

Optimal Portfolio Choice Under Loss Aversion

Arjan Berkelaar* and Roy Kouwenberg

Econometric Institute Report EI 2000-08/A

Econometric Institute
Faculty of Economics
Erasmus University Rotterdam
P.O.Box 1738, 3000 DR Rotterdam
The Netherlands

Department of Finance
Faculty of Economics
Erasmus University Rotterdam
P.O.Box 1738, 3000 DR Rotterdam
The Netherlands

Abstract

Prospect theory and loss aversion play a dominant role in behavioral finance. In this paper we derive closed-form solutions for optimal portfolio choice under loss aversion. When confronted with gains a loss averse investor behaves similar to a portfolio insurer. When confronted with losses, the investor aims at maximizing the probability that terminal wealth exceeds his aspiration level. Our analysis indicates that a representative agent model with loss aversion cannot resolve the equity premium puzzle. We also extend the martingale methodology to allow for more general utility functions and provide a simple approach to incorporate skewed and fat-tailed return distributions.

Keywords: optimal asset allocation, behavioral finance, loss aversion.

JEL Classifications Codes: G11,G12.

*Corresponding author: Erasmus University Rotterdam, Econometric Institute, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands. Tel.nr.: +31-104082388, Fax.nr.: +31-104089162, Email: berkelaar@few.eur.nl

1 Introduction

In standard consumption-investment models agents are assumed to be fully rational expected utility maximizers. Under the paradigm of constant relative risk aversion (CRRA) and a constant investment opportunity set, Samuelson (1969) and Merton (1969) have shown that myopic policies are optimal. In other words: (i) risk aversion does not affect the optimal mix of risky assets and (ii) the investment horizon is irrelevant. Recently, several authors (Brennan, Schwartz & Lagnado 1997, Kim & Omberg 1996, Campbell & Viceira 1999) have studied the impact of stochastic opportunity sets on optimal portfolio choice. However, investors are assumed to behave rationally and maximize expected utility.

This fact ignores the huge amount of evidence for anomalies in human behavior found by psychologists. The field of behavioral finance has evolved attempting to understand and explain how emotions and cognitive errors influence investors and the decision-making process. The common belief in this field is that the study of psychology and other social sciences can shed light on the efficiency of financial markets as well as explain many stock market anomalies, market bubbles, and crashes. Probably the most important exponent of behavioral finance is prospect theory, originating from the work of Kahneman & Tversky (1979).

Kahneman & Tversky (1979) found that contrary to expected utility theory, people treat gains and losses differently and typically over- or underweight true probabilities. They also found that individuals are much more distressed by prospective losses than they are happy with equivalent gains. Moreover, individuals respond differently to equivalent situations depending on whether the situation is presented in the context of losses or gains. Finally, they have found that people will take more risks to avoid losses than to realize gains. Faced with gains investors are risk averse, however when facing losses investors become risk seeking.

Although prospect theory has been around for at least two decades, relatively little is known about the implications of loss aversion for optimal portfolio choice. A rigorous analysis is complicated by the fact that the prospective value function displays non-differentiability (first-order risk aversion) and is convex for losses and concave for gains. As a result, the traditional dynamic programming approach based on the Hamilton-Jacobi-Bellman equation cannot be applied. We extend the martingale methodology of Cox & Huang (1989) in order to derive closed-form solutions for the optimal portfolio choice of a loss averse investor.

In particular, we are interested in the following questions:

1. What does loss aversion imply for portfolio choice? Does the pain experienced when stocks perform poorly make a loss averse investor reluctant to allocate heavily to stocks?
2. Does the optimal portfolio choice of a loss averse investor exhibit time diversification effects?
3. How does a loss averse investor behave when confronted with negatively skewed and leptokurtic asset returns?
4. Can loss aversion provide an explanation for the equity premium puzzle?

Our results are as follows: first, under general security price processes, we show that the optimal

strategy for a loss averse investor can be decomposed in a probability maximizing strategy and a growth strategy. A loss averse investor aims at maximizing the probability of reaching his aspiration level (the reference point distinguishing gains and losses) and desires some additional upside by investing part of his wealth in a growth strategy. The probability maximizing strategy is equivalent to an investment in a binary cash-or-nothing call. The growth strategy is related to the growth optimal portfolio.

Second, when asset prices follow a geometric Brownian motion, we derive closed-form solutions for the optimal portfolio choice. When confronted with gains a loss averse investor behaves similar to a portfolio insurer, aiming to retain wealth above the aspiration level. When confronted with losses the investor maximizes the probability that terminal wealth exceeds the aspiration level. This behavior is denoted as a break-even effect. Due to this effect a loss averse investor always prefers a gamble over a sure rate of return when he is below his aspiration level. These break-even effects demonstrate that a loss averse investor allocates more heavily to stocks when behind on his investment.

Break-even effects are prevalent in human behavior. Let us illustrate this with a simple example. Suppose an individual is to meet a liability due next week, e.g. pay the rent for housing. Assume his current funds are insufficient to meet this obligation. Furthermore, we assume that the individual does not receive any exogenous income over the course of the week (such as labor income, or a donation from mom and dad). As he knows that currently he simply cannot fulfill his obligation, such an individual is inclined to gamble, e.g. by going to a casino. During the week, as the individual still lacks the means to meet his obligation, he is likely to gamble even more. The individual gambles since he “has not made peace with his losses”.

In the original description of loss aversion risk seeking behavior in the domain of losses is always accompanied by the possibility to break even. We extend the original formulation of Kahneman & Tversky (1979) by distinguishing between gambles that allow the investor to break even and gambles that lack this opportunity. We assume that the investor becomes risk averse again for large losses, as the likelihood to break even becomes very small. This behavior is also documented by Thaler & Johnson (1990) based on experiments. In the extended version of loss aversion, the investor decreases his exposure to stocks again in bad states. In intermediate states the fraction invested in stocks is increased as break-even effects dominate the behavior of the investor.

Finally, we study the effect of skewness and kurtosis in asset returns on the optimal portfolio choice under loss aversion. In standard economic models for portfolio choice skewness and kurtosis in asset returns are ignored. However, there is convincing evidence that stock returns are fat-tailed and negatively skewed. We rely on a Gram-Charlier expansion of the lognormal density function, rather than making explicit distributional assumptions. We investigate whether loss aversion in combination with a skewed and fat-tailed return distribution can explain the equity premium puzzle. Our findings indicate that break-even effects are amplified in the presence of skewness and kurtosis in stock returns.

In conclusion, our partial equilibrium results indicate that loss aversion *can not* explain the participation puzzle. As a loss averse individual is confronted with losses, break-even effects induce him to allocate heavily to stocks. Moreover, as relative risk aversion over gains is quite

low, a loss averse investor even allocates heavily to stocks when confronted with gains. Our findings contradict the intuition of Benartzi & Thaler (1995) that myopic loss aversion makes the investor “reluctant to allocate heavily to stocks, due to the pain experienced when stocks perform poorly.” The main reason is that break-even effects persuade loss averse investors to accept any gamble over a sure rate of return.

This paper provides three contributions. Besides deriving closed-form solutions for the optimal portfolio choice under loss aversion we extend the martingale methodology of Cox & Huang (1989) to allow for pseudoconcave utility functions. In this paper we apply this extended approach to the portfolio choice of a loss averse individual, however, other situations where the utility function is pseudoconcave can be analyzed with our approach as well. Furthermore, we are the first to derive closed-form solutions for the optimal portfolios when asset returns are skewed and fat-tailed. The Gram-Charlier expansion utilized in this paper can also be used in other situations.

Apart from methodological contributions, our paper also contributes to the literature on loss aversion and the equity premium puzzle. Surprisingly, there is little research on optimal portfolio choice under loss aversion, although it is a celebrated model in behavioral finance. Benartzi & Thaler (1995) were the first to consider the implication of loss aversion on portfolio choice. They study a one-period model, and perform simulations to calculate the optimal mix of stocks and bonds. Benartzi & Thaler (1995) reason that myopic loss aversion makes investors reluctant to invest in stocks, even in the face of a substantial equity premium. Unfortunately, they do not provide closed-form solutions for the optimal portfolio choice.

Barberis, Huang & Santos (2000) recently tried to make plausible that loss aversion can explain the equity premium puzzle. Barberis et al. (2000) only derive the first-order conditions for equilibrium prices and do not derive the optimal asset allocation under loss aversion. Furthermore, they consider an infinite horizon model and measure loss averse preferences intertemporally in a discrete-time model. Moreover, they assume that the value function is piecewise linear, ignoring risk aversion over gains. Finally, our results demonstrate that their explanation for the equity premium puzzle stems mainly from the house money effect of Thaler & Johnson (1990) and not so much from loss aversion.

Independently, Gomes (2000) also studied the portfolio choice problem under loss aversion. He concludes that a loss averse investor will not hold stocks unless the expected equity premium is quite high. The model studied by Gomes is quite different from ours. He studies a single period model, with only two states of the world. In a multi-state setting he has to rely on numerical techniques. Moreover, he does not solve the problem to optimality but rather constructs the solution part by part. Finally, he does not use the original value function proposed by Kahneman & Tversky (1979) but an adjusted version.

This paper is organized as follows. In Section 2 we consider the general continuous-time economy and discuss myopic loss aversion. In Section 3 we study the behavior of a loss averse investor under general price uncertainty. Section 4 derives closed-form solutions for the optimal portfolio choice under lognormal asset returns. We also study the impact of skewness and kurtosis on the optimal portfolio choice. Section 5 discusses whether loss aversion may explain the equity

premium puzzle. Section 6 concludes this paper.

2 Economic Setting

2.1 The Economy

In this section we formulate our economy and the dynamic investment problem. We consider a finite-horizon, $[0, T]$, economy.¹ We assume that the investor trades $K + 1$ assets continuously in a market without transaction costs. The zero-th asset is a riskless money market account $S_0(t)$:

$$dS_0(t) = r(t)S_0(t)dt,$$

The prices of the remaining assets $S = \{S_k(t)\}_{k=1}^K$ follow Ito processes with drift rate $\mu_k(t)$ and volatility $\sigma_k(t)$:

$$dS_k(t) = \mu_k(t)S_k(t)dt + \sigma_k(t)S_k(t)dB_t, \quad k = 1, \dots, K, \quad (1)$$

where the interest rate $r(\cdot)$, the drift rates $\mu(\cdot)$ and the volatility matrix $\sigma(\cdot)$ are adapted process (possibly path-dependent).

In order to meet his investment goals the investor chooses a portfolio consisting of the K risky assets and the riskless money market account. We denote the fraction invested in risky asset k at time t by $w_k(t)$ and the fraction invested in the riskless asset at t by $w_0(t)$. For any self-financing portfolio the wealth W_t of the investor can be expressed as the following stochastic process (using vector notation):

$$dW_t = r(t)W_tdt + (\mu(t) - r(t))'w(t)W_tdt + \sigma(t)'w(t)W_tdB_t. \quad (2)$$

where we substituted $w_0(t) = 1 - \sum_i w_i(t)$. The initial wealth of the investor is denoted by W_0 .

Markets are assumed to be complete, implying the existence of a unique state price density (or pricing kernel) ξ_t , given by

$$\xi_t = \exp\left(-\int_0^t r(s)ds\right)Z_t,$$

where Z_t denotes the Radon-Nikodym derivative of a change of probability measure defined by

$$Z_t = \frac{dQ}{dP} = \exp\left(-\frac{1}{2}\int_0^t \|\kappa(s)\|^2 ds - \int_0^t \kappa(s)'dB_s\right),$$

and $\kappa(t) = \sigma^{-1}(t)(\mu(t) - r(t))$ denotes the market price of risk process. Note that σ is an invertible matrix, as we assumed that markets are complete.

¹All stated processes are assumed to be well-defined and satisfy the appropriate regularity conditions. For technical details the reader is referred to Karatzas & Shreve (1998).

Alternatively, we may write the state price density process as

$$\frac{d\xi_t}{\xi_t} = -r(t)dt - \kappa(t)'dB_t, \quad \xi_0 = 1. \quad (3)$$

The pricing kernel ξ_t relates future cash-flows $C(X_s, s)$, $s \in (t, T]$ to today's price P_t :

$$P(X_t, \tau) = E_t \left(\int_t^T \frac{\xi_s C(X_s, s)}{\xi_t} ds \right),$$

where $\tau = T - t$. For a zero-coupon bond with maturity $\tau = T - t$ we have $P(X_t, \tau) = E_t \left(\frac{\xi_T}{\xi_t} \right)$. The state price density process (or pricing kernel) will play an important role in deriving the optimal trading strategies.

Another important interpretation of the pricing kernel is due to Merton (1990). Merton shows that the inverse of the pricing kernel is equal to the value of the growth-optimum portfolio. The growth optimum portfolio, π_t , is the strategy that maximizes the expected growth rate of the portfolio (or the average continuously compounded return on the portfolio), i.e. $\pi_t = \arg \max_{\pi_t} E_t(\log(W_T))$. Moreover, this portfolio is mean-variance efficient. When the investment opportunity set is constant and utility is concave and differentiable, any optimal portfolio can be represented in terms of simple combinations of the growth optimum portfolio and the riskless asset (so-called two fund separation).

