# Bailouts, the Incentive to Manage Risk, and Financial Crises

Stavros Panageas*

*University of Chicago -Booth School of Business and NBER*

April 2009

**Abstract**

A firm's termination leads to bankruptcy costs. This may create an incentive for outside stakeholders or the firm's debtholders to bail out the firm as bankruptcy looms. Because of this implicit guarantee, firm shareholders have an incentive to increase volatility in order to exploit the implicit protection. However, if they increase volatility too much they may induce the guarantee-extending parties to "walk away". I derive the optimal risk management rule in such a framework and show that it allows high volatility choices, while net worth is high. However, risk limits tighten abruptly when the firm's net worth declines below an endogenously determined threshold. Hence, the model reproduces the qualitative features of existing risk management rules, and can account for phenomena such as "flight to quality".

**Keywords:** Continuous time methods, Default, Implicit Guarantees, Risk management, Bailouts

*Contact: Stavros Panageas, 5807 S. Woodlawn Avenue, Chicago IL, 60637 USA. email: stavros.panageas@chicagogsb.edu. I would like to thank the editor, an anonymous referee, and Andy Abel, Andy Atkeson, Peter DeMarzo, Nicolae Garleanu, Rich Kihlstrom, Dirk Krueger, George Pennachi, Michael Roberts and participants of seminars, lunches and conference sessions at Chicago GSB, MIT, Wharton, Univ. of Tokyo, the Minneapolis FED, the New York FED, the BIRS center on Financial Mathematics, the NBER Summer Institute (2006), and the Western Finance Association Meetings (2008), for useful comments and discussions. Jianfeng Yu provided exceptional research assistance.

# 1 Introduction

In debating the charter for the Bank of England in 1840, Sir Robert Peel (the Prime Minister of Britain at the time) used the following words:

> While the charter is well-designed and while we are taking all precautions which legislation can prudently take against the recurrence of a monetary crisis, a crisis may occur despite of our precautions. If it does, and if it be necessary to assume grave responsibility for the purpose of meeting it, I dare say men will be found willing to assume such responsibility.

Sir Robert Peel's words are as relevant today as they were 168 years ago. As the Unites States is going through one of the worst financial crises of the last decades, it seems important to recall that the current crisis is not unique in its features. During the last few decades, the world has seen several financial crises (Asian crisis, Russian crisis etc.) that all shared a common theme: Periods of increased risk appetites, as typically evidenced by high leverage ratios, led financial institutions to the brink of bankruptcy. Bailouts and restructuring followed. At the same time, large liquidations of risky positions - sometimes referred to as "flight to quality" - exacerbated the initial negative shocks and led to prolonged periods of depressed asset valuations.

The subprime lending crisis that the United States is experiencing these days provides a reconfirmation of this general pattern: The quest for higher expected returns in the years 2004-2006 led to increased leverage ratios and lending to subprime borrowers. These developments left little margin for error when house values declined and delinquencies increased. Government sponsored bailouts followed once some of the financial instiutions were considered "too big to fail". At the same time, risky markets that attracted several participants between 2004-2006 (such as the market for collateralized debt obligations) were abandoned in a quite dramatic fashion in favor of simpler and safer investment forms.

The commonality of the structure of financial crises suggests the possibility of an economic mechanism that can simultaneously explain their recurrent features. Two phenomena seem to be of first order importance: a) the pattern of high initial risk taking followed by rapid reversals of risk appetite around the onset of a crisis and b) the prevalence of bailouts and restructuring during a crisis.

Pre-existing research has suggested that the first phenomenon may have a simple, almost mechanical explanation[1]: It is the very nature of the risk management practices followed by financial institutions that makes them prone to risk appetite reversals. Indeed, existing risk management rules (such as Value at Risk) allow high volatility choices in good times and automatically tighten the risk limits in response to declining market values. This tends to exacerbate the effects of negative shocks. Then why do such risk management rules exist in the first place? This question is important both for positive as well as normative reasons.

The present paper proposes an answer to this question. It develops a model where risk management rules are derived as optimal responses to the adverse risk taking incentives created by bailouts. Additionally, the incentives to undertake a bailout are endogenously determined, making it possible to provide a joint explanation for both the observed risk appetite reversals and the prevalence of bailouts.

Specifically, the baseline version of the model features three agents: the firm's shareholders, its debtholders and a stakeholder (such as the parent company of the firm, an insurer that guarantees principal repayment to debtholders, junior claimants, the government etc.). The stakeholder incurs a discrete cost or externality if the firm is terminated, and hence may be willing to bail out the firm, if bankruptcy looms. As one might expect, the presence of such an implicit guarantee makes the shareholders inclined to raise the volatility of the projects that they undertake. However, the stakeholder's guarantee to the shareholders is implicit and the benefit from the firm's continued presence is bounded. Therefore, high volatility choices could make it prohibitively costly for the stakeholder to bail out the firm.

In reality, this tension leads to the adoption of regulations, self-regulations, covenants, laws etc. that I will refer to as "risk management rules" or commitments. Such rules place limits on the risks that firms can take and hence serve the purpose of reassuring the

[1]See e.g. Basak and Shapiro (2001). Papers that make related points include Grossman and Zhou (1996), Basak (1995), Pavlova and Rigobon (2005), Gromb and Vayanos (2002).

stakeholder. A new aspect of the model is that rules, regulations and commitments are allowed to be imperfect, as they are likely to be in reality. The imperfection stems from the fact that future shareholders may choose to renege by paying a cost[2]. The imperfection of commitment implies that the credibility of a risk management rule is not taken as given. Instead, adherence to the rule has to be dynamically consistent.

Within this framework, I analyze the optimal choice of a risk management rule and show that it has a particularly simple form: undertake projects with high risk levels when net worth (defined as assets minus liabilities) is sufficiently high and switch to projects with low risk levels when net worth falls below an endogenously determined threshold.

The intuition for this result is simple. An optimal risk management rule should induce the stakeholder to bail out the firm, in order to avoid the deadweight cost of bankruptcy. Simultaneously, it should provide future shareholders with high continuation values, in order to reduce the temptation to renege. The optimal risk management rule achieves both of these objectives. By tightening the risk limits when net worth is low, it becomes possible to allow projects with high volatilities when the firm's assets safely exceed its liabilities. By postponing the high volatilities for the times when net worth is high, the anticipated growth rate of shareholder value is maximized. This "backloading" effect is common in many dynamic contracting contexts.[3]

This paper belongs to the continuous time literature that analyzes capital structure via contingent claim methods. This literature was initiated by the seminal Merton (1974) paper. Duffie (2001) presents a textbook treatment.[4] This literature takes the cash flow and control rights of debt and equity claims as given and uses the risk neutral pricing approach of Cox and Ross (1976) in a continuous time framework to price claims on a firm (including implicit guarantees) by option valuation techniques. The present paper contributes to this literature

[2] This helps capture situations where firms can circumvent risk management rules by undertaking costly activities such as setting up offshore, off-balance sheet entities etc.

[3] See e.g. DeMarzo and Fishman (2007), DeMarzo and Sannikov (2006).

[4] A representative sample of papers in this voluminous literature includes Ronn and Verma (1986), Leland (1994), Leland and Toft (1996), Anderson and Sundaresan (1996), Lucas and McDonald (2005), Constantinides, Donaldson, and Mehra (2002), Pennacchi and Lewis (1994).

by explicitly modeling the incentives of the shareholders to take risk and the incentives of the stakeholder to undertake a bailout.

Leland (1998) also models endogenous volatility choice. The present paper supports the results in Leland (1998), in that it shows analytically the optimality of simple Markovian "bang-bang" type volatility policies. However, the two papers have a different focus and consider different frictions and choices, so that the optimal volatility process takes a different form. Specifically, in Leland (1998) shareholders have an incentive to *increase* rather than *decrease* volatility as net worth declines and termination looms. The reason is that in Leland (1998) there are no bailouts or debt renegotiations, so that the terminal nature of bankruptcy removes the incentives to mitigate risk that are present in this paper. Therefore, in Leland (1998) the only incentives to mitigate risk result from the callability of debt. For parsimony, and in order to illuminate the new insights of the present paper, I abstract from taxes, callability, and the endogenous choice of capital structure, so that the only reason to mitigate risk is the participation constraint of the stakeholder.

The model is also related to a literature in financial economics that studies how commonly observed risk management practices can lead to variations in institutional risk taking. See e.g. Grossman and Zhou (1996), Basak (1995) Pavlova and Rigobon (2005), Basak and Shapiro (2001), Gromb and Vayanos (2002). The contribution of this paper is to understand *why* the prevailing risk management rules dictate risk limits that tighten as firm's net worth declines. In the context of optimal option exercise with idiosyncratic risk, Miao and Wang (2007) show that the precautionary motive introduces concavity into the agent's objective, and hence an incentive to mitigate volatility. In the present paper, the value of the guarantee is always a convex function of assets, and were it not for the stakeholder's participation constraint, the firm would always set volatility to its highest possible level. Motivated mostly by the Asian crisis, a literature in international economics considers the effects of bailouts for understanding crises in developing economies by taking a general equilibrium perspective.[5] However, this literature does not consider the risk-shifting incentives of shareholders

[5] See e.g. Schneider and Tornell (2004), and references therein. See also Calvo (1999), Kaminsky and Reinhart (2001), Caballero and Panageas (2005) on the related issue of contagion due to international capital

and the bailout-extension incentives of the stakeholders jointly. The paper also relates to the voluminous literature on debt, allocation of control and cash flow rights, default and reorganization, that I will not attempt to summarize here.[6] The main differences between this paper and that literature is that a) for the most part, the present paper studies the incentives to inject "new money" into a company, as opposed to splitting the existing cash flows, and b) the present paper focuses on the risk taking[7] and risk management incentives of bailouts, in an intertemporal framework with potentially imperfect commitment[8]. Methodologically, the paper uses continuous time methods to analyze an intertemporal incentive problem. Continuous time methods allow a close and explicit characterization of the solution to dynamic incentive problems. However, the present paper differs with the dynamic contracting literature,[9] since the goal is not to study the optimal design of debt and equity or the dynamic evolution of a firm's capital structure. Instead, this paper takes the capital structure as given, and focuses exclusively on the incentives to take risk and the incentives to undertake bailouts within a dynamic framework.

The structure of the paper is as follows. Section 2 presents the setup of the basic model. In order to expedite the presentation of the main result, Section 3 restricts attention to Markovian policies and derives the optimal volatility policy in that class assuming the pres-

flows.

[6] See Hart and Moore (1998) for a seminal contribution in this literature.

[7] In the context of banking, Ritchken, Thompson, DeGennaro, and Li (1993) show that charter value can create risk management incentives. However, in their simpler setup there are no commitment or strategic issues.

[8] As Leland (1998) points out, commitment (and the lack thereof) is a central issue behind the asset substitution problem of Jensen and Meckling (1976). Recent literature in economic theory and monetary economics has made advances in terms of making commitment an endogenous choice rather than imposing it as an assumption. See e.g. Caruana and Einav (2008) for game theoretic applications and Giannoni and Woodford (2002) for applications to monetary economics. In the context of the asset substitution problem studied in this paper, section 5 introduces a new approach to modeling endogenous commitment. Specifically, I let the involved parties choose both the risk management rule and how large will be the cost if the commitment is abandoned. Furthermore, I assume that higher costs of reneging (more stringent regulations or self-regulations) are associated with higher distortions. Surprisingly, it turns out that endogeneizing commitment in this way implies two results: a) Simple Markovian policies are optimal, since current decisionmakers have an incentive to choose a commitment that limits future shareholders' temptation to renege. b) The qualitative nature of the optimal risk management rule (high volatility in high net worth states, low volatility in low net worth states) is not affected by whether commitment is imposed as a regulation by the stakeholder or is voluntarily chosen by the shareholders.

[9] See e.g. DeMarzo and Fishman (2007),DeMarzo and Sannikov (2006).

ence of full commitment. Section 4 presents several realistic extensions of the baseline model and a discussion of its real world implications. Section 5 introduces the notion of imperfect/costly commitment and shows that Markovian commitments are optimal even after allowing for general (potentially history dependent) commitment policies. Assuming imperfect/costly commitment, section 6 establishes that the qualitative features of the optimal risk management rule are the same, irrespective of whether the rule is determined by the shareholders or by the stakeholder through regulation. Section 7 concludes. All proofs are relegated to the appendix.

# 2 Model

The baseline model makes a number of simplifying assumptions for expositional reasons. Several of the simplifying assumptions are relaxed in subsequent sections.

## 2.1 Lenders and the outside stakeholder

There are three types of agents in the baseline model: a continuum of competitive lenders, a continuum of anonymous shareholders and an outside stakeholder, who derives some benefit from the firm's continued existence.

The lenders hold a fixed liability of the firm in the amount $L$. This liability remains constant throughout time for simplicity. The firm also owns assets in the amount $W_t$, so that the firm's net worth at time $t$ is $W_t - L$. The assets of the firm satisfy $W_0 > L$ at time 0.

