Risk-taking under a Punishing Bailout

Alexander K. Zentefis*

The University of Chicago Booth School of Business
azentefi@ChicagoBooth.edu

August 19, 2014

Abstract

A government commitment to bail out a large investment fund may encourage the fund manager to take excessive risks ex ante. But what if the bailout were harsh? The loan would have a penalty rate and the manager would have to sacrifice part of the fund’s equity. I show that as the value of the fund declines, the manager actually reduces risk in order to avoid the government 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.

JEL classification: G11, G23, G28
Keywords: risk-taking, portfolio choice, government policy, bailout

*I am grateful for very helpful comments from John Cochrane, Douglas Diamond, Lars Hansen, Bora Keskin, Stavros Panageas, Aaron Pancost, Fabrice Tourre, Pietro Veronesi, the Chicago Booth Finance Brownbag seminar, and especially Varun Gupta.
The latest version of this paper is posted at https://phd.chicagoboosth.edu/alexander.zentefis/.
1 Introduction and Overview

This paper analyzes portfolio choice when there is an implicit, but punishing, bailout guarantee by the government. An investment fund manager borrows money and invests in a risky asset. If the fund is at the brink of default, the manager knows the government is committed to bailing out the fund with emergency lending. However, the government loans are supplied at a penalty rate, and in exchange, the manager must forfeit a fraction of equity ownership of the fund to the government. The manager cannot refuse assistance if the government grants it.

These government bailout terms have their grounding in the real-world. A penalty rate of interest and the relinquishing of equity ownership are what the U.S. Federal Reserve enforced for the emergency credit facility extended to AIG in 2008 to prevent the company’s collapse. Whether AIG knew it would be bailed out, and under those specific terms, is not something I explore. My goal is not to match the investment behavior of AIG. Instead, I ask generally: if an asset manager knows he would be bailed out, but on punishing terms, how would his ex-ante investment policy be distorted?

Because of this government guarantee, one might suspect the manager would have an incentive to increase portfolio risk at all times in order to exploit the implicit protection. This will not be the case. The manager’s investment policy will depend on the state of the fund’s wealth. In fact, at lower wealth levels, the manager actually curtails risk to avoid the bailout because of its punishing terms.

The more the fund loses, the more the manager reduces risk. However, if portfolio wealth declined far enough so that the manager’s remaining interest in the portfolio net of his obligations (including to the government) approaches zero, then limited liability and the prospect of no further loss from the bailout incite him to taken on maximal risk. The threshold at which he initiates the extreme risk-taking is greater than his total indebtedness, so while the fund is not yet insolvent, his behavior increases the probability the fund will reach insolvency 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 outside creditors and the government for bailout money lead the manager to mitigate risk as the value of the fund declines. The fund’s total indebtedness becomes a form of “subsistence” or “habit” wealth, which the manager, being risk averse, 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, so as wealth decreases, the
fraction invested in the risky asset declines. This explains why the portfolio manager in the model increases risk as portfolio wealth increases, but reduces his risky holdings as the fund value declines.

Multiple bailouts are possible if the manager has reached default multiple times. The manager’s subsistence amount in the model, therefore, is not fixed, but stochastic, increasing with the cumulative amount of government loans he has received. Because these loans reflect the historical path of portfolio wealth, another interpretation of this subsistence amount is a habit stock of “bad habits.” The manager’s poor historical performance that drove the fund to a bailout raises the bad habit stock. Given his risk aversion, he 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 reached the subsistence level, the manager would have completely divested from the risky asset to ensure that wealth could decline no further. In the model, however, the manager possess inalienable wealth that is protected from the fund’s creditors, including the government. This wealth may be housing, land, boats, or personal financial assets, for instance. This inalienable wealth stops the manager from shedding risk completely as the fund declines in value. Instead, his incentive to mitigate risk gradually wanes as his remaining interest in the fund net of his obligations declines to zero. At that point, anticipation of unlimited gains with no risk of loss stirs him to take on maximal risk.

Repeated bailouts increase the manager’s indebtedness to the government, raising his subsistence level. This means that portfolio wealth would have to fall by less in the future to dip below this higher subsistence and trigger his extreme risk-taking. The greater risk exposure makes the fund more vulnerable to loss, insolvency, and additional government assistance. Hence, current bailouts foreshadow future bailouts.

One might look to the terms of the government 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 stop him when portfolio value is low enough. This suggests that additional terms, such as direct portfolio allocation constraints would be necessary.

I model the manager’s portfolio problem in the wake of a government bailout as a stochastic optimal control problem. To obtain the manager’s optimal investment policy, I numerically approximate the problem using a Markov chain approximation method. This numerical technique, expounded in Kushner and Dupuis (1992), approximates the original control problem with a simpler one using a Markov chain that resembles the original controlled stochastic process. Under suitable conditions, the solution to the simpler problem converges to that of the original control problem, so that one can be assured to obtain an optimal policy
from the approximation that is very close to the true one.

A related paper on investment policy in the presence of bailouts is Panageas (2010). There, the portfolio manager is risk neutral, instead of risk averse. The manager is not obligated to repay the bailout money, and the bailout is not guaranteed. In that setting the manager takes on maximal risk until wealth falls below an endogenous threshold, at which point he completely reverses and minimizes risk. As in this paper, risk taking there increases in good times and falls in bad times, but does so for a completely different reason, which is to avoid having the bailout reneged upon.

2 The Model

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

2.1 Investment Opportunities

There are two distinct assets 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 then given by

\[
\frac{dP^0_t}{P^0_t} = r dt \\
\frac{dP^1_t}{P^1_t} = \mu dt + \sigma dZ_t
\]

where \( \{Z_t, t \geq 0\} \) is a standard one-dimensional Wiener process on a complete probability space \( \{\Omega, \mathcal{F}, \mathbb{P}\} \), while \( \mu > r \) and \( \sigma > 0 \) are constants. Let \( \mathbb{F} = \{\mathcal{F}_t, t \geq 0\} \) 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 \( \{\alpha_t, t \geq 0\} \), which denotes the fraction of total assets invested in the risky asset. The remaining fraction of wealth \( (1 - \alpha_t) \) is invested in the risk-free asset. The process \( \alpha_t \) is considered admissible if it is progressively measurable with respect to the filtration \( \mathbb{F} \) and satisfies the integrability condition \( \int_0^T |\alpha_s| ds < \infty \) almost surely for all \( T > 0 \). Together with the uniform Lipschitz continuity in \( \alpha \) of the drift and diffusion, this condition ensures that there exists a unique strong solution to the manager’s
wealth process defined below.

So that the results will be independent of time, the fund will liquidate 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 his continuation value is zero.

