18 can be rewritten as

$$
\begin{aligned}
&\theta_i\Bigg(\mu_i + \tfrac{\nu\zeta_i}{\nu-2} - R_f\Bigg) - \tfrac{\hat{\gamma}}{2}\Bigg(\theta_i^2\sigma_i^2 + 2\theta_i(\beta_i\sigma_M^2 - \theta_{M,i}\sigma_i^2)\Bigg) \\
&\quad + \tfrac{\hat{\phi}}{6}\Bigg(\theta_i^3 skew_i\sigma_i^3 + 3\theta_i^2\Big(\beta_{\big((R_i-\bar{R}_i)^2,(R_M-\bar{R_M})\big)}\sigma_M^2 - \theta_{M,i}skew_i\sigma_i^3\Big) \\
&\qquad + 3\theta_i\Big(\beta_{\big((R_i-\bar{R}_i),(R_M-\bar{R_M})^2\big)}\sigma^2_{(R_M-\bar{R_M})^2} - \theta_{M,i}^2 skew_i\sigma_i^3 \\
&\qquad\quad - 2\theta_{M,i}\big(\beta_{\big((R_i-\bar{R}_i)^2,(R_M-\bar{R_M})\big)}\sigma_M^2 - \theta_{M,i}skew_i\sigma_i^3\big)\Big)\Bigg) \\
&\quad - \lambda\hat{b}\int_{-\infty}^{R_f-\theta_{i,-1}g_i/\theta_i}(\theta_i(R_f - R_i) - \theta_{i,-1}g_i)^\alpha dw(P(R_i)) \\
&\quad - \hat{b}\int_{R_f-\theta_{i,-1}g_i/\theta_i}^{\infty}(\theta_i(R_i - R_f) + \theta_{i,-1}g_i)^\alpha dw(1 - P(R_i)),
\end{aligned}
\tag{20}
$$

where

$$
\hat{\gamma} = \gamma W_0\Theta_{M,R}, \quad \hat{\phi} = \phi W_0^2\Theta_{M,R}^2, \quad \hat{b} = bW_0^{\alpha-1}\Theta_{M,R}^{\alpha-1}. \tag{21}
$$

The β_i is regression coefficient between return of stock i and market return; $\beta_{\big((R_i-\bar{R}_i)^2,(R_M-\bar{R_M})\big)}$ is regression coefficient between $(R_i - \bar{R}_i)^2$ and $R_M - \bar{R_M}$; and $\beta_{\big((R_i-\bar{R}_i),(R_M-\bar{R_M})^2\big)}$ is regression coefficient between $(R_i - \bar{R}_i)$ and $(R_M - \bar{R_M})^2$. The σ_i and $skew_i$ are the standard deviation and skewness of stock i's return, and σ_M is the standard deviation of market return. Same as equation 17, this objective function has either a unique global maximum at $\theta_i = \theta_{M,i}$ or two global maxima that straddle $\theta_{M,i}$. In Appendix A, we show the process to derive equation 20 from equation 17.

4 Model Implication

Using the model, we can restrict the inputs to see how a specific difference in stock's characteristics affects the related premium and how the characteristics together influence the premium. In this section, we look at coskewness premium based on PSSR and idiosyncratic skewness premium based on Expected IdioSkew, and discuss the implication of the model. We also analyze the mechanism of the implication from the model.

4.1 Parameter Value and Model Input

To see the implication of our model on anomaly premium, we proceed as follows. Since the prediction of stock returns depends on the stock's weight in the market portfolio, we consider an economy with $N = 1,000$ stocks rather than $N = 10$; then, each anomaly decile contains 100 stocks. We sort the stocks in ascending order by the characteristics we want to test, such as coskewness, so the stocks 1 to 100 that has the lowest value of this characteristics belong to decile 1; similarly, stocks 101 to 200 belong to decile 2, and so on. To make the calculation simpler and more representative, we assume all stocks in a given decile are identical, which means stocks in a given decile have the same characteristics, namely, the empirical characteristics of the typical stock in that anomaly decile. For each decile, we randomly pick up a stock and compute its expected return using our model. Since all stocks in a given decile are identical, the expected return is the same one for all stocks in that decile, and it is also the average return of that anomaly decile. By calculating the expected return difference between decile 1 and decile 10, we can know the premium of the anomaly, and our implications come from the process of varying skewness preference term weight and prospect theory term weight.

Equation 20 shows that, in order to determine the location parameter μ_i and therefore the expected return of stock i, we need to set four groups of model inputs and parameters. The value of these inputs and parameters for benchmark results are shown in Table 1 and

Table 2 respectively. First, we need to set parameters related to the market returns, including σ_M and $\sigma_{(R_M-\bar{R_M})^2}$. We set $\sigma_M = 0.25$ following Barberis et al. (2021) to have reasonable Sharpe ratio and market premium within the range of weights of each part of aversion or preference[18], and we use $\sigma_{(R_M-\bar{R_M})^2}$ calculated using monthly market return data from July 1963 to December 2021.

[insert Table 1 here]

Second, we also need to set the values for weights of each part of aversion or preference, which are $\hat{\gamma}$, $\hat{\phi}$, and $\hat{b}$. By varying these weight, we can learn the relationship between the terms. To have a reasonable range for the weights, we set the weights that are comparable to the ones used in literature. In Mitton and Vorkink (2007) model, which only includes mean, risk-aversion, and skewness preference, $\hat{\gamma}$ is 0.8 and $\hat{\phi}$ is 2.4. In Barberis et al. (2021) mode, which only includes mean, risk-aversion, and prospect theory term, $\hat{\gamma}$ is 0.6 and $\hat{b}$ is 0.6. In our benchmark results, we set $\hat{\gamma} = 0.6$, $\hat{\phi} \in \{0, 0.6, 1.2, 1.8, 2.4, 3\}$, and $\hat{b} \in \{0, 0.2, 0.4, 0.6, 0.8, 1\}$ so that the weight for each term is comparable to the weight used in literature. We also try parameters in broader range, such as greater weights for each parts, but they do not affect the implication of our model.

Third, we set the values for other parameters that depend on assumptions and are fixed in the model, including ν, α, λ, and δ. Following Barberis et al. (2021), we set $\nu = 7.5$, which provide a reasonable degree of fat-tailedness in stock returns. We also set $(\alpha, \lambda, \delta) = (0.7, 0.65, 1.5)$ to reflect the various findings in literature, such as Tversky and Kahneman (1992), Chapman et al. (2018), and Walasek et al. (2018). One potential issue is that Hu et al. (2022) find probability weighting parameter $\delta < 1$ and $\delta > 1$, which represents overweight and underweight of tail events, could lead to different implications to stock returns. We focus on the $\delta < 1$ case as classic prospect theory literature.

[18]In Barberis et al. (2021), when $\hat{\gamma} = 0.6$, $\hat{\phi} = 0$, and $\hat{b} = 0$, the market premium is about 0.6 and the Sharpe ratio is 12.5% lower than the market portfolio. We also use $\sigma_M = 0.1543$, which is calculated using monthly market return data from July 1963 to December 2021. The results are close and the implication is same.