2.2 Myopic Loss Aversion

Optimal portfolio choice traditionally relies on rational behavior and risk aversion. Over the past 30 years psychologists have found compelling evidence that people treat gains and losses differently, and in particular that losses loom larger than gains. This behavior was formalized in Kahneman and Tversky's prospect theory (Kahneman & Tversky 1979). Behavioral finance has promoted this theory as a descriptive theory for decision-making under uncertainty. Recently, prospect theory has received a lot of attention in the financial literature. Benartzi & Thaler (1995), Shumway (1997), and Barberis et al. (2000) try to explain the equity premium puzzle with loss aversion, while Odean (1998) focuses on the disposition effect.²

Prospect theory was originated by the work of Kahneman & Tversky (1979) who presented a number of choice problems to students and university faculty and found several violations of expected utility theory.³ The main conclusions from these experiments are:

1. people care about changes in wealth rather than wealth itself;
2. people care about small risks;

²The disposition effect refers to the tendency of investors to sell winning investments too soon and hold losing investments too long, see e.g. Shefrin & Statman (1985).

³This group of people may not seem representative for the average population. However, after the original publication of Kahneman and Tversky similar experiments have been performed with other groups of people as well.

3. people are risk seeking in the domain of losses.

These findings are formalized in prospect theory: economic agents maximize an S-shaped value function that is concave for gains, convex for losses and steeper for losses than for gains, and apply decision weights over- and underweighting true probabilities (prospective rather than expected values). In particular the value function is modeled as:

$$U(x) = \begin{cases} -A(\theta - x)^{\gamma_1}, & \text{for } x \leq \theta \\ +B(x - \theta)^{\gamma_2}, & \text{for } x > \theta \end{cases} \quad (4)$$

where $A\gamma_1 > 0$ and $B\gamma_2 > 0$ to ensure that $U(\cdot)$ is an increasing function and $0 < \gamma_1, \gamma_2 \leq 1$. An illustration of the value function can be found in Figure 1.

Critical for the value function is the reference point distinguishing gains and losses. This reference point may represent the status quo of the investor (e.g. his current assets) or an aspiration level. Another reasonable reference point may be some liability: an investor concerned with his balance sheet might only experience a loss as the value of his assets falls below the value of his liabilities. Odean (1998) uses the original purchase price of an asset as reference point. We interpret the reference point as an aspiration level throughout and assume that it is constant. Incorporating a stochastic reference point is straightforward and does not alter our conclusions qualitatively.

In order to study the investment problem under loss aversion we refrain from using subjective decision weights (Kahneman & Tversky 1979), and assume that agents maximize expected values rather than prospective values. This assumption is also made by Shumway (1997) and Barberis et al. (2000). Consequently, we model loss averse agents as expected utility maximizers. Furthermore, we ignore consumption decisions since one of the implications of the framework of Kahneman & Tversky (1979) is that investors have preferences over returns, rather than over the consumption profile that these returns help provide.

A crucial notion in loss aversion is the evaluation frequency. We illustrate the impact of the evaluation frequency on optimal asset allocation with a simple example. Consider an investor who can invest in a risky asset in a discrete two-period economy. We assume that the investor has the following utility function (which is a special case of (4))

$$U(x) = \begin{cases} -2.25(-x)^{0.88} & x < 0 \\ x^{0.88} & x > 0 \end{cases}$$

These parameter values are estimated by Kahneman & Tversky (1992) by conducting experiments and using a nonlinear regression.

Suppose there is a risky asset worth \$100 today. Over the course of the first year the price either goes up to \$110 with probability 1/2 or the price goes down to \$95 with probability 1/2. It is straightforward to verify that expected utility is negative, hence the investor will not buy this asset. Suppose however, that over the second year returns are 10% and -5%, with probability 1/2 each, again. Over the two year horizon the stock has either gone up to \$121 with probability 1/4, gone up to \$104.50 with probability 1/2, or gone down to \$90.25 with probability 1/4.

Expected utility over the two year horizon is positive, hence the investor will want to buy the security. This simple example shows that investment decisions are influenced by the evaluation frequency.

As this example illustrates, loss aversion implies that investors take different decisions when confronted with a single gamble or a sequence of gambles. The investor in our example is not willing to accept a one-shot gamble, although he agrees to buy the asset if confronted with the two-period gamble.⁴ Benartzi & Thaler (1999) conducted experiments in the context of retirement savings decisions to study repeated investment decisions over time. The study of Benartzi and Thaler demonstrates empirically that, when investors are loss-averse they are willing to take more risk if they evaluate their performance infrequently.

How often do investors evaluate their performance? As Benartzi & Thaler (1995) point out people file taxes each year and many institutional investors are evaluated annually by regulating authorities. An evaluation period of one year seems very plausible. Even when saving for retirement a young investor may experience utility from gains and losses on his investments as he checks his investment account. Benartzi & Thaler (1995) labeled the combination of loss aversion and a short evaluation period *myopic loss aversion*. In their attempt to explain the equity premium puzzle they consider myopic loss aversion as a crucial ingredient.

Our main purpose is to study the effect of loss-averse preferences on the demand for risky assets and analyze whether loss aversion can explain the equity premium puzzle. We assume that the horizon of our investor under consideration coincides with his evaluation period. This implies that we use a rather short horizon and we only measure utility at the evaluation horizon. This assumption allows us to express the utility function of a loss-averse agent in terms of wealth at the first evaluation date. In the next we discuss and solve the portfolio optimization problem of a loss-averse investor.

3 Portfolio Optimization Under Loss Aversion

In this section we solve the portfolio optimization problem of a loss averse investor and discuss and analyze the properties of the solution. We make no assumptions on the distribution of the asset returns throughout this section as we derive the optimal wealth profile of a loss averse investor at his evaluation horizon. This general derivation provides insight in the structure of the optimal strategy regardless of distributional assumptions. In the next section we specialize to the case where asset prices follow geometric Brownian motions with constant interest rate and market price of risk and we study some of the analytic properties of the optimal strategies in more detail.

As the utility function of a loss-averse investor is nonconcave and nondifferentiable the traditional stochastic dynamic programming approach and its associated Hamilton-Jacobi-Bellman (HJB) equation cannot be utilized to derive the optimal portfolio. We shall apply the martin-

⁴Kahneman & Tversky (1981) stressed that the behavior captured by loss aversion only applies to one-shot gambles.

gale methodology to reduce the dynamic portfolio problem to an equivalent static optimization problem. The martingale methodology was independently developed by Karatzas, Lehoczky & Shreve (1987) and Cox & Huang (1989) under the complete market paradigm and time-additive utility.⁵ The utility function is assumed to be continuous, increasing, and strictly concave. Duffie & Skiadas (1994) extended this approach to a more general setting with non-additive preferences given through stochastic differential utility, however the utility function is still assumed to be smooth.⁶

We extend the martingale methodology to allow for non-concave and non-differentiable utility functions. This is especially important as the value function in loss aversion is neither concave nor differentiable. We rely on pseudoconcavity of the value function (see Appendix A). Pseudoconcavity is sufficient to guarantee existence of an optimal solution. Strict pseudoconcavity ensures that the solution is unique. We apply our extended martingale methodology to study the optimal portfolio selection problem under loss aversion, and we derive closed-form solutions for the optimal dynamic trading strategies.

We explicitly incorporate a non-negativity constraint on the investor's wealth, as the utility function does not exhibit infinite marginal utility at zero wealth.⁷ This complicates the derivation of the optimal solution and renders an approach based on solving the HJB equation analytically impossible. The investor aims at solving the following dynamic portfolio problem:

$$\begin{aligned} \max \quad & E[U(T, W_T)] \\ \text{s.t.} \quad & dW_t = rW_t dt + (\mu - \iota r)' w_t W_t dt + \sigma' w_t W_t dB_t \\ & W_t \geq 0, \forall t \in [0, T] \end{aligned} \quad (5)$$

The martingale methodology allows the problem to be restated as the following static optimization problem.

$$\begin{aligned} \max \quad & E[U(T, W_T)] \\ \text{s.t.} \quad & E[\xi_T W_T] \leq \xi_0 W_0 \\ & W_T \geq 0 \end{aligned} \quad (6)$$

The assumption of complete markets, a constant opportunity set and the absence of exogenous sources of income (e.g. labor income) ensures that the optimally invested wealth will never reach zero before time T (Cox & Huang 1989).⁸ As a result it suffices to restrict terminal wealth only.

Proposition 1 characterizes the optimal solution. As the utility function is pseudoconcave, existence of this solution is guaranteed.

⁵He & Pearson (1991) and Karatzas, Lehoczky, Shreve & Xu (1991) extended the martingale approach to an incomplete market setting.

⁶Stochastic differential utility can be interpreted as a continuous-time limit of recursive utility studied by Kreps & Porteus (1978) and Epstein & Zin (1989).

⁷Merton (1971) derived linear policies for HARA-utility functions. Sethi & Taksar (1988) have shown that some of Merton's solutions are incorrect since bankruptcy problems are ignored. Explicit constraints are necessary to prevent bankruptcy, and closed-form solutions are more difficult to derive.

⁸When an exogenous source of income is taken into account the optimally invested wealth may reach zero prior to the horizon. For an analysis of this problem the reader is referred to El-Karoui & Jeanblanc-Picqué (1998).

Proposition 1 *The time T optimal wealth of a loss averse investor with $0 < \gamma_1 < 1$ and $0 < \gamma_2 < 1$ is*

$$W(T) = \begin{cases} \theta + \left(\frac{y\xi_T}{B\gamma_2}\right)^{1/(\gamma_2-1)} & \xi_T < \bar{\xi} \\ 0 & \xi_T \geq \bar{\xi} \end{cases} \quad (7)$$

where $\bar{\xi}$ solves $f(\bar{\xi}) = 0$ with

$$f(x) = \frac{1-\gamma_2}{\gamma_2} \left(\frac{1}{yx}\right)^{\gamma_2/(1-\gamma_2)} (B\gamma_2)^{1/(1-\gamma_2)} - \theta yx + A\theta^{\gamma_1}$$

and $y \geq 0$ satisfies $E(\xi_T W(T)) = \xi_0 W_0$.

In Figure 2 we display the optimal terminal wealth of a loss-averse (LA) agent and compare it with a HARA agent and a constant relative risk averse (CRRA) agent. The HARA agent has utility

$$U(W_T) = B(W_T - \theta)^{\gamma_2}, \quad (8)$$

with initial wealth $W_0 \geq \theta \exp(-rT)$. The HARA agent can always ensure that his terminal wealth exceeds the aspiration level θ , as he fully invests in the riskfree asset when $W_t = \theta \exp(-rT)$. The HARA investor, therefore, behaves similar to a portfolio insurer. The benchmark CRRA investor has a relative risk aversion coefficient $\gamma = \gamma_2$ and always invests a constant fraction in stocks.

Note that terminal wealth is a discontinuous function of the growth optimal portfolio $1/\xi_T$. In good states (low ξ_T) the loss-averse agent behaves like the CRRA agent. In moderate states the investor behaves similar to a portfolio insurer (the optimal strategy for our HARA agent) with floor θ , as first-order risk aversion makes the investor averse against small changes in wealth. In bad states ($\xi_T \geq \bar{\xi}$) the investor ends up with zero wealth. Since the investor is mostly concerned with small changes in wealth (first-order risk aversion) relative to the threshold the gambling behavior below the threshold (the investor is risk seeking below θ) causes the investor to incur large losses in these bad states.

The location of $\bar{\xi}$ depends on the preferences of the investor. As the reference point θ increases $\bar{\xi}$ decreases, and consequently the bad-states region increases. Accordingly, wealth in the good states region should be increased to meet the higher floor. If the investor becomes more risk averse over gains (i.e. γ_2 is decreased) $\bar{\xi}$ increases, and consequently the bad-states region shrinks. Accordingly, wealth in the good states should be decreased. As first-order risk aversion becomes more pronounced, i.e. $(A\gamma_1/B\gamma_2)$ increases, $\bar{\xi}$ increases, and consequently the bad-states region shrinks at the expense of the intermediate region. This can be understood by realizing that, as first-order risk aversion becomes more pronounced, the investor cares more about small risks and the intermediate region becomes more attractive.

We can interpret the optimal strategy of the LA investor as a combination of a binary option and an investment in the growth optimal portfolio with a nonnegative, nonlinear payoff. The binary option is a cash-or-nothing call on the growth optimal portfolio paying θ , with a strike

price equal to $1/\bar{\xi}$. This binary option is the optimal strategy for an investor that maximizes the probability that terminal wealth exceeds θ (or equivalently the investor minimizes the probability of falling short of θ), as the following proposition shows.⁹

Proposition 2 *Assume that an investor aims at maximizing the probability of beating θ . Then, the optimal payoff function of the investor is a binary cash-or-nothing call option on the growth optimal portfolio with strike price $1/\bar{\xi}$, i.e. the time T optimal wealth of the investor is:*

$$W(T) = \begin{cases} \theta & \xi_T \leq \frac{A}{\theta} \frac{1}{y} \\ 0 & \xi_T > \frac{A}{\theta} \frac{1}{y} \end{cases} \quad (9)$$

where $y \geq 0$ satisfies $E(\xi_T W(T)) = \xi_0 W_0$.