The manager services the debt by paying a flow rate of interest $rL$. The interest rate on the loan is the risk-free rate since 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 portfolio manager will have an unconditional bailout from the government. 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 $W_t \geq L$ (see Karatzas and Shreve (1991) (p. 210-211), and Panageas (2010)) is given by:

$$\int_{s=0}^{t} \frac{dG_s}{L} = \max \left\{ 0, \max_{0 \leq s \leq t} \left\{ \log (L) - \left( \log (W_0) + \int_{u=0}^{s} [r + \alpha_u (\mu - r) - rL] du - \frac{1}{2} \int_{u=0}^{s} \alpha_u^2 \sigma^2 du + \int_{u=0}^{s} \alpha_u \sigma dZ_u \right) \right\} \right\}.$$  

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 given by

$$dW_t = [W_t (\alpha_t \mu + (1 - \alpha_t) r) - rL] dt + W_t \alpha_t \sigma dZ_t + dG_t. \quad (2.1)$$

The bailout transfers are incremental loans distributed to the manager, 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 government at the termination date $\tau$. Moreover, at the instant he receives the first government loan, the manager sacrifices a fraction $0 < x < 1$ in equity ownership in the fund to the government. No matter the total amount it ultimately lends to the fund, the government keeps only this fraction of equity.

The manager is not restricted to keep net capital $W_t - L - G_t$ strictly positive at all times. The government transfers ensure that $W_t \geq L$, so the original liabilities are riskless. The government’s loans, however, are subordinate to the original liabilities, and there is no guarantee that the government will be repaid in full for its transfers. The bailout shifts default risk from the fund’s creditors to the government. As compensation for the enhanced risk, the government charges a penalty rate on the loan and is granted an equity stake in the
fund. To see the distribution of funds at the liquidating date $\tau$ explicitly, consider Table (1). The gross penalty interest rate $R_g = (1 + r_g)$.

<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>Government</td>
<td>$\min {W_\tau - L, R_g G_\tau} + \max {0, x (W_\tau - L - R_g G_\tau)}$</td>
<td>$0$</td>
</tr>
<tr>
<td>Manager</td>
<td>$\max {0, (1 - x) (W_\tau - L - R_g G_\tau)}$</td>
<td>$W_\tau - L$</td>
</tr>
</tbody>
</table>

Table 1: Distribution of Payoffs at $\tau$

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

2.3 Portfolio Decision Problem

The goal of the manager is to choose the investment policy $\{a_t, t \geq 0\}$ to maximize his utility over net wealth. He has CRRA utility with relative risk aversion coefficient $\gamma > 0, \gamma \neq 1$. Mathematically, his optimal expected value function for all $t \leq \tau$ is

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

(2.2)

where $(\cdot)^+$ denotes the positive part of $(\cdot)$, and where $1_{\{G_\tau > 0\}}$ is an indicator function that is 1 if government transfers at liquidation $G_\tau$ are positive so that the manager’s position is diluted, and zero otherwise, while $0 < a < \infty$ is a constant amount of inalienable wealth available to the manager that is separate from the assets in the portfolio and is protected from outside creditors. The presence of inalienable wealth is designed to ensure that a government bailout could actually occur. Otherwise, if the manager’s utility were CRRA strictly over $(1 - x \cdot 1_{\{G_\tau > 0\}}) (W_\tau - L - R_g G_\tau)$, then the amount $x \cdot 1_{\{G_\tau > 0\}} W_t + (1 - x \cdot 1_{\{G_\tau > 0\}}) (L + R_g G_\tau)$ could be considered as a form of “subsistence wealth” that is stochastic and increasing in wealth if $G_t > 0$ for any $t \leq \tau$; otherwise, fixed at $L$. In this case, since marginal utility at the level $W_t = x \cdot 1_{\{G_\tau > 0\}} W_t + (1 - x \cdot 1_{\{G_\tau > 0\}}) (L + R_g G_\tau)$ is
infinite, the manager would ensure that wealth never declined to this amount (see Dybvig (1995); Ingersoll (1992)). And since \( G_t \geq 0 \) for all \( t \), this would mean that wealth would never decline to \( W_t = L \). The original liabilities would be riskless already and there would be no need for a government bailout.

Since \( \tau \sim \exp(\lambda) \), re-write (2.2) so that

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

(2.3)

where

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

(2.4)

Alternatively,

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

(2.5)

where \( Q(G_t, W_t) = \frac{V(G_t, W_t)}{\lambda} \). The discounted value function \( Q \) is finite due to discounting.

Now, take a small time interval \( \delta > 0 \). The Bellman principle of optimality suggests:

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

(2.6)

The formulation of the manager’s problem in (2.6) is what I use to solve the model.

3 Solving the Model

To determine the manager’s discounted value function \( Q \) and optimal investment policy \( \alpha(G_t, W_t)^* \), I apply numerical methods to a discrete state version of the control problem in (2.6). I follow a Markov chain approximation method (see Kushner and Dupuis (1992)). The essence of a Markov chain approximation method is to approximate the original control problem with a simpler problem by approximating the original controlled stochastic process, here \( \{(G_t, W_t), t \geq 0\} \), by a suitable Markov chain process controlled on a finite state space. The finite state space is a discretization of the original state space. The criterion that must be satisfied by the approximating chain is “local consistency,” which means that the conditional mean and variance of the changes in the Markov chain process at a local level must “resemble” the local mean drift and variance of the original process \( W_t \). This “resemblance” gets closer and closer as the approximating chain converges to the original process.

The numerical problem is then to solve the control problem for the approximating controlled chain. An advantage of this approach is that proving the value function of the
approximating problem actually converges to the original value function is done by probabilistic methods without requiring the actual analytical solution to the original problem.

It turns out that the approximation method implies that the original continuous time dynamic investment problem of the manager from (2.6) can be conveniently represented as a discrete state dynamic programming problem in the following matrix form:

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

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 one of optimally controlling the transition probabilities between the Markov states in wealth and government transfers. It turns out that under some technical conditions, the solution $Q^h$ of this dynamic program converges to the value function from the original control problem in (2.5); 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>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu$</td>
<td>0.11</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>0.20</td>
</tr>
<tr>
<td>$r$</td>
<td>0.05</td>
</tr>
<tr>
<td>$L$</td>
<td>0.5</td>
</tr>
<tr>
<td>$r_p$</td>
<td>0.085</td>
</tr>
<tr>
<td>$x$</td>
<td>0.8</td>
</tr>
<tr>
<td>$a$</td>
<td>0.1</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>0.01</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>3</td>
</tr>
</tbody>
</table>

Table 2: Baseline Parameters

The annualized volatility $\sigma$ on the risky asset was selected to match a Sharpe Ratio of 0.3 when the risk-free interest rate $r$ is set to an historical average of 5% and the annualized mean return on the asset implies a risk premium of 6%. This Sharpe Ratio is similar to that of the historical average for the U.S. equity market. The equity dilution percentage $x$ was put to 80%, while the penalty rate above risk-free interest is 8.5%. These parameters were chosen to match the terms the US Federal Reserve enforced with AIG. The indebtedness $L$ of the fund was set to yield a leverage ratio of 2 at starting fund wealth $W_0 = 1$. The
absolute level of the fund’s debt will turn out not to matter much in the manager’s portfolio choice, but rather, leverage $\frac{W}{W-L}$. The manager’s inalienable wealth $a$ represents 10% of the starting value of the fund. Finally, his coefficient of relative risk aversion is 3 while the mean rate of fund liquidation implies a subjective rate of time preference of 1%. The impact on the optimal investment policy from changes in the risky asset’s volatility $\sigma$, the manager’s risk aversion $\gamma$, the level of his inalienable wealth $a$, and the bailout terms available to the government 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 (4.1).

