Risk-taking under a Punishing Bailout

Alexander K. Zentefis*

Yale School of Management
alexander.zentefis@yale.edu

September 17, 2022

Abstract

An outside stakeholder’s commitment to bail out a levered investment fund might encourage the fund manager to take extreme risks. But what if the bailout were harsh? The loan would have a penalty rate and the manager would have to sacrifice part equity interest in the fund. I show that as the fund value declines, the manager actually reduces risk to avoid the rescue—unless failure is imminent. In that case, the manager increases risk immensely to profit from no further loss, which makes the bailout more likely. The paper informs which terms can be successful in bailout policy design.

JEL classification: G11, G23, G28

Keywords: risk-taking, portfolio choice, bailout

*Corresponding author. Address: 165 Whitney Ave, New Haven, CT 06511. Email: alexander.zentefis@yale.edu. Tel: +1 203 432 3628. I am grateful for very helpful comments from John Cochrane, Doug Diamond, Varun Gupta, Gary Gorton, Lars Hansen, Yunzhi Hu, Paymon Khorrami, Bora Keskin, Stavros Panageas, Aaron Pancost, Jung Sakong, Fabrice Tourre, Pietro Veronesi, and seminar participants at Chicago Booth.
1 Introduction

This paper analyzes the optimal portfolio choice of a levered investment fund—such as a hedge fund or a financial institution’s proprietary trading desk—that has an implicit, but punishing, bailout guarantee. If the fund is at the brink of default, the manager knows that an outside stakeholder—such as the parent institution, the parent fund, or the national government—is committed to bailing out the levered fund with emergency lending. However, the loans are supplied at a penalty rate, and in exchange for the bailout, the manager must forfeit a fraction of his or her equity interest in the fund to the outside stakeholder.

Famous examples of levered investment fund bailouts are those of LTCM in 1998, a Bear Stearns hedge fund in 2007, and AIG in 2008. For LTCM: the Federal Reserve Bank of New York organized over a dozen lenders to provide $3.6 billion to save the hedge fund. In exchange, the fund partners sacrificed ninety-percent of their share to the participating lenders (Lowenstein, 2000). For the Bear Stearns hedge fund: the parent investment bank pledged up to $3.2 billion in loans to prevent the fund’s collapse after bad investments in subprime loans. The senior executive responsible for the fund was soon ousted (Creswell and Bajaj, 2007; Bajaj, 2007). For AIG: the U.S. Federal Reserve enforced a penalty rate of interest and the relinquishing of equity ownership as terms for the emergency credit facility extended to the firm in 2008 to avert the company’s bankruptcy (Sjostrom Jr., 2009).

In this paper, I ask: if a levered fund manager knows that he or she will be bailed out, but on punishing terms, what investment policy will the person adopt ex-ante? Because of the bailout guarantee, one might suspect that the manager would shift risk to the outside stakeholder by raising portfolio volatility at all times. This is not the case. The manager’s policy depends on the fund’s state: When the fund performs well, the manager moves wealth into a risky asset, but not by a drastic amount. When the fund declines in value, the manager actually reduces risky holdings to avoid the bailout. The more the fund loses, the more the manager curtails risk. Despite the bailout guarantee, the manager does not
engage in moral hazard.

However, if portfolio wealth declines substantially to near default, the manager does shift risk. At that point, the prospect of no further loss from the bailout incites the manager to take on maximal risk. The threshold at which the manager initiates the extreme risk-taking is greater than the fund’s total liabilities. Hence, while the fund has not yet defaulted, the manager’s behavior increases the probability the fund will default and require a bailout.

Why does the manager display this kind of investment behavior? Risk aversion, the sacrifice of equity ownership, and the obligation to repay the costly bailout money lead the manager to mitigate risk as the value of the fund declines. The fund’s total indebtedness becomes a form of endogenous “subsistence” or “habit” wealth, which the manager is compelled to prevent the fund from falling below. Similar to dynamic investment models under habit formation preferences (Ingersoll, 1992) or subsistence consumption (Dybvig, 1995; Presman and Taksar, 1992), the dollar amount invested in the risky asset is proportional to wealth less the subsistence amount. As wealth declines, the fraction invested in the risky asset decreases. This strategy explains why the money manager in the model increases risk gradually as portfolio wealth increases, but reduces risky holdings as the fund value shrinks.

Because multiple bailouts are possible—the fund can default multiple times—the manager’s habit amount in the model is not fixed, but stochastic. This stochastic habit increases in the cumulative amount of bailout loans. Because these loans reflect the historical path of portfolio wealth, another interpretation of this habit amount is a stock of “bad habits.” The manager’s poor historical performance that drove the fund to a bailout raises the bad habit stock. Given the manager’s risk aversion, he or she manages the fund to trim risk-taking in bad times to avoid another bailout and “falling back into bad habits.”

In standard habit or subsistence models, if portfolio wealth comes very close to the subsistence level, the manager would have divested from the risky asset to ensure that wealth declines no further. In the model, however, the manager possess inalienable wealth that is separated from the fund. This external wealth may be housing, land, boats, or
personal financial assets. The outside wealth stops the manager from shedding risk completely as the fund declines in value. Instead, the incentive to mitigate risk gradually wanes as the manager’s remaining interest in the fund net of liabilities falls to zero. At that point, anticipation of unlimited gains with no risk of loss stirs the manager to take on maximal risk.

Repeated bailouts increase the manager’s indebtedness to the outside stakeholder, which raises the bad habit level. This relation implies that portfolio wealth would have to fall by less in the future to dip below the higher threshold and trigger the extreme risk-taking. This greater risk-taking at higher portfolio wealth levels makes the fund more vulnerable to loss, default, and additional bailouts. A bailout today makes a bailout in the future more likely.

The bailout guarantee in the model creates a convex payoff for the fund manager, which connects this research to the literature on investment behavior of money managers who face convex compensation structures (e.g., Chevalier and Ellison, 1997; Carpenter, 2000; Goetzmann, Ingersoll Jr and Ross, 2003; Basak, Pavlova and Shapiro, 2007; Hodder and Jackwerth, 2007; Panageas and Westerfield, 2009). The bailout setting and the nuanced investment strategy displayed in this paper is distinct from existing work. A notable comparison is to Hodder and Jackwerth (2007), which also features a manager taking high-risk bets near an endogenous lower boundary. However, risk-taking near that boundary diminishes as the manager’s investment problem lengthens in time—unlike in this paper, where high risk-taking is present for wealth near the subsistence boundary even in an infinite horizon problem.

One might look to the terms of the bailout to learn if they are effective at stopping the extreme risk-taking as the portfolio value nears the manager’s subsistence. It turns out that both the equity dilution and the penalty rate of interest do little to curtail the high-risk strategy when portfolio value is low enough. This suggests that additional terms, such as direct portfolio constraints when the fund suffers large losses, would be necessary.
2 Model

A risk-averse levered fund manager engages in a dynamic investment problem. The manager’s goal is to maximize total expected utility of final wealth at a random stopping time. The fund has a fixed liability $L$ that is constant throughout time. The manager pays interest continuously on the liability with principal due at the fund’s termination. The fund begins with initial wealth $W_0 > L$ at time 0 and the manager makes investment decisions in continuous time.

2.1 Investment opportunities

Two distinct assets are available to the manager. One has rate of return equal to the risk-free interest rate, which has constant rate of return $r$. The other asset is risky, whose price per share obeys a geometric Brownian motion. The instantaneous returns of the two assets are

$$\frac{dP_t^0}{P_t^0} = r dt$$

$$\frac{dP_t^1}{P_t^1} = \mu dt + \sigma dZ_t,$$

where $\{Z_t, t \geq 0\}$ is a standard one-dimensional Brownian process on a complete probability space $\{\Omega, \mathcal{F}, \mathbb{P}\}$, while $\mu > r$ and $\sigma > 0$ are constants. Let $\mathcal{F} = \{\mathcal{F}_t, t \geq 0\}$ denote the right-continuous filtration of the sigma-algebra $\mathcal{F}$ generated by $\{Z_t, t \geq 0\}$.

