Portfolio Choice under Loss Aversion

Kai Li[1], Jun Liu[2] and Jiahui Shui[2]

[1]Macquarie Business School, Macquarie University, NSW 2109, Australia

[2]Rady School of Management, University of California San Diego, La Jolla, California 92093

kai.li@mq.edu.au, junliu@ucsd.edu, jishui@ucsd.edu

January 31, 2024

Abstract

This paper studies optimal portfolio choice under the loss aversion utility proposed in Tversky and Kahneman (1992). The loss aversion utility imposes lower penalty for large losses than the risk aversion utility functions. As a result, it tends to generate large risky positions and sometimes unbounded optimal portfolios, especially when asset returns are highly skewed (either positive or negative). Loss averse investors may hold assets with a zero risk premium and hold a positive amount of assets with a negative risk premium. Loss aversion can affect the optimal portfolio only when the reference point is sufficiently high, a case that is less studied in the literature, and the loss aversion utility becomes the standard HARA utility when the reference point is low.

Key words: Prospect theory, loss aversion, reference point, optimal portfolio choice, solution existence.

JEL Classification: C61, G11

1 Introduction

Prospect theory (Kahneman and Tversky, 1979; Tversky and Kahneman, 1992) is one of the most influential descriptive theories of decision-making under uncertainty. It has been widely used to explain different investor behaviors and financial market anomalies.[1] At the core of this theory is a utility function that features loss aversion, an observation that people are more sensitive to losses than to equivalent gains. This utility function is defined over gains and losses relative to a reference point, and it is concave for gains and convex for losses with an S shape ("diminishing sensitivity"). These features lead to several difficulties in applying prospect theory. First, the reference point is difficult to be identified (e.g., Barberis, 2013). In the literature, different referent points are used to explain different behaviors, and there is a lack of systematic understanding of their effects. Second, maximization problems under loss aversion may not be well formulated due to the nonconcavity of the utility function. The literature usually imposes certain constraints or utility variations to ensure bounded solutions. However, the optimal choice under the original loss aversion utility function of Tversky and Kahneman (1992) is largely overlooked.

In this paper, we take the loss aversion utility of Tversky and Kahneman (1992) at face value without any modification and study its implications for optimal choice. However, we acknowledge that in real world people face various frictions, and the constraints as imposed in the literature are important for practical interest. We study a standard static portfolio choice problem in complete markets, and solve for the optimal portfolios using the Cox and Huang's (1989) approach. We provide a comprehensive analysis of the effects of the reference point and the conditions under which loss aversion leads to bounded optimal portfolios. The starkly different implications of different reference points provide a systematic understanding

[1] Benartzi and Thaler (1995) find that loss aversion helps explain the equity premium puzzle due to the reluctance of agents to invest in stocks. Barberis, Huang and Santos (2001) show that loss aversion produces excess return volatility and low correlation between stock returns and consumption growth. Gomes (2005) uses loss aversion to explain the disposition effect and low equity market participation rates. Grinblatt and Han (2005) find that prospect theory helps understand the momentum effect as investors tend to hold on to their losing stocks. Barberis and Xiong (2009) find that loss aversion preferences can predict a disposition effect when the preferences are defined over realized gains and losses rather than over paper gains and losses. Li and Yang (2013) find that diminishing sensitivity predicts a disposition effect, price momentum, a reduced return volatility, and a positive return-volume correlation; loss aversion predicts the opposite.

of the loss aversion utility and the reference point and theoretical guidance for choosing and testing the reference point.

The first set of analyses in this paper is to analytically derive the existence conditions of the optimal portfolios under loss aversion. We show that there is no bounded optimal solution for half of the parameter space in which loss aversion is low. With insufficient penalty for large losses, an investor can achieve arbitrarily large expected utility by increasing terminal wealth in one state and at the same time decreasing in another under the budget constraint, leading to unbounded utility. However, there always exist (bounded) optimal solutions for the other half of the parameter space with sufficiently high loss aversion, even though the investor can be risk seeking. Under budget preserving, large gain in one state must be associated with large losses in other states. High loss aversion ensures that the utility loss in the latter states surpasses the utility gain in the former state, preventing the expected utility from approaching infinity.

The above results show that the nonexistence of optimal solution is essentially due to insufficient loss penalty of the loss aversion utility. In fact, with diminishing sensitivity, the loss aversion utility function imposes much lower penalty for losses than do risk aversion utility functions. For example, with the same reference point, the hyperbolic absolute risk aversion (HARA) utility as studied in Merton (1971) is identical to the loss aversion utility when wealth is higher than the reference point (gains) but has negative infinite (rather than finite) utility for losses. A loss averse investor is actually less averse to losses and hence trades more aggressively than the corresponding risk averse investor.

Furthermore, we show that the minimum value of the loss aversion coefficient above which optimal solutions exist increases without bound as asset returns are more skewed (either positive or negative) or as the number of states increases. A highly skewed asset allows the investor to achieve a high utility gain at certain states (by either longing or shorting the asset), and large numbers of states under complete markets provide the investor with more potential opportunities. As a result, in both cases, a high loss aversion is required to ensure finite expected utility. We find that the optimal portfolio choice problem under incomplete markets are more likely to have bounded solutions than under complete markets, since the former allow less investment opportunities. Our results show that a loss aversion parameter greater than one (investors are more sensitive to losses than to gains), as usually assumed in the literature, is a necessary but not a sufficient condition for existence of solutions.

Next, we examine the effects of the reference point on the optimal portfolios. We find that when the optimal solutions exist, their properties depend crucially on the ratio of the reference point to the investor's initial wealth level. Hereafter, we simply call this ratio the reference point since we can normalize initial wealth to one when we study a single-period model. First, when the reference point is lower than the gross return of the riskless asset ("the gross riskless rate"), the optimal terminal wealth is greater than the reference point in all states. In fact, within the risk-averse region, the expected utility has a unique local maximum, which is also the global maximum. More importantly, we show that a loss averse investor with such a reference point behaves identically to a standard HARA (risk averse) investor, since the two utility functions are the same over the gain domain. In this case, she always longs/shorts a risky asset with a positive/negative risk premium, and the expected utility is always positive. Furthermore, the optimal portfolio problem under loss aversion can be transformed into a problem with a reference point of zero by redefining the initial wealth. Then the above results read that the optimal portfolio under loss aversion is identical to that in a standard portfolio problem under the CRRA utility with positive (redefined) initial wealth. These results hold true, regardless of the completeness of the markets.

Reference points lower than the gross riskless rate are commonly studied in the literature. For example, Tversky and Kahneman (1992) suggest that one candidate for the reference point is the initial wealth level. In this case, the ratio of the reference point to initial wealth is one, which is lower than the gross riskless rate. This reference point is used in, e.g., Benartzi and Thaler (1995) who study static portfolio choice under loss aversion and use loss aversion to explain the equity premium puzzle. Our results suggest that the investors studied in Benartzi and Thaler (1995) are similar to a standard HARA investor.

Second, we study the case when the reference point equals the gross riskless rate, which is another popular value of reference point used in the asset pricing literature. Unlike this literature where the loss aversion utility is usually studied under certain portfolio/utility constraints, we focus on the implications of the loss aversion utility without any constraints or changes to the functional form. We show that the investor takes infinite positions in the risky asset, if the asset return is highly skewed or if the investor has low loss aversion. Otherwise, the investor does not invest in the risky asset, since under loss aversion preferences, first-order risk aversion (Segal and Spivak, 1990) applies at the reference point, causing the investor to be reluctant to take on small risks when her wealth is at the reference point. In the above two cases when the reference point is lower than or equal to the gross riskless rate, the optimal portfolio (if exists) is not affected by the coefficients of loss aversion and curvatural sensitivity.

Finally, we look at the case when the reference point is higher than the gross riskless rate. This case is where loss aversion truly has effects and somehow less studied in the literature. We find that the optimal portfolios in this case starkly differ from those under the risk aversion utility functions, and they exhibit significantly different properties under complete markets and under incomplete markets. When the markets are complete, the optimal wealth is lower than the reference point in the worst state, against which it is most expensive to protect, but higher in all the other states. Intuitively, it is not optimal to set wealth below the reference point in more than two states because the investor would exhibit risk seeking in the loss domain. As a result, the investor's wealth tends to be negatively skewed when she leaves all utility losses in a single state.[2] The "pain" in this single state dominates the "pleasure" in all other states, and as a consequence the value function is always negative. However, when the markets are incomplete, the investor divides all states into two groups, one with positive excess returns and the other with negative excess returns, and she assigns positive wealth in all states in one group and negative wealth in all states in the other group.

We further show that in this case when the reference point is higher than the gross riskless rate, the investor is forced to take large risky positions, either leverage or shorting the risky asset. As the reference point increases, the investor tends to assign more extremely positive and negative wealth levels across states, resulting in more dispersed distribution of optimal wealth. Moreover, the sign of the optimal portfolio weight is not determined by the risk premium but by the relative levels of pricing kernel across states, which determine the tradeoff between losses in one state and gains in another. As a result, the investor may hold assets with a zero risk premium and can long (short) an asset with an arbitrarily low

[2] The negative skewness of wealth may be helpful for understanding insurance policy. Azevedo and Gottlieb (2012) show that a risk-neutral firm can extract arbitrarily high expected values from consumers with prospect theory preferences since such consumers accept gambles with arbitrarily large negative expected values. As a result, there is no solution to the firm's problem. If the consumers are subject to wealth constraints, the firm can extract all the consumers' wealth with probability approaching one.

(high) risk premium. These results are in sharp contrast with those under the risk aversion (concave) utility functions.

When the reference point is higher than the gross riskless rate, the expected utility function is nonconcave. This leads to multiple local maximums for the expected utility function, each for one partition of states, and the greatest of them are the global maximums. This further indicates a discontinuity in the parameter space. A small parameter change can cause a big jump in the optimal portfolio and the value function. This poses great challenges to the numerical solution methods as usually used in this literature. In addition, there can be multiple global maximums at the same time; thus, the investor can achieve the same highest expected utility by either longing or shorting the risky asset.

When the reference point is lower than or equal to the gross riskless rate, the optimal portfolio exhibits very similar properties under complete markets and incomplete markets. Although incomplete markets reduce potential investment opportunities, the investor can assign her wealth at least as the gross riskless rate, which is higher than the reference point, by holding the riskless asset. As a result, the optimal wealth will not drop in the loss domain in any state, regardless of the completeness of the markets. However, the optimal portfolio can be very different under the two types of markets if the reference point is higher than the gross riskless rate. Incomplete markets allow less opportunities, and as a consequence the investor may not be able to allocate her wealth to be lower than the reference point only in a single (the most expensive) state. In the loss aversion literature, both complete markets (e.g., Berkelaar, Kouwenberg and Post, 2004; Barberis and Xiong, 2009) and incomplete markets (e.g., Benartzi and Thaler, 1995; He and Zhou, 2011) have been widely studied.

Diminishing sensitivity, loss aversion, and reference dependence are three major psychological concepts captured by the loss aversion utility. Diminishing sensitivity in the loss domain (or the risk seeking behavior) actually leads to lower penalty for large losses than that imposed by the risk aversion (concave) utility functions. As a result, relative to the corresponding risk averse investor, a loss averse investor tends to hold more lottery-like assets with positively skewed returns. This result complements the literature that finds that due to the probability weighting that overweighs the tails of outcome distribution, prospect theory investors desire positive skewness (e.g., Barberis and Huang, 2008). This low loss penalty also leads to other key differences from the risk aversion utility functions regarding portfolio implications, such as unbounded optimal portfolios, large risky positions, and insensitivity of portfolio to risk premium.

However, loss aversion does not drive the above key differences from the risk aversion utility functions. In fact, most (if not all) risk aversion utility functions, or even risk neutral utility, imply greater utility losses within the loss domain than the loss aversion utility function. The loss aversion coefficient has even no impact on the optimal portfolio and wealth when the reference point is lower than or equal to the gross riskless rate.[3] Regarding reference dependence, we have shown that the relative level of the reference point and the gross riskless rate qualitatively determines the properties of the optimal portfolios (if exist). While loss aversion affects large-scale properties of the optimal portfolio, such as boundedness of the optimal portfolio, reference dependence determines local properties.

The above results are for the case without constraints (apart from the budget constraint). When imposing a positive wealth constraint, which is sometimes assumed in the literature, the optimal solutions always exist; however, they are often given by corner solutions (extreme solutions). In this case, the constraint is the key determinant of the optimal portfolio. We find that the optimal wealth can be lower than the reference point in more than one states, distinctly different from that in the unconstraint case. Furthermore, wealth constraint can still have effects in the case when the unconstrained choice has bounded optimal solutions. Imposing constraints can even qualitatively change the optimal strategy. Notably, imposing wealth constraint redefines the utility function: it sets the utility to be minus infinity for wealth level beyond the constraint.

The reference point plays an important role in the loss aversion utility. However, the choice of the reference point is a key challenge in the application of prospect theory as pointed out by Barberis (2013), and the results on it are mixed in the literature.[4] As noted by Baillon et al.

[3]The loss aversion coefficient can impact the optimal portfolios only when the reference point is higher than the gross riskless rate. In this case, high loss aversion lowers both the optimal holdings of stocks and the deviation of the optimal wealth from the reference point, due to the aversion to big losses.

[4]Kahneman and Tversky (1979) suggest that the reference point is current wealth or expectation levels. Tversky and Kahneman (1991) argue that the reference point can be affected by aspirations, expectations, norms, and social comparisons. Köszegi and Rabin (2006) postulate that the reference point is the rational expectations about future outcomes. In the context of portfolio choice, one reference point is the price at which the stock was initially purchased, e.g., the dispersion effect (Shefrin and Statman, 1985). Gneezy (2005) find that participants are most likely to use the historical stock price peak as the reference point.

(2020), "*This lack of clarity [of the reference point] is undesirable, because it creates extra freedom in deriving predictions, making it impossible to rigorously test reference-dependent theories empirically.*" Our analytical results provide a comprehensive understanding of its effects in the context of portfolio choice. We demonstrate that the coefficients of loss aversion and risk seeking can impact the properties of the optimal portfolios (if exist) only if the reference point is higher than the gross riskless rate, providing theoretical guidance for choosing and testing the reference point.

To focus on loss aversion, our paper studies a static model. Our analysis is general: given a reference point, we analytically solve for the optimal portfolio. In a dynamic setting, the reference point can be time-varying, and its effects can be more complex. The literature studies both exogenous and endogenous reference points and usually under portfolio constraints. Berkelaar, Kouwenberg and Post (2004) consider an exogenous (time-varying) reference point and solve for the dynamic optimal portfolios under positive wealth constraint.[5] Barberis and Xiong (2009, 2012) and Ingersoll and Jin (2013) study realization utility where an investor receives a utility burst from realized gains/losss by setting the purchase price of risky assets as the reference point. They show that realization utility explains the disposition effect. Dai, Qin and Wang (2024) add to the realization utility literature by relaxing the key and common assumption of binary stock holdings. Meng and Weng (2018) further show that the key driving force behind the explanation of the disposition effect using prospect theory is the level of the reference point, regardless of whether the reference point is updated or not, and that any reference point that is sufficiently higher than initial wealth can lead to the disposition effect. Another natural reference point is consumption. For example, Dybvig and Rogers (2013) study preferences with endogenous references that generalize the recursive utility and are axiomatically motivated. Choi, Jeon and Koo (2022) study consumption irreversibility where the previous consumption level is the reference point.

Baucells, Weber and Welfens (2011) show that the reference point of subjects is most heavily influenced by the purchase and the current stock prices. Baillon, Bleichrodt and Spinu (2020) find that the most common reference points are the status quo and a security-based level representing the maximum outcome that one can reach for sure.

[5]An exogenous reference point is similar to a constant reference point in a static setting. Furthermore, the loss averse investor with an exogenous reference point will behave like a risk averse investor when the investment horizon is long, since the utility at low wealth levels is not important when the investor adheres to the expanding market, as predicted by portfolio turnpike theorems (e.g., Dybvig, Rogers and Back, 1999).

The loss aversion utility function is not concave. This leads the maximization problem to be rather delicate. To ensure that the optimal portfolios are bounded under loss aversion, the literature usually either imposes certain portfolio constraints, such as short-sales constraints (Aït-Sahalia and Brandt, 2001), a nonnegativity of wealth constraint (Berkelaar et al., 2004; Barberis and Xiong, 2009), binary stock holdings (Li and Yang, 2013), and maximum tolerable loss (similar to a wealth constraint) (Ingersoll, 2016), or introduces different restrictions or variations to the S-shaped utility function. For example, Gomes (2005) imposes concavity when wealth is low. Barberis et al. (2001) consider that investors' utility function consists of a standard risk aversion utility component, in addition to the loss aversion utility function. Ingersoll (2016) imposes extreme-risk avoidance. However, few papers explicitly examine the conditions for the existence of optimal solutions. These conditions are important since they help understand under what market conditions/restrictions loss aversion is applicable to studying the behavior of investors and asset prices. Our paper adds to this literature by analytically deriving the solution existence conditions without either constraints or modifications of utility function.

He and Zhou (2011) is amongst the few studies that analytically examine the existence of optimal solutions under prospect theory. They demonstrate that a prospect theory model with general functional forms of utility and probability weighting, which include that proposed in Tversky and Kahneman (1992) and following variations, is "easily ill-posed" (in the sense of no bounded optimal solution). Our paper has two major differences from He and Zhou (2011). While He and Zhou (2011) study a single risky asset, our paper generally studies multiple assets. In discrete time with more than two states, complete markets require more than two assets. In addition to the existence of solutions, He and Zhou (2011) explicitly solve for the optimal portfolio for the case with the reference point only being the gross riskless rate. Our results on the optimal portfolio apply to any reference point.

In addition to loss aversion, the literature also exploits other deviations from expected utility theory, such as disappointment aversion (Gul, 1991). Disappointment aversion generally implies aversion to losses and has an endogenous reference point. While prospect theory has no axiomatic basis, disappointment aversion is grounded in axioms. Different from loss aversion, with which the risk aversion is first-order only at the reference point, the risk aversion is first-order at every level with disappointment aversion preferences. As a result, a disappointment averse investor holds no equity in a non-participation region, instead of a single point, as shown in Ang, Bekaert and Liu (2005).

The paper is organized as follows. Section 2 discusses the loss aversion preferences and the optimization problem. Sections 3 studies portfolio choice under loss aversion in complete markets with two states, and Section 4 studies extensions, e.g., with multiple states and in incomplete markets. Section 5 concludes. Calculation details are included in the appendices.

2 The Setup

2.1 The Portfolio Problem

We study a standard static optimal portfolio selection problem for an investor who has loss aversion preferences that will be specified shortly in Section 2.2. The investment opportunity set of the investor with an initial wealth w_0 contains $N+1$ assets, including N risky assets and 1 riskless asset. Denote by $\mathbf{R}$ and r the gross returns of the risky assets and the riskless asset, respectively. The optimization problem for the investor is given by

$$\max_{\boldsymbol{x}} \mathbb{E}\big[U(w)\big], \tag{1}$$

subject to the budget constraint $w = w_0 r + \boldsymbol{x}'(\mathbf{R} - r\mathbf{1})$, where $\boldsymbol{x}$ is a vector of the values of the investor's holdings of the risky assets, $U(\cdot)$ is the loss aversion utility function that will be specified shortly in (2), and w is the investor's terminal wealth at the end of the period. This paper focuses on the optimization problem (1) without wealth constraints. We also study the case with wealth constraints in Section 3.4. The loss aversion utility generally allows any wealth levels, and accordingly, our results developed in this paper also hold for any initial wealth w_0.

2.2 Loss Aversion

The utility function in cumulative prospect theory of Tversky and Kahneman (1992) is defined over gains and losses relative to a reference point θ:

$$U(w) = \begin{cases} \frac{1}{1-\gamma_+}(w-\theta)^{1-\gamma_+} & \text{if } w \geq \theta; \\ -A\frac{1}{1-\gamma_-}(\theta-w)^{1-\gamma_-} & \text{if } w < \theta, \end{cases} \tag{2}$$

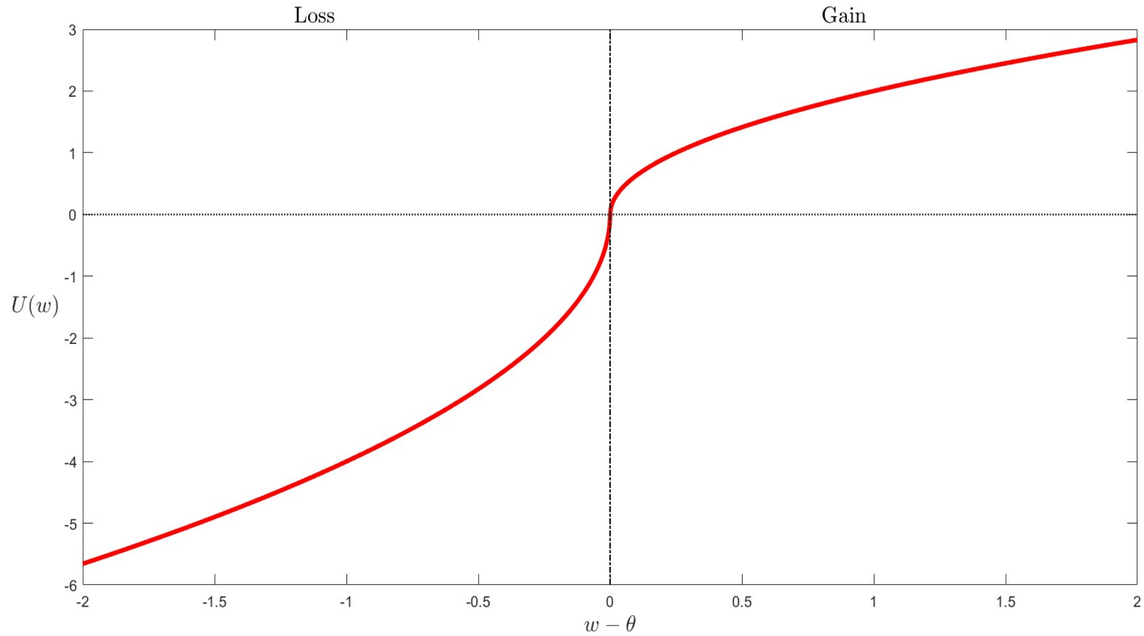

Figure 1: This figure illustrates the S-shaped utility function. Here, $A = 2$ and $\gamma_\pm = 0.5$.

where w is the investor's wealth, θ is the reference point, $\gamma_\pm \in [0, 1)$ controls the curvature,[6] and $A > 0$ determines the degree of loss aversion. With a single period, as in Tversky and Kahneman (1992), the reference point θ is a constant. The utility function $U(w)$ is increasing and is concave for $w \geq \theta$ and convex for $w \leq \theta$, as illustrated in Figure 1.

Under loss aversion (2), first-order risk aversion (Segal and Spivak, 1990) (for large A) or first-order risk seeking (for small A) applies at the reference point θ. For the other points, second-order risk aversion ($w > \theta$) or risk seeking ($w < \theta$) applies. As a result, an investor with loss aversion is reluctant to take on small risks when her wealth is close to the reference point θ, but always accepts a gamble with a positive mean when her wealth is above θ. This is different from Knightian uncertainty and disappointment aversion (Gul, 1991), with which the risk aversion is first-order at every level.

When $\gamma_+ = \gamma_-$, which is estimated by Tversky and Kahneman (1992) and widely assumed in the portfolio choice and asset pricing literature, the utility (2) becomes

$$U(w) = \begin{cases} \frac{1}{1-\gamma}(w-\theta)^{1-\gamma} & \text{if } w \geq \theta; \\ -A\frac{1}{1-\gamma}(\theta - w)^{1-\gamma} & \text{if } w < \theta. \end{cases} \tag{3}$$

[6]If $\gamma_\pm \geq 1$, the utility approaches minus infinity when w approaches θ from above, and hence the utility function is not increasing.

In this paper, we focus mainly on this case. We also study $\gamma_+ \neq \gamma_-$ in Section 3.5.

Cumulative prospect theory of Tversky and Kahneman (1992) incorporates two complementary components: an S-shaped loss aversion utility function as discussed above and a probability weighting function. In this paper, we focus mainly on the effects of the S-shaped utility function by assuming the investor's subjective probability measure is the same as the physical measure. We will discuss the impacts of probability weighting in Appendix B.

2.2.1 Comparison with HARA Utility

Merton (1971) studies the optimal portfolio and consumption choice problem under utility functions in the hyperbolic absolute risk aversion (HARA) family. The HARA family is given by $U(w) = \frac{\gamma}{1-\gamma}(\frac{\beta w}{\gamma} + \eta)^{1-\gamma}$. When $0 < \gamma < 1$ as in the loss aversion utility function, the HARA utility can be written as

$$U(w) = \begin{cases} \frac{A}{1-\gamma}(w-\theta)^{1-\gamma}, & \text{for } w \geq \theta; \\ -\infty, & \text{for } w < \theta, \end{cases} \tag{4}$$

where $A = \beta^{1-\gamma}\gamma^{\gamma} > 0$ and $\theta = -\frac{\gamma\eta}{\beta}$, to ensure uniform concavity.

Equation (4) shows that with a "reference point" θ, the HARA (risk aversion) utility has a positive and finite utility gain in the "gain domain" (i.e., $w > \theta$) and a utility loss of minus infinity in the "loss domain" (i.e., $w < \theta$). Therefore, a HARA investor is more sensitive to losses than to equivalent gains, suggesting that like the loss aversion utility, the HARA utility (4) also features loss aversion.