[insert Table 2 here]

Fourth, we need to know a typical stock's characteristics in each decile, which are S_i, ζ_i, σ_i, $skew_i$, β_i, $\beta_{\left((R_i-\bar{R_i})^2,(R_M-\bar{R_M})\right)}$, $\beta_{\left((R_i-\bar{R_i}),(R_M-\bar{R_M})^2\right)}$, g_i, and $\theta_{M,i}$. The value of S_i and ζ_i are calculated using equation 15 and 16 given a fixed degree of freedome ν and model inputs, σ_i and $skew_i$. Other inputs are computed as follows. For a specific premium, such as coskewness, we rank all stocks listed on the NYSE, NASDAQ, and Amex on their predicted systematic skewness (PSS) (or Expected IdioSkew for idiosyncratic skewness)[19] and group them into deciles in each month from July 1963 to December 2021. As a result, Decile 1 corresponds to stocks with the lowest PSS and decile 10 corresponds to stocks with the highest PSS. Then, we calculate the following stock-level characteristics. The σ_i and $skew_i$ are the annualized standard deviation and skewness and calculated using stock i's past five years monthly data[20]. The β_i, $\beta_{\left((R_i-\bar{R_i})^2,(R_M-\bar{R_M})\right)}$, and $\beta_{\left((R_i-\bar{R_i}),(R_M-\bar{R_M})^2\right)}$ are regression coefficient between return of stock i and market return, regression coefficient between $(R_i-\bar{R_i})^2$ and $R_M-\bar{R_M}$ and regression coefficient between $(R_i-\bar{R_i})$ and $(R_M-\bar{R_M})^2$, respectively, and are calculated using stock i's future twelve months daily data[21]. The g_i is the capital gain overhang and are calculated following Barberis et al. (2021)[22]. The θ_i is the market holding and is calculated as stock i's last month market capitalization divided by the sum of market capitalization for all stocks. To get a typical stock's characteristics in each decile, we average the stock-level characteristics to different levels to achieve different purposes. For characteristics that are studied, we average stock-level results within each decile in each month and then average the decile results in all months, which in turn gives

[19] The calculation of the coskewness measure follows Langlois (2020). The calculation of the idiosyncratic skewness measure follows Boyer et al. (2010).

[20] The cross-section pattern is same if we use future twelve months daily data or future three years monthly data. The magnitudes of standard deviation and skewness are close for past five years monthly data and future three years monthly data. We do not use future twelve month daily data because the independent assumption used in annualization will erase almost all skewness.

[21] One may use past five years monthly data or future three years monthly data as well. The difference among deciles will be larger than our inputs but will not affect the results much. The future twelve months daily data is also a better representative for investors expectation according to Barberis et al. (2021).

[22] There are other ways to calculate capital gain overhangs, such as Grinblatt and Han (2005) and Frazzini (2006). We follow Barberis et al. (2021) to compute it as $(P_i - R_i)/R_i$, where P_i is stock i's current price and R_i is the average purchase price, because it a more precise match for g_i in the model.

us a typical stock's characteristics as inputs that are different among deciles (decile-level). For characteristics that are set in the same value for all deciles, we average stock-level results for all stocks in each month and then average the time series to get a typical stock's characteristics as inputs that are same in all deciles (market-level). Market holdings are a little bit different. Since we assume $N = 1,000$ stocks in economy, which leads to 100 stocks in each decile, our inputs for each decile is 0.1. We divide it by 100 in the model computation process, so the $\theta_i = 0.001$ for each stock.

4.2 Coskewness Premium and Idiosyncratic Skewness Premium

To study the premiums in the environment that the premiums are drived only by the corresponding characteristics, we allow only the key characteristics to be different among deciles. Specifically, coskewness premium is ideally a premium that is caused by the co-moment between stock i's return and the square of market return, so we keep $\beta_{\left((R_i-\bar{R_i}),(R_M-\bar{R_M})^2\right)}$ originally from the empirical results at decile-level but set other inputs at market-level. Idiosyncratic skewness premium is ideally a premium that is from individual skewness, so we only keep $skew_i$ originally from the empirical results at decile-level but set other inputs at market-level. The inputs for coskewness premium and idiosyncratic skewness premium are reported in Table 2. Using the parameters and inputs described in Section 4.1, we calculate the model-predicted return spreads between decile 1 and decile 10 for coskewness premium and idiosyncratic skewness.

Coskewness premium is paid for the co-moment of individual stocks that boost the portfolio return skewness due to the existence of skewness preference. Empirically, it is obtained by sorting stocks based on measure of coskewness, so we keep the original empirical results for the related co-moment column only. As shown in Table 2, the empirical method does work well in sorting the stock in ascending order by the co-moment.

[insert Figure 3 here]

Figure 3 presents the model-predicted return spreads for coskewness premium by varying $\hat{\phi}$ and $\hat{b}$. Each column of the symbols, which are in different colors and shapes, stands for the predicted return spreads with different $\hat{\phi}$ but same $\hat{b}$. In contrast, each row of the symbols, which are in same colors and shapes, presents for the predicted return spreads with same $\hat{\phi}$ but different $\hat{b}$. As expected, holding $\hat{b}$ in constant, the spread returns increase with $\hat{\phi}$, which indicates that skewness preference under expected utility theory help to explain coskewness premium. In addition, holding $\hat{\phi} = 0$, the spread returns does not change for different $\hat{\phi}$, meaning that prospect theory term does not solely generate coskewness premium if individual stock skewness are same. This outcome should be true intuitively since if investors do not care about portfolio skewness while the individual skewness is same among stocks, the coskewness premium will be zero for all weights of prospect theory term even if the co-moments are different. However, interestingly, holding the weights of $\hat{\phi}$ in constant at a non-zero level, the increase of $\hat{b}$ decreases the implied return spread. For example, if we keep $\hat{\phi} = 2.4$, the increase of $\hat{b}$ from 0 to 1.0 decrease the return spread by about 50 percent. This outcome means that even the individual stock skewness are same among stocks, the existence of prospect theory term lowers the premium for coskewness as long as the investors also evaluate the utility from a stock under prospect theory to some extent. This result implies that skewness preference under expected utility theory and prospect theory work against with each other in explaining coskewness premium when both of them exist. We will have further discussion about the mechanism of this result in section 4.3.

Idiosyncratic skewness premium is paid for the utility from individual stock's skewness only without contribution from co-moments to other assets[23]. Thus, we keep the original empirical results for individual skewness column only and set other inputs at the same level, especially for the co-moments. The empirical results shown in Table 2 also indicate that the empirical method works well in sorting the stock in ascending order by individual skewness.

[insert Figure 4 here]

[23]The idiosyncratic skewness is empirically computed using the residue of Fama and French (1993) to remove the impact of systematic risks. In our model, such systematic risks come from co-moments

Figure 4 displays the model-predicted return spreads for idiosyncratic skewness premium. The interpretation of columns and rows of symbols are same as the interpretation in Figure 3. As expected, the model-predicted premium increases with the weight of prospect theory term. Moreover, the return spread increase very slightly with the weight of skewness preference. This small increase is reasonable since a stock can not only increase the portfolio skewness by its co-moments with other assets but also increase the portfolio skewness by its individual skewness. Nonetheless, the small increase also illustrates that individual skewness does not play critical role in boosting utility from portfolio skewness.

The model implies that higher weight on skewness preference term leads to greater coskewness premium, but higher weight on prospect theory term reduces coskewness premium, indicating a competing relationship between the two terms in explaining coskewness premium. The model also implies that both higher weight on prospect theory term and higher weight on skewness preference can lead to greater idiosyncratic skewness premium, but the prospect theory term dominate the increase.

4.3 Mechanism

While the results of idiosyncratic skewness premium is intuitive, the result of coskewness premium is weird: if investors evaluate risks of individual stock by prospect theory to some extent, why do they pay smaller coskewness premium even if the individual stock skewness are same among stocks? We now answer this question in this section.

In our equilibrium, investors determine the holdings of an asset in the following way. First, they evaluate the utility from each components in the objective function. Second, they add up all the components to get a function of utility with respect to holdings. Last, they maximize the total utility by searching the optimal holdings. However, this process is not achieved at once: the optimal holdings and the market clearing return affect each other, so the equilibrium is achieved by doing the process many times to find a balance point. The process is relatively easy if investors have only mean, variance, and skewness in the system as the later two can be set without influence from the mean part. However, it becomes more complicated as we add the prospect theory term into the objective function since it is also affected by the expected return. We show the process in the following examples.