2.2 Investment policy and wealth process

The manager controls a portfolio investment policy $\{\alpha_t \in [0, \infty), t \geq 0\}$, which denotes the fraction of total fund assets invested in the risky asset. The remaining fraction of wealth $(1 - \alpha_t)$ is invested in the risk-free asset. So that the results are independent of time, the fund liquidates at an exogenous random stopping time, denoted by $\tau$. The stopping time is exponentially distributed with constant intensity $\lambda > 0$. At the time of liquidation, funds net of liabilities are distributed to the manager, the decision problem ends, and the continuation value is zero.
The manager services the debt by paying a flow rate of interest \( r_L \). The interest rate on the loan is the risk-free rate because the bailout provision makes the debt riskless. The creditors can terminate the fund if assets under management fall below liabilities, or \( W_t < L \). In this circumstance, however, the fund manager has an unconditional bailout from the outside stakeholder. The bailout takes the form of incremental transfers \( dG_t \geq 0 \) to the manager once \( W_t = L \). These transfers ensure that \( W_t \geq L \) for all \( t \geq 0 \). The unique minimal process to enforce \( W_t \geq L \) is

\[
\int_{s=0}^{t} \frac{dG_s}{L} = \max \left[ 0, \max_{0 \leq s \leq t} \left[ \log L - (\log W_0 + H_s) \right] \right],
\]

where the process \( H_s \) is defined as

\[
H_s \equiv \int_{u=0}^{s} \left( r + \alpha_u (\mu - r) - rL \right) du - \frac{1}{2} \int_{u=0}^{s} \alpha_u^2 \sigma^2 du + \int_{u=0}^{s} \alpha_u \sigma dZ_u.
\]

Karatzas and Shreve (1991) (p. 210-211) discuss the minimal process \( G_t \). Intuitively, the transfers \( \{G_t, t \geq 0\} \) support the wealth process in such a way that each time \( W_t \) falls by an amount \( \varepsilon > 0 \) below \( L \), an offsetting amount \( \varepsilon \) is transferred to the fund, so that \( W_t \geq L \) for all \( t \). The process defined above is the unique minimal one to accomplish this. This transfer process \( G_t \) takes the value of 0 at time zero, is non-decreasing and can only increase when wealth is on the liability boundary \( L \), i.e. \( W_t = L \).

With the bailout transfers available to the fund, the evolution of assets is

\[
dW_t = \left[ W_t (\alpha_t \mu + (1 - \alpha_t) r) - rL \right] dt + W_t \alpha_t \sigma dZ_t + dG_t.
\]

The bailout transfers are incremental loans distributed to the fund, secured by the assets under management. The cost of the bailout loans to the fund is at a penalty rate of interest \( r_g = r + r_p \), with \( 0 < r_p < 1 \). Interest and principal \( (1 + r_g) G_\tau \) are payable to the stakeholder at the termination date \( \tau \). At the instant the manager receives the first bailout loan, the manager sacrifices a fraction \( 0 < x < 1 \) of equity interest in the fund to the outside
stakeholder.

The manager is not restricted to keep net capital $W_t - L - G_t$ strictly positive at all times. The stakeholder transfers ensure that $W_t \geq L$, so the original liabilities are riskless. The bailout loans, however, are subordinate to the original liabilities, and there is no guarantee that the stakeholder is repaid in full for its transfers. The bailout shifts default risk from the fund’s creditors to the stakeholder. Table 1 presents the distribution of funds at the liquidating date $\tau$. The gross penalty interest rate $R_g = \left(1 + r_g\right)$.

Table 1: Distribution of Payoffs at $\tau$

<table>
<thead>
<tr>
<th>Entity</th>
<th>Payoff (if bailout)</th>
<th>Payoff (if no bailout)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Creditors</td>
<td>$L$</td>
<td>$L$</td>
</tr>
<tr>
<td>Stakeholder</td>
<td>$\min\left{W_\tau - L, R_g G_\tau\right}$</td>
<td>$0$</td>
</tr>
<tr>
<td></td>
<td>$\left(\text{subordinated debt payoff}\right)$</td>
<td>$\left(\text{equity payoff}\right)$</td>
</tr>
<tr>
<td></td>
<td>$+ \max{0, x \left(W_\tau - L - R_g G_\tau\right)}$</td>
<td></td>
</tr>
<tr>
<td>Manager</td>
<td>$\max{0, (1-x) \left(W_\tau - L - R_g G_\tau\right)}$</td>
<td>$W_\tau - L$</td>
</tr>
</tbody>
</table>

Notes: The table presents the payoffs to creditors, the outside stakeholder, and the manager at the end of the investment period $\tau$ when both a bailout occurs and a bailout does not occur. In both scenarios, creditors receive their principal amount. The outside stakeholder receives an equity and subordinated debt payoff if it bails out the manager and nothing if it does not. The manager receives the full net worth of the fund if bailed out and a fraction of that amount if not bailed out.

The fund’s creditors always and only receive $L$. If there is a bailout, the outside stakeholder receives a subordinated debt plus equity payoff. The manager keeps the residual value of the fund. Should $W_t > L$ for all $t \leq \tau$—and thus there is no bailout—then at time $\tau$, the creditors receive $L$, the manager receives the residual $W_\tau - L$, and the stakeholder receives nothing.

All of the manager’s time-varying wealth is in the fund. For hedge funds, Fung and Hsieh (1999) document that managers have a substantial personal investment in their funds. Having the manager in the model be compensated instead with a fraction of equity ownership is more realistic. Nevertheless, adjusting the compensation structure that way
makes no difference—doing so is isomorphic to changing the parameter value for the sacrificed equity fraction \( x \) after a bailout (or imposing a sacrificed fraction even when \( G_\tau = 0 \)). In section 6, I present the optimal investment strategy results for varying levels of \( x \), including when \( G_\tau = 0 \).

### 2.3 Portfolio problem

The manager chooses an investment policy \( \{\alpha_t, \ t \geq 0\} \) to maximize utility over net worth at time \( \tau \). The manager has CRRA utility with relative risk aversion coefficient \( \gamma > 0, \gamma \neq 1 \). The manager’s value function for all \( t \leq \tau \) is

\[
V(G_t, W_t) = \sup_{\alpha(G_t, W_t)} \mathbb{E}_t \left[ \left( a + (1 - x \cdot 1_{\{G_\tau > 0\}}) (W_\tau - L - R_g G_\tau)^+ \right)^{1-\gamma} \right],
\]

where \((\cdot)^+\) denotes the positive part of \((\cdot)\), and where \(1_{\{G_\tau > 0\}}\) is an indicator function that is 1 if bailout transfers \( G_\tau \) are positive and zero otherwise.

The constant \( 0 < a < \infty \) is the manager’s inalienable wealth that is separate from the assets in the portfolio and protected from outside creditors (e.g., personal housing or investment accounts). The inalienable wealth is necessary for a bailout to occur at all. To see why, suppose \( a = 0 \). In this case, the manager’s utility can be written as

\[
U(G_t, W_t) = \mathbb{E}_t \left[ \frac{(W_\tau - \omega_\tau^*)^{1-\gamma}}{1-\gamma} \right],
\]

where

\[
\omega_\tau^* \equiv x \cdot 1_{\{G_\tau > 0\}} W_\tau + (1 - x \cdot 1_{\{G_\tau > 0\}}) (L + R_g G_\tau).
\]

The amount \( \omega_\tau^* \) can be treated as a “subsistence wealth” that is stochastic. Because marginal utility at the subsistence level is infinite, the manager would ensure that the portfolio value never declined to this amount (see Dybvig, 1995; Ingersoll, 1992). Because \( G_t \geq 0 \) for all \( t \), this investment policy would even maintain the portfolio value above \( W_t = L \). Hence, the fund would never default, and there would never be a need for a bailout.
For a small time interval $\delta > 0$, the value function in (6) can be recast as

$$Q(G_t, W_t) = \sup_{\alpha(G_t, W_t)} \mathbb{E}_t \left[ \int_0^\delta e^{-\lambda u} C(G_{t+u}, W_{t+u}) du \right] + e^{-\lambda \delta} \mathbb{E}_t [Q(G_{t+\delta}, W_{t+\delta})],$$

(9)

with

$$C(G, W) \equiv \frac{\left( a + (1 - x \cdot 1_{\{G > 0\}}) \left( W - L - R_g G \right) \right)^{1-\gamma}}{1 - \gamma},$$

(10)

and $Q(G_t, W_t) = \frac{V(G_t, W_t)}{\lambda}$ is the discounted value function. I solve the model using the formulation of the manager’s problem in (9).

### 3 Solution

To determine the manager’s discounted value function $Q$ and optimal investment policy $\alpha^* (G_t, W_t)$, I numerically solve a discrete state version of the control problem in (9) whose solution converges to the solution of the original problem. I use a Markov chain approximation method akin to the one described in Kushner and Dupuis (1992). The numerical method transforms the original control problem into a simpler one by approximating the stochastic process $\{(G_t, W_t), t \geq 0\}$ with a suitable Markov chain process controlled on a finite state space. This finite state space is a discretization of the original state space.

This approximation method converts the original continuous time dynamic investment problem of the manager into a discrete state dynamic programming problem in the following matrix form:

$$Q^h = \sup_{\alpha(G, W) \in A} R^h(\alpha) Q^h + J^h,$$

(11)

where the vector $Q^h$ is the value function of the discrete states, the matrix $R^h(\alpha)$ is a discounted probability transition matrix over the states, the vector $J^h$ is the “flow” utility of the manager over an incremental time period independent of the investment policy, and $A$ is a compact set. Each object is indexed by a discretization level $h > 0$ of the state space. The manager’s problem reduces to optimally controlling the transition probabilities between the Markov states in wealth and bailout transfers. The solution $Q^h$ of this dynamic
program converges to the value function from the original control problem in (9) under conditions the manager’s problem satisfies. Hence, the optimal investment policy from the Markov chain problem is a good approximation to that of the original problem. The full details of this representation of the manager’s problem and the numerical procedure used to solve it are presented in the Appendix.