The second fund represents an investment in a binary asset-or-nothing call on the growth optimal portfolio with a nonlinear payoff and a strike price equal to $1/\bar{\xi}$. The decomposition of the optimal terminal wealth of the LA investor into a binary cash-or-nothing call option and an investment in the growth optimal portfolio gives a nice interpretation to the optimal strategy of a loss averse investor. The binary cash-or-nothing call gives rise to a *probability maximizing strategy* as discussed above. The investment in the growth optimal portfolio (i.e. a binary asset-or-nothing call on the growth optimal portfolio) will be labeled the *growth strategy*.

In this section we discussed the characteristics of the optimal solution without making assumptions about the security price processes. We showed that the optimal strategy, under general security price processes, can be decomposed into a probability maximizing strategy and a growth strategy. In the next section we specialize to lognormal state prices with constant interest rate and market price of risk (i.e. we assume that asset prices follow geometric Brownian motions with constant coefficients). We derive closed-form solutions for the optimal wealth and portfolio policies and analyze the characteristics of these solutions. In particular we study the two strategies derived in this section, the probability maximizing strategy and the growth strategy, in more detail.

4 Optimal Portfolio Choice under Loss Aversion

4.1 Properties of the LA Strategy

In the previous section we characterized the optimal terminal wealth of a loss-averse investor under general price processes. In this section we derive closed-form solutions for the optimal policies when asset prices follow a geometric Brownian motion. We assume that interest rates and the market price of risk are constant. We study whether loss aversion can explain the equity premium puzzle, and in particular the participation puzzle.¹⁰ We also consider whether

⁹The probability maximizing strategy was recently studied by Browne (1999).

¹⁰The participation puzzle refers to the stylized fact that a large number of people do not hold stocks. Mankiw & Zeldes (1991) show that the participation puzzle forms a large part of the equity premium puzzle.

the claim of Benartzi & Thaler (1995) that a loss averse investor with a shorter evaluation period allocates less to stocks holds true.

We first consider the characteristics of the optimal strategy of a loss averse investor. When applying the martingale methodology the optimal strategies are not given in feedback form as with stochastic dynamic programming. Instead, the optimal strategies are derived as a function of the pricing kernel ξ_t (or equivalently as a function of the growth optimal portfolio $Z_t = 1/\xi_t$). Proposition 3 presents closed-form expressions for the optimal wealth and portfolio strategies at time $t < T$.

Proposition 3 *Consider a loss-averse investor. Assume, $0 < \gamma_1 < 0$, $0 < \gamma_2 < 1$, and r and κ are constants. Then:*

(i) *The time t optimal wealth is given by*

$$W(t) = \theta e^{-r(T-t)} N(d_1(\bar{\xi})) + \left(\frac{B\gamma_2}{y\xi_t} \right)^{1/(1-\gamma_2)} e^{\Gamma(t)} N(d_2(\bar{\xi})) \quad (10)$$

where $N(\cdot)$ denotes the cumulative standard normal distribution function and

$$\begin{aligned} d_1(x) &= \frac{\log\left(\frac{x}{\xi_t}\right) + \left(r - \frac{1}{2}\|\kappa\|^2\right)(T-t)}{\|\kappa\|\sqrt{(T-t)}} \\ d_2(x) &= d_1(x) + \frac{\|\kappa\|\sqrt{(T-t)}}{1-\gamma_2} \\ \Gamma(t) &= \frac{\gamma_2}{1-\gamma_2} \left(r + \frac{1}{2}\|\kappa\|^2\right)(T-t) + \frac{1}{2} \left(\frac{\gamma_2}{1-\gamma_2}\right)^2 \|\kappa\|^2(T-t) \end{aligned}$$

and where $\bar{\xi}$ solves $f(\bar{\xi}) = 0$ with

$$f(x) = \frac{1-\gamma_2}{\gamma_2} \left(\frac{1}{yx}\right)^{\gamma_2/(1-\gamma_2)} (B\gamma_2)^{1/(1-\gamma_2)} - \theta yx + A\theta^{\gamma_1}$$

(ii) *The fraction of wealth invested in the risky assets is*

$$w(t) = \frac{(\sigma')^{-1}\kappa}{W(t)} \left(\frac{\theta e^{-r(T-t)} \phi(d_1(\bar{\xi}))}{\|\kappa\|\sqrt{(T-t)}} + \left(\frac{B\gamma_2}{y\xi_t} \right)^{1/(1-\gamma_2)} e^{\Gamma(t)} \left(\frac{\phi(d_2(\bar{\xi}))}{\|\kappa\|\sqrt{(T-t)}} + \frac{N(d_2(\bar{\xi}))}{1-\gamma_2} \right) \right) \quad (11)$$

where $N(\cdot)$ denotes the cumulative standard normal distribution function and $\phi(\cdot)$ denotes the standard normal density function.

Note that as ξ_t goes to infinity (and equivalently wealth goes to zero) the fraction invested in stocks goes to zero (to avoid bankruptcy). As ξ_t goes to zero (and equivalently wealth goes to infinity) the optimal fraction invested in stocks tends to the optimal policy of the constant relative risk averse investor with $\gamma = \gamma_2$.

Figure 3 compares the optimal wealth at time $t < T$ (10) prior to the investment horizon and the relative exposure to stocks for the LA investor with the portfolio insurance policy and the benchmark CRRA case. We assume that there is one risky asset (stocks). Figure 3(a) shows that the optimal time- t wealth exhibits concavity in the intermediate region as the investor begins to insure himself. Figure 3(b) reveals that the optimal fraction invested in stocks is V-shaped. In good states the investor behaves similar to a portfolio insurer, investing more heavily in the riskfree asset to establish the floor. However, as risk aversion ($\gamma = 0.88$) is very low, the benchmark investor holds a large portion of wealth in equities. Consequently, even the loss averse agent invests heavily in stocks. Moreover, in bad states the investor increases his exposure to stocks, as risk seeking over losses becomes more pronounced.

In the previous section we have shown that the optimal LA strategy can be decomposed into a probability maximizing strategy and a growth strategy. Figure 4 shows the optimal fraction invested in stocks as a function of the pricing kernel for each of these strategies. The growth strategy dictates a diminishing stock position as the pricing kernel increases. The risk-taking behavior is entirely caused by the probability maximizing strategy. As the investor only cares about the probability that terminal wealth exceeds his aspiration level, stocks are more attractive despite their higher volatility. We label this behavior as a break-even effect. The probability maximizing strategy illustrates that “a person who has not made peace with his losses is likely to accept gambles that would be unacceptable to him otherwise” (Kahneman & Tversky 1979).

Due to the probability maximizing behavior, the investor bets on a favorable realization of a large equity investment, in bad states, to break even with his reference point. Since, in the description of myopic loss aversion, risk seeking in the presence of losses is always accompanied by an opportunity to break even with the reference point, the LA investor always bets on this favorable realization. As the probability that terminal wealth exceeds the aspiration level increases by accepting a gamble, the investor prefers this gamble over a sure rate of return. These break-even effects contradict the intuition of Benartzi & Thaler (1995) that loss averse investors shun stocks as the horizon shortens and that they are reluctant to allocate heavily to stocks, due to the pain experienced when stocks perform poorly.

Figure 5 displays the optimal fraction invested in stocks for different values of risk aversion over gains γ_2 and the Sharpe-ratio κ . The effect of higher risk aversion (Figure 5(a)) over gains is obvious: a more risk averse investor allocates less to stocks. When γ_2 goes to zero the optimal fraction invested in stocks tends to the optimal policy of the probability maximizer (Figure 4). Figure 5(b) shows the effect of the Sharpe ratio on the optimal fraction invested in stocks. If the Sharpe ratio declines the investor desires more protection as stocks yield less return per unit of risk. In good states the investor decreases the exposure to stocks more heavily to gain protection. Consequently, the break-even point (minimum fraction in stocks) occurs at a lower value of the pricing kernel (higher wealth level). In bad states the investor more aggressively increases his allocation to stocks to break even when confronted with a lower Sharpe-ratio.

As optimal wealth is a strictly decreasing function of the pricing kernel ξ we may also derive the optimal strategies in feedback form. The optimal fraction invested in stocks as function of wealth is again V-shaped. The fraction invested in stocks is minimal at some critical wealth level. Figure 6 shows these critical wealth levels as a function of time. If wealth falls below this

critical value break-even effects dominate the behavior of the loss averse investor. When wealth remains above the critical value the investor behaves prudent. In the presence of gains, the demand for stocks increases as the investor becomes more wealthy (as in the portfolio insurance strategy). In the presence of losses, the investor desires to break even with his reference point by increasing his demand for stocks.

It should be noted that the critical path for small t is below the satiation level, $\theta e^{-r(T-t)}$, of a probability maximizer. In this region the growth strategy obviously provides an additional cushion. When the time to maturity shrinks the loss averse investor becomes more risk seeking than a probability maximizer (as the critical level is above $\theta e^{-r(T-t)}$) and the growth strategy no longer provides a cushion. The critical wealth level represents the kink in the derived utility function at time t . The martingale methodology of Cox & Huang (1989) also allows us to determine this indirect utility function, even though it does not have the desired derivatives to satisfy the HJB equation. Proposition 4 presents the indirect utility function for a loss-averse agent.

Proposition 4 *Consider a loss-averse investor. Assume, $0 < \gamma_1 < 0$, $0 < \gamma_2 < 1$, and r and κ are constants. Denote the time- t optimal wealth $W_t = F(t, \xi_t)$. Then, the indirect utility function $J(W, t)$ is given by:*

$$J(W, t) = \hat{J}(\xi_t, t) = -A\theta^{\gamma_1} N(-d_3(\bar{\xi})) + B \left(\frac{B\gamma_2}{y\xi_t} \right)^{\gamma_2/(1-\gamma_2)} e^{\Gamma(t)} N(d_2(\bar{\xi})) \quad (12)$$

where

$$d_3(x) = d_1(x) + \|\kappa\| \sqrt{(T-t)}$$

Figure 7 shows the indirect utility function and its corresponding relative risk aversion function. The relative risk aversion function (Figure 7(b)) is clearly discontinuous as the derived utility function (Figure 7(a)) inherits the kink of the value function. Risk aversion is increasing below a critical wealth level inducing the investor to increase the optimal fraction in stocks as wealth decreases. Above the critical wealth level the relative risk aversion function is decreasing, inducing the investor to increase the fraction invested in stocks as wealth increases.

The optimal LA strategy can be fully understood from the behavior of the relative risk aversion function. When wealth remains above the critical level, the investor is prudent. The optimal policy exhibits time diversification: at longer horizons the investor invests more in stocks. As wealth falls below the critical level prior to the evaluation horizon, the investor's behavior is dominated by a gambling attitude. The investor aims to break even with his aspiration level. The optimal policy displays a reverse time diversification effect: at shorter horizons the investor allocates more heavily to stocks. As the investor's wealth is below his aspiration level, a shorter horizon leaves less time to break even, and consequently the loss averse investor increases his exposure to stocks.

4.2 Break-even Effects

We have seen that an LA investor always gambles at low levels of wealth in order to break even. This can be attributed to the fact that in the description of loss aversion risk seeking in the presence of losses is always accompanied by an opportunity to break even with the reference point. Thaler & Johnson (1990) demonstrate empirically that there is an important distinction between gambles that offer the opportunity to break even and gambles that lack this opportunity. Gambles that offer the opportunity to break even may be acceptable to an investor, whereas gambles that lack this opportunity are avoided. Thaler & Johnson (1990) find empirical evidence that investors are risk seeking over gambles that allow them to break even, whereas they are risk averse over gambles that lack this opportunity.

In this section we consider an extension of the value function of (Kahneman & Tversky 1979) that takes this evidence into account. We assume that the investor becomes risk averse again for large losses as the likelihood to break even is very small. In particular, we consider the following modification to (4):¹¹

$$U(W) = \begin{cases} \frac{1}{\gamma}W^\gamma - \frac{1}{\gamma}\underline{W}^\gamma - A(\theta - W)^{\gamma_1} & W \leq \underline{W} \\ -A(\theta - W)^{\gamma_1} & \underline{W} < W < \theta \\ +B(W - \theta)^{\gamma_2} & W \geq \theta \end{cases} \quad (13)$$

where $\gamma < 1$ ensuring that the utility function is concave when $W \leq \underline{W}$. Moreover, we choose the utility function below \underline{W} to display constant relative risk aversion. The additional constant is chosen to ensure that the utility function is continuous. The utility function is pseudoconcave again.

Proposition 5 characterizes the optimal solution

Proposition 5 *Consider a loss-averse investor with utility (13). Assume, $0 < \gamma_1 < 0$, $0 < \gamma_2 < 1$, $\gamma < 1$ and r and κ are constants. Then:*

(i) *The time T optimal wealth is given by*

$$W(T) = \begin{cases} \theta + \left(\frac{y\xi_T}{B\gamma_2}\right)^{1/(\gamma_2-1)} & \xi_T \leq \bar{\xi} \\ (y\xi_T)^{1/(\gamma-1)} & \xi_T > \bar{\xi} \end{cases} \quad (14)$$

where $\bar{\xi}$ solves $f(\bar{\xi}) = 0$ with

$$f(x) = \frac{1-\gamma_2}{\gamma_2} \left(\frac{1}{yx}\right)^{\gamma_2/(1-\gamma_2)} (B\gamma_2)^{1/(1-\gamma_2)} - \theta yx + A(\theta - \underline{W})^{\gamma_1} + \frac{1}{\gamma}\underline{W}^\gamma - \left(\frac{1-\gamma}{\gamma}\right) (yx)^{\gamma/(\gamma-1)}$$

and $y \geq 0$ satisfies $E(\xi_T W(T)) = \xi_0 W_0$.