---

Figure 4.1: Optimal Investment Policy, Baseline Parameters
Illustration of the optimal fraction invested in the risky asset at different values of portfolio wealth for two separate cumulative amounts of government transfers. 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 cumulative government transfers $G$ are 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 indebtedness of the fund: $L$ and $L + G'$, corresponding to the two values of $G$, are marked in the figure with thicker ticks on the x-axis in wealth.

For levels of wealth sufficiently far from total indebtedness, the optimal fraction invested in the risky asset is increasing in wealth. Hence, in good times, the manager engages in greater risk-taking. However, the optimal policy is to invest less in the risky asset than the fraction determined in Merton (1969), which is noted by a solid horizontal line.

In bad times, as 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 200%. In the case of $G = G'$, he does this well before wealth is even close to the bailout threshold $L$, which is where further government transfers would be only be made.

### 4.1 Interpreting the Results

The manager faces a trade-off between honoring the repayment of his total debt and taking advantage of expected profit opportunities in the risky asset. Because he is risk averse, his indebtedness becomes a sort of subsistence amount which he normally (if he lacked the inalienable wealth) would prevent the fund to dip below. Although not proved analytically, it would appear that the optimal policy here is similar to the one under habit formation preferences (Ingersoll (1992)) or 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, increasing with the cumulative amount of government transfers. This can explain 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

Interestingly, the dependence of $G$ on the path of wealth can be interpreted as introducing a kind of habit formation preference over “bad habits”. One can see this from the manager’s utility function:

$$U (G, W) = \left( a + (1 - x \cdot 1_{G>0}) (W - L - R_g G)^+ \right)^{1-\gamma}.$$
The amount $G$ is the cumulative total the fund has needed with urgency to avoid insolvency because of previous poor performance. By the terms of the bailout, this amount must eventually be deducted from portfolio wealth. Hence, there is an inter-temporal complementarity between past and current wealth. The worse the manager performed in the past by dragging the fund to insolvency, the greater is $G$, and the more regret does this past performance bring, with the penalty rate of interest measuring the strength of this regret.

Under this type of preference, the manager would invest in a way to avoid “re-living” his previous bad habits by ensuring that wealth $W$ always exceeded his subsistence amount through an adequate investment in the riskless asset.

From Figure (4.1), this subsistence amount (taken to be the level of wealth immediately before $\alpha$ abruptly increases) appears to be greater than the indebtedness $L + R_g G$. The reason for this is the following: any current level of transfers $G$ affects transfers at later times (by affecting wealth), and hence, later subsistence amounts as well. Therefore, the current subsistence amount would have to be the present value of all expected future subsistence amounts until the exogenous fund termination, given the current value of $G$. Additionally, the subsistence amount should increase with the manager’s inalienable wealth $a$. The more private wealth the manager possesses away from the fund, the less the portfolio value must fall before he would be willing to engage in the extreme risk-taking. Considering these factors, one would expect the subsistence amount to be larger than total indebtedness.

As in other subsistence or habit models, the dollar amount invested in the risky asset is proportional to wealth less the subsistence amount, so as wealth decreases, the fraction $\alpha$ invested in the risky asset declines. This is what we see here. Conversely, as wealth increases, $\alpha$ increases. And although not proved analytically, for very large values of $W$, where utility is close to CRRA with relative risk aversion $\gamma$, it would appear that the proportion in the risky asset would approach 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 can explain this behavior.

4.1.2 Roles of Inalienable Wealth, Limited Liability, and No Further Loss

Firstly, the inalienable wealth assures 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 to maintain wealth above the subsistence amount. Wealth would never decline to insolvency to trigger a bailout. With the inalienable
wealth, there is no longer this compulsion. Secondly, limited liability lets the manager even possess a form of inalienable wealth protected from creditors. Finally, the government bailout guarantees the portfolio value can bear no further decline. Any loss which would make the fund insolvent would be borne by the government. At this point, an incremental investment in the risky asset can yield the manager zero at worst, with no limit on gains. So, the manager shifts the risk to the government and invests maximally in the risky asset.

What is particularly interesting about the manager’s behavior is that he starts increasing risk after the fund has declined to near insolvency, but before actually reaching it. This behavior is owed to the fund wealth already having declined below the manager’s subsistence amount. Moreover, the threshold of wealth at which he commences the excessive risk-taking increases with greater government assistance (larger $G$). This high-risk investment raises the probability his wealth gains enough to surpass the subsistence amount, which would return the manager’s risk-taking to normal levels. A good outcome would put him far above insolvency as well as the need of further government assistance. But the policy also raises the chance the fund bears massive losses, forcing the government to step in again with more transfers. Current bailouts therefore foreshadow future bailouts.

Overall, the required repayment on any emergency bailout lending plus the sacrifice of equity ownership in the event of a bailout do induce the portfolio manager to reduce risk-taking as wealth declines. The lower risk-taking mitigates the chances of an initial bailout or further government assistance. However, if wealth declined to a point where insolvency is imminent or where the manager’s remaining equity in the fund is virtually zero, then one can expect him to abruptly take enormous risk, thus increasing the chances the government 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 to that in Panageas (2010). But the reason is very different. There, a risk neutral manager whose fund carries a bailout guarantee adopts a bang-bang investment policy: he takes on maximal risk until net worth declines to a particular threshold, at which point he reverses, and tightens risk-taking to its lowest limit. This threshold is determined endogenously so that the bailout does not become prohibitively costly for the outside stakeholder bailing out the fund to renege on its commitment. The optimal investment policy in that paper maximizes the value of the bailout guarantee while maintaining the outside stakeholder on his bailout participation constraint. A depiction of the qualitative behavior of the optimal investment policy in Panageas (2010) is given in Figure (4.2).
Figure 4.2: Optimal Investment Policy, Comparison to Panageas (2010)
Illustration of 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 is to take maximal risk in good times, but switch to minimal risk, here set to zero, if wealth declines to some endogenous level, here denoted $W^*$.

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. Moreover, the curtailing of risk in bad times is undertaken voluntarily by the manager despite an accessibility to bailout money guaranteed indefinitely, or at least until the exogenous liquidation of the fund. 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 on his own, so long as wealth stays above his subsistence level.

The manager’s optimal investment behavior as wealth approaches the bailout threshold also differs greatly between the two papers. In Panageas (2010), risk-taking would most likely have already plummeted by this point to keep the government from reneging on the bailout. Here, once wealth falls below a point granting him subsistence in the fund, he bears no further loss, and the credible promise of a fully committed bailout prompts him to take maximal risk.

Ultimately, the two papers offer alternative perspectives on investment behavior in the
presence of bailouts. If the investor is risk-neutral and there is no bailout commitment, then one should expect high risk-taking in good times when wealth is high, and abrupt reversals as wealth declines and a bailout looms, so as to safeguard against a reneging on the bailout promise (as in Panageas (2010)). Alternatively, if the manager is risk averse and the bailout is perpetually guaranteed (or he believes it to be), but the manager must repay the assistance at a penalty, one should expect risk-taking to increase in proportion to wealth, but less extremely so (as in here). And the prospect of having to pay greater penalties from accepting bailout money leads the manager to gradually taper risk as wealth declines. But once wealth falls so much that it is below his minimal “subsistence position” in the fund, then one should expect him to increase risk dramatically to benefit from the one-sided gains on his investments.

