Leverage Choice and Credit Spread Dynamics when Managers Risk Shift

Murray Carlson and Ali Lazrak

Sauder School of Business
University of British Columbia*

October 28, 2005

Abstract

We develop a structural model of the leverage choices of risk-averse managers who are compensated with cash and stock. We further characterize credit spread dynamics over the life of the debt. Managers optimally balance the tax benefits of debt with the utility cost that results from their ex-post asset substitution choices. Our model predicts the existence of a U-shaped relationship between the cash component of pay and leverage levels: when cash compensation is low, safe debt with a high face value is issued and when cash compensation is high, risky debt with a high face value is issued. At moderate levels of the cash-to-stock value ratio low leverage is chosen but credit spreads can be significant and again relate to compensation terms. The model illustrates the quantitative importance of including agency costs in the tradeoff theory of capital structure.

*2053 Main Mall, Vancouver, BC, V6T 1Z2; murray.carlson@sauder.ubc.ca; ali.lazrak@sauder.ubc.ca. We thank seminar participants at UBC, the 2005 Northern Finance Association Meetings and Harjoat Bhamra, Adlai Fisher, Ron Giammarino, Rob Heinkel, Marcin Kacperczyk, Hernan Ortiz-Molina for helpful comments.

1 Introduction

Corporate finance theory seeks to describe the behavior of decision makers within a firm and this behavior often interacts directly with the design of corporate securities and their valuation. Structural models incorporate ideas from both the corporate finance and asset pricing literatures to explicitly model these connections, thereby relating primitive assumptions that describe managers and their operating context to security choice and asset price dynamics. This structural approach has had a particularly long history in the analysis of capital structure and corporate debt pricing, dating back to Black and Scholes (1973), Kraus and Litzenberger (1973), and Merton (1974).[1] These models postulate that managers maximize shareholder value, an assumption that has been recently questioned and is difficult to verify empirically (see, for example, the survey by Baker, Ruback and Wurgler (2005)). Characteristics of managerial compensation can be observed, however, and may or may not induce shareholder value maximization (John and John (1993)). Hence, a natural question to address is: How do compensation terms interact with debt choices and their price dynamics? Our paper addresses this issue and formalizes these relationships. In so doing, our theoretical exercise provides empirical predictions that can be translated into new tests of the theories of compensation, debt choice, and debt dynamics.

More specifically, our paper extends structural models of capital structure and credit spreads by considering the additional agency problem that is created when the firm is run by self-interested risk averse managers with realistic compensation contracts. In particular, we assume that managers, rather than shareholders, operate the firm's financial and real assets. We also assume that the manager's compensation includes both cash and stock components, consistent with the empirical finding that company stock holdings are a significant component of CEO pay (see, for example, Hall and

[1] More recent structural models include Fisher, Heinkel and Zechner (1989), Leland (1994), Leland and Toft (1996), Goldstein, Ju and Leland (2001), Collin-Dufresne and Goldstein (2001) and Titman and Tsyplakov (2003). Managers in these models are assumed to maximize value, either of the firm or of the firm's shareholders, and their decisions are typically restricted to choosing the debt amount, when to refinance, and when to default. Notable exceptions are Mauer and Triantis (1994), Leland (1998), and Hennessy and Whited (2005) who allow shareholder value maximizing managers to also make operating decisions. Their models thus incorporate elements of both the tradeoff theory of capital structure (Modigliani and Miller (1963)) and of the agency costs of debt (Jensen and Meckling (1976)) in a dynamic setting.

Liebman (1998)).[2] As a consequence, when debt is present the manager's pay is convex in firm value and he has the incentive to increase firm volatility by engaging in asset substitution. Debt in our setting also has a direct impact on firm value since we assume that it gives rise to tax benefits and default costs. Managers, therefore, face an interesting tradeoff when choosing leverage. Their wealth increases when leverage increases, due to the effect of tax shields on the value of his stock holdings, but firm risk also increases because of ex-post asset substitution. The optimal leverage choice balances the wealth benefit against the utility cost associated with higher risk.

Importantly, we find a non-monotonic relationship relating initial leverage to the relative size of the cash component of compensation. Leverage is highest when the cash/stock value ratio is either very high or very low. Credit spreads of newly issued debt are very different at these extremes, however. Safe debt is issued when the cash component is relatively low, but risky debt is issued when the cash component is relatively high. We also show that when the cash component is at moderate levels managers issue less debt and leverage choice is less sensitive to the contract terms. Within this range, leverage ratios are similar but credit spreads are increasing in the level of cash compensation. Thus, our findings highlight the importance of controlling for the structure of managerial pay when valuing new debt.

Our theory also suggests that compensation terms can provide an economic mechanism through which the quantitative predictions of debt choice and credit spreads from structural models can be enhanced. This finding addresses the empirical weaknesses of existing structural models. First, they predict that firms will choose higher leverage levels than are observed in practice (see, for example, Graham (2000)). Second, calibrated versions of these models predict counterfactually low yields given any measured level of firm leverage, especially for debt that is relatively safe (see Eom, Helwege and Huang (2003)). Our work shows that these empirical shortcomings can be addressed in a structural framework by introducing an asset substitution incentive. We are able to provide parameterizations of our model that generate both low initial leverage and high initial credit spreads by varying the manager's compensation terms.

Our work is related to Leland (1998) who also considers the impact of asset substitution on leverage choice and credit spreads. The key difference be-

[2] For other evidence see Murphy (1999) who provides a recent survey of the compensation literature.

tween our models derives from the assumed objectives of the decision makers. The focus of his model is on the leverage effects of stockholder-bondholder conflicts and he conciously chooses to ignore manager-stockholder conflicts (see his footnote 4). His managers, therefore, maximize shareholder value. We extend his work by structurally modeling the impact of our risk-averse manager's compensation terms. In addition, unlike in his setting where managers are restricted to choose among only two pre-specified levels of volatility, our framework allows for the consideration of general forms of volatility choice. Differences between our models lead to different conclusions regarding the importance of agency costs since we find a large quantitative impact on leverage and credit spreads.

There are several recent papers that model operating decisions made by risk-averse managers. Our paper is complementary to the work on optimal contracting by Cadenillas, Cvitanic and Zapatero (2004) in which a manager, who is compensated only with stock of a levered firm, makes a volatility and effort choice. Shareholders in their model choose debt and stock grant amounts so as to maximize firm value net of the cost of retaining the manager and their optimal contract balances the benefits of reducing moral hazard against the compensation cost. Different economic forces give rise to optimal leverage in our model, as previously described. Furthermore, our managers are paid, in part, with cash. By including this component of pay we generate dramatically different predictions regarding credit spreads since, unlike in our model, the debt issued in Cadenillas, Cvitanic and Zapatero (2004) is always riskless.