The firm is a productive entity that can never fully eliminate the risks associated with its operation. However, it can control the total volatility of its assets by costlessly adjusting the fractions it invests in projects involving high or low risk. Under the risk neutral measure[10] both high and low risk projects yield an expected return equal to the interest rate $r$ per unit

[10] Roughly speaking, pricing under the risk neutral measure means that the implied Arrow Debreu prices in the market are used to determine the value of the firm. For more details see Duffie (2001).

of time $dt$. By adjusting its portfolio of high and low risk projects, a firm can attain any level of overall asset volatility in the interval $[\sigma_1, \sigma_2]$ for all $t \geq 0$. As a result, its assets follow a geometric Brownian Motion under the risk neutral measure[11]

$$\frac{dW_t}{W_t} = rdt + \sigma_t dZ_t,\ \ \sigma_t \in [\sigma_1, \sigma_2]\,,\ \ \sigma_1 > 0 \tag{1}$$

where the drift $r > 0$ is the prevailing (real) interest rate in the economy, $dZ_t$ is a standard Brownian motion, and $\sigma_t$ presents the volatility of total assets.

To keep the analysis simple, in the baseline model the firm can pay no intermediate dividends to its shareholders until a random time $\tau$, at which time it pays a liquidating dividend in the amount of $W_\tau - L$. Intermediate dividends that are a constant fraction of the firm's assets can easily be introduced, without affecting any of the insights[12]. The firm also pays a flow of $rL$ to its lenders, up to the time of its liquidation[13].

Liquidation occurs either exogenously or endogenously. Exogenous liquidation happens at a random exponentially distributed time $\tau$ with constant hazard $\lambda > 0$. This will facilitate the use of infinite horizon optimization techniques by making all solutions independent of time. In addition to this exogenous arrival of termination, lenders can terminate the firm prior to $\tau$ : By covenant, (or because lending is secured by the assets of the firm, or is short term) they can enforce liquidation if the assets of the firm fall below its liabilities, i.e. if $W_t < L$. This assumption and the associated simplicity of the bankruptcy trigger will expedite the presentation of the results without affecting the conclusions.

When the firm gets terminated by its lenders ("endogenous liquidation"), the outside stakeholder incurs a monetary cost $B$. Purely for simplicity, I assume that this cost does not arise if liquidation is exogenous[14]. The source of the cost $B$ typically depends on whether

[11] The fact that $\sigma_1 > 0$ implies that the nature of the firm's business is such, that it can never fully eliminate risk. The idea that there is always some risk in a productive entity is a common assumption in production economies (see e.g. Cox, Ingersoll, and Ross (1985)).

[12] A previous version of this paper shows that intermediate dividends only affect the definition of certain constants.

[13] As Leland (1994), I assume that equity issuance can be used to finance the payments to the debtholders, as long as shareholder value is positive.

[14] This assumption can be easily relaxed without affecting any of the results.

the firm is non-financial or financial. For the first type of companies, the cost $B$ could have political origins (e.g. the political cost associated with increased unemployment in a region). For financial companies, the cost $B$ could be interpreted as a fire sale or bankruptcy cost due to rapid liquidation of the assets. For instance, assume that the outside stakeholder is an insurer to the lenders and has committed to incur any fire sale or bankruptcy costs in the event of a liquidation, so that debtholders do not experience any principal losses. In that case, if the firm's assets drop to $L$ and the debtholders force liquidation of the firm, there will be bankruptcy and fire sale costs in the amount of $B$, that the stakeholder will have to incur[15]. A further interpretation of $B$ as an externality occurs when a firm has claimants of different seniority that could be hit asymmetrically by bankruptcy costs[16]. $B$ could also be the result of systemic risk or it could have reputational origins. For instance, at the onset of the recent subprime lending crisis, several major investment banks bailed out structure investment vehicles or hedge funds they were sponsoring, so as to shield their claimholders from losing their invested capital.

Before proceeding, it should be noted that even though the assumption of a discrete bankruptcy cost or externality $B$ is critical for the results, the assumption about the existence of an outside stakeholder isn't. Section 4.1 presents a variant of the basic model where it is the debtholders who incur the cost $B$ rather than some outside stakeholder.

Whatever the reason for the cost or externality $B$, the outside stakeholder has the option of making transfers to the firm in order to keep its assets above $L$, and hence prevent liquidation by the lenders. In mathematical terms

$$dW_t = rW_t dt + \sigma_t W_t dZ_t + dG_t, \tag{2}$$

[15] For instance, a standard practice of major investment banks was to provide their structured investment vehicles (SIV's) with a guarantee to purchase their short term paper at fixed rates, if the need presented itself. Economically this is identical to providing a guarantee to the debtholders of the fund. Similarly, the deposit insurance agency of a given country might have to incur such bankruptcy costs in order to protect the bank's lenders.

[16] As an example, consider a firm that has debt in the amount $L = L^S + L^J$, whrere $L^S$ is senior debt and $L^J$ is junior debt. If there are bankruptcy costs in the amount $B \leq L^J$ and senior debtholders can request liquidation once the firm's assets reach $L$, then they can impose an externality on junior debtholders by requesting liquidation.

where $dG_t \geq 0$ represents incremental transfers that can be used once $W_t = L$ in order to enforce $W_t \geq L$ for all $t$. Intuitively, one should think of these injections as follows: Each time the firm's assets $W_t$ fall by an amount $\varepsilon > 0$ below $L$, the stakeholder transfers $\varepsilon$ to the firm. Since the stakeholder has no incentive to make transfers to the firm beyond the ones that are absolutely necessary to ensure its existence, one can focus on the *minimal* process that is required to keep $W_t > L$. Karatzas and Shreve (1991) (p.210-211) show that the unique minimal process for $G_t$ that will safeguard $W_t \geq L$ for all $t$ is given by

$$\frac{\int_0^t dG_s}{L} = \max\left[0, \max_{0\leq s\leq t}\left\{-\left(\log\left(W_0\right) - \log\left(L\right) + rs - \frac{1}{2}\int_0^s \sigma_u^2 du + \int_0^s \sigma_u dZ_u\right)\right\}\right]. \quad (3)$$

In the baseline model the stakeholder injects funds without receiving a share of the firm's dividends (or other form of security) in exchange for the transfers $G_t$. Section 4 enriches the model to allow for this realistic extension.

A key assumption of the model is that the stakeholder has a choice on whether to bail out the firm or not. In particular, once the assets of the firm become equal to its liabilities, the stakeholder can decide whether to make the transfers $dG_t$ or to just let the lenders seize the assets and terminate the firm.

Defining $\tau^l$ to be the time of firm liquidation (be it exogenous or lender-induced), a sufficient condition for the stakeholder to always prefer to bail out the firm is that the net present value of the costs associated with keeping the firm alive is less than the benefit of doing so

$$E_t\left(\int_t^{\tau^l} e^{-r(s-t)} dG_s | W_t = L\right) \leq B. \quad (4)$$

The expectation is taken under the risk neutral measure, and so are all expectations in the rest of the paper. Since the firm controls the volatility process $\sigma_t$, it also influences the net present value of the transfers on the left hand side of this equation.

## 2.2 Shareholders

The volatility choices of the firm are determined by its shareholders. Therefore, I use the terms "the firm" and "the shareholders" interchangeably.

To determine the value of the firm to shareholders, observe that the total value of the firm is given by $W_t + P_t$, where $P_t$ is the value of the implicit option that the stakeholder extends to the firm

$$P_t = E_t \left( \int_t^{\tau^l} e^{-r(s-t)} dG_s \right) \tag{5}$$

The total value of the firm is just equal to the sum of the claims that debtholders and shareholders hold. Letting $V_t$ denote shareholder value and $D_t$ denote debtholder value, one obtains $W_t + P_t = V_t + D_t$. Since debtholders can always induce liquidation once $W_t = L$, principal repayment is guaranteed. Accordingly[17] $D_t = L$. Using this observation, shareholder value is

$$V(W_t) = W_t - L + P_t \tag{6}$$

Equation (6) has two implications. First, different volatility processes will affect shareholder value through their effect on the value of the guarantee $P_t$. Since $P_t \geq 0$, shareholders always have an incentive to induce the stakeholder to extend the guarantee once $W_t = L$.

To check intuition, it is also useful to confirm that the firm has an incentive to set high levels of volatility in order to exploit the guarantee provided by the stakeholder. To be more specific, ignoring temporarily the constraint (4) and assuming that the stakeholder

[17] Note that for any liquidation time $\tau^l$ one obtains

$$\begin{aligned} D_t &\equiv E_t \int_t^{\tau^l} e^{-r(s-t)} rL ds + E_t e^{-r\left(\tau^l - t\right)} L \\ &= L - E_t e^{-r\left(\tau^l - t\right)} L + E_t e^{-r\left(\tau^l - t\right)} L = L, \end{aligned}$$

where the second equality follows after integrating by parts. The above equation asserts that the present value of interest payments from time $t$ and until the random time of liquidation $\tau^l$ plus the (guaranteed) repayment of $L$ at the random time of liquidation $\tau^l$ add up to the value of debt, namely $L$.

unconditionally guarantees the perpetual continuation of the firm until the time of exogenous termination, the following result holds:

**Lemma 1** *Assume that $\tau^l = \tau$ in expression (5). Assume furthermore that volatility is constant at the level $\overline{\sigma}$ for all $t \geq 0$ and define $\alpha$ as:*

$$\alpha\left(\overline{\sigma}\right) = \frac{-(r - \frac{1}{2}\overline{\sigma}^2) - \sqrt{\left(r - \frac{\overline{\sigma}^2}{2}\right)^2 + 2\overline{\sigma}^2\left(r + \lambda\right)}}{\overline{\sigma}^2} < 0 \tag{7}$$

*Then, the value of the guarantee is given by:*

$$P\left(W_t; \overline{\sigma}\right) = \frac{L}{\left|\alpha\left(\overline{\sigma}\right)\right|}\left(\frac{W}{L}\right)^{\alpha(\overline{\sigma})} \tag{8}$$

It is also straightforward to show the following result

**Lemma 2** *Assume that $\tau^l = \tau$ in expression (5). Then the volatility choice that maximizes $P_t$ is given by $\sigma_t = \sigma_2$.*

In light of the above result, if the stakeholder extended an unconditional and perpetual guarantee to the firm, then the shareholder value maximizing choice of volatility would be to set $\sigma_t$ equal to its upper bound $\sigma_2$ for all $t > 0$. This captures the standard asset substitution intuition of unconditional guarantees.

The above two Lemmas only apply if the guarantee is unconditional. The focus of this paper, however, is on guarantees that are implicit, i.e. guarantees that will only be extended if (4) is satisfied. In order to make the problem interesting, I make the following assumption:

**Assumption 1**

$$P\left(L; \sigma_1\right) = \frac{L}{\left|\alpha\left(\sigma_1\right)\right|} < B < \frac{L}{\left|\alpha\left(\sigma_2\right)\right|} = P\left(L; \sigma_2\right) \tag{9}$$

In light of Lemma 1 and equation (4), assumption 1 has two implications: a) to ensure that the firm has at least one feasible choice of volatility that will make it possible to satisfy the constraint (4) (namely by setting $\sigma_t = \sigma_1$) and b) to impose that setting volatility equal to the upper bound $\sigma_2$ for all $t > 0$ will violate the constraint (4).

## 2.3 Commitment and risk management rules

The above discussion illustrates the tension that is at the core of this paper. On the one hand, shareholders would like to set high volatility levels in order to increase the value of the implicit guarantee. On the other hand, if volatility choices are too large, then it will become too expensive for the stakeholder to extend the guarantee.

To resolve this tension, I introduce commitment via some form of regulation that I will refer as a "risk management rule". Commitment serves the purpose of reassuring the stakeholder that the firm will not exploit the implicit protection.

To expedite the presentation of the key results, this section makes several simplifying assumptions: Specifically, once the firm is started, shareholders have the ability to precommit costlessly and perfectly as to how future volatility will be determined. Furthermore, shareholders can only formulate Markovian commitments, i.e. the promised volatility choice depends exclusively on the state variable $W_t$. Finally, the risk management rule is determined in a shareholder-value maximizing way.

Section 5 relaxes all these simplifying assumptions by allowing arbitrary adapted policies (i.e. not necessarily Markovian policies). Furthermore, that section allows for the possibility of imperfect and costly commitment, in the sense that the risk management rule can be circumvented at a cost. Section 6 discusses the case where the stakeholder can impose the risk management rule on the shareholders via regulation or law, so that shareholders (rather than the stakeholder) are put against their participation constraint. The main result of sections 5 and 6 is that the features of the optimal risk management rule are not altered by these extensions.

In light of the simplifying assumptions introduced in the present section, determining the optimal risk management rule amounts to solving the following problem.

**Problem 1** *Let $\mathcal{M}$ denote the class of Markovian policies, i.e. policies of the form $\sigma_t = f(W_t)$ for some $f : [L, \infty) \to [\sigma_1, \sigma_2]$. Then for any $W_t$, choose $\sigma$ so as to maximize*

$$\max_{\sigma \in \mathcal{M}} P\left(W_t; \sigma\right) \tag{10}$$

*subject to the constraint*

$$P(L) \leq B. \tag{11}$$

In light of (6) maximizing $P\left(W_t; \sigma\right)$ is equivalent to maximizing shareholder value. Hence, the objective (10) is the familiar shareholder value maximization objective, while (11) simply re-states the stakeholder's participation constraint (4) taking into account that the Markovian nature of the volatility policies makes $P_t$ also Markovian.