5 Comparative Statics

Here I examine the extent to which the fund manager’s optimal portfolio policy changes under different values of the model parameters. I consider changes to the volatility of the risky asset \( \sigma \), his relative risk aversion \( \gamma \), and his level of inalienable wealth \( a \).

The optimal investment policies under varying levels of risky asset volatility are depicted in Figure (5.1). Total transfers \( G \) are set to zero, so this figure illustrates the optimal investment behavior before any government bailout has taken place. The optimal policy over wealth appears to behave very similarly to the baseline case: decreasing risk as wealth declines until spiking it rapidly if wealth falls below a subsistence amount near the insolvency and bailout point \( L \). The investment fraction for sufficiently high wealth is decreasing with \( \sigma \), in accordance with the diminishing attractiveness of the risky asset investment. Though the horizontal Merton lines are not shown, each optimal policy approaches its respective Merton fraction as wealth increases. One can see an unusual uptick, however, in the case \( \sigma = 60\% \) that propels \( \alpha \) above the Merton fraction when wealth is near the upper boundary. This end behavior theoretically should not be the case and is likely owed to numerical error.

Interestingly, the increase in risk-taking occurs earliest for the highest volatility case. This likely is due to a larger subsistence threshold being met earlier as wealth declines than in the other cases. Despite total indebtedness being the same in all three cases, the prospect of future bailouts is higher for any fixed investment \( \alpha \) in the risky asset from the higher volatility. This would make the present value of all expected future government transfers larger, and hence, the subsistence amount greater and surrendered earliest as wealth declines.
Figure 5.1: **Optimal Investment Policy, Varying σ**

Illustration of the optimal fraction invested in the risky asset at different values of portfolio wealth for 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, while maintaining 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 appears to occur earlier in the highest volatility case as wealth approaches the bailout threshold.

As seen in Figure (5.2), the variation in the optimal investment policy under different levels of risk aversion follows intuition. The fraction in the risky asset for wealth above subsistence is decreasing in risk aversion. Furthermore, the more risk averse the manager, the later he engages in the maximal risk-taking as wealth approaches the bailout threshold. This delay is owed to a smaller subsistence amount, which presumably is from a lower present value of expected future government transfers because of a less risky portfolio.
Figure 5.2: **Optimal Investment Policy, Varying $\gamma$**

Illustration of the optimal fraction invested in the risky asset at different values of portfolio wealth for 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-bearing decreases with risk aversion, and the subsistence wealth threshold at which the manager accelerates risk-taking appears to come later the more risk averse the manager.

Finally, Figure (5.3) depicts variation in the optimal investment policy as the manager’s amount of inalienable wealth $a$ changes. The less private wealth the manager has “tied-up” in the fund, the earlier would he engage in high risk-taking as the portfolio value declines. This result corroborates the interpretation that his subsistence threshold in the fund is increasing in his inalienable wealth $a$. The portfolio wealth need not decline as much before passing this higher subsistence point, inducing him to take on more risk earlier. If the portfolio value is sufficiently high, then the optimal policies converge to the shared Merton line no matter the amount of inalienable wealth.

While not obvious from the figure, for the cases $a = 0.5$ and $a = 1$, the manager does in fact reduce risk as wealth approaches the subsistence point. But the region in which he does taper risk-taking within the depicted wealth space is small. Very soon does the high inalienable wealth encourage him to increase risk.
Figure 5.3: **Optimal Investment Policy, Varying a**
Illustration of the optimal fraction invested in the risky asset at different values of portfolio wealth for 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 the more inalienable wealth the manager possesses.

6 Government Equity Stake and Penalty Rate of Interest

The bailout terms available to the government are the penalty rate of interest charged on the emergency credit and the equity stake the portfolio manager must sacrifice in exchange for the bailout. In this section, I examine the impact on optimal investment when these terms change.

Figure (6.1) presents the optimal investment policy for varying rates of equity dilution. The case of no equity dilution ($x = 0$) is also observed. The optimal policy is seen to be independent of the dilution rate. This comes as no surprise given the scale invariance of the manager’s utility function over net portfolio wealth and ignoring the inalienable wealth. Provided he is not completely diluted after the initial bailout (the government takes over the
fund), then he will choose the same investment strategy, irrespective of his remaining equity stake.

![Diagram](image)

**Figure 6.1: Optimal Investment Policy, Varying \( x \)**

Illustration of the optimal fraction invested in the risky asset at different values of portfolio wealth for three separate values of the government 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. This behavior holds even for larger values of transfers \( G \).

Changes in the penalty rate of interest are seen to have some impact on the manager’s investment policy, but not in a very helpful direction. Figure (6.2) depicts the optimal policy for varying penalty spreads above the risk-free rate on the emergency credit extended in the bailout. Even a loan with payments at the risk-free interest rate is shown. So that the government interest rate can have the largest effect, the figure presents the policy over wealth when government transfers are already right below the upper bound in the space for \( G \), \( G' = 0.8 \).
Figure 6.2: **Optimal Investment Policy, Varying** $r_p$
Illustration of the optimal fraction invested in the risky asset at different values of portfolio wealth for three separate values of the government 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.

For sufficiently high levels of wealth, the investment policies look very similar. As wealth declines, they all induce the manager to shed risk from the portfolio. The higher penalty rate is seen to encourage slightly less risk-taking, but not by much. Should wealth decline enough to reach the manager’s subsistence amount, he escalates risk more as the penalty rate of interest increases. This behavior is not surprising given that a larger penalty rate will bring him to his threshold of extreme risk-taking sooner.

Perhaps a lower penalty rate of interest, then, is best, since it decreases the chance the manager takes on extreme risk? A lower rate of interest, or even an interest free loan, however, would not remove the probability of excessive risk-taking. And a loan without penalty interest extended to a failing fund would likely be politically infeasible.

The obligation alone that the manager repay the bailout money does have the effect of inducing him to reduce risk as portfolio value declines. But neither the equity dilution rate nor the penalty rate of interest are effective bailout terms at the disposable of the government to curtail the manager’s extreme risk-taking when the portfolio value has declined significantly.
If the fund has suffered many losses such that an additional bailout awaits, then he will take on the maximal risk.

This result suggests that additional bailout terms might be required to constrain the manager’s risk-taking. If the government is committed to rescuing an investment fund from insolvency, and the fund manager knows this and understands he must repay the bailout money, then explicit portfolio constraints in the bailout terms might be more effective. Constraints for the initial bailout would have to trigger prior to actual insolvency of the fund. Then they must kick-in and be enforced at increasing values of portfolio wealth the more assistance the fund has received.