Lewellen (2003) undertakes an empirical investigation the relationship between managerial compensation and capital structure. Our formal model builds on her insight that stock compensation gives managers a tendency to avoid debt and our model's predictions are consistent with her findings that managerial preferences help explain firm debt-equity choices. Our study could also explain the empirical findings of Ortiz-Molina (2005) in which low pay-for-performance sensitivity is associated with high leverage. Morellec (2003) models manager-shareholder conflicts in a context where managers derive direct utility from investing in and retaining control of new projects. His model is capable of explaining the low leverage ratios observed in the data, but he does not explore the model's implications for credit spreads. Ju, Parrino, Poteshman and Weisbach (2003) consider the leverage choice of a risk-averse manager but where managers do not make ex-post operating decisions. They undertake a careful calibration exercise and show that the model generates low optimal leverage ratios. Parrino, Poteshman and Weisbach (2005) take the leverage choice as given and examine the decision of a manager to accept or reject a project that may increase firm value and alter the unlevered asset dynamics. They again calibrate the model and find that investment distortions are sensitive to the proportion of the manager's wealth in stock. Neither this paper nor the previous one analyzes credit spread dynamics.[3]

Our paper is also related to Carpenter (2000) who examines the portfolio choices made by fund managers with compensation contracts and objective functions like those we model. Although we share the same basic modeling framework, the focus of the two papers is quite different. First, we consider optimal leverage choice in the presence of taxes and bankruptcy costs. This is outside the context of Carpenter (2000) because by choosing the initial leverage level, our managers endogenously alter the terms of their contract. Second, we analyze the relationships between asset substitution and bond yields, thereby extending her insights to the literature on credit spread modeling.

The remainder of the paper is structured as follows. Section 2 introduces the model. Section 3 develops closed-form expressions for the asset substitution behavior, credit spreads, and security values. Section 4 characterizes the optimal leverage choice and links this choice to the manager's compensation terms. Section 5 concludes.

2 The Model

This section describes our assumptions and formally specifies the manager's optimization problems.

2.1 Market and Firm Value Dynamics

Our model is developed in a partial equilibrium, complete markets setting. We specify a pricing kernel with dynamics

$$\frac{dM_t}{M_t} = -rdt - \alpha dz_t$$

[3]Several other papers examine the quantitative impact of agency costs resulting from stockholder-bondholder conflicts, including Mello and Parsons (1992), Parrino and Weisbach (1999), Moyen (2000) and Hennessey (2003).

where $r > 0$ is the instantaneous risk-free rate, $\alpha > 0$ is the market risk premium, z is a standard Brownian motion and where $M_0 = 1$.

We consider a firm with a finite life T that generates a liquidating cashflow of V_T at the terminal date.[4] This cash payment will be shared among the firm's equity- and debt-holders, and will also serve to pay taxes. A manager has the ability to dynamically control the risk of this payment. We assume that all risk is systematic and that investors can observe the manager's risk choice. As a result, the terminal dividend will evolve according to

$$\frac{dV_t}{V_t} = (r + \alpha\nu_t)dt + \nu_t dz_t \tag{1}$$

where ν_t is the manager's time-t choice of market risk and initial unlevered firm value is normalized to $V_0 = 1$. The process for ν must be based on observable information which, in this setting with only one source of uncertainty, is fully contained in the path of the Brownian motion z. We will refer to this process as the present value of terminal cashflows and note that it also represents the process for unlevered firm value.[5]

2.2 Taxes and Bankruptcy Costs

Taxes are assumed to be paid only at the corporate level and can be reduced if the firm pays interest on outstanding debt.[6] We assume that the firm's managers set debt levels by issuing a zero-coupon claim with face value L and issue price $B_0(L)$ at the initial date. The proceeds from the issue are used exclusively to redeem outstanding shares, so that the only motivation for their issue is to generate a tax shield. We thus model debt choice within the context of a tradeoff theory of capital structure.

Tax shields accrue at maturity but only if the firm is solvent. The firm is considered solvent if bondholders and tax obligations can be paid in full, in which case the tax amount is

$$\mathcal{T} = \tau(V_T - (L - B_0(L))$$

[4] One may also view V_T as the before-tax value of assets-in-place at the terminal date. This view requires that the assets not be levered up at that date, since in that case the terminal value would change. We will not consider the effects of this type of dynamic leverage choice, and resort to our simpler but more static view of V_T as a terminal payment.

[5] This follows because taxes of an unlevered firm are proportional to the terminal dividend.

[6] We ignore the non-debt tax shields associated with the capital required to generate unlevered firm value so as to focus our attention exclusively on the tax effects of leverage.

where $\tau \in [0, 1)$ is the firm's tax rate. The expression $L - B(L)$ represents interest paid to bondholders and reduces taxable income in these states, hence the tax shield amount is $\tau(L - B_0(L))$.[7] In case of insolvency, taxes are simply

$$\mathcal{T} = \tau V_T.$$

Our state-contingent tax payments are consistent with Kim (1978) who specifies that the tax claim is senior to the debt claim and that partial payments of interest do not reduce taxable income when the firm is bankrupt.

Having specified the tax payments, we can define the bankruptcy threshold, V^b by the equality $V^b = L + \tau(V^b - (L - B_0(L)))$ which states that the terminal cashflow is just sufficient to cover payments to bondholders and the tax authority. Rearranging the equation yields

$$V^b = L + \frac{\tau}{1-\tau} B_0(L). \tag{2}$$

Note that when taxes are non-zero the bankruptcy threshold is strictly above the debt face value. Solvent states are defined by $V_T \geq V^b$ and, conversely, insolvent states by $V_T < V^b$.

Insolvency is also assumed to have a direct negative impact on terminal cashflows due to a proportional deadweight bankruptcy cost $\delta_f(1 - \tau)V_T$.

We summarize this subsection by expressing the terminal state-contingent cashflows of the firm net of taxes an bankruptcy costs as follows:

$$C_T(L) = \begin{cases} (1-\tau)V_T + \tau(L - B_0(L)) & \text{if } V_T \geq V^b, \\ (1-\delta_f)(1-\tau)V_T & \text{otherwise.} \end{cases} \tag{3}$$

This quantity is commonly referred to as free cashflow and represents funds available to be paid to any of the firm's claimants.

2.3 Valuation

We seek to derive values at any date prior to T for the firm's terminal cashflows, equity claims and debt claims. These claims are all non-linear functions of the terminal unlevered cashflow V_T and we price them using the pricing kernel M. In this subsection, we take as given an arbitrary unlevered terminal firm value V_T satisfying the condition $V_0 = 1 = E(M_T V_T)$ and proceed with the valuation of the these securities.