3.1 Baseline parameters

Baseline parameters of the model are provided in Table 2.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Parameter</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r$</td>
<td>0.05</td>
</tr>
<tr>
<td>$r_p$</td>
<td>0.085</td>
</tr>
<tr>
<td>$\mu$</td>
<td>0.11</td>
</tr>
<tr>
<td>$x$</td>
<td>0.8</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>0.20</td>
</tr>
<tr>
<td>$a$</td>
<td>0.1</td>
</tr>
<tr>
<td>$L$</td>
<td>0.5</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>.01</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>3</td>
</tr>
</tbody>
</table>

Notes: The table presents the baseline parameter values of the model. The interest rate $r$ is set to 5%, whereas the expected return $\mu$ and volatility $\sigma$ of the risky asset are set to 11% and 20%, respectively. Together, these values generate a Sharpe ratio of 0.3, which is similar to the historical average for the U.S. equity market. The principal on the fund’s debt $L$ is set to make the fund levered 2-to-1 at the start. The manager’s risk aversion $\gamma$ is 3. The penalty rate of interest on the stakeholder bailout loan $r_p$ and the equity dilution fraction $x$ match the terms of the AIG bailout as described in Sjostrom Jr. (2009). Inalienable wealth is set to 10% of the fund starting value, and the manager’s implied subjective rate of time preference $\lambda$ is 1%.

The interest rate $r$ is set to 5%, which matches the historical average of the U.S. 1-year Treasury rate. The annualized mean return $\mu$ on the risky asset implies a risk premium of 6%. The annualized volatility $\sigma$ of the risky asset matches a Sharpe ratio of 0.3, which is similar to the historical average for the U.S. equity market. The equity dilution percentage $x$ is 80%, whereas the penalty rate above risk-free interest is 8.5%. These parameters match the terms the U.S. Federal Reserve enforced with AIG, which were made public (Sjostrom Jr., 2009).

The indebtedness $L$ of the fund yields a leverage ratio of 2 at starting fund wealth $W_0 = 1$. (The absolute level of the fund’s debt will not matter in the manager’s portfolio choice, but...
rather, the portfolio leverage \( \frac{W}{W - L} \). The manager’s inalienable wealth \( a \) represents 10% of the starting value of the fund. Finally, the coefficient of relative risk aversion is 3, whereas the mean rate of fund liquidation \( \lambda \) implies a subjective rate of time preference for the manager, which I set to 1%. The effect on the optimal investment policy from changes in the risky asset’s volatility \( \sigma \), the manager’s risk aversion \( \gamma \), the level of inalienable wealth \( a \), and the bailout terms available to the outside stakeholder in both \( x \) and \( r_p \) are given in Sections 5 and 6.

### 4 Main Results

The numerical approximation to the portfolio manager’s optimal investment policy under the baseline parameters is presented in Figure 1.

Figure 1: Optimal Investment Policy, Baseline Parameters

![Figure 1: Optimal Investment Policy, Baseline Parameters](image)

**Notes:**
- The figure illustrates the optimal fraction invested in the risky asset at different values of portfolio wealth for two separate cumulative amounts of bailout money. The value \( \alpha^* \) is decreasing in wealth, but rapidly increases once wealth has fallen below a threshold approaching the total indebtedness of the fund.
- The figure depicts the manager’s optimal investment fraction \( \alpha^* \) in the risky asset for
varying levels of fund wealth $W$ when the cumulative bailout money $G$ is at two separate values, $G = 0$ and $G = G' = 0.8$. The value of $G'$ is right below the maximum transfers $G$ established in the bounded state space used in the numerical routine. The total obligations of the fund ($L$ and $L + G'$) are marked in the figure with thicker ticks on the x-axis in wealth.

For levels of wealth sufficiently far from total obligations, the optimal fraction invested in the risky asset is increasing in wealth. In good times, the manager engages in greater risk-taking. Even so, the optimal policy is to invest less in the risky asset than the fraction $\frac{\mu - r}{\gamma \sigma^2}$ determined in Merton (1969), which is marked in the figure by a solid horizontal line.

On the other hand, in bad times when portfolio wealth declines, the manager reduces exposure to the risky asset. If portfolio wealth declines far enough to near the fund’s total indebtedness, the manager reverses, and rapidly increases the proportion of the fund in the risky asset to the allowable upper bound, which here is imposed to be $\alpha = 200\%$. Moral hazard only appears when the fund is near default (i.e., $W = L$). However, when the fund has already been bailed out in the past ($G = G'$), the manager begins shifting risk to the outside stakeholder well before wealth is even close to default. The manager’s history of bailouts influences the investment policy in bad times when the fund does poorly.

4.1 Interpreting the results

The manager faces a trade-off between honoring the repayment of total debt and taking advantage of expected profit opportunities in the risky asset. Because the manager is risk averse, indebtedness becomes a kind of subsistence amount which the manager would prevent the fund from dipping below if it were not for the inalienable wealth. The optimal policy here is similar to that of subsistence consumption (Dybvig, 1995; Presman and Taksar, 1992), which is to invest a sufficient quantity in the risk-free asset that would strictly maintain wealth above the subsistence threshold. Here, this quantity is not fixed, but stochastic and increasing with the cumulative amount of bailout transfers $G$. This investment policy explains why the fraction in the risky asset is lower for $G = G'$ than for $G = 0$. The amount necessary to satisfy subsistence is larger in the former case, requiring a
larger share invested in the riskless asset for any fixed $W$.

### 4.1.1 Endogenous habit preferences

The dependence of $G$ on the path of wealth can be interpreted as introducing a kind of habit formation preference akin to Ingersoll (1992). But here, the preference is with respect to avoiding “bad habits.” One can see this connection from the manager’s utility function:

$$
U(G, W) = \left( a + (1 - x \cdot 1_{[G>0]})(W - L - R_g G)^{+}\right)^{1-\gamma}.
$$

(12)

The amount $G$ is the cumulative total the fund has needed with urgency to avoid default because of previous poor performance. By the terms of the bailout, this amount must eventually be deducted from portfolio wealth. An intertemporal complementarity exists between past and current wealth. The worse the manager performed in the past by dragging the fund to need a bailout, the greater is $G$, and the more regret this past performance brings. The penalty rate of interest measures the strength of this regret. Under this habit preference, the manager invests in a way to keep clear of “re-living” the previous bad habits by ensuring that wealth $W$ always exceeded the habit amount through an adequate investment in the riskless asset.

Interestingly, this habit amount—the level of wealth immediately before $\alpha$ abruptly increases in Figure 1—is greater than the total indebtedness $L + R_g G$. There are two reasons for this difference. The first reason is that the current level of bailout transfers $G$ affects transfers in the future. In fact, in the next section I describe how greater transfers today imply additional transfers in the future are more likely. Therefore, the habit amount reflects not only the current level of transfers, but also the present value of all expected future transfers. The second reason is that the manager’s inalienable wealth $a$ distorts the habit amount upward. One can see this distortion in the manager’s utility function (7) by adding $a$ to $\omega^+$. The manager has less at stake with greater outside wealth, so the extreme risk-taking occurs earlier.
As in other habit and subsistence models, the dollar amount invested in the risky asset is proportional to wealth less the subsistence amount; as wealth decreases, the fraction $\alpha$ invested in the risky asset declines. That behavior is displayed in Figure 1. Conversely, as wealth increases, $\alpha$ increases. For very large values of wealth $W$ the proportion in the risky asset approaches the fraction from Merton (1969).

In the standard habit or subsistence model, as wealth reached the subsistence level, the manager would have completely divested from the risky asset to ensure that wealth could decline no further. This is not what we see here. The manager instead abruptly reverses and increases risk to its limit. The combination of inalienable wealth, limited liability, and a floor on further losses explains this risk-taking behavior, which I describe next.

### 4.1.2 Roles of inalienable wealth, limited liability, and no further loss

First, the inalienable wealth ensures a minimum level of utility above zero so that marginal utility at wealth equaling subsistence is not infinite. Without the inalienable wealth, the risk aversion of the manager would induce him or her to maintain wealth above the subsistence amount. Wealth would never decline enough to trigger a bailout. With the inalienable wealth, this compulsion is no longer present. Second, limited liability permits the manager to possess inalienable wealth protected from creditors. As described earlier, this inalienable wealth makes the fund susceptible to failure. Third, the bailout guarantees the portfolio value can bear no further decline. At this point, an incremental investment in the risky asset yields the manager no loss and unlimited gains. Thus, the manager shifts the risk to the outside stakeholder and invests maximally in the risky asset.