7 Concluding Remarks

The key assumptions of the model that lead the portfolio manager to cut down on risk as the portfolio value declines are his requirement to repay his debt obligations and his risk aversion. His indebtedness $L + R_g G$ creates a form of subsistence wealth. This subsistence wealth, in the wake of his risk aversion, compels him to maintain the fund value above a level that also honors his obligations. Since the cumulative government transfers $G$ reflect the historical path of wealth—for they equal the amount his wealth in the past has flirted with insolvency—an alternative interpretation is that the obligation to payback with interest the bailout money $G$ equips him with a kind of habit formation preferences. But in this setting, his habit stock is the sum of his past “bad habits,” measured by poor investment performance that drove the fund wealth to insolvency, necessitating a bailout. As long as times are not that bad, and the portfolio wealth is sufficiently high, the manager will trim risk-taking when wealth declines in order to avoid another bailout and a “fall back into bad habits.”

But if wealth falls to very low levels, the incentive to mitigate risk gradually wanes as his remaining interest in the fund net of his obligations declines to zero. Once this happens, then limited liability, the possession of private wealth divorced from the fund, and the prospect of unlimited gains with no risk of loss incite him to take on maximal risk. Before any bailout has commenced, this point of extreme risk-taking is close to, but above, the outside liability amount $L$, and it is increasingly farther above $L$ the more the government grants assistance. The high risk-taking before insolvency increases the probability the fund actually becomes insolvent and requires an initial bailout. And every bailout makes a subsequent one more likely.

Direct constraints on portfolio allocation as he nears insolvency can prevent the manager from pursuing the high risk-taking, and thus reduce the chances of needing further government support. Such constraints would have to be enforced well in advance of insolvency and become increasingly stringent the more assistance the fund has received.
As a final comment, the combination of CRRA utility over net wealth and a level of subsistence discourages the manager from setting the fraction in risky holdings above the proportion in Merton (1969), unless wealth has fallen to the extreme-risk region. The risky asset proportion \( \frac{\mu - r}{\gamma \sigma^2} \) is a limiting case to which the manager tends as wealth becomes very large. Therefore, in the best of times, when portfolio wealth is high, investment funds that carry implicit unconditional bailout guarantees from the government would not bear wild amounts of risk in the model unless the Sharpe ratios of their risky assets were quite large. Only when portfolio wealth has declined substantially in bad times would one observe the extreme risk-taking.
References


A Markov Chain Approximation Method

The following procedure is borrowed heavily from Kushner and Dupuis (1992).

A.1 Bounding the State Space

As mentioned in the text, a Markov chain approximation method is adopted to numerically approximate the manager’s value function $Q$ and optimal investment policy $\alpha(t)^*$ from the stochastic control problem of (2.5). For numerical purposes, one must bound the state space of the original problem in some way to a compact set $B$. This requires that $B$ is specified along with the behavior of the system on the boundary $\partial B$. Natural lower boundaries from the manager’s problem 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 $\{G_t, W_t, t \geq 0\}$ will pass them with only small probability when initiated from their starting values and stopped at the liquidation time $\tau$. Moreover, the upper boundary should be set so that the optimal investment policy at points away from that boundary are not seriously distorted.

I suppose that the compact set $B$ is 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. As for the behavior of the state variables on the lower boundary, one of “reflection” for wealth $W$ at the value $L$ is a consequence of the manager’s problem itself. If $W$ is on the boundary of $L$ and attempts to escape below $L$, then it is instantaneously reflected back from $L$ by a compensating government transfer amount. The compensator $G$ is simultaneously incremented by this same amount. Thus, the behavior of $G$ at the lower boundary of $L$ is scheduled by the behavior of $W$.

Given this intimate relation between $W$ and $G$ at the reflecting boundary it is important to specify a vector $r(G, W) \in \mathbb{R}^2$ that captures the direction of reflection in the state space for each point within the reflecting boundary set $\partial B^R = \{(G, W) \in B : W = L\} \subset \partial B$. A natural choice is $r(G, W) = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$, which is a vector of unit length with direction $(1, 1)$. Thus, whenever $W = L$, and is about to dip below $L$, the vector $(G, W)$ is “pushed” in the direction $r$, which appropriately restores $W$ back to $L$ and increments $G$ by the restoring amount.

The behavior of the system at the “upper boundary” is now considered. A natural behavior would be that the first time $W = \bar{W}$ or $G = \bar{G}$, the two processes remain at their points of first contact with either upper boundary. Thus, the boundary behavior is “absorption”. At that point, the manager’s decision problem ends and the wealth net of liabilities are dispensed to him. 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 $(G, W) \in \partial B^\bar{W} \cup \partial B^G$ the manager’s value function would be

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

with the function $C$ defined in (2.4). The upper boundaries $\bar{W}$ and $\bar{G}$ should be set large enough so that the time it takes to reach either is large with high probability in order to avoid letting the numerical modification significantly impact the solution of the manager’s problem. Moreover, given the exogenous liquidation time $\tau$ present in the problem, one would want this time to be greater than $\tau$ with high probability.

### A.2 Discretization of the State Space

Having bounded the state space to the set $B$ and described the behavior at the boundary $\partial B$, we are now ready for its discretization.

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 $\partial B^R_h = \{(G, W) \in B_h : W = L\}$. With finite grid points, it is useful to add a set of states to which the approximating Markov chain enters when it exits $B_h$ and from which it is reflected back instantaneously into $B_h$. So, let $\partial B^+_h$ denote the additional set of points which are not in $B_h$ but are in the $h$–grid of the rectangle $B$ enlarged by $h$ on the “bottom” side of $B_h$, where $W = L$ for all $G$ in the $h$–grid. Anytime the Markov chain takes value $(G, W) \in \partial B^+_h$, it instantly reflects back into $B_h$. Therefore, we can take $\partial B^+_h$ to be the reflecting boundary of the finite state space. The upper boundaries of $\partial B^W_h = \{(G, W) \in B_h : W = \bar{W}\}$ and $\partial B^G_h = \{(G, W) \in B_h : G = \bar{G}\}$ require no such grid extension because they are absorbing. Neither does the lower boundary $\{(G, W) \in B_h : G = 0\}$ since the compensating process $G$ begins but never exits there since the process is non-decreasing. Define $S_h = B_h \cup \partial B^+_h$. This shall be the state space of the approximating controlled Markov chain.