[7]To aid exposition of the economic forces at play, we ignore the fact that, in practice, the tax shield would be received as an amortized amount over the life of the bond.

We first value the firm's debt. A circularity arises because the bankruptcy threshold, V^b, depends on the initial debt value, $B_0(L)$, and the initial debt value depends on the bankruptcy threshold. We begin by expressing the payoff to bondholders at maturity assuming V^b is known:

$$B_T(L) = \begin{cases} L & \text{if } V_T \geq V^b, \\ (1-\delta_f)(1-\tau)V_T & \text{otherwise.} \end{cases}$$

This follows because funds sufficient to fully repay bondholders are available only when V_T is above V^b (by the definition of solvency in equation (2)), and bondholders receive all of the terminal cashflows otherwise. The initial bond price $B_0(L) = E\left(M_T B_T(L)\right)$ can then be expressed as

$$B_0(L) = E\left[M_T L \mathbf{1}_{V_T \geq V^b} + (1-\delta_f)(1-\tau)M_T V_T \mathbf{1}_{V_T < V^b}\right]$$

where the notation $\mathbf{1}_{\mathbb{S}}$ is an indicator function for the event $\mathbb{S}$. To make the bankruptcy threshold consistent with the initial bond price it is required that V^b solves the non-linear equation (2):

$$V^b = L + \frac{\tau}{1-\tau} E\left[M_T L \mathbf{1}_{V_T \geq V^b} + (1-\delta_f)(1-\tau)M_T V_T \mathbf{1}_{V_T < V^b}\right]. \qquad (4)$$

This condition uniquely identifies V^b which will, in turn, allow us to easily express equity value.

Equity claimants only receive payment in solvent states. Formally, the equity value S_T at maturity is given by

$$S_T = \begin{cases} C_T - L & \text{if } V_T \geq V^b, \\ 0 & \text{otherwise.} \end{cases}$$

Recall that the solvency states are defined by $C_T \geq L$ so that the equity payoff is equal to that of a call option on terminal free cashflow with strike price equal to the face value of the bond, i.e. $S_T = (C_T - L)^+$.[8]

Alternatively, using equations (2) and (3) one can verify that terminal equity value is equivalently given by

$$S_T = (1-\tau)(V_T - V^b)^+.$$

This equation shows that the equity payoff can be expressed as the payoff on $(1-\tau)$ units of a call option on the unlevered terminal firm cashflow, V_T, with strike price V^b. Initial equity value is given by $S_0 = (1-\tau)E\left[M_T(V_T - V^b)^+\right]$.

[8] We adopt the common notation $X^+ = \max\{0, X\}$.

2.4 Managerial Compensation

Managers are paid at the terminal date with cash and stock. Their terminal wealth is given by:

$$W_T = A + pS_T,$$

where $A > 0$ is the predetermined cash component of the compensation and $p \in [0, 1]$ is the proportion of equity granted to the manager.[9] Consistent with the view that moral hazard issues preclude borrowing against ones future income, we impose the strong restriction that managers cannot undo their compensation by trading bonds or the index. This friction will play a primary role in driving the managers choices and serves to model the agency costs associated with their inability to allocate wealth without restriction.[10]

We further assume that managers are risk averse and derive utility from terminal wealth:

$$U(W_T) = \frac{W_T^{1-\gamma}}{1-\gamma}$$

where $\gamma > 0$ is the coefficient of relative risk aversion.[11]

2.5 The Manager's Choices

We assume that self-interested managers are free to choose leverage and firm risk in order to maximize their derived utility from compensation. It is convenient to consider these choices as occurring in two distinct stages. Figure 1 summarizes the sequence of events.

In the second stage of the problem the manager has pre-committed to a fixed level of debt L, which has been sold for its fair market value $B_0(L)$, and he now controls the standard deviation of unlevered firm value ν. This problem is formally stated as:

$$J(p, L) \equiv \sup_{\nu} E\left[U(A + p(1-\tau)(V_T - V^b)^+\right], \tag{5}$$

[9] We assume that external funds are available for the riskless payment of the cash component of the compensation, A. Alternatively, we could view the firms assets as segregated into two accounts: one that the manager can control, worth V_0, and another that he cannot, that is riskless and worth $e^{-rT}A$.

[10] This condition can be weakened if we work in a more realistic setting where firm risk has an idiosyncratic component that the managers cannot hedge and that only partial borrowing against the terminal payout is permitted. The economic intuition from our model will be hold in such a setting but the added complexity would obscure the results.

[11] Utility is defined by $U(W_T) = \log(W_T)$ if $\gamma = 1$.

where V_T is the terminal value of dividends defined by (1) and where the constant V^b is given by (2). We also impose the restriction $V_T \geq 0$ which is consistent with limited liability for all firm claimants.

In the first stage of the decision problem managers announce the face value of debt (which cannot be later changed), sell it to outside investors, and repurchase outstanding equity with the proceeds. Managers cannot sell their own stock, so an increase in debt levels will be accompanied by an increase in their proportion of the firm's outstanding equity. We assume that investors with rational expectations correctly forecast the second stage volatility choice of managers so that debt and equity are fairly priced. Hence, the value of the the firm after the leverage announcement L is $C_0(L) = E(M_T C_T(L))$ where $C_T(L)$ is defined by equation (3) and V_T in that equation is the correctly anticipated outcome of the manager's second stage optimization.

We assume that the manager is endowed with an initial equity proportion in the unlevered firm, p_0, and now determine his final equity proportion in the levered firm, p. Note that his initial dollar holdings of firm equity can be represented by $p_0 C_0(L)$ and that after the debt issue and equity repurchase his dollar holdings can be expressed as $p(C_0(L) - B_0(L))$. Under our assumptions, these values are equal and serve to define p:

$$p_0 C_0(L) = p(C_0(L) - B_0(L)). \tag{6}$$

The manager's first stage problem can now be defined formally:

$$\sup_{p,L} J(p, L) \tag{7}$$

subject to the non-linear equation (6) relating p and L. Two opposing forces are at play in this problem. Risk averse managers wish to tradeoff risk and return so as to optimize their terminal wealth. They are restricted to do this by controlling only leverage and firm volatility. Increasing leverage provides a tax benefit that directly increase firm value, and because managers hold a proportion p of the firm this increases their wealth. Compensation is convex, however, and will provide an incentive to increase risk in certain states. Managers anticipate these actions and so have an incentive to to choose low leverage in the first stage. These tradeoffs give rise to predictions relating the manager's compensation to leverage choice, risk choice and credit spreads that we explore in the next two sections.