Interestingly, the manager starts increasing risk after the fund has declined to near bankruptcy, but before actually reaching it. This behavior is due to the fund wealth already having declined below the manager’s subsistence amount. In addition, the threshold of wealth at which the manager commences the excessive risk-taking increases with greater bailout assistance (larger $G$). This high-risk investment raises the probability that the portfolio earns enough to surpass the subsistence amount, which would return the manager’s risk-taking to normal levels. The manager gambles for resurrection: a
good outcome would put the manager far above default and cancel any need for further assistance; a bad outcome bears large losses, which forces the outside stakeholder to step in again with more transfers. Because the high risk-taking starts well before the point of default, current bailouts would foreshadow future bailouts.

Overall, the terms of the punishing bailout—required repayment on any emergency bailout lending plus a sacrifice of equity ownership—induce the portfolio manager to reduce risk-taking as wealth declines. The lower risk-taking mitigates the chances of an initial bailout or further bailout money. However, if wealth declined to make bankruptcy imminent and the manager’s remaining equity in the fund is virtually zero, then one can expect the manager to abruptly accelerate risk-taking. This strategy increases the chance the stakeholder would have to step in—for the first time or again.

4.2 Comparison to Panageas (2010)

The result that in the presence of bailouts, portfolio risk-taking increases in good times, but decreases in bad times is similar in spirit to Panageas (2010). But the reason is very different. There, a risk neutral manager whose fund carries a bailout guarantee from an outside stakeholder adopts a bang-bang investment policy: the manager takes on maximal risk until net worth declines to a particular threshold, at which point the manager reverses, and tightens risk-taking to its lowest limit. This threshold is determined endogenously so that the bailout does not become so costly that the outside stakeholder reneges on its commitment. A depiction of the qualitative behavior of the optimal investment policy in Panageas (2010) is given in Figure 2.

In this paper, risk aversion and an obligation to repay the bailout money (missing in Panageas, 2010) induce the manager to make smoother adjustments in risk-taking policies between good and bad times. Curtailing risk in bad times is also undertaken voluntarily by the manager despite an accessibility to committed bailout money. In Panageas (2010), an unconditional bailout would incite maximal risk-taking at all times. Here, the prospect of paying a penalty for the assistance encourages the manager to attenuate risk, so long as
Notes: The figure illustrates the optimal fraction invested in the risky asset at different values of portfolio wealth. In this paper, the value $\alpha^*$ decreases in wealth, but rapidly increases once wealth has fallen below a threshold approaching the total indebtedness of the fund. In Panageas (2010), the optimal investment policy takes maximal risk in good times, but switches to minimal risk—here set to zero—if wealth declines to some endogenous level—here denoted $W^*$.

A major distinction between this paper and Panageas (2010) is the manager’s optimal investment behavior as wealth approaches the bailout threshold. In Panageas (2010), risk-taking would have already plummeted by this point to keep the stakeholder from reneging on the bailout. Here, once wealth falls below the subsistence threshold, the fund bears no further loss, and the credible promise of a fully committed bailout prompts the manager to take maximal risk. In scenarios where both sides are fairly confident that a bailout is guaranteed—such as saving a firm’s reputation in the case of Bear Stearns or averting systemic externalities in the case of LTCM and AIG—one should anticipate the investment policy in this paper rather than that in Panageas (2010).
5 Comparative Statics

The optimal investment policies under varying levels of risky asset volatility $\sigma$ are depicted in Figure 5. Total transfers $G$ are set to zero, so this figure illustrates the optimal investment behavior before any bailout has taken place. The optimal policy over wealth behaves very similarly to the baseline case: risk decreases as wealth declines until spiking rapidly once wealth falls below a subsistence amount near the bailout point $L$. Investment in the risky asset decreases in $\sigma$ because that asset becomes less and less attractive for a risk averse manager.

The increase in risk-taking occurs earliest for the highest volatility case. This behavior is due to a larger subsistence threshold being met sooner than in the other cases as wealth declines. Despite total indebtedness being the same in all three cases, the prospect of future bailouts is greater for any fixed investment $\alpha$ in the risky asset when volatility is higher. The present value of expected future bailout transfers is therefore larger, which increases the subsistence amount. A higher subsistence in turn means wealth drops by less before triggering extreme risk-taking.

Figure 5 presents the optimal investment policy under different levels of risk aversion $\gamma$. The fraction in the risky asset is decreasing in risk aversion. Furthermore, the more risk averse the manager, the later he or she engages in maximal risk-taking as wealth approaches the bailout threshold. A safer portfolio by a more risk averse manager reduces the present value of expected transfers, which shrinks the subsistence amount and delays the extreme risk-taking.

Finally, Figure 5 depicts the optimal investment policy as the manager’s inalienable wealth $a$ changes. Having more private wealth outside the fund leads the manager to increase risk earlier as the portfolio value declines. Greater inalienable wealth raises the subsistence threshold, which implies portfolio wealth need not decline as much before inducing the extreme risk-taking. For sufficiently high portfolio values, the optimal policies converge to the shared Merton line no matter the amount of inalienable wealth.
Notes: The figure presents the optimal fraction invested in the risky asset for different values of portfolio wealth across three separate values of risky asset volatility \( \sigma \). Cumulative transfers \( G \) are zero. Volatilities are 0.12, 0.2, and 0.6 to yield Sharpe ratios of 0.5, 0.3, and 0.1, respectively, assuming a risk premium of 6%. The value \( \alpha^* \) maintains the general behavior over wealth as in the baseline case for each \( \sigma \), though greater risk-taking occurs earlier in the highest volatility case as wealth approaches the bailout threshold.

6 Bailout Terms

The bailout terms available to the outside stakeholder are the penalty rate of interest charged on the emergency credit \( (r_p) \) and the equity stake the portfolio manager must sacrifice in exchange for the bailout (\( x \)). In this section, I examine how the optimal investment adjusts with these terms. Figure 6 presents the investment policy for varying rates of equity dilution. The case of no equity dilution (\( x = 0 \)) is also depicted. The optimal policy is independent of the dilution rate. The scale invariance of the manager’s utility function—for moderate amounts of inalienable wealth—explains this behavior. Provided the manager is not completely diluted after the initial bailout (i.e., the stakeholder takes a complete equity share), the manager chooses the same investment strategy, irrespective of the remaining
Figure 4: Optimal Investment Policy, Varying Risk Aversion

Notes: The figure illustrates the optimal fraction invested in the risky asset for different values of portfolio wealth across three separate values of risk aversion \( \gamma \). Cumulative transfers \( G \) are zero. The value \( \alpha^* \) maintains the general behavior over wealth as in the baseline case for each \( \gamma \). Risk-taking decreases with risk aversion, and the wealth threshold at which the manager accelerates risk-taking is lower the more risk averse the manager.

equity stake.

Figure 6 depicts the optimal policy for varying penalty spreads above the risk-free rate on the emergency credit extended in the bailout. A loan with no penalty rate is also shown. So that the interest rate can have the largest effect on investment behavior, the figure presents the policy over wealth when bailout transfers \( G' = 0.8 \), which is immediately below the upper bound in the state space.

For sufficiently high levels of wealth, the investment policies look very similar. As wealth declines, all penalty rates lead the manager to reduce portfolio risk. The higher penalty rate encourages slightly less risk-taking, but not by much. On the other hand, if the portfolio wealth declines enough to reach the subsistence amount, the manager escalates risk by more as the penalty rate increases. A higher penalty rate increases the amount
Notes: The figure presents the optimal fraction invested in the risky asset for different values of portfolio wealth across three separate values of inalienable wealth \( a \). Cumulative transfers \( G \) are zero. The value \( \alpha^* \) maintains the general behavior over wealth as in the baseline case for each \( a \). For sufficiently large values of wealth, the investment policy is the same irrespective of the manager’s outside wealth. As wealth declines, the extreme risk-taking initiates earlier when the manager has greater inalienable wealth.

ultimately owed to the stakeholder, which shrinks the manager’s remaining interest in the fund. When the fund is near default, the manager takes advantage of the bailout by increasing risk to reap the benefits from a good investment outcome. Lending at a high penalty rate backfires if the portfolio continues to perform poorly after an initial bailout.

The obligation to repay the bailout money leads the manager to reduce risk as the fund value declines. But neither the equity dilution rate nor the penalty rate of interest are effective terms to curtail the manager’s extreme risk-taking when the portfolio value is very low. If the fund has suffered severe losses such that an additional bailout awaits, the manager will take the maximal risk.

These results suggest that additional bailout terms are likely required to constrain the manager’s risk-taking. If the outside stakeholder is committed to rescuing the levered fund
Figure 6: Optimal Investment Policy, Varying Equity Dilution

Notes: The figure illustrates the optimal fraction invested in the risky asset for different values of portfolio wealth across three separate values of the equity dilution $x$. Cumulative transfers $G$ are zero. The value $\alpha^*$ maintains the general behavior over wealth as in the baseline case, but is unaffected by changes in the dilution rate. Though not shown, this behavior holds even for larger values of transfers $G$.

from bankruptcy—and the fund manager knows this and understands that the fund must eventually repay the bailout money—then explicit portfolio constraints might be more effective. Portfolio constraints prior to an initial bailout would have to trigger as the fund approached default. These constraints must activate and be enforced at increasing values of portfolio wealth the more assistance the fund has received.