¹¹Independently, Gomes (2000) proposed a similar utility function, however, he assumes that utility displays increasing relative risk aversion for small levels of wealth.

(ii) The time t optimal wealth is given by

$$W(t) = \theta e^{-r(T-t)} N(d_1(\bar{\xi})) + \left(\frac{B\gamma_2}{y\xi_t} \right)^{1/(1-\gamma_2)} e^{\Gamma(t)} N(d_2(\bar{\xi})) \quad (15)$$

$$+ (y\xi_t)^{1/(\gamma-1)} e^{\Gamma_2(t)} (1 - N(d_3(\bar{\xi}))) \quad (16)$$

where $N(\cdot)$ denotes the cumulative standard normal distribution function and

$$d_1(x) = \frac{\log\left(\frac{x}{\xi_t}\right) + \left(r - \frac{1}{2}\|\kappa\|^2\right)(T-t)}{\|\kappa\|\sqrt{(T-t)}} \quad (17)$$

$$d_2(x) = d_1(x) + \frac{\|\kappa\|\sqrt{(T-t)}}{1-\gamma_2} \quad (18)$$

$$\Gamma(t) = \frac{\gamma_2}{1-\gamma_2} \left(r + \frac{1}{2}\|\kappa\|^2 \right) (T-t) + \frac{1}{2} \left(\frac{\gamma_2}{1-\gamma_2} \right)^2 \|\kappa\|^2 (T-t) \quad (19)$$

and d_3 and Γ_2 equal d_2 and Γ respectively with γ_2 replaced by γ .

(iii) The fraction of wealth invested in the risky assets is

$$w(t) = \frac{(\sigma')^{-1}\kappa}{W(t)} \left(\frac{\theta e^{-r(T-t)} \phi(d_1(\bar{\xi}))}{\|\kappa\|\sqrt{(T-t)}} + \left(\frac{B\gamma_2}{y\xi_t} \right)^{1/(1-\gamma_2)} e^{\Gamma(t)} \left(\frac{\phi(d_2(\bar{\xi}))}{\|\kappa\|\sqrt{(T-t)}} + \frac{N(d_2(\bar{\xi}))}{1-\gamma_2} \right) \right) \\ + \frac{(\sigma')^{-1}\kappa}{W(t)} (y\xi_t)^{1/(\gamma-1)} e^{\Gamma_2(t)} \left(\frac{1 - N(d_3(\bar{\xi}))}{1-\gamma} - \frac{\phi(d_3(\bar{\xi}))}{\|\kappa\|\sqrt{(T-t)}} \right)$$

where $N(\cdot)$ denotes the cumulative standard normal distribution function and $\phi(\cdot)$ denotes the standard normal density function.

Note that as ξ_t goes to infinity (and equivalently wealth goes to zero) the fraction invested in stocks converges to the optimal policy of a constant relative risk averse investor with risk aversion coefficient γ . As ξ_t goes to zero (and equivalently wealth goes to infinity) the fraction invested in stocks tends to the optimal policy of a constant relative risk averse investor with risk aversion coefficient γ_2 .

Figure 8 shows the optimal fraction invested in stocks relative to the benchmark CRRA investor for both the standard LA investor and the extended LA investor. In intermediate states the extended loss averse investor starts to gamble as break-even effects dominate the policy. In bad states the investor decreases his exposure to stocks as break-even effects are no longer sensible and risk aversion dominates again. It should be noted, however, that the equity premium puzzle cannot be explained by this extended description of loss aversion. In bad states, the utility function is equivalent to a power utility function. It is widely known that constant relative risk aversion cannot explain the equity premium puzzle, unless the relative risk aversion coefficient is extremely large (Mehra & Prescott 1985).

4.3 Impact of skewness and kurtosis

In the previous section we derived closed-form solutions for the portfolio choice of a loss-averse investor under the assumption that asset prices are i.i.d. lognormally distributed. There is a vast amount of empirical evidence that asset returns are negatively skewed and leptokurtic. Fama (1965) was the first to show that stock returns (at a daily basis) exhibit fat tails. Sample estimates of skewness for stock index returns tend to be negative, while sample estimates of excess kurtosis for stock returns on indexes are positive. Table 1 shows sample statistics for monthly returns on the S&P500 from January 1990 to December 1999.

Table 1: Summary Statistics for S&P Returns, 1990-1999

	<i>Mean</i>	<i>Std. Dev.</i>	<i>Skewness</i>	<i>Excess Kurtosis</i>
S&P 500	14.25%	46.58%	-0.82	2.21

Statistics are based on monthly returns (annualized).

In this section we study the impact of skewness and kurtosis on the optimal portfolio choice of a loss averse investor. Rather than relying on distributional assumptions such as a Student-t distribution or on other departures from normality such as jump-diffusion models, we adopt a Gram-Charlier expansion of the lognormal density function. The Gram-Charlier expansion was first introduced in finance by Jarrow & Rudd (1982). The advantages of using a Gram-Charlier expansion are:

1. we do not have to make distributional assumptions, which can be hard to defend and estimate from historical data; instead we can rely on sample estimates for skewness and excess kurtosis;
2. we are able to derive closed-form solutions for the optimal portfolio choice of a loss averse investor when asset returns are skewed and fat-tailed.

The Gram-Charlier expansion provides a parsimonious representation of a distribution with skewness and kurtosis. It generates an approximate density function for a standardized random variable. For the lognormal density function, the Gram-Charlier expansion is given by:

$$f(z; \beta, k) = (1 + \beta(z^3 - 3z) + \delta(z^4 - 6z^2 + 3))\phi(z), \quad (20)$$

where $z = (\log(x) - \mu)/\sigma$, $\beta = \zeta/6$, $\delta = \chi/24$ and ζ, χ denote skewness and kurtosis of $\log(x)$ respectively, and $\phi(\cdot)$ denotes the standard normal density function. Equation (20) is often viewed as an approximation to an arbitrary density function with nonzero higher moments, however, for moderate values of ζ and χ it is a density function in itself. An extensive treatment of the Gram-Charlier expansion can be found in Johnson, Kotz & Balakrishnan (1994).

First, we first consider the impact of skewness and kurtosis on the optimal policy for our HARA investor.

Proposition 6 *Consider a HARA investor with utility (8). Let β and δ denote scaled skewness and kurtosis respectively. Then*

(i) The time T optimal wealth is given by

$$W(T) = \theta + \left(\frac{B\gamma_2}{y\xi_T} \right)^{1/(1-\gamma_2)} \quad (21)$$

where $y \geq 0$ satisfies $E(\xi_T W(T)) = \xi_0 W_0$.

(ii) The time t optimal wealth is given by

$$W(t) = \theta e^{-r(T-t)} Q_1(\sigma, \beta, \delta) + \left(\frac{B\gamma_2}{y\xi_t} \right)^{1/(1-\gamma_2)} e^{\Gamma(t)} Q_1(\gamma_2\sigma/(1-\gamma_2), \beta, \delta) \quad (22)$$

where $Q_1(\sigma, \beta, \delta) = 1 + \beta\sigma^3 + \delta\sigma^4$ and

$$\Gamma(t) = \frac{\gamma_2}{1-\gamma_2} \left(r + \frac{1}{2} \|\kappa\|^2 \right) (T-t) + \frac{1}{2} \left(\frac{\gamma_2}{1-\gamma_2} \right)^2 \|\kappa\|^2 (T-t)$$

$N(\cdot)$ denotes the cumulative standard normal distribution function and $\phi(\cdot)$ denotes the standard normal density function.

(iii) The fraction of wealth invested in the risky assets is (in feedback form):

$$w_t = \frac{(\sigma')^{-1}\kappa}{1-\gamma_2} \left(1 - \frac{\theta e^{-r(T-t)}}{W(t)} Q_1(\sigma, \beta, \delta) \right) \quad (23)$$

Since $Q_1(\sigma, \beta, \delta) = 1 + \beta\sigma^3 + \delta\sigma^4$, the optimal policy is influenced by the skewness and kurtosis of asset returns only through the products $\beta\sigma^3$ and $k\sigma^4$ respectively. As these values are quite small, skewness and kurtosis have hardly any impact on the optimal portfolio choice for the HARA investor. For the benchmark CRRA investor the optimal portfolio is even independent of the skewness and kurtosis of asset returns.¹² When the investor exhibits loss averse preferences this is no longer true and skewness and kurtosis have a substantial impact on the optimal strategy. The next proposition presents a closed-form expression for the optimal fraction invested in wealth when asset returns exhibit skewness and kurtosis.

Proposition 7 *Assume that the investor is loss averse with $0 < \gamma_1 < 1$. Let β and k denote scaled skewness and kurtosis respectively. Then,*

(i) The time t optimal wealth is given by

$$W(t) = \theta e^{-r(T-t)} G(d_1(\bar{\xi}), \sigma) + \left(\frac{B\gamma_2}{y\xi_t} \right)^{1/(1-\gamma_2)} e^{\Gamma(t)} G(d_2(\bar{\xi}), \lambda\sigma) \quad (24)$$

¹²The optimal policy for the benchmark CRRA investor is given by Proposition 6 where $\theta = 0$.

where $\lambda = \gamma_2/(1 - \gamma_2)$

$$\begin{aligned}
d_1(x) &= \frac{\log\left(\frac{x}{\bar{\xi}_t}\right) + \left(r - \frac{1}{2}\|\kappa\|^2\right)(T-t)}{\|\kappa\|\sqrt{(T-t)}} \\
\alpha(x) &= d_1(x) + \sigma, \quad d_2(x) = \alpha(x) - \lambda\sigma \\
\Gamma(t) &= \frac{\gamma_2}{1 - \gamma_2} \left(r + \frac{1}{2}\|\kappa\|^2\right)(T-t) + \frac{1}{2} \left(\frac{\gamma_2}{1 - \gamma_2}\right)^2 \|\kappa\|^2(T-t) \\
G(x, \sigma) &= N(x)Q_1(\sigma, \beta, \delta) - \phi(x)Q_2(\alpha(x), \sigma, \beta, \delta) \\
Q_1(\sigma, \beta, \delta) &= 1 + \beta\sigma^3 + \delta\sigma^4 \\
Q_2(\alpha, \sigma, \beta, \delta) &= \beta[(\alpha + \sigma)^2 - \alpha\sigma - 1] + \delta[\alpha((\alpha + \sigma)^2 - \alpha\sigma - 3) + \sigma(\sigma^2 - 1)]
\end{aligned}$$

and where $\bar{\xi}$ solves $f(\bar{\xi}) = 0$ with

$$f(x) = \frac{1 - \gamma_2}{\gamma_2} \left(\frac{1}{x}\right)^{\gamma_2/(1-\gamma_2)} (B\gamma_2)^{1/(1-\gamma_2)} - \theta x + A\theta\gamma_1$$

(ii) The fraction of wealth invested in the risky assets is

$$\begin{aligned}
w(t) &= \frac{(\sigma')^{-1}\kappa}{W(t)} \left(\left(\frac{B\gamma_2}{y\xi_t}\right)^{1/(1-\gamma_2)} e^{\Gamma(t)} \left(\frac{G(d_2(\bar{\xi}), \lambda\sigma)}{(1-\gamma_2)} + \frac{\Psi(\alpha(\bar{\xi}), \lambda\sigma, \beta, \delta)}{\|\kappa\|\sqrt{(T-t)}} \right) + \right. \\
&\quad \left. + \frac{\theta e^{-r(T-t)}}{\|\kappa\|\sqrt{(T-t)}} \Psi(\alpha(\bar{\xi}), \sigma, \beta, \delta) \right) \quad (25) \\
&\quad \quad \quad (26)
\end{aligned}$$

where $\phi(\cdot)$ denotes the standard normal density function and

$$\begin{aligned}
\Psi(\alpha, \sigma, \beta, \delta) &= \phi(\alpha - \sigma)(Q_1(\sigma, \beta, \delta) - Q_2'(\alpha, \sigma, \beta, \delta) + (\alpha - \sigma)Q_2(\alpha, \sigma, \beta, \delta)) \\
&= \phi(\alpha - \sigma)(1 + \beta(\alpha^3 - 3\alpha) + \delta(\alpha^4 - 3\alpha^2 - 3\alpha + 3) + 3\delta\sigma(1 - \alpha))
\end{aligned}$$

Figure 9(a) shows the fraction invested in stocks for a loss averse investor relative to the benchmark CRRA investor for different values of negative skewness. The main conclusion is that the break-even effects are more pronounced as negatively skewed returns make the occurrence of bad states more likely. In the good states region the impact of skewness is relatively small and induces the investor to lower his exposure to stocks. Note that the break-even effect is no longer monotonic as in the lognormal case discussed in the previous section. This non-monotonic effect is caused by the fact that the density function is no longer unimodal.

Figure 9(b) shows the fraction invested in stocks for a loss averse investor relative to the benchmark CRRA investor for different values of excess kurtosis. We may conclude that excess kurtosis amplifies the break-even effects. Again, in the good states region the impact of excess kurtosis is relative small and induces the investor to lower his exposure to stocks. A similar decomposition of the optimal strategy as in the previous section, shows that the excess kurtosis and negative skewness mainly influence the probability maximizing strategy, whereas the growth strategy is almost independent of skewness and kurtosis in asset returns.