# 3 Solution

## 3.1 The set of feasible payoffs

The first step towards solving problem 1 is to characterize the set of payoff functions $P(W)$ that can be attained by $\sigma(W) \in \mathcal{M}$, while also satisfying (11). This is the purpose of the next Lemma.

**Lemma 3** *Let the payoff function $P$ be defined as in (5), and assume that it satisfies constraint (11). Then the following results hold for any $\sigma(W) \in \mathcal{M}$ :*

1. *In the domain $(L, \infty)$, $P$ satisfies the ordinary differential equation*

$$\frac{\sigma^2(W)}{2} W^2 P_{WW} + r P_W W - (r + \lambda) P = 0 \tag{12}$$

2. *P is within the bounds* $0 \leq P(W) \leq B$ *for all* $W \in [L, \infty)$. *At* $+\infty$ *the function* $P$ *satisfies* $\lim_{W \to \infty} P(W) = 0$

3. $P \in \mathcal{C}^1$ *and the derivatives of* $P$ *satisfy* $P_W(L) = -1$, $P_W < 0, P_{WW} > 0$.

Lemma 3 states several properties of any feasible payoff function. The first property is a familiar Black-Scholes type differential equation. Heuristically, it can be derived by observing that $P_t$ is a "claim" whose rate of appreciation in the domain $[L, \infty)$ is equal to the sum of the interest rate $r$ and the hazard rate of termination $\lambda$

$$\frac{dE(P_t)}{dt} = (r + \lambda) P_t \tag{13}$$

Using Ito's Lemma, $\frac{dE(P_t)}{dt}$ can be expressed as $\frac{\sigma^2(W)}{2} W^2 P_{WW} + r P_W W$. Combining Ito's Lemma with (13) leads to (12).

Property 2 in Lemma 3 places upper and lower bounds on the set of feasible payoffs.[18] Property 3 has a somewhat more intricate proof, which is given in the appendix. It is however straightforward to give a heuristic intuition for the first claim, namely $P_W(L) = -1$. Assuming participation compatibility, if the firm's assets were to ever fall below $L$ by some (infinitesimal) $\varepsilon > 0$, then the stakeholder would inject resources equal to $\varepsilon$ which would reset the assets back to $L$. In mathematical terms $P(L - \varepsilon) = \varepsilon + P(L)$. Expanding the left hand side of this equation in a Taylor fashion around $L$, cancelling $P(L)$ from both sides and dividing by $\varepsilon$ gives $P_W(L) = -1$.

[18] To see why $P$ will always be between those two bounds, fix a $t \geq \tau_0$ and let $\tau^L$ be the first time (after $t$) such that $W_{\tau^L} = L$. Then

$$P(W_t) = E_t\left(e^{-(r+\lambda)(\tau^L - t)} P_{\tau^L}\right) \leq E_t\left(e^{-(r+\lambda)(\tau^L - t)} B\right) \leq B \tag{14}$$

The first equality in (14) follows from $dG_s = 0$ for all $s \in [t, \tau^L]$ . The first inequality in (14) follows by constraint (11) and the second inequality follows since $e^{-(r+\lambda)(\tau^L - t)} \leq 1$.

## 3.2 The optimization problem as a deterministic optimal control problem

Having characterized the set of all feasible payoffs in Lemma 3, it is now possible to re-write the optimization problem 1 as a standard (deterministic) optimal control problem. Note that $P(W_t)$ in the maximization problem 1 can be re-written as

$$P(W_t) = P(L) + \int_L^{\infty} P'(x) 1\{x < W_t\} dx,$$

where $1\{x < W_t\}$ is an indicator function taking the value 1 if $x < W_t$ and 0 otherwise. Assuming that the participation constraint binds, $P(L) = B$. (Lemma 4 in section 5 verifies that this constraint optimally binds). Furthermore, using the characterization of all attainable payoffs from Lemma 3, one can write the optimization problem 1 as follows:

$$\max_{\sigma(x)} \int_L^{\infty} P'(x) 1\{x < W_t\} dx \tag{15}$$

$$\begin{bmatrix} P' \\ P'' \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ \frac{2(r+\lambda)}{\sigma^2} \frac{1}{x^2} & -\frac{2r}{\sigma^2} \frac{1}{x} \end{bmatrix} \begin{bmatrix} P \\ P' \end{bmatrix} \tag{16}$$

$$\begin{bmatrix} P(L) \\ P'(L) \end{bmatrix} = \begin{bmatrix} B \\ -1 \end{bmatrix}, \lim_{x \to \infty} P(x) = 0 \tag{17}$$

Equation (16) is simply a transformation of the second order equation (12) into a system of two first order ordinary differential equations, while equation (17) gives the boundary conditions of the state variables $(P, P')$ at $L$ and $\infty$.

Having formulated problem 1 as a deterministic optimal control problem facilitates the use of standard techiques to solve it (see e.g. Leonard and Van Long (1992)). The exact details of the solution are relegated to the appendix. Here, I only give a brief sketch of the main steps.

The first step is to set up the Hamiltonian associated with the optimal control problem:

$$H = 1\{x < W_t\}P'(x) + \pi_1(x)P'(x) + \pi_2(x)\frac{2}{\sigma^2}\left((r+\lambda)\,P(x)\frac{1}{x^2} - rP'(x)\frac{1}{x}\right), \qquad (18)$$

where $\pi_1(x), \pi_2(x)$ denote the co-state variables for the two state variables $(P, P')$. Standard optimal control theory implies that the optimal volatility policies should maximize $H$. Furthermore, the fact that $P'' > 0$ (by Lemma 3) implies that $((r+\lambda)\,P(x)\frac{1}{x^2} - rP'(x)\frac{1}{x}) > 0$ (by [12]). Hence, maximizing $H$ w.r.t. $\sigma$ gives the optimal policy

$$\sigma^*(x) = \begin{cases} \sigma_1 \text{ if } \pi_2(x) > 0 \\ \sigma_2 \text{ if } \pi_2(x) < 0 \end{cases} \qquad (19)$$

Equation (19) implies that even though the firm is free to choose any $\sigma_t \in [\sigma_1, \sigma_2]$, one can restrict attention to "bang-bang" type policies, i.e. policies where only the extreme points $\sigma_1$ or $\sigma_2$ are chosen at an optimum[19]. Furthermore, if $\pi_2(x)$ is a non-increasing function of $x$, then (19) suggests that the optimal policy switches from $\sigma_1$ to $\sigma_2$ at some point $W^*$ where $\pi_2$ changes sign. Motivated by this observation, a reasonable conjecture is that, for an appropriately chosen constant $W^*$, the optimal policy is of the form

$$\sigma^*(x) = \begin{cases} \sigma_1 \text{ if } x < W^* \\ \sigma_2 \text{ if } x \geq W^* \end{cases}. \qquad (20)$$

Lemma 5 in the appendix uses policy (20) to determine a closed form solution for $P(W_t; \sigma^*)$ taking an arbitrary $W^* \geq L$ as given. It then shows that the (unique) value of $W^*$ that ensures that $P(L; \sigma^*) = B$ is given by

$$W^* = L\left[\left(\frac{\alpha_2^- - \alpha_1^+}{\alpha_2^- - \alpha_1^-}\right)\frac{\left(1 + \frac{B}{L}\alpha_1^-\right)}{\left(1 + \frac{B}{L}\alpha_1^+\right)}\right]^{\frac{1}{\alpha_1^- - \alpha_1^+}}, \qquad (21)$$

[19] This supports analytically the approach that was taken in Leland (1998), who constrained attention to two-valued policies.

where the constants $\alpha_1^{\pm}, \alpha_2^{\pm}$ are defined as

$$\alpha_1^{\pm} = \frac{-\left(r - \frac{\sigma_1^2}{2}\right) \pm \sqrt{\left(r - \frac{\sigma_1^2}{2}\right)^2 + 2\sigma_1^2 (r+\lambda)}}{\sigma_1^2} \text{ and} \tag{22}$$

$$\alpha_2^{\pm} = \frac{-\left(r - \frac{\sigma_2^2}{2}\right) \pm \sqrt{\left(r - \frac{\sigma_2^2}{2}\right)^2 + 2\sigma_2^2 (r+\lambda)}}{\sigma_2^2}. \tag{23}$$

Given this explicit expression for $W^*$, the appendix uses the conjectured policy (20) to solve the appropriate differential equations associated with the co-state variables $\pi_1(x), \pi_2(x)$ and verify that they indeed satisfy (19), which implies the optimality of (20). This leads to Proposition 1

**Proposition 1** *Let $\sigma_t^*$ be defined as in (20) with $W^*$ given by (21). Then*

$$P(W_t; \sigma_t^*) \geq P(W_t; \sigma_t)$$

*for any volatility policy $\sigma_t \in \mathcal{M}$* and for any $W_t \geq L$. *The inequality becomes an equality if $\sigma_t = \sigma_t^*$.*

In summary, this proposition implies that the firm will always follow a simple policy: keep volatility at the lower bound $\sigma_1$ while $W_t \leq W^*$, and then switch to maximal volatility $\sigma_2$ if current assets $W_t$ exceed $W^*$.

## 3.3 The intuition behind the optimal policy

Why is it optimal for the firm to lower, instead of raise volatility as its net worth declines? To see intuitively why, let $\tau_0$ be a time when $W_{\tau_0} = L$, fix a level $W_1 > L$ and let $\tau_1$ be the first time after $\tau_0$ such that $W_{\tau_1} = W_1$. Because of (4), the continuation value $P_{\tau_1}$ must satisfy the constraint

$$E_{\tau_0}\left(\int_{\tau_0}^{\tau_1} e^{-(r+\lambda)(s-\tau_0)} dG_s\right) + P_{\tau_1} E_{\tau_0}\left(e^{-(r+\lambda)(\tau_1-\tau_0)}\right) \leq B. \tag{24}$$

Equation (24) implies a trade-off between volatility choices in the time interval $[\tau_0, \tau_1]$ and volatility choices thereafter. The easiest way to see this is to re-arrange equation (24) in order to derive an upper bound for $P_{\tau_1}$:

$$P_{\tau_1} \leq \frac{B - E_{\tau_0}\left(\int_{\tau_0}^{\tau_1} e^{-(r+\lambda)(s-\tau_0)} dG_s\right)}{E_{\tau_0}\left(e^{-(r+\lambda)(\tau_1-\tau_0)}\right)} \leq \frac{B - \min_{\sigma_{s\in[\tau_0,\tau_1]}} E_{\tau_0}\left(\int_{\tau_0}^{\tau_1} e^{-(r+\lambda)(s-\tau_0)} dG_s\right)}{\min_{\sigma_{s\in[\tau_0,\tau_1]}} E_{\tau_0}\left(e^{-(r+\lambda)(\tau_1-\tau_0)}\right)}. \quad (25)$$

The rightmost expression of equation (25) gives an upper bound to the continuation value that can be assigned at time $\tau_1$. By using a similar proof as in Lemma 2, it is possible to show that the solution to the two minimization problems of equation (25) is obtained by simply setting $\sigma_{s\in[\tau_0,\tau_1]} = \sigma_1$. This is intuitive: By setting volatility at the lowest level between $\tau_0$ and $\tau_1$, it becomes possible to obtain the highest possible guarantee value at time $\tau_1$, while still satisfying (4). Hence, by "backloading" the high volatilities for the states of the world where the firm's assets are sufficiently high, one can maximize the value of the guarantee "looking forward". After a certain level of assets has been reached, the optimal risk management rule needs to switch to high levels of volatility in order to "deliver" on these high continuation values to the shareholders[20].

# 4 Extensions and Discussion

## 4.1 Absence of a stakeholder and debt forgiveness

It seems reasonable to ask if the model's predictions carry through even in cases where no stakeholder is present. To answer this question, this subsection presents a simple and stylized variant of the previous model, whereby optimal principal writedowns can produce effects that are similar to bailouts.[21]

[20] There is an interesting analogy here to results in the dynamic contracting literature. A common prediction in that literature is that optimal contracts will involve "backloading" of payoffs in order to ensure that the growth rate of continuation values is maximized. (See e.g. DeMarzo and Fishman (2007),DeMarzo and Sannikov (2006)).

[21] For a rich and tractable model that studies allocation of cash flows and strategic debt service within a dynamic valuation framework, see Anderson and Sundaresan (1996). Sundaresan and Wang (2007) study

Specifically, this subsection keeps all the assumptions made sofar with the exceptions that 1) there is no stakeholder and hence $dG_s = 0$, 2) endogenous liquidation makes the firm's assets decline by a fraction $b < 1$, i.e. $W_{\tau^+} = (1-b)\,W_{\tau^-}$, and most importantly 3) whenever $W_\tau < L_\tau$ the debtholders either have the option to declare the firm bankrupt and obtain its assets that are worth $(1-b)\,W_\tau$, or they can forgive debt $\Delta L_\tau < 0$ so as to ensure that $L_\tau + \Delta L_\tau = W_\tau$. Under these assumptions, if a firm has survived by time $t$, it means that $L_t$ is given by $L_t = \min\{L_0, \min_{0\leq s\leq t} W_s\}$.