7 Conclusion

This paper studies a levered fund manager’s optimal investment policy when the fund is subject to a guaranteed, but punishing, bailout from an outside stakeholder. The bailout rescues the fund from bankruptcy, but in exchange, the manager must sacrifice a fraction of equity and pay a penalty interest rate on the emergency loan. Multiple bailouts can
Figure 7: Optimal Investment Policy, Varying Penalty Interest Rate

Notes: The figure gives the optimal fraction invested in the risky asset for different values of portfolio wealth across three separate values of the loan penalty rate $r_p$. Cumulative transfers $G = G' = 0.8$. The value $\alpha^*$ maintains the general behavior over wealth as in the baseline case, but extreme risk-taking commences earlier the higher the penalty rate.

occur if necessary. Despite a committed bailout, the manager does not shift risk toward the stakeholder at all times. Instead, as the fund value declines, the manager reduces risk to avoid a bailout. Only when the fund nearly defaults on its debt does the manager escalate risk to exploit the rescue. Moral hazard appears only when the fund is about to fail.

The manager’s risk aversion and the obligation to repay the bailout money with interest explains this non-linear investment behavior. Once issued, the total bailout loans owed to the stakeholder becomes a kind of endogenous subsistence amount or habit stock that the manager desperately tries to prevent the fund from falling below. This habit stock can be considered a stock of bad habits because it captures the manager’s past poor performance that drove the fund to require a bailout. As the value of the fund declines, the manager reduces risk-taking to avoid “falling back into bad habits” and receiving another bailout.
If wealth falls to very low levels, the incentive to mitigate risk gradually wanes as the manager’s remaining interest in the fund net of obligations declines to zero. At this point, the manager’s limited liability, his or her possession of private wealth divorced from the fund, and the prospect of unlimited gains with no risk of loss incites the manager to take on maximal risk.

The more assistance the stakeholder has granted, the more the manager ultimately must repay, which reduces the manager’s personal stake in the fund. This larger indebtedness encourages the manager to switch sooner to maximal risk-taking, even for relatively high values of portfolio wealth. This high risk-taking increases the probability the fund loses enough to require another bailout. Every bailout therefore makes a subsequent one more likely.

The penalties of the bailout are shown to have limited success at curtailing the manager’s extreme risk-taking when the fund is near default. Instead, terms that impose explicit portfolio constraints are likely to be more effective. These constraints must reduce risk as the fund value declines but relax as the fund value does well. Such constraints would have to be enforced well in advance of default and become increasingly stringent the more assistance the fund has received.

As a final comment, the manager’s optimal investment in the risky asset converges to the fraction $\frac{\mu - r}{\gamma \sigma^2}$ as the value of the fund increases. Therefore, in good times when portfolio wealth is high, levered investment funds that carry implicit bailout guarantees would not bear wild amounts of risk in the model unless the Sharpe ratios of their risky assets were large. Only when portfolio wealth has declined substantially in bad times would one observe the extreme risk-taking.
References


A Online Appendix: Markov Chain Approximation

This section describes the numerical procedure that solves the model and is supplemental. The procedure follows Kushner and Dupuis (1992).

A.1 Bounding the State Space

I use a Markov chain approximation method to numerically approximate the manager’s value function $Q$ and optimal investment policy $\alpha^*(t)$ from the stochastic control problem of (10). To use this method, one must bound the state space of the original problem to a compact set $B$. Both the set $B$ and the behavior of the system on the boundary $\partial B$ must be specified. The state space of the manager’s problem are permissible values for the two state variables $G$ and $W$.

Natural lower boundaries for $G$ and $W$ are zero and the liability level $L$, respectively. The upper boundaries require some assumptions. Upper values $\bar{G} > 0$ and $\bar{W} > L$ must be set so that the state variables $\{G_t, W_t, t \geq 0\}$ exceed the upper boundaries with small probability when the variables are initiated from their starting values and stopped at the liquidation time $\tau$. The upper boundaries must also ensure that the optimal investment policy $\alpha^*(t)$ at locations away from those boundaries are not seriously distorted. I set the compact set $B$ as a rectangle $B = \{(G, W) : 0 \leq G \leq \bar{G}, L \leq W \leq \bar{W}\}$. Confinement to a rectangle is not necessary for the approximation method to work. My choices for $(\bar{G}, \bar{W})$ and other computational parameters are in Section A.8.

A.1.1 Reflection and absorption

At the lower boundary $W = L$, the state variable $W$ reflects, so that it always remains at or above the boundary. If $W$ is on the boundary $L$ and attempts to escape below $L$, it is instantaneously raised back to $L$ by a compensating bailout transfer amount. The compensator $G$ is simultaneously incremented by this same amount. The behavior of $G$ at the lower boundary $L$ is regulated by the behavior of $W$.

The close relation between $W$ and $G$ at the reflecting boundary $W = L$ requires specifying a vector $r(G, W) \in \mathbb{R}^2$. This vector captures the direction of reflection in the state space and applies at each point within the reflecting boundary $\partial B^R = \{(G, W) \in B : W = L\} \subset \partial B$. A natural choice is $r(G, W) = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$, which is a vector of unit length having direction $(1,1)$. When $W$ is ready to dip below $L$, the state variables $(G, W)$ are “pushed” in the direction $r$, which appropriately restores $W$ back to $L$ and increments $G$ by the restoring amount.

At the upper boundary, on the other hand, the natural behavior for the system is absorption: the first time $W$ reaches $\bar{W}$ (or $G$ reaches $\bar{G}$), the stochastic process remains there. At that moment, the decision problem ends, and the portfolio wealth net of liabilities is dispensed to the manager. Let the upper boundaries $\partial B^W = \{(G, W) \in B : W = \bar{W}\}$ and $\partial B^G = \{(G, W) \in B : G = \bar{G}\}$. For $(G, W) \in \partial B^W \cup \partial B^G$ the manager’s value function is

$$Q(G, W) = C(G, W),$$

with the function $C$ defined in (10).
A.2 Discretizing the State Space
The state space just described must be discretized for the numerical routine. For a discretization level \( h > 0 \), suppose (without loss of generality) that the sides of \( B \) are integer multiples of \( h \), and let \( B_h \) be the \( h \)-grid on \( B \). There should be a single reflecting barrier “right below” the set of grid points \( \partial B_h^R = \{(G, W) \in B_h : W = L\} \). With finite grid points, it is useful to add a set of states to which the approximating Markov chain is from the original continuous state problem, so denote \( \partial B_h^\ast \) as the finite state controlled stochastic chain on \( S_h \). The control parameter \( \xi^h_n \) is the actual control action for the chain at discrete time \( \Delta t^h_n \). The goal is to get an approximating Markov chain that is “locally consistent” (described below) with the manager’s wealth process given in (5). This property is essentially all that is needed to have the Markov chain converge to the original stochastic process \( \{(G_t, W_t) , t \geq 0\} \).

A.3 Local Consistency of the Approximating Markov Chain

Denote \( \{\xi^h_n, n < \infty\} \) as a finite state controlled stochastic chain on \( S_h \). The set \( S_h \) is the state space of the approximating controlled Markov chain.

A.3.1 Local Consistency Away from the Reflecting Boundary

Let us first consider the points of the compact state space away from the reflecting boundary, i.e., \( \{(G, W) \in B - \partial B^R\} \). Here, \( dG = 0 \). In this region, the wealth process is

\[
\, dW = b(W, \alpha) \, dt + s(W, \alpha) \, dZ_t,
\]

where \( b(W, \alpha) \) and \( s(W, \alpha) \) are abbreviations for the state-dependent drift and diffusion, respectively, of the controlled wealth process in (5). Denote by \( p^h((G, W), (G', W'))|\alpha\) the transition probability of the chain between states \( (G, W) \) and \( (G', W') \). The control parameter \( \alpha \) is from the original continuous state problem, so denote \( \alpha^h_n \) as the random variable which is the actual control action for the chain at discrete time \( n \).

Suppose there are time intervals \( \Delta t^h((G, W), \alpha) > 0 \) and define \( \Delta t^h_n = \Delta t^h(\xi^h_n, \alpha^h_n) \). Let \( \sup_{(G, W), \alpha} \Delta t^h((G, W), \alpha) \to 0 \) as \( h \to 0 \), but \( \inf_{(G, W), \alpha} \Delta t^h((G, W), \alpha) > 0 \) for each \( h > 0 \). Define the difference \( \Delta \xi^h_n = \xi^h_{n+1} - \xi^h_n \). Let \( \mathbb{E}^{h\alpha}_{(W,G)n} \) denote the conditional expectation given \( \{\xi^h_i, \alpha^h_i, i \leq n, \xi^h_n = (G, W), \alpha^h_n = \alpha\} \). The chain \( \{\xi^h_n, n < \infty\} \) satisfies a local consistency
with the process for \( G \) wealth process of (13).