By decomposing the total utility into different parts based on the preferences and aversion, Figure 5, 6, and 7 show the decomposition of the coskewness premium for three cases respectively: $\hat{\gamma} = 0.6 \quad \& \quad \hat{\phi} = 2.4 \quad \& \quad \hat{b} = 0.6$, $\hat{\gamma} = 0.6 \quad \& \quad \hat{\phi} = 2.4 \quad \& \quad \hat{b} = 0$, and $\hat{\gamma} = 0.6 \quad \& \quad \hat{\phi} = 0 \quad \& \quad \hat{b} = 0.6$. All the three figures are interpreted as follows. In each panel, the red solid line in each panel presents the utility values for decile 1 and the blue dash line presents the ones for decile 10. The red circles present the equilibria for decile 1 and the blue cicles present the equilibria for decile 10. The x-axis measures the holdings (θ_i) for decile 1 or decile 10 and the y-axis in each panel measures the utility value from the components. The top left panel is for the total utility from the objective function. The top middle panel is for the utility from the sum of variance, skewness, and prospect theory parts. The top right panel is for the utility from expected return (mean). The bottom left is for the variance (risk aversion). The bottom middle is for the preference of skewness under expected utility theory. The bottom right is for the preference of skewness under prospect theory. We discuss about the mechanism of the relationship between skewness preference term and prospect theory term by comparing the three figures.

[insert Figure 5 here]

[insert Figure 6 here]

[insert Figure 7 here]

Figure 5 is a typical example that investors have non-zero weights on all the components. Since variance part and skewness preference part have nothing to do with the mean, the shape of the two parts can be determined easily based on the inputs. The two curves in variance part are overlapping because we set the standard deviation and CAPM beta at the same level for both deciles. The two curves in skewness part are different because the co-moments are different. Decile 10 has greater co-moments, so it generates greater utility when the holdings are greater than zero and less utility when the holdings are less than zero. If we don't have prospect theory term, then the equilibrium turns into the one as in Figure 6, and the curve for the sum of the two components is concave in reasonable domain. By varying the expected return, which rotate the straight line in mean part, the utility can be rotated to a specific degree that the homogeneous holding is same as the market holding, which clears the market. From Figure 6, we know that decile 10 in general provide more utility and hence has a lower expected return. Nonetheless, with a sufficient large weight on prospect theory term, the curve for the sum of variance, skewness, and prospect theory loses the concavity, and investors could have heterogeneous holdings to clear the market. Thus, the expected returns for the two deciles will change. From Figure 5, the curve for prospect theory in decile 10 is actually lower than the curve in decile 1, so the return spread (decile 1 minus decile 10) between the two deciles decreases to adjust to this change. Why the curve for prospect theory in decile 10 has to be lower than the one in decile 1? Let's compare Figure 7 and Figure 5. In Figure 7, the weight on skewness preference is zero, so all inputs used in the computation are same. As a result, although we have heterogeneous holdings, the expected return for the two deciles are same as expected. Now, we add the skewness preference as in Figure 5. For decile 10, we can think of the way that the additional positive utility is absorbed by rotating the curve for mean part first[24]: the more utility that skewness preference gives, the less expected return that the stock will have. As the expected return decreases, the μ_i in equation 14 becomes smaller as the other parameters hold at the same level[25]. It means that the return distribution have the same shape but is shifted parallel to the left. Then, more proportion of the distribution are treated as losses, so the utility from prospect theory part is smaller. However, this channel does not work alone since the rotating of the curve of prospect theory also decrease the utility. That is the reason why the

[24]Note that from equation 20, we know that the expected return is the slope of the mean part curve.
[25]Other parameters are decided in equation 15 and equation 16, where these two inputs do not change.

holdings also changes. The new equilibrium is achieved when all the parts reach to a new balance as we finally see in Figure 5: investors are more under-diversified (further away from market holding) and the return spread is smaller. This technical analysis can also induce the intuition about why the coskewness premium decreases with the increase of weight on prospect theory term.

The intuition of the competing effect is simple. First, investors have to balance among all terms they care, so the increase of the importance of one term will make the other one relatively less important under same holdings. Second, how utility from skewness preference under expected utility theory affect utility from prospect theory? The skewness preference changes the relative attractiveness of the stock evaluated by prospect theory. If the portfolio has already "gambling" enough, a specific gamble will not be very tempting anymore. Third, how utility from prospect theory term affect utility from skewness preference? The preference based on prospect theory significantly increase the underdiversification degree to the portfolio, so the utility from portfolio skewness changes with holdings, making the premium varies. Due to these channels, the coskewness premium displays the pattern we seen in the figure.

5 Conclusion

Expected utility theory and prospect theory are two important frameworks to study skewness preference, from investor's rationality and irrationality respectively in evaluating individual assets. With the development of mean-variance-skewness utility, prospect theory, and narrow framing in literature, we are able to incorporate both frameworks into one homogeneous preference model by two separate terms.

In this paper, we use such model to study the coskewness premium and idiosyncratic skewness premium. We show that the two terms play different roles in each premium, and more importantly, they can compete with each other in explaining the coskewness premium that intuitively should mainly attribute to portfolio skewness.

Appendix A Process to derive equation 20

Here, we derive equation 20 from equation 17. Substituting 19 into 17 and applying definitions of weights in 21, we obtain

$$
\begin{aligned}
&\theta_i\Big(\mu_i + \tfrac{\nu\zeta_i}{\nu-2} - R_f\Big) - \tfrac{\hat{\gamma}}{2}\Big(\theta_i^2 E[(R_i-\bar{R}_i)^2] + 2\theta_i \sum_{j\neq i}\theta_{M,j}E[(R_i-\bar{R}_i)(R_j-\bar{R}_j)]\Big)\\
&\quad + \tfrac{\hat{\phi}}{6}\Big(\theta_i^3 E[(R_i-\bar{R}_i)^3] + 3\theta_i^2\sum_{j\neq i}\theta_{M,j}E[(R_i-\bar{R}_i)^2(R_j-\bar{R}_j)]\\
&\qquad + 3\theta_i\sum_{j\neq i}\theta_{M,j}^2 E[(R_i-\bar{R}_i)(R_j-\bar{R}_j)^2]\\
&\qquad + 3\theta_i\sum_{j\neq i}\sum_{k\neq\{j,i\}}\theta_{M,j}\theta_{M,k}E[(R_i-\bar{R}_i)(R_j-\bar{R}_j)(R_k-\bar{R}_k)]\Big)\\
&\quad - \lambda\hat{b}\int_{-\infty}^{R_f-\theta_{i,-1}g_i/\theta_i}(\theta_i(R_f-R_i)-\theta_{i,-1}g_i)^\alpha dw(P(R_i))\\
&\quad - \hat{b}\int_{R_f-\Theta_{i,-1}g_i/\theta_i}^{\infty}(\theta_i(R_i-R_f)+\theta_{i,-1}g_i)^\alpha dw(1-P(R_i)).
\end{aligned} \tag{22}
$$