### A.3 Local Consistency of the Approximating Markov Chain

Denote $\{\xi_n^h, n < \infty\}$ as a finite state controlled stochastic chain on $S_h$. The goal is to get an approximating Markov chain that is “locally consistent” (described below) with the manager’s wealth process given in (2.1). 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.3.1 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., \( \{(G, W) \in B - \partial B^h \} \). Here, \( dG = 0 \). In this region, the wealth process is given by

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

where \( b(W, \alpha) \) and \( s(W, \alpha) \) are abbreviations for the state-dependent drift and diffusion, respectively, of the controlled wealth process in (2.1). Denote by \( p_{h}^{\alpha} \left( (G, W), (G', W') \mid \alpha \right) \) the transition probability of the chain between states \( (G, W) \) and \( (G', W') \). The control parameter \( \alpha \) is of the original continuous state problem, so denote \( \alpha_{n}^{h} \) 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 \( \Delta t_{n}^{h} = \Delta t^{h} \left( \xi_{n}, \alpha_{n}^{h} \right) \). 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 \( \Delta \xi_{n}^{h} = \xi_{n+1}^{h} - \xi_{n}^{h} \). Let \( \mathbb{E}^{h,\alpha}_{(W,G),n} \) denote the conditional expectation given \( \{ \xi_{i}^{h}, \alpha_{i}^{h}, i \leq n, \xi_{n}^{h} = (G, W), \alpha_{n}^{h} = \alpha \} \). The chain \( \{ \xi_{n}^{h}, n < \infty \} \) satisfies a local consistency requirement (Kushner and Dupuis (1992), pp. 71), if

\[
\mathbb{E}^{h,\alpha}_{(W,G),n} \left[ \Delta \xi_{n}^{h} \right] = b(W, \alpha) \Delta t^{h} (W, \alpha) + o \left( \Delta t^{h} (W, \alpha) \right)
\]

\[
\mathbb{E}^{h,\alpha}_{(W,G),n} \left[ \Delta \xi_{n}^{h} - \mathbb{E}^{h,\alpha}_{(W,G),n} \Delta \xi_{n}^{h} \right]^{2} = s(W, \alpha)^{2} \Delta t^{h} (W, \alpha) + o \left( \Delta t^{h} (W, \alpha) \right)
\]

\[
\lim_{h \to 0} \sup_{\omega, n} \left| \Delta \xi_{n}^{h} \right| = 0,
\]

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 (A.2) are seen to match the “local properties” of the drift and diffusion in the wealth process of (A.1).

So that the chain is Markov, define a control policy \( \alpha^{h} = \{ \alpha_{n}^{h}, n < \infty \} \) as admissible if the chain has the Markov property under that policy, i.e.,

\[
P \left\{ \xi_{n+1}^{h} = (G', W') \mid \xi_{i}^{h}, \alpha_{i}^{h}, i \leq n \right\} = p^{h} \left( \xi_{n}, (G', W') \mid \alpha_{n}^{h} \right).
\]

A.3.2 Local Consistency on the Reflecting Boundary

The aim of the reflecting boundary is 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:

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

(A.3)

with the process for \( G \) explicitly given in the text. The government transfers \( G \) are not controlled by the manager, as they simply offset exactly the amount \( W \) would have dipped below \( L \). Hence, their direction are not controlled either. Thus we may use \( p^h ((G, W), (G', W')) \) to denote the transition probability function of the Markov chain \( \{ s^n, n < \infty \} \) for points \( (G, W) \in \partial B_h^+ \). This transition function is considered locally consistent with the reflection direction \( r(G, W) = \left( \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right) \) if there are constants \( \epsilon > 0, c_1 > 0 \) and function \( c_2(h) \to 0 \) as \( h \to 0 \), such that for all \( (G, W) \in \partial B_h^+ \) and all \( h \), you have (Kushner and Dupuis (1992), pp. 137):

\[
\begin{align*}
E^{h, \alpha}_{(W, G), n} [ \Delta s^n ] & \in \left\{ \theta \left( \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right) + o(h) : c_2(h) \geq \theta \geq c_1 h \right\} \\
\text{cov}^{h, \alpha}_{(W, G), n} (\Delta s^n) & = O(h^2) \\
p^h ((G, W), B_h) & \geq \epsilon_1, \forall h \text{ and } (G, W) \in \partial B_h^+ .
\end{align*}
\]  

(A.4)

Essentially, the conditions in (A.4) require the approximating chain to have an average change in direction that is the “allowable reflection” \( r(G, W) \) plus a “small” error. Moreover, the transition probability \( p^h \) of points in the reflecting boundary to points in the rest of the compact state space must be positive.

If a Markov chain is locally consistent in terms of (A.2) and also locally consistent with the reflection direction \( r(G, W) \), as defined above, then the chain is locally consistent with the reflected diffusion process for wealth in (A.3). 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 (A.3).

For the manager’s investment problem, getting a probability transition function which is locally consistent with the boundary reflection in terms of (A.4) 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 (A.4).

### 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 \(\{\xi^h_n, n < \infty\}\). Therefore, let \(\{\alpha^h_n, n < \infty\}\) be an admissible control for the chain and define the interpolated time \(t^h_n = \sum_{i=0}^{n-1} \Delta t^h_i\). Define the continuous time interpolations \(\xi^h(\cdot)\) and \(\alpha^h(\cdot)\) by:

\[
\xi^h(t) = \xi^h_n, \quad \alpha^h(t) = \alpha^h_n, \quad t \in \left[ t^h_n, t^h_{n+1} \right).
\]

The interpolated process defined by (A.5) is piecewise constant, and if it satisfies (A.2) and (A.4), it is an approximation to 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 that \(A\) is a compact set in which we have restricted the manager’s investment policy to take values. Recall that \(\partial B^\bar{W}_h\) and \(\partial B^\bar{G}_h\) are the “upper” boundaries in \(B_h\) when \(W = \bar{W}\) and \(G = \bar{G}\), respectively. Denote by \(S^+\) as the points in the \(h\)-grid away from the absorbing and reflective boundaries, i.e., \(S^+ = S_h - \partial B^\bar{W}_h - \partial B^\bar{G}_h - \partial B^+_h\). From the manager’s dynamic program in (2.6), a natural analogue using the approximating Markov chain is

\[
Q^h(G, W) = \left\{ \sup_{\alpha(\cdot) \in A_1} \left\{ e^{-\Delta t^h((G, W), \alpha)} \sum_{(G', W')} p^h((G, W), (G', W') | \alpha) Q^h(G', W') \ight. \right. \\
\left. \left. + C(G, W) \Delta t^h((G, W), \alpha) \right\} \right\} \\
C(G, W) \quad (G, W) \in S^- \quad (G', W') \in S^+ \\
Q^h(G + h, W + h) \quad (G, W) \in \partial B^\bar{W}_h \cup \partial B^\bar{G}_h
\]

(A.6)

When the Markov chain is at points within \(S_h\) but away from the absorbing boundaries \(\partial B^\bar{W}_h\) and \(\partial B^\bar{G}_h\) 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, \(p^h((G, W), (G', W') | \alpha) = 0\) for \((G', W') \in S_h\) and \(G' \neq G\).