This transition function is considered locally consistent with the reflection direction \( r \) if it never attempts to leave. Use \( \partial B^+_h \) to denote the reflecting boundary for the approximating chain. The transition probabilities at the states in \( \partial B^+_h \) are chosen to resemble the behavior of the reflecting wealth diffusion:

\[
dW = b(W, \alpha) \, dt + s(W, \alpha) \, dZ + dG, \tag{15}
\]

where \( o(\cdot) \) is little-o notation and the supremum in the third equation is taken over all sample paths \( \omega \) and discrete times \( n \) of the chain. The “local properties” of the chain expressed in (14) are seen to match the “local properties” of the drift and diffusion in the wealth process of (13).

So that the chain is Markov, define a control policy \( \alpha^h \) as admissible if the chain has the Markov property under that policy, i.e.,

\[
P \left\{ \Omega_{n+1}^h = (\gamma', W') \mid \Omega^h_i, \alpha^h_i, i \leq n \right\} = p^h \left( \Omega_n, (\gamma', W') \mid \alpha^h_n \right).
\]

### A.3.2 Local Consistency on the Reflecting Boundary

The reflecting boundary is meant to keep the wealth process in the set \( B \) if it ever attempts to leave. Use \( \partial B^+_h \) to denote the reflecting boundary for the approximating chain. The transition probabilities at the states in \( \partial B^+_h \) are chosen to resemble the behavior of the reflecting wealth diffusion:

\[
dW = b(W, \alpha) \, dt + s(W, \alpha) \, dZ + dG, \tag{15}
\]

where \( o(\cdot) \) is little-o notation and the supremum in the third equation is taken over all sample paths \( \omega \) and discrete times \( n \) of the chain. The “local properties” of the chain expressed in (14) are seen to match the “local properties” of the drift and diffusion in the wealth process of (13).

So that the chain is Markov, define a control policy \( \alpha^h \) as admissible if the chain has the Markov property under that policy, i.e.,

\[
P \left\{ \Omega_{n+1}^h = (\gamma', W') \mid \Omega^h_i, \alpha^h_i, i \leq n \right\} = p^h \left( \Omega_n, (\gamma', W') \mid \alpha^h_n \right).
\]

### A.3.2 Local Consistency on the Reflecting Boundary

The reflecting boundary is meant to keep the wealth process in the set \( B \) if it ever attempts to leave. Use \( \partial B^+_h \) to denote the reflecting boundary for the approximating chain. The transition probabilities at the states in \( \partial B^+_h \) are chosen to resemble the behavior of the reflecting wealth diffusion:

\[
dW = b(W, \alpha) \, dt + s(W, \alpha) \, dZ + dG, \tag{15}
\]

where \( o(\cdot) \) is little-o notation and the supremum in the third equation is taken over all sample paths \( \omega \) and discrete times \( n \) of the chain. The “local properties” of the chain expressed in (14) are seen to match the “local properties” of the drift and diffusion in the wealth process of (13).

So that the chain is Markov, define a control policy \( \alpha^h \) as admissible if the chain has the Markov property under that policy, i.e.,

\[
P \left\{ \Omega_{n+1}^h = (\gamma', W') \mid \Omega^h_i, \alpha^h_i, i \leq n \right\} = p^h \left( \Omega_n, (\gamma', W') \mid \alpha^h_n \right).
\]

The conditions in (16) require the approximating chain to have an average change in direction that is the “allowable reflection” \( r(G, W) \) plus a “small” error. In addition, the transition probability \( p^h \) of states in the reflecting boundary to states in the rest of the compact state space must be positive.

If a Markov chain is locally consistent in terms of (14) and also locally consistent with the reflection direction \( r(G, W) \), as defined in (16), then the chain is locally consistent with
the reflected diffusion process for wealth in (15). This property is what we need for proper convergence of both the Markov chain and the discrete state value function, which is defined below.

For points \((G, W) \in \partial B_h^+\), the interpolation interval \(\Delta t^h (G, W)\) for the Markov chain is defined to be zero. This property captures the instantaneous character of the reflection inherent in the original reflected diffusion (15).

For the manager’s investment problem, getting a probability transition function that is locally consistent on the boundary reflection is quite easy. If \(W\) breaches the lower value \(L\) by the amount \(h\), then \(G\) must subsequently increase by \(h\) as \(W\) is also incremented by \(h\). Therefore, the choice for the transition function is simply:

\[
p^h ((G, W), (G, W) + (h, h)) = 1
\]

for all \((W, G) \in \partial B_h^+\), which satisfies the conditions of (16). In section A.5, I give the locally consistent transition probabilities for the Markov chain when it is away from the reflecting boundary.

### A.4 Approximating the Original Control Problem

In order to approximate the continuous time process \((G_t, W_t)\), we need to use a continuous time interpolation of the discrete time process \(\{x_n^h, n < \infty\}\). Therefore, let \(\{\alpha_n^h, n < \infty\}\) be an admissible control for the chain and define the interpolated time \(t_n^h = \sum_{i=0}^{n-1} \Delta t_i^h\). Define the continuous time interpolations \(\xi^h (\cdot)\) and \(\alpha^h (\cdot)\) by:

\[
\begin{align*}
\xi^h (t) &= \xi_{n^h}^h, \\
\alpha^h (t) &= \alpha_{n^h}, & t \in \left[t_{n^h}, t_{n^h + 1}\right).
\end{align*}
\]

The interpolated process defined by (17) is piecewise constant, and if it satisfies (14) and (16), it is an approximation of the reflected diffusion that preserves its local properties both at and away from the boundary. Note that the values of the states at the moments of the reflection do not appear in the definition of the interpolation. These states are “instantaneous” and the reflection is instantaneous for the continuous time interpolation.

Suppose \(A\) is a compact set in which the manager’s investment policy can only take values. Recall that \(\partial B^W_h\) and \(\partial B^G_h\) are the “upper” boundaries in \(B_h\) when \(W = \bar{W}\) and \(G = \bar{G}\), respectively. Denote by \(S^-_h\) as the points in the \(h\)–grid away from the absorbing and reflective boundaries; i.e., \(S^-_h = S_h - \partial B^W_h - \partial B^G_h - \partial B^+_h\). From the manager’s dynamic program in (9), a natural analogue using the approximating Markov chain is

\[
Q^h (G, W) = \begin{cases}
\sup_{\alpha (\cdot) \in A} \left\{ e^{-1 \Delta t^h ((G, W), \alpha)} \sum_{(G', W')} p^h ((G, W), (G', W') | \alpha) Q^h (G', W') \\
+ C (G, W) \Delta t^h ((G, W), \alpha) \right\} & (G, W) \in S^-_h, (G', W') \in S_h \\
C (G, W) & (G, W) \in \partial B^W_h \cup \partial B^G_h \\
Q^h (G + h, W + h) & (G, W) \in \partial B^+_h.
\end{cases}
\]

When the Markov chain is at points within \(S_h\) but away from the absorbing boundaries \(\partial B^W_h\) and \(\partial B^G_h\) and reflecting boundary \(\partial B^+_h\), the Bellman equation takes a familiar form. The value function \(Q^h (G, W)\) equals a flow \(C (G, W)\) over the small time step \(\Delta t^h ((G, W), \alpha)\).
plus the discounted mean of the value function over the possible increments in the Markov chain. The points in this region have transition probabilities over the whole space \( S_h \). However, points in this region can only take new values in \( W \). The transfer component \( G \) must remain fixed. Thus, \( p^h ((G, W), (G', W') | a) = 0 \) for \((G', W') \in S_h \) and \( G' \neq G \).

If the Markov chain reaches an absorbing boundary so that \((G, W) \in \partial B_h^W \cup \partial B_h^G \), then the decision problem of the manager ends and the manager receives the value \( C(G, W) \).

Lastly, if the state reaches the reflecting boundary \( \partial B_h^+, \) then the manager’s value function should equal with probability one the value function when both \( W \) and \( G \) are incremented by \( h \). The terms of the bailout trigger this increment instantaneously with probability one. Because of this instantaneous reflection, there is no time increment \( \Delta t^h ((G, W)) \).