Expanding $\sum_{j\neq i}\theta_{M,j}E[(R_i-\bar{R}_i)(R_j-\bar{R}_j)]$, $\sum_{j\neq i}\theta_{M,j}E[(R_i-\bar{R}_i)^2(R_j-\bar{R}_j)]$, and $\sum_{j\neq i}\theta_{M,j}^2E[(R_i-\bar{R}_i)(R_j-\bar{R}_j)^2]$ in equation 22 by using market return that subtracted R_i, we have

$$
\begin{aligned}
&\theta_i\Big(\mu_i + \tfrac{\nu\zeta_i}{\nu-2} - R_f\Big) - \tfrac{\hat{\gamma}}{2}\Big(\theta_i^2 E[(R_i-\bar{R}_i)^2] + 2\theta_i\Big(E\left[(R_i-\bar{R}_i)(R_M-\bar{R_M})\right] - \theta_{M,i}E\left[(R_i-\bar{R}_i)^2\right]\Big)\Big)\\
&\quad + \tfrac{\hat{\phi}}{6}\Big(\theta_i^3E[(R_i-\bar{R}_i)^3] + 3\theta_i^2\Big(E\left[(R_i-\bar{R}_i)^2(R_M-\bar{R_M})\right] - \theta_{M,i}E\left[(R_i-\bar{R}_i)^3\right]\Big)\\
&\qquad + 3\theta_i\Big(E\left[(R_i-\bar{R}_i)(R_M-\bar{R_M})^2\right] - \theta_{M,i}^2E\left[(R_i-\bar{R}_i)^3\right]\\
&\qquad\quad - 2\theta_{M,i}\big(E\left[(R_i-\bar{R}_i)^2(R_M-\bar{R_M})\right] - \theta_{M,i}E\left[(R_i-\bar{R}_i)^3\right]\big)\Big)\Big)\\
&\quad - \lambda\hat{b}\int_{-\infty}^{R_f-\theta_{i,-1}g_i/\theta_i}(\theta_i(R_f-R_i)-\theta_{i,-1}g_i)^\alpha dw(P(R_i))\\
&\quad - \hat{b}\int_{R_f-\Theta_{i,-1}g_i/\theta_i}^{\infty}(\theta_i(R_i-R_f)+\theta_{i,-1}g_i)^\alpha dw(1-P(R_i)).
\end{aligned} \tag{23}
$$

Since

$$
\begin{aligned}
&Cov\left((R_i-\bar{R_i})^2,(R_M-\bar{R_M})\right)\\
&=E\left[\left((R_i-\bar{R_i})^2-\overline{(R_i-\bar{R_i})^2}\right)\left((R_M-\bar{R_M})-\overline{R_M-\bar{R_M}}\right)\right]\\
&=E\left[\left((R_i-\bar{R_i})^2-\overline{(R_i-\bar{R_i})^2}\right)(R_M-\bar{R_M})\right]\\
&=E\left[(R_i-\bar{R_i})^2(R_M-\bar{R_M})\right]-E\left[\overline{(R_i-\bar{R_i})^2}(R_M-\bar{R_M})\right]\\
&=E\left[(R_i-\bar{R_i})^2(R_M-\bar{R_M})\right]-\overline{(R_i-\bar{R_i})^2}E[R_M-\bar{R_M}]\\
&=E\left[(R_i-\bar{R_i})^2(R_M-\bar{R_M})\right]-\overline{(R_i-\bar{R_i})^2}E[R_M]+\overline{(R_i-\bar{R_i})^2}E[\bar{R_M}]\\
&=E\left[(R_i-\bar{R_i})^2(R_M-\bar{R_M})\right],
\end{aligned}
\tag{24}
$$

and

$$
\begin{aligned}
&Cov\left((R_i-\bar{R_i}),(R_M-\bar{R_M})^2\right)\\
&=E\left[\left((R_i-\bar{R_i})-\overline{R_i-\bar{R_i}}\right)\left((R_M-\bar{R_M})^2-\overline{(R_M-\bar{R_M})^2}\right)\right]\\
&=E\left[(R_i-\bar{R_i})\left((R_M-\bar{R_M})^2-\overline{(R_M-\bar{R_M})^2}\right)\right]\\
&=E\left[(R_i-\bar{R_i})(R_M-\bar{R_M})^2)\right]-E\left[(R_i-\bar{R_i})\overline{(R_M-\bar{R_M})^2}\right]\\
&=E\left[(R_i-\bar{R_i})(R_M-\bar{R_M})^2)\right]-\overline{(R_M-\bar{R_M})^2}E[R_i-\bar{R_i}]\\
&=E\left[(R_i-\bar{R_i})(R_M-\bar{R_M})^2)\right]-\overline{(R_M-\bar{R_M})^2}E[R_i]+\overline{(R_M-\bar{R_M})^2}E[\bar{R_i}]\\
&=E\left[(R_i-\bar{R_i})(R_M-\bar{R_M})^2)\right],
\end{aligned}
\tag{25}
$$

equation 23 becomes

$$
\begin{aligned}
&\theta_i\Big(\mu_i + \tfrac{\nu\zeta_i}{\nu-2} - R_f\Big) - \tfrac{\hat{\gamma}}{2}\Big(\theta_i^2 E[(R_i-\bar{R}_i)^2] + 2\theta_i\Big(Cov(R_i, R_M) - \theta_{M,i}E\left[(R_i-\bar{R}_i)^2\right]\Big)\Big) \\
&\quad + \tfrac{\hat{\phi}}{6}\Big(\theta_i^3 E[(R_i-\bar{R}_i)^3] + 3\theta_i^2\Big(Cov\left((R_i-\bar{R}_i)^2, (R_M-\bar{R_M})\right) - \theta_{M,i}E\left[(R_i-\bar{R}_i)^3\right]\Big) \\
&\qquad + 3\theta_i\Big(Cov\left((R_i-\bar{R}_i), (R_M-\bar{R_M})^2\right) - \theta_{M,i}^2 E\left[(R_i-\bar{R}_i)^3\right] \\
&\qquad\quad - 2\theta_{M,i}\big(Cov\left((R_i-\bar{R}_i)^2, (R_M-\bar{R_M})\right) - \theta_{M,i}E\left[(R_i-\bar{R}_i)^3\right]\big)\Big)\Big) \\
&\quad - \lambda\hat{b}\int_{-\infty}^{R_f-\theta_{i,-1}g_i/\theta_i}(\theta_i(R_f-R_i) - \theta_{i,-1}g_i)^\alpha dw(P(R_i)) \\
&\quad - \hat{b}\int_{R_f-\Theta_{i,-1}g_i/\theta_i}^{\infty}(\theta_i(R_i-R_f) + \theta_{i,-1}g_i)^\alpha dw(1-P(R_i)).
\end{aligned}
\tag{26}
$$

Plug

$$
\begin{aligned}
&E[(R_i-\bar{R}_i)^2] = \sigma_i^2 \\
&E[(R_i-\bar{R}_i)^3] = skew_i\sigma_i^3 \\
&Cov(R_i, R_M) = \beta_i\sigma_M^2 \\
&Cov\left((R_i-\bar{R}_i)^2, (R_M-\bar{R_M})\right) = \beta_{(R_i-\bar{R}_i)^2,(R_M-\bar{R_M})}\sigma_M^2 \\
&Cov\left((R_i-\bar{R}_i), (R_M-\bar{R_M})^2\right) = \beta_{(R_i-\bar{R}_i),(R_M-\bar{R_M})^2}\sigma_{(R_M-\bar{R_M})^2}
\end{aligned}
\tag{27}
$$

into equation 26, we get equation 20, which is used in computation.

References

Baele, L., Driessen, J., Ebert, S., Londono, J., Spalt, O., 2019. Cumulative prospect theory, option returns, and the variance premium. Review of Financial Studies 32, 3667–3723.

Bakshi, G., Kapadia, N., Madan, D., 2003. Stock return characteristics, skew laws, and the differential pricing of individual equity options. The Review of Financial Studies 16, 101–143.

Bali, T., Murray, S., 2013. Does risk-neutral skewness predict the cross section of equity option portfolio returns? Journal of Financial and Quantitative Analysis 48, 1145–1171.

Barberis, N., 2012. A model of casino gambling. Management Science 58, 35–51.

Barberis, N., 2013. Thirty years of prospect theory in economics: A review and assessment. Journal of Economic Perspectives 27, 173–196.

Barberis, N., Huang, M., 2001. Mental accounting, loss aversion, and individual stock returns. Journal of Finance 56, 1247–1292.

Barberis, N., Huang, M., 2008a. The loss aversion/narrow framing approach to the equity premium puzzle, in: Handbook of the Equity Risk Premium. Elsevier.

Barberis, N., Huang, M., 2008b. Stocks as lotteries: The implications of probability weighting for security prices. American Economic Review 98, 2066–2100.