5 Loss Aversion and The Equity Premium Puzzle

One of the most striking puzzles in finance, is the question why, historically, the equity premium has been so high and interest rates so low. Standard economic models fail to explain this high equity premium, as was first observed by Mehra & Prescott (1985). This stylized fact is termed the equity premium puzzle, and over the last decade a stream of literature has attempted to explain it. Possible explanations include market imperfections (Basak & Cuoco 1998), incompleteness (Constantinides & Duffie 1996), and alternative assumptions on preferences such as habit formation (Campbell & Cochrane 1999) and loss aversion (Barberis et al. 2000, Benartzi & Thaler 1995).

Campbell & Cochrane (1999) and Barberis et al. (2000) have recently come up with plausible explanations for the equity premium puzzle. Campbell & Cochrane (1999) study a model with habit formation. As people form consumption habits over time, they become averse against a decline in their standard of living. This might induce an individual to be reluctant to investing in the stock market, as he fears a recession. However, as most of the stock market is owned by institutional investors, consumption habits may not produce the complete picture. The model of Campbell & Cochrane (1999) is based on a complete and perfect market, and can replicate the equity premium, the predictability of asset returns, and a nearly constant interest rate.

Barberis et al. (2000) consider a combination of prospect theory and the house money effect of Thaler & Johnson (1990) to resolve the equity premium puzzle. Apart from consumption the investor additionally has preferences over gains and losses from investing in the stock market. The model of Barberis et al. (2000) is related to the model studied by Campbell & Cochrane (1999) as both imply time-varying risk aversion. Such an economy is labeled a moody investor economy by Bekaert & Grenadier (1999). The model of Barberis et al. (2000) reproduces the equity premium, predictability of asset returns, and a nearly constant interest rate. Crucial in the analysis of Barberis et al. (2000) is the house money effect observed by Thaler & Johnson (1990).

Thaler & Johnson (1990) extend the work of Kahneman & Tversky (1979) to study how risk aversion is effected by prior gains and losses. Using real money experiments they find evidence for (i) increased risk aversion after prior losses, (ii) risk seeking after prior gains, and (iii) changes in risk-taking behavior when one outcome can allow decision-makers to break even. The first two observations are labeled as the house money effect referring to the ‘gambling parlance of “playing with the house money”’. Barberis et al. (2000) conclude that “loss aversion cannot *by itself* explain the equity premium”, without time-varying risk aversion as induced by the house money effect.

Mankiw & Zeldes (1991) show that the equity premium puzzle can largely be attributed to a participation puzzle. The participation puzzle refers to the stylized fact that a large fraction of households do not hold stocks. Can loss aversion explain the equity premium or the participation puzzle? Based on our closed-form solutions the answer to this question is negative. As a loss averse individual is confronted with losses break-even effects induce him to allocate heavily to stocks. Moreover, as relative risk aversion over gains is quite low, a loss averse investor even allocates heavily to stocks when confronted with gains. Therefore, we believe that a

representative agent model with loss aversion cannot explain the equity premium, given the estimated parameter of Kahneman & Tversky (1992).

Our extended description of loss aversion cannot resolve the equity premium puzzle either as the optimal policy in bad states tends to the optimal policy under constant relative risk aversion. It is widely known that constant relative risk aversion cannot explain the equity premium, unless the relative risk aversion coefficient is extremely large. Even in the presence of negatively skewed and fat-tailed stock returns the equity premium puzzle (or the participation puzzle) cannot be resolved. Although skewness and kurtosis have a substantial effect on the optimal portfolio choice of a loss averse investor, they mainly aggravate the break-even effects causing the investor to drastically increase his exposure to stocks.

Our analysis of the optimal portfolios of loss averse investors contradict the intuition of Benartzi & Thaler (1995) that myopic loss aversion makes the investor “reluctant to allocate heavily to stocks, due to the pain experienced when stocks perform poorly.” In a one period portfolio optimization model with loss averse preferences, Benartzi & Thaler (1995) find, based on simulations, that myopic loss averse investors do not want to invest in stocks, even in the face of a substantial equity premium. Additionally, they claim that a shorter evaluation period makes stocks less attractive to a loss averse investor. These results seem to indicate that myopic loss aversion may explain the participation puzzle.

How can we resolve the seeming discrepancy between the conclusions of Benartzi & Thaler (1995) versus Barberis et al. (2000) and our closed-form solutions? It should be noted that our model is somewhat different from the one-period model considered by Benartzi & Thaler (1995). In our continuous-time model the investor can rebalance his portfolio continuously, while in the one-period model of Benartzi & Thaler (1995) the investor can only decide on an investment strategy today, without rebalancing afterwards. Benartzi & Thaler (1995) determine the optimal fraction invested in stocks for a loss averse investor in their one-period model, using historical stock returns from 1926 to 1990. They conclude that portfolios with 30 percent to 55 percent stocks are optimal, apparently resolving the equity premium puzzle.

We are concerned about the robustness of this result. Rather than solving the one-period model of Benartzi & Thaler (1995) once for a certain historical data set, we consider the impact of different equity premiums. As in Benartzi & Thaler (1995) the evaluation period is one year and the initial wealth equals the aspiration level θ . For a fixed Sharpe-ratio we draw 10,000 stock returns from a lognormal distribution. We determine the optimal fraction invested in stocks numerically, and repeat this procedure for different values of the Sharpe-ratio. To avoid bankruptcy in the buy-and-hold model, we do not allow for borrowing and consequently the investor cannot allocate more than 100% to stocks.

The results are given in Figure 10. Note the clear discontinuity at a Sharpe-ratio of 0.27. For larger values of the Sharpe ratio the investor allocates 100% of his wealth to stocks. This discontinuity should not be surprising as we have shown previously in our continuous-time model, that terminal wealth is a discontinuous function of ξ_T . Here we obtain a similar result. However, as we only consider buy-and-hold strategies, the fraction invested in stocks is also discontinuous. Since historically the Sharpe-ratio has been around 0.50 in the US, myopic loss aversion cannot

explain the equity premium puzzle, as the investor would dedicate his entire wealth to stocks.

Given our discussion in this section and the analysis in this paper, we conclude that myopic loss aversion in itself is not likely to explain the equity premium, given the parameters of Kahneman & Tversky (1979). Loss aversion induces break-even effects. These break-even effects provide an incentive to allocate heavily to stocks, as the investor aims to maximize the probability that his wealth at the evaluation horizon exceeds his aspiration level. Another concern is the low risk aversion over gains, $\gamma_2 = 0.88$. Even if the investor is confronted with gains, this low risk aversion implies a large equity position. It should be noted that the parameters for utility function (4) estimated by Kahneman & Tversky (1992) apply to an experimental situation rather than observed investment behavior.

6 Conclusions

In this paper we studied the investment behavior of a loss averse investor. We derived closed-form solutions for the optimal portfolio choice under lognormal returns and when returns exhibit skewness and kurtosis. This paper provides three contributions:

1. we derive closed-form solutions for the optimal portfolio choice under loss aversion;
2. we extend the martingale methodology of Cox & Huang (1989) to allow for pseudoconcave utility functions;
3. we derive closed-form solutions for optimal portfolio choice when returns are skewed and fat-tailed.

Our main conclusions are:

1. When confronted with gains a loss averse investor behaves similar to a portfolio insurer trying to keep wealth from falling below his aspiration level.
2. When confronted with losses a loss averse investor behaves similar to a probability maximizer, caring only about the probability that terminal wealth exceeds his aspiration level. This probability maximizing behavior can be interpreted as a break-even effect.
3. The equity premium puzzle cannot be explained in a representative agent model with loss aversion, due to (i) break-even effects, and (ii) small risk aversion over gains

We believe that our findings may shed light on the discussion about loss aversion and the equity premium puzzle, initiated by Benartzi & Thaler (1995) and recently further analyzed by Barberis et al. (2000). Loss averse investors behave similar to probability maximizers when confronted with losses. Accordingly, such investors are only concerned about the probability that terminal wealth exceeds the aspiration level regardless of the volatility of asset returns. Consequently, since stock returns are more volatile than bond returns, a substantial investment in stocks yields a higher probability that terminal wealth will exceed the aspiration level.

These break-even effects are implicitly contained in the description of prospect theory by Kahneman & Tversky (1979). However, such effects were not anticipated by Benartzi & Thaler (1995) and Barberis et al. (2000). Furthermore, we believe that these break-even effects may help explain the disposition effect (Shefrin & Statman 1985), although we do not discern between realized returns and expected returns. The desire to break even may explain why investors hold losing investments too long. A first step in this direction is taken by Odean (1998) who performs an empirical study. Further research is necessary to conclude that loss aversion can explain the disposition effect.

References

- Avriel, M., Diewert, W., Schaible, S. & Zang, I. (1988), *Generalized Concavity*, Plenum Press, New York, USA.
- Barberis, N., Huang, M. & Santos, T. (2000), Prospect theory and asset prices, forthcoming in quarterly journal of economics, University of Chicago.
- Basak, S. & Cuoco, D. (1998), 'An equilibrium model with restricted stock market participation', *Review of Financial Studies* **11**, 309–341.
- Bekaert, G. & Grenadier, S. (1999), Stock and bond pricing in an affine economy, Working paper, Graduate School of Business, Stanford University.
- Benartzi, S. & Thaler, R. (1995), 'Myopic loss aversion and the equity premium puzzle', *Quarterly Journal of Economics* **110**(1), 73–92.
- Benartzi, S. & Thaler, R. (1999), 'Risk aversion or myopia? choices in repeated gambles and retirement investments', *Management Science* **45**(3), 364–381.
- Brennan, M., Schwartz, E. & Lagnado, R. (1997), 'Strategic asset allocation', *Journal of Economic Dynamics and Control* **21**, 1377–1403.
- Browne, S. (1999), 'The risk and rewards of minimizing shortfall probability', *Journal of Portfolio Management* **Summer 1999**, 76–85.
- Campbell, J. & Viceira, L. (1999), 'Consumption and portfolio decisions when expected returns are time-varying', *Quarterly Journal of Economics* **114**, 433–495.
- Campbell, J. Y. & Cochrane, J. H. (1999), 'By force of habit: A consumption-based explanation of aggregate stock market behavior', *Journal of Political Economy* **107**, 205–251.
- Constantinides, G. & Duffie, D. (1996), 'Asset pricing with heterogeneous consumers', *Journal of Political Economy* **104**, 219–240.
- Cox, J. & Huang, C. (1989), 'Optimum consumption and portfolio policies when asset prices follow a diffusion process', *Journal of Economic Theory* **49**, 33–83.
- Diewert, W. (1981), Alternative characterizations of six kinds of quasiconcavity in the nondifferentiable case with applications to nonsmooth programming, in S. Schaible & W. Ziemba, eds, 'Generalized Concavity in Optimization and Economics', Academic Press, New York, USA.
- Duffie, D. & Skiadas, C. (1994), 'Continuous-time security pricing: A utility gradient approach', *Journal of Mathematical Economics* **23**, 107–132.

- El-Karoui, N. & Jeanblanc-Picqué, M. (1998), 'Optimization of consumption with labor income', *Finance and Stochastics* **2**(4), 409–440.
- Epstein, L. & Zin, S. (1989), 'Substitution, risk-aversion, and the temporal behavior of consumption and asset returns: A theoretical framework', *Econometrica* **57**, 937–969.
- Fama, E. (1965), 'The behavior of stock prices', *Journal of Business* **47**, 244–280.
- Gomes, F. (2000), Loss aversion and the demand for risky assets, Working paper, Harvard University.
- He, H. & Pearson, N. (1991), 'Consumption and portfolio policies with incomplete markets and short sale constraints: The infinite-dimensional case', *Journal of Economic Theory* **54**, 259–304.
- Jarrow, R. & Rudd, A. (1982), 'Approximate option valuation for arbitrary stochastic processes', *Journal of Financial Economics* **10**, 347–369.
- Johnson, N., Kotz, S. & Balakrishnan, N. (1994), *Continuous Univariate Distributions, Vol. 1*, Wiley, New York, USA.
- Kahneman, D. & Tversky, A. (1979), 'Prospect theory: An analysis of decision under risk', *Econometrica* **47**, 263–290.
- Kahneman, D. & Tversky, A. (1981), 'The framing of decisions and the psychology of choice', *Science* **211**, 453–458.
- Kahneman, D. & Tversky, A. (1992), 'Advances in prospect theory: Cumulative representation of uncertainty', *Journal of Risk and Uncertainty* **5**, 297–323.
- Karatzas, I., Lehoczky, J. & Shreve, S. (1987), 'Optimal portfolio and consumption decisions for a 'small investor' on a finite horizon', *SIAM Journal on Control and Optimization* **25**, 1157–1186.
- Karatzas, I., Lehoczky, J., Shreve, S. & Xu, G. (1991), 'Martingale and duality methods for utility maximization in an incomplete market', *SIAM Journal on Control and Optimization* **29**, 702–730.
- Karatzas, I. & Shreve, S. (1998), *Methods of Mathematical Finance*, Springer Verlag.
- Kim, T. & Omberg, E. (1996), 'Dynamic nonmyopic portfolio behavior', *Review of Financial Studies* **9**, 141–161.
- Kreps, D. & Porteus, E. (1978), 'Temporal resolution of uncertainty and dynamic choice theory', *Econometrica* **46**, 185–200.
- Mangasarian, O. (1965), 'Pseudo-convex functions', *SIAM Journal on Control* **3**, 281–290.
- Mankiw, N. & Zeldes, S. (1991), 'The consumption of stockholders and non-stockholders', *Journal of Financial Economics* **29**, 97–112.
- Mehra, R. & Prescott, E. (1985), 'The equity premium puzzle', *Journal of Monetary Economics* pp. 145–161.
- Merton, R. (1969), 'Lifetime portfolio selection under uncertainty: The continuous-time case', *Review of Economics and Statistics* **51**, 247–257.
- Merton, R. (1971), 'Optimum consumption and portfolio rules in a continuous time model', *Journal of Economic Theory* **3**, 373–413.
- Merton, R. (1990), *Continuous Time Finance*, Basil Blackwell, Oxford.