A *sufficient* condition to induce debtholders to always prefer to write down principal is that for every time $\tau$ such that $W_\tau = L_\tau$, the value of a bankrupt firm is less than the anticipated present value of interest and principal payments. Mathematically,

$$(1-b)\,W_\tau \leq E_\tau \int_\tau^{\tau^L} e^{-r(s-\tau)}(rL_s)ds + E_\tau\left[e^{-r(\tau^L-\tau)}L_{\tau^L}\right] \text{ for all } \tau : W_\tau = L_\tau \tag{26}$$

Applying integration by parts on the right hand side of (26), using the fact that $W_\tau = L_\tau$, and simplifying gives the simpler condition:

$$-E_\tau \int_\tau^{\tau^L} e^{-r(s-\tau)}dL_s \leq bL_\tau. \tag{27}$$

Notice that (27) has a form that is quite similar to (4) except that $dG_s$ is replaced by $-dL_s$ (recall that $dL_s < 0$) and the right hand side is proportional to $L_\tau$. Assuming that equation (27) will always be satisfied, shareholder value is given by

$$V(W_t, L_t) = W_t - E_t \int_t^{\tau^L} e^{-r(s-t)}(rL_s)ds + E_t\left[e^{-r(\tau^L-t)}\left(W_{\tau^L} - L_{\tau^L}\right)\right] = W_t - L_t + \widehat{P}(W_t, L_t)$$

where $\widehat{P}(W_t, L_t)$ is defined as

$$\widehat{P}(W_t, L_t) \equiv -E_t \int_t^{\tau^L} e^{-r(s-t)}dL_s. \tag{28}$$

strategic debt service in the presence of real options.

Constraining attention to volatility policies that set $\sigma_t$ as a function of the asset to liability ratio $w_t \equiv \frac{W_t}{L_t}$, and repeating the same steps of section 3, one obtains that the optimal policy is given by[22]

$$\sigma^*(x) = \begin{cases} \sigma_1 \text{ if } x \in \left[1, \left\{\left(\frac{\alpha_2^- - \alpha_1^+}{\alpha_2^- - \alpha_1^-}\right)\left(\frac{1+b\alpha_1^- - b}{1+b\alpha_1^+ - b}\right)\right\}^{\frac{1}{\alpha_1^- - \alpha_1^+}}\right] \\ \sigma_2 \text{ if } x \geq \left\{\left(\frac{\alpha_2^- - \alpha_1^+}{\alpha_2^- - \alpha_1^-}\right)\left(\frac{1+b\alpha_1^- - b}{1+b\alpha_1^+ - b}\right)\right\}^{\frac{1}{\alpha_1^- - \alpha_1^+}}, \end{cases}$$

which has the same familiar form as in the baseline model.

## 4.2 Bailouts and Security Issuance

In the baseline model bailouts have taken the form of direct transfers. In reality, the stakeholder undertaking the bailout often obtains some form of security in exchange for injecting funds. To provide a simple example, return to the baseline model and suppose that the first time that $W_t = L$, the firm gives the stakeholder a claim to a share $x < 1$ of the firm's liquidating dividends in exchange for receiving the transfer process $G_t$ as described by equation (3).

Since the shareholder value of the firm is given by (6), and $W_t = L$, the total value of the firm (including the newly issued shares) is $P_t$. By obtaining a fraction $x$ of the firm's shares, the cost of the bailout to the shareholder is reduced to $(1-x)\,P_t$ and hence the constraint (4) becomes

$$(1-x)\,E_t\left(\int_t^{\tau^l} e^{-r(s-t)}dG_s|W_t = L\right) \leq B \tag{29}$$

More importantly, even though the value of the firm to the original shareholders is now $xP_t$, the same volatility commitments that maximize $P_t$ will maximize $xP_t$, and all the analysis of the paper can be repeated after replacing the constraint (4) with (29)[23].

[22] Details are given in the working paper version.

[23] An alternative way to think about equation (29) is as follows: Through the bailouts, the stakeholder "pays" an amount $P_t$ (this is the net present value of the transfers) to purchase shares that are worth

### 4.3 Temporary Externalities

Realistically, it is sometimes observed that firms wind down after being bailed out. The present framework allows for this, if one assumes that externalities are temporary. For instance, if the externality $B$ is due to some transient dislocation or illiquidity in markets that could become zero with some hazard rate $\beta > 0$, then all the analysis goes through after modifying equation (12) to

$$\frac{\sigma^2(W)}{2} W^2 P_{WW} + rP_W W - (r + \lambda + \beta)P = 0. \tag{30}$$

Equation (30) simply implies that the overall hazard rate of guarantee termination is $\lambda + \beta$ instead of $\lambda$.

### 4.4 Bailouts through mergers and acquisitions

Especially in government sponsored bailouts, there is pressure for the government to not bail out the existing shareholders for reasons of "fairness". However, in many such cases the government may still try to salvage the company by finding a buyer, who acquires the company with all of its assets and its liabilities. The joint surplus that the government and the buyer can obtain is $B$, i.e. the cost of liquidating the firm. In the resulting negotiations the government may agree (implicitly or explicitly) to provide transfers to the buyer, in order to incentivize him or her to acquire the company[24], and the buyer will have an incentive to give a volatility promise that satisfies the constraint (4). From this point on, one can simply repeat the analysis of the baseline version of the model. Another case where existing

$(1 - x)\, P_t$. Hence, she effectively purchases over-priced shares, which results in a net transfer to the original shareholders. Whether the transfer takes the form of outright cash injections, or purchases of "over-priced" shares has no material consequence for the analysis. It should also be clear that the above argument does not depend on the type of claim that stakeholders obtain. As long as a) the firm survives once $W_t = L$, b) the shareholder retains some non-zero cash flow rights on the liquidating dividend, and c) the debtholder gets repaid capital and interest in all states of the world, then the stakeholder must be purchasing overpriced securities.

[24]For instance in the recent bailout of Bear Stearns by JP Morgan, the government agreed to provide a muti-billion dollar credit line to JP Morgan in order to provide incentives for the acquisition.

shareholders could receive a continuation value that sets them against their participation constraint is when the government chooses the risk management rule, as in section 6.

# 5 Imperfect and Costly commitment

So far, the paper has only considered Markovian policies. A voluminous literature restricts attention to Markovian policies on a priori informational grounds.[25] However, a more important reason why the analysis sofar has focused attention on such policies, is that Markovian policies turn out to be optimal in the space of all (possibly history-dependent) policies as long as commitment is imperfect and costly rather than perfect and costless.

To introduce the notion of imperfect and costly commitment, let $\tau_0$ be the first time that $W_{\tau_0} = L$, so that the stakeholder needs to form a view as to how the firm will set volatility in the future. Importantly, from this point onward, shareholders will be able to choose arbitrary adapted volatility policies.

In reality, commitment is likely to be imperfect and costly. To model these notions, I assume that any risk management rule can be circumvented (i.e. abandoned) by future shareholders at a cost $I > 0$. For instance, in a world of imperfect accounting, shareholders can pay a cost and create legal entities or invest in off-balance sheet items that make it hard to observe or assess the value of the firm's assets, thus allowing the firm to take extra risk.

To endogeneize the extent of commitment, I assume that the parties currently involved in the formulation of a risk management rule can choose ex-ante the penalty $I$ that future shareholders will have to pay if they choose to deviate. In most real-world examples such a choice is achieved by restricting the firm's ability to invest in certain instruments. By making the restrictions more stringent, the cost $I$ of circumventing them is likely to become higher. In that sense, current shareholders and stakeholders can choose how costly it will be

[25] See e.g. Chapter 13 in the textbook of Fudenberg and Tirole (1991) for a list of papers that restrict players' actions to be Markovian. In the case where the players are firms, this restriction to "memory-less" strategies is routinely motivated by the fact that firms -unlike individuals- are run by continuously changing managers who may not have full knowledge of policies or commitments of their predecessors. However, they do know the current state variables.

to circumvent the risk management rule.

Clearly, the higher $I$, the more credible any risk management rule will become. However, it also seems plausible that rules that impose a high $I$ may have other unintended distortions: For instance, by making it very costly for a firm to engage in certain types of transactions (such as derivatives, off-balance sheet and off-shore transactions) it also becomes hard for the firm to use these instruments for tax-planning or risk sharing. To capture the idea that more "draconian" commitment devices lead to more distortions, I assume that if future shareholders' cost of reneging is set to some level $I$, then current shareholders will incur a monetary deadweight cost equal to $kI$, where $k \in (0,1)$. This is the "implementation cost" associated with the risk management rule.

A novel implication of this setup is that since commitment is costly, both the rules and the cost of reneging are jointly and endogenously determined. Different choices of $I$ between zero and infinity span the spectrum between no commitment and limitless commitment.

The following definitions formalize the notions described above. The notation $\sigma_{s>t}$ refers to the volatility process that is adopted after time $t$. Importantly, $\sigma_{s>t}$ can be an arbitrary adapted process (i.e. not necessarily Markovian).

**Definition 1** *Let $\tau^L$ be any time such that $W_{\tau^L} = L$. Then a volatility process $\sigma_s$ is* participation compatible *if*

$$P\left(L; \sigma_{s \geq \tau^L}\right) \leq B \tag{31}$$

Clearly, a stakeholder will never agree to bail out a firm at time $\tau^L$ unless constraint (31) is satisfied. The next definition captures the idea of commitment credibility.

**Definition 2** *Let $\tau^0$ be the time at which the commitment is entered. Fix a $t > \tau_0$ and let $\chi$ be the first time after $t$, such that $W_\chi = L$. For a given level of $I$, a volatility process $\sigma_{s \geq \tau^0}$*

*is* credible *if for all $t$ and $W_t$*

$$P\left(W_t;\sigma_{s\geq t}\right)\geq\sup_{\sigma_{s\in(t,\chi)}}Ee^{-(r+\lambda)(\chi-t)}\left[\sup_{\sigma_{s\geq\chi}}P(L;\sigma_{s\geq\chi})-kI\right]-I \tag{32}$$

*and $\sigma_{s\geq\chi}$ is participation compatible.*

Definition 2 captures the simple notion that the value of the guarantee under commitment and/or the cost $I$ should always be large enough, so that future shareholders will not find it optimal to pay the cost $I$ and then reset the volatility from that point on, ignoring past commitments. The term inside square brackets is the value that the shareholder can obtain by re-entering a new promise at some future time $\chi>t$, at which time the shareholders and the stakeholder will have to contemplate a new bailout. The term $\sup_{\sigma_{s\in(t,\chi)}}Ee^{-(r+\lambda)(\chi-t)}$ captures the idea that shareholders can choose volatility freely between $t$ and $\chi$, if they choose to renege.

Definition 2 implies that for any given volatility process $\sigma_{s\geq t}$, there exists a minimal amount $\widehat{I}\left(\sigma_{s\geq t}\right)$ that will make that volatility process credible.[26]

With all these definitions in hand, it is now possible to give the definition of an optimal volatility process.

**Definition 3** *Let $\widehat{I}\left(\sigma\right)=\min I\in[0,\infty)$ such that $\sigma_{s\geq\tau_0}$ is credible. A volatility process $\sigma^*_{s\geq\tau^0}$ is optimal if it is participation compatible, and*

$$P\left(L;\sigma^*_{s\geq\tau^0}\right)-k\widehat{I}\left(\sigma^*_{s\geq\tau^0}\right)\geq P\left(L;\sigma_{s\geq\tau^0}\right)-k\widehat{I}\left(\sigma_{s\geq\tau^0}\right) \tag{33}$$

*for any other participation compatible $\sigma_{s\geq\tau^0}$.*

According to this definition, a process $\sigma^*_{s\geq\tau^0}$ is optimal if it is participation compatible and maximizes the value of the implicit guarantee net of the costs that are required to ensure its credibility.

[26] Since the volatilities are bounded, both the left hand side and the second term on the right hand side are bounded.

It will be useful at this stage to make a conjecture, that is verified later. Let $\tau^L$ denote any time at which $W_{\tau^L} = L$. Then

$$P(L; \sigma^*_{s \geq \tau^L}) = B \tag{34}$$

In particular the conjecture (34) applies to time $\tau^0$.

Given conjecture (34), the search for the optimal commitment $\sigma^*_{s\geq\tau^0}$ amounts to minimizing $\widehat{I}\left(\sigma_{s\geq\tau^0}\right)$ over all policies that satisfy (34). Intuitively, shareholders would like to put the stakeholder against her participation constraint, while keeping the implementation cost $k\widehat{I}\left(\sigma_{s\geq\tau^0}\right)$ as low as possible. An additional implication of (34) is that equation (32) becomes

$$P\left(W_t; \sigma_{s\geq t}\right) \geq \left[B - kI\right] \sup_{\sigma_{s\in(t,\chi)}} Ee^{-(r+\lambda)(\chi - t)} - I = \left[B - kI\right]\left(\frac{W_t}{L}\right)^{\alpha(\sigma_2)} - I \tag{35}$$

The rightmost equality of (35) asserts that once shareholders renege, they will set volatility to the highest possible level, until they have to negotiate with the stakeholder again.[27] Given this constant volatility choice, the expression $Ee^{-(r+\lambda)(\chi-t)}$ has a simple closed form expression[28] given by $\left(W_t/L\right)^{\alpha(\sigma_2)}$.

Re-arranging equation (35) and recognizing that $\widehat{I}$ will have to be determined so that (35) holds at all times and for all levels of $W_t$ gives

$$\widehat{I}\left(\sigma_{s\geq\tau^0}\right) = \max_{W_t \geq L} \frac{B\left(\frac{W_t}{L}\right)^{\alpha(\sigma_2)} - \inf_{t\geq\tau_0} P\left(W_t; \sigma_{s\geq t}\right)}{1 + k\left(\frac{W_t}{L}\right)^{\alpha(\sigma_2)}} \tag{36}$$

Under conjecture (34), the aim of the shareholders is to choose a volatility process that will minimize $\widehat{I}$, while satisfying (34).