It is convenient to write the dynamic programming problem expressed in (18) using matrix notation. Define the vector \( Q^h = \{ Q^h (G, W), (G, W) \in S_h \} \) and the gain vector \( J^h (a) = \{ J^h (G, W, a), (G, W) \in S_h - \partial B_h^W - \partial B_h^G \} \) with components

\[
J^h (G, W, a) = C(G, W) \Delta t^h ((G, W), a),
\]

for points \((G, W) \in S_h^- \) and \( J^h (G, W, a) = 0 \) for points \((G, W) \in \partial B_h^+ \) on the reflecting boundary. Next, define the matrix

\[
R^h (a) = \left\{ r^h ((G, W), (G', W') | a); (G, W) \in S_h - \partial B_h^W - \partial B_h^G, (G', W') \in S_h \right\},
\]

where

\[
r^h ((G, W), (G', W') | a) = e^{-\lambda \Delta t^h ((G, W), a)} p^h ((G, W), (G', W') | a)
\]

holds for \((G, W) \in S_h^- \); i.e. points away from both absorbing and reflecting boundaries. For points on the reflecting boundary \((G, W) \in \partial B_h^+ \), the terms \( r^h ((G, W), (G', W') | a) = 1 \) for \((G', W') = (G + h, W + h) \) and zero otherwise. With these constructions, we can write the Bellman equation from (18) in the compact form for all values of the state space \( S_h \) of the controlled Markov chain:

\[
Q^h = \sup_{a(t) \in A} \left\{ R^h (a) Q^h + J^h (a) \right\} \quad \begin{cases} 
(G, W) \in S_h - \partial B_h^W - \partial B_h^G \\
C \quad (G, W) \in \partial B_h^W \cup \partial B_h^G.
\end{cases} \tag{19}
\]

### A.5 Constructing the Markov Chain

To solve the dynamic program in (19), one needs suitable interpolation intervals \( \Delta t^h ((G, W), a), \) as well as transition probabilities \( p^h ((G, W), (G', W') | a) \). These two objects will formally construct the approximating Markov chain we require, provided they satisfy the local property conditions of (16) and (18). These objects for states on the reflecting boundary were given in section A.3.2.

For states away from the reflecting boundary, one method to obtain the interpolation intervals and the transition probability functions is “finite differencing.” Both \( \Delta t^h ((G, W), a) \) and \( p^h ((G, W), (G, W') | a) \) can be generated automatically from a carefully chosen finite difference approximation to the familiar differential operator of the controlled (but unreflected) wealth process. This derived Markov chain would satisfy the local consistency
properties of (14).

On the original bounded state space, let \((G, W) \in B - dB^W - dB^G - dB^R\) so that the wealth process \(W\) is given by (13) and \(dG = 0\). Define the reward functional \(I\) for the manager’s problem as

\[
I(G, W, \alpha) = \mathbb{E} \left[ \int_0^\infty e^{-\lambda u} C(G_u, W_u) \, du \right].
\]

Notice that \(Q(G, W) = \sup_{\alpha} I(G, W, \alpha)\). Assuming that \(I\) is sufficiently smooth, apply Ito’s formula to obtain

\[
\mathcal{L}^W I(G, W, \alpha) - \lambda I(G, W, \alpha) + C(G, W) = 0,
\]

where \(\mathcal{L}^W\) is the infinitesimal generator of \(W\). Expanding the above expression yields

\[
I_W b(W, \alpha) + \frac{1}{2} I_{WW}s^2(W, \alpha) - \lambda I + C = 0. \tag{20}
\]

Note the standard second order approximation for a function \(f\):

\[
f_{ww} = \frac{f(w + h) - 2f(w) + f(w - h)}{h^2}.
\]

Define the one-sided difference approximations

\[
f_w = \begin{cases} 
\frac{f(w + h) - f(w)}{h} & \text{if } b(w, \alpha) \geq 0, \\
-\frac{f(w) - f(w - h)}{h} & \text{if } b(w, \alpha) < 0.
\end{cases}
\]

The one-sided difference uses the forward difference if the drift of wealth at any point is non-negative and the backward difference otherwise.

The term \(\lambda I\) in (20) is due to discounting can be ignored for now in the finite difference method. This is acceptable, as local consistency of the Markov chain does not depend on whether there is discounting. A discounting-related term will be added back once the probabilities have been established.

Denote \(b^+ \equiv \max(b, 0)\) and \(b^- \equiv \max(-b, 0)\), and note that \(b^+ - b^- = b\) and \(b^+ + b^- = |b|\).

Apply the difference approximations to \(I(G, W, \alpha)\) in \(W\), collect terms, multiply by \(h^2\), and divide all terms by the coefficient of \(I\) to obtain the approximating equation

\[
I^h(W, \alpha) = \frac{s^2 + h b^+}{s^2 + h |b|} I^h(W + h, \alpha) + \frac{s^2 + h b^-}{s^2 + h |b|} I^h(W - h, \alpha) + C \frac{h^2}{s^2 + h |b|}, \tag{21}
\]

where \(I^h\) represents the approximated function to \(I\) on the \(h\)-grid \(S_h\). (I suppressed the other variable \(G\) upon which the function \(I\) depends.) One can re-write (16) as

\[
I^h(W, \alpha) = p^h(W, W + h|\alpha) I^h(W + h) + p^h(W, W - h|\alpha) I^h(W - h) + C \Delta^h(W, \alpha), \tag{22}
\]

where \(p^h\) and \(\Delta^h(W)\) are defined by their corresponding terms in (21). The terms
\( p^h (W, W \pm h | \alpha) \) are non-negative and sum to one. Therefore, they can be considered transition probabilities for a Markov chain on the subset of the state space: \( S_h - \partial B^W_h - \partial B^C_h - \partial B^r_h \).

Because these probabilities depend on \( \alpha \), computing them can be challenging when optimizing the value function. A way around this is to define adjusted transition probabilities that are independent of \( \alpha \). Let

\[
D^h (W) = \max_{a \in A} \left[ s^2 (W, \alpha) + h \mid b (W, \alpha) \right].
\]

Define the new transition probabilities as

\[
\bar{p}^h (W, W \pm h | \alpha) = \frac{s^2 (W, \alpha) /2 + hb^\pm (W, \alpha)}{D^h (W)}
\]

\[
\bar{p}^h (W, W | \alpha) = 1 - \sum_{W' \neq W} \bar{p}^h (W, W \pm h | \alpha)
\]

\[
= \frac{D^h (W) - [s^2 (W, \alpha) + h \mid b (W, \alpha)]}{D^h (W)} ,
\] (23)

with the time increment

\[
\Delta \bar{t}^h (W) = \frac{h^2}{D^h (W)}.
\] (24)

Let \( \bar{I}^h \) denote the reward under the adjusted probabilities \( \bar{p}^h \). Then

\[
\bar{I}^h (W, \alpha) = \sum_{W' \neq W} \bar{p}^h (W, W' | \alpha) \bar{I}^h (W', \alpha) + \bar{p}^h (W, W | \alpha) \bar{I}^h (W, \alpha) + C \Delta \bar{t}^h (W) .
\] (25)

This equation equals

\[
\bar{I}^h (W, \alpha) = \sum_{W' \neq W} \frac{\bar{p}^h (W, W' | \alpha)}{1 - \bar{p}^h (W, W | \alpha)} \bar{I}^h (W', \alpha) + C \frac{\Delta \bar{t}^h (W)}{1 - \bar{p}^h (W, W | \alpha)}
\]

\[
= \sum_{W' \neq W} p^h (W, W' | \alpha) \bar{I}^h (W', \alpha) + C \Delta \bar{t}^h (W, \alpha) .
\] (26)

Comparing (16) to (22), one can see that \( I^h (W, \alpha) = \bar{I}^h (W, \alpha) \) for all controls \( \alpha \) for which (22) or (26) has a unique solution. Thus, we can use (26), which eliminates the dependence of the transition probabilities on the control, and get the same results.

The probabilities \( \bar{p}^h \) of (23) can be considered the transition probabilities for a Markov chain on the space \( S_h - \partial B^W_h - \partial B^C_h - \partial B^r_h \) with \( \Delta \bar{t}^h (W) \) as the associating time interval. Notice that both the transition probabilities \( \bar{p}^h \) of the Markov chain and time interval \( \Delta \bar{t}^h (W) \) depend on the drift and diffusion of the original wealth process. A larger drift \( |b (W, \alpha)| \) and diffusion \( s^2 (W, \alpha) \) raises the probability that the chain moves to a different step in the time increment. As the magnitudes of these coefficients increase, the interval decreases in size. Finally, only if the “positive” drift \( b^+ \) exceeds the “negative” drift \( b^- \), is there a higher chance the chain will move to a larger \( W \) in the state space than a lesser one. The diffusion
Thus, for $G$ the controlled Markov chain with an interpolation interval determined by programming equation using the controlled chain for all points $(s, \alpha)$ satisfied. Brownian process in the manager’s problem (see Kushner (1990)), and the others we have approximating Markov chain $\xi_s^{(C)}$ local consistency requirements in (14) and (16), existence of a solution to the stochastic differential equation of (15) in the weak sense, continuity and boundedness of the function $C(G, W)$, and continuity of the first hitting time of $\partial B$ by the limit process $W$, of the approximating Markov chain $\xi^h(\cdot)$. This last condition is satisfied by the properties of the Brownian process in the manager’s problem (see Kushner (1990)), and the others we have satisfied.