3 Managerial Compensation, Asset Substitution and Valuation

In this section, we take as given a fixed level of debt and analyze the impact of the manager's compensation package on his risk choice in the second stage optimization problem (5). Our problem admits a closed form solution, and we explicitly demonstrate how managers engage in asset substitution, whereby firm variance rises when value falls. This action affects all asset values and we will focus our analysis on the implications for credit spread dynamics. We show that compensation terms can be set so that anticipated asset substitution generates high bond yields even when current levels of leverage and firm asset volatility are low.

3.1 Characterizing Asset Substitution

In our complete market setting, dynamically controlling the volatility process ν is equivalent to choosing the state-contingent terminal cashflow V_T. The manager's dynamic choice problem is then equivalent to the static choice problem[12]

$$\sup_{V_T} E\left[\frac{\left(A + p(1-\tau)(V_T - V^b)^+\right)^{1-\gamma}}{1-\gamma}\right], \tag{8}$$

subject to the budget constraint $E\left(M_T V_T\right) \leq V_0$, the positivity constraint $V_T \geq 0$, and where V^b is defined by (2).

The objective in problem (8) is non-standard due to the kink at $V_T = V^b$ induced by the compensation function. Fortunately, this class of problems has been previously studied by Carpenter (2000) who shows that a solution can be characterized once the objective function is "concavified" (see Aumann and Perles (1968)). The following proposition formulates the optimal second stage choice:

Proposition 1 *Let $\overline{V}$ be the unique solution to the nonlinear equation*

$$\left[\frac{[A + p(1-\tau)(\overline{V} - V^b)]^{1-\gamma}}{1-\gamma} - \frac{A^{1-\gamma}}{1-\gamma}\right]\overline{V}^{\,-1}$$
$$= p(1-\tau)\left[A + p(1-\tau)(\overline{V} - V^b)\right]^{-\gamma}. \tag{9}$$

[12]See Karatzas *et al* (1987) and Cox and Huang (1989).

Then optimal firm value is given by

$$V_T = \left[\overline{V} + \left(\overline{V} - V^b + \frac{A}{p(1-\tau)} \right) \left(\left(\frac{M_T}{\overline{M}} \right)^{-\frac{1}{\gamma}} - 1 \right) \right] \mathbf{1}_{\{M_T \leq \overline{M}\}}, \tag{10}$$

where $\overline{M} \in (0, \infty)$ *is the unique scalar such that* $E(M_T V_T) = V_0$.

This proposition gives the state-contingent optimal choice of terminal dividends V_T as a function of the pricing kernel at that date, M_T (which may be viewed as a summary statistic for the state). To understand the manager's actions begin by substituting $M_T = \overline{M}$ in equation (10) to see that in this state the terminal dividend is $V_T = \overline{V}$. Equation (10) therefore shows that the optimal terminal cashflow will never lie in the interval $(0, \overline{V})$. It is intuitive that the manager does not value dividends below V^b because his compensation is flat in that region. Proposition 1 follows because convexity in the compensation schedule makes the manager risk neutral over terminal payoffs in the range $[0, \overline{V}]$ and although it is not immediately apparent from equation (9), $\overline{V}$ is larger than V^b. An immediate consequence of this local risk-neutrality is that dividends lying within this region are dominated by dividends with the same present value but divided between a zero payoff state and states with payoffs larger than $\overline{V}$.

Figure 2 illustrates the impact of concavification on the manager's optimal choice by plotting the probability density function of the terminal cashflow.[13] Two motives give rise to this density. The first is the manager's inherent attitude toward risk, summarized by γ, which drives the (translated) lognormal behavior in good times. The second is the asset substitution induced by the convexity of the compensation. The manager has an incentive to effectively gamble in bad times to provide the possibility to end his tenure with valuable stock. One important implication of the gambling behavior is that the firm assets may end up worthless ($V_T = 0$) with positive probability.

The manager's choice of terminal cashflow is second-best when compared to his allocation decision when he is not constrained from trading his compensation package. We now consider his first-best choice if given the same

[13]See the appendix for The formula for this pdf is in the appendix. The figure is generated using the parameters $r = 0.05$, $\alpha = 0.33$, $V_0 = 1$, $A = 0.005$, $p = 0.01$, $L = 0.5$, $\gamma = 2$, $T = 5$ and $\tau = 0.3$. Its qualitative features are robust to arbitrary parameterizations with $A > 0$ and $L > 0$.

amount of wealth, but in the form of compensation comprised solely of a proportion $\hat{p} > p$ of the equity of an unlevered firm:

$$\sup_{V_T} U(\hat{p}(1-\tau)V_T) \tag{11}$$

subject to $V_0 = E(M_T V_T)$.[14] We refer to this case as the "Merton Benchmark" because the manager's optimal terminal wealth choice is provided by Merton (1971). The state-contingent dividend can be written as $V_T = (\lambda M_T)^{-1/\gamma}$ where λ solves the problem's constraint.

Figure 3 plots the manager's terminal wealth, along with the Merton Benchmark allocation, as a function of the pricing kernel level at the terminal date. The two wealth schedules cross twice at the points $\overline{M}$ and M^*, consistent with the fact that they have the same present value. We again see the effect of asset substitution activity, where in states with high levels of the pricing kernel the manager has set $V_T = 0$ so that terminal wealth is A. To achieve first-best, he would like to sell claims from states where Arrow-Debreu prices are high (i.e. when $M_T > M^*$). This would require him to sell the fixed component of his pay, however, and this is not permitted. This highlights that the important friction generating asset substitution derives from the assumption that the manager cannot borrow against the cash component of his compensation.

We now explicitly characterize risk taking behavior in the second stage:

Proposition 2 *For any time $t \in [0, T)$, the manager's optimal volatility choice is given by*

$$\begin{aligned} \nu_t &= \left(\overline{V} - V^b + \frac{A}{p(1-\tau)}\right) e^{-(r+\alpha^2/(2\gamma))\gamma^*(T-t)} \left(\frac{M_t}{\overline{M}}\right)^{-1/\gamma} \\ &\quad \times \left[\frac{\alpha}{\gamma}\mathcal{N}(d(\gamma^*, M_t/\overline{M})) + \frac{n(d(\gamma^*, M_t/\overline{M}))}{\sqrt{T-t}}\right] \frac{1}{V_t} \\ &\quad + \left(V^b - \frac{A}{p(1-\tau)}\right) e^{-r(T-t)} \frac{n(d(1, M_t/\overline{M}))}{V_t\sqrt{T-t}}, \end{aligned} \tag{12}$$

where $\gamma^* = 1 - 1/\gamma$,

$$d(x, m) = \left(-\ln m + r(T-t) - \frac{\alpha^2}{2}(1 - 2(1-x))(T-t)\right) / (\alpha\sqrt{T-t}),$$

[14] To hold wealth in the two problems constant, the equity proportion $\hat{p}$ must solve $\hat{p}(1-\tau)V_0 = e^{-rT}A + pE(M_T S_T)$ where S_T is the optimal terminal stock value associated with the second stage optimization with compensation terms (A, p).