Equation (36) implies that $\widehat{I}$ is a decreasing function of the continuation values $P\left(W_t; \sigma_{s\geq t}\right)$. Between two participation compatible commitments, the one that implies higher continua-

[27] The proof of this fact follows the same steps as Lemma 2 and is omitted.
[28] See Øksendal (2003), p. 217.

tion values at each point in time, will be preferred since shareholders in the future will be less tempted to renege. Accordingly, $\widehat{I}$ will be lower.

An important implication of (36) is that shareholders at two times $t_1$ and $t_2$ such that $W_{t_1} = W_{t_2}$ are treated symmetrically; equation (36) implies that it is only $\min(P_{t_1}, P_{t_2})$ that matters for the determination of $\widehat{I}\,(\sigma_{s\geq\tau^0})$, whether $t_1 < t_2$ or $t_2 > t_1$. This time invariance makes Markovian policies (that set by definition $P_{t_1} = P_{t_2}$) optimal. Formally, this is shown in the next proposition.

**Proposition 2** *For* any *(potentially non-Markovian)* participation compatible *policy* $\sigma$, *there exists a lower bound* $\widehat{I}^*$, *such that* $\widehat{I}^* \leq \widehat{I}(\sigma)$. *Finally, for the Markovian policy* $\sigma^*$ *of equation (20), one obtains* $\widehat{I}^* = \widehat{I}(\sigma^*)$.

Proposition 2 shows that the policy of equation (20) is optimal, since it attains the lower bound $\widehat{I}^*$, while satisfying the constraint (34).[29][30]

The last step is to verify the conjecture (34).

**Lemma 4** *It is always optimal for constraint (31) to hold as an equality.*

# 6 Allocation of bargaining power

In several realistic situations, risk management rules are imposed by the stakeholder via regulation. Assuming costly commitment, this section shows that the distinction between regulation and self-regulation affects the rents of the two parties, but not the qualitative features of the optimal rule.

[29] A corollary of Proposition 2 that can be proven in a similar way, is that it is impossible to find any participation compatible strategy unless the firm pays at least $\widehat{I}^*$. From a practical perspective this shows that the conclusions of the model are robust to how exactly one models credibility. Specifically, definition 2 requires that a commitment be credible at *all* states and dates. This assumption may seem too strong at first. However, the above argument implies that participation compatibility *alone and by itself* places a lower bound on $I$. Hence, as long as one requires merely participation compatibility, the proposed policy of this paper is optimal.

[30] Proposition 2 does *not* assert that $\sigma^*$ is the unique policy that attains the lower bound $\widehat{I}^*$. However, it does assert that no other participation compatible (potentially non-Markovian) policy can improve on the markovian policy $\sigma^*$.

To be precise, suppose that the shareholders of the firm have some outside option when $W_\tau = L$. Such an outside option could be the result of legal difficulties in enforcing absolute priority, or more simply it could result from some scarce expertise that the shareholders can use elsewhere, if the firm gets liquidated. Whatever the source, suppose that the monetary value of this outside option is $\widetilde{v} \in (\max(P(L;\sigma_1) - \widehat{I}^*, 0), B)$. Since shareholders can always "walk away" with $\widetilde{v}$, it must be the case that

$$P(L;\sigma) - k\widehat{I}(\sigma) \geq \widetilde{v} \tag{37}$$

Now suppose that the stakeholder can determine the risk management rule and the associated punishments. Since the stakeholder is trying to minimize the value of the implicit guarantee, the stakeholder has an incentive to impose a risk management rule that will make equation (37) hold as an equality. However, as long as the assumptions of section 5 still hold, and shareholders could deviate from the prescribed policy at some cost $I$, then the optimal policy that minimizes $P(L;\sigma)$ is the one that minimizes the implementation cost $k\widehat{I}(\sigma)$. To see this, note that since (37) has to hold as an equality it follows that $P(L;\sigma) = \widetilde{v} + k\widehat{I}(\sigma)$ for any policy and hence $\min_\sigma P(L;\sigma) = \widetilde{v} + k \min_\sigma \widehat{I}(\sigma)$. (It should also be noted here, that it doesn't matter if the implementation cost $k\widehat{I}(\sigma)$ is "levied" on the shareholders or the stakeholder, since it simply reduces the joint surplus.)

From this point on, the entire analysis of the paper is applicable, with the only exception that the binding constraint $P(L) = \widetilde{v} + k\widehat{I}(\sigma)$ replaces the binding constraint (34). Since the optimal policy has always the same qualitative form irrespective of the value of $P(L)$, it follows that the optimal risk management rule remains qualitatively intact: choose low values of $\sigma_1$ when $W_t$ is lower than some threshold, and choose $\sigma_2$ when $W_t$ is above that threshold. However, the exact magnitude of the threshold, and hence the distribution of rents, does depend on whether it is the stakeholder or the shareholders that choose the optimal risk management rule.

To conclude this section, I also note that similar arguments can be used to show that if the creation of the firm creates a positive externality for the stakeholder (reduction of unemployment, increased tax revenue, increased efficiency due to additional competition etc.) then the risk-management rule will be entered at the time of the firm's creation and the distribution of rents will depend on the bargaining power of the two parties (shareholders and stakeholder)[31].

# 7 Conclusion

This paper presented a model, whereby a firm is bailed out so as to avoid costs associated with bankruptcy. The optimal actions for the stakeholder, the firm and the lenders were derived endogenously. Even though the presence of an implicit guarantee increases the shareholders' incentives to take risk, it also makes it more and more costly for the stakeholder to continue providing the implicit protection.[32]

The optimal risk management rule is to increase volatility when the firm's net worth is high and reduce volatility when its net worth declines. This policy reduces future shareholders' temptation to renege, when assets are safely above liabilities.

The predictions of the model seem to be qualitatively in line with existing risk management practices that tighten risk limits in response to declining net worth. Therefore, the model provides a potential justification for existing risk management rules, and is consistent with empirical phenomena such as flight to quality.

[31] The working paper version contains additional details.

[32] For a paper that shows that moral hazard is attenuated in an infinite horizon setting, see e.g. Panageas and Westerfield (2005).

# A Appendix

**Proof of Lemma 1.** Let $\tau^{\overline{W}}$ denote the first passage time to some $\overline{W} > W_t > L$, defined as $\tau^{\overline{W}} = \inf_{s \geq t}\{s : W_s \geq \overline{W}\}$. Consider the price of a guarantee that is terminated at either the exogenous liquidation time $\tau^l$ or $\tau^{\overline{W}}$, whichever comes first

$$P^{(\overline{W})}\left(W_t; \overline{\sigma}\right) = E_t\left(\int_t^{\tau^{\overline{W}} \wedge \tau^l} e^{-r(s-t)} dG_s\right) \tag{38}$$

It is easiest to price this claim first and then take the limit as $\overline{W} \to \infty$ in order to arrive at (8). One can use standard results to express $P^{(\overline{W})}$ as

$$P^{(\overline{W})}\left(W_t; \overline{\sigma}\right) = E_t\left(\int_t^{\tau^{\overline{W}}} e^{-(r+\lambda)(s-t)} dG_s\right) \tag{39}$$

In order to construct $P^{(\overline{W})}$ it is simplest to start by searching for a function that satisfies the following properties

$$\frac{\overline{\sigma}^2 W_s^2}{2} P_{WW}^{(\overline{W})} + r P_W^{(\overline{W})} W_s - (r+\lambda) P^{(\overline{W})} = 0 \tag{40}$$

$$P_W^{(\overline{W})}(L) = -1 \tag{41}$$

$$P^{(\overline{W})}(\overline{W}) = 0 \tag{42}$$

$$P_W^{(\overline{W})} < \infty \text{ for all } W \in \left[L, \overline{W}\right] \tag{43}$$

Finding a $P_W^{(\overline{W})}$ that satisfies (40), along with the boundary conditions (41) and (42) is straightforward. With $\alpha$ given by (7) and $\alpha^+$ defined as

$$\alpha^+ = \frac{-(r - \frac{1}{2}\overline{\sigma}^2) + \sqrt{\left(r - \frac{\overline{\sigma}^2}{2}\right)^2 + 2\overline{\sigma}^2 (r+\lambda)}}{\overline{\sigma}^2} > 0,$$

the general solution to (40) is:

$$P^{(\overline{W})}(W_t) = C_1 W_t^{\alpha} + C_2 W_t^{\alpha^+}$$

where $C_1, C_2$ are arbitrary constants. One needs to determine the constants $C_1, C_2$ so that (41) and (42) hold. Carrying out this computation, yields the following unique solution to (40), that satisfies (41), (42) and (43):

$$P^{(\overline{W})}(W_t) = \frac{\frac{L}{\alpha}\left(\frac{\overline{W}}{L}\right)^{\alpha} W_t^{\alpha^+} - \frac{L}{\alpha}\left(\frac{\overline{W}}{L}\right)^{\alpha} \overline{W}^{\alpha^+-\alpha} W_t^{\alpha}}{\overline{W}^{\alpha^+} - \frac{\alpha^+}{\alpha} L^{\alpha^+}\left(\frac{\overline{W}}{L}\right)^{\alpha}} \tag{44}$$

It is now straightforward to verify that (44) is the solution to (39), as follows: Applying Ito's Lemma to $P^{(\overline{W})}$ and taking expectations yields

$$\begin{aligned} 0 \;=\;& P^{(\overline{W})}(W_t) \\ & -E_t\left[e^{-(r+\lambda)(\tau^{\overline{W}}-t)} P^{(\overline{W})}(\overline{W})\right] + E_t\left[\int_t^{\tau^{\overline{W}}} e^{-(r+\lambda)(s-t)} \overline{\sigma} P_W^{(\overline{W})} W_s dB_s\right] \\ & +E_t\left[\int_t^{\tau^{\overline{W}}} e^{-(r+\lambda)(s-t)} \left(\frac{\overline{\sigma}^2 W_s^2}{2} P_{WW}^{(\overline{W})} + r P_W^{(\overline{W})} W_s - (r+\lambda) P^{(\overline{W})}\right) ds\right] \\ & +E_t\left[\int_t^{\tau^{\overline{W}}} e^{-(r+\lambda)(s-t)} P_W^{(\overline{W})}(L) dG_s\right] \end{aligned} \tag{45}$$

The second line in (45) is zero because of (42) and because $\overline{\sigma} P_W^{(\overline{W})} W_s$ is bounded for all $W \in [L, \overline{W}]$ by (43). The third line is zero because of (40). Hence (45) reduces to

$$P^{(\overline{W})}(W_t) = -E_t\left[\int_t^{\tau^{\overline{W}}} e^{-(r+\lambda)(s-t)} P_W^{(\overline{W})}(L) dG_s\right] \tag{46}$$

Combining (41) and (46) leads to (39). To conclude the proof, let $\overline{W} \to \infty$ in equation (44) and apply the monotone convergence theorem to obtain $\lim_{\overline{W}\to\infty} P^{(\overline{W})}(W_t) = \; P(W_t) = -\frac{L}{\alpha}\left(\frac{W_t}{L}\right)^{\alpha}$. ■

**Proof of Lemma 2.** By Lemma 1, $P^{(\sigma_2)} = P(W; \sigma_2) = -\frac{L}{\alpha}\left(\frac{W_t}{L}\right)^{\alpha}$ is convex in $W$, because $\alpha(\sigma_2) < 0$. Hence $P^{(\sigma_2)}$ satisfies the Hamilton Jacobi Bellman equation

$$\max_{\sigma \in [\sigma_1, \sigma_2]} \left\{\frac{\sigma^2}{2} W^2 P_{WW}^{(\sigma_2)}\right\} + rW P_W^{(\sigma_2)} - (r+\lambda) P^{(\sigma_2)} = 0. \tag{47}$$

The boundary conditions at $L$ and at $+\infty$ are the same as in Lemma 1. Given the continuous differentiability of $P^{(\sigma_2)}$, a classical verification theorem along the lines of Fleming and Soner (1993) implies that setting $\sigma_t = \sigma_2$ is optimal. ■

**Proof of Lemma 3.** To show result 1, let $\mathcal{U}$ be any domain of the form: $(L, W_2)$ for arbitrarily large $W_2$ such that $W_t < W_2 < \infty$. Consider now any stopping time $\tau^{\mathcal{U}}$ before $W_t$ exits the domain $\mathcal{U}$. Then, by the definition of $P$ and for any volatility process $\overline{\sigma}_t$ :

$$e^{-(r+\lambda)t}P\left(W_t\right) = E_t\left[e^{-(r+\lambda)\tau^{\mathcal{U}}}P\left(W_{\tau^{\mathcal{U}}}\right)\right]$$

This local martingale property of $e^{-(r+\lambda)t}P\left(W_t\right)$ in the domain $\mathcal{U}$ implies that (12) holds and that $P \in C^1$ (for details see Øksendal (2003), Chapter 9). The first part of the proof of result 2 is contained in the text (see equation [14]). To see why $\lim_{W\to\infty} P(W) = 0$, define $\tau^L = \inf_{s\geq t}\{s : W_s = L\}$ and note that for arbitrary $x > t$ :

$$\begin{aligned} P(W_t) &= E\left(e^{-(r+\lambda)(\tau^L-t)}E\left(\int_{\tau^L}^{\tau} e^{-r(s-\tau^L)}dG_s|W_{\tau^L}=L\right)\right) \leq E\left(e^{-(r+\lambda)(\tau^L-t)}B\right) = \\ &= \Pr\left(\tau^L < x\right)E\left(e^{-(r+\lambda)\left(\tau^L-t\right)}B|\tau^L < x\right) + \Pr\left(\tau^L \geq x\right)e^{-(r+\lambda)(x-t)}E\left(e^{-(r+\lambda)(\tau^L-x)}B|\tau^L \geq x\right) \\ &\leq B\left[\Pr\left(\tau^L < x\right) + \Pr\left(\tau^L \geq x\right)e^{-(r+\lambda)(x-t)}\right] \end{aligned} \tag{48}$$

Now, fix an arbitrary $\varepsilon > 0$ and choose large $x$ such that $e^{-(r+\lambda)(x-t)} = \frac{\varepsilon}{2B}$. The properties of Brownian motion imply that there always exists $W_t$ large enough such that $\Pr\left(\tau^L < x\right) < \frac{\varepsilon}{2B}$. In light of (48), this then implies that $P(W_t) < \varepsilon$. Since $\varepsilon$ can be chosen arbitrarily small, the result follows.