- Odean, T. (1998), ‘Are investors reluctant to realize their losses?’, *Journal of Finance* **53**, 1775–1798.
- Samuelson, P. (1969), ‘Lifetime portfolio selection by dynamic stochastic programming’, *Review of Economics and Statistics* **51**, 239–246.
- Sethi, S. & 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. & Statman, M. (1985), ‘The disposition to sell winners too early and ride losers too long: Theory and evidence’, *Journal of Finance* **40**(3), 777–790.
- Shumway, T. (1997), Explaining returns with loss aversion, Working paper, University of Michigan Business School.
- Thaler, R. & Johnson, E. (1990), ‘Gambling with the house money and trying to break even: The effects of prior outcomes on risky choice’, *Management Science* **36**, 643–661.

A Pseudoconcave Functions

In this appendix we briefly discuss pseudoconcavity and show that utility function (4) is a pseudoconcave function. Pseudoconcavity is a useful concept in economics as it allows a generalization of many results originally derived for differentiable and concave utility functions. Pseudoconcavity of the utility function is sufficient to guarantee existence of an optimal solution. Strict pseudoconcavity ensures uniqueness of the optimal solution. For a general treatment of pseudoconcavity and applications in economics the interested reader is referred to Avriel, Diewert, Schaible & Zang (1988).

Definition 1 introduces the notion of pseudoconcavity.

Definition 1 *Let $f : C \rightarrow \mathbb{R}$, where $C \subset \mathbb{R}^n$ is an open convex set. The function f is called pseudoconcave if for $x, y \in C$, and $0 < \lambda < 1$:*

$$f(x) > f(y) \Rightarrow f(\lambda x + (1 - \lambda)y) \geq f(y) + (1 - \lambda)\lambda b(x, y),$$

where $b(x, y)$ is a positive number depending, in general, on x and y .

If f is pseudoconcave, then $-f$ is pseudoconvex. A function which is both pseudoconcave and pseudoconvex is called pseudomonotonic (or pseudolinear). It is easy to prove the following lemma.

Lemma 1 *Let $f : (a, b) \rightarrow \mathbb{R}$ be a one-dimensional strictly increasing function. Then, the function f is pseudomonotonic.*

Proof: Choose $x, y \in (a, b)$ such that $f(x) > f(y)$. Since f is strictly increasing we conclude that $x > y$. Take $0 < \lambda < 1$ and define $b(x, y) = (x - y)\epsilon$, where $\epsilon > 0$ and

$$\epsilon \leq \frac{f(\lambda x + (1 - \lambda)y) - f(y)}{\lambda(1 - \lambda)(x - y)},$$

hence $b(x, y)$ is a positive number. Then it follows that

$$f(y) + (1 - \lambda)\lambda b(x, y) \leq f(\lambda x + (1 - \lambda)y),$$

proving that f is pseudoconcave. The proof that f is pseudoconvex is analogous. \square

It is now straightforward to prove the following result.

Proposition 8 *Utility function (4) is pseudomonotonic.*

Proof: Since (4) is strictly increasing, we conclude from Lemma 1 that the function (4) is pseudomonotonic. \square

Since the function (4) is pseudomonotonic it is, in particular, pseudoconcave.

Let us briefly discuss the implications of pseudoconcavity in convex optimization. Consider the following convex optimization problem:

$$\begin{aligned} \max \quad & f(x) \\ \text{s.t.} \quad & g_i(x) \leq 0, \quad i = 1, \dots, m, \\ & h_j(x) = 0, \quad j = 1, \dots, p. \end{aligned}$$

Mangasarian (1965) (in case of smooth functions) and Diewert (1981) (in case of nonsmooth functions) have shown that if f is pseudoconcave, g_i is pseudoconvex (for all i), and h_j is pseudomonotonic (for all j) then the first order conditions for the Lagrangian of the above problem imply a global optimum. If the function f is strictly pseudoconcave this global optimum is unique.

B Mathematical Proofs

Proof of Proposition 1

Problem (6) is a convex optimization problem with a pseudoconcave objective function. The convex-duality approach (see, e.g. Karatzas & Shreve (1998)) can easily be adapted to allow for pseudoconcave utility functions. The Legendre-Fenchel transform (or convex conjugate) for (6) is defined by

$$U^*(\xi_T) = \max_{W \geq 0} \{U(W) - y\xi_T W\}, \tag{27}$$

where $\xi_T \geq 0$ denotes the pricing kernel. We first consider the solution to this pointwise maximization problem for all ξ_T , and show that (7) solves this problem. Then, we show that the solution to the pointwise maximization problem (27) also solves (6).

Let us denote the part of the utility function below θ by $U_1(W)$ and the part above θ by $U_2(W)$. If $W \leq \theta$, the optimally invested wealth W^* should satisfy the following Karush-Kuhn-Tucker (KKT) conditions

$$\begin{aligned} U_1'(W^*) &= y\xi_T - \lambda, \quad W^* \geq 0, \\ \lambda W^* &= 0, \quad \lambda \geq 0, \end{aligned}$$

where λ denotes the Lagrange multiplier associated with the nonnegativity constraint on wealth. Solving for the KKT conditions we obtain

$$W^* = \max \left\{ \theta - \left(\frac{A\gamma_1}{y\xi_T} \right)^{1/1-\gamma_1}, 0 \right\} := W_1^*.$$

Similarly, if $W \geq \theta$, the optimally invested wealth W^* should satisfy $U_2'(W^*) = y\xi_T$ and we obtain

$$W^* = \theta + \left(\frac{B\gamma_2}{y\xi_T} \right)^{1/1-\gamma_2} := W_2^*$$

Since $U(W) - y\xi_T W$ is not concave, we need to compare the two local maxima W_1^* and W_2^* to determine the global maximum. If

$$U(W_1^*) - y\xi_T W_1^* \leq U(W_2^*) - y\xi_T W_2^* \quad (28)$$

the optimal solution is given by W_2^* . Since $0 < \gamma_1 < 1$ and $0 < \gamma_2 < 1$ it is not difficult to verify W_2^* is the optimal solution if $\xi_T \leq \frac{A\gamma_1}{y}\theta^{\gamma_1-1}$. For $\xi_T > \frac{A\gamma_1}{y}\theta^{\gamma_1-1}$ we should compare the values of $W_1^* = 0$ and W_2^* . From (28) we may conclude that W_2^* is optimal if $\xi_T \leq \bar{\xi}$, where $\bar{\xi}$ satisfies $f(\bar{\xi}) = 0$, with

$$f(x) = \frac{1 - \gamma_2}{\gamma_2} \left(\frac{1}{yx} \right)^{\gamma_2/(1-\gamma_2)} (B\gamma_2)^{1/(1-\gamma_2)} - \theta yx + A\theta^{\gamma_1}.$$

Let us denote the optimal solution of (27) by $W^*(T)$. Now let $W(T)$ be any candidate optimal solution, satisfying the static budget equation in (6). Then, we have

$$\begin{aligned} E[U(T, W_T^*)] - E[U(T, W_T)] &= \\ E[U(T, W_T^*)] - E[U(T, W_T)] - y\xi_0 W_0 + y\xi_0 W_0 &\geq \\ E[U(T, W_T^*)] - E[U(T, W_T)] - yE[\xi_T W_T^*] + yE[\xi_T W_T] &= \\ E[U^*(\xi_T)] - E[U^*(\xi_T)] &\geq 0 \end{aligned}$$

where the first inequality follows from the fact that the static budget equation holds with equality for W_T^* and with inequality for W_T . The second inequality follows from the fact that W_T^* is the optimal solution to (27). Hence, we conclude that $W^*(T)$ is the optimal solution of the static problem (6). This concludes the proof. \square

Proof of Proposition 2

Consider the probability maximizing problem

$$\begin{aligned} \max \quad & AP(W_T \geq \theta) \\ \text{s.t.} \quad & E[\xi_T W_T] \leq \xi_0 W_0, \\ & W_T \geq 0, \end{aligned} \quad (29)$$

where A is some positive constant. We can reformulate this problem as a utility maximization problem with utility

$$U(T, W_T) = \begin{cases} 0 & \text{if } W_T < \theta \\ A & \text{if } W_T \geq \theta. \end{cases}$$

The Legendre-Fenchel transform is given $U^*(\xi_T) = \max_{W \geq 0} \{A \mathbf{1}_{\{W_T \geq \theta\}} - y\xi_T W\}$, and y solves $E(\xi_T W_T) = \xi_0 W_0$. Solving for this problem, it is not difficult to conclude that the optimum is given by:

$$W^* = \begin{cases} \theta & \text{if } \xi_T < \frac{A}{\theta} \frac{1}{y} \\ 0 & \text{if } \xi_T \geq \frac{A}{\theta} \frac{1}{y} \end{cases}$$

This terminal wealth profile is the payoff of a binary cash or nothing call on the growth optimal portfolio $Z_t = 1/\xi_t$ with strike price $1/\bar{\xi} = A/(\theta y)$. A similar argument as in the proof of Proposition 1 demonstrates that W^* is the optimal solution to (29). This concludes the proof. \square

Proof of Proposition 3

(i) Applying Ito's Lemma it is straightforward to show that $\xi_t W_t$ is a martingale. Therefore

$$W_t = \frac{1}{\xi_t} E_t[\xi_T W_T] = \frac{1}{\xi_t} E_t \left[\xi_T \left(\theta + \left(\frac{y \xi_T}{B \gamma_2} \right)^{1/(\gamma_2-1)} \right) \mathbf{1}_{\{\xi_T \leq \bar{\xi}\}} \right], \quad (30)$$

where we substituted (7). Since r and κ are constant, and the asset prices follow geometric Brownian motions, the pricing kernel $\log(\xi_T)$ is normally distributed with mean

$$\log(\xi_t) - \left(r + \frac{1}{2} \|\kappa\|^2 \right) (T - t)$$

and variance $\|\kappa\|^2 (T - t)$. After some straightforward calculus we obtain (10).

(ii) We have two alternative characterizations for wealth at time t . Reformulating (10) as a stochastic process and equating the diffusion part of this stochastic process with the diffusion part of the wealth process (2) we obtain an explicit expression for the fraction of wealth invested in the risky assets.

For ease of notation we define wealth as $W_t =: F(t, \xi_t)$. Using Ito's Lemma and (3) we obtain:

$$dW_t = G(t, \xi_t) dt - \frac{\partial F(t, \xi_t)}{\partial \xi_t} \xi_t \kappa' dB_t, \quad (31)$$

for some $G(t, \xi_t)$ (note that we are only interested in the diffusion part). Comparing the diffusion part of (2) with the diffusion part of (31) we obtain the following expression for the optimal fraction invested in the risky assets:

$$w_t = - \frac{(\sigma')^{-1} \kappa'}{W_t} \left(\frac{\partial F(t, \xi_t)}{\partial \xi_t} \xi_t \right). \quad (32)$$

Substituting (10) in (32) yields the expression in (11). □

Proof of Proposition 4

Denote time- t wealth as a function of the pricing kernel ξ_t by $W_t = F(t, \xi_t)$. Cox & Huang (1989) showed that the indirect utility function $J(W, t)$ is given by

$$J(W, t) := \hat{J}(F^{-1}(W, t), t) = E_t[U(W_T^*)], \quad (33)$$

where W_T^* denotes the optimal terminal wealth of the investor. Substituting (4) and (7) in (33) we have

$$\hat{J}(F^{-1}(W, t), t) = E_t \left[-A \theta^{\gamma_1} \mathbf{1}_{\{\xi_T > \bar{\xi}\}} + B \left(\frac{y \xi_T}{B \gamma_2} \right)^{\gamma_2/(\gamma_2-1)} \mathbf{1}_{\{\xi_T \leq \bar{\xi}\}} \right].$$

Since the pricing kernel $\log(\xi_T)$ is normally distributed with mean

$$\log(\xi_t) - \left(r + \frac{1}{2} \|\kappa\|^2 \right) (T - t)$$

and variance $\|\kappa\|^2(T-t)$, it is straightforward to derive the expression in (12). \square

Proof of Proposition 5

- (i) The proof is similar to the proof of Proposition 1. Solving for the pointwise maximization problem

$$U^*(\xi_T) = \max_{W \geq 0} \{U(W) - y\xi_T W\}, \tag{34}$$

we obtain the following local maxima:

$$\begin{aligned} W_0^* &= (y\xi_T)^{1/(\gamma-1)} \\ W_1^* &= \max \left\{ \theta - \left(\frac{A\gamma_1}{y\xi_T} \right)^{1/1-\gamma_1}, \underline{W} \right\} \\ W_2^* &= \theta + \left(\frac{B\gamma_2}{y\xi_T} \right)^{1/1-\gamma_2} \end{aligned}$$

Comparing the corresponding optimal values, by substituting in 34, we may conclude that W_2^* is optimal when $\xi_T \leq \bar{\xi}$, where $\bar{\xi}$ satisfies $f(\bar{\xi}) = 0$ with

$$f(x) = \frac{1-\gamma_2}{\gamma_2} \left(\frac{1}{yx} \right)^{\gamma_2/(1-\gamma_2)} (B\gamma_2)^{1/(1-\gamma_2)} - \theta yx + A(\theta - \underline{W})^{\gamma_1} + \frac{1}{\gamma} \underline{W}^\gamma - \left(\frac{1-\gamma}{\gamma} \right) (yx)^{\gamma/(\gamma-1)}$$

If $\xi_T \leq \bar{\xi}$ the optimal solution is characterized by W_0^* . Applying a similar line of reasoning as in the proof of Proposition 1 we conclude that the optimal solution to the static optimization problem with utility (13) is given by (14).