$\mathcal{N}(.)$ and $n(.)$ are the standard normal cumulative distribution and density functions, and where the present value at time t of the terminal cashflow is given by

$$V_t = \left(\overline{V} - V^b + \frac{A}{p(1-\tau)}\right) e^{-(r+\alpha^2/(2\gamma))\gamma^*(T-t)} \left(\frac{M_t}{\overline{M}}\right)^{-1/\gamma} \mathcal{N}(d(\gamma^*, M_t/\overline{M})) + \left(V^b - \frac{A}{p(1-\tau)}\right) e^{-r(T-t)} \mathcal{N}(d(1, M_t/\overline{M})).$$

Risk choice is illustrated in Figure 4, where time-t standard deviation of unlevered asset value, ν_t is plotted against the present value of firm cashflows V_t. In the graph, we also plot the standard deviation chosen in the Merton Benchmark case which can be shown to be equal to a constant level $\nu_t^M = \alpha/\gamma$. We see that when V_t is low, firm volatility is higher than the Merton level. If stock is deep out-of-the-money, asset substitution incentives can become very strong and the firm can become extremely risky. As V_t increases firm stock becomes in-the-money, the benefit of deviating from the Merton strategy falls. In the limit when stock is deep in-the-money, the manager volatility choice converges to the Merton volatility. This is intuitive because stock becomes the dominant component of his compensation in these states and the probability of bankruptcy becomes small.

Figure 4 also shows that volatility choice is not necessarily a monotonic function of V_t. In particular, if the cash component is relatively low we observe a U-shaped function whereby volatility can be reduced below even the Merton level.[15] This reflects the fact that his compensation can actually induce conservative behavior when the manager acts to keep his firm's stock in-the-money and secure final pay above A.[16]

3.2 Security Valuation with Asset Substitution

We now provide closed-form expressions for the market price of the firm's bonds and levered assets. We demonstrate the impact of asset substitution on these values and show how it affects credit spreads and their dynamics.

[15] The function can be shown to be non-monotonic if $A < p(1-\tau)V^b$.

[16] For another perspective on this behavior in a portfolio allocation setting, see Carpenter (2000).

3.2.1 Bond Prices, Credit Spreads and their Dynamics

The bond value provides the basis for valuation of all other firm claims and we begin by deriving its price. As we demonstrated in the previous subsection, at the terminal date the firm will either be solvent with $V_T \geq \overline{V} > V^b$, or insolvent with $V_T = 0$. Bondholders will thus receive the face value L when the firm is solvent and zero otherwise. Application of the valuation formula $B_t(L) = E\left(\frac{M_T}{M_t} B_T(L) \mid \mathcal{F}_t\right)$ results in the time-t bond price

$$B_t(L) = Le^{-r(T-t)}\, \mathcal{N}(d(1, M_t/\overline{M})), \tag{13}$$

where the function d is defined in Proposition 2.

The last term in the bond pricing expression is the risk-neutral probability that the firm will be solvent. When discounted, it represents the value of an Arrow-Debreu security paying \$1 in the solvent state. This intuition accounts for the particularly simple form of the pricing relationship.

Bond yields in the model can be immediately deduced from the bond price:

$$y_t \equiv -\frac{\ln\left(B_t(L)/L\right)}{T-t} = r - \frac{1}{T-t}\ln(\mathcal{N}(d(1, M_t/\overline{M}))).$$

The corresponding credit spread is:

$$\rho_t \equiv y_t - r = -\frac{1}{T-t}\ln(\mathcal{N}(d(1, M_t/\overline{M}))). \tag{14}$$

This equation shows that the credit spread is proportional to the logarithm of the risk-neutral probability of solvency since the firm is solvent if $M_T < \overline{M}$ and the term $\mathcal{N}(d(1, M_t/\overline{M}))$ gives the risk-neutral probability of this event contingent on the current level of the pricing kernel, M_t.

We will compare these credit spreads to spreads from the Merton benchmark model. Bond prices in that setting, where volatility is held constant at $\nu_t^M = \alpha/\gamma$, are given by a slightly more complicated formula because of the fact that bondholders receive a non-zero payoff in the insolvent states:

$$\begin{aligned} B_t^M(L) &= Le^{-r(T-t)}\mathcal{N}\left(\gamma d(\tilde{\gamma}_1, V^b/V_t^M)\right) \\ &+ V_t^M(1-\delta_f)(1-\tau)\mathcal{N}\left(-\gamma d(\tilde{\gamma}_2, V^b/V_t^M)\right), \end{aligned} \tag{15}$$

where

$$\tilde{\gamma}_1 = \frac{1}{2}\left(1 + \frac{1}{\gamma^2}\right),\ \tilde{\gamma}_2 = 1 - \tilde{\gamma}_1 \text{ and } V_t^M = V_0 e^{rt}\frac{e^{-\left(\frac{r}{\gamma}+\frac{\alpha^2}{\gamma^2}\right)t}}{M_t^{1/\gamma}}. \tag{16}$$

This is essentially the formula provided in Merton (1974) but with non-zero taxes and bankruptcy costs. As in the Merton model, the bond can be replicated by a long position in a riskless bond with face value L and short positions in options. In our context, one must write two option contracts to complete the hedge. The first contract is $(1-\delta_f)(1-\tau)$ units of puts on the terminal cashflow, V_T, with strike price V^b. The second contract is a binary option that pays the foregone tax shield $\tau(L-B_0(L))$ and the bankruptcy cost $\delta_f(1-\tau)V^b$ when the firm is insolvent. The net payoff from the hedge accounts for the fact that bondholders are residual claimants to non-zero firm value in solvent states, but themselves incur the costs of bankruptcy. They of course anticipate this payoff structure, as reflected in the pricing function (15). This formula can be used to calculate yields and spreads for this benchmark.