Barberis, N., Huang, M., 2009. Preferences with frames: A new utility specification that allows for the framing of risks. Journal of Economic Dynamics & Control 33, 1555–1576.

Barberis, N., Huang, M., Santos, T., 2001. Prospect theory and asset prices. Quarterly Journal of Economics 116, 1–53.

Barberis, N., Huang, M., Thaler, R., 2006. Individual preferences, monetary gambles, and stock market participation: A case for narrow framing. American Economic Review 96, 1069–1090.

Barberis, N., Jin, L., Wang, B., 2021. Prospect theory and stock market anomalies. Journal of Finance 76, 2639–2687.

Barberis, N., Mukherjee, A., Wang, B., 2016. Prospect theory and stock returns: An empirical test. Review of Financial Studies 29, 3068–3107.

Barone-Adesi, G., 1985. Arbitrage equilibrium with skewed asset returns. Journal of Financial and Quantitative Analysis 20, 299–313.

Beedles, W., 1979. Return, dispersion, and skewness: Synthesis and investment strategy. Journal of Financial Research 2, 71–80.

Benartzi, S., Thaler, R., 1995. Myopic loss aversion and the equity premium puzzle. Quarterly Journal of Economics 110, 73–92.

Black, F., Jensen, M., Scholes, M., 1972. The capital asset pricing model: Some empirical tests, in: Studies in the Theory of Capital Markets. Praeger Publishers Inc., New York.

Blume, M., Friend, I., 1973. A new look at the capital asset pricing model. Journal of Finance 28, 19–33.

Boyer, B., Mitton, T., Vorkink, K., 2010. Expected idiosyncratic skewness. Review of Financial Studies 23, 169–202.

Brockett, P., Kahane, Y., 1992. Risk, return, skewness and preference. Management Science 38, 851–866.

Brunnermeier, M., Gollier, C., Parker, J., 2007. Optimal beliefs, asset prices, and the preference for skewed returns. American Economic Review 92, 159–165.

Brunnermeier, M., Parker, J., 2005. Optimal expectations. American Economic Review 95, 1092–1118.

Chapman, J., Snowberg, E., Wang, S., Camerer, C., 2018. Loss attitudes in the u.s. population: Evidence from dynamically optimized sequential experimentation. California Institute of Technology.

Christoffersen, P., Fournier, M., Jacobs, K., Karoui, M., 2021. Option-based estimation of the price of coskewness and cokurtosis risk. Journal of Financial and Quantitative Analysis 56, 65–91.

Chunhachinda, P., Dandapani, K., Hamid, S., Prakash, A., 1997. Portfolio selection and skewness: Evidence from international stock markets. Journal of Banking & Finance 21, 143–167.

Conine Jr, T., Tamarkin, M., 1981. On diversification given asymmetry in returns. Journal of Finance 36, 1143–1155.

Conrad, J., Dittmar, R., Ghysels, E., 2013. Ex ante skewness and expected stock returns. Journal of Finance 68, 85–124.

Dittmar, R., 2002. Nonlinear pricing kernels, kurtosis preference, and evidence from the cross section of equity returns. Journal of Finance 57, 369–403.

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

Frazzini, A., 2006. The disposition effect and underreaction to news. Journal of Finance 61, 2017–2046.

Friend, I., Blume, M., 1970. Measurement of portfolio performance under uncertainty. American Economic Review 60, 561–575.

Friend, I., Westerfield, R., 1980. Co-skewness and capital asset pricing. Journal of Finance 35, 897–913.

Grinblatt, M., Han, B., 2005. Prospect theory, mental accounting, and momentum. Journal of financial economics 78, 311–339.

Harvey, C., Siddique, A., 2000. Conditional skewness in asset pricing tests. Journal of Finance 55, 1263–1295.

Hu, J., Li, J., Zhao, F., 2022. Cumulative prospect theory, time-varying probability weighting and stock returns. University of Texas at Dallas.

Kahneman, D., Tversky, A., 1979. Prospect theory: An analysis of decision under risk. Econometrica 47, 263–291.

Kane, A., 1982. Skewness preference and portfolio choice. Journal of Financial and Quantitative Analysis 17, 15–25.

Kozhan, R., Neuberger, A., Schneider, P., 2013. The skew risk premium in the equity index market. Review of Financial Studies 26, 2174–2203.

Kraus, A., Litzenberger, R., 1976. Skewness preference and the valuation of risk assets. Journal of Finance 31, 1085–1100.

Langlois, H., 2020. Measuring skewness premia. Journal of Financial Economics 135, 399–424.

Markowitz, H., 1959. Portfolio Selection: Efficient Diversfication of Investments. Yale University Press, New Haven.

Mitton, T., Vorkink, K., 2007. Equilibrium underdiversification and the preference for skewness. Review of Financial Studies 20, 1255–1288.

Prakash, A., Chang, C., Pactwa, T., 2003. Selecting a portfolio with skewness: Recent evidence from us, european, and latin american equity markets. Journal of Banking & Finance 27, 1375–1390.

Rubinstein, M., 1973. The fundamental theorem of parameter-preference security valuation. Journal of Financial and Quantitative Analysis 8, 61–69.

Scott, R., Horvath, P., 1980. On the direction of preference for moments of higher order than the variance. Journal of Finance 35, 915–919.

Simaan, Y., 1993. Portfolio selection and asset pricing-three-parameter framework. Management Science 39, 568–577.

Simkowitz, M., Beedles, W., 1978. Diversification in a three-moment world. Journal of Financial and Quantitative Analysis 13, 927–941.

Tversky, A., Kahneman, D., 1981. The framing of decisions and the psychology of choice. Science 211, 453–458.

Tversky, A., Kahneman, D., 1992. Advances in prospect theory: Cumulative representation of uncertainty. Journal of Risk and Uncertainty 5, 297–323.

Walasek, L., Mullett, T., Stewart, N., 2018. A meta-analysis of loss aversion in risky contexts. University of Warwick.

Table 1
Model parameter values

Parameter	Variable	Value
SD of market return	σ_M	0.25
SD of $(R_M - \bar{R_M})^2$	$\sigma_{(R_M-\bar{R_M})^2}$	0.0135
Degree of fat-tailedness in stock returns	ν	7.5
Curvature of utility function	α	0.7
Degree of probability weighting	δ	0.65
Sensitivity to losses	λ	1.5
Risk-aversion coefficient	$\hat{\gamma}$	0.6
Portfolio-level skewness preference coefficient	$\hat{\phi}$	0(0.6)3
Stock-level skewness preference coefficient	$\hat{b}$	0(0.2)1

The table presents the parameter values used in the numerical solutions of the model. Standard deviations are calculated using market returns from July 1967 to December 2021. Portfolio-level skewness preference coefficient is used from 0 to 3 increased by 0.6 in each computation. Stock-level skewness preference coefficient is used from 0 to 1 increased by 0.2 in each computation.