Assertion 3 contains three specific statements. The first statement is that $P_W(L) = -1$. To see why this is so, take any $\overline{W}$ and define $\overline{\tau} = \inf\{s \geq t : W_s \geq \overline{W}\}$. Applying Ito's Lemma to $P$ gives:

$$\begin{aligned} e^{-(r+\lambda)(T\wedge\overline{\tau}-t)}P\left(W_{T\wedge\overline{\tau}}\right) &= P(W_t) + \int_t^{T\wedge\overline{\tau}} e^{-(r+\lambda)(s-t)}\left(\frac{\sigma^2(W_s)}{2}W_s^2P_{WW} + rP_WW_s - (r+\lambda)P\right)ds \\ &+ \int_t^{T\wedge\overline{\tau}} e^{-(r+\lambda)(s-t)}P_W\sigma(W_s)W_sdB_s + \int_t^{T\wedge\overline{\tau}} e^{-(r+\lambda)(s-t)}P_W(L)dG_s \end{aligned}$$

Taking expectations on both sides and using equation (12) leads to:

$$\begin{aligned} P(W_t) &= -E_t\left(\int_t^{T\wedge\overline{\tau}} e^{-(r+\lambda)(s-t)}P_W(L)dG_s\right) \\ &\quad -E_t\left(\int_t^{T\wedge\overline{\tau}} e^{-(r+\lambda)(s-t)}P_W\sigma(W_s)W_s dB_s\right) \\ &\quad +E_t\left[e^{-(r+\lambda)(T\wedge\overline{\tau}-t)}P\left(W_{T\wedge\overline{\tau}}\right)\right] \end{aligned} \tag{49}$$

Since $P(W_t)$ represents the payoff of strategy $\sigma(W)$ it follows that:

$$P(W_t) = E_t\left(\int_t^{T\wedge\overline{\tau}} e^{-(r+\lambda)(s-t)}dG_s\right) + E_t\left[e^{-(r+\lambda)(T\wedge\overline{\tau}-t)}P\left(W_{T\wedge\overline{\tau}}\right)\right] \tag{50}$$

for any stopping time $\overline{\tau}$. Combining (50) and (49), it follows that:

$$E_t\left(\int_t^{T\wedge\overline{\tau}} e^{-(r+\lambda)(s-t)}[1+P_W(L)]dG_s\right) = -E_t\left(\int_t^{T\wedge\overline{\tau}} e^{-(r+\lambda)(s-t)}P_W\sigma(W_s)W_s dB_s\right) \tag{51}$$

As the differential equation (12) has a classical solution[33], $P_W$ is a continuous and hence bounded function in the closed interval $\left[L,\overline{W}\right]$.Therefore, $P_W\sigma(W_s)W_s$ is bounded in $\left[L,\overline{W}\right]$. Hence, the integrand on the right hand side of equation (51) is a martingale. Therefore, the right hand side of equation (51) is 0, and so must be the left side. This can only be the case if $P_W(L) = -1$.

The proof that $P_W < 0$ proceeds by contradiction. Assume otherwise. In particular assume that there exist a $W^{***} > L$ such that $P_W\left(W^{***}\right) > 0$. Since $P_W(L) = -1$ and the differential equation (12) has a continuous first derivative, there must be a point $\widehat{W} > L$ such that $P_W(\widehat{W}) = 0$. Since equation (12) holds at $\widehat{W}$, one obtains $(\sigma^2(\widehat{W})/2)\widehat{W}^2 P_{WW}\left(\widehat{W}\right) = (r+\lambda)P\left(\widehat{W}\right) > 0$, since $P > 0$. Hence $P_{WW}\left(\widehat{W}\right) > 0$. Therefore, at $\widehat{W}$ the function $P$ must have a local minimum. Since $P > 0$ for all $W \geq L$ and $\lim_{W\to\infty} P(W) = 0$, the function $P$ must also have a local maximum at some point $\widetilde{W} > \widehat{W}$, so that $P_W\left(\widetilde{W}\right) = 0$, and $P_{WW}\left(\widetilde{W}\right) < 0$. But this is impossible, by equation (12), since at $\widetilde{W}$ it would have to be the case that $(\sigma^2(\widetilde{W})/2)\widetilde{W}^2 P_{WW}\left(\widetilde{W}\right) = (r+\lambda)P\left(\widetilde{W}\right) > 0$, which is a contradiction to $P_{WW}\left(\widetilde{W}\right) < 0$. Hence it must be the case that $P_W(W) \leq 0$ for all $W$. Given that $P_W \leq 0$ it is now straightforward to use (12) to establish that $P_{WW} = \frac{2}{\sigma^2(W)W^2}\left[-rP_W W + (r+\lambda)P\right] > 0$. In turn, $P_{WW} > 0$ implies that $P_W$ is increasing

[33] See Øksendal (2003), Chapter 9

throughout. Moreover it can never cross 0. Hence it must be bounded between $P_W(L) = -1$ and 0, as was asserted above. ■

**Proof of Proposition 1.** The first step towards proving proposition 1 is to compute the value of the guarantee, assuming that policy (20) is followed.

**Lemma 5** *Let the constants $\alpha_1^{\pm}, \alpha_2^{\pm}$ be defined as in (22) and also let the constant $A$ be defined as*

$$A = \frac{1}{\alpha_1^- - \alpha_1^+ \frac{\alpha_2^- - \alpha_1^-}{\alpha_2^- - \alpha_1^+} \left(\frac{L}{W^*}\right)^{\alpha_1^+ - \alpha_1^-}}.$$

*With these definitions, take an arbitrary $W^* > L$ and suppose that shareholders adopt policy $\sigma^*$ of equation (20). Then $P(W_t; \sigma^*)$ is given by*

$$\frac{P(W_t;\sigma^*)}{L} = \begin{cases} A\left[\left(\frac{W_t}{L}\right)^{\alpha_1^+}\left(\frac{\alpha_2^- - \alpha_1^-}{\alpha_2^- - \alpha_1^+}\left(\frac{L}{W^*}\right)^{\alpha_1^+ - \alpha_1^-}\right) - \left(\frac{W_t}{L}\right)^{\alpha_1^-}\right] & \text{if } L \le W_t \le W^* \\ A\left[\frac{\alpha_2^- - \alpha_1^-}{\alpha_2^- - \alpha_1^+} - 1\right]\left(\frac{L}{W^*}\right)^{\alpha_2^- - \alpha_1^-}\left(\frac{W_t}{L}\right)^{\alpha_2^-} & \text{if } W_t > W^* \end{cases} \tag{52}$$

*Specifically, $P(L;\sigma^*) = B$ if and only if $W^*$ is chosen as in equation (21)*

**Proof of Lemma 5.** A detailed proof of this Lemma would replicate the same steps as Lemma 1. To save space, I only give a sketch. Applying the same logic as in Lemma 1, $P$ should satisfy:

$$0 = \begin{cases} \frac{\sigma_2^2 W^2}{2} P_{WW} + rP_W W - (r+\lambda)P \text{ if } W > W^* \ge L \\ \frac{\sigma_1^2 W^2}{2} P_{WW} + rP_W W - (r+\lambda)P \text{ if } L \le W \le W^* \end{cases}$$

The general solution to this equation is

$$P(W) = \begin{cases} C_{21} W^{\alpha_2^+} + C_{22} W^{\alpha_2^-} \text{ if } W > W^* \ge L \\ C_{11} W^{\alpha_1^+} + C_{12} W^{\alpha_1^-} \text{ if } L \le W \le W^* \end{cases}$$

where the constants $\alpha_1^+, \alpha_1^-, \alpha_2^+, \alpha_2^-$ are defined in (22). In order to be able to replicate the same steps as in Lemma 1, $P(W)$ must be continuous and continuously differentiable[34] at $W^*$. This

[34] In particular, these conditions will make it possible to apply Ito's Lemma as in Lemma 1.

implies:

$$C_{21}\left(W^*\right)^{\alpha_2^+} + C_{22}\left(W^*\right)^{\alpha_2^-} = C_{11}\left(W^*\right)^{\alpha_1^+} + C_{12}\left(W^*\right)^{\alpha_1^-} \tag{53}$$

$$\alpha_2^+ C_{21}\left(W^*\right)^{\alpha_2^+-1} + \alpha_2^- C_{22}\left(W^*\right)^{\alpha_2^--1} = \alpha_1^+ C_{11}\left(W^*\right)^{\alpha_1^+-1} + \alpha_1^- C_{12}\left(W^*\right)^{\alpha_1^--1} \tag{54}$$

To enforce $\lim_{W\to\infty} P(W) = 0$, it is also necessary to impose $C_{21} = 0$. Finally, the condition $P_W(L) = -1$ implies:

$$\alpha_1^+ C_{11}\left(L\right)^{\alpha_1^+-1} + \alpha_1^- C_{12}\left(L\right)^{\alpha_1^--1} = -1 \tag{55}$$

Solving for $C_{11}, C_{12}, C_{22}$ from equations (53), (54), (55) leads to (52). Equation (21) follows immediately by setting $P(L) = B$ and solving for $W^*$. ■

The next step is to use standard optimal control theory to derive the evolution of $\pi_1, \pi_2$ as

$$\dot{\pi}_1 = -\frac{2\left(r+\lambda\right)}{\left[\sigma^*(x)\right]^2}\pi_2\frac{1}{x^2} \tag{56}$$

$$\dot{\pi}_2 = -\left(\pi_1 + 1\{x < W_t\}\right) + r\frac{2}{\left[\sigma^*(x)\right]^2}\pi_2\frac{1}{x} \tag{57}$$

The key step towards proving Proposition 1 is to establish the existence of a solution to the system of equations (56) and (57), satisfying $\pi_2(W^*) = 0$ and

$$\pi_2(x)\begin{cases}\geq 0 \text{ if } x < W^* \\ \leq 0 \text{ if } x > W^*\end{cases} \tag{58}$$

with at least one of the two inequalities being strict for some values $x$. Furthermore, to provide sufficient conditions for the optimality of policy (20), the following properties will also be required:

$$\lim_{x\to\infty} |\pi_1(x)| < \infty \tag{59}$$

$$\lim_{x\to\infty} |\pi_2(x)| < \infty \tag{60}$$

The next Lemma constructs an explicit continuous solution to $\pi_1, \pi_2$ that satisfies (56), (57), (58), (59), (60) and $\pi_2(W^*) = 0$.