Figure 5 presents the relationship between credit spreads and the present value of terminal cashflows for various levels of the cash component of compensation, A. The Merton spreads for low ($\delta_f = 0$) and high ($\delta_f = 0.5$) bankruptcy costs are also depicted. The figure shows that when asset substitution effects are present, the credit spread is uniformly higher than that produced by the Merton model with zero bankruptcy costs. Credit spreads increase as A increases, again consistent with the intuition that increasing insurance against bad outcomes enhances risk taking incentives. These spreads can be substantial even for high levels of V_t where debt to equity ratios are low. Furthermore, our model is capable of generating yields higher than in the Merton benchmark even when unlevered asset volatility is low. Referring to Figure 4, observe that with $A = 0.0005$ the current level of volatility is approximately equal to the Merton level if $V_t = 0.75$, yet our yields at this level of V_t are considerably higher than the Merton benchmark with zero bankruptcy costs (see Figure 5).

It is interesting to note that asset substitution does not necessarily increase yields. Observe that in Figure 5 credit spreads from the Merton model with high bankruptcy costs can exceed credit spreads generated by our model. This occurs because although risk shifting serves to reallocate a portion of the terminal cashflows to zero, it also reallocates cashflows from insolvent states (where bankruptcy costs would be incurred) to the solvent states. Bondholders will approve of this activity when the benefit associated with the higher probability of solvency exceeds the cost of potentially receiving zero payment.

In addition to its effect on credit spreads, asset substitution also has a direct impact on bond return volatility. Using Itô's Lemma, the instantaneous standard deviation η_t of the bond return defined by $dB_t(L)/B_t(L)$ can be shown to be

$$\eta_t = \frac{1}{\sqrt{T-t}} \frac{n(d(1, M_t/\overline{M})}{\mathcal{N}(d(1, M_t/\overline{M}))}. \tag{17}$$

This function is closely related to equation (14) since credit spreads are proportional to the risk-neutral probability of solvency, $\mathcal{N}(d(1, M_t/\overline{M}))$. In fact, the model predicts a relationship between credit spreads and bond return volatility that is independent of the terms of the manager's contract, (A, p), and the firm's leverage, L. To see this, combine equations (14) and (17) to obtain

$$\eta_t = \frac{1}{\sqrt{T-t}e^{-\rho_t(T-t)}} n\left(\mathcal{N}^{-1}(e^{-\rho_t(T-t)})\right), \tag{18}$$

where $\mathcal{N}^{-1}$ is the inverse function of the standard normal cumulative distribution function.

Bond return standard deviation is plotted against credit spreads for bonds with a fixed time to maturity in Figure 6. The figure also plots the relationships predicted by the Merton benchmark model with low and high bankruptcy costs. The figure shows that bond return volatilities are uniformily higher in the presence of asset substitution. High bankruptcy costs are required in the Merton model to generate comparable levels of bond return volatility at low yield levels. The figure also highlights that equation (18) is a strictly increasing and unbounded function of credit spreads. On the other hand, the Merton bond return volatilities are bounded. This can be understood by considering the behavior of the Merton bond prices at extreme levels of V_t. When firm value is very high and credit spreads are near zero, the bonds are essentially riskless and volatility is zero. When firm value is very low and credit spreads are very high, the bonds are like equity claims on firm value net of taxes and bankruptcy costs, $(1-\delta_f)(1-\tau)V_T$. The cashflows that bondholders receive in this case are proportional to unlevered firm value and, as a consequence, have the same volatility. In this particular case, bond return standard deviation asymptotically approaches the limit α/γ which is the Merton standard deviation of the firm's unlevered assets.

Equation (18) also provides advice for hedging credit risk. It shows that the proportional amount of a hedging instrument in the risky bond's replicating portfolio can be determined with knowledge of only the bond's credit spread.

3.2.2 Levered Firm Value

We now provide an expression for the present value of the firm's terminal free cashflows at any date:

Proposition 3 *For any time $t \in [0, T)$, total firm value is given by:*

$$C_t = (1-\tau)V_t + \tau(L - B_0(L))e^{-r(T-t)}\,\mathcal{N}(d(1, M_t/\overline{M})), \tag{19}$$

where the function d is defined in Proposition 2 and $B_0(L)$ is given by equation (13) with $t = 0$.

Firm value is composed of three parts. The quantity $(1-\tau)V_t$ is the unlevered firm value. The next term in equation (19) is the present value of tax shields. The difference between the face value payment and the initial debt proceeds, $L - B_0(L)$, represents the interest expense that is deductible if the firm is solvent. When multiplied by the tax rate, τ, this represents the tax savings. Tax shields are then discounted by multiplying the terminal savings by the value of the binary option paying off in solvent states, $e^{-r(T-t)}\,\mathcal{N}(d(1, M_t/\overline{M}))$.

The final part of firm value is the bankruptcy cost. Bankruptcy costs are assumed to be proportional to the terminal cashflow V_T when the firm is insolvent. The manager optimally avoids having positive value in these states, however, so these costs are never incurred. As a result, there is no term relating bankruptcy cost to firm value in equation (19).[17]

The manager's contract terms (A, p) affect his risk-taking behavior and, as a consequence, firm value. To understand the link consider the effect of an increase in A. The cash component of the compensation represents insurance for bad outcomes which are more tolerable when A is high. Bond prices thus decrease due to the increase in asset substitution. In our model, firm value is affected only through the impact on the tax shield. There are two effects. First, as the bond price falls, the amount of interest expense, $L - B_0(L)$, increases. On the other hand, the probability of receiving the reduction in taxes falls. These two opposing forces give rise to the non-monotonic relationship between A and firm value, C_0, as demonstrated in Figure 7.

[17] This counterfactual implication can be addressed but at the expense of further complicating the analysis in this paper. For example, our model could be augmented to include a managerial bankruptcy penalty of the form $\delta_M(V^b - V_T)^+$. This would introduce a set of insolvent states in which firm value is positive.

A related argument holds for changes in p. An increase in the manager's stock compensation leads to more risk aversion as a result of increased sensitivity to V_T (see Ross (2004)). Bond prices are, therefore, an increasing function of p. We again see that a non-monotonic relationship exists between p and firm value. For low levels of p, debt is risky and the tax shield amount at maturity, $\tau(L-B_0(L))$, is high. It is not likely to be received, however, so the value of the binary option paying off in the solvent state is low. For high levels of p the opposite is true. Intermediate levels of the variable give rise to the highest present value of tax shields and, consequently, to the highest levels of firm value.

4 The Leverage Choice

In this section, we analyze the manager's leverage choice in the first stage optimization problem (7). We begin by undertaking a comparative static analysis of the effect of leverage on firm value in both our model and the Merton benchmark. We then study how our risk-averse manager chooses optimal leverage. This choice problem relates to the manager's contract terms and we provide explicit testable implications linking firm leverage to managerial compensation.

4.1 The Effect of Leverage on Firm Value

In our setting, firm value is sensitive to leverage levels because of its effects on tax shields and bankruptcy costs. We first demonstrate these tradeoffs in the Merton benchmark model. We then contrast these predictions to those of our model to highlight the role of asset substitution undertaken by the manager in the second-stage.