Table 2
Model inputs

Decile	σ_i	$skew_i$	g_i	β_i	$\theta_{M,i} * 100$	$\beta_{\left((R_i-\bar{R_i})^2,(R_M-\bar{R_M})\right)}$	$\beta_{\left((R_i-\bar{R_i}),(R_M-\bar{R_M})^2\right)}$
Coskewness premium							
1	0.475	0.176	-0.017	0.988	0.100	-0.002	-12.099
2	0.475	0.176	-0.017	0.988	0.100	-0.002	-10.563
3	0.475	0.176	-0.017	0.988	0.100	-0.002	-9.301
4	0.475	0.176	-0.017	0.988	0.100	-0.002	-8.174
5	0.475	0.176	-0.017	0.988	0.100	-0.002	-7.495
6	0.475	0.176	-0.017	0.988	0.100	-0.002	-6.838
7	0.475	0.176	-0.017	0.988	0.100	-0.002	-6.270
8	0.475	0.176	-0.017	0.988	0.100	-0.002	-5.591
9	0.475	0.176	-0.017	0.988	0.100	-0.002	-4.866
10	0.475	0.176	-0.017	0.988	0.100	-0.002	-3.092
Idiosyncratic skewness premium							
1	0.475	0.056	-0.017	0.988	0.100	-0.002	-6.830
2	0.475	0.064	-0.017	0.988	0.100	-0.002	-6.830
3	0.475	0.083	-0.017	0.988	0.100	-0.002	-6.830
4	0.475	0.107	-0.017	0.988	0.100	-0.002	-6.830
5	0.475	0.136	-0.017	0.988	0.100	-0.002	-6.830
6	0.475	0.165	-0.017	0.988	0.100	-0.002	-6.830
7	0.475	0.195	-0.017	0.988	0.100	-0.002	-6.830
8	0.475	0.233	-0.017	0.988	0.100	-0.002	-6.830
9	0.475	0.291	-0.017	0.988	0.100	-0.002	-6.830
10	0.475	0.410	-0.017	0.988	0.100	-0.002	-6.830

The table reports the model inputs to calculate the corresponding premium. To save space in the table, we round the number to 0.001, but we use more digits in our model computation process. Standard Deviation and Skewness are annualized values.

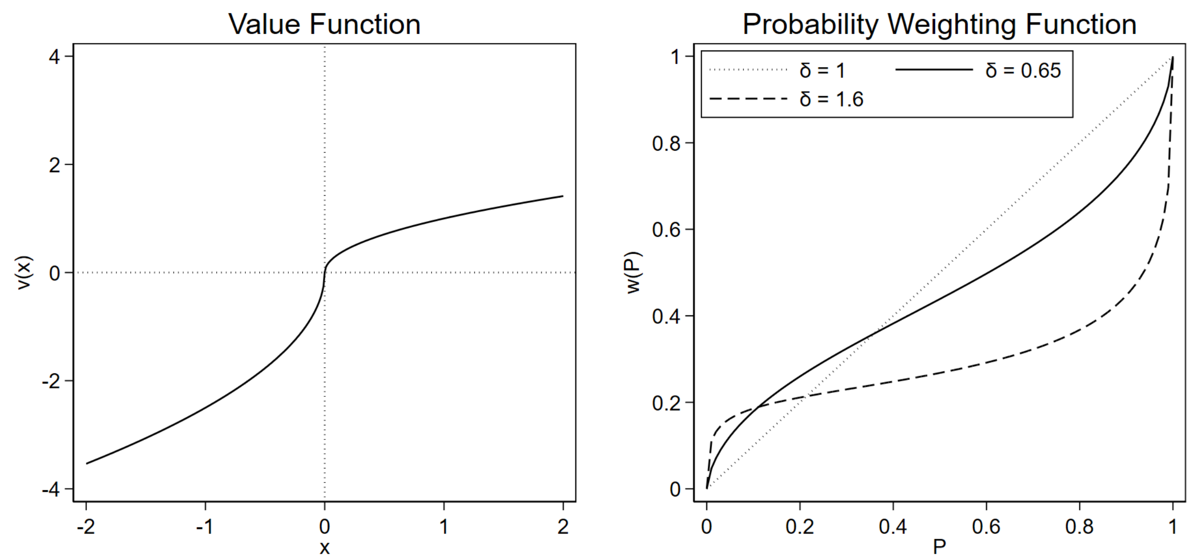

Figure 1
The prospect theory value function and probability weighting function
The left panel plots the value function proposed by Tversky and Kahneman (1992) as part of their cumulative prospect theory, namely, $v(x) = x^{\alpha}$ for $x \geq 0$ and $v(x) = -\lambda(-x)^{\alpha}$ for $x < 0$, for $\alpha = 0.5$ and $\lambda = 2.5$. The right panel plots the probability weighting function they propose, namely, $w(P) = P^{\delta}/(P^{\delta} + (1-P)^{\delta})^{1/\delta}$, for three different values of δ. The dotted line corresponds to $\delta = 1$, the solid line to $\delta = 0.65$, and the dash line to $\delta = 0.4$.

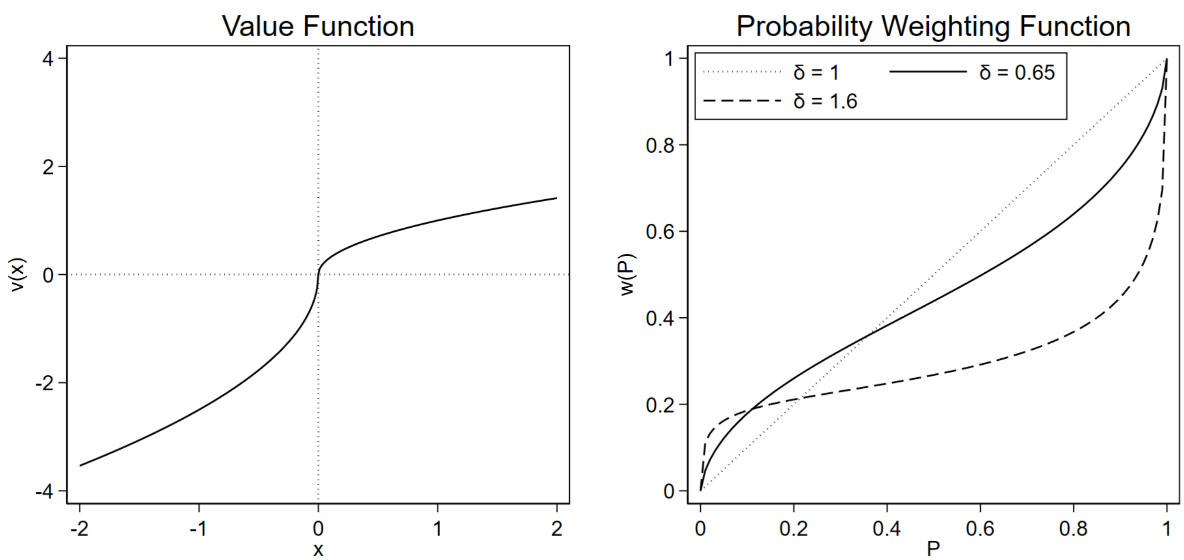

Figure 2
The prospect theory value function and probability weighting function
The left panel plots the value function proposed by Tversky and Kahneman (1992) as part of their cumulative prospect theory, namely, $v(x) = x^{\alpha}$ for $x \geq 0$ and $v(x) = -\lambda(-x)^{\alpha}$ for $x < 0$, for $\alpha = 0.5$ and $\lambda = 2.5$. The right panel plots the probability weighting function they propose, namely, $w(P) = P^{\delta}/(P^{\delta} + (1-P)^{\delta})^{1/\delta}$, for three different values of δ. The dotted line corresponds to $\delta = 1$, the solid line to $\delta = 0.65$, and the dash line to $\delta = 0.4$.