- (ii) The proof is analogous to the proof of part (i) in Proposition 3. Wealth at time t satisfies

$$W_t = \frac{1}{\xi_t} E_t[\xi_T W_T] = \frac{1}{\xi_t} E_t \left[\xi_T \left(\theta + \left(\frac{y\xi_T}{B\gamma_2} \right)^{1/(\gamma_2-1)} \right) \mathbf{1}_{\{\xi_T \leq \bar{\xi}\}} + (y\xi_T)^{1/(\gamma-1)} \mathbf{1}_{\{\xi_T > \bar{\xi}\}} \right], \tag{35}$$

where we substituted (14). Furthermore, the pricing kernel $\log(\xi_T)$ is normally distributed. After some straightforward calculus we obtain (15).

- (iii) The proof is analogous to the proof of part (ii) in Proposition 3. Denote wealth by $W_t = F(t, \xi_t)$, the optimal fraction invested is given by:

$$w_t = -\frac{(\sigma')^{-1}\kappa'}{W_t} \left(\frac{\partial F(t, \xi_t)}{\partial \xi_t} \xi_t \right). \tag{36}$$

Substituting (15) in (36) yields the expression in (17). \square

Proof of Proposition 6

- (i) The proof is similar to the proof of Proposition 1. Solving for the pointwise maximization problem

$$U^*(\xi_T) = \max_{W \geq 0} \{U(W) - y\xi_T W\}, \quad (37)$$

we obtain the following optimum:

$$W^* = \theta + \left(\frac{B\gamma_2}{y\xi_T}\right)^{1/(1-\gamma_2)}. \quad (38)$$

Applying a similar line of reasoning as in the proof of Proposition 1 we conclude W^* is the optimal solution to the static optimization problem with utility (8) is given by (21).

- (ii) The proof is analogous to the proof of part (i) in Proposition 3. Wealth at time t satisfies

$$W_t = \frac{1}{\xi_t} E_t[\xi_T W_T] = \frac{1}{\xi_t} E_t \left[\xi_T \theta + \left(\frac{B\gamma_2}{y\xi_T}\right)^{1/(1-\gamma_2)} \right], \quad (39)$$

where we substituted (21). As we used a Gram-Charlier expansion of the lognormal density function, the pricing kernel $\log(\xi_T)$ has the following density function:

$$f(z_T; \beta, k) = (1 + \beta(z_T^3 - 3z_T) + \delta(z_T^4 - 6z_T^2 + 3))\phi(z_T), \quad (40)$$

where $z_T = (\log(\xi_T) - \mu)/\sigma$, $\beta = \zeta/6$, $\delta = \chi/24$ and ζ, χ denote skewness and kurtosis of $\log(\xi_T)$ respectively, and $\mu = \log(\xi_t) - (r + \frac{1}{2}\|\kappa\|^2)(T - t)$, $\sigma = \|\kappa\|\sqrt{T - t}$ denote the mean and standard deviation of $\log(\xi_T)$. After some tedious but straightforward calculus we obtain (22).

- (iii) The proof is analogous to the proof of part (ii) in Proposition 3. Denote wealth by $W_t = F(t, \xi_t)$. The optimal fraction invested in the risky assets is given by:

$$w_t = -\frac{(\sigma')^{-1}\kappa'}{W_t} \left(\frac{\partial F(t, \xi_t)}{\partial \xi_t} \xi_t \right). \quad (41)$$

Substituting (22) in (41) yields the expression in (23). □

Proof of Proposition 7

- (i) The proof is analogous to the proof of part (i) in Proposition 3. Wealth at time t satisfies

$$W_t = \frac{1}{\xi_t} E_t[\xi_T W_T] = \frac{1}{\xi_t} E_t \left[\xi_T \left(\theta + \left(\frac{y\xi_T}{B\gamma_2}\right)^{1/(\gamma_2-1)} \right) \mathbf{1}_{\{\xi_T \leq \bar{\xi}\}} \right], \quad (42)$$

where we substituted (7). As we used a Gram-Charlier expansion of the lognormal density function, the pricing kernel $\log(\xi_T)$ has the following density function:

$$f(z_T; \beta, k) = (1 + \beta(z_T^3 - 3z_T) + \delta(z_T^4 - 6z_T^2 + 3))\phi(z_T), \quad (43)$$

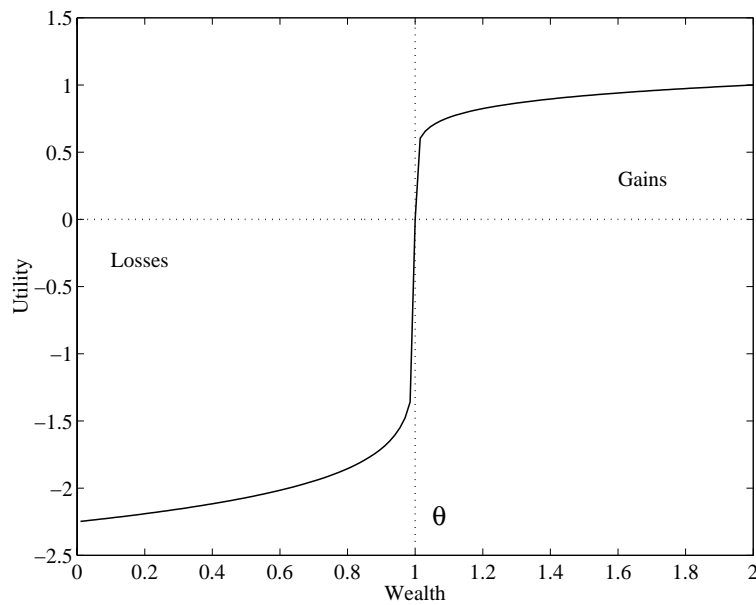
where $z_T = (\log(\xi_T) - \mu)/\sigma$, $\beta = \zeta/6$, $\delta = \chi/24$ and ζ, χ denote skewness and kurtosis of $\log(\xi_T)$ respectively, and $\mu = \log(\xi_t) - (r + \frac{1}{2}\|\kappa\|^2)(T - t)$, $\sigma = \|\kappa\|\sqrt{T - t}$ denote the mean and standard deviation of $\log(\xi_T)$. After some tedious but straightforward calculus we obtain (24).

- (ii) The proof is analogous to the proof of part (ii) in Proposition 3. Denote wealth by $W_t = F(t, \xi_t)$. The optimal fraction invested in the risky assets is given by:

$$w_t = -\frac{(\sigma')^{-1}\kappa'}{W_t} \left(\frac{\partial F(t, \xi_t)}{\partial \xi_t} \xi_t \right). \quad (44)$$

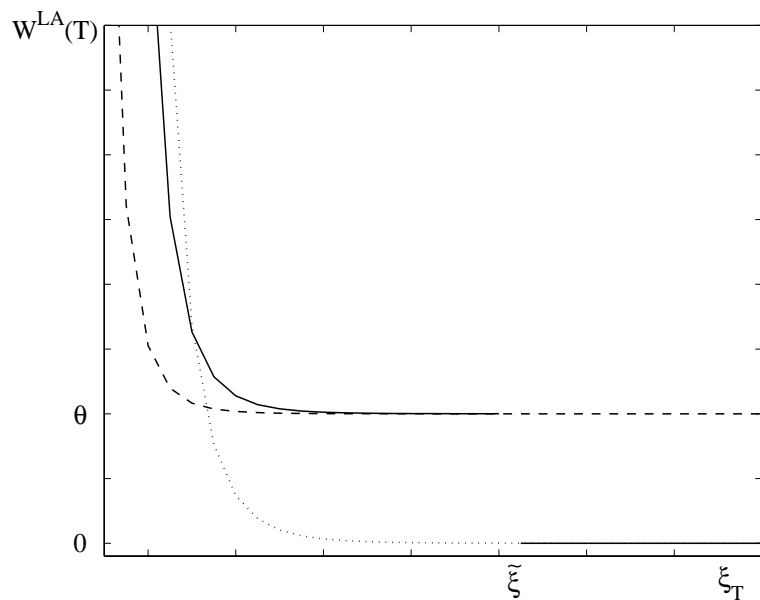
Substituting (24) in (44) yields the expression in (25). \square

Figure 1: Utility for Loss Averse Agent



This figure shows the value function of a loss-averse agent. The parameters used are $\gamma_1 = \gamma_2 = 0.88$, $A = 2.25$, $B = 1.0$, and $\theta = 1.0$.

Figure 2: Optimal Terminal Wealth



This figure shows the optimal terminal wealth of a loss-averse agent (solid plot), a HARA agent (dashed plot) and a CRRA agent (dotted plot).

Figure 3: Optimal Intermediate Wealth and Optimal Portfolio

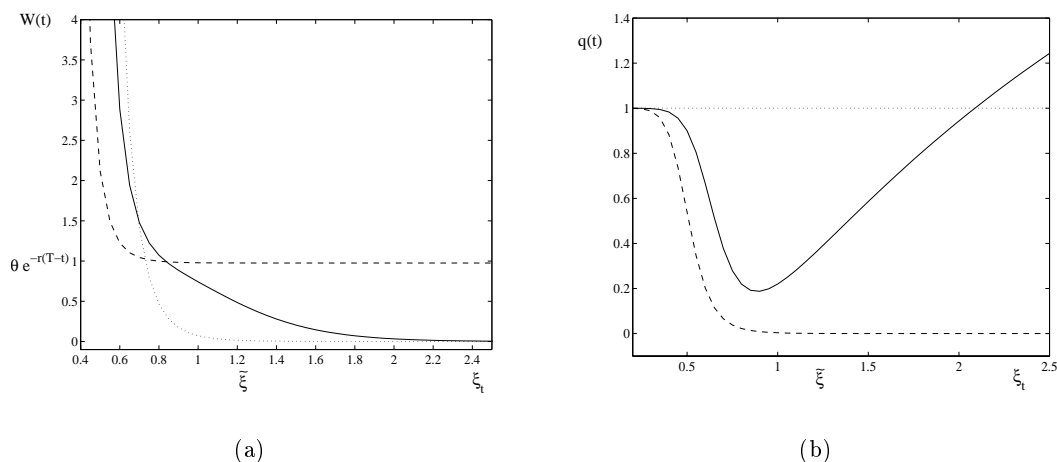
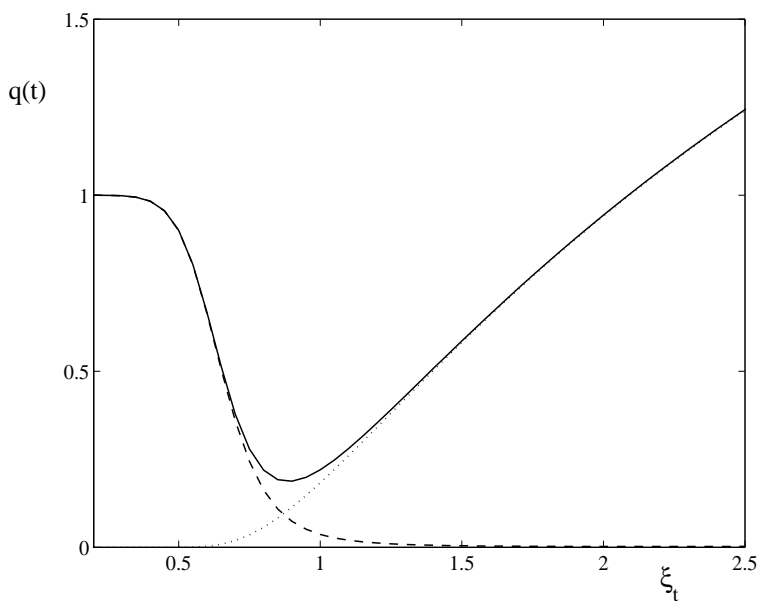


Figure (a) shows the optimal intermediate wealth of a loss averse agent (solid plot), a CRRA agent (dotted plot), and a HARA investor (dashed plot). Figure (b) shows the optimal fraction invested in stocks for a loss-averse investor relative to the fraction invested by the benchmark CRRA agent. The parameters used are $\gamma = \gamma_1 = \gamma_2 = 0.88$, $A = 2.25$, $B = 1.0$, $W_0 = 1.0$, $\theta = 1.0$, $r = 0.05$, $\|k\| = 0.4$, $T = 1$, $t = 0.5$, $\xi_0 = 1$. Then, $\tilde{\xi} = 1.21$.