4.1.1 Firm Value and Debt in the Merton Benchmark

Firm value in the Merton benchmark model is given by

$$C_t^M = (1-\tau)V_t^M + \tau(L - B_0^M(L))e^{-r(T-t)}\mathcal{N}\left(\gamma d(\tilde{\gamma}_1, V^b/V_t^M)\right) - \delta_f(1-\tau)V_t^M\mathcal{N}\left(-\gamma d(\tilde{\gamma}_2, V^b/V_t^M)\right), \qquad (20)$$

where $\tilde{\gamma}_1$, $\tilde{\gamma}_2$ and V_t^M were previously defined in equation (16). This expression has three terms: the unlevered firm asset value, the present value of tax shields, and the present value of bankruptcy costs. Debt levels affect firm value through the last two terms.

Figure 8 shows that a firm value maximizing leverage level exists even without bankruptcy costs. The intuition for this result is as follows. The tax shield amount at maturity, $\tau(L - B_0^M(L))$, is increasing in the face value of debt.[18] Its present value is obtained by multiplying by the term $e^{-r(T-t)}\mathcal{N}(\gamma d(\tilde{\gamma}_1))$, which is directly proportional to the risk-neutral probability of solvency. When leverage increases, this probability falls. In the Merton benchmark, neither of these two opposing effects is dominant for all levels of leverage. Thus, when L increases, firm value initially rises above its unlevered value, $(1-\tau)V_t$, and eventually falls back to the same value when L is large.

Firm value is reduced when bankruptcy costs are present. Furthermore, increasing bankruptcy costs reduces the leverage level at which firm value is maximized (see Figure 8).

4.1.2 Firm Value and Debt with Asset Substitution

We now examine the sensitivity of firm value to debt levels when managers can control the firm value process in the second stage. We undertake a simple comparative static analysis whereby all parameters in the model are held constant except the face value of debt, L.

Figure 9 plots firm value from equation (19) as a function of leverage. This figure shows that the manager's second stage risk choices act to make firm value a strictly increasing function of leverage. Unlike in the Merton benchmark, this occurs because an increase in leverage causes the tax shield amount to rise at a faster rate than the risk neutral solvency probability falls.

An important implication of our analysis is that asset substitution will not necessarily reduce firm value. In fact, our model shows that at high leverage levels firm value can be enhanced relative to the Merton benchmark where no risk shifting occurs.

4.2 Optimal Leverage

We now study the properties of the optimal leverage choice. In order to undertake this analysis we utilize standard numerical methods to approximate

[18]This follows from the fact that $B_0^M(L)/L$ represents the value of a claim paying \$1 in solvency and must decrease when bankruptcy becomes more likely.

the solution to (7).[19]

Managers trade off two opposing effects when selecting firm leverage. On one hand, their wealth is increasing in leverage due to the tax effects previously described. Managers are risk averse, however, and they internalize the fact that their leverage decision will pre-commit them to risk shift, thereby penalizing their derived utility.[20] The purpose of our numerical exercise is to quantify which of the two forces is dominant under various compensation terms.

The top panel in Figure 10 depicts the relationship between optimal leverage and the cash component of the manager's compensation, A. This relationship is U-shaped. The cash component of compensation provides the manager with insurance, so when A is high the manager's risk aversion is relatively low. In such cases he perceives the cost of increasing the face value of debt to be low. Optimal leverage is, therefore, increasing in A in this region. The bottom panel in Figure 10 relates credit spreads for newly issued debt to A. Note that this function is upward sloping and that the magnitude of credit spreads are large when A is large. This implies that managers with compensation that is relatively safe are issuing very risky debt.

An interesting implication of Figure 10 is that when the cash component of compensation is low, managers choose to issue high levels of safe debt. This behavior can be understood by considering the special case in which $A = 0$. Managers will avoid running their firm into bankruptcy at all costs, since their marginal utility is infinite at zero wealth. It seems counterintuitive that they would then add leverage. This intuition does not account for the fact that they control volatility and have the ability to ensure firm value achieves a fixed lower bound at the terminal date. In selecting the optimal debt level, managers will thus tradeoff the cost of achieving this lower bound through their asset substitution policy against the tax shields associated with safe debt.

Following this logic, when A is low, the manager's second stage choice of terminal cashflows results in a firm with low risk. In these cases, his derived

[19]Numerical solutions to the first stage problem are straightforward to compute due to existence of the closed form expressions derived in Section 3. In our particular implementation we have used the one-dimensional function optimization routine from the Matlab Optimization Toolbox.

[20]We prove in the appendix that in the absence of taxation managers optimally choose not to issue debt. This is a direct consequence of their risk aversion, since in that case no wealth effects are present.

utility of wealth is relatively high. As a result, his inherent risk aversion does not have as strong a dampening effect on leverage.

An important empirical implication of our model is that low firm leverage can be associated with low or high credit spreads. For example, again referring to Figure 10 we see that if a manager chooses to issue debt with face value around 0.2, the credit spread can be as low as 2% or as high as 15%. This highlights that compensation terms interact with the leverage choice and asset substitution. The riskiness of debt is not only determined by firm characteristics, but also by observable aspects of the manager's pay.

5 Conclusion

In this paper, we have demonstrated the relevance of the agency costs of Jensen and Meckling (1976) for structural models of leverage choice and credit spreads. In particular, we have assumed a realistic compensation structure for risk-averse managers, consisting of cash and stock. We showed that managers will optimally choose to lever the firm and that their resulting wealth will be convex in unlevered asset value. This convexity induces asset substitution, leading to riskier payouts and higher credit spreads than predicted by the prior literature. Finally, we demonstrated that optimal leverage choice is the result of a balance between tax benefits and the utility cost of ex-post risk shifting.

Our model can be extended in many interesting directions. One possibilty is to consider the optimal contract from the shareholders' perspective. This would require that we consider one more prior optimization problem in which the amount of cash compensation and the number of unlevered shares are determined, subject to a participation constraint.

A more ambitious extension would place our model in a dynamic context. The model in this paper could be considered one stage of a multi-period problem in which periodic capital structure and default decisions could be made. Dynamic contracting as in Heinkel and Stoughton (1994) could also be incorporated. These changes, if analytically tractible, would add significant realism and allow for empirical tests of a dynamic capital structure model in which risk shifting plays an important role.

Appendix: Proofs

Proof of Proposition 1. Proposition 1 is a reformulation of Theorem 1 in Carpenter (2000). Here, we provide a new proof that applies techniques from convex analysis and generalizes Carpenter's Theorem to a larger class of compensation schedules. For additional mathematical detail see, for example, Rockafellar (1970) or Chapter 3 of Ekeland and Turnbull (1983).