Furthermore, by letting $\beta = \gamma^{\frac{\gamma}{\gamma-1}}$, the HARA utility (4) is identical to the loss aversion utility (3) for $w \geq \theta$. However, the two utility functions are different when wealth lower than the reference point $w < \theta$. When $w < \theta$, the loss aversion utility function (3) leads to a low but finite utility, and the HARA utility function (4) leads to a minus infinity utility. As a result, a loss averse investor is less averse to losses and hence tends to trade more aggressively than the corresponding HARA investor.[7]

[7]The HARA utility function becomes the exponential utility function in the limiting case $\gamma = -\infty$ and $\eta = 1$. Due to diminishing sensitivity, the loss aversion utility function also has lighter penalty for large losses than the exponential utility function. However, the exponential utility has no reference point and may not be directly comparable with the loss aversion utility.

The identical utility of the two utility functions in the gain domain further leads to the following results on the optimal portfolio.

Proposition 1. *Assume that the markets are complete and that the portfolio problem under loss aversion has bounded solutions (conditions given by Proposition 2).*

When the reference point is lower than or equal to initial wealth multiplied by the gross riskless rate $\theta \leq w_0 r$, the optimization problem (1) *under the loss aversion utility* (3) *is identical to that under the HARA utility* (4).

Proposition 1 shows that when the reference point is sufficiently low, the optimal portfolio under loss aversion is identical to that under the standard HARA utility as studied in Merton (1971). In fact, the portfolio problem under the HARA utility (4) with initial wealth w_0 can be transformed into a standard problem under a CRRA utility with initial wealth $\hat{w}_0 = w_0 - r^{-1}\theta$. With positive initial wealth for the latter problem (i.e., $\hat{w}_0 > 0$ or $\theta < w_0 r$), the optimal wealth is always in the "gain domain" ($\hat{w}^* > 0$); here we consider zero as a reference point for the CRRA utility. This means that the optimal wealth for the original HARA utility problem is always higher than the reference point $w^* > \theta$. In this case, the loss aversion utility is identical to the HARA utility and hence leads to the same optimal portfolio.

Low reference points are commonly studied in the literature. For example, one popular reference point is the initial wealth level, which is suggested in Tversky and Kahneman (1992) and used in, e.g., Benartzi and Thaler (1995). In this case, the ratio of the reference point to initial wealth is one, which is lower than the gross riskless rate ($\theta \leq w_0 r$).

The results in Proposition 1 may remain valid under incomplete markets. For example, we show in Section 4.1 that these results can hold true in a simple case of incomplete markets with three states.

3 Binomial Model

To provide clear economic insights, in this section we study a binomial model under complete markets. There are two assets: a riskless asset with a constant gross return r and a risky asset. The gross return R of the risky asset is either u, with probability p, or d ($< u$), with probability $1 - p$. We choose $d < r < u$ to guarantee no arbitrage.

In the remainder of this section, we first solve for the optimal wealth and then resolve the optimal trading strategy by using the martingale approach (Cox and Huang, 1989). With two assets and two states, the markets are complete. There exist unique strictly positive state prices. The state prices for state u and d are given, respectively, by $p\xi_u$ and $(1-p)\xi_d$, where the pricing kernel ξ satisfies $\xi_u = \frac{r-d}{p(u-d)r}$ and $\xi_d = \frac{u-r}{(1-p)(u-d)r}$. As a result, the optimization problem (1) can be rewritten in terms of state prices:

$$\max_{w_u, w_d} pU(w_u) + (1-p)U(w_d), \tag{5}$$

subject to the budget constraint

$$p\xi_u w_u + (1-p)\xi_d w_d = w_0, \tag{6}$$

where w_u and w_d are the investor's portfolio wealth in states u and d, respectively.

3.1 Existence of Bounded Optimal Solutions

Using the budget constraint (6), we express w_d in terms of w_u: $w_d = \frac{w_0 - p\xi_u w_u}{(1-p)\xi_d}$. This allows us to rewrite the expected utility function EU in terms of w_u:

$$EU(w_u) = pU(w_u) + (1-p)U\Big(\frac{w_0 - p\xi_u w_u}{(1-p)\xi_d}\Big). \tag{7}$$

The following lemma states the asymptotic behavior of the expected utility (7).

Lemma 1. *(Asymptotic behavior.) When $w_u \to +\infty$, the EU satisfies*

$$EU \to \begin{cases} +\infty, & \text{if} \quad A < \big(\frac{p}{1-p}\big)^{\gamma}\big(\frac{\xi_d}{\xi_u}\big)^{1-\gamma}; \\ 0, & \text{if} \quad A = \big(\frac{p}{1-p}\big)^{\gamma}\big(\frac{\xi_d}{\xi_u}\big)^{1-\gamma}; \\ -\infty, & \text{if} \quad A > \big(\frac{p}{1-p}\big)^{\gamma}\big(\frac{\xi_d}{\xi_u}\big)^{1-\gamma}. \end{cases} \tag{8}$$

When $w_u \to -\infty$, the EU satisfies

$$EU \to \begin{cases} +\infty, & \text{if} \quad A < \big(\frac{1-p}{p}\big)^{\gamma}\big(\frac{\xi_u}{\xi_d}\big)^{1-\gamma}; \\ 0, & \text{if} \quad A = \big(\frac{1-p}{p}\big)^{\gamma}\big(\frac{\xi_u}{\xi_d}\big)^{1-\gamma}; \\ -\infty, & \text{if} \quad A > \big(\frac{1-p}{p}\big)^{\gamma}\big(\frac{\xi_u}{\xi_d}\big)^{1-\gamma}. \end{cases} \tag{9}$$

Both (8) and (9) show that the expected utility can be unbounded from above if loss aversion A is sufficiently low. In this case, the low loss aversion leads to insufficient penalty for extremely negative wealth. As a result, the investor can achieve arbitrarily large expected utility by assigning her terminal wealth to be positive infinity in one state and negative infinity in the other under the budget constraint. This is different from standard expected utility functions that are uniformly concave. The expected utility functions impose sufficiently heavy penalty for strongly low level of wealth, while the utility gains in good states are relatively smaller, preventing the EU from approaching positive infinity.

To have a well-posed problem (5), the EU should be bounded from above. Lemma 1 leads to the following conditions of the existence of optimal solutions.

Proposition 2. *(Existence of solutions.) Define*

$$\underline{A} = \max\left\{\left(\frac{p}{1-p}\right)^{\gamma}\left(\frac{\xi_d}{\xi_u}\right)^{1-\gamma}, \quad \left(\frac{1-p}{p}\right)^{\gamma}\left(\frac{\xi_u}{\xi_d}\right)^{1-\gamma}\right\}. \tag{10}$$

1. *When $A > \underline{A}$, the optimization problem* (5) *has bounded solutions.*

2. *When $A < \underline{A}$, the optimization problem* (5) *does not have bounded solutions.*

3. *When $A = \underline{A}$, the optimization problem* (5) *has bounded solutions for $\theta \le w_0 r$ and does not have bounded solutions for $\theta > w_0 r$.*

Proposition 2 shows that there is a lower bound $\underline{A}$ for the loss aversion parameter A, below which the optimal portfolio choice problem (5) has no solution. This is essentially because the penalty of loss is too light. Consider a budget preserving change

$$\Delta w = \left(\frac{\Delta}{p\xi_u}, \ \frac{-\Delta}{(1-p)\xi_d}\right). \tag{11}$$

It says that if the investor wants to increase wealth in, e.g., state u by $\frac{\Delta}{p\xi_u}$, she has to give up $\frac{-\Delta}{(1-p)\xi_d}$ in state d. When $A < \underline{A}$, the utility gain in one state (state u) is always higher than the utility loss in the other state (state d). As a result, the investor can always increase her expected utility without bound by assigning more wealth in the first state and at the same time less wealth in the second state, leading to infinite expected utility.

Proposition 2 further shows that there always exist optimal solutions for $A > \underline{A}$. With a high level of loss aversion A, the utility gain when wealth exceeds the reference point is lower than the utility loss when wealth is lower than the reference point by the same amount.

This renders extreme assignment of wealth suboptimal. The above results therefore suggest that the driver of the nonexistence of optimum is insufficient loss penalty rather than risk seeking preferences. With sufficient penalty on losses (large loss aversion A), there always exist global optimum even under risk seeking.

Most papers generally require that loss aversion A is great than 1 (investors are more sensitive to losses than to equivalent gains), which is consistent with the estimate in Tversky and Kahneman (1992). Condition (10) shows that this condition appears to be a necessary condition for existence of solutions. It is not sufficient to guarantee interior solutions, depending on the market parameters p, ξ_u, and ξ_d, as well as utility parameter γ. Figure 2 plots the lower bound $\underline{A}$ against p (the left panel) and against pricing kernel ratio ξ_u/ξ_d (the right panel). It shows that $\underline{A}$ is low when asset return distribution is symmetric. When the probability is close to zero (positively skewed return) or 1 (negatively skewed return),[8] or the pricing kernel ratio is close to zero, the lower bound $\underline{A}$ approaches infinity, and it is unlikely to have bounded optimal solutions. Intuitively, an asset with highly skewed returns (either left or right) allows the investor to achieve a high utility gain at certain states (by either shorting or longing the asset); thus, a high loss aversion is required to ensure finite expected utility.

3.2 Optimal Wealth

We have derived the conditions under which the optimal solutions exist. Now we study the optimal portfolio under these conditions. We first study the case $A > \underline{A}$ and then describe the extreme case $A = \underline{A}$ at the end of Section 3.3. The following proposition describes the results of the optimal wealth.

Proposition 3. *(Optimal wealth and value function.) Suppose $A > \underline{A}$.*

1. *When $\theta < w_0 r$, the optimal wealth in the two states is given by*

$$\begin{aligned} w_u^* - \theta &= \left(w_0 - r^{-1}\theta\right)\left[p\xi_u^{1-\frac{1}{\gamma}} + (1-p)\xi_d^{1-\frac{1}{\gamma}}\right]^{-1}\xi_u^{-\frac{1}{\gamma}} > 0, \\ w_d^* - \theta &= \left(w_0 - r^{-1}\theta\right)\left[p\xi_u^{1-\frac{1}{\gamma}} + (1-p)\xi_d^{1-\frac{1}{\gamma}}\right]^{-1}\xi_d^{-\frac{1}{\gamma}} > 0, \end{aligned} \tag{12}$$

   and the value function is given by $J = \frac{(w_0 - r^{-1}\theta)^{1-\gamma}}{1-\gamma}[p\xi_u^{1-\frac{1}{\gamma}} + (1-p)\xi_d^{1-\frac{1}{\gamma}}]^{\gamma}$.

[8] Return skewness is $p(1-p)(1-2p)(\frac{u-d}{\sigma})^3$, which is positive for $p < 0.5$ and negative for $p > 0.5$.

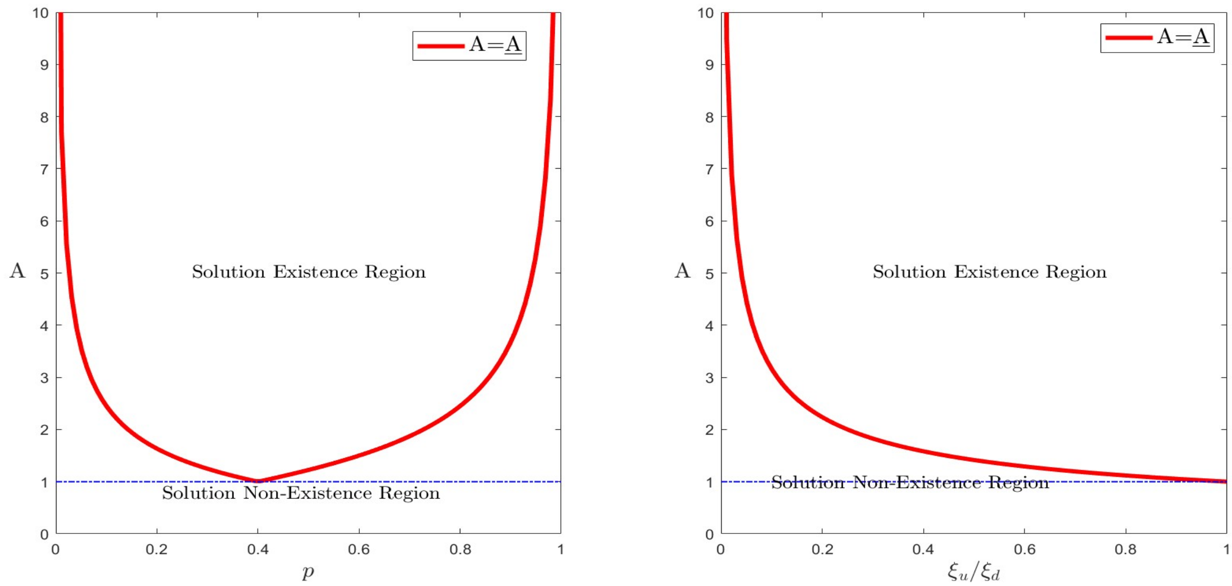

Figure 2: The left panel plots the lower bound $\underline{A}$ against probability p, and the right panel plots $\underline{A}$ against the pricing kernel ratio ξ_u/ξ_d (we assume $\xi_u \le \xi_d$ without loss of generality). Here, $\gamma = 0.5$, $\xi_u = 1$ and $\xi_d = 1.5$ for the left panel and $\gamma = 0.5$ and $p = 0.5$ for the right panel.

2. *When $\theta = w_0 r$, the optimal wealth in the two states is given by*

$$w_u^* = w_d^* = \theta, \tag{13}$$

and the value function is given by $J = 0$.

3. *When $\theta > w_0 r$, the EU has two local maximums, and the global maximum is the greater of them. Specifically, one local maximum occurs for $w_d < \theta < w_u$, at which the portfolio wealth is given by*

$$\begin{aligned} w_u^* - \theta &= \left(r^{-1}\theta - w_0\right)\left[-p\xi_u^{1-\frac{1}{\gamma}} + (1-p)A^{\frac{1}{\gamma}}\xi_d^{1-\frac{1}{\gamma}}\right]^{-1}\xi_u^{-\frac{1}{\gamma}} > 0, \\ \theta - w_d^* &= \left(r^{-1}\theta - w_0\right)\left[-p\xi_u^{1-\frac{1}{\gamma}} + (1-p)A^{\frac{1}{\gamma}}\xi_d^{1-\frac{1}{\gamma}}\right]^{-1}A^{\frac{1}{\gamma}}\xi_d^{-\frac{1}{\gamma}} > 0, \end{aligned} \tag{14}$$

and the EU is given by $EU_+ = -\frac{(r^{-1}\theta - w_0)^{1-\gamma}}{1-\gamma}[-p\xi_u^{1-\frac{1}{\gamma}} + (1-p)A^{\frac{1}{\gamma}}\xi_d^{1-\frac{1}{\gamma}}]^{\gamma}$; another local maximum occurs for $w_u < \theta < w_d$, at which the wealth is given by

$$\begin{aligned} \theta - w_u^{**} &= \left(r^{-1}\theta - w_0\right)\left[pA^{\frac{1}{\gamma}}\xi_u^{1-\frac{1}{\gamma}} - (1-p)\xi_d^{1-\frac{1}{\gamma}}\right]^{-1}A^{\frac{1}{\gamma}}\xi_u^{-\frac{1}{\gamma}} > 0, \\ w_d^{**} - \theta &= \left(r^{-1}\theta - w_0\right)\left[pA^{\frac{1}{\gamma}}\xi_u^{1-\frac{1}{\gamma}} - (1-p)\xi_d^{1-\frac{1}{\gamma}}\right]^{-1}\xi_d^{-\frac{1}{\gamma}} > 0, \end{aligned} \tag{15}$$

and the EU is given by $EU_- = -\frac{(r^{-1}\theta - w_0)^{1-\gamma}}{1-\gamma}[pA^{\frac{1}{\gamma}}\xi_u^{1-\frac{1}{\gamma}} - (1-p)\xi_d^{1-\frac{1}{\gamma}}]^{\gamma}$.

Although the reference point θ does not affect the existence of solutions in Proposition 2, it significantly affects the properties of optimal portfolio (if exists). Proposition 3 shows that the properties of the optimal solution depend crucially on the relative level of the initial wealth w_0 to the reference point discounted by the gross riskless rate $r^{-1}\theta$. The optimal portfolio problem under loss aversion can be transformed into a problem with a reference point of zero by redefining the initial wealth as $\hat{w}_0 = w_0 - r^{-1}\theta$ (Appendix A.1). Then Proposition 3 reads that the optimal portfolio for this transformed problem exhibits starkly different properties when the sign of the initial wealth $\hat{w}_0$ is different.

We first consider the case $\theta < w_0 r$. Because the gross riskless rate is determined by the state prices: $r^{-1} = \mathbb{E}[\xi]$, in this case we have $\mathbb{E}[\xi]\theta < w_0$. Together with the budget constraint $\mathbb{E}[\xi w] = w_0$, we obtain

$$\mathbb{E}[\xi(w-\theta)] > 0,$$

which shows that the terminal wealth w must be greater than the reference point θ in at least one of the two states u and d. Proposition 3 confirms this and further shows that it is not optimal to have the terminal wealth higher than the reference point in one state but lower in the other state, and the optimal wealth is greater than the reference point in both states. In fact, in this gain domain ($w > \theta$), the EU has a local maximum, which is also the global maximum. Proposition 3 also shows the investor allocates more wealth in the better state with low pricing kernel and that the expected utility is always positive.

When $\theta = w_0 r$, it is optimal to assign wealth in both states to be at the reference point, $w^*_{u,d} = \theta$. In this case, the expected utility is zero. Intuitively, consider a budget preserving change (11) starting from $w_u = w_d = \theta$. If wealth in one state is greater than the reference point, the wealth in the other state must be lower than the reference point. Due to the strong loss penalty (for $A > \underline{A}$), the utility gain in the former state is lower than the utility loss in the latter, leading to negative expected utility.

When $\theta > w_0 r$, due to the budget constraint, the investor can either assign her terminal wealth to be lower than the reference point at one state and at the same time greater at the other state, or assign her terminal wealth to be lower than the reference point at both states. Proposition 3 shows that the optimal solution occurs only in the former case: either $w^*_d < \theta < w^*_u$ or $w^*_u < \theta < w^*_d$. In fact, if the investor's terminal wealth is lower than the reference point in both states, she is risk seeking. She can always increase EU by choosing a more extreme wealth allocation across states, eventually pushing wealth above the reference point in one of the two states. Proposition 3 also shows that in this case the expected utility is always negative; thus, the value function is negative.

Proposition 3 shows that loss aversion does not impact the optimal wealth when $\theta < w_0 r$ or $\theta = w_0 r$. It takes effects only when $\theta > w_0 r$. In the case of $\theta > w_0 r$, the deviation of wealth from the reference point $|w^* - \theta|$ decreases with the loss aversion A. Especially, when $A \to \infty$, the optimal wealth in one state converges to the reference point ($w_u^* \to \theta$ in (14) or $w_d^{**} \to \theta$ in (15)). In fact, when $\theta > w_0 r$, the investor's initial wealth is relatively low, and she has to allow her wealth to be lower than the reference point in at least one state. When $A \to \infty$, the investor is extremely averse to losses. She wants to have as small as possible losses in this state. In doing so, she sets her wealth in the other state at zero.

The above results show that it cannot be optimal to assign wealth to be lower than the reference point at both states. We summarize the results on the relative levels of optimal terminal wealth to the reference point in the following corollary.

Corollary 1. *Suppose* $A > \underline{A}$.

1. *When* $\theta < w_0 r$, *the optimal wealth is greater than the reference point at both states.*
2. *When* $\theta = w_0 r$, *the optimal wealth is equal to the reference point.*
3. *When* $\theta > w_0 r$, *the optimal wealth is greater than the reference point at one state and lower than the reference point at the other state.*

The S-shaped utility is not concave, leading to nonconcave EU (in terms of the control variable). To understand the effect of the nonconcavity of the utility, Figure 3 plots the EU against w_u. The left panel is for the case $\theta < w_0 r$ and the right panel the case $\theta > w_0 r$. We first look at the left panel of Figure 3. We divide the interval of w_u into three subintervals, $(-\infty, \theta)$, $[\theta, \hat{w}_u]$, and $(\hat{w}_u, +\infty)$, where $\hat{w}_u \equiv \frac{w_0 - (1-p)\xi_d \theta}{p \xi_u}$ corresponds to the case with the wealth in state d being at the referent point ($\hat{w}_d = \theta$). The three intervals correspond to the cases of $w_u < \theta < w_d$, $w_{u,d} \geq \theta$, and $w_d < \theta < w_u$, respectively. The left panel shows that the EU is convex and increases with w_u over $w_u \in (-\infty, \theta]$, is concave and has a local maximum

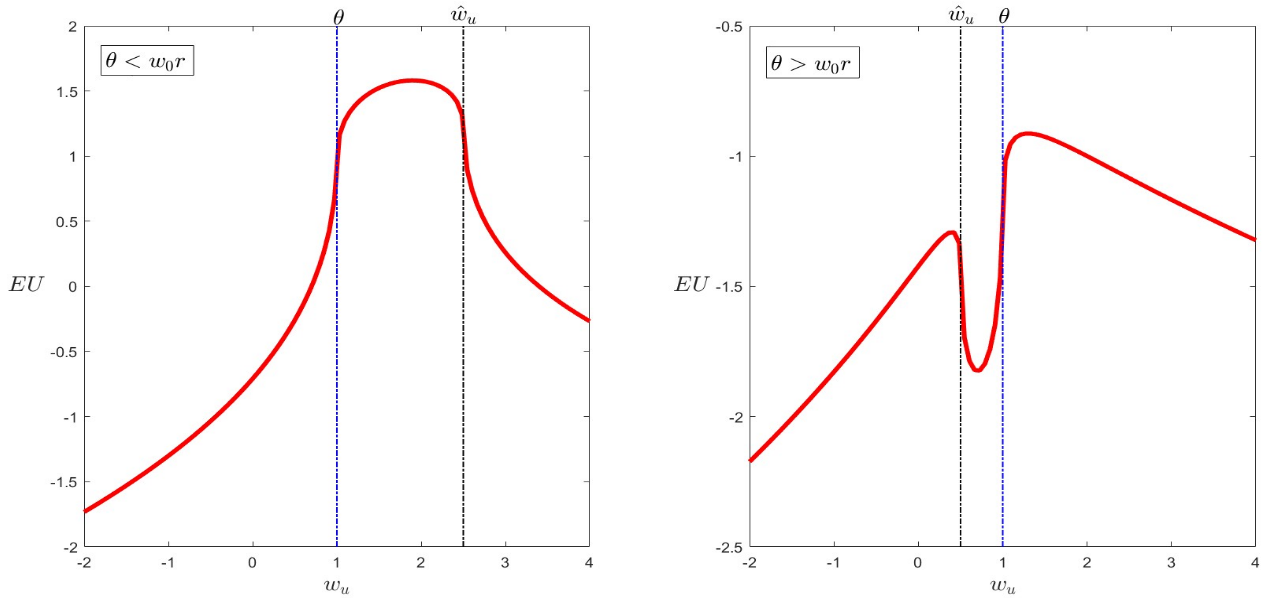

Figure 3: The figure plots the expected utility function EU against wealth w_u for $\theta < w_0 r$ (the left panel) and $\theta > w_0 r$ (the right panel). Here, $A = 2$, $\gamma = 0.5$, $\xi_u = 1$, $\xi_d = 1.5$, $p = 0.5$, and $\hat{w}_u \equiv \frac{w_0 - (1-p)\xi_d \theta}{p \xi_u}$ is the value of w_u such that $\hat{w}_d = \theta$. We set $\theta / w_0 = 0.5$ ($< r$) for the left panel and $\theta / w_0 = 1$ ($> r$) for the right panel.

over $w_u \in [\theta, \hat{w}_u]$,[9] and is convex and decreases with w_u over $w_u \in [\hat{w}_u, \infty]$. The nonconcavity does not take effect when wealth is higher than the reference point in both states, i.e., when $w_u \in [\theta, \hat{w}_u]$. These results together show that the unique local maximum over the interval $w_u \in [\theta, \hat{w}_u]$ is the global maximum.

The right panel illustrates the case $\theta > w_0 r$. It shows that the expected utility function is not concave and has multiple local maximums. We still discuss the results in terms of three subintervals of w_u. For $w_u \in (-\infty, \hat{w}_u)$ or $w_u \in (\theta, +\infty)$, which corresponds to $w_u < \theta < w_d$ or $w_d < \theta < w_u$, respectively, the EU is concave and has a local maximum in each of the two intervals. For $w_u \in [\hat{w}_u, \theta]$, the wealth at both states is lower than the reference point ($w_{u,d} \leq \theta$). In this interval, the EU is convex, and there is no local maximum. Therefore, there are two local maximums, one in $(-\infty, \hat{w}_u)$ and one in $(\theta, +\infty)$, and one of them (the first one under our parameters) is the global maximum.

Proposition 3 shows that when $\theta > w_0 r$, the global maximum of the EU is the greater

[9]In particular, if $\gamma = 0$, the EU is linear not concave in the interval $w_u \in [\theta, \hat{w}_u]$, and the optimum occurs at the boundary.

of the two local maximum. The following corollary further provides the conditions that determine which one is the global maximum.