**Lemma 6** *Let $W^*$ be given by (21). Then, there exist continuous functions $\pi_1$ and $\pi_2$ that solve the pair of differential equations (56), (57) and satisfy $\pi_2(W^*) = 0$, (58), (59), (60).*

**Proof of Lemma 6.** The proof proceeds by explicitly constructing two functions that satisfy all the stated properties. Assume first that $W > W^*$. By the form of the conjectured optimal policy, one needs to distinguish 3 sub-regions for $x$ :

**(a)** $L \leq x < W^*$

**(b)** $W^* \leq x \leq W$

**(c)** $x > W$

Define the four constants $\beta_1^+, \beta_1^-, \beta_2^+, \beta_2^-$ as

$$\beta_1^{\pm} = \frac{-\left(\frac{\sigma_1^2}{2} - r\right) \pm \sqrt{\left(\frac{\sigma_1^2}{2} - r\right)^2 + 2\sigma_1^2 (r+\lambda)}}{\sigma_1^2} \text{ and } \beta_2^{\pm} = \frac{-\left(\frac{\sigma_2^2}{2} - r\right) \pm \sqrt{\left(\frac{\sigma_2^2}{2} - r\right)^2 + 2\sigma_2^2 (r+\lambda)}}{\sigma_2^2}$$

In light of the conjectured optimal policy, in region (a) the differential equation (56), (57) has the general solution:

$$\begin{aligned} \pi_1(x) &= D_{11} x^{\beta_1^+} + D_{21} x^{\beta_1^-} - 1 \\ \pi_2(x) &= -\frac{\sigma_1^2 \beta_1^+}{2(r+\lambda)} D_{11} x^{\beta_1^+ + 1} - \frac{\sigma_1^2 \beta_1^-}{2(r+\lambda)} D_{21} x^{\beta_1^- + 1} \end{aligned}$$

for appropriate constants $D_{11}, D_{21}$. Similarly, in region (b) the general solution is:

$$\begin{aligned} \pi_1(x) &= D_{12} x^{\beta_2^+} + D_{22} x^{\beta_2^-} - 1 \\ \pi_2(x) &= -\frac{\sigma_2^2 \beta_2^+}{2(r+\lambda)} D_{12} x^{\beta_2^+ + 1} - \frac{\sigma_2^2 \beta_2^-}{2(r+\lambda)} D_{22} x^{\beta_2^- + 1} \end{aligned}$$

and in region (c):

$$\begin{aligned} \pi_1(x) &= D_{13} x^{\beta_2^+} + D_{23} x^{\beta_2^-} \\ \pi_2(x) &= -\frac{\sigma_2^2 \beta_2^+}{2(r+\lambda)} D_{13} x^{\beta_2^+ + 1} - \frac{\sigma_2^2 \beta_2^-}{2(r+\lambda)} D_{23} x^{\beta_2^- + 1} \end{aligned}$$

It remains to determine the six constants in the above equations in order to obtain the solution to $\pi_1, \pi_2$. Starting with region (c), it is clear that (59), (60) can only hold if $D_{13} = 0$, since $\beta_2^+ > 0$. To ensure continuity of $\pi_1(x), \pi_2(x)$ at point $W$, the constants $D_{23}, D_{12}, D_{22}$ need to satisfy (after some straightforward cancellations):

$$\begin{aligned} D_{12} W^{\beta_2^+} + (D_{22} - D_{23}) W^{\beta_2^-} &= 1 && (61) \\ -\beta_2^+ D_{12} W^{\beta_2^+ + 1} - \beta_2^- (D_{22} - D_{23}) W^{\beta_2^- + 1} &= 0 && (62) \end{aligned}$$

Similarly, continuity of $\pi_1(x), \pi_2(x)$ at $W^*$ implies that

$$\begin{aligned} D_{11} (W^*)^{\beta_1^+} + D_{21} (W^*)^{\beta_1^-} &= D_{12} (W^*)^{\beta_2^+} + D_{22} (W^*)^{\beta_2^-} \\ -\beta_1^+ D_{11} (W^*)^{\beta_1^+ + 1} - \beta_1^- D_{21} (W^*)^{\beta_1^- + 1} &= -\left(\frac{\sigma_2}{\sigma_1}\right)^2 \left[\beta_2^+ D_{12} (W^*)^{\beta_2^+ + 1} + \beta_2^- D_{22} (W^*)^{\beta_2^- + 1}\right] \end{aligned}$$

Finally, to ensure $\pi_2(W^*) = 0$ it must also be the case that

$$-\beta_1^+ D_{11} (W^*)^{\beta_1^+ + 1} - \beta_1^- D_{21} (W^*)^{\beta_1^- + 1} = 0 \qquad (63)$$

Solving this system of equations leads to the following solution for $\pi_1, \pi_2$ :

**(a)** $L \leq x < W^*$

$$\begin{aligned} \pi_1(x) &= \left(\frac{W^*}{W}\right)^{\beta_2^+} \frac{1}{(\beta_1^+ - \beta_1^-)} \left(\frac{x}{W^*}\right)^{\beta_1^-} \left[\beta_1^+ - \beta_1^- \left(\frac{x}{W^*}\right)^{\beta_1^+ - \beta_1^-}\right] - 1 \\ \pi_2(x) &= -\frac{\sigma_1^2}{2(r+\lambda)} \left(\frac{W^*}{W}\right)^{\beta_2^+} \frac{\beta_1^- \beta_1^+}{(\beta_1^+ - \beta_1^-)} \left(\frac{x}{W^*}\right)^{\beta_1^-} \left[1 - \left(\frac{x}{W^*}\right)^{\beta_1^+ - \beta_1^-}\right] x \end{aligned}$$

**(b)** $W^* \leq x \leq W$

$$\begin{aligned} \pi_1(x) &= \frac{1}{(\beta_2^+ - \beta_2^-)} \left(\frac{x}{W}\right)^{\beta_2^-} \left[\beta_2^+ \left(\frac{W^*}{W}\right)^{\beta_2^+ - \beta_2^-} - \beta_2^- \left(\frac{x}{W}\right)^{\beta_2^+ - \beta_2^-}\right] - 1 \\ \pi_2(x) &= -\frac{\sigma_2^2}{2(r+\lambda)} \frac{\beta_2^+ \beta_2^-}{(\beta_2^+ - \beta_2^-)} \left(\frac{x}{W}\right)^{\beta_2^-} \left[\left(\frac{W^*}{W}\right)^{\beta_2^+ - \beta_2^-} - \left(\frac{x}{W}\right)^{\beta_2^+ - \beta_2^-}\right] x \end{aligned}$$

**(c)** $x > W$

$$\pi_1(x) = \frac{\beta_2^+}{(\beta_2^+ - \beta_2^-)} \left(\frac{W^*}{W}\right)^{\beta_2^-} \left[\left(\frac{W^*}{W}\right)^{\beta_2^+ - \beta_2^-} - 1\right] \left(\frac{x}{W^*}\right)^{\beta_2^-}$$

$$\pi_2(x) = -\frac{\sigma_2^2}{2(r+\lambda)} \frac{\beta_2^+ \beta_2^-}{(\beta_2^+ - \beta_2^-)} \left(\frac{W^*}{W}\right)^{\beta_2^-} \left[\left(\frac{W^*}{W}\right)^{\beta_2^+ - \beta_2^-} - 1\right] \left(\frac{x}{W^*}\right)^{\beta_2^-} x$$

By construction, $\pi_1(x), \pi_2(x)$ are continuous and satisfy $\pi_2(W^*) = 0$, (59), (60). It remains to verify that this solution also satisfies (58). This follows from $\beta_2^+ > 0, \beta_2^- < 0$ and also $\beta_1^+ > 0, \beta_1^- < 0$. The proof for $W < W^*$ follows similar steps and is therefore omitted. ■

**Proof of Proposition 1 continued.** Given the existence of an appropriate pair of co-state variables $\pi_1, \pi_2$ that solve the pair of differential equations (56), (57) and satisfy $\pi_2(W^*) = 0$, (58), (59), (60), it is now possible to verify optimality by using a standard sufficiency theorem of optimal control (see e.g. Leonard and Van Long (1992), p. 289). ■

**Proof of Proposition 2.** Let $\Pi$ be the set of all participation compatible policies $\sigma$. The first step in order to obtain $\widehat{I}^*$ is to find a function $g(W_t)$ such that

$$g(W_t) \geq \inf_{t \geq \tau_0} P(W_t; \sigma_{s \geq t}) \text{ for all } \sigma \in \Pi \tag{64}$$

Constructing such an upper bound is straightforward. First, fix a level $W_1 > L$ and let $\tau_1$ be the first time after $\tau_0$ such that $W_t = W_1$. An upper bound to $\inf_{t \geq \tau_0} P(W_1; \sigma_{s \geq t})$ is given by the highest possible value $P_{\tau_1}$ that can be assigned by any participation compatible policy. Two observations are useful in order to determine $P_{\tau_1}$. The first observation is that $P_{\tau_1}$ must satisfy the constraint $E_{\tau_0}\left(\int_{\tau_0}^{\tau_1} e^{-(r+\lambda)(s-\tau_0)} dG_s\right) + P_{\tau_1} E_{\tau_0}\left(e^{-(r+\lambda)(\tau_1-\tau_0)}\right) \leq B$, which can be rewritten as

$$P_{\tau_1} \leq \frac{B - E_{\tau_0}\left(\int_{\tau_0}^{\tau_1} e^{-(r+\lambda)(s-\tau_0)} dG_s\right)}{E_{\tau_0}\left(e^{-(r+\lambda)(\tau_1-\tau_0)}\right)} \tag{65}$$

Using an argument similar to the proof of Lemma 2, one can show that[35] setting $\sigma_s = \sigma_1$ for all $s \in [\tau_0, \tau_1]$ will minimize both $E_{\tau_0}\left(\int_{\tau_0}^{\tau_1} e^{-(r+\lambda)(s-\tau_0)} dG_s\right)$ and $E\left(e^{-(r+\lambda)(\tau_1-\tau_0)}\right)$, and hence

[35] The proof of this fact follows steps similar to Lemma 2 and is ommited.

will maximize the right hand side of (65). This is intuitive. In order to have the highest possible flexibility to promise a high level of $P_{\tau_1}$, one needs to set volatility prior to $\tau_1$ as low as possible. More importantly, this simple observation suggests that it is possible to find an explicit expression for the right hand side of equation (65). In particular, let $u(W)$ be the solution to the differential equation $\frac{\sigma_1^2}{2} u_{WW} W^2 + r u_W W - (r+\lambda) u = 0$, subject to the boundary conditions $u(L) = B$, and $u_W(L) = -1$. There is a unique solution to this equation which is given by

$$u(W) = L\left[\frac{\alpha_1^+ \frac{B}{L} + 1}{\alpha_1^+ - \alpha_1^-}\left(\frac{W}{L}\right)^{\alpha_1^-} - \frac{\alpha_1^- \frac{B}{L} + 1}{\alpha_1^+ - \alpha_1^-}\left(\frac{W}{L}\right)^{\alpha_1^+}\right]. \tag{66}$$

By arguments similar to the ones used in the proof of Lemma 3, the function $u(W)$ satisfies:

$$u(W_1) = \frac{B - E_{\tau_0}\left(\int_{\tau_0}^{\tau_1} e^{-(r+\lambda)(s-\tau_0)} dG_s | \sigma_{s\in[\tau_0,\tau_1]} = \sigma_1\right)}{E_{\tau_0}\left(e^{-(r+\lambda)(\tau_1-\tau_0)} | \sigma_{s\in[\tau_0,\tau_1]} = \sigma_1\right)} \tag{67}$$

and hence it gives a closed form expression for the right hand side of (65). Letting $\chi$ be the first time after $\tau_1$ such that $W_\chi = L$, a second observation about $P_{\tau_1}$ is that it is bounded above by

$$P_{\tau_1} \le \max_{\sigma} E\left(e^{-(r+\lambda)(\chi-\tau_1)}\right) B = E\left(e^{-(r+\lambda)(\chi-\tau_0)} | \sigma_{s\in[\tau_1,\chi]} = \sigma_2\right) B = B\left(\frac{W_1}{L}\right)^{\alpha_2^-}. \tag{68}$$

The above observations, together with the fact that $W_1$ is arbitrary, imply that the function $g(W) \equiv \min\left[u(W), B\left(\frac{W}{L}\right)^{\alpha_2^-}\right]$ satisfies the equation (64). In turn, this implies that for any participation compatible policy

$$\widehat{I}(\sigma) \ge \widehat{I}^* \equiv \max_{W>L}\left[\frac{B\left(\frac{W}{L}\right)^{\alpha_2^-} - g(W)}{1 + k\left(\frac{W}{L}\right)^{\alpha(\sigma_2)}}\right] = \max_{W>L}\left[\frac{B\left(\frac{W}{L}\right)^{\alpha_2^-} - g(W)}{1 + k\left(\frac{W}{L}\right)^{\alpha_2^-}}\right], \tag{69}$$

where the rightmost equality follows from $\alpha(\sigma_2) = \alpha_2^-$. It will be useful to establish a few properties of the expression inside the square brackets of (69). To this end define $n(W) = \frac{B\left(\frac{W}{L}\right)^{\alpha_2^-} - g(W)}{1+k\left(\frac{W}{L}\right)^{\alpha_2^-}}$. By its definition $g(L) = B$, and hence $n(L) = 0$. Also, the definition of $g(W)$ implies that $n(W) \ge 0$. Moreover, $u_W(L) = -1$ and assumption (9) implies that $\frac{d\left[B\left(\frac{W}{L}\right)^{\alpha_2^-}\right]}{dW_{|W_t=L}} = \alpha_2^- \frac{B}{L} > -1$. These two last facts can be used to show that $n_W(L) > 0$, and hence $n > 0$ in a neighborhood of $L$.