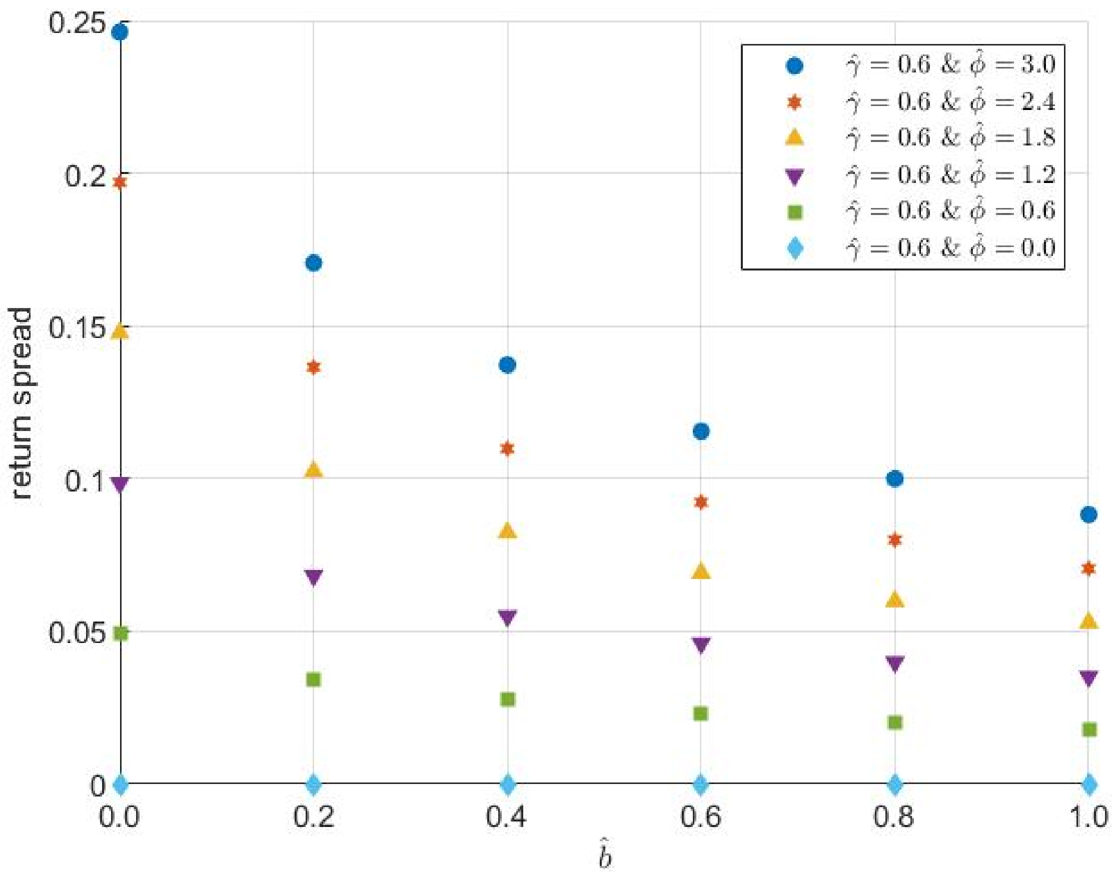

Figure 3
Model-predicted coskewness premium using uni-variate input
The figure plots the model-predicted coskewness premium using uni-variate input by varying the weights of portfolio-level skewness preference ($\hat{\phi}$) and stock-level skewness preference ($\hat{b}$). The x-axis measures the weight of stock-level skewness preference ($\hat{b}$) and the y-axis measures the model-predicted return spread in percent.

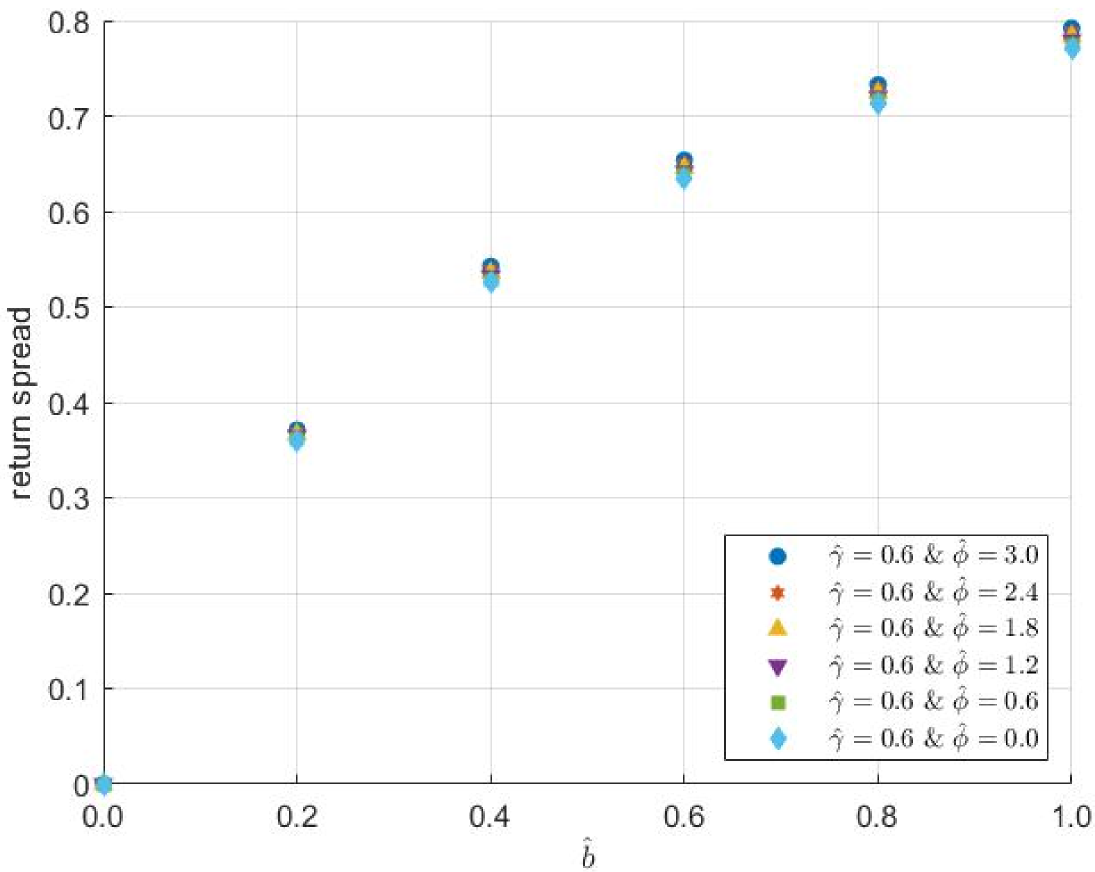

Figure 4
Model-predicted idiosyncratic skewness premium using uni-variate input
The figure plots the model-predicted idiosyncratic skewness premium using uni-variate input by varying the weights of portfolio-level skewness preference ($\hat{\phi}$) and stock-level skewness preference ($\hat{b}$). The x-axis measures the weight of stock-level skewness preference ($\hat{b}$) and the y-axis measures the model-predicted return spread in percent.

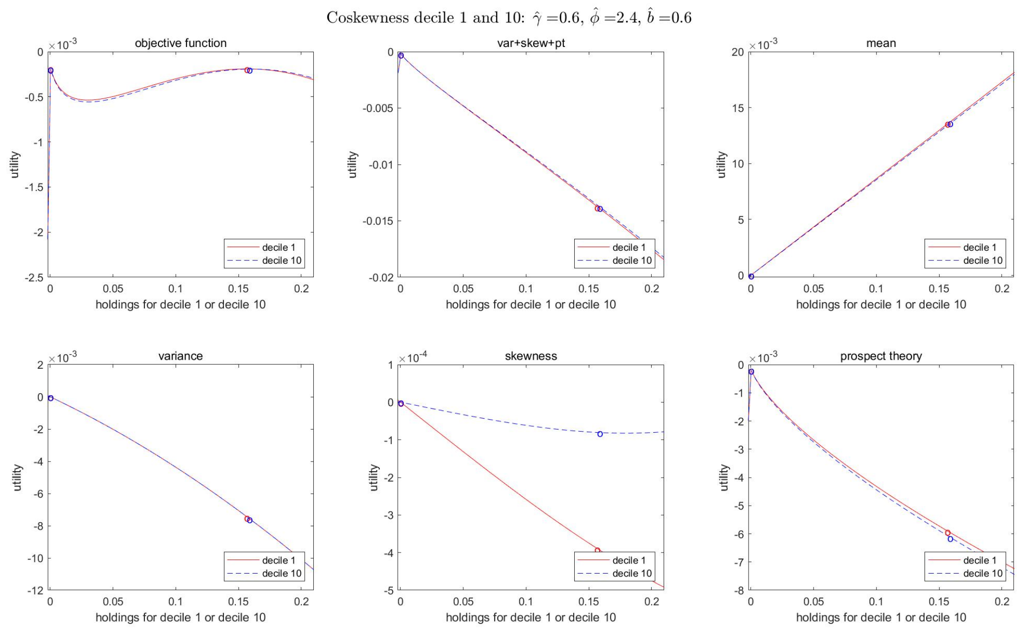

Figure 5
Decomposing objective function for coskewness premium with $\hat{\gamma} = 0.6$, $\hat{\phi} = 2.4$, and $\hat{b} = 0.6$