Corollary 2. *Suppose* $A > \underline{A}$ *and* $\theta > w_0 r$*.*

1. *When* $(1-p)\xi_d^{1-\frac{1}{\gamma}} < p\xi_u^{1-\frac{1}{\gamma}}$*, the global maximum occurs for* $w_d^* < \theta < w_u^*$*.*

2. *When* $(1-p)\xi_d^{1-\frac{1}{\gamma}} > p\xi_u^{1-\frac{1}{\gamma}}$*, the global maximum occurs for* $w_u^{**} < \theta < w_d^{**}$*.*

3. *When* $(1-p)\xi_d^{1-\frac{1}{\gamma}} = p\xi_u^{1-\frac{1}{\gamma}}$*, there are two global maximums, one for* $w_d^* < \theta < w_u^*$ *and one for* $w_u^{**} < \theta < w_d^{**}$*.*

When $(1-p)\xi_d^{1-\frac{1}{\gamma}} < p\xi_u^{1-\frac{1}{\gamma}}$, $EU_- < EU_+$, and thus EU_+ is the global maximum.[10] In this case, Proposition 3 shows that at the global maximum $w_d^* < \theta < w_u^*$. To understand this result, we consider a case with $p = 0.5$ and $\xi_u < \xi_d$. Because a change in w_d is associated with a larger change in w_u under the budget constraint $0.5\xi_u w_u + 0.5\xi_d w_d = w_0$, allocating more wealth in state u relative to state d will yield higher utility. As a result, the optimal wealth is lower than the reference point in the bad state d, against which it is more expensive to protect, but higher in the good state u.

The result is the opposite for $(1-p)\xi_d^{1-\frac{1}{\gamma}} > p\xi_u^{1-\frac{1}{\gamma}}$. In this case, the global maximum occurs for $w_u^* < \theta < w_d^*$, and the value function is given by EU_-. Particularly, when $(1-p)\xi_d^{1-\frac{1}{\gamma}} = p\xi_u^{1-\frac{1}{\gamma}}$, we have $EU_- = EU_+$, and both are global maximums.

The nonconcavity of the S-shaped utility causes the optimal portfolios to be sensitive to parameters and yields jumps in the optimal portfolios. To see this, we examine the impacts of the pricing kernel on the optimal portfolios. Figure 4 illustrates the optimal wealth w_u^* as a function of the pricing kernel ξ_u. When the pricing kernel is too low or too high, there is no optimal solution as shown in Proposition 2. Within the interval in which there exist optimal solutions, there are two thresholds, ξ_u^1 and ξ_u^2, for ξ_u. The first threshold ξ_u^1 is such that $w_0 r = \theta$, and the second ξ_u^2 is such that $(1-p)\xi_d^{1-\frac{1}{\gamma}} = p\xi_u^{1-\frac{1}{\gamma}}$. When $\xi_u < \xi_u^1$, the optimal wealth w_u^* is given by (12) and is positive. When $\xi_u > \xi_u^1$, we have $\theta > w_0 r$. In this case, if $\xi_u^1 < \xi_u < \xi_u^2$, the optimal wealth w_u^* is given by (14) and is also positive; if $\xi_u > \xi_u^2$, the optimal wealth w_u^* is given by (15) and is negative. In particular, at $\xi_u = \xi_u^2$, there are two different optimal wealth w_u^*, one is higher than the reference point and one is lower. When

[10]Note that in this case, we have $\frac{1-p}{[(1-p)\xi_d]^{1-\gamma}} < \frac{p}{(p\xi_u)^{1-\gamma}}$.

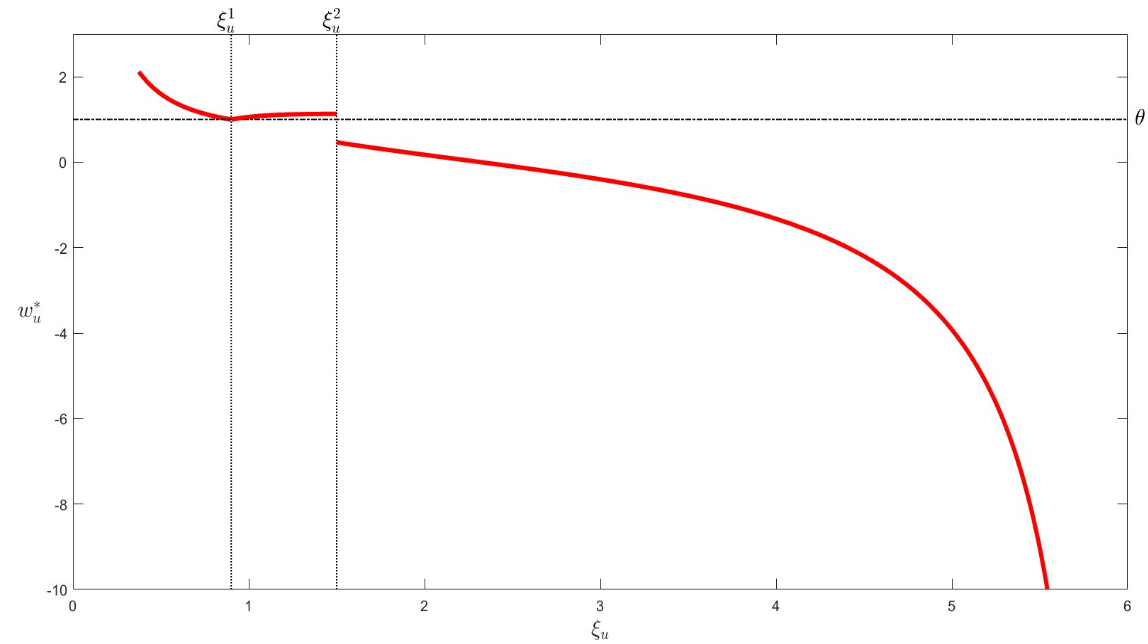

Figure 4: This figure illustrates the optimal wealth w_u^* against pricing kernel ξ_u. Here, $\xi_d = 0.5$, $\gamma = 0.5$, $A = 2$, $w_0 = 1.2$, $\theta = 1$, and $p = 0.5$.

$\xi_u > \xi_u^2$, the global maximum switches from EU_+ to EU_-, associated with a jump in the optimal wealth around $\xi_u = \xi_u^2$, which is a discontinuity in the parameter space. As a result, a small change in a parameter can cause a big jump in the optimal portfolio and expected utility. This poses significant challenges to numerical solution methods as usually used in the literature.

Proposition 3 shows that the sign of the value function varies significantly across the different cases. We summarize the sign of the value function in the following corollary. It shows that the sign is determined by the initial wealth level. When the initial wealth level is high $\theta < w_0 r$, the value function is positive. When initial wealth is low $\theta > w_0 r$, although the optimal wealth can be positive at certain state in Figure 4, the value function is always negative.

Corollary 3. *Suppose $A > \underline{A}$. The value function is positive for $\theta < w_0 r$, negative for $\theta > w_0 r$, and equal to zero for $\theta = w_0 r$.*

3.3 Optimal Portfolio Allocation

Given the optimal wealth, we can further resolve the optimal portfolio allocation.

Proposition 4. *(Optimal portfolio allocation.) Suppose $A > \underline{A}$.*

1. *When $\theta < w_0 r$, the value of the optimal holdings of the risky asset is given by*

$$x^* = \frac{(k-1)(w_0 r - \theta)}{(u-r) + k(r-d)}, \tag{16}$$

where $k = [\frac{p(u-r)}{(1-p)(r-d)}]^{\frac{1}{\gamma}}$.

2. *When $\theta = w_0 r$, the optimal holdings of the risky asset is given by*

$$x^* = 0. \tag{17}$$

3. *When $\theta > w_0 r$, the optimal holdings of the risky asset is given by*

$$x^* = \begin{cases} x_+, & \text{if} \quad \frac{p(u-r)^{1-\gamma}}{(1-p)(r-d)^{1-\gamma}} > 1; \\ x_-, & \text{if} \quad \frac{p(u-r)^{1-\gamma}}{(1-p)(r-d)^{1-\gamma}} < 1; \\ x_+ \text{ and } x_-, & \text{if} \quad \frac{p(u-r)^{1-\gamma}}{(1-p)(r-d)^{1-\gamma}} = 1, \end{cases} \tag{18}$$

where

$$x_+ = \frac{(k_+ + 1)(\theta - w_0 r)}{(u-r) - k_+(r-d)} > 0, \qquad x_- = \frac{(k_- + 1)(\theta - w_0 r)}{(u-r) - k_-(r-d)} < 0,$$

with $k_+ = [\frac{p(u-r)}{A(1-p)(r-d)}]^{\frac{1}{\gamma}}$ and $k_- = [\frac{Ap(u-r)}{(1-p)(r-d)}]^{\frac{1}{\gamma}}$.

When $\theta < w_0 r$, (16) shows that the value of optimal holdings of stock x^* is positive if and only if the risk premium is positive ($k > 1$). This result is typically found with standard expected utility. In this case, the optimal holdings are independent of loss aversion A.

When $\theta = w_0 r$, the investor does not participate into the equity market regardless of the level of the risk premium. In fact, for $A > \underline{A}$, first-order risk aversion is applying at $w = \theta$ but second-order risk aversion at all $w > \theta$. As a result, an investor with loss aversion is reluctant to take on small risks when her wealth is close to θ but always accepts a gamble with a positive mean when her wealth is above θ. Proposition 4 shows that when the reference point equals the investor's initial wealth level scaled up by the gross riskless rate, the investor is reluctant to invest in stocks. This is consistent with the literature that shows that loss aversion helps explain investors' reluctance to participate in the equity markets. Proposition 4 also shows that when the reference point is different from this breakpoint ($\theta \neq w_0 r$), the investor will always participate into the equity market.[11]

When $\theta > w_0 r$, we have shown that the EU has two local maximums. Proposition 4 shows that if $\frac{p(u-r)^{1-\gamma}}{(1-p)(r-d)^{1-\gamma}} > 1$, that is, state u is a better state relative to d,[12] the global maximum occurs for $w_d < \theta < w_u$. In this case, $x_+ > \frac{\theta - w_0 r}{u-r} > 0$: the investor assigns more wealth in state u than in d by longing the risky asset. As a result, the optimal holdings x^* are always positive, even if the risk premium is negative. Indeed, under the S-shaped utility, the sign of the optimal holdings are mainly determined by the relative levels of pricing kernel at states u and d, which determine the tradeoff between losses in one state and gains in the other state. Because proportional scales in the expected returns at both states do not affect the condition $\frac{p(u-r)^{1-\gamma}}{(1-p)(r-d)^{1-\gamma}} > 1$, the investor can even go long an asset with an *arbitrarily* low risk premium. These results are in sharp contrast with that under risk aversion utility functions. When the investment opportunity set contains a risky asset and a riskless asset, a risk averse investor is always long the risky asset if it has a positive risky premium.

In the case of $\theta > w_0 r$, if $\frac{p(u-r)^{1-\gamma}}{(1-p)(r-d)^{1-\gamma}} < 1$ (d is a better state), the global maximum occurs for $w_u < \theta < w_d$. To assign more wealth in state d, the investor must be short the risky asset: $x_- < -\frac{\theta - w_0 r}{r-d} < 0$. As a result, the optimal holdings are always negative, even with a positive risk premium. In the limiting case $\frac{p(u-r)^{1-\gamma}}{(1-p)(r-d)^{1-\gamma}} = 1$, the two local maximums have the same expected utility, and both x_u and x_d are optimal.

The following corollary summarizes the sign of the optimal holdings of stock.

Corollary 4. *1. When $\theta < w_0 r$, $x^* > 0$ if and only if the risk premium is positive.*

2. When $\theta = w_0 r$, the optimal holdings of stock are zero $x^* = 0$.

3. When $\theta > w_0 r$, $x^* > 0$ if and only if $\frac{p(u-r)^{1-\gamma}}{(1-p)(r-d)^{1-\gamma}} > 1$, independent of the sign of the risk premium.

[11] Second-order risk aversion (seeking) applies at all wealth levels greater (lower) than the reference point, and as a result, outside this breakpoint, the investor always participates into the equity markets. Particularly, when the reference point is higher than the gross riskless rate as discussed below, the investor tends to even take a large position, either long or short, in the risky asset.

[12] In this case, $\frac{1-p}{[(1-p)\xi_d]^{1-\gamma}} < \frac{p}{(p\xi_u)^{1-\gamma}}$, which corresponds to Case 1 in Corollary 2.

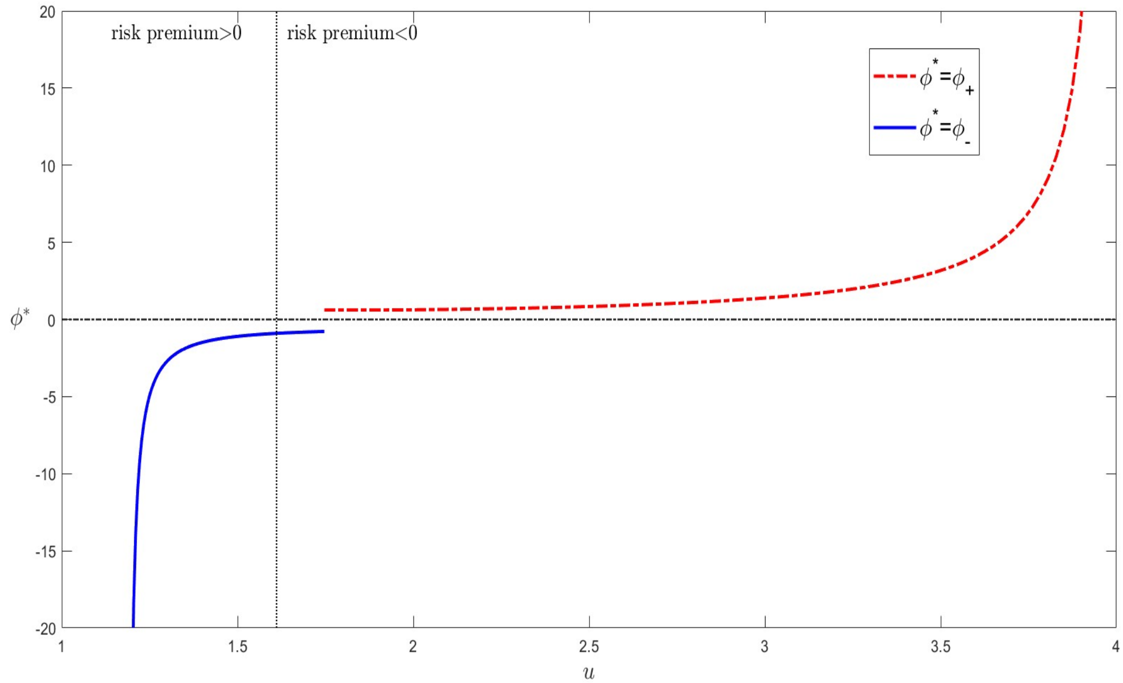

Figure 5: This figure illustrates the optimal portfolio weight $\phi^* \equiv x^*/w_0$ against u for the case $\theta > w_0 r$. Here, $A = 2$, $\gamma = 0.5$, $d = 0.5$, $r = 1$, $w_0 = 0.8$, $\theta = 1$, and $p = 0.45$.

Figure 5 illustrates the optimal portfolio weight $\phi^* \equiv x^*/w_0$ as a function of u for the case $\theta > w_0 r$.[13] When $\frac{p(u-r)^{1-\gamma}}{(1-p)(r-d)^{1-\gamma}} < 1$,[14] that is, d is a better state, the optimal portfolio weight is illustrated by the blue solid line. In this case, the optimal portfolio weight is always negative, independent of the sign of risk premium. For example, the vertical dotted line illustrates the value of u at which the risk premium is 0. The risk premium is positive when u is greater than this value. Figure 5 shows that the optimal portfolio weight can be still negative even when the risk premium is positive. As u decreases and approaches $r + k_-(r-d)$, the optimal portfolio approaches negative infinity. When $\frac{p(u-r)^{1-\gamma}}{(1-p)(r-d)^{1-\gamma}} > 1$, that is, u is a better state, the optimal portfolio weight is illustrated by the red dash-dot line. In this case, the optimal portfolio weight is always positive. As u increases and approaches $r + k_+(r-d)$, the optimal portfolio approaches positive infinity.

In the extreme case of $\frac{p(u-r)^{1-\gamma}}{(1-p)(r-d)^{1-\gamma}} = 1$, there are two different optimal portfolio weights

[13]Note that there is no optimal solution for the optimal choice problem when u is too small ($u < r + k_+(r - d)$), or too large ($u > r + k_-(r-d)$), as shown in Proposition 2.

[14]It follows that $u < \hat{u} \equiv r + [(1-p)/p]^{\frac{1}{1-\gamma}}(r-d)$.

$\phi^* = \phi_-(\equiv x_-/w_0)$ and $\phi^* = \phi_+(\equiv x_+/w_0)$, with one (ϕ_-) being negative and one (ϕ_+) positive. Both of them yield the same EU. This means that the investor can achieve the highest expected utility by either buying the risky asset or short-selling it. As a result, the optimal portfolio weight can jump from positive to negative around this extreme point $u = \hat{u}$, for a small change in parameters (including both market parameters and utility parameters). We will show in Section 3.6 that with multiple states, there can be multiple local maximums of the EU and hence multiple jumps in the optimal portfolio weight.

Due to this jump, the optimal portfolio weight is either $\phi^* \geq 0.61$ or $\phi^* \leq -0.78$ as illustrated in Figure 5. the portfolio weight cannot be optimal if it is small in magnitude $(-0.78 < \phi^* < 0.61)$. Investors with a high reference point tend to take a large position, either long or short, in the risky asset, always participating into the stock market.

Propositions 3–4 directly lead to the following results on comparative statics.

Corollary 5. *(Comparative statics.)*

1. *The optimal investment in stocks x^* and the optimal wealth depends linearly on the reference point θ.*

   The magnitudes of the optimal investment in stocks $|x^*|$, the deviation of the optimal wealth from the reference point $|w^* - \theta|$, and the value function J increases with $|\theta - w_0 r|$.

2. *When $\theta < w_0 r$ or $\theta = w_0 r$, the optimal stock investment x^* is independent of A.*

   When $\theta > w_0 r$, the optimal investment in stocks x^*, the deviation of the optimal wealth from the reference point $|w^* - \theta|$, and the value function J decrease with A.

The sign of the optimal holdings of stock has been described in Corollary 4. Corollary 5 Part 1 further describes its magnitude. Recall that when the reference point is the gross riskless rate $\theta = w_0 r$, the optimal stock investment is zero. Corollary 5 shows that as the reference point deviates from the gross riskless rate, the magnitude of stock investment increases. Because the utility is defined over the deviation of wealth from the reference point θ, both the optimal wealth and the optimal holdings depend linearly on θ.

The reference point θ significantly affects the wealth distribution when the reference point is higher than the gross riskless rate $\theta > w_0 r$. In this case, as θ increases, the wealth distribution becomes more dispersed and volatile. The investor assigns extremely positive and negative wealth levels across states.

The second part of Corollary 5 states the effect of loss aversion A on the optimal holdings of stock. When the reference point is low $\theta \leq w_0 r$, the optimal demand is independent of A. When the reference point is high $\theta > w_0 r$, as loss aversion A increases, the optimal holdings of stock always decrease for both the case $\frac{p(u-r)^{1-\gamma}}{(1-p)(r-d)^{1-\gamma}} > 1$ and $\frac{p(u-r)^{1-\gamma}}{(1-p)(r-d)^{1-\gamma}} \leq 1$, since the investor is averse to big losses. This further lowers the value function. Note that the results in Corollary 5 still hold true for more than 2 states as shown shortly in Section 3.6.

The above results show that although the loss aversion significantly affects the existence of optimal solutions (Proposition 2), it has a less significant effect on the optimal portfolios if they exist. On the other hand, the reference point does not impact the existence of optimal solutions but is the most important factor in driving the optimal portfolios.

In the above analyses, we focus on the case $A > \underline{A}$. The following proposition also describes the results for the extreme case $A = \underline{A}$,

Proposition 5. *Suppose* $A = \underline{A}$.

1. *When* $\theta < w_0 r$*, the optimal solutions are the same as those in Propositions 3–4.*
2. *When* $\theta = w_0 r$*, any* $x \geq 0$ *is optimal if* $(\frac{p}{1-p})^{\gamma}(\frac{\xi_d}{\xi_u})^{1-\gamma} \geq 1$*, and any* $x \leq 0$ *is optimal if* $(\frac{p}{1-p})^{\gamma}(\frac{\xi_d}{\xi_u})^{1-\gamma} \leq 1$*. The value function always equals zero.*
3. *When* $\theta > w_0 r$*, there is no optimal solution.*

Proposition 5 shows that the optimal solution in the case $A = \underline{A}$ exhibits very different properties from the case $A > \underline{A}$. In particular, when $\theta = w_0 r$, utility gain in one state is identical to the utility loss in the other state. As a result, the expected utility always equals zero, independent of the level of stock holdings (thus, only the sign of stock holdings is relevant).

3.4 Positive Wealth Constraint

We have shown that with the S-shaped utility function, there exist optimal portfolios for half of the parameter space, but there is no optimal solution for the other half of the parameter space. The lack of solutions over a substantial fraction of the parameter space represents a limitation of the loss aversion preferences. To have solutions, the literature sometimes imposes wealth constraint. In this section, we examine its effects. Without loss of generality, we consider positive wealth constraint:[15]

$$w \geq 0. \tag{19}$$

We show that imposing constraints often leads to corner solutions.

Assume the initial wealth is positive $w_0 > 0$.[16] The wealth constraint $w_u \geq 0$ for state u together with the budget constraint $w_d = \frac{w_0 - p\xi_u w_u}{(1-p)\xi_d}$ leads to an upper bound of wealth in state d: $w_d \leq \frac{w_0}{(1-p)\xi_d}$. Similarly, there is also an upper bound for w_u due to constraint $w_d \geq 0$. As a result, the wealth in the two states lies in the intervals:

$$w_u \in \left[0,\ \frac{w_0}{p\xi_u}\right], \qquad w_d \in \left[0,\ \frac{w_0}{(1-p)\xi_d}\right]. \tag{20}$$

With positive wealth constraint, the EU always has maximums. When $A < \underline{A}$, there is no optimal wealth for the unconstraint case. After imposing the constraint, there exists optimal wealth and it is given by corner solutions (extreme solutions):

$$\begin{cases} w_u^* = \frac{w_0}{p\xi_u} \\ w_d^* = 0 \end{cases} \quad \text{or} \quad \begin{cases} w_u^{**} = 0 \\ w_d^{**} = \frac{w_0}{(1-p)\xi_d} \end{cases}. \tag{21}$$

With a sufficiently high reference point (or low initial wealth), the optimal wealth in both states can be lower than the reference point. This significantly differs from the case without the wealth constraint, in which it cannot be optimal to assign wealth to be lower than the reference point at both states.

Now we consider $A > \underline{A}$. In this case, the unconstraint problem has optimal solutions. Especially, if $\theta > w_0 r$, the optimal wealth is negative at one of the two states. However, imposing constraint in this state binds the wealth to be zero. These results show that wealth constraint can still have effects even if the choice has optimal solution in the unconstraint case. Imposing constraints can even qualitatively change the optimal strategy.

In addition, Proposition 3 implies that when w_0 is sufficiently low or when the reference point is sufficiently high: $w_0 < \theta \min\{p\xi_u[A^{-\frac{1}{\gamma}}(\xi_u/\xi_d)^{-\frac{1}{\gamma}} + 1], (1-p)\xi_d[A^{-\frac{1}{\gamma}}(\xi_d/\xi_u)^{-\frac{1}{\gamma}} + 1]\}$,

[15] Imposing other levels of bound (either lower or upper bound) of wealth can be understood from our results by adjusting θ and w_0 by noting that our analyses generally hold for any values of them and that imposing a lower wealth bound is equivalent to certain upper wealth bound due to the budget constraint.

[16] Although the loss aversion preferences in Tversky and Kahneman (1992) allows negative wealth, there is no solution for $w_0 \leq 0$ under the positive wealth constraint.

the optimal solutions are always corner solutions, as one of the two bounds of w_u in (20) must be reached. When $A > \underline{A}$ and $\theta \leq w_0 r$, the optimal wealth tends to be interior solutions ($w_{u,d} > 0$).[17] In this case, the wealth constraint takes no effect.

Remark 1. *With wealth constraints, the optimal portfolio weights always exist but are often given by corner solutions.*

The above results highlight the stark differences in the optimal portfolios with and without wealth constraint. It is worth noting that imposing wealth constraint is similar to changing utility function. For example, imposing positive wealth constraint is equivalent to setting utility as $-\infty$ for $w \in (-\infty, 0)$. However, the original loss aversion utility is defined over all wealth levels in Tversky and Kahneman (1992).

Berkelaar et al. (2004) study optimal portfolio choice under loss aversion in a continuous-time setting. In Berkelaar et al. (2004), the markets are complete. By the Cox and Huang approach, the dynamic problem can be rewritten as a static problem ((10) in Berkelaar et al. (2004)), in which there are infinitely many states indexed by a Brownian motion. Problem (10) in Berkelaar et al. (2004) imposes a positive wealth constraint.[18] The problem studied in this section can be viewed as a single-period two-state version of Berkelaar et al. (2004). Under the positive wealth constraint, the optimal solution always exists in Berkelaar et al. (2004). Our results further suggest that this solution is unlikely to be an internal one. In addition, Berkelaar et al. (2004) find that the optimal wealth is either above the reference point or equal to zero. Our results show that the wealth level depends crucially on the levels of the reference point and loss aversion.