Figure 4: Optimal Fraction Invested in Stocks



This figure shows the optimal fraction invested in stocks for a loss averse agent (solid plot), the probability maximizing strategy (dotted plot), and the growth strategy (dashed plot). The parameters used are $\gamma = \gamma_1 = \gamma_2 = 0.88$, $A = 2.25$, $B = 1.0$, $W_0 = 1.0$, $\theta = 1.0$, $r = 0.05$, $\|k\| = 0.4$, $T = 1$, $t = 0.5$, $\xi_0 = 1$.

Figure 5: Impact of Risk Aversion and Sharpe Ratio

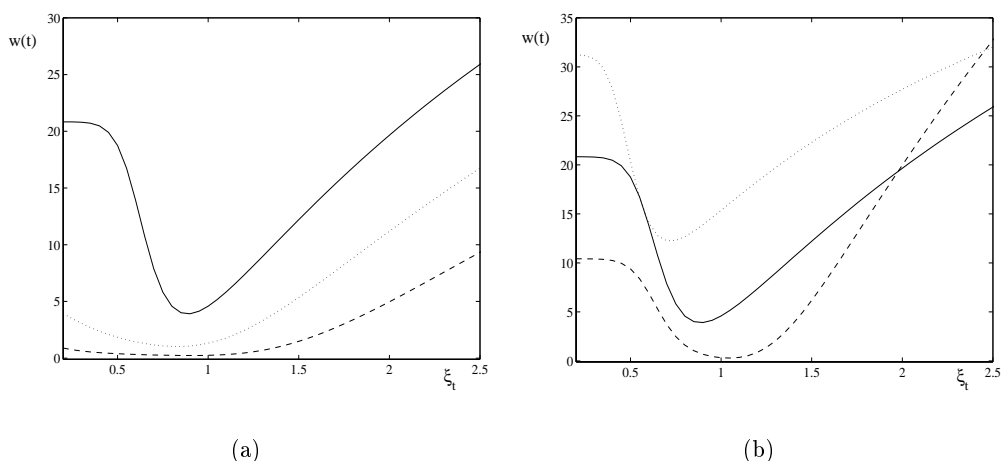
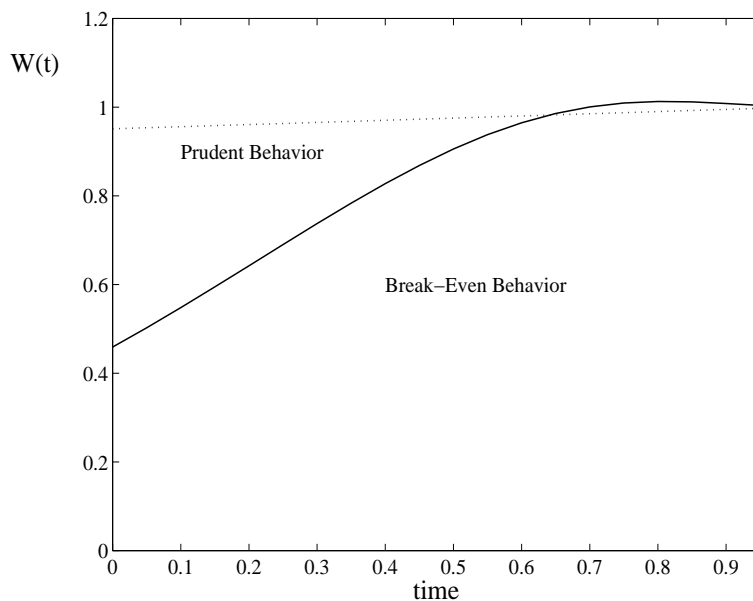


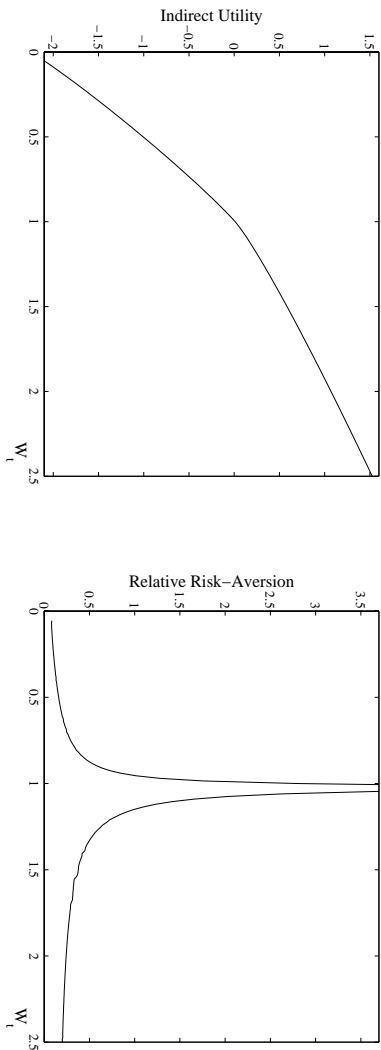
Figure (a) shows the optimal fraction invested in stocks for different risk aversion levels over gains: $\gamma_2 = 0.12$ (dashed plot), $\gamma_2 = 0.5$ (dotted plot), and $\gamma_2 = 0.88$ (solid plot). Figure (b) shows the optimal fraction invested in stocks for different Sharpe ratios: $\kappa = 0.2$ (dashed plot), $\kappa = 0.4$ (solid plot), and $\kappa = 0.6$ (dotted plot). The parameters used are $\gamma = \gamma_1 = \gamma_2 = 0.88$, $A = 2.25$, $B = 1.0$, $W_0 = 1.0$, $\theta = 1.0$, $r = 0.05$, $T = 1$, $t = 0.5$, $\xi_0 = 1$.

Figure 6: Critical Path



This figure shows the critical path (solid plot) and the satiation level of a probability maximizer (dotted plot). If wealth falls below the critical value break-even effects dominate the behavior of the loss averse investor. When wealth remains above the critical value the loss averse investor behaves as a portfolio insurer. The parameters used are $\gamma = \gamma_1 = \gamma_2 = 0.88$, $A = 2.25$, $B = 1.0$, $W_0 = 1.0$, $\theta = 1.0$, $r = 0.05$, $\|k\| = 0.4$, $T = 1$, $\xi_0 = 1$.

Figure 7: Indirect Utility

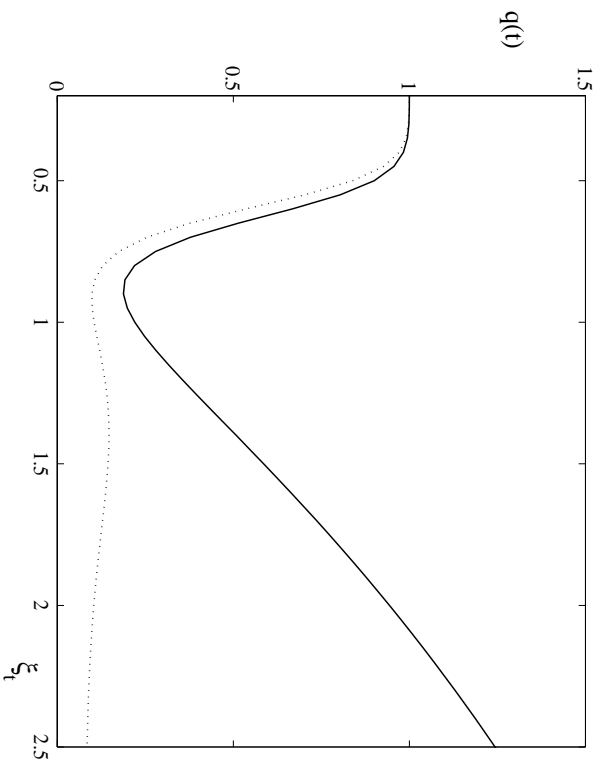


(a)

(b)

Figure (a) shows the indirect utility function at time $t = 0.9$. Figure (b) shows the relative risk aversion function of the indirect utility function at time $t = 0.9$. The parameters used are $\gamma = \gamma_1 = \gamma_2 = 0.88$, $A = 2.25$, $B = 1.0$, $W_0 = 1.0$, $\theta = 1.0$, $r = 0.05$, $\|k\| = 0.4$, $T = 1$, $\xi_0 = 1$.

Figure 8: Optimal Fraction Invested in Stocks: Extended Loss Aversion



This figure shows the optimal fraction invested in stocks for a loss averse agent (solid plot), and an extended loss averse agent (dotted plot). The parameters used are $\gamma = -1.0$, $\gamma_1 = \gamma_2 = 0.88$, $A = 2.25$, $B = 1.0$, $W_0 = 1.0$, $\theta = 1.0$, $\underline{W} = 0.5$, $r = 0.05$, $\|k\| = 0.4$, $T = 1$, $t = 0.5$, $\xi_0 = 1$.

Figure 9: Impact of Skewness and Kurtosis

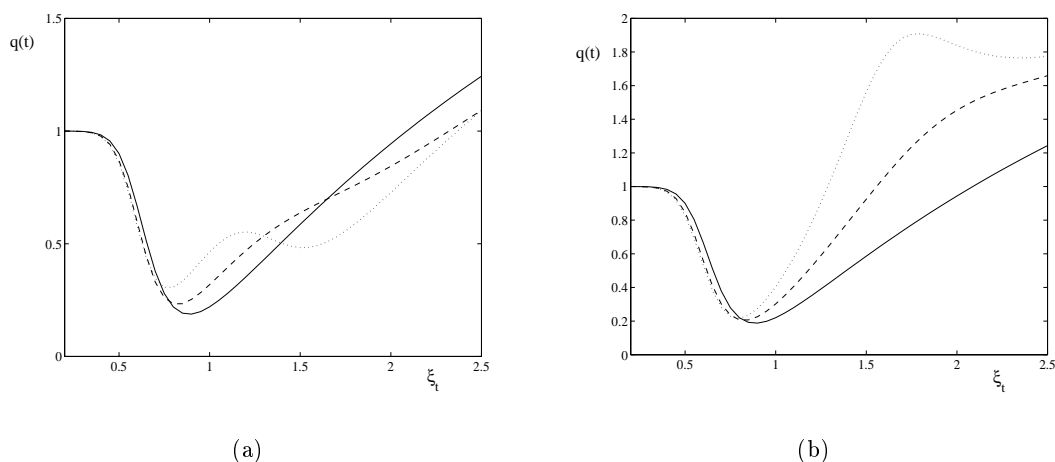
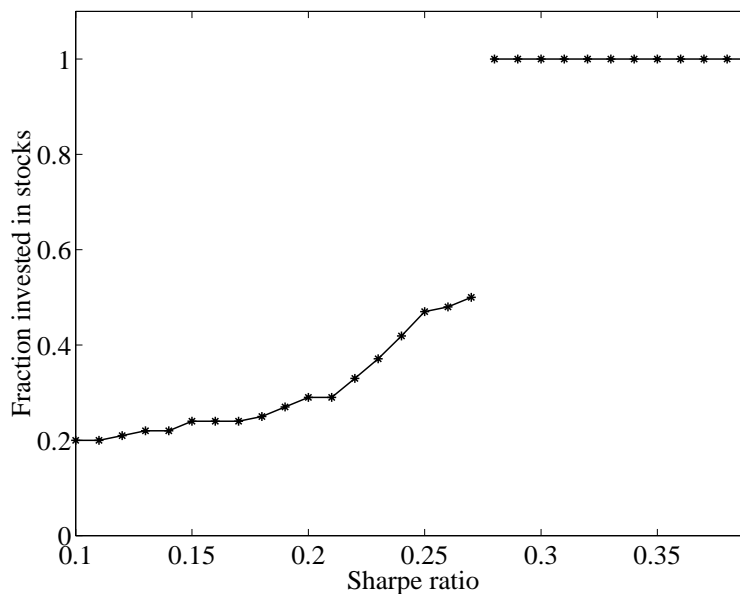


Figure (a) shows the optimal fraction invested in stocks relative to the fraction invested by the benchmark CRRRA agent for different values for skewness: $\zeta = 0.0$ (solid plot), $\zeta = -0.4$ (dashed plot), $\zeta = -1.2$ (dotted plot). Figure (b) shows the optimal fraction invested in stocks relative to the fraction invested by the benchmark CRRRA agent for different values of excess kurtosis: $\chi = 0.0$ (solid plot), $\chi = 0.7$ (dashed plot), $\chi = 1.7$ (dotted plot). The parameters used are $\gamma = \gamma_1 = \gamma_2 = 0.88$, $A = 2.25$, $B = 1.0$, $W_0 = 1.0$, $\theta = 1.0$, $r = 0.05$, $\|k\| = 0.4$, $T = 1$, $\xi_0 = 1$.

Figure 10: Optimal Fraction Invested in Stocks: Buy and Hold Strategy



This figure shows the optimal fraction invested in stocks in a buy-and-hold strategy for a loss averse agent as a function of the Sharpe ratio. The parameters used are $\gamma_1 = \gamma_2 = 0.88$, $A = 2.25$, $B = 1.0$, $W_0 = 1.0$, $\theta = 1.0$, $r = 0.05$, $T = 1$.