If the Markov chain reached an absorbing boundary so that \((G, W) \in \partial B^\bar{W}_h \cup \partial B^\bar{G}_h\), then the decision problem of the manager ends and he receives the value \(C(G, W)\). Lastly, if the
state were at the reflecting boundary $\partial B_h^+$, then the manager’s value function should equal with probability one his value function when both $W$ and $G$ are incremented by $h$, since that is what happens 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 (A.6) using matrix notation. Define the vector $Q^h = \{Q^h (G, W), (G, W) \in S_h \}$ and the gain vector $J^h (\alpha) = \{J^h (G, W, \alpha), (G, W) \in S_h - \partial B_h^{\bar{W}} - \partial B_h^{\bar{G}} \}$ with components

$J^h (G, W, \alpha) = C (G, W) \Delta t^h ((G, W), \alpha)$

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

$R^h (\alpha) = \{r^h ((G, W), (G', W') | \alpha); (G, W) \in S_h - \partial B_h^{\bar{W}} - \partial B_h^{\bar{G}}, (G', W') \in S_h \}$,

where

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

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 $r^h ((G, W), (G', W') | \alpha) = 1$ for $(G', W') = (G + h, W + h)$ and zero otherwise. With these constructions, we can write the Bellman equation from (A.6) in the compact form for all values of the state space $S_h$ of the controlled Markov chain:

$Q^h = \begin{cases} 
\sup_{\alpha(.) \in A} \left[ R^h (\alpha) Q^h + J^h (\alpha) \right] & \quad (G, W) \in S_h - \partial B_h^{\bar{W}} - \partial B_h^{\bar{G}} \\
C & \quad (G, W) \in \partial B_h^{\bar{W}} \cup \partial B_h^{\bar{G}}. 
\end{cases}$ \quad (A.7)

### A.5 Constructing the Markov Chain

To solve the dynamic program in (A.7), one needs suitable interpolation intervals $\Delta t^h ((G, W), \alpha)$, as well as transition probabilities $p^h ((G, W), (G', W') | \alpha)$ for the approximating Markov chain away from the absorbing and reflecting boundaries. These two objects will formally construct the approximating Markov chain we require, provided they satisfy the local property conditions of (A.4) and (A.6). A method to obtain these intervals and transition functions is provided in Kushner and Dupuis (1992) and is known as the “finite difference” method. It turns out that both $\Delta t^h ((G, W), \alpha)$ and $p^h ((G, W), (G, W') | \alpha)$ can be generated auto-
matically from a carefully chosen finite difference approximation to the familiar differential operator of the controlled but unreflected wealth process. Moreover, this derived Markov chain would satisfy the local consistency properties of (A.2).

On the original bounded state space, let \((G, W) \in B - \partial B W - \partial B G - \partial B R\) so that the wealth process \(W\) be given by (A.1), and \(dG = 0\). Define the reward functional \(I\) for the manager’s problem as

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

Notice that \(Q (G, W) = \sup_{\alpha(\cdot)} I (G, W, \alpha)\). Assuming that \(I\) is sufficiently smooth, apply Itô’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

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

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

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

Thus, if the drift of wealth at any point is non-negative, use the forward difference, and if the drift is negative, then use the backward difference.

The term \(\lambda I\) in (A.8) that is due to discounting can be ignored for now in the finite difference method. This is acceptable since 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

$$I^h (W, \alpha) = \frac{s^2/2 + hb^+}{s^2 + h|b|} I^h (W + h, \alpha) + \frac{s^2/2 + hb^-}{s^2 + h|b|} I^h (W - h, \alpha) + C \frac{h^2}{s^2 + h|b|}, \quad (A.9)$$

where $I^h$ represents the approximated function to $I$ on the $h-$grid $S_h$ and I have suppressed the other variable $G$ upon which the function $I$ depends. One can re-write this as

$$I^h (W, \alpha) = p^h (W, W + h|\alpha) I^h (W + h) + p^h (W, W - h|\alpha) I^h (W - h) + C \Delta t^h (W, \alpha), \quad (A.10)$$

where $p^h$ and $\Delta t^h (W)$ are defined by their corresponding terms in (A.9). The terms $p^h (W, W \pm h|\alpha)$ are non-negative, and sum to unity. Therefore, they can be considered transition probabilities for a Markov chain on the subset of the state space: $S_h - \partial B^\bar{\alpha}_h - \partial B^\bar{\alpha}_h$.

However, the dependence of these probabilities in their denominators on $\alpha$ may lead to computational problems when maximizing for the value function and investment policy. A way around this is to define adjusted transition probabilities. Let

$$D^h (W) = \max_{\alpha \in A} [s^2 (W, \alpha) + h |b(W, \alpha)|].$$

Define the new transition probabilities

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

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

$$= \frac{D^h (W') - [s^2 (W, \alpha) + h |b(W, \alpha)|]}{D^h (W)}, \quad (A.11)$$

with the time increment

$$\Delta \bar{t}^h (W) = \frac{h^2}{D^h (W)}. \quad (A.12)$$

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

$$\tilde{I}^h (W, \alpha) = \sum_{W' \neq W} \bar{p}^h (W, W'|\alpha) \tilde{I}^h (W', \alpha) + \bar{p}^h (W, W|\alpha) \tilde{I}^h (W, \alpha) + C \Delta \tilde{t}^h (W). \quad (A.13)$$
This equation equals

$$ I^h(W, \alpha) = \sum_{W' \neq W} \frac{p^h(W, W'|\alpha)}{1 - \bar{p}^h(W, W|\alpha)} \bar{I}^h(W', \alpha) + C \frac{\Delta \bar{I}^h(W)}{1 - \bar{p}^h(W, W|\alpha)} $$

$$ = \sum_{W' \neq W} p^h(W, W'|\alpha) \bar{I}^h(W', \alpha) + C \Delta \bar{I}^h(W, \alpha). \quad \text{(A.14)} $$

Comparing this equation to (A.10), one can see that $I^h(W, \alpha) = \bar{I}^h(W, \alpha)$ for all controls $\alpha$ for which (A.10) or (A.14) has a unique solution. Thus we can use (A.14), which eliminates the dependence of the transition probabilities on the control, and get the same results.

The probabilities $\bar{p}^h$ of (A.11) can be considered the transition probabilities for a Markov chain on the space $S_h - \partial B^W_h - \partial B^G_h - \partial B^+ h$ with $\Delta \bar{I}^h(W)$ the associating time interval. Notice that both the transition probabilities $p^h$ of the Markov chain and time interval $\Delta \bar{I}^h(W)$ depend on the drift and diffusion of the original wealth process. The larger are the drift $|b(W, \alpha)|$ and diffusion $s^2(W, \alpha)$, the greater probability the chain moves to a different step in the time increment. Moreover, 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 coefficient $s^2$ does not affect the relative probabilities. Such characteristics of the Markov chain make it resemble rather well the local behavior of the diffusion, which is what we require. Checking for local consistency, we have

$$ E^{h,\alpha}_{(W,G),n} \left[ \Delta \xi^h_n \right] = h \frac{s^2(W, \alpha) / 2 + hb^+(W, \alpha)}{D^h(W)} - h \frac{s^2(W, \alpha) / 2 + hb^-(W, \alpha)}{D^h(W)} = b(W, \alpha) \Delta \bar{I}^h(W) $$

$$ E^{h,\alpha}_{(W,G),n} \left[ \Delta \xi^h_n - E^{h,\alpha}_{(W,G),n} \Delta \xi^h_n \right]^2 = s(W, \alpha)^2 \Delta \bar{I}^h(W) + o \left( \Delta \bar{I}^h(W) \right) $$

$$ \lim_{h \downarrow 0} \sup_{n,\omega} \left| \Delta \xi^h_n \right| = 0. $$

Thus, the controlled Markov chain with an interpolation interval determined by $\Delta \bar{I}^h(W)$ is locally consistent with the wealth process defined by (A.1).