3.5 The case $\gamma_+ \neq \gamma_-$

The above analyses focus on the case with the same curvature parameters, $\gamma_+ = \gamma_- \equiv \gamma$. This is commonly considered in most studies. In this section, we consider the case with different γ_+ and γ_-, as generally allowed in Tversky and Kahneman (1992).[19]

[17] If $\theta < 0$, the optimal solutions can be given by corner solutions.

[18] In a continuous-time economy where the risky assets follow geometric Brownian motions, investors' wealth could be negative with a positive probability (Sethi and Taksar, 1988); thus, the wealth constraint is binding.

[19] Tversky and Kahneman (1992) estimates the same values of them: $\gamma_+ = \gamma_- = 0.12$. Wu and Gonzalez (1996) find γ_- is higher ($\gamma_+ = 0.48$, $\gamma_- = 0.63$). Abdellaoui (2000) estimate a higher γ_+ ($\gamma_+ = 0.11$,

When $\gamma_+ \neq \gamma_-$, the loss aversion utility (2) still features diminishing sensitivity and reference dependence. However, whether the investor is loss averse depends on the size of positions as shown in the following lemma.

Lemma 2. *1. When $\gamma_- < \gamma_+$, loss aversion applies when $|w - \theta| > (\frac{1-\gamma_+}{1-\gamma_-}A)^{\frac{1}{\gamma_- - \gamma_+}}$.*

2. When $\gamma_+ < \gamma_-$, loss aversion applies when $|w - \theta| < (\frac{1-\gamma_+}{1-\gamma_-}A)^{\frac{1}{\gamma_- - \gamma_+}}$.

Lemma 2 shows that when $\gamma_- < \gamma_+$, the investor is loss averse only for large positions: $|w - \theta| > (\frac{1-\gamma_+}{1-\gamma_-}A)^{\frac{1}{\gamma_- - \gamma_+}}$; however, when wealth is close to the reference point, the investor is actually more sensitive to gains than to equivalent losses, opposite to loss aversion. When $\gamma_- < \gamma_+$, the investor is loss averse only for small positions. In contrast, in the case of $\gamma_+ = \gamma_-$ as studied above, loss aversion is completely controlled by the coefficient A, and the investor is always loss averse when $A > 1$.

The following proposition discusses the existence of bounded optimal solutions for problem (5) if $\gamma_+ \neq \gamma_-$.

Proposition 6. *(Existence of bounded solutions when $\gamma_+ \neq \gamma_-$.)*

1. When $\gamma_- < \gamma_+$, the optimization problem has bounded solutions.

2. When $\gamma_+ < \gamma_-$, the optimization problem does not have bounded solutions.

Although the magnitude of loss aversion A determines the existence of optimal solutions for the case $\gamma_- = \gamma_+$, Proposition 6 shows that it does not affect the existence when $\gamma_+ \neq \gamma_-$. If $\gamma_- < \gamma_+$, the penalty for losses is strong and the optimal portfolio weights always exist. However, if $\gamma_+ < \gamma_-$, there is no internal optimal portfolio weight, independent of A. When $\gamma_- \neq \gamma_+$, the large-scale properties of the optimal portfolio are determined by both A and the relative level of γ_- and γ_+; however, the latter has a dominating effect when losses/gains are large and completely determines the boundedness of the optimal portfolio.

Proposition 7. *(Optimal wealth.) Suppose that $\gamma_- < \gamma_+$.*

1. Case 1: $\theta \leq (w_0 - 1)r$.

$\gamma_- = 0.08$).

Suppose that $A > \underline{A}^*$*, where* $\underline{A}^* = \max\{\frac{p^{-\gamma_-}}{(1-p)^{-\gamma_+}}\frac{\xi_u^{1-\gamma_-}}{\xi_d^{1-\gamma_+}}, \frac{(1-p)^{-\gamma_-}}{p^{-\gamma_+}}\frac{\xi_d^{1-\gamma_-}}{\xi_u^{1-\gamma_+}}\}$*. In this case, the optimal wealth in the two states is given by*

$$\begin{aligned} w_u^* - \theta &= \left(w_0 - r^{-1}\theta\right)\left[p\xi_u^{1-\frac{1}{\gamma_+}} + (1-p)\xi_d^{1-\frac{1}{\gamma_+}}\right]^{-1}\xi_u^{-\frac{1}{\gamma_+}} > 0, \\ w_d^* - \theta &= \left(w_0 - r^{-1}\theta\right)\left[p\xi_u^{1-\frac{1}{\gamma_+}} + (1-p)\xi_d^{1-\frac{1}{\gamma_+}}\right]^{-1}\xi_d^{-\frac{1}{\gamma_+}} > 0, \end{aligned} \tag{22}$$

and the value function is given by $J = \frac{(w_0 - r^{-1}\theta)^{1-\gamma_+}}{1-\gamma}[p\xi_u^{1-\frac{1}{\gamma_+}} + (1-p)\xi_d^{1-\frac{1}{\gamma_+}}]^{\gamma_+} > 0$.

2. *Case 2:* $(w_0 - 1)r < \theta \le w_0 r$.

   The EU has a local maximum over each of the three intervals, $w_u \in (-\infty, \theta]$*,* $w_u \in [\theta, \hat{w}_u]$*, and* $w_u \in [\hat{w}_u, +\infty)$*, and the global maximum is the greatest of them.*

3. *Case 3:* $\theta > w_0 r$.

   The EU has two local maximums, and the global maximum is the greater of them. Specifically, one local maximum occurs for $w_d < \theta < w_u$*, at which the portfolio wealth is governed by* (A.4)*, and the other local maximum occurs for* $w_u < \theta < w_d$*, at which the wealth is governed by* (A.3).

Proposition 7 part 1 shows that when the coefficient A is large and the reference point is sufficiently low (or the initial wealth is sufficiently large $w_0 \ge r^{-1}\theta + 1$), the results on the optimal wealth in this case of $\gamma_- \ne \gamma_+$ are the same as those in the case $\gamma_- = \gamma_+$ as stated in Proposition 3. In both cases, the expected utility has a unique local maximum in $w_u \in [\theta, \hat{w}_u]$, which is also the global maximum, and hence the optimal wealth is higher than the reference point at both states.

However, Proposition 7 part 2 shows that when A is small or when the reference point is equal to or lower than but close to $w_0 r$, the local maximum in $w_u \in [\theta, \hat{w}_u]$ may not be the global maximum. As a result, the optimal wealth can be lower than the reference point even when $\theta < w_0 r$, and the optimal wealth can deviate from the reference point (nonzero stock holdings) when $\theta = w_0 r$.[20] These results are significantly different from those in the case $\gamma_- = \gamma_+$ as stated in Proposition 3. In fact, for $\gamma_- < \gamma_+$, loss aversion does not apply when wealth is close to the reference point. As a result, the *EU* becomes more complex and can have multiple local maximums.

[20] The wealth is higher than the reference point only for $w_u \in [\theta, \hat{w}_u]$.

Proposition 7 part 3 shows that when $\theta > w_0 r$, the expected utility has two local maximums, and the global maximum is the greater of them. This result for the case of $\gamma_- \neq \gamma_+$ is consistent with that in the case $\gamma_- = \gamma_+$ as stated in Proposition 3.

3.6 Multiple States

We extend the results in Section 3 to a general case with S states in the world. Assume that there are $N+1 = S$ assets and the markets are complete.

3.6.1 Existence of Optimal Solutions

We divide all states into two sets, $\mathbb{S}_+$ and $\mathbb{S}_-$:

$$\mathbb{S}_+ = \{s : w_s \geq \theta\}, \qquad \mathbb{S}_- = \{s : w_s < \theta\}.$$

For notational simplicity, we use 1_+ to denote $1_{s\in\mathbb{S}_+}$ and 1_- to denote $1_{s\in\mathbb{S}_-}$. We extend the results on the existence condition for the binomial model in Proposition 2 to this general case with S states.

Lemma 3. *Given a non-trivial partition* $\{\mathbb{S}_+, \mathbb{S}_-\}$*, if*

$$\mathbb{E}[\xi^{1-1/\gamma}(1_+ - A^{1/\gamma}1_-)] > 0,$$

the EU *has no global maximum.*

Lemma 3 shows that if there exists a global maximum, we must have

$$\mathbb{E}\big[\xi^{1-1/\gamma}(1_+ - A^{1/\gamma}1_-)\big] \leq 0, \tag{23}$$

for any partition $\{\mathbb{S}_+, \mathbb{S}_-\}$. For each partition, condition (23) leads to a lower bound for A:

$$A \geq \left(\frac{\mathbb{E}[\xi^{1-1/\gamma}1_+]}{\mathbb{E}[\xi^{1-1/\gamma}1_-]}\right)^{\gamma} = \left(\frac{\sum_{s\in\mathbb{S}_+} p_s \xi_s^{1-1/\gamma}}{\sum_{s\in\mathbb{S}_-} p_s \xi_s^{1-1/\gamma}}\right)^{\gamma}, \qquad \forall \mathbb{S}_+, \mathbb{S}_-, \tag{24}$$

where p_s is the probability for state s, $s = 1, ..., S$. To find the maximum of these lower bounds, we only need to consider the cases where there is only 1 state with $w_s < \theta$ by noting that the state prices are positive. Thus, (24) is equivalent to

$$A \geq \left(\frac{\sum_{t\neq s} p_t \xi_t^{1-1/\gamma}}{p_s \xi_s^{1-1/\gamma}}\right)^{\gamma}, \qquad \forall s \in \{1, \cdots, S\}. \tag{25}$$

By further noting that the EU is a continuous function on a compact set, we have the following proposition on the conditions for the existence of global maximums.

Proposition 8. *Define*

$$\underline{A} = \max_s \left\{ \left(\frac{\sum_{t \neq s} p_t \xi_t^{1-1/\gamma}}{p_s \xi_s^{1-1/\gamma}} \right)^\gamma \right\}. \tag{26}$$

1. *When $A > \underline{A}$, the optimization problem has bounded solutions.*

2. *When $A < \underline{A}$, the optimization problem has bounded solutions.*

3. *When $A = \underline{A}$, the optimization problem has bounded solutions for $\theta \leq w_0 r$ and does not have hounded solutions for $\theta > w_0 r$.*

Proposition 8 shows that there is a lower bound $\underline{A}$ for A given by (26). The optimal portfolio choice problem has interior solutions only if A is higher than this bound.

With S states, it follows from (26) that the lower bound $\underline{A}$ must satisfy

$$\underline{A} \geq (S-1)^\gamma,$$

and the equality holds only if $p_s \xi_s^{1-1/\gamma} = p_t \xi_t^{1-1/\gamma}$, $\forall s, t$. When there are two states $S = 2$, the bound (26) reduces to (10) in Proposition 2, and it is greater than or equal to 1. As the number of states S increases, the lower bound $\underline{A}$ increases. This is because more states provide more potential opportunities that are exploited with more assets, with which the expected utility is more likely to approach infinity. When S is large, to have interior solutions, the agent must have an extremely high aversion to losses that imposes large penalty for losses. Because $\underline{A}$ increases with S without bound, when S is very large, e.g., a limiting case with an infinitely many states, the EU has no global maximum.

Corollary 6. *As the number of states S increases, the conditions for the existence of interior solutions becomes stricter.*

Proposition 8 also implies that if the EU has global optimums, we must have

$$A > \frac{p_t^\gamma \xi_t^{\gamma-1}}{p_s^\gamma \xi_s^{\gamma-1}}, \qquad \forall s \in \{1, \cdots, S\}. \tag{27}$$

To understand this result, we consider the asymptotic behavior with budget preserving changes in terminal wealth over any two states s, t, that is, $(\Delta w_s, \Delta w_t) = (\frac{\Delta}{p_s \xi_s}, -\frac{\Delta}{p_t \xi_t})$. If condition (27) is violated, the investor can always increase her expected utility without bound by assigning more wealth at state t and less wealth at state s accordingly. If there are two states $S = 2$, then (27) can be a necessary and sufficient condition for global maximums as shown in Proposition 2. When $S > 2$, it is not sufficient to guarantee the existence of global maximums and we need a higher bound $\underline{A}$.

3.6.2 Optimal Solutions

In this section, we study the optimal solutions. To this end, we assume that the *EU* has global maximums. The optimal solutions are summarized in the following proposition.

Proposition 9. *(Optimal wealth and value function.) Suppose $A > \underline{A}$.*

1. *When $\theta < w_0 r$, the optimal wealth is given by*

$$w^* - \theta = \frac{w_0 - r^{-1}\theta}{\mathbb{E}[\xi^{1-1/\gamma}]}\xi^{-1/\gamma}, \tag{28}$$

and the value function is given by

$$J = \frac{\left(w_0 - r^{-1}\theta\right)^{1-\gamma}}{1-\gamma}\left(\mathbb{E}[\xi^{1-1/\gamma}]\right)^{\gamma} > 0. \tag{29}$$

2. *When $\theta = w_0 r$, the optimal wealth is given by*

$$w^* - \theta = 0, \tag{30}$$

and the value function is given by $J = 0$.

3. *When $\theta > w_0 r$, for each non-trivial partition $\{\mathbb{S}_+, \mathbb{S}_-\}$, the EU has a local maximum, at which the wealth is given by*

$$\begin{aligned} (w-\theta)1_+ &= \frac{r^{-1}\theta - w_0}{\mathbb{E}[\xi^{1-1/\gamma}(-1_+ + A^{1/\gamma}1_-)]}\xi^{-1/\gamma}1_+, \\ (\theta - w)1_- &= \frac{r^{-1}\theta - w_0}{\mathbb{E}[\xi^{1-1/\gamma}(-1_+ + A^{1/\gamma}1_-)]}A^{1/\gamma}\xi^{-1/\gamma}1_-, \end{aligned} \tag{31}$$

and the expected utility is given by

$$EU = -\frac{(r^{-1}\theta - w_0)^{1-\gamma}}{1-\gamma}\Big(\mathbb{E}[\xi^{1-1/\gamma}(-1_+ + A^{1/\gamma}1_-)]\Big)^{\gamma} < 0. \tag{32}$$

The global maximum is the greatest of these local maximums.

Proposition 9 shows again that the optimal solutions exhibit starkly different properties for $\theta < w_0 r$, $\theta = w_0 r$, and $\theta > w_0 r$. When $\theta < w_0 r$, the *EU* has a unique local maximum, which is the global maximum. At the global maximum, the terminal wealth in all states is higher than the reference point $w^* > \theta$ (if $w_0 r > \theta$). In this case, the expected utility is always positive.

When $\theta = w_0 r$, the terminal wealth in all states is equal to the reference point $w^* = \theta$, and the expected utility equals zero.

When $\theta > w_0 r$, there are multiple local maximums, one for each non-trivial partition $\{\mathbb{S}_+, \mathbb{S}_-\}$. The global maximum is achieved at the partition that leads to the highest EU, and (32) shows that this partition leads to the lowest $\mathrm{E}[\xi^{1-1/\gamma} 1_-]$.[21] Furthermore, this lowest must be achieved for partitions with only 1 state in $\mathbb{S}_-$. That is, it is optimal to set wealth to be lower than the reference point in the state with the lowest $\xi_s^{1-1/\gamma} p_s$ and higher than the reference point in all the other states.

Corollary 7. *1. When $\theta < w_0 r$, the optimal wealth is higher than the reference point $w^* > \theta$ in all states.*

2. When $\theta = w_0 r$, the optimal wealth is equal to the reference point $w^* = \theta$ in all states.

3. When $\theta > w_0 r$, the optimal wealth is lower than the reference point in the states with the lowest $\xi_s^{1-1/\gamma} p_s$ and higher than the reference point in all the other states.

Corollary 7 shows that when the reference point is high ($\theta > w_0 r$), the optimal wealth is lower than the reference point in one state but higher in all the other states. In this case, the investor tends to pursuit negative skewness by leaving all pains in a single state.

Figure 6 illustrates the impacts of return skewness on the optimal portfolio when the reference point is high ($\theta > w_0 r$). We consider a case with three states and three assets, including a riskless asset and two risky assets (assets A and B). To examine investor's demand for skewness, we let the gross returns of the two risky assets have the same mean 1.2 and same variance 0.2. We further fix asset A's return skewness to be 0 and let asset B's return skewness vary. Panel (a) illustrates asset B's gross returns R_b in the three states. Asset B's return $R_{b,d}$ ($R_{b,u}$) in state d (u) satisfies $R_{b,d} < r < R_{b,u}$, and as $R_{b,m}$ decreases, its return skewness changes from negative to positive. Panel (b) shows that for different levels of return skewness of asset B, the optimal wealth return tends to be very negative. The skewness of the optimal wealth is much lower than the skewness of all individual assets. This is because the optimal wealth is higher than the reference point in two states (u and m) but

[21] Equation (32) shows that this partition leads to the smallest $\mathbb{E}[\xi^{1-1/\gamma}(-1_+ + A^{1/\gamma} 1_-)]$. Note that $\mathbb{E}[\xi^{1-1/\gamma}(-1_+ + A^{1/\gamma} 1_-)] = (A^{1/\gamma} + 1)\mathbb{E}[\xi^{1-1/\gamma} 1_-] - \mathbb{E}[\xi^{1-1/\gamma}]$. So the global maximum occurs for the partition with the lowest $\mathrm{E}[\xi^{1-1/\gamma} 1_-]$.

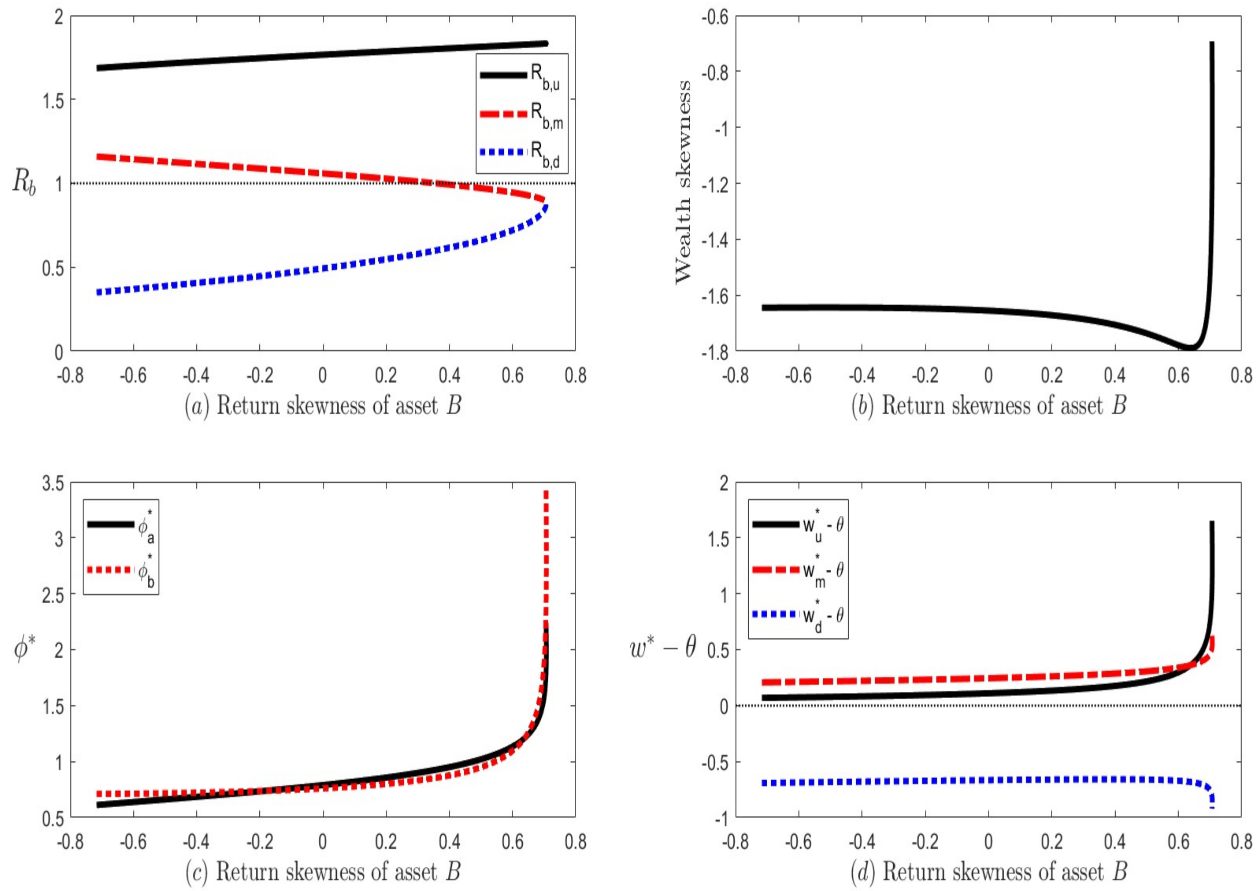

Figure 6: This figure illustrates the impacts of return skewness on the optimal portfolio for the case $\theta > w_0 r$. The investment opportunity set contains three assets: a riskless asset and two risky assets (A and B). The gross returns of the two risky assets have the same mean 1.2 and same variance 0.2. We set asset A's return skewness to be 0 and let asset B's skewness vary. Panel (a) plots asset B's returns in the three states (u, m, and d) against its return skewness. Panel (b) illustrates the return skewness of the optimal portfolio wealth. Panel (c) illustrates optimal portfolio weights of the two risky assets. Panel (d) illustrates the difference between the optimal portfolio wealth and the reference point ($w^* - \theta$) in the three states. Here, $A = 8$, $\gamma = 0.5$, $r = 1$, $w_0 = 0.8$, $\theta = 1$, and the probabilities of the three states $(p_1, p_2, p_3) = (1/2, 1/3, 1/6)$.

much lower than the reference point in the other (state d) as illustrated in panel (d). As a result, with low initial wealth, the agent actually seeks negative skewness.

Corollary 7 further shows that the "uninsured" state with wealth lower than the reference point is the one with the lowest $\xi_s^{1-1/\gamma} p_s$. It tends to be the worst state of the world with the highest state price, against which it is most expensive to protect.

4 Incomplete Markets

Section 3 focuses on portfolio choice problems under complete markets. In this section, we consider incomplete markets. In this case, there are infinitely many strictly positive state prices, which lead to infinitely many constraints. We transform the problem to one with finitely many constraints that can be solved with the Lagrangian approach. We find that the optimal solution exhibits similar properties to that under complete markets.

With incomplete markets, there are infinitely many risk neutral measures. The maximization problem is:

$$\begin{aligned} &\max && \mathbb{E}[U(w)] \\ &\text{subject to} && \mathbb{E}^{\mathbb{Q}}[w/w_0] = r, \quad \forall \mathbb{Q} \in \mathcal{M}, \end{aligned} \tag{33}$$

where $\mathcal{M}$ is an infinite set of all risk neutral measures. Therefore, there are infinitely many constraints in problem (33). Indeed, these infinitely many restrictions can be reduced to finite ones. The intuitive reason is that there are finite states, so we can always find a finite base to represent the set $\mathcal{M}$. For simplicity, we use the following notations:

Definition 1. *Let $W = \{(q_1, \cdots, q_n) \in \mathbb{R}^n \big| q_1 R_1 + \cdots + q_n R_n = r\}$, where R_i is the gross return of an asset at the ith state, and $\mathcal{P}^+ = \{(q_1, \cdots, q_n) | q_1 + \cdots + q_n = 1, q_i \geq 0, \forall i = 1, 2, \cdots, n\}$. Then the set of risk-neutral measures is given by*

$$\mathcal{M} = W \cap \mathcal{P}^+. \tag{34}$$

We can always find a finite subset of $\mathcal{M}$, say $\mathcal{Q}$, such that every element in $\mathcal{M}$ can be represented as a linear combination of $\mathcal{Q}$. We denote $\mathcal{Q} = \{\mathbb{Q}_1, \cdots, \mathbb{Q}_l\}$, where $l \leq n$ and $\mathbb{Q}_i$ satisfies $\mathbb{E}^{\mathbb{Q}_i}[R] = r$, $i = 1, \cdots, l$. Therefore, for any $\mathbb{Q} \in \mathcal{M}$, it can be written as a linear combination of $\mathbb{Q}_i$, that is,

$$\mathbb{Q} = \alpha_1 \mathbb{Q}_1 + \cdots + \alpha_l \mathbb{Q}_l, \quad \alpha_1 + \cdots + \alpha_l = 1. \tag{35}$$

Hence $\mathbb{E}^{\mathbb{Q}}[R] = \alpha_1 \mathbb{E}^{\mathbb{Q}_1}[R] + \cdots + \alpha_l \mathbb{E}^{\mathbb{Q}_l}[R] = \alpha_1 r + \cdots + \alpha_l r = r$.

Lemma 4. *Maximization problem* (33) *with infinitely many constraints can be reduced to the following one with finitely many constraints:*

$$\begin{aligned} &\max && \mathbb{E}[U(w)] \\ &\textit{subject to} && \mathbb{E}^{\mathbb{Q}}[w] = w_0 r, \quad \forall \mathbb{Q} \in \mathcal{Q}. \end{aligned} \tag{36}$$

The Lagrangian is given by $\mathcal{L} = \mathbb{E}[U(w)] - \sum_{i=1}^{l} \lambda_i (\mathbb{E}^{\mathbb{Q}_i}[w] - w_0 r)$.

We reduce the portfolio choice problem under incomplete markets to one with a finite number of constraints. As a result, this problem can be solved with the Lagrangian approach.

4.1 Trinomial Model

We consider a trinomial model. Suppose that there are only two assets, one risky and one riskless assets, and three different states, $\omega_1, \omega_2, \omega_3$. The return of the risky asset equals u, m, d at the three states, respectively. Without loss of generality, we assume $d < m < u$.