Finally, by assumption (9), $\alpha_1^- \frac{B}{L} + 1 < 0$. Hence, $u(W) \to \infty$ and, since $\alpha_1^+ > 1$, $u_W \to 0$ as

$W \to \infty$. By contrast $B\left(\frac{W}{L}\right)^{\alpha_2^-} \to 0$ as $W \to \infty$ and the derivative of $B\left(\frac{W}{L}\right)^{\alpha_2^-}$ is always negative. Hence there always exists a value $W^u$, such that $g(w) = B\left(\frac{w}{L}\right)^{\alpha_2^-}$ for all $w \geq W^u$. Therefore $n(w) = 0$ for all $w \geq W^u$. Since the function $n$ starts at 0 when $W = L$, and becomes 0 for all $W \geq W^u$, and is positive and continuous for $W \in [L, W^u]$, it must attain a maximum at some point $W^{**}$ between $L$ and $W^u$. To compute this maximum it is easiest to take the log of $n(W)$, differentiate with respect to $W$ and set the resulting expression equal to 0 to obtain

$$\frac{\alpha_2^- \frac{B}{L}\left(\frac{W^{**}}{L}\right)^{\alpha_2^- -1} - \left[\alpha_1^- \frac{\alpha_1^+ \frac{B}{L}+1}{\alpha_1^+ - \alpha_1^-}\left(\frac{W^{**}}{L}\right)^{\alpha_1^- -1} - \alpha_1^+ \frac{\alpha_1^- \frac{B}{L}+1}{\alpha_1^+ - \alpha_1^-}\left(\frac{W^{**}}{L}\right)^{\alpha_1^+ -1}\right]}{\frac{B}{L}\left(\frac{W^{**}}{L}\right)^{\alpha_2^-} - \left[\frac{\alpha_1^+ \frac{B}{L}+1}{\alpha_1^+ - \alpha_1^-}\left(\frac{W^{**}}{L}\right)^{\alpha_1^-} - \frac{\alpha_1^- \frac{B}{L}+1}{\alpha_1^+ - \alpha_1^-}\left(\frac{W^{**}}{L}\right)^{\alpha_1^+}\right]} = \frac{\alpha_2^- k \left(\frac{W^{**}}{L}\right)^{\alpha_2^- -1}}{1 + k\left(\frac{W^{**}}{L}\right)^{\alpha_2^-}} \tag{70}$$

Straightforward, but tedious algebra shows that this equation has a unique root. Having obtained $\widehat{I}^*$ as a lower bound on $\widehat{I}(\sigma)$, it is now possible to verify the optimality of the policy $\sigma^*$ of equation (20), by showing that $\widehat{I}(\sigma^*) = \widehat{I}^*$. As a first step towards showing this, I use the quantity $W^*$ as defined in equation (21) and show that $W^* < W^{**}$. After some manipulations one can verify that

$$\frac{\alpha_2^- \frac{B}{L}\left(\frac{W^*}{L}\right)^{\alpha_2^- -1} - \left[\alpha_1^- \frac{\alpha_1^+ \frac{B}{L}+1}{\alpha_1^+ - \alpha_1^-}\left(\frac{W^*}{L}\right)^{\alpha_1^- -1} - \alpha_1^+ \frac{\alpha_1^- \frac{B}{L}+1}{\alpha_1^+ - \alpha_1^-}\left(\frac{W^*}{L}\right)^{\alpha_1^+ -1}\right]}{\frac{B}{L}\left(\frac{W^*}{L}\right)^{\alpha_2^-} - \left[\frac{\alpha_1^+ \frac{B}{L}+1}{\alpha_1^+ - \alpha_1^-}\left(\frac{W^*}{L}\right)^{\alpha_1^-} - \frac{\alpha_1^- \frac{B}{L}+1}{\alpha_1^+ - \alpha_1^-}\left(\frac{W^*}{L}\right)^{\alpha_1^+}\right]} = \alpha_2^- \left(\frac{W^*}{L}\right)^{-1} < \frac{\alpha_2^- k \left(\frac{W^*}{L}\right)^{\alpha_2^- -1}}{1 + k\left(\frac{W^*}{L}\right)^{\alpha_2^-}} \tag{71}$$

where the equality follows from (53)-(54) and the inequality follows from $\alpha_2^- < 0$, $k < 1$, $\frac{W^*}{L} > 1$. Combining (70) and (71) shows that $n_W(W^*) < 0$. Hence it must be the case that $W^{**} < W^*$.

Since the functions $P(W)$ of equation (52) and $u(W)$ coincide between $L$ and $W^*$, and $W^{**} < W^*$, it follows that

$$\widehat{I}^* = \max_{L<W}\left[\frac{B\left(\frac{W}{L}\right)^{\alpha_2^-} - g(W)}{1 + k\left(\frac{W}{L}\right)^{\alpha_2^-}}\right] = \max_{L<W<W^*}\left[\frac{B\left(\frac{W}{L}\right)^{\alpha_2^-} - u(W)}{1 + k\left(\frac{W}{L}\right)^{\alpha_2^-}}\right] = \max_{L<W<W^*}\left[\frac{B\left(\frac{W}{L}\right)^{\alpha_2^-} - P(W)}{1 + k\left(\frac{W}{L}\right)^{\alpha_2^-}}\right]. \tag{72}$$

The first equation in (72) is the definition of $\widehat{I}^*$, the second and third equations follow from the fact that $u(W) = P(W) < B(W/L)^{\alpha_2^-}$ for all $W \in (L, W^*]$. The final step of the proof is to verify that

$$\max_{L\leq W\leq W^*}\left[\frac{B\left(\frac{W}{L}\right)^{\alpha_2^-}-P(W)}{1+k\left(\frac{W}{L}\right)^{\alpha_2^-}}\right]=\max_{L<W}\left[\frac{B\left(\frac{W}{L}\right)^{\alpha_2^-}-P(W)}{1+k\left(\frac{W}{L}\right)^{\alpha_2^-}}\right]=\widehat{I}(\sigma^*). \tag{73}$$

This follows from the fact that $\left[B\left(\frac{W}{L}\right)^{\alpha_2^-}-P(W)\right]/\left[1+k\left(\frac{W}{L}\right)^{\alpha_2^-}\right]$ is a declining function of $W$ for $W>W^*$, and hence

$$\max_{W>W^*}\left[\frac{B\left(\frac{W}{L}\right)^{\alpha_2^-}-P(W)}{1+k\left(\frac{W}{L}\right)^{\alpha_2^-}}\right]=\frac{B\left(\frac{W^*}{L}\right)^{\alpha_2^-}-P(W^*)}{1+k\left(\frac{W^*}{L}\right)^{\alpha_2^-}}\leq\max_{L\leq W\leq W^*}\left[\frac{B\left(\frac{W}{L}\right)^{\alpha_2^-}-P(W)}{1+k\left(\frac{W}{L}\right)^{\alpha_2^-}}\right] \tag{74}$$

Using (74), it follows that (73) holds. Finally combining (73) and (72) implies that $\widehat{I}^*=\widehat{I}(\sigma^*)$. ■

**Proof of Lemma 4.** Lemma 2 has established that for any level of $B$, the policy $\sigma^*$ of equation (20) is optimal, in the sense that it attains the lower bound $\widehat{I}^*$ (which also depends on $B$). Since the optimal policy $\sigma^*$ assigns the same value $P(L)=B$ every time that $W_t=L$, it suffices to check that is is optimal to set $P_{\tau_0}=B$. To verify this, note that the shareholders' value if they set $P_{\tau_0}=B$, is given by $V=B-k\widehat{I}^*$. Differentiating $V$ with respect to $B$, combining (69) with (66) and using the envelope theorem shows that

$$V_B=1-\left[\frac{k\left(\frac{W^{**}}{L}\right)^{\alpha_2^-}}{1+k\left(\frac{W^{**}}{L}\right)^{\alpha_2^-}}\right]\frac{\alpha_1^+\left[1-\left(\frac{W^{**}}{L}\right)^{\alpha_1^--\alpha_2^-}\right]-\alpha_1^-\left[1-\left(\frac{W^{**}}{L}\right)^{\alpha_1^+-\alpha_2^-}\right]}{\alpha_1^+-\alpha_1^-} \tag{75}$$

The second term on the right hand side of (75) is smaller than 1, because $1-\left(\frac{W^{**}}{L}\right)^{\alpha_1^--\alpha_2^-}<1$ (because $\alpha_1^-<\alpha_2^-<0$) and $1-\left(\frac{W^{**}}{L}\right)^{\alpha_1^+-\alpha_2^-}<0$. Hence $V_B>0$, and therefore it is optimal to set $P(L)=B$. ■

# References


Anderson, R. W., and S. Sundaresan (1996): "Design and Valuation of Debt Contracts," *Review of Financial Studies*, 9(1), 37–68.

Basak, S. (1995): "A General Equilibrium Model of Portfolio Insurance," *Review of Financial Studies*, 8(4), 1059–1090.

Basak, S., and A. Shapiro (2001): "Value-at-Risk-Based Risk Management: Optimal Policies and Asset Prices," *Review of Financial Studies*, 14(2), 371–405.

Caballero, R., and S. Panageas (2005): "A Quantitative Model of Sudden Stops and External Liquidity Management," NBER Working Paper 11293.

Calvo, G. (1999): "Contagion in Emerging Markets: When Wall Street is a Carrier.," Working Paper, University of Maryland.

Caruana, G., and L. Einav (2008): "A Theory of Endogenous Commitment," *Review of Economic Studies*, 75(1), 99 – 116.

Constantinides, G. M., J. B. Donaldson, and R. Mehra (2002): "Junior Must Pay: Pricing the Implicit Put in Privatizing Social Security," NBER Working Paper 8906.

Cox, J. C., J. E. Ingersoll, and S. A. Ross (1985): "A Theory of the Term Structure of Interest Rates," *Econometrica*, 53(2), 385–407.

Cox, J. C., and S. A. Ross (1976): "The Valuation of Options for Alternative Stochastic Processes," *Journal of Financial Economics*, 3(1/2), 145–166.

DeMarzo, P. M., and M. J. Fishman (2007): "Optimal Long-Term Financial Contracting," *Review of Financial Studies*, 20(6), 2079 – 2128.

DeMarzo, P. M., and Y. Sannikov (2006): "Optimal Security Design and Dynamic Capital Structure in a Continuous-Time Agency Model," *Journal of Finance*, 61(6), 2681–2724.

Duffie, D. (2001): *Dynamic asset pricing theory*. Princeton University Press, Princeton and Oxford.

Fleming, W. H., and H. M. Soner (1993): *Controlled Markov processes and viscosity solutions*, vol. 25. Springer-Verlag, New York.

Fudenberg, D., and J. Tirole (1991): *Game theory*. MIT Press, Cambridge MA.

Giannoni, M. P., and M. Woodford (2002): "Optimal Interest Rate Rules I: General Theory," Working Paper, Columbia University.

Gromb, D., and D. Vayanos (2002): "Equilibrium and Welfare in Markets with Financially Constrained Arbitrageurs," *Journal of Financial Economics*, 66, 361–407.

Grossman, S. J., and Z. Zhou (1996): "Equilibrium Analysis of Portfolio Insurance," *Journal of Finance*, 51(4), 1379–1403.

Hart, O., and J. Moore (1998): "Default and Renegotiation: A Dynamic Model of Debt," *Quarterly Journal of Economics*, 113(1), 1–41.

Jensen, M. C., and W. H. Meckling (1976): "Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure," *Journal of Financial Economics*, 3(4), p305 – 360.

Kaminsky, G. L., and C. M. Reinhart (2001): "Bank Lending and Contagion: Evidence From the Asian Crisis," in Takatoshi Ito and Anne Krueger, eds. Regional and Global Capital Flows: Macroeconomic Causes and Consequences (Chicago: University of Chicago Press), 73-99.

Karatzas, I., and S. E. Shreve (1991): *Brownian motion and stochastic calculus*, vol. 113. Springer-Verlag.

Leland, H. E. (1994): "Corporate Debt Value, Bond Covenants, and Optimal Capital Structure," *Journal of Finance*, 49(4), 1213–1252.

——— (1998): "Agency Costs, Risk Management, and Capital Structure," *Journal of Finance*, 53(4), 1213 – 1243.

Leland, H. E., and K. B. Toft (1996): "Optimal Capital Structure, Endogenous Bankruptcy, and the Term Structure of Credit Spreads," *Journal of Finance*, 51(3), 987–1019.

Leonard, D., and N. Van Long (1992): *Optimal Control Theory and Static Optimization in Economics*. Cambridge University Press, Cambridge.

Lucas, D., and R. McDonald (2005): "An Options-Based Approach to Evaluating the Risk of Fannie Mae and Freddie Mac," Working Paper, Northwestern University (Kellogg).

Merton, R. C. (1974): "On the Pricing of Corporate Debt: The Risk Structure of Interest Rates," *Journal of Finance*, 29(2), 449–470.

Miao, J., and N. Wang (2007): "Investment, Consumption and Hedging under Incomplete Markets," *Journal of Financial Economics*, (86), p608 – 642.

Øksendal, B. (2003): *Stochastic differential equations*. Springer-Verlag.

Panageas, S., and M. Westerfield (2005): "High Water Marks, High Risk Appetites? Convex Compensation, Long Horizons and Portfolio Choice," forthcoming, Journal of Finance.

Pavlova, A., and R. Rigobon (2005): "Wealth Transfers, Contagion and Portfolio Constraints," C.E.P.R. Discussion Paper 5117.

Pennacchi, G., and C. Lewis (1994): "The Value of Pension Benefit Guarantee Corporation Insurance," *Journal of Money, Credit and Banking*, 26(3).

Ritchken, P., J. Thompson, R. DeGennaro, and A. Li (1993): "On Flexibility, Capital Structure and Investment Decisions for the Insured Bank," *Journal of Banking and Finance*, 17, 1133–1146.

Ronn, E. I., and A. K. Verma (1986): "Pricing Risk-Adjusted Deposit Insurance: An Option-Based Model," *Journal of Finance*, 41(4), 871–895.

Schneider, M., and A. Tornell (2004): "Balance Sheet Effects, Bailout Guarantees and Financial Crises," *Review of Economic Studies*, 71(3), 883–913.

Sundaresan, S., and N. Wang (2007): "Investment under Uncertainty with Strategic Debt Service," *American Economic Review*, 97(2), 256–261.