Having all the necessary components of the Markov chain, we can express the dynamic programming equation using the controlled chain for all points $(G, W) \in S_h - \partial B^W_h - \partial B^G_h - \partial B^+ h$

$$ Q^h(G, W) = \sup_{\alpha(\cdot) \in A} \left\{ e^{-\lambda \Delta \bar{I}^h((G, W))} \sum_{G', W'} p^h((G, W), (G', W') |\alpha) Q^h(G', W') + C(G, W) \Delta \bar{I}^h(W) \right\}, \quad \text{(A.15)} $$
where I have included the appropriate discount factor $e^{-\lambda \Delta \bar{t}^h((G, W))}$. Recall that in this subset of the state space, the Markov chain can only transition between states in $W$, but not $G$. Thus, for $G' \neq G$

$$\bar{p}^h \left( (G, W), (G', W') | \alpha \right) = 0.$$  

Note that states $(G, W) \in S_h^-$ away from the absorbing and reflecting boundaries have only three non-zero transition probabilities: $\bar{p}^h \left( (G, W), (G, W \pm h) | \alpha \right)$ and $\bar{p}^h \left( (G, W), (G, W) | \alpha \right)$.

The representation of the Bellman equation in (A.15) matches (A.6), but here we have now properly specified the locally consistent transition probabilities and time intervals.

### A.6 Convergence of the Approximating Chain and the Value Function

Formal proofs that the continuous time interpolated Markov chain $\xi^h (\cdot)$ from (A.5) has a subsequence as $h \to 0$ which converges in weak measure to the controlled diffusion of (A.3) and that $Q^h$ in (A.7) 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 (A.2) and (A.4), existence of a solution to the stochastic differential equation of (A.3) 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 $W_t$ of the approximating Markov chain $\xi^h (\cdot)$. This last condition is satisfied by the properties of the Wiener process (see Kushner (1990)) and the others we have satisfied.

### A.7 Computational Method to Approximate 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 (A.7), which are re-printed below, but here emphasizing the dependence on the locally consistent transition probabilities $\bar{p}^h$ and time intervals $\Delta \bar{t}^h$ determined in the last section:

$$Q^h = \left\{ \begin{array}{ll}
\sup_{\alpha(\cdot) \in A} R^h \left( \bar{p}, \Delta \bar{t}^h, \alpha \right) Q^h + J^h \left( \Delta \bar{t}^h \right) & (G, W) \in S_h - \partial B^W_h - \partial B^G_h \\
C. & (G, W) \in \partial B^W_h \cup \partial B^G_h
\end{array} \right. \quad (A.16)$$

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 (A.16) (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 \( \left\{ \alpha_m^h (\cdot) \right\} \) for the Markov chain are computed, and the control \( \alpha_{m+1}^h (\cdot) \) is obtained after 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 \( \left\{ \alpha_m^h (\cdot), m \geq 1 \right\} \), and let \( I^h (\alpha) \) be the reward function from the matrix equation

\[
I^h (\alpha) = \begin{cases} 
R^h \left( \bar{p}^h, \Delta \bar{t}^h, \alpha \right) I^h (\alpha) + J^h \left( \Delta \bar{t}^h \right) & (G, W) \in S_h - \partial B^W_h - \partial B^G_h \\
C & (G, W) \in \partial B^W_h \cup \partial B^G_h 
\end{cases}
\]

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 \( \alpha_{m+1}^h (\cdot) \) as

\[
\alpha_{m+1}^h ((G, W)) = \arg \max_{\alpha \in A} \left[ R^h \left( \bar{p}^h, \Delta \bar{t}^h, \alpha \right) I^h (\alpha_m^h) + J^h \left( \Delta \bar{t}^h \right) \right],
\]

for \( (G, W) \in S_h - \partial B^W_h - \partial B^G_h \) then \( I^h (\alpha_m) \to Q^h \). Moreover, 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

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

until convergence is met by some criterion. This could require a great deal of calculation for each updating step of the control sequence. Instead, one could just use an integer \( K \) 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 (A.17). Note that \( K = 1 \) and \( K = \infty \) are the standard value and policy iteration methods, respectively.

The change in the modification 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).

The structure of the manager’s problem itself also furnishes an opportunity to save on computation time through a kind of domain decomposition. The manager’s value function \( Q^h \) is known at the absorbing barriers \( \partial B^W_h \) and \( \partial B^G_h \). It equals \( C (G, W) \). Moreover, the Markov chain \( \xi_n^h \) 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 $g \in \mathbb{Z}^+\text{,}$ the value function $Q^h$ and optimal control $\alpha(\cdot)$ have been determined. Begin the modified policy iteration algorithm in the $(i - 1)$ column of $G$. Search for the optimal policy over all wealth $W$ in this column using the sequence $\{\alpha^h_m\}$ until the reward function has converged, 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 $\epsilon$. This would include values of $W$ for which the Markov chain tries to escape the absorbing and reflecting boundaries. In the first case, the Markov chain is terminated and the reward function is known, while in the second case, the chain transitions with probability one back to the $i$-th column of $G$, where the reward function has already been determined. Next, move to the $(i - 2)$ column of $G$, where $G = (g - 1)h$, and repeat the algorithm again over this column of wealth values. Continue migrating left in the state space until it 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.

The choices of relevant computational-related objects are presented in Table (3).

<table>
<thead>
<tr>
<th>Object</th>
<th>Object</th>
<th>Object</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A$</td>
<td>$[0, 2]$</td>
<td>$G$</td>
</tr>
<tr>
<td>$W_0$</td>
<td>1</td>
<td>$h$</td>
</tr>
<tr>
<td>$W$</td>
<td>3</td>
<td>$\epsilon$</td>
</tr>
</tbody>
</table>

Table 3: Computational-Related Objects

The lower bound of the compact set $A$ of the policy space was set to prevent short sales of the risky asset. Since $\mu - r > 0$, the manager would never choose a short position in the risky asset in any case. The upper bound was set to 2 in order to allow him to borrow against the risky asset. It turns out that the manager only reaches the upper bound of the policy space when he is nearing his subsistence amount. The manager’s optimal policy is to increase riskiness as much as possible, so he will reach whichever upper bound is in place. Hence, the actual value of the upper bound matters little.

The manager’s starting wealth $W_0$ was set to 1 as a normalization. The upper bounds $\bar{W}$ and $\bar{G}$ of the state space were set so that they would have a low probability of being reached after, on average, $\frac{1}{\lambda} = 100$ steps of the Markov chain before the liquidation time when starting wealth $W_0 = 1$. Moreover, they were set so that the optimal policy near them was not seriously distorted. In the case of $\bar{G}$, the manager’s policy of tapering risk as wealth
declines is true for any $G$; the only difference across $G$ is at what level of wealth does he reverse course and increase risk dramatically. And the investment policy of increasing risk as $W$ increases is true for all wealth sufficiently away from the extreme risk-taking threshold. With this in mind, $\bar{W}$ was set to 3 and $\bar{G}$ was set to 1. This way, the state space could contain levels of $W$ both near and sufficiently far from the manager’s subsistence amount.

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 $\epsilon$ was set to $1e^{-6}$.