For any risk neutral measure $\mathbb{Q} = (q_1, q_2, q_3)$, it must satisfies

$$\begin{cases} u q_1 + m q_2 + d q_3 = r, \\ q_1 + q_2 + q_3 = 1, \end{cases} \tag{37}$$

and $q_i \geq 0$, $i = 1, 2, 3$. The fundamental set of solutions is given by

$$\mathbb{Q} = k \left(\frac{m-d}{u-m}, \frac{d-u}{u-m}, 1 \right) + \left(\frac{r-m}{u-m}, \frac{u-r}{u-m}, 0 \right), \tag{38}$$

where the range of k is determined by the following inequalities:

$$\begin{cases} 0 \leq k \dfrac{m-d}{u-m} + \dfrac{r-m}{u-m} \leq 1, \\ 0 \leq k \dfrac{d-u}{u-m} + \dfrac{u-r}{u-m} \leq 1, \\ 0 \leq k \leq 1. \end{cases} \tag{39}$$

That is,

$$\max \left(0, \frac{m-r}{m-d} \right) \leq k \leq \frac{u-r}{u-d}. \tag{40}$$

Therefore, any risk neutral measure $\mathbb{Q}$ can be written as $\mathbb{Q} = \alpha\mathbb{Q}_1 + (1-\alpha)\mathbb{Q}_2$, where $\alpha \in [0,1]$, and the bases are given by

$$\begin{cases} \mathbb{Q}_1 = \left(\frac{r-m}{u-m}, \frac{u-r}{u-m}, 0\right), \quad \mathbb{Q}_2 = \left(\frac{r-d}{u-d}, 0, \frac{u-r}{u-d}\right), \quad \text{if} \quad m < r, \\ \mathbb{Q}_1 = \left(0, \frac{r-d}{m-d}, \frac{m-r}{m-d}\right), \quad \mathbb{Q}_2 = \left(\frac{r-d}{u-d}, 0, \frac{u-r}{u-d}\right), \quad \text{if} \quad m \geq r. \end{cases} \tag{41}$$

Corollary 8. *The optimal portfolio choice problem* (33) *with three states and two assets can be rewritten as the following one with two constraints:*

$$\begin{aligned} \max \quad & p_1 U(w_u) + p_2 U(w_m) + p_3 U(w_d) \\ \text{subject to} \quad & \mathbb{E}^{\mathbb{Q}_1}[w] = w_0 r, \\ & \mathbb{E}^{\mathbb{Q}_2}[w] = w_0 r, \end{aligned} \tag{42}$$

where $\mathbb{Q}_1$ *and* $\mathbb{Q}_2$ *are given by* (41).

With three states, Corollary 8 shows that the optimization problem involves two constraints. In contrast, under complete markets, only one constraint is needed (see, e.g., (A.5)). As a result, the solution existence condition in incomplete markets tends to be weaker than in complete markets. The following proposition states the conditions under which the optimization problem (42) has bounded solutions.

Proposition 10. *(Existence of solutions.) Define*

$$\underline{A}^{incomplete} = \max\left\{ \frac{\mathbb{E}\big[[(R-r)^+]^{1-\gamma})\big]}{\mathbb{E}\big[[(r-R)^+]^{1-\gamma})\big]}, \frac{\mathbb{E}\big[[(r-R)^+]^{1-\gamma})\big]}{\mathbb{E}\big[[(R-r)^+]^{1-\gamma})\big]} \right\}, \tag{43}$$

where $(a)^+$ *equals* a *if* $a > 0$*, and 0 otherwise.*

1. *When* $A > \underline{A}^{incomplete}$*, the optimization problem* (42) *has bounded solutions.*
2. *When* $A < \underline{A}^{incomplete}$*, the optimization problem does not have bounded solutions.*

In (43), all states are divided into two groups, one with positive excess returns and the other with negative excess returns. Proposition 10 shows that bounded optimal portfolios exist when the loss aversion coefficient A is higher than the ratio of the probability-weighted sums of utility gains between the two groups. These solution existence conditions under incomplete markets are weaker than those under complete markets as stated in Proposition 8. Under complete markets, the investor can trade the Arrow-Debreu security for any state, and the solution existence conditions depend on the utility gain in any one state relative to that in all the other states.

The solution existence conditions in Proposition 10 are consistent with the general "well-posedness" condition that is defined in terms of the large-loss aversion degree (LLAD) in He and Zhou (2011).[22] However, the conditions under complete markets as stated in Proposition 8 are different from and stricter than the condition in He and Zhou (2011).

Proposition 11. *(Optimal wealth.) Suppose* $A > \underline{A}^{incomplete}$.

1. *When* $\theta < w_0 r$, *if* $A \geq \frac{p_1}{p_3}(\frac{u-r}{r-d})^{1-\gamma}$ *and* $A \geq \frac{p_3}{p_1}(\frac{r-d}{u-r})^{1-\gamma}$,[23] *the EU has a unique global maximum, at which the optimal wealth is positive and governed by* (A.9) *and* (A.12) *when* $m < r$ *and governed by* (A.10) *and* (A.16) *when* $m \geq r$. *The value function is always positive.*

2. *When* $\theta = w_0 r$, *the optimal wealth in the three states is given by* $w_u^* = w_m^* = w_d^* = \theta$, *and the value function equals zero.*

3. *When* $\theta > w_0 r$, *the EU has two local maximums, and the global maximum is the greater of them. Specifically, when* $m < r$, *one local maximum occurs for* $w_u < \theta$ *and* $w_m, w_d > \theta$, *at which the wealth is governed by* (A.9) *and* (A.13); *another local maximum occurs for* $w_u > \theta$ *and* $w_m, w_d < \theta$, *at which the wealth is governed by* (A.9) *and* (A.14). *When* $m \geq r$, *one local maximum occurs for* $w_d < \theta$ *and* $w_u, w_m > \theta$, *at which the wealth is governed by* (A.10) *and* (A.17); *another local maximum occurs for* $w_d > \theta$ *and* $w_u, w_m < \theta$, *at which the wealth is governed by* (A.10) *and* (A.18).

Proposition 11 shows that when $\theta < w_0 r$, the expected utility has a unique global maximum, at which the optimal wealth is positive across all states. When $\theta = w_0 r$, the investor

[22] He and Zhou (2011) define LLAD $\equiv \lim_{x\to\infty}[u(-x)/u(x)]$. For utility (2), LLAD equals the loss aversion A if $\gamma_- = \gamma_+$, ∞ if $\gamma_- < \gamma_+$, and 0 if $\gamma_- > \gamma_+$. Our results in Proposition 10 are consistent with their general well-posedness condition. Conditions in Proposition 10 differ from the extreme-risk avoidance (XRA) condition imposed in Ingersoll (2016) that is used to prevent unbounded utility from above. Ingersoll (2016) defines that a utility has XRA if $\lim_{x\to\infty}[u(x)/u(-kx)]$, $\forall k > 0$. The XRA requires $\gamma_- < \gamma_+$, which is consistent with our results in Proposition 6.

[23] These are technical assumptions that guarantee $EU < 0$ over $w_u \in [\tilde{w}_u, \hat{w}_u]$ in the case $m < r$ and over $w_d \in [\hat{w}_d, \tilde{w}_d]$ in the case $m \geq r$, respectively. However, numerical simulations show that EU is decreasing in both intervals under typical parameters, and theoretically $\frac{\partial EU}{\partial w_u} = -\infty$ at the four boundary points. Comparing with the bounded solution existence condition (43), these assumptions tend to require higher A.

holds only the riskless asset and the optimal wealth is at the reference point in all states. These results for the cases $\theta < w_0 r$ and $\theta = w_0 r$ under incomplete markets are the same as those under complete markets as stated in Proposition 9. Although incomplete markets reduce some potential investment opportunities, the investor can assign her wealth at least as $w_0 r$, which is higher than the reference point, by simply holding the riskless asset. As a result, the optimal wealth will not drop in the loss domain in any state, regardless of the completeness of the markets.

However, when $\theta > w_0 r$, the optimal portfolio are significantly different under complete markets and under incomplete markets. Under complete markets, the investor assigns wealth to be lower than the reference point in only one state, against which it is most expensive to protect, but higher in all the other states, as shown in Proposition 9. However, under incomplete markets, the investor divides all states into two groups, one with positive excess returns and the other with negative excess returns, and she assigns positive wealth in all the states in one group and negative wealth in all the states in the other group. In fact, incomplete markets allow less opportunities, and the investor may not be able to allocate her wealth to be lower than the reference point only in a single (the most expensive) state.

The above results are further illustrated in the following example.

Example 1. *Suppose that* $r = 1.03$, $u = 1.10, m = 1.00$, *and* $d = 0.95$. *Then the risk-neutral measure* $\mathbb{Q} = (q_1, q_2, q_3)$ *satisfies*

$$\begin{cases} 1.10q_1 + 1.00q_2 + 0.95q_3 = 1.03, \\ q_1 + q_2 + q_3 = 1. \end{cases}$$

So the risk-neutral measure has the form $\mathbb{Q} = (8/15+k, -3k, 7/15+2k)$, *where* $k \in [-7/30, 0]$. *Taking* $k = -7/30$ *and* 0, *we can get the bases:*

$$\mathbb{Q}_1 = (3/10, 7/10, 0), \quad \mathbb{Q}_2 = (8/15, 0, 7/15).$$

Therefore, our optimization problem is

$$\begin{aligned} \max \quad & p_1 U(w_u) + p_2 U(w_m) + p_3 U(w_d) \\ \text{subject to} \quad & 3/10 w_u + 7/10 w_m = w_0 r, \\ & 8/15 w_u + 7/15 w_d = w_0 r, \end{aligned} \tag{44}$$

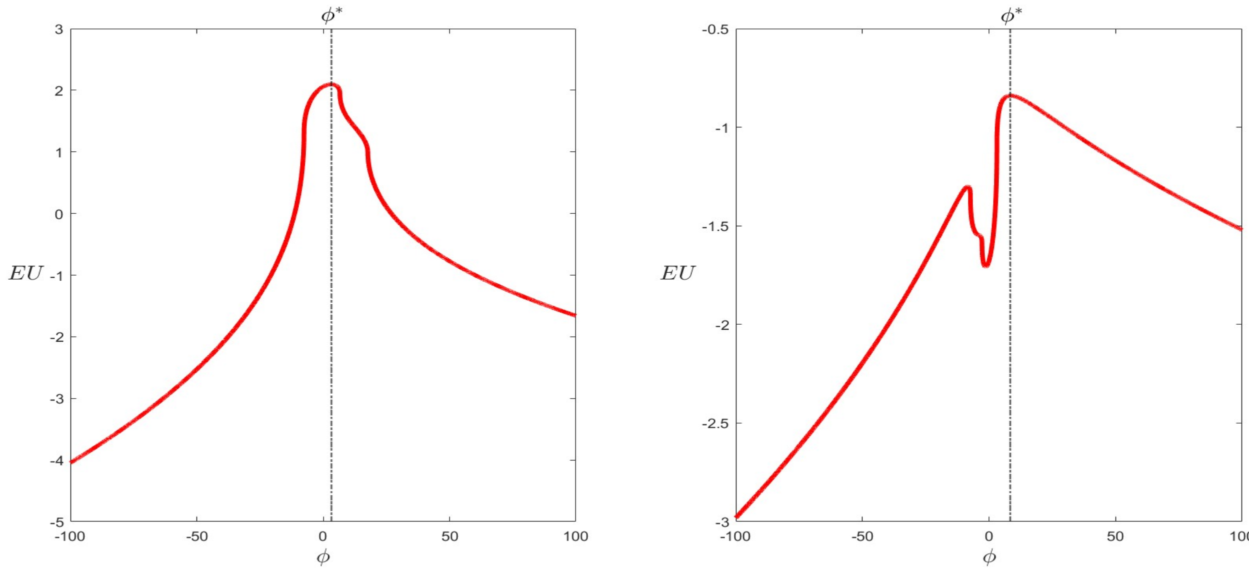

Figure 7: The figure plots the expected utility function EU against the portfolio weight $\phi(\equiv x/w_0)$ for $\theta \leq w_0 r$ (the left panel) and $\theta > w_0 r$ (the right panel). Here, $A = 2$, $\gamma = 0.5$, $\theta = 1$, $r = 1.03$, $u = 1.1$, $m = 1$, $d = 0.95$, and $\mathbb{P} = (1/2, 1/3, 1/6)$. We also set $w_0 = 2$ $(> r^{-1}\theta)$ for the left panel and $w_0 = 0.8$ $(< r^{-1}\theta)$ for the right panel.

where $\mathbb{P} = (p_1, p_2, p_3)$ is the physical probability measure. Solving this optimization problem is sufficient to get optimal wealth.

Figure 7 illustrates the expected utility function EU as a function of the portfolio weight ϕ. Here we set a high value of $A = 2$ to have global maximum.[24] *In the left panel, $\theta \leq w_0 r$. The EU has a unique local maximum (at $\phi^* = 3.14$), which is also the global maximum. In this case, the value function is positive. In the right panel, $\theta > w_0 r$. There are two local maximums, and the global maximum is the greater of them (at $\phi^* = 8.51$). In this case, the value function is negative.*

4.2 Lognormally Distributed Returns

When the risky asset returns follow lognormal distributions (e.g., Merton, 1973), the markets are incomplete in a static portfolio choice problem. The solution existence conditions are

[24] If the portfolio choice problem under complete markets has optimal solutions, the corresponding problem under incomplete markets (i.e., with less assets available) should also have optimal solutions, since the EU in the former case should be higher than or equal to that in the latter case.

given by He and Zhou (2011, Theorem 2). Although there may not exist bounded optimal solutions in this case, the condition of nonexistence of solutions is stricter than the case of complete markets as stated in Section 3.6. This is because incomplete markets restrict investment opportunities and largely prevent investor's utility from approaching infinity.

5 Conclusion

Loss aversion is one of the central topics in behavioural finance. When studying its implications for portfolio choice and asset pricing, the literature often applies certain variations to the utility function proposed in Tversky and Kahneman (1992) to ensure that the optimal portfolio is bounded and computes the optimal portfolio numerically. However, the properties of the optimal policy under loss aversion have not been rigorously understood. Furthermore, the reference point plays a key role in decision making under loss aversion, and different reference points have been used in the literature. However, there is a lack of systematic understanding of its effects. This paper studies the optimal choice under pure loss aversion. We analytically solve for the existence conditions of optimal portfolios and provide a general analysis of the reference point.

We show that there is no optimal solution for half of the parameter space in which the loss aversion is low. This is essentially due to insufficient penalty for losses. However, there always exist optimal solutions for the other half of the parameter space. The lower bound of loss aversion above which bounded optimal solutions exist increases without bound as asset returns are more skewed or the number of states of the world increases. Bounded solutions are more likely to exist under incomplete markets than complete markets. Under wealth constraint, optimal solutions always exist but are often given by corner solutions. The lack of solution over a substantial fraction of the parameter space and in some economically relevant cases represents a limitation of loss aversion in the application in economics.[25]

We find that when the optimal solutions exist, the reference point plays a key role in determining the properties of the optimal portfolios. One popular reference point chosen in the literature is the investor's current wealth. We find that in this case, the investor behaves identically to a standard risk averse investor, and her holdings of the risky asset are

[25]For example, De Giorgi, Hens and Levy (2004) find that a CAPM equilibrium does not exist when investors have prospect theory preferences with heterogeneous preference parameters.

determined by the risk premium. Another commonly used reference point is the investor's current wealth scaled up by the gross riskless rate. Our results show that a loss averse investor in this case refrains from investing in the risky asset, irrespective of its return distribution. In the above two cases, the optimal portfolio (if exists) does not depend on the coefficients of loss aversion and risk seeking. This result starkly differs from that in the models that use certain portfolio constraints or impose certain restrictions/variations to the utility function.

Loss aversion truly has effects on the optimal portfolio only when the investor has a reference point higher than her wealth multiplied by the gross riskless rate, which is, however, less studied in the literature. In this case, the coefficients of loss aversion and risk seeking start to affect the optimal solutions (if exist), and the optimal portfolios distinctly differ from those under the risk aversion utility functions. We show that the optimal stock holdings are not determined by the risk premium. The investor can go long (short) an asset with an arbitrarily low (high) risk premium. The investor trades more aggressively than a risk averse investor and tends to take a large position, either long or short, in the risky asset, especially when the asset's returns are highly skewed. The nonconcavity of the loss aversion utility function leads to discontinuity in the parameter space. A small parameter change can result in a big jump in the optimal portfolio and the value function, which poses significant challenges to numerical solution methods as usually used in this literature.

A Proofs

A.1 Proof of Proposition 1

Suppose that the terminal wealth is higher than or equal to the reference point at all states. Then the value function under the HARA utility function (4) is given by

$$\begin{aligned} &\max_{\boldsymbol{x}} \mathbb{E}\Big[\frac{(w-\theta)^{1-\gamma}}{1-\gamma}\Big] \\ =&\max_{\boldsymbol{x}} \mathbb{E}\Big[\frac{[(w_0-\boldsymbol{x}'\mathbf{1})r+\boldsymbol{x}'\mathbf{R}-\theta]^{1-\gamma}}{1-\gamma}\Big] \\ =&\max_{\boldsymbol{x}} \mathbb{E}\Big[\frac{[(\hat{w}_0-\boldsymbol{x}'\mathbf{1})r+\boldsymbol{x}'\mathbf{R}]^{1-\gamma}}{1-\gamma}\Big], \end{aligned} \tag{A.1}$$

where $\hat{w}_0 = w_0 - r^{-1}\theta$. Equation (A.1) shows that when $\hat{w}_0 = w_0 - r^{-1}\theta \geq 0$, the optimal portfolio problem is equivalent to the standard portfolio problem under the CRRA utility with nonnegative initial wealth. In the latter problem, the optimal wealth is always nonnegative, $\hat{w}^* = (\hat{w}_0 - \boldsymbol{x}^{*\prime}\mathbf{1})r + \boldsymbol{x}^{*\prime}\mathbf{R} \geq 0$, which is equivalent to $w^* - \theta \geq 0$, where the equality holds when $\hat{w}_0 = 0$, and $\boldsymbol{x}^*$ is the optimal holdings of risky assets.

On the other hand, the optimal wealth under the loss aversion utility function (3) is also higher than or equal to the reference point when $w_0 - r^{-1}\theta \geq 0$, as shown in Proposition 9. Therefore, the HARA utility and the loss aversion utility, which are concave and identical in the gain domain, produce the same optimal portfolio in this case.

A.2 Proof of Lemma 1

Suppose w_u is sufficiently large. Then w_d is low by the budget constraint (6), and the expected utility becomes

$$EU = \frac{1}{1-\gamma}\Big[p(w_u-\theta)^{1-\gamma} - A(1-p)\Big(\theta - \frac{w_0 - p\xi_u w_u}{(1-p)\xi_d}\Big)^{1-\gamma}\Big].$$

When $w_u \to +\infty$,

$$EU \approx \Big[p - A(1-p)\Big(\frac{p\xi_u}{(1-p)\xi_d}\Big)^{1-\gamma}\Big]\frac{w_u^{1-\gamma}}{1-\gamma},$$

leading to the asymptotic behavior (8).

When w_u is sufficiently low, the expected utility becomes

$$EU = \frac{1}{1-\gamma}\Big[-Ap(\theta-w_u)^{1-\gamma} + (1-p)\Big(\frac{w_0 - p\xi_u w_u}{(1-p)\xi_d} - \theta\Big)^{1-\gamma}\Big].$$

When $w_u \to -\infty$,

$$EU \approx \Big[- Ap + (1-p)\Big(\frac{p\xi_u}{(1-p)\xi_d}\Big)^{1-\gamma}\Big] \frac{(-w_u)^{1-\gamma}}{1-\gamma},$$

leading to the asymptotic behavior (9).

A.3 Proof of Proposition 2

Lemma 1 shows that when $A > \underline{A}$, the EU goes to $-\infty$ as $w_u \to \pm\infty$. Because the EU is a continuous function on a compact set, the EU has a global maximum.

When $A < \underline{A}$, the EU goes to $+\infty$ as $w_u \to +\infty$ or as $w_u \to -\infty$. In this case, there is no optimal solution for problem (5).

When $A = \underline{A}$, whether there exist optimal solutions depends on the relative level of the reference point and initial wealth. When $\theta > w_0 r$, the EU is negative as shown in Appendix A.4. The EU approaches zero as $x \to +\infty$ if $(\frac{p}{1-p})^{\gamma}(\frac{\xi_d}{\xi_u})^{1-\gamma} \geq 1$, or as $x \to -\infty$ if $(\frac{p}{1-p})^{\gamma}(\frac{\xi_d}{\xi_u})^{1-\gamma} \leq 1$. In either case, the optimal holdings are infinite. When $\theta < w_0 r$, the optimal holdings are bounded as shown Appendix A.4. When $\theta = w_0 r$, if $(\frac{p}{1-p})^{\gamma}(\frac{\xi_d}{\xi_u})^{1-\gamma} \geq 1$, any $x \geq 0$ is optimal (the EU is zero), and if $(\frac{p}{1-p})^{\gamma}(\frac{\xi_d}{\xi_u})^{1-\gamma} \leq 1$, any $x \leq 0$ is optimal (the EU is zero).

A.4 Proof of Proposition 3

The EU as a function of w_u is given by

$$EU = \begin{cases} \frac{1}{1-\gamma}\big[p(w_u-\theta)^{1-\gamma} + (1-p)(\frac{w_0 - p\xi_u w_u}{(1-p)\xi_d} - \theta)^{1-\gamma}\big], & \text{if} \quad w_u \geq \theta \text{ and } w_u \leq \hat{w}_u; \\ \frac{1}{1-\gamma}\big[p(w_u-\theta)^{1-\gamma} - A(1-p)(\theta - \frac{w_0 - p\xi_u w_u}{(1-p)\xi_d})^{1-\gamma}\big], & \text{if} \quad w_u \geq \theta \text{ and } w_u \geq \hat{w}_u; \\ \frac{1}{1-\gamma}\big[-Ap(\theta - w_u)^{1-\gamma} + (1-p)(\frac{w_0 - p\xi_u w_u}{(1-p)\xi_d} - \theta)^{1-\gamma}\big], & \text{if} \quad w_u \leq \theta \text{ and } w_u \leq \hat{w}_u; \\ -\frac{A}{1-\gamma}\big[p(\theta - w_u)^{1-\gamma} + (1-p)(\theta - \frac{w_0 - p\xi_u w_u}{(1-p)\xi_d})^{1-\gamma}\big], & \text{if} \quad w_u \leq \theta \text{ and } w_u \geq \hat{w}_u, \end{cases}$$

where $\hat{w}_u \equiv \frac{w_0 - (1-p)\xi_d \theta}{p\xi_u}$.

A.4.1 $w_0 \geq r^{-1}\theta$

Note that $r^{-1} = \mathbb{E}[\xi] = p\xi_u + (1-p)\xi_d$. When $w_0 \geq r^{-1}\theta$, we have $\hat{w}_u \geq \theta$. Thus, the case $w_u < \theta$ and $w_u > \hat{w}_u$ cannot occur. The EU becomes

$$EU = \begin{cases} \frac{1}{1-\gamma}\big[p(w_u-\theta)^{1-\gamma} + (1-p)(\frac{p\xi_u}{(1-p)\xi_d})^{1-\gamma}(\hat{w}_u - w_u)^{1-\gamma}\big], & \text{if} \quad w_u \in [\theta, \hat{w}_u]; \\ \frac{1}{1-\gamma}\big[p(w_u-\theta)^{1-\gamma} - A(1-p)(\frac{p\xi_u}{(1-p)\xi_d})^{1-\gamma}(w_u - \hat{w}_u)^{1-\gamma}\big], & \text{if} \quad w_u \in [\hat{w}_u, +\infty); \\ \frac{1}{1-\gamma}\big[-Ap(\theta - w_u)^{1-\gamma} + (1-p)(\frac{p\xi_u}{(1-p)\xi_d})^{1-\gamma}(\hat{w}_u - w_u)^{1-\gamma}\big], & \text{if} \quad w_u \in (-\infty, \theta]. \end{cases}$$

The EU is concave when $w_u \in [\theta, \hat{w}_u]$; thus, it has a local maximum between θ and $\hat{w}_u$.[26] In addition,

$$EU = \begin{cases} \frac{w_u^{1-\gamma}}{1-\gamma}\big[p - A(1-p)(\frac{p\xi_u}{(1-p)\xi_d})^{1-\gamma}\big], & \text{when} \quad w_u \to +\infty; \\ \frac{(-w_u)^{1-\gamma}}{1-\gamma}\big[-Ap + (1-p)(\frac{p\xi_u}{(1-p)\xi_d})^{1-\gamma}\big], & \text{when} \quad w_u \to -\infty. \end{cases}$$

If $p - A(1-p)(\frac{p\xi_u}{(1-p)\xi_d})^{1-\gamma} > 0$, the EU approaches $+\infty$ when $w_u \to +\infty$, the EU has a local minimum in the interval $(\hat{w}_u, +\infty)$ and there is no global maximum.

If $-Ap + (1-p)(\frac{p\xi_u}{(1-p)\xi_d})^{1-\gamma} > 0$, the EU approaches $+\infty$ when $w_u \to -\infty$, the EU has a local minimum in the interval $[\theta, +\infty)$ and there is no global maximum.

Thus, to have a global maximum, we focus on the case $p - A(1-p)(\frac{p\xi_p u}{(1-p)\xi_d})^{1-\gamma} \leq 0$ and $-Ap + (1-p)(\frac{p\xi_u}{(1-p)\xi_d})^{1-\gamma} \leq 0$.