A6 Convergence of the Approximating Chain and the Value Function

Formal proofs that the continuous time interpolated Markov chain $\xi^h(\cdot)$ from (17) has a subsequence as $h \to 0$ which converges in weak measure to the controlled diffusion of (15) and that $Q^h$ in (19) converges to $Q$ are given in Kushner and Dupuis (1992), Ch. 11, and Kushner (1990), Section 8. Loosely speaking, the conditions for this convergence are the local consistency requirements in (14) and (16), existence of a solution to the stochastic differential equation of (15) in the weak sense, continuity and boundedness of the function $C(G, W)$, and continuity of the first hitting time of $\partial B$ by the limit process $W$, of the approximating Markov chain $\xi^h(\cdot)$. This last condition is satisfied by the properties of the Brownian process in the manager’s problem (see Kushner (1990)), and the others we have satisfied.
A.7 Approximating the Value Function and Optimal Policy

Computationally, the goal is to solve for a policy function \( \alpha (G, W) \) that satisfies the dynamic programming matrix equations from (19), which are re-printed below:

\[
Q^h = \left\{ \sup_{\alpha \in A} R^h \left( \bar{p}^h, \Delta \bar{p}^h, \alpha \right) Q^h + J^h \left( \Delta \bar{p}^h \right) \right\} \\
(G, W) \in S_h - \partial Z_h - \partial Z_h \\
(G, W) \in \partial Z_h \cup \partial Z_h
\]  

(28)

Loosely, if the elements of \( R^h \) and \( J^h \) are continuous functions of \( \alpha \) and the states, and if \( R^h \) is a contraction for at least one control, which we have, then there is a unique solution to (28) (see Kushner and Dupuis (1992), ch. 6). Normally, some kind of fixed-point iteration is followed to solve for \( Q^h \). I use a modified policy iteration algorithm. This method is a blend of the standard policy iteration and value iteration approaches. In the standard policy iteration, a sequence of optimizing control policies \( \{ \alpha_m^h (\cdot) \} \) for the Markov chain are computed, and the control \( \alpha_m^h (\cdot) \) is obtained after the convergence of the reward function solution under \( \alpha_m^h \). However, the modified algorithm differs from the standard policy iteration by stopping short of convergence in the reward function iteration for each control \( \alpha_m^h (\cdot) \), which can be onerous. Instead, at every step in updating the control, it quits iterating after a “good” approximation of the reward function is obtained.

Formally, denote the sequence of controls \( \{ \alpha_m^h (\cdot), m \geq 1 \} \), and let \( I^h (\alpha) \) be the reward function from the matrix equation:

\[
I^h (\alpha) = \left\{ R^h \left( \bar{p}^h, \Delta \bar{p}^h, \alpha \right) I^h (\alpha) + J^h \left( \Delta \bar{p}^h \right) \right\} \\
(G, W) \in S_h - \partial Z_h - \partial Z_h \\
(G, W) \in \partial Z_h \cup \partial Z_h
\]

Suppose we have a “candidate” optimal control for the chain, denoted \( \alpha_m^h ((G, W)) \). If we define the subsequent control in the sequence \( \alpha_m^h (\cdot) \) as

\[
\alpha_m^h ((G, W)) = \arg \max_{\alpha \in A} \left\{ R^h \left( \bar{p}^h, \Delta \bar{p}^h, \alpha \right) I^h (\alpha) + J^h \left( \Delta \bar{p}^h \right) \right\},
\]

(29)

for \( (G, W) \in S_h - \partial Z_h - \partial Z_h \) then \( I^h (\alpha_m) \to Q^h \). In addition, the local consistency of the Markov chain would give \( Q^h \to Q \) weakly in measure. One can see here that for every update to the control sequence, an estimate of \( I^h (\alpha_m) \) is required. In standard policy iteration, this estimate would follow from the iteration:

\[
I_{k+1}^h (\alpha_{m+1}) = R^h \left( \bar{p}^h, \Delta \bar{p}^h, \alpha_{m+1} \right) I_k^h (\alpha_{m+1}) + J^h \left( \Delta \bar{p}^h \right)
\]

until convergence is met by some criterion. Satisfying convergence could require a great deal of calculation for each updating step of the control sequence. Instead, one could just use an integer \( K \) of reward function iterations, and define \( I^h (\alpha_{m+1}) \) to be the value at the end of the iterations for use in the subsequent update in the control \( (\alpha_{m+2}) \) via (29). Note that \( K = 1 \) and \( K = \infty \) are the standard value and policy iteration methods, respectively. Modifying the iteration algorithm in this fashion can save considerable computation time,
yet still provide the necessary convergence. Proofs for convergence in the discounted problem, as we have, are available in Puterman and Shin (1978).

Not only does the modified policy iteration save computation time, but so too does the structure of the manager’s problem itself. The manager’s problem allows use of a domain decomposition. The manager’s value function $Q^h$ is known at the absorbing barriers $\partial B^N_h$ and $\partial B^C_h$; it equals $C(G, W)$. In addition, the Markov chain $\xi^h_n$ can transition to the “right” in the grid state space $B_h$ only if it has hit the “bottom” reflecting boundary $\partial B^+_h$. Otherwise, the chain only transitions vertically between different values of $W$ at a fixed $G$. Hence, the problem’s environment strongly suggests that the efficiency of the modified policy algorithm could benefit greatly if it began on the right, where the value function is known and “migrate” to the left. This is the direction I follow.

Specifically, suppose that in the $i$-th “column” of the state space, where $G = gh$ for some $g \in \mathbb{Z}_{++}$, the value function $Q^h$ and optimal control $\alpha(\cdot)$ is determined. From there, the modified policy iteration algorithm begins in the $(i - 1)$ column of $G$. It then searches for the optimal policy over all wealth $W$ in this column using the sequence $\{\alpha^h_m\}$ until the reward function converges, i.e.

$$\| I^h(W, G = gh, \alpha^h_{m+1}) - I^h(W, G = gh, \alpha^h_m) \|_{\infty} < \epsilon$$

for a convergence criterion $\epsilon$. (Searching for the optimal policy also considers values of $W$ for which the Markov chain tries to escape the absorbing and reflecting boundaries. In the first case, the Markov chain terminates and the reward function is known; whereas in the second case, the chain transitions with probability one back to the $i$-th column of $G$, where the reward function was already determined.) Next, the algorithm moves to the $(i - 2)$ column of $G$, where $G = (g - 1)h$. From there, the search for the optimal policy repeats over this column of wealth values. The algorithm continues migrating left in the state space until the state space has been exhausted. Decomposing the state space in this fashion exploits the natural flow of the system and avoids having to iterate over the entire space in every step.

### A.8 Choice of computational parameters

The computational parameter values are presented in Table 3.
Table 3: Computational Parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Parameter</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A$</td>
<td>[0, 2]</td>
</tr>
<tr>
<td>$W_0$</td>
<td>1</td>
</tr>
<tr>
<td>$\bar{W}$</td>
<td>3</td>
</tr>
<tr>
<td>$G$</td>
<td>1</td>
</tr>
<tr>
<td>$h$</td>
<td>0.2</td>
</tr>
<tr>
<td>$\epsilon$</td>
<td>$1e^{-6}$</td>
</tr>
</tbody>
</table>

Notes: The table presents the parameters for the numerical procedure. The compact set $A$ is a closed interval with lower bound 0 (to prevent short selling) and upper bound 2, which allows the manager to borrow off the risk-free asset. Starting fund value $W_0$ is set to 1, whereas the upper bound of the state space for the fund value is three times that amount. The upper bound for transfers $G$ is 1. (The lower bound is zero.) These bounds for the state space were set so that the state variables $(W, G)$ had a low probability of reaching them, and so that the bounds did not distort the optimal policy. Finally, the discretization level of the state space $h$ is 0.2 and the convergence criterion $\epsilon$ is $1e^{-6}$. Other values for these two parameters were used as well, and the qualitative aspects of the investment policy were unchanged.

The lower bound of the compact set $A$ for the policy space was set to prevent short sales of the risky asset. In any case, because $\mu - r > 0$, the manager would never short the risky asset. The upper bound was set to 2 so that the manager could borrow. It turned out that the manager only reached the upper bound of the policy space when the portfolio value neared the subsistence amount. The optimal policy increased risk as much as possible, so the manager would reach whichever upper bound was in place.

The manager’s starting wealth $W_0$ was set to 1 as a normalization. The upper bounds of the state space were set to $\bar{W} = 3$ and $\bar{G} = 1$. These values were selected to ensure that the state variables had a low probability of reaching the upper boundaries after 100 steps of a Markov chain that began at $(W_0, G_0) = (1, 0)$. The 100 steps was the expected termination time of the fund $(\frac{1}{\lambda})$. In addition, the values were set so that the optimal policy near the upper boundaries was not distorted. For $\bar{G}$, the manager’s policy to taper risk as wealth declined was true for any $G$; the only difference across $G$ was the level of wealth at which the manager reversed to drastically increase risk. Hence, the policy was not distorted for the $\bar{G}$ chosen. For $\bar{W}$, the investment policy to gradually increase risk as $W$ rose was true for all wealth sufficiently away from the subsistence amount. Also, the risky asset investment neared $\alpha = \frac{\mu - r}{\gamma \sigma^2}$ as wealth reached $\bar{W}$. Hence, this boundary did not distort the policy either.

The discretization level $h$ of the state space was set low enough to capture the qualitative aspects of the manager’s optimal policy, while still leaving the computation time reasonable. Finally, the value function convergence criterion $\epsilon$ is set to $1e^{-6}$.