The figure plots the total utility from objective function and its components for coskewness premium with $\hat{\gamma} = 0.6$, $\hat{\phi} = 2.4$, and $\hat{b} = 0.6$. In each panel, the red solid line in each panel presents the utility values for decile 1 and the blue dash line presents the ones for decile 10. The red circles present the equilibria for decile 1 and the blue cicles present the equilibria for decile 10. The x-axis measures the holdings (θ_i) for decile 1 or decile 10 and the y-axis in each panel measures the utility value from the components. The top left panel is for the total utility from the objective function. The top middle panel is for the utility from the sum of variance, portfolio-level skewness, and stock-level skewness (prospect theory) parts. The top right panel is for the utility from expected return (mean). The bottom left is for the variance (risk aversion). The bottom middle is for the preference of portfolio-level skewness. The bottom right is for the preference of stock-level skewness (prospect theory value).

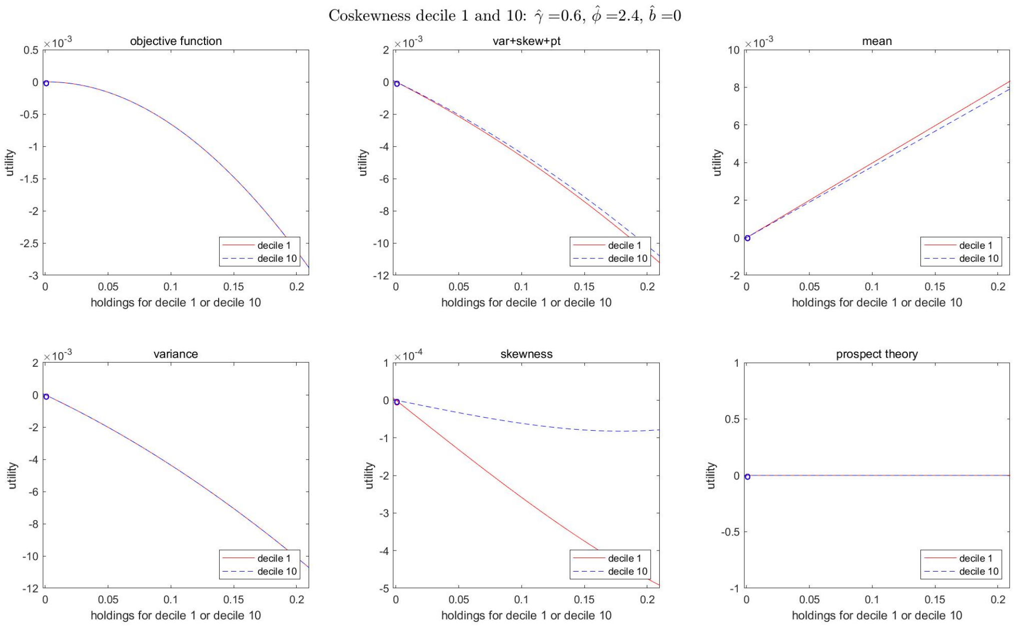

Figure 6
Decomposing objective function for coskewness premium with $\hat{\gamma} = 0.6$, $\hat{\phi} = 2.4$, and $\hat{b} = 0$

The figure plots the total utility from objective function and its components for coskewness premium with $\hat{\gamma} = 0.6$, $\hat{\phi} = 2.4$, and $\hat{b} = 0$. In each panel, the red solid line in each panel presents the utility values for decile 1 and the blue dash line presents the ones for decile 10. The red circles present the equilibria for decile 1 and the blue cicles present the equilibria for decile 10. The x-axis measures the holdings (θ_i) for decile 1 or decile 10 and the y-axis in each panel measures the utility value from the components. The top left panel is for the total utility from the objective function. The top middle panel is for the utility from the sum of variance, portfolio-level skewness, and stock-level skewness (prospect theory) parts. The top right panel is for the utility from expected return (mean). The bottom left is for the variance (risk aversion). The bottom middle is for the preference of portfolio-level skewness. The bottom right is for the preference of stock-level skewness (prospect theory value).

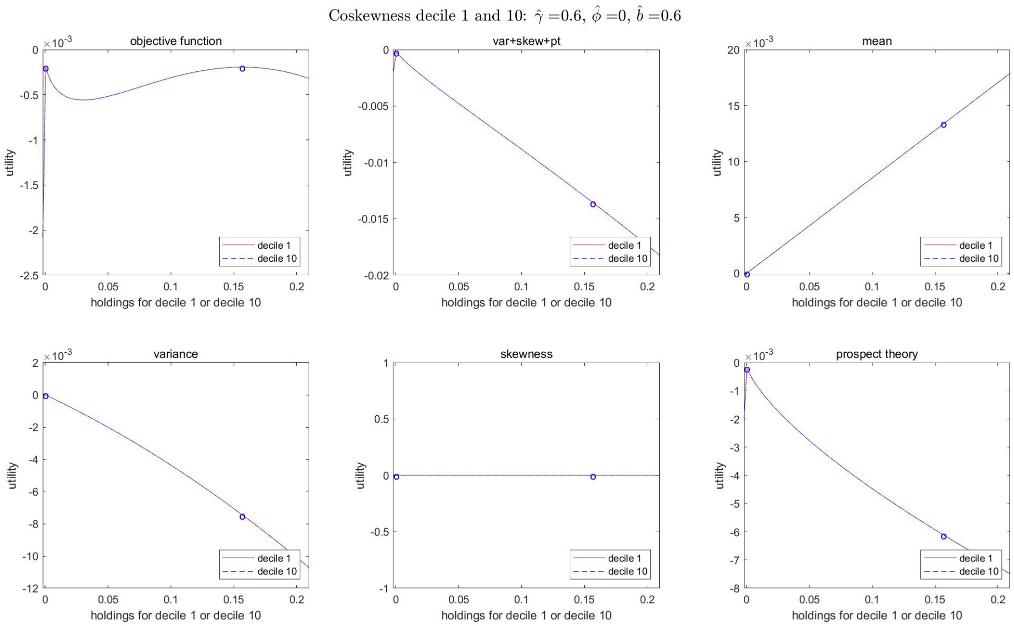

Figure 7
Decomposing objective function for coskewness premium with $\hat{\gamma} = 0.6$, $\hat{\phi} = 0$, and $\hat{b} = 0.6$
The figure plots the total utility from objective function and its components for coskewness premium with $\hat{\gamma} = 0.6$, $\hat{\phi} = 0$, and $\hat{b} = 0.6$. In each panel, the red solid line in each panel presents the utility values for decile 1 and the blue dash line presents the ones for decile 10. The red circles present the equilibria for decile 1 and the blue cicles present the equilibria for decile 10. The x-axis measures the holdings (θ_i) for decile 1 or decile 10 and the y-axis in each panel measures the utility value from the components. The top left panel is for the total utility from the objective function. The top middle panel is for the utility from the sum of variance, portfolio-level skewness, and stock-level skewness (prospect theory) parts. The top right panel is for the utility from expected return (mean). The bottom left is for the variance (risk aversion). The bottom middle is for the preference of portfolio-level skewness. The bottom right is for the preference of stock-level skewness (prospect theory value).