Let us first, define the function $\varphi : \mathbb{R} \longrightarrow \mathbb{R} \cup \{\infty\}$ by

$$\varphi(x) = \begin{cases} -\frac{\left(A+p(1-\tau)(x-V^b)^+\right)^{1-\gamma}}{1-\gamma} & \text{if } x \geq 0, \\ +\infty & \text{otherwise.} \end{cases}$$

The function $\varphi^* : \mathbb{R} \longrightarrow \mathbb{R}$ denotes the Legendre-Fenchel transform of φ defined by

$$\varphi^*(y) = \sup_{x \in \mathbb{R}} \left(xy - \varphi(x) \right).$$

Similarly, the function $\varphi^{**} : \mathbb{R} \longrightarrow \mathbb{R}$ denotes the double Legendre-Fenchel transform of φ defined as

$$\varphi^{**}(y) = \sup_{x \in \mathbb{R}} \left(xy - \varphi^*(x) \right).$$

We will use two basic results from convex analysis.

*R*1 The function φ^{**} is the convex envelope of φ, that is, the function φ^{**} is the largest convex function dominated by φ;

*R*2 If φ^* admits a supporting line[21] at $x \in \mathbb{R}$ with slope k, then φ^{**} admits a supporting line at k with a slope x.

[21] The function φ^* admits a supporting line at x if there exists $k \in \mathbb{R}$ such that

$$\varphi^*(y) \geq \varphi^*(x) + k(y - x),$$

for all $y \in \mathbb{R}$. We will say then that φ^* admits a supporting line at x with slope k. The geometric meaning of a supporting line is intuitive: the supporting line must be uniformly below the graph of the function φ^*. For instance, a convex and differentiable function admits a supporting line at any point and the supporting line, in this case, is just the familiar tangent line. Supporting lines are also called subdifferentials.

Now, the second stage optimization problem (8) may be equivalently formulated as[22]

$$\mathcal{P}: \quad \begin{array}{c} \inf_{V_T} E\left[\varphi(V_T)\right], \\ \text{subject to } E(M_T V_T) \leq V_0. \end{array}$$

The problem $\mathcal{P}$ is not standard because we minimize a non convex function. The strategy of the proof goes as follow: We first solve a "convexified" version of problem $\mathcal{P}$ defined as

$$\tilde{\mathcal{P}}: \quad \begin{array}{c} \inf_{V_T} E\left[\varphi^{**}(V_T)\right], \\ \text{subject to } E(M_T V_T) \leq V_0, \end{array}$$

and then show that the optimum for $\tilde{\mathcal{P}}$ is actually an optimum for $\mathcal{P}$.

Before proceeding with this, we first compute φ^*:

$$\varphi^*(y) = \begin{cases} +\infty & \text{if } y > 0, \\ y\left(V^b - \frac{A}{p(1-\tau)}\right) - \frac{\gamma}{1-\gamma}\left(\frac{-y}{p(1-\tau)}\right)^{-\frac{1-\gamma}{\gamma}} & \text{if } y \in [\varphi'(\overline{V}), 0], \\ -\varphi(0) = \frac{A^{1-\gamma}}{1-\gamma} & \text{if } y \leq \varphi'(\overline{V}), \end{cases}$$

where the constant $\overline{V}$ is defined in Proposition 1. Therefore, φ^* is differentiable on $(0, \varphi'(\overline{V})) \cup (\varphi'(\overline{V}), +\infty)$ with a derivative

$$\nabla\varphi^*(y) = \begin{cases} \left(V^b - \frac{A}{p(1-\tau)}\right) + \frac{1}{p(1-\tau)}\left(\frac{-y}{p(1-\tau)}\right)^{-\frac{1}{\gamma}} & \text{if } y \in (\varphi'(\overline{V}), 0), \\ 0 & \text{if } y < \varphi'(\overline{V}). \end{cases}$$

Furthermore, at the point $y = \varphi'(\overline{V})$, the function φ^* has many supporting lines with slopes in the interval $[0, \overline{V}]$ since $\overline{V}$ is the right derivative of φ^* at $\varphi'(\overline{V})$. Finally, the double Legendre-Frenchel transform φ^{**} can also be derived and is given by

$$\varphi^{**}(x) = \begin{cases} +\infty & \text{if } x < 0, \\ -\frac{A^{1-\gamma}}{1-\gamma} - p(1-\tau)\left(A + p(1-\tau)(\overline{V} - V^b)\right)^{-\gamma} x & \text{if } x \in [0, \overline{V}] \\ \varphi(x) & \text{otherwise.} \end{cases}$$

Let us now proceed with the actual proof of Proposition 1. First, let us consider problem $\tilde{\mathcal{P}}$. The problem $\tilde{\mathcal{P}}$ is convex since the function φ^{**} is

[22] Note that the positivity constraint is now incorporated in the definition of φ and that the maximization has been replaced with a minimization because φ reverses the sign of the utility function.

convex (this is an implication of the result $R1$). So, by standard result from optimization (Luenberger (1969)) the first order conditions

$$\nabla\varphi^{**}(V_T) = -\lambda M_T, \quad \lambda > 0 \text{ is such that } E(M_T V_T) = V_0,$$

are both necessary and sufficient. We see then that the first order conditions for problem $\tilde{P}$ stipulate that φ^{**} admits a supporting line at V_T with slope $-\lambda M_T$ (almost surely). Given the supporting line duality result $R2$, we see that the the first order conditions for problem $\tilde{P}$ may be restated by saying that φ^* admits a supporting line at $-\lambda M_T$ with slope V_T (almost surely). Given the closed form expression of φ^* we see that the first order conditions for $\tilde{P}$ can be expressed as[23]

$$V_T = \begin{cases} \left(V^b - \frac{A}{p(1-\tau)}\right) + \frac{1}{p(1-\tau)}\left(\frac{-\lambda M_T}{p(1-\tau)}\right)^{-\frac{1}{\gamma}} & \text{if } -\lambda M_T > \varphi'(\overline{V}), \\ 0 & \text{if } -\lambda M_T \leq \varphi'(\overline{V}), \end{cases}$$

with the budget restriction $E(M_T V_T) = V_0$. Expression (10) is then obtained by substituting $\overline{M} = -\frac{\varphi'(\overline{V})}{\lambda}$ in the above expression for V_T.

The last step of the proof is to establish that the optimal V_T for $\tilde{\mathcal{P}}$ is also optimal for $\mathcal{P}$. To see this, observe first that φ