1. if $p - A(1-p)(\frac{p\xi_u}{(1-p)\xi_d})^{1-\gamma} < 0$, when $w_u \to +\infty$, the EU approaches $-\infty$.
2. if $-Ap + (1-p)(\frac{p\xi_u}{(1-p)\xi_d})^{1-\gamma} < 0$, when $w_u \to -\infty$, the EU approaches $-\infty$.
3. if $p - A(1-p)(\frac{p\xi_u}{(1-p)\xi_d})^{1-\gamma} = 0$, when $w_u \to +\infty$, the EU approaches 0. In this case, we have $-Ap + (1-p)(\frac{p\xi_u}{(1-p)\xi_d})^{1-\gamma} < 0$.
4. if $-Ap + (1-p)(\frac{p\xi_u}{(1-p)\xi_d})^{1-\gamma} = 0$, when $w_u \to -\infty$, the EU approaches 0. In this case, we have $p - A(1-p)(\frac{p\xi_u}{(1-p)\xi_d})^{1-\gamma} < 0$.

Finally, if $p - A(1-p)(\frac{p\xi_u}{(1-p)\xi_d})^{1-\gamma} \leq 0$, we have $\frac{\partial EU}{\partial w_u} < 0$ over $w_u \in [\hat{w}_u, +\infty)$, and if $-Ap + (1-p)(\frac{p\xi_u}{(1-p)\xi_d})^{1-\gamma} \leq 0$, we have $\frac{\partial EU}{\partial w_u} > 0$ over $w_u \in (-\infty, \theta]$. Therefore, the local maximum between θ and $\hat{w}_u$ is the global maximum.

[26]Here we assume $\gamma > 0$. If $\gamma = 0$, the EU is linear not concave in the interval $w_u \in [\theta, \hat{w}_u]$, and the optimum occurs at the boundary.

Now, we derive the optimal wealth and value function. We use the Lagrangian approach to find local maximum. The Lagrangian is given by

$$\mathcal{L} = p\frac{(w_u-\theta)^{1-\gamma}}{1-\gamma} + (1-p)\frac{(w_d-\theta)^{1-\gamma}}{1-\gamma} - \lambda\big[p\xi_u w_u + (1-p)\xi_d w_d - w_0\big],$$

where $\lambda > 0$ is the Lagrange multiplier. The FOC for the local maximum is

$$(w_u-\theta)^{-\gamma} = \lambda\xi_u, \qquad (w_d-\theta)^{-\gamma} = \lambda\xi_d.$$

So

$$w_u - \theta = (\xi_d/\xi_u)^{\frac{1}{\gamma}}(w_d-\theta).$$

The budget constraint leads to

$$w_0 - r^{-1}\theta = p\xi_u(w_u-\theta) + (1-p)\xi_d(w_d-\theta) = \big[p\xi_u(\xi_d/\xi_u)^{\frac{1}{\gamma}} + (1-p)\xi_d\big](w_d-\theta).$$

The value function is

$$\begin{aligned} J &= \frac{1}{1-\gamma}\big[p(\xi_d/\xi_u)^{\frac{1-\gamma}{\gamma}} + (1-p)\big](w_d-\theta)^{1-\gamma} \\ &= \frac{(w_0-r^{-1}\theta)^{1-\gamma}}{1-\gamma}\big[p(\xi_d/\xi_u)^{\frac{1-\gamma}{\gamma}} + (1-p)\big](w_d-\theta)^{1-\gamma}\big[p\xi_u(\xi_d/\xi_u)^{\frac{1}{\gamma}} + (1-p)\xi_d\big]^{\gamma-1} \\ &= \frac{(w_0-r^{-1}\theta)^{1-\gamma}}{1-\gamma}\Big[p\xi_u^{1-\frac{1}{\gamma}} + (1-p)\xi_d^{1-\frac{1}{\gamma}}\Big]^{\gamma}. \end{aligned}$$

A.4.2 $w_0 < r^{-1}\theta$

In this case, $\hat{w}_u < \theta$; thus, the case $w_u > \theta$ and $w_u < \hat{w}_u$ cannot occur. So the EU is

$$EU = \begin{cases} \frac{1}{1-\gamma}\big[p(w_u-\theta)^{1-\gamma} - A(1-p)(\frac{p\xi_u}{(1-p)\xi_d})^{1-\gamma}(w_u-\hat{w}_u)^{1-\gamma}\big], & \text{if} \quad w_u \in [\theta, +\infty); \\ \frac{1}{1-\gamma}\big[-Ap(\theta-w_u)^{1-\gamma} + (1-p)(\frac{p\xi_u}{(1-p)\xi_d})^{1-\gamma}(\hat{w}_u-w_u)^{1-\gamma}\big], & \text{if} \quad w_u \in (-\infty, \hat{w}_u]; \\ -\frac{1}{1-\gamma}\big[A(p(\theta-w_u)^{1-\gamma} + (1-p)(\frac{p\xi_u}{(1-p)\xi_d})^{1-\gamma}(w_u-\hat{w}_u)^{1-\gamma}\big], & \text{if} \quad w_u \in [\hat{w}_u, \theta]. \end{cases}$$

The utility function is convex when $w_u \in [\hat{w}_u, \theta]$ and

$$EU = \begin{cases} \frac{w_u^{1-\gamma}}{1-\gamma}\big[p - A(1-p)(\frac{p\xi_u}{(1-p)\xi_d})^{1-\gamma}\big], & \text{when} \quad w_u \to +\infty; \\ \frac{(-w_u)^{1-\gamma}}{1-\gamma}\big[-Ap + (1-p)(\frac{p\xi_u}{(1-p)\xi_d})^{1-\gamma}\big], & \text{when} \quad w_u \to -\infty. \end{cases}$$

The EU has a local minimum between θ and $\hat{w}_u$.

1. If $p - A(1-p)(\frac{p\xi_u}{(1-p)\xi_d})^{1-\gamma} > 0$, when $w_u \to +\infty$, the EU approaches $+\infty$, and there is no global maximum.

2. If $p - A(1-p)(\frac{p\xi_u}{(1-p)\xi_d})^{1-\gamma} < 0$, when $w_u \to +\infty$, the EU approaches $-\infty$, the EU has a local maximum for $w_u \in (\theta, \infty)$ (and thus $w_d < \theta$).

   The FOC for the local maximum is

$$(w_u - \theta)^{-\gamma} = \lambda \xi_u, \qquad A(\theta - w_d)^{-\gamma} = \lambda \xi_d.$$

   So

$$w_u - \theta = \frac{(1/\xi_u)^{\frac{1}{\gamma}}}{(A/\xi_d)^{\frac{1}{\gamma}}}(\theta - w_d).$$

   The budget constraint leads to

$$r^{-1}\theta - w_0 = -p\xi_u(w_u - \theta) + (1-p)\xi_d(\theta - w_d) = \left(- p\xi_u \frac{(1/\xi_u)^{\frac{1}{\gamma}}}{(A/\xi_d)^{\frac{1}{\gamma}}} + (1-p)\xi_d \right)(\theta - w_d).$$

   The quasi-value function is

$$\begin{aligned} J &= \frac{1}{1-\gamma}\left[p\left(\frac{(1/\xi_u)^{\frac{1}{\gamma}}}{(A/\xi_d)^{\frac{1}{\gamma}}}\right)^{1-\gamma} - A(1-p)\right](\theta - w_d)^{1-\gamma} \\ &= \frac{(r^{-1}\theta - w_0)^{1-\gamma}}{1-\gamma}\left[p\left(\frac{(1/\xi_u)^{\frac{1}{\gamma}}}{(A/\xi_d)^{\frac{1}{\gamma}}}\right)^{1-\gamma} - A(1-p)\right]\left[(1-p)\xi_d - p\xi_u \frac{(1/\xi_u)^{\frac{1}{\gamma}}}{(A/\xi_d)^{\frac{1}{\gamma}}}\right]^{\gamma-1} \\ &= -\frac{(r^{-1}\theta - w_0)^{1-\gamma}}{1-\gamma}\left((1-p)A^{\frac{1}{\gamma}}\xi_d^{1-\frac{1}{\gamma}} - p\xi_u^{1-\frac{1}{\gamma}}\right)^{\gamma}. \end{aligned}$$

3. If $-Ap + (1-p)(\frac{p\xi_u}{(1-p)\xi_d})^{1-\gamma} > 0$, when $w_u \to -\infty$, the EU approaches $+\infty$, and there is no global maximum.

4. If $-Ap + (1-p)(\frac{p\xi_u}{(1-p)\xi_d})^{1-\gamma} < 0$, when $w_u \to -\infty$, the EU approaches $-\infty$, and there is a local maximum in the interval $[-\infty, \hat{w}_u]$. This case is symmetric to case (2).

5. If $p - A(1-p)(\frac{p\xi_u}{(1-p)\xi_d})^{1-\gamma} = 0$, when $w_u \to +\infty$, we have $EU \to 0$, and $\frac{\partial EU}{\partial w_1} > 0$ over $w_u \in [\theta, +\infty)$; thus, there is no local maximum over $w_u \in [\theta, +\infty)$.

6. If $-Ap + (1-p)(\frac{p\xi_u}{(1-p)\xi_d})^{1-\gamma} = 0$, when $w_u \to -\infty$, we have $EU \to 0$, and $\frac{\partial EU}{\partial w_1} < 0$ over $w_u \in (-\infty, \hat{w}_u]$; thus, there is no local maximum over $w_u \in (-\infty, \hat{w}_u]$.

Finally, if $p - A(1-p)(\frac{p\xi_u}{(1-p)\xi_d})^{1-\gamma} < 0$ and $-Ap + (1-p)(\frac{p\xi_u}{(1-p)\xi_d})^{1-\gamma} < 0$, there are two local maximum, one in $[\theta, +\infty)$ and one in $(-\infty, \hat{w}_u]$, and one of them is the global maximum.

A.5 Proof of Corollary 2

When $w_0 r < \theta$ and $A > \underline{A}$, the EU has two local maximum, and the global maximum is the greater of EU_+ and EU_-. Proposition 3 shows that $J_- < J_+$ if and only if $(1-p)\xi_d^{1-\frac{1}{\gamma}} < p\xi_u^{1-\frac{1}{\gamma}}$.

A.6 Proof of Proposition 4

The expected utility (1) becomes

$$p\Big[(w_u-\theta)^{1-\gamma}1_{\{w_u\geq\theta\}}-A(\theta-w_u)^{1-\gamma}1_{\{w_u<\theta\}}\Big]+(1-p)\Big[(w_d-\theta)^{1-\gamma}1_{\{w_d\geq\theta\}}-A(\theta-w_d)^{1-\gamma}1_{\{w_d<\theta\}}\Big],$$

where w_u and w_d are the terminal wealth at states u and d, respectively, given by

$$w_u = w_0 r + x(u-r), \qquad w_d = w_0 r + x(d-r). \tag{A.2}$$

First consider the case when the optimal wealth is positive in both states. The FOC is

$$p[w_0 r - \theta + x(u-r)]^{-\gamma}(u-r) = -(1-p)[w_0 r - \theta + x(d-r)]^{-\gamma}(d-r),$$

Let $k = [\frac{p(u-r)}{(1-p)(r-d)}]^{\frac{1}{\gamma}}$. The FOC becomes $w_0 r - \theta + x(u-r) = k[w_0 r - \theta + x(d-r)]$. The optimal holding x is

$$x^* = \frac{(k-1)(w_0 r - \theta)}{(u-r)+k(r-d)}.$$

Next we consider the case when the optimal wealth is negative in one of the states. The FOC is

$$\begin{cases} p[w_0 r-\theta+x_+(u-r)]^{-\gamma}(u-r) = -(1-p)A[\theta - w_0 r - x_+(d-r)]^{-\gamma}(d-r), & \text{if} \quad w_0 r + x_+(d-r) < \theta; \\ pA[\theta - w_0 r - x_-(u-r)]^{-\gamma}(u-r) = -(1-p)[w_0 r - \theta + x_-(d-r)]^{-\gamma}(d-r), & \text{if} \quad w_0 r + x_-(u-r) < \theta. \end{cases}$$

Let $k_+ = [\frac{p(u-r)}{A(1-p)(r-d)}]^{\frac{1}{\gamma}}$ and $k_- = [\frac{Ap(u-r)}{(1-p)(r-d)}]^{\frac{1}{\gamma}}$. The FOC becomes

$$\begin{cases} w_0 r - \theta + x_+(u-r) = -k_d[w_0 r - \theta + x_+(d-r)], & \text{if} \quad w_0 r + x_+(d-r) < \theta; \\ -[w_0 r - \theta + x_-(u-r)] = k_u[w_0 r - \theta + x_-(d-r)], & \text{if} \quad w_0 r + x_-(u-r) < \theta. \end{cases}$$

The optimal stock holdings x are

$$x_\pm = \frac{(k_\pm + 1)(w_0 r - \theta)}{(u-r) - k_\pm(r-d)}.$$

A.7 Proof of Proposition 5

The proof follows from Appendices A.3–A.4.

A.8 Proof of Proposition 6

When w_u is sufficiently large, the expected utility becomes

$$EU = p\frac{(w_u - \theta)^{1-\gamma_+}}{1-\gamma_+} - (1-p)\frac{A}{1-\gamma_-}\left[\theta - \frac{w_0 - p\xi_u w_u}{(1-p)\xi_d}\right]^{1-\gamma_-}.$$

When $w_u \to +\infty$,

$$EU \approx p\frac{(w_u)^{1-\gamma_+}}{1-\gamma_+} - (1-p)A\frac{(w_u)^{1-\gamma_-}}{1-\gamma_-}\left[\frac{p\xi_u}{(1-p)\xi_d}\right]^{1-\gamma_-}.$$

It goes to $+\infty$ if $\gamma_+ < \gamma_-$ and goes to $-\infty$ if $\gamma_- < \gamma_+$.

When w_u is sufficiently low, the expected utility becomes

$$EU = -pA\frac{(\theta - w_u)^{1-\gamma_-}}{1-\gamma_-} + (1-p)\frac{1}{1-\gamma_+}\left[\frac{w_0 - p\xi_u w_u}{(1-p)\xi_d} - \theta\right]^{1-\gamma_+}.$$

When $w_u \to -\infty$,

$$EU \approx -pA\frac{(-w_u)^{1-\gamma_-}}{1-\gamma_-} + (1-p)\frac{(-w_u)^{1-\gamma_+}}{1-\gamma_+}\left[\frac{p\xi_u}{(1-p)\xi_d}\right]^{1-\gamma_+}.$$

It also goes to $+\infty$ if $\gamma_+ < \gamma_-$ and goes to $-\infty$ if $\gamma_- < \gamma_+$.

A.9 Proof of Proposition 7

The EU as a function of w_u is given by

$$EU = \begin{cases} \frac{1}{1-\gamma_+}\left[p(w_u-\theta)^{1-\gamma_+} + (1-p)(\frac{w_0 - p\xi_u w_u}{(1-p)\xi_d} - \theta)^{1-\gamma_+}\right], & \text{if} \quad w_u \geq \theta \text{ and } w_u \leq \hat{w}_u; \\ \frac{1}{1-\gamma_+}p(w_u-\theta)^{1-\gamma_+} - A\frac{1}{1-\gamma_-}(1-p)(\theta - \frac{w_0 - p\xi_u w_u}{(1-p)\xi_d})^{1-\gamma_-}, & \text{if} \quad w_u \geq \theta \text{ and } w_u \geq \hat{w}_u; \\ -A\frac{1}{1-\gamma_-}p(\theta - w_u)^{1-\gamma_-} + \frac{1}{1-\gamma_+}(1-p)(\frac{w_0 - p\xi_u w_u}{(1-p)\xi_d} - \theta)^{1-\gamma_+}, & \text{if} \quad w_u \leq \theta \text{ and } w_u \leq \hat{w}_u; \\ -\frac{A}{1-\gamma_-}\left[p(\theta - w_u)^{1-\gamma_-} + (1-p)(\theta - \frac{w_0 - p\xi_u w_u}{(1-p)\xi_d})^{1-\gamma_-}\right], & \text{if} \quad w_u \leq \theta \text{ and } w_u \geq \hat{w}_u, \end{cases}$$

where $\hat{w}_u \equiv \frac{w_0 - (1-p)\xi_d\theta}{p\xi_u}$.

A.9.1 $w_0 \geq r^{-1}\theta$

Note that $r^{-1} = \mathbb{E}[\xi] = p\xi_u + (1-p)\xi_d$. When $w_0 \geq r^{-1}\theta$, we have $\hat{w}_u \geq \theta$. Thus, the case $w_u < \theta$ and $w_u > \hat{w}_u$ cannot occur. The EU becomes

$$EU = \begin{cases} \frac{1}{1-\gamma_+}\left[p(w_u-\theta)^{1-\gamma_+} + (1-p)(\frac{w_0-p\xi_u w_u}{(1-p)\xi_d} - \theta)^{1-\gamma_+}\right], & \text{if} \quad w_u \in [\theta, \hat{w}_u]; \\ \frac{1}{1-\gamma_+}p(w_u-\theta)^{1-\gamma_+} - A\frac{1}{1-\gamma_-}(1-p)(\theta - \frac{w_0-p\xi_u w_u}{(1-p)\xi_d})^{1-\gamma_-}, & \text{if} \quad w_u \in [\hat{w}_u, +\infty); \\ -A\frac{1}{1-\gamma_-}p(\theta-w_u)^{1-\gamma_-} + \frac{1}{1-\gamma_+}(1-p)(\frac{w_0-p\xi_u w_u}{(1-p)\xi_d} - \theta)^{1-\gamma_+}, & \text{if} \quad w_u \in (-\infty, \theta]. \end{cases}$$

The EU is concave when $w_u \in [\theta, \hat{w}_u]$; thus, it has a local maximum between θ and $\hat{w}_u$. Now, we derive the wealth at this local maximum. We use the Lagrangian approach to find local maximum. The Lagrangian is given by

$$\mathcal{L} = p\frac{(w_u-\theta)^{1-\gamma_+}}{1-\gamma_+} + (1-p)\frac{(w_d-\theta)^{1-\gamma_+}}{1-\gamma_+} - \lambda\left[p\xi_u w_u + (1-p)\xi_d w_d - w_0\right],$$

where $\lambda > 0$ is the Lagrange multiplier. The FOC for the local maximum is

$$(w_u-\theta)^{-\gamma_+} = \lambda\xi_u, \qquad (w_d-\theta)^{-\gamma_+} = \lambda\xi_d.$$

So

$$w_u - \theta = (\xi_d/\xi_u)^{\frac{1}{\gamma_+}}(w_d-\theta).$$

Together with the budget constraint, the wealth at this local maximum given by (22).

Next, we look at $w_u \in (-\infty, \theta]$. When $w_u \to -\infty$, $EU = -Ap\frac{(-w_u)^{1-\gamma_-}}{1-\gamma_-} \to -\infty$. In this interval,

$$\frac{\partial EU}{\partial w_u} = p\Big[A(\theta-w_u)^{-\gamma_-} - \frac{p^{-\gamma_+}}{(1-p)^{-\gamma_+}}\frac{\xi_u^{1-\gamma_+}}{\xi_d^{1-\gamma_+}}(\hat{w}_u-w_u)^{-\gamma_+}\Big].$$

To have $\frac{\partial EU}{\partial w_u} > 0$ (i.e., the EU is increasing over $w_u \in (-\infty, \theta]$), we need to demonstrate $A > \underline{A}_+\frac{(\theta-w_u)^{\gamma_-}}{(\hat{w}_u-w_u)^{\gamma_+}}$, where $\underline{A}_+ \equiv \frac{p^{-\gamma_+}}{(1-p)^{-\gamma_+}}\frac{\xi_u^{1-\gamma_+}}{\xi_d^{1-\gamma_+}}$. In fact,

$$\begin{aligned} &\underline{A}_+\frac{(\theta-w_u)^{\gamma_-}}{(\hat{w}_u-w_u)^{\gamma_+}} = \underline{A}_+\Big(\frac{\theta-w_u}{\hat{w}_u-w_u}\Big)^{\gamma_-}(\hat{w}_u-w_u)^{\gamma_--\gamma_+} \\ <&\underline{A}_+ \times 1 \times (\hat{w}_u-\theta)^{\gamma_--\gamma_+} = \frac{p^{-\gamma_-}}{(1-p)^{-\gamma_+}}\frac{\xi_u^{1-\gamma_-}}{\xi_d^{1-\gamma_+}}(w_0-r^{-1}\theta)^{\gamma_--\gamma_+} \leq A, \end{aligned}$$

where the last inequality holds true if $w_0 - r^{-1}\theta \geq 1$ and $A \geq \frac{p^{-\gamma_-}}{(1-p)^{-\gamma_+}}\frac{\xi_u^{1-\gamma_-}}{\xi_d^{1-\gamma_+}}$.

Finally, we consider $w_u \in [\hat{w}_u, +\infty)$. When $w_u \to \infty$, $EU =\to -\infty$. In this interval,

$$\frac{\partial EU}{\partial w_u} = p\Big[(w_u-\theta)^{-\gamma_+} - A\frac{p^{-\gamma_-}}{(1-p)^{-\gamma_-}}\frac{\xi_u^{1-\gamma_-}}{\xi_d^{1-\gamma_-}}(w_u-\hat{w}_u)^{-\gamma_-}\Big].$$

To have $\frac{\partial EU}{\partial w_u} < 0$, we need to demonstrate $A > \underline{A}_- \frac{(w_u-\theta)^{\gamma_+}}{(w_u-\hat{w}_u)^{\gamma_-}}$, where $\underline{A}_- \equiv \frac{(1-p)^{-\gamma_-}}{p^{-\gamma_-}} \frac{\xi_d^{1-\gamma_-}}{\xi_u^{1-\gamma_-}}$. In fact,

$$\begin{aligned}
&\underline{A}_- \frac{(w_u-\theta)^{\gamma_+}}{(w_u-\hat{w}_u)^{\gamma_-}} = \underline{A}_- \left(\frac{w_u-\hat{w}_u}{w_u-\theta}\right)^{\gamma_-} (w_u-\theta)^{\gamma_- - \gamma_+} \\
<&\underline{A}_- \times 1 \times (\hat{w}_u-\theta)^{\gamma_- - \gamma_+} = \frac{(1-p)^{-\gamma_-}}{p^{-\gamma_+}} \frac{\xi_d^{1-\gamma_-}}{\xi_u^{1-\gamma_+}} (w_0 - r^{-1}\theta)^{\gamma_- - \gamma_+} \leq A,
\end{aligned}$$

where the last inequality holds true if $w_0 - r^{-1}\theta \geq 1$ and $A \geq \frac{(1-p)^{-\gamma_-}}{p^{-\gamma_+}} \frac{\xi_d^{1-\gamma_-}}{\xi_u^{1-\gamma_+}}$.

Therefore, when $w_0 - r^{-1}\theta \geq 1$, $A \geq \frac{p^{-\gamma_-}}{(1-p)^{-\gamma_+}} \frac{\xi_u^{1-\gamma_-}}{\xi_d^{1-\gamma_+}}$, and $A \geq \frac{(1-p)^{-\gamma_-}}{p^{-\gamma_+}} \frac{\xi_d^{1-\gamma_-}}{\xi_u^{1-\gamma_+}}$, the local maximum over $w_u \in [\theta, \hat{w}_u]$ is also the global maximum.

It is worth noting that $\frac{\partial EU}{\partial w_u}$ can be negative in both $w_u \in (-\infty, \theta]$ and $w_u \in [\hat{w}_u, +\infty)$ if A is sufficiently small or if $w_0 - r^{-1}\theta$ is close to zero. Especially, when $w_0 = r^{-1}\theta$, $\frac{\partial EU}{\partial w_u}$ is negative and the EU is decreasing in a left neighborhood of $w_u = \theta$, and $\frac{\partial EU}{\partial w_u}$ is positive and the EU is increasing in a right neighborhood of $w_u = \hat{w}_u$.

A.9.2 $w_0 < r^{-1}\theta$

In this case, $\hat{w}_u < \theta$; thus, the case $w_u > \theta$ and $w_u < \hat{w}_u$ cannot occur. So the EU is

$$EU = \begin{cases}
\frac{1}{1-\gamma_+} p(w_u-\theta)^{1-\gamma_+} - A\frac{1}{1-\gamma_-}(1-p)(\theta - \frac{w_0 - p\xi_u w_u}{(1-p)\xi_d})^{1-\gamma_-}, & \text{if} \quad w_u \in [\theta, +\infty); \\
-A\frac{1}{1-\gamma_-} p(\theta - w_u)^{1-\gamma_-} + \frac{1}{1-\gamma_+}(1-p)(\frac{w_0 - p\xi_u w_u}{(1-p)\xi_d} - \theta)^{1-\gamma_+}, & \text{if} \quad w_u \in (-\infty, \hat{w}_u]; \\
-\frac{A}{1-\gamma_-}\left[p(\theta - w_u)^{1-\gamma_-} + (1-p)(\theta - \frac{w_0 - p\xi_u w_u}{(1-p)\xi_d})^{1-\gamma_-}\right], & \text{if} \quad w_u \in [\hat{w}_u, \theta].
\end{cases}$$

The utility function is convex when $w_u \in [\hat{w}_u, \theta]$, and $EU \to -\infty$ as $w_u \to \pm\infty$.

The EU has a local maximum over $w_u \in (-\infty, \hat{w}_u]$. The FOC for the local maximum is

$$A(\theta - w_u)^{-\gamma_-} = \lambda \xi_u, \qquad (w_d - \theta)^{-\gamma_+} = \lambda \xi_d.$$

Together with the budget constraint, we obtain the following equation governing w_u:

$$(\theta - w_u)^{-\gamma_-} = \frac{1}{A} \frac{p^{-\gamma_+}}{(1-p)^{-\gamma_+}} \frac{\xi_u^{1-\gamma_+}}{\xi_d^{1-\gamma_+}} (\hat{w}_u - w_u)^{-\gamma_+}. \tag{A.3}$$

Then w_d is solved for using the budget constraint.

When $w_u \in (-\infty, \hat{w}_u]$, the FOC for the local maximum leads to the following equation governing w_u:

$$(w_u - \theta)^{-\gamma_+} = A \frac{p^{-\gamma_-}}{(1-p)^{-\gamma_-}} \frac{\xi_u^{1-\gamma_-}}{\xi_d^{1-\gamma_-}} (w_u - \hat{w}_u)^{-\gamma_-}. \tag{A.4}$$

A.10 Proof of Lemma 3

The budget constraint is given by $\mathbb{E}[\xi\hat{w}] = w_0$, where w is a wealth distribution across the S states. Given a starting wealth distribution $\hat{w}$ that satisfies the budget constraint, any wealth w under the budget constraint should satisfy the following budget-preserving change: $w - \hat{w} = \xi^{-1/\gamma}(1_+ - a1_-)\Delta$, with $a = \mathbb{E}[\xi^{1-1/\gamma}1_+]/\mathbb{E}[\xi^{1-1/\gamma}1_-]$. The EU is given by

$$\begin{aligned} EU =& \frac{1}{1-\gamma}\mathbb{E}\big[(w-\theta)^{1-\gamma}1_+ - A(\theta - w)^{1-\gamma}1_-\big] \\ =& \frac{1}{1-\gamma}\mathbb{E}\Big[\big(\hat{w} + \xi^{-1/\gamma}\Delta - \theta\big)^{1-\gamma}1_+ - A\big(\theta - \hat{w} + a\xi^{-1/\gamma}\Delta\big)^{1-\gamma}1_-\Big]. \end{aligned}$$

When $\Delta \to +\infty$,

$$\begin{aligned} EU \approx& \frac{\Delta^{1-\gamma}}{1-\gamma}\big(\mathbb{E}[\xi^{1-1/\gamma}1_+] - Aa^{1-\gamma}\mathbb{E}[\xi^{1-1/\gamma}1_-]\big) \\ =& \frac{\Delta^{1-\gamma}}{1-\gamma}\big(\mathbb{E}[\xi^{1-1/\gamma}1_+]\big)^{1-\gamma}\big[(\mathbb{E}[\xi^{1-1/\gamma}1_+])^{\gamma} - A(\mathbb{E}[\xi^{1-1/\gamma}1_-])^{\gamma}\big]. \end{aligned}$$

This goes to $+\infty$ if $\mathbb{E}[\xi^{1-1/\gamma}(1_+ - A^{1/\gamma}1_-)] > 0$.

A.11 Proof of Proposition 9

The objective function is given by

$$\max_{\{w_s\}} \mathbb{E}[U(w)] = \max_{\{w_s\}} \sum_s p_s U(w_s),$$

such that $w_0 = \sum_s p_s \xi_s w_s = \mathbb{E}[\xi w]$. We will consider interior solutions for the case. To this end, we use Lagrangian approach. The Lagrangian is given by

$$\mathbb{E}[U(w)] - \lambda\big(\mathbb{E}[\xi w] - w_0\big), \tag{A.5}$$

where $\lambda \geq 0$ is the Lagrange multiplier.

When $w_0 \geq \mathbb{E}[\xi]\theta$, the FOC is given by $(w-\theta)^{-\gamma} = \lambda\xi$. The budget constraint leads to $w_0 - \mathbb{E}[\xi]\theta = \mathbb{E}[\xi(w-\theta)] = \mathbb{E}\big[\xi(\lambda\xi)^{-1/\gamma}\big] = \lambda^{-1/\gamma}\mathbb{E}\big[\xi^{1-1/\gamma})\big]$. Therefore, the Lagrange multiplier is given by $\lambda = \Big(\frac{w_0 - \mathbb{E}[\xi]\theta}{\mathbb{E}[\xi^{1-1/\gamma}]}\Big)^{-\gamma}$, leading to the optimal wealth and the value function given by (28) and (29).

When $w_0 < \mathbb{E}[\xi]\theta$, assume a local optimum occurs at one partition S_+ and S_-. The FOC is given by

$$(w-\theta)^{-\gamma}1_+ = \lambda\xi 1_+, \qquad A(\theta - w)^{-\gamma}1_- = \lambda\xi 1_-.$$

The budget constraint leads to

$$w_0-\mathbb{E}[\xi]\theta = \mathbb{E}[\xi(w-\theta)] = \mathbb{E}\left[\xi\Big((\lambda\xi)^{-1/\gamma}1_+ - (\lambda\xi/A)^{-1/\gamma}1_-\Big)\right] = \lambda^{-1/\gamma}\mathbb{E}\big[\xi^{1-1/\gamma}(1_+-A^{1/\gamma}1_-)\big].$$

Therefore, the Lagrange multiplier is given by

$$\lambda = (\mathbb{E}[\xi]\theta - w_0)^{-\gamma}\left(\mathbb{E}[\xi^{1-1/\gamma}(-1_+ + A^{1/\gamma}1_-)]\right)^{\gamma}. \tag{A.6}$$

Note that $\lambda > 0$, so $w_0 - \mathbb{E}[\xi]\theta > 0$ and $\mathbb{E}[\xi^{1-1/\gamma}(1_+ - A^{1/\gamma}1_-)] > 0$ has the same sign.

The expected utility function EU is given by

$$\begin{aligned} EU &= \frac{1}{1-\gamma}\mathbb{E}\Big[\big((\lambda\xi)^{-1/\gamma}\big)^{1-\gamma}1_+ - A\big((\lambda\xi/A)^{-1/\gamma}\big)^{1-\gamma}1_-\Big] \\ &= \frac{\lambda^{1-1/\gamma}}{1-\gamma}\mathbb{E}[\xi^{1-1/\gamma}(1_+ - A^{1/\gamma}1_-)]. \end{aligned} \tag{A.7}$$

By substituting (A.6) into (A.7), the EU becomes (32).

The global maximum is achieved only for partition with only 1 state in $\mathbb{S}_-$. Now consider partitions with only 1 state in S_-. Suppose $\mathbb{S}_- = \{t\}$, that is, state t is the only state with wealth lower than θ. In this case, $\mathbb{E}[\xi^{1-1/\gamma}1_-] = \xi_t^{1-1/\gamma}p_t$, and we can explicitly solve the budget constraint:

$$w_0 - \mathbb{E}[\xi]\theta = \sum_{s\neq t} p_s\xi_s(w_s-\theta) + p_t\xi_t(w_t-\theta).$$

Therefore,

$$\theta - w_t = \frac{\sum_{s\neq t} p_s\xi_s(w_s-\theta)}{p_t\xi_t} - (w_0 - \mathbb{E}[\xi]\theta). \tag{A.8}$$

The EU can be express in unconstrained variables

$$EU = \sum_{s\neq t} p_s\frac{(w_s-\theta)^{1-\gamma}}{1-\gamma} - Ap_t\frac{(\theta-w_t)^{1-\gamma}}{1-\gamma},$$

where $\theta - w_t$ is given by equation (A.8).

FOC for w_s^*, $s\neq t$, is

$$p_s(w_s^*-\theta)^{-\gamma} = Ap_t(\theta-w_t^*)^{-\gamma}\frac{p_s\xi_s}{p_t\xi_t}.$$

The Hessian matrix is

$$H_{s,s'} = -\gamma p_s(w_s^*-\theta)^{-\gamma-1}\delta_{s,s'} + Ap_t(\theta-w_t^*)^{-\gamma-1}\frac{p_s\xi_s p_{s'}\xi_{s'}}{(p_t\xi_t)^2}.$$

Using the FOC condition, the Hessian can be written as

$$H_{s,s'} = -\gamma Ap_t(\theta-w_t^*)^{-\gamma}\frac{p_s\xi_s}{p_t\xi_t}(w_s^*-\theta)^{-1}\delta_{s,s'} + \gamma Ap_t(\theta-w_t^*)^{-\gamma-1}\frac{p_s\xi_s p_{s'}\xi_{s'}}{(p_t\xi_t)^2}$$

$$= \gamma \frac{Ap_t}{\xi_t}(\theta - w_t^*)^{-\gamma} \frac{1}{p_t \xi_t} \left(-\frac{p_s \xi_s}{w_s^* - \theta} \delta_{s,s'} + \frac{p_s \xi_s p_{s'} \xi_{s'}}{p_t \xi_t (\theta - w_t^*)} \right).$$

We should be able to show that H is negative definite. Define $\frac{d\mathbb{Q}}{d\mathbb{P}} = \frac{\xi}{\mathbb{E}[\xi]}$, thus $q_s = p_s \xi_s / \sum_s (p_s \xi_s)$, which are the risk-neutral probabilities. The Hessian can be written as

$$H_{s,s'} = \gamma \frac{Ap_t}{\xi_t q_t}(\theta - w_t^*)^{-\gamma} \left(-\frac{q_s}{w_s^* - \theta} \delta_{s,s'} + \frac{q_s q_{s'}}{q_t (\theta - w_t^*)} \right)$$

$$= \gamma \frac{Ap_t}{\xi_t q_t^2}(\theta - w_t^*)^{-\gamma - 1} \left(-\frac{\theta - w_t^*}{w_s^* - \theta} q_s q_t \delta_{s,s'} + q_s q_{s'} \right).$$

Only terms in the bracket are relevant for determining whether H is negative-definite,

$$h_{s,s'} = \left(-\frac{\theta - w_t^*}{w_s^* - \theta} q_s q_t \delta_{s,s'} + q_s q_{s'} \right) = \left(-a_s q_s q_t \delta_{s,s'} + q_s q_{s'} \right),$$

where $a_s = \frac{\theta - w_t^*}{w_s^* - \theta} q_s q_t \delta_{s,s'}$. Thus,

$$h = \begin{pmatrix} -a_1 q_1 q_t + q_1 q_1 & q_1 q_2 & \dots & q_1 q_S \\ q_1 q_2 & -a_2 q_2 q_t + q_2 q_2 & \dots & q_2 q_S \\ \vdots & \vdots & \vdots & \vdots \\ q_1 q_S & q_2 q_S & \dots & -a_S q_S q_t + q_S q_S \end{pmatrix}$$

$$= \begin{pmatrix} -a_1 q_t + q_1 & q_1 & \dots & q_1 \\ q_2 & -a_2 q_t + q_2 & \dots & q_2 \\ \vdots & \vdots & \vdots & \vdots \\ q_S & q_S & \dots & -a_S q_t + q_S \end{pmatrix} \Pi_s q_s.$$

Subtracting column 1 from column s for all $s \neq 1$, we get

$$h' = \begin{pmatrix} -a_1 q_t + q_1 & a_1 q_t & \dots & a_1 q_t \\ q_2 & -a_2 q_t & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ q_S & 0 & \dots & -a_S q_t \end{pmatrix}$$

The determinant of h' is

$$(-a_1 q_t + q_1) \Pi_{s \neq 1}(-a_s q_t) - q_2 \Pi_{s \neq 2}(-a_s q_t) + \dots = \left(\Pi_s (-a_s q_t) \right) \left(1 - \sum_{s=1}^{S} \frac{q_s}{q_t a_s} \right).$$

If $w_0 - \mathbb{E}[\xi]\theta < 0$, then

$$\left(1 - \sum_{s=1}^{S} \frac{q_s}{q_t a_s} \right) = \frac{1}{q_t (\theta - w_t^*)} \left(q_t (\theta - w_t^*) - \sum_{s \neq t} q_s (w_s^* - \theta) \right) = \frac{-(\mathbb{E}[\xi]\theta - w_0)}{q_t (\theta - w_t^*)} > 0.$$

The above equation can be true for all principle minors of dimension u, $u = 1, ..., S-1$, so the sign of their determinant is $(-1)^u$. This implies that the eigenvalues of h are all negative; thus h is negative definite. Therefore, the Hessian is negative definite, and the EU has a local maximum at w^*.

A.12 Proof of Proposition 10

We first consider the case $m < r$. In this case,

$$\mathbb{Q}_1 = (q_{11}, q_{12}, 0) \equiv \left(\frac{r-m}{u-m}, \frac{u-r}{u-m}, 0\right), \quad \mathbb{Q}_2 = (q_{21}, 0, q_{23}) \equiv \left(\frac{r-d}{u-d}, 0, \frac{u-r}{u-d}\right).$$

The two constraints in problem (42) lead to

$$w_m = \frac{w_0 r - q_{11} w_u}{q_{12}}, \qquad w_d = \frac{w_0 r - q_{21} w_u}{q_{23}}. \tag{A.9}$$

When $w_u \to +\infty$, we have

$$\begin{aligned} EU &\approx p_1 \frac{1}{1-\gamma} w_u^{1-\gamma} - A p_2 \frac{1}{1-\gamma}\Big(\frac{q_{11}}{q_{12}} w_u\Big)^{1-\gamma} - A p_3 \frac{1}{1-\gamma}\Big(\frac{q_{21}}{q_{23}} w_u\Big)^{1-\gamma} \\ &= \Big[p_1 - A p_2 \Big(\frac{q_{11}}{q_{12}}\Big)^{1-\gamma} - A p_3 \Big(\frac{q_{21}}{q_{23}}\Big)^{1-\gamma}\Big] \frac{w_u^{1-\gamma}}{1-\gamma} \\ &= \Big[p_1 - A p_2 \Big(\frac{r-m}{u-r}\Big)^{1-\gamma} - A p_3 \Big(\frac{r-d}{u-r}\Big)^{1-\gamma}\Big] \frac{w_u^{1-\gamma}}{1-\gamma}. \end{aligned}$$

It leads to

$$EU \to \begin{cases} -\infty, & \text{if} \quad A > \underline{A}^*; \\ 0, & \text{if} \quad A = \underline{A}^*; \\ +\infty, & \text{if} \quad A < \underline{A}^*, \end{cases}$$

where

$$\underline{A}^* = \frac{p_1 (u-r)^{1-\gamma}}{p_2 (r-m)^{1-\gamma} + p_3 (r-d)^{1-\gamma}} = \frac{\mathbb{E}\big[[(R-r)^+]^{1-\gamma})\big]}{\mathbb{E}\big[[(r-R)^+]^{1-\gamma})\big]}.$$

When $w_u \to -\infty$, we have

$$\begin{aligned} EU &\approx -A p_1 \frac{1}{1-\gamma}(-w_u)^{1-\gamma} + p_2 \frac{1}{1-\gamma}\Big(-\frac{q_{11}}{q_{12}} w_u\Big)^{1-\gamma} + p_3 \frac{1}{1-\gamma}\Big(\frac{-q_{21}}{q_{23}} w_u\Big)^{1-\gamma} \\ &= \Big[-A p_1 + p_2 \Big(\frac{q_{11}}{q_{12}}\Big)^{1-\gamma} + p_3 \Big(\frac{q_{21}}{q_{23}}\Big)^{1-\gamma}\Big] \frac{(-w_u)^{1-\gamma}}{1-\gamma} \\ &= \Big[-A p_1 + p_2 \Big(\frac{r-m}{u-r}\Big)^{1-\gamma} + p_3 \Big(\frac{r-d}{u-r}\Big)^{1-\gamma}\Big] \frac{(-w_u)^{1-\gamma}}{1-\gamma}. \end{aligned}$$

It leads to

$$EU \to \begin{cases} -\infty, & \text{if} \quad A > (\underline{A}^*)^{-1}; \\ 0, & \text{if} \quad A = (\underline{A}^*)^{-1}; \\ +\infty, & \text{if} \quad A < (\underline{A}^*)^{-1}. \end{cases}$$

Therefore, the optimization problem (42) has bounded solutions when $A > \underline{A}^{incomplete}$ and does not have bounded solutions when $A < \underline{A}^{incomplete}$, where

$$\underline{A}^{incomplete} \equiv \max\left\{\frac{\mathbb{E}[[(R-r)^+]^{1-\gamma})]}{\mathbb{E}[[(r-R)^+]^{1-\gamma})]}, \frac{\mathbb{E}[[(r-R)^+]^{1-\gamma})]}{\mathbb{E}\big[[(R-r)^+]^{1-\gamma})\big]}\right\}.$$

Next, we consider the case $m \geq r$. In this case,

$$\mathbb{Q}_1 = (0, q_{12}, q_{13}) \equiv \left(0, \frac{r-d}{m-d}, \frac{m-r}{m-d}\right), \quad \mathbb{Q}_2 = (q_{21}, 0, q_{23}) \equiv \left(\frac{r-d}{u-d}, 0, \frac{u-r}{u-d}\right).$$

The two constraints in problem (42) lead to

$$w_u = \frac{w_0 r - q_{23} w_d}{q_{21}}, \qquad w_m = \frac{w_0 r - q_{13} w_d}{q_{12}}. \tag{A.10}$$

Following the same analysis above, one can show that $A > \underline{A}$ ($A < \underline{A}$) is also the condition under which $EU \to -\infty$ ($EU \to +\infty$) as $w_u \to \pm\infty$.

A.13 Proof of Proposition 11

A.13.1 $m < r$

We first consider the case $m < r$. Define

$$\hat{w}_u = \frac{w_0 r - q_{12}\theta}{q_{11}}, \qquad \tilde{w}_u = \frac{w_0 r - q_{23}\theta}{q_{21}}, \tag{A.11}$$

which is the value of w_u when $w_m = \theta$ and $w_d = \theta$, respectively.

We present our results for $w_0 \geq r^{-1}\theta$ and $w_0 < r^{-1}\theta$, respectively. First, when $w_0 \geq r^{-1}\theta$, we have $\theta \leq \tilde{w}_u \leq \hat{w}_u$, which leads to four intervals for w_u, corresponding to four different cases as discussed below.

Case 1: $w_u < \theta$, $w_m > \theta$, $w_d > \theta$.

The expected utility is given by

$$EU = -p_1 A \frac{1}{1-\gamma}(\theta - w_u)^{1-\gamma} + p_2 \frac{1}{1-\gamma}\Big(\frac{q_{11}}{q_{12}}\Big)^{1-\gamma}(\hat{w}_u - w_u)^{1-\gamma} + p_3 \frac{1}{1-\gamma}\Big(\frac{q_{21}}{q_{23}}\Big)^{1-\gamma}(\tilde{w}_u - w_u)^{1-\gamma}.$$

Under the condition $A > \underline{A}^{incomplete}$, $EU \to -\infty$ as $w_u \to \pm\infty$. In addition,

$$\begin{aligned}\frac{\partial EU}{\partial w_u} &= p_1 A(\theta - w_u)^{-\gamma} - p_2\Big(\frac{q_{11}}{q_{12}}\Big)^{1-\gamma}(\hat{w}_u - w_u)^{-\gamma} - p_3\Big(\frac{q_{21}}{q_{23}}\Big)^{1-\gamma}(\tilde{w}_u - w_u)^{-\gamma} \\ &> \Big[Ap_1 - p_2\Big(\frac{q_{11}}{q_{12}}\Big)^{1-\gamma} - p_3\Big(\frac{q_{21}}{q_{23}}\Big)^{1-\gamma}\Big](\theta - w_u)^{-\gamma} \geq 0,\end{aligned}$$

showing that EU is increasing over the interval $w_u \in (-\infty, \theta)$.

Case 2: $\theta \leq w_u < \tilde{w}_u$, $w_m \geq \theta$, $w_d \geq \theta$.

In this case, wealth in the three states is in the gain domain, and EU is concave, $\frac{\partial^2 EU}{\partial w_u^2} < 0$, and has a local maximum over this interval.

Case 3: $\tilde{w}_u \leq w_u < \hat{w}_u$, $w_m > \theta$, $w_d \leq \theta$.

The expected utility is given by $EU = p_1\frac{1}{1-\gamma}(w_u - \theta)^{1-\gamma} + p_2\frac{1}{1-\gamma}(\frac{q_{11}}{q_{12}})^{1-\gamma}(\hat{w}_u - w_u)^{1-\gamma} - Ap_3\frac{1}{1-\gamma}(\frac{q_{21}}{q_{23}})^{1-\gamma}(w_u - \tilde{w}_u)^{1-\gamma}$. It leads to

$$\begin{aligned}\frac{\partial EU}{\partial w_u} &= p_1(w_u - \theta)^{-\gamma} - p_2\Big(\frac{q_{11}}{q_{12}}\Big)^{1-\gamma}(\hat{w}_u - w_u)^{-\gamma} - Ap_3\Big(\frac{q_{21}}{q_{23}}\Big)^{1-\gamma}(w_u - \tilde{w}_u)^{-\gamma} \\ &< p_1(w_u - \theta)^{-\gamma} - Ap_3\Big(\frac{q_{21}}{q_{23}}\Big)^{1-\gamma}(w_u - \tilde{w}_u)^{-\gamma} \\ &< \Big[p_1 - Ap_3\Big(\frac{q_{21}}{q_{23}}\Big)^{1-\gamma}\Big](w_u - \theta)^{-\gamma} \leq 0,\end{aligned}$$

where the last equality is based on the assumption $A \geq \frac{p_1}{p_3}(\frac{q_{23}}{q_{21}})^{1-\gamma} = \frac{p_1}{p_3}(\frac{u-r}{r-d})^{1-\gamma}$. Then EU is decreasing over the interval $w_u \in [\theta, \tilde{w}_u)$.

Case 4: $\hat{w}_u \leq w_u$, $w_m \leq \theta$, $w_d < \theta$.

The expected utility is given by $EU = p_1\frac{1}{1-\gamma}(w_u - \theta)^{1-\gamma} - Ap_2\frac{1}{1-\gamma}(\frac{q_{11}}{q_{12}})^{1-\gamma}(w_u - \hat{w}_u)^{1-\gamma} - Ap_3\frac{1}{1-\gamma}(\frac{q_{21}}{q_{23}})^{1-\gamma}(w_u - \tilde{w}_u)^{1-\gamma}$. It leads to

$$\begin{aligned}\frac{\partial EU}{\partial w_u} &= p_1(w_u - \theta)^{-\gamma} - Ap_2\Big(\frac{q_{11}}{q_{12}}\Big)^{1-\gamma}(w_u - \hat{w}_u)^{-\gamma} - Ap_3\Big(\frac{q_{21}}{q_{23}}\Big)^{1-\gamma}(w_u - \tilde{w}_u)^{-\gamma} \\ &< \Big[p_1 - Ap_2\Big(\frac{q_{11}}{q_{12}}\Big)^{1-\gamma} - Ap_3\Big(\frac{q_{21}}{q_{23}}\Big)^{1-\gamma}\Big](w_u - \hat{w}_u)^{-\gamma} \leq 0,\end{aligned}$$

showing that EU is decreasing over the interval $w_u \in [\tilde{w}_u, +\infty)$.

Therefore, the local maximum of EU over $\theta \leq w_u < \tilde{w}_u$ in Case 2 is the global maximum. Now, we derive the optimal wealth and value function. We use the Lagrangian approach to find this local maximum. The Lagrangian is given by

$$\begin{aligned}\mathcal{L} &= \mathbb{E}[U(w)] - \lambda_1(\mathbb{E}^{\mathbb{Q}_1}[w] - w_0 r) - \lambda_2(\mathbb{E}^{\mathbb{Q}_2}[w] - w_0 r) \\ &= p_1\frac{(w_u - \theta)^{1-\gamma}}{1-\gamma} + p_2\frac{(w_m - \theta)^{1-\gamma}}{1-\gamma} + p_3\frac{(w_d - \theta)^{1-\gamma}}{1-\gamma} \\ &\quad - \lambda_1(q_{11}w_u + q_{12}w_m - w_0 r) - \lambda_2(q_{21}w_u + q_{23}w_d - w_0 r).\end{aligned}$$

The FOC for the local maximum is

$$p_1(w_u-\theta)^{-\gamma}=\lambda_1 q_{11}+\lambda_2 q_{21}, \qquad p_2(w_m-\theta)^{-\gamma}=\lambda_1 q_{12}, \qquad p_3(w_d-\theta)^{-\gamma}=\lambda_2 q_{23},$$

which, together with the budget constraint, leads to the following equation governing the optimal wealth w_u^*:

$$p_1(u-r)^{1-\gamma}(w_u-\theta)^{-\gamma}-p_2(r-m)^{1-\gamma}(\hat{w}_u-w_u)^{-\gamma}-p_3(r-d)^{1-\gamma}(\tilde{w}_u-w_u)^{-\gamma}=0. \quad \text{(A.12)}$$

Then we can solve for w_m^* and w_d^* from (A.9). The value function is $J=p_1\frac{(w_u-\theta)^{1-\gamma}}{1-\gamma}+p_2\frac{(w_m-\theta)^{1-\gamma}}{1-\gamma}+p_3\frac{(w_d-\theta)^{1-\gamma}}{1-\gamma}\geq 0$.

In particular, when $w_0=r^{-1}\theta$, we have $\hat{w}_u=\tilde{w}_u=\theta$, $w_u^*=w_m^*=w_d^*=\theta$, and $Eu=0$.

Finally, we consider the case $w_0<r^{-1}\theta$. In this case, we have $\hat{w}_u<\tilde{w}_u<\theta$, which leads to four intervals for w_u, corresponding to four different cases as discussed below.

Case 1: $w_u<\hat{w}_u$, $w_m>\theta$, $w_d>\theta$.

In this case, $\frac{\partial EU}{\partial w_u}<0$ when w_u is in a left neighbourhood of $\hat{w}_u$, and $EU\to-\infty$ as $w_u\to-\infty$. Thus, EU has a local maximum in the interval $w_u\in(-\infty,\hat{w}_u)$.

Case 2: $\hat{w}_u\leq w_u<\tilde{w}_u$, $w_m\leq\theta$, $w_d>\theta$.

In this case, EU is decreasing and convex when w_u is small and decreasing and concave when w_u is large.

Case 3: $\tilde{w}_u\leq w_u<\theta$, $w_m<\theta$, $w_d\leq\theta$.

In this case, EU is convex.

Case 4: $\theta\leq w_u<$, $w_m\leq\theta$, $w_d\leq\theta$.

In this case, EU is increasing when w_u is small (higher but close to θ) and tends to $-\infty$ as $w_u\to+\infty$. Thus, EU has a local maximum over $w_u\in[\theta,+\infty)$.

Using the Lagrangian approach, the wealth w_u at the local maximum over $w_u\in(-\infty,\hat{w}_u)$ in Case 1 is governed by

$$p_1(u-r)^{1-\gamma}A(\theta-w_u)^{-\gamma}-p_2(r-m)^{1-\gamma}(\hat{w}_u-w_u)^{-\gamma}-p_3(r-d)^{1-\gamma}(\tilde{w}_u-w_u)^{-\gamma}=0, \quad \text{(A.13)}$$

and the wealth w_u at the local maximum over $w_u\in[\theta,+\infty)$ in Case 4 is governed by

$$p_1(u-r)^{1-\gamma}(w_u-\theta)^{-\gamma}-p_2(r-m)^{1-\gamma}A(w_u-\hat{w}_u)^{-\gamma}-p_3(r-d)^{1-\gamma}A(w_u-\tilde{w}_u)^{-\gamma}=0. \quad \text{(A.14)}$$

A.13.2 $m \geq r$

The case of $m \geq r$ is symmetric to the case of $m < r$.

Define

$$\hat{w}_d = \frac{w_0 r - q_{21}\theta}{q_{23}}, \qquad \tilde{w}_d = \frac{w_0 r - q_{12}\theta}{q_{13}}, \tag{A.15}$$

which is the value of w_d when $w_u = \theta$ and $w_m = \theta$, respectively.

When $w_0 \geq r^{-1}\theta$, we have $\theta \leq \hat{w}_d \leq \tilde{w}_d$, and EU is increasing over $w_d \in (-\infty, \theta)$, concave over $w_d \in [\theta, \hat{w}_d)$, decreasing over $w_d \in [\theta, \hat{w}_d)$ if $A \geq \frac{p_3}{p_1}(\frac{r-d}{u-r})^{1-\gamma}$, and decreasing over $w_d \in [\hat{w}_d, +\infty)$. As a result, the local maximum of EU over $w_d \in [\theta, \hat{w}_d)$ is the global maximum, at which the Lagrangian approach leads to w_d^* governed by

$$p_3(r-d)^{1-\gamma}(w_d - \theta)^{-\gamma} - p_2(m-r)^{1-\gamma}(\tilde{w}_d - w_d)^{-\gamma} - p_1(u-r)^{1-\gamma}(\hat{w}_d - w_d)^{-\gamma} = 0. \tag{A.16}$$

and w_u^* and w_m^* are determined by (A.10).

When $w_0 < r^{-1}\theta$, we have $\tilde{w}_d < \hat{w}_d < \theta$, and EU has a local maximum over $w_d \in (-\infty, \tilde{w}_d)$ and $w_d \in [\theta, +\infty)$, at which the wealth is governed, respectively, by

$$p_3(r-d)^{1-\gamma}A(\theta - w_d)^{-\gamma} - p_2(m-r)^{1-\gamma}(\tilde{w}_d - w_d)^{-\gamma} - p_1(u-r)^{1-\gamma}(\hat{w}_d - w_d)^{-\gamma} = 0, \tag{A.17}$$

and

$$p_3(r-d)^{1-\gamma}(w_d - \theta)^{-\gamma} - p_2(m-r)^{1-\gamma}A(w_d - \tilde{w}_d)^{-\gamma} - p_1(u-r)^{1-\gamma}A(w_d - \hat{w}_d)^{-\gamma} = 0. \tag{A.18}$$

B Portfolio Choice under Decision Weights

B.1 Decision Weights under Portfolio Choice

In addition to the S-shaped utility function (2) that characterizes loss aversion, probability weighting is another major component of prospect theory, which is a form of subjective probabilities. Suppose that there are S states in the world. Denote by $\tilde{w}_s$ the outcome in state s and p_s the associated physical probability satisfying $\sum_s p_s = 1$. Under cumulative prospect theory (Tversky and Kahneman, 1992), the S outcomes are first sorted in ascending order:[27]

$$\tilde{w}_{-m} < \tilde{w}_{-m+1} < \cdots < \tilde{w}_{-1} < \tilde{w}_{(0)} = 0 < \tilde{w}_1 < \cdots < \tilde{w}_n, \tag{B.1}$$

[27] In a portfolio choice context, if the portfolio outcomes in two states are equal, they should be merged into one state.

where $\tilde{w}_{-m}, \cdots, \tilde{w}_{-1}$ are losses, and $\tilde{w}_1, \cdots, \tilde{w}_n$ are gains relative to the reference point. With a reference point θ, the losses/gains equal $\tilde{w}_s = w_s - \theta$, where w_s is the level of wealth in state s. Then the decision weights π_s for losses and gains are separately determined by the cumulative probability weighting functions $\Omega^{\pm}$:

$$\begin{cases} \pi_{-i} = \Omega^-(p_{-m} + \cdots + p_{-i}) - \Omega^-(p_{-m} + \cdots + p_{-i-1}), & -i = -m, \cdots, -1, \\ \pi_j = \Omega^+(p_j + \cdots + p_n) - \Omega^+(p_{j+1} + \cdots + p_n), & j = 1, \cdots, n. \end{cases} \tag{B.2}$$

Tversky and Kahneman (1992) propose weighting functions of the form:

$$\Omega^{\pm}(P) = \frac{P^{\delta^{\pm}}}{[P^{\delta^{\pm}} + (1-P)^{\delta^{\pm}}]^{1/\delta^{\pm}}}. \tag{B.3}$$

These functions tend to inflate the probabilities for both tails. The decision weights in (B.2)–(B.3) may not sum up to 1. In this paper, we further assume $\sum_s \pi_s = 1$ that allows us to interpret π_s as subjective probability. As a result, the expectation $\mathbb{E}[\cdot]$ in the optimization problem (1) is taken under the investor's subjective probabilities (B.2).

In the decision problems studied in Tversky and Kahneman (1992), the set of outcomes is predetermined. As a result, the decision weights are determined prior to decision making. However, in portfolio choice problems as we studied, the outcomes depend on investor's decisions, which depends on the decision weights. As a result, portfolio choice under prospect theory is a fixed-point problem.

Prior to studying the optimal portfolio choice problem, we discuss an implication of our previous results under loss aversion for the decision weight. In a portfolio choice problem, the decision weight for the status quo ($\tilde{w}_{(0)} = 0$) as defined in (B.1) is endogenously determined. In fact, Proposition 3 and Proposition 9 show that the case of $\tilde{w}(\equiv w - \theta) = 0$ occurs if and only if $\theta = w_0 r$, independent of probabilities. In this case, the investor holds only the riskless asset, and there is only one (deterministic) outcome $w = \theta$, implying that $\pi_0 \equiv 1$ is endogenously determined. This differs from the original decision problems with predetermined outcomes studied in Tversky and Kahneman (1992), where all decision weights, including the weight for outcome $\tilde{w}_{(0)}$, are predetermined by the probability weighting functions (B.2).

Lemma 5. *In a portfolio choice problem, the decision weight π_0 for the outcome $\tilde{w}_{(0)} = 0$ is endogenously determined and equals $\pi_0 \equiv 1$.*

Lemma 5 implies that in portfolio choice, if one state has non-zero outcome ($\tilde{w}_{(0)} \neq 0$), then the outcomes in all states are not zero. Lemma 5 further implies that a pre-specified π_0 in a portfolio choice/asset pricing problem may lead to inconsistency.

B.2 Optimal Portfolio Choice

Under prospect theory, the optimal portfolio choice is a fixed-point problem. We solve this problem in two steps. First, given a fixed ordering of outcomes (portfolio values across states), our results in Sections 3–4 can be used, leading to the optimal portfolios conditional on this ordering. In the second step, we can derive the optimal portfolios by maximizing the expected utility across all orderings.

To elaborate the effects of probability weighting on the optimal choice, we consider a binomial model. Section 3 shows that there can be five different orderings of outcomes:

$$\begin{cases} (a) \quad w_d < \theta < w_u; \\ (b) \quad w_u < \theta < w_d; \\ (c) \quad \theta < w_d < w_u; \\ (d) \quad \theta < w_u < w_d; \\ (e) \quad w_u = w_d = \theta, \end{cases} \tag{B.4}$$

and each leads to a different set of decision weights. Denote by $\pi_u^{(o)}$ and $\pi_d^{(o)}$ the decision weights for states u and d, respectively, in ordering (o), where $o = a, \cdots, e$. We further denote the pricing kernel in this ordering by $\xi_u^{(o)}$ and $\xi_d^{(o)}$, which satisfy $\xi_u^{(o)} = \frac{r-d}{\pi_u^{(o)}(u-d)r}$ and $\xi_d = \frac{u-r}{\pi_d^{(o)}(u-d)r}$.

B.2.1 Existence of Optimal Solutions

Under probability weighting, a necessary condition for the existence of optimal solutions is that there exist (local) optimal solutions conditional on every ordering. Lemma 1 in Section 3 further shows that to derive the asymptotic behavior, we only need to consider orderings (a) and (b), under which the EU may approach positive infinity, since the EU always has a finite upper bound under the other orderings (c)–(e). As a result, Proposition 2 leads to the following results on the solution existence conditions.

Proposition 12. *The optimal portfolio choice problem under prospect theory has no internal solution if* $A < \underline{A}^{pw}$*, where*

$$\underline{A}^{pw} = \max\left\{ \left(\frac{\pi_u^{(a)}}{\pi_d^{(a)}}\right)^{\gamma}\left(\frac{\xi_d^{(a)}}{\xi_u^{(a)}}\right)^{1-\gamma},\ \left(\frac{\pi_d^{(a)}}{\pi_u^{(a)}}\right)^{\gamma}\left(\frac{\xi_u^{(a)}}{\xi_d^{(a)}}\right)^{1-\gamma},\ \left(\frac{\pi_u^{(b)}}{\pi_d^{(b)}}\right)^{\gamma}\left(\frac{\xi_d^{(b)}}{\xi_u^{(b)}}\right)^{1-\gamma},\ \left(\frac{\pi_d^{(b)}}{\pi_u^{(b)}}\right)^{\gamma}\left(\frac{\xi_u^{(b)}}{\xi_d^{(b)}}\right)^{1-\gamma} \right\}. \tag{B.5}$$

By comparing with the lower bound in the case without probability weighting (10), we have $\underline{A}^{pw} \geq \underline{A}$. Here the superscript "*pw*" stands for "probability weighting". It shows that the probability weighting tends to lead to a stricter lower bound for A, since $\underline{A}^{pw}$ should be the maximum of $\underline{A}$ across all orderings.

Furthermore, under probability weighting, there may not exist optimal solution even if $A > \underline{A}^{pw}$ as shown shortly in Example 3. In contrast, $A > \underline{A}$ is a sufficient condition for the existence of internal solutions without probability weighting (Proposition 2). These results thus show probability weighting intensifies the conditions for the existence of solutions.

Corollary 9. *Probability weighting tends to lead to stricter conditions for the existence of optimal solutions.*

B.2.2 Optimal Solutions

In the following analyses, we suppose $A > \underline{A}^{pw}$ and study the properties of the solutions.

First, we consider $\theta < w_0 r$. In this case, only orderings (c) and (d) can occur as shown in Proposition 3. Conditioning on an ordering ((c) or (d)), the decision weights $\pi_{u,d}^{(o)}$ are determined by the weighting functions (B.2), and then we can solve for the optimal solutions. Denote by $w_{u,d}^{(o)}$ the optimal wealth conditioning on ordering (o). Proposition 3 shows that this conditional optimal wealth is given by:

$$\begin{aligned} w_u^{(o)} - \theta &= \left(w_0 - r^{-1}\theta\right)\left[\pi_u^{(o)}(\xi_u^{(o)})^{1-\frac{1}{\gamma}} + \pi_d^{(o)}(\xi_d^{(o)})^{1-\frac{1}{\gamma}}\right]^{-1}(\xi_u^{(o)})^{-\frac{1}{\gamma}}, \\ w_d^{(o)} - \theta &= \left(w_0 - r^{-1}\theta\right)\left[\pi_u^{(o)}(\xi_u^{(o)})^{1-\frac{1}{\gamma}} + \pi_d^{(o)}(\xi_d^{(o)})^{1-\frac{1}{\gamma}}\right]^{-1}(\xi_d^{(o)})^{-\frac{1}{\gamma}}, \end{aligned} \quad \text{for} \quad o = c, d. \tag{B.6}$$

and the conditional subjective expected utility is given by

$$EU^{(o)} = \frac{(w_0 - r^{-1}\theta)^{1-\gamma}}{1-\gamma}[\pi_u^{(o)}(\xi_u^{(o)})^{1-\frac{1}{\gamma}} + \pi_d^{(o)}(\xi_d^{(o)})^{1-\frac{1}{\gamma}}]^{\gamma}. \tag{B.7}$$

The unconditionally optimal solution is obtained by comparing the conditional EU across the orderings. The optimal wealth is summarized in the following proposition.[28]

[28]The case with $w_u^{(c)} < w_d^{(c)}$ and $w_u^{(d)} > w_d^{(d)}$ cannot occur in Proposition 13. In fact, the decision weights

Proposition 13. *Suppose that* $A > \underline{A}^{pw}$ *and* $\theta < w_0 r$.

1. *When* $w_u^{(c)} > w_d^{(c)}$ *and* $w_u^{(d)} < w_d^{(d)}$, *the value function is given by* $J = \max\{EU^{(c)}, EU^{(d)}\}$, *and the optimal wealth is given by* (B.6) *with the ordering that leads to the higher EU.*

2. *When* $w_u^{(c)} > w_d^{(c)}$ *and* $w_u^{(d)} > w_d^{(d)}$,[29] *the value function is* $J = EU^{(c)}$, *and the optimal wealth is given by* $w_u^* = w_u^{(c)}$ *and* $w_d^* = w_d^{(c)}$.

3. *When* $w_u^{(c)} < w_d^{(c)}$ *and* $w_u^{(d)} < w_d^{(d)}$, *the value function is* $J = EU^{(d)}$, *and the optimal wealth is given by* $w_u^* = w_u^{(d)}$ *and* $w_d^* = w_d^{(d)}$.

The following example further illustrates Case 1 in Proposition 13.

Example 2. *Suppose that* $A = 2$, $\gamma = 0.5$, $\theta = 1$, $w_0 = 1$, $r = 1.1$, $u = 1.65$, $d = 0.65$, $p = 0.45$, *and* $\delta^+ = 0.61$.[30]

Because $\theta < w_0 r$, *orderings* (*c*) *and* (*d*) *can happen. If ordering* (*c*) *applies, it follows from* (B.3) *that the decision weights are* $\pi_u^{(c)} = 0.553$ *and* $\pi_d^{(c)} = 0.447$. *The optimal holding is* $x^{(c)} = 0.08$, *and the subjective expected utility is* $EU^{(c)} = 0.65$. *In this case, the optimal wealth is* $w_u^{(c)} = 1.14$ *and* $w_d^{(c)} = 1.06$, *satisfying* $\theta < w_d^{(c)} < w_u^{(c)}$, *i.e., ordering* (*c*). *If ordering* (*d*) *applies, the decision weights are* $\pi_u^{(d)} = 0.447$ *and* $\pi_d^{(d)} = 0.553$. *In this case,* $x^{(d)} = -0.003$, $EU^{(d)} = 0.63$, $w_u^{(d)} = 1.098$, $w_d^{(d)} = 1.101$, *satisfying* $\theta < w_u^{(d)} < w_d^{(d)}$.

The subjective expected utility in ordering (*c*) *is higher than that in* (*d*)*; thus,* $x^* = x^{(c)}$ *and* $J = EU^{(c)}$.

Two observations follow from Example 2. First, the asset exhibits positive skewness under the physical measure since $p < 0.5$. Probability weighting inflates the probabilities for the right tail (state u). As a result, the investor overestimates the stock expected return, holding more stocks in her portfolio. This result is consistent with the results of Barberis and Huang (2008) that prospect theory investors prefer lottery-like assets.

---

in orderings (c) and (d) satisfy $\pi_u^{(d)} < p < \pi_u^{(c)}$. In this case, the subjective risk premium in ordering (c) is higher than that in ordering (d). However, the case with $w_u^{(c)} < w_d^{(c)}$ and $w_u^{(d)} > w_d^{(d)}$ implies that the subjective risk premium is negative in ordering (c) and is positive in ordering (d).

[29] In this case, ordering (d) cannot occur, which requires $w_u < w_d$.

[30] In this example with $\theta < w_0 r$, decision weights are determined by weighting function $\Omega^+(P) = \frac{P^{\delta^+}}{[P^{\delta^+} + (1-P)^{\delta^+}]^{1/\delta^+}}$, independent of the coefficient δ^-, since all outcomes are positive.

Second, without probability weighting, the physical risk premium of the stock is zero, and the optimal holdings of the stock are zero $x^* = 0$ as shown in Proposition 3. However, with probability weighting, conditional on either longing or shorting the stock, there is an optimal solution, which "confirms" the investor's initial choice in the first step.

Proposition 13 Case 2 tends to occur when the asset return has high skewness. In this case, the investor overestimates the right tail (state u) probability, leading to higher holdings of stocks. In contrast, Proposition 13 Case 3 tends to occur when the asset return is more negatively skewed. These results again show that probability weighting cause investors to prefer assets with positive skewness and dislike assets with negative skewness.

It follows from Proposition 13 that when the optimal solution exists, the value function tends to be higher than the case without probability weighting since probability weighting allows the investor to compare different orderings of portfolio values when making the optimal choice. In addition, Corollaries 1–4 in Section 3 still hold true under probability weighting.

Next, we consider the case $\theta = w_0 r$. Proposition 3 shows that the investor holds only the riskless asset when $\theta = w_0 r$. In this case, there is only 1 outcome: $w^{(e)} = \theta$, i.e., ordering (e). Then the decision weight is $\pi^{(e)} \equiv 1$, and the results on the optimal portfolios and value function are the same as the case without probability weighting.

Proposition 14. *Suppose that $A > \underline{A}^{pw}$ and $\theta = w_0 r$. Under prospect theory preferences, the optimal stock holdings are $x^* = 0$, the optimal wealth is given by $w_{u,d} = \theta$, and the value function is $J = 0$.*

Finally, we can also derive the optimal wealth for the case $\theta > w_0 r$ following the same discussions.[31] We discuss an example in this case that shows that $A > \underline{A}^{pw}$ is not sufficient to guarantee the existence of optimal solutions.

Example 3. *Suppose that $\theta > w_0 r$, $A = 2$, $\gamma = 0.5$, $r = 1.1$, $u = 1.2$, $d = 1$, $p = 0.5$, $\delta^+ = 0.61$, and $\delta^- = 0.5$. In this case, orderings (a) and (b) can happen. If ordering (a) applies, the decision weights are $\pi_u^{(c)} = 0.46$ and $\pi_d^{(c)} = 0.54$. In this case, the optimal wealth satisfies $w_u^{(a)} < \theta < w_d^{(a)}$, which is inconsistent with ordering (a). Thus, this ordering cannot occur. If ordering (b) applies, $\pi_u^{(b)} = 0.54$ and $\pi_d^{(b)} = 0.46$. In this case, $w_d^{(b)} < \theta < w_u^{(b)}$, which is inconsistent with ordering (b). Thus, this ordering cannot occur either.*

[31] In this case, there will be eight different situations. Results are available upon request.

Therefore, the fixed point problem of optimal choice has no solution even if $A > \underline{A}^{pw}$. *However, under the parameters but without probability weighting, the optimal choice has solutions.*

References

Abdellaoui, M. (2000), 'Parameter-free elicitation of utility and probability weighting functions', *Management Science* **46**, 1497–1512.

Aït-Sahalia, Y. and Brandt, M. (2001), 'Variable selection for portfolio choice', *Journal of Finance* **56**, 1297–1351.

Ang, A., Bekaert, G. and Liu, J. (2005), 'Why stocks may disappoint', *Journal of Financial Economics* **76**, 471–508.

Azevedo, E. and Gottlieb, D. (2012), 'Risk-neutral firms can extract unbounded profits from consumers with prospect theory preferences', *Journal of Economic Theory* **147**, 1291–1299.

Baillon, A., Bleichrodt, H. and Spinu, V. (2020), 'Searching for the reference point', *Management Science* **66**, 93–112.

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

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

Barberis, N., Huang, M. and Santos, T. (2001), 'Prospect theory and asset prices', *Quarterly Journal of Economics* **116**, 1–53.

Barberis, N. and Xiong, W. (2009), 'What drives the disposition effect? An analysis of a long-standing preference-based explanation', *Journal of Finance* **64**, 751–784.

Barberis, N. and Xiong, W. (2012), 'Realization utility', *Journal of Financial Economics* **104**, 251–271.

Baucells, M., Weber, M. and Welfens, F. (2011), 'Reference-point formation and updating', *Management Science* **57**, 506–519.

Benartzi, S. and Thaler, R. (1995), 'Myopic loss aversion and the equity premium puzzle', *Quarterly Journal of Economics* **100**, 73–92.

Berkelaar, A., Kouwenberg, R. and Post, T. (2004), 'Optimal portfolio choice under loss aversion', *Review of Economics and Statistics* **86**, 973–987.

Choi, K., Jeon, J. and Koo, H. (2022), 'Intertemporal preference with loss aversion: Consumption and risk-attitude', *Journal of Economic Theory* **200**, 105380.

Cox, J. and Huang, C. (1989), 'Optimal consumption and portfolio policies when asset process follow a diffusion process', *Journal of Economic Theory* **49**, 33–83.

Dai, M., Qin, C. and Wang, N. (2024), 'Dynamic trading with realization utility', *Journal of Finance* **forthcoming**.

De Giorgi, E., Hens, T. and Levy, H. (2004), Existence of CAPM equilibria with prospect theory preferences, working paper, University of Zurich.

Dybvig, P. and Rogers, L. (2013), High hopes and disappointment, working paper, Washington University in Saint Louis.

Dybvig, P., Rogers, L. and Back, K. (1999), 'Portfolio turnpikes', *Review of Financial Studies* **12**, 165–195.

Gneezy, U. (2005), *Updating the reference level: Experimental evidence,*, Springer US, pp. 263–284. in *Experimental Business Research: Marketing, Accounting and Cognitive Perspectives Volume III*, Eds. G.I. Bischi, C. Chiarella and I. Sushko.

Gomes, F. (2005), 'Portfolio choice and trading volume with loss-averse investors', *Journal of Business* **78**, 675–706.

Grinblatt, M. and Han, B. (2005), 'Prospect theory, mental accounting, and momentum', *Journal of Financial Economics* **78**, 311–339.

Gul, F. (1991), 'A theory of disappointment aversion', *Econometrica* **59**, 667–686.

He, X. and Zhou, X. (2011), 'Portfolio choice under cumulative prospect theory: An analytical treatment', *Management Science* **57**, 315–331.

Ingersoll, J. (2016), 'Cumulative prospect theory, aggregation, and pricing', *Critical Finance Review* **5**, 305–350.

Ingersoll, J. and Jin, L. (2013), 'Realization utility with reference-dependent preferences', *Review of Financial Studies* **26**, 723–767.

Kahneman, D. and Tversky, A. (1979), 'Prospect theory: An analysis of decision under risk', *Econometrica* **47**, 263–292.

Köszegi, B. and Rabin, M. (2006), 'A model of reference-dependent preferences', *Quarterly Journal of Economics* **121**, 1133–1165.

Li, Y. and Yang, L. (2013), 'Prospect theory, the disposition effect, and asset prices', *Journal of Financial Economics* **107**, 715–739.

Meng, J. and Weng, X. (2018), 'Can prospect theory explain the disposition effect? A new perspective on reference points', *Management Science* **64**, 3331–3351.

Merton, R. (1971), 'Optimum consumption and portfolio rules in a continuous-time model', *Journal of Economic Theory* **3**, 373–413.

Merton, R. (1973), 'An intertemporal capital asset pricing model', *Econometrica* **41**, 867–887.

Segal, U. and Spivak, A. (1990), 'First order versus second order risk aversion', *Journal of Economic Theory* **51**, 111–125.

Sethi, S. and Taksar, M. (1988), 'A note on Merton's "Optimum Consumption and Portfolio Rules in a Continuous-Time Model"', *Journal of Economic Theory* **46**, 395–401.

Shefrin, H. and Statman, M. (1985), 'The disposition to sell winners too early and ride losers too long: Theory and evidence', *Journal of Finance* **40**, 777–790.

Tversky, A. and Kahneman, D. (1991), 'Loss aversion in riskless choice: A reference-dependent model', *Quarterly Journal of Economics* **106**, 1039–1061.

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

Wu, G. and Gonzalez, R. (1996), 'Curvature of the probability weighting function', *Management Science* **42**, 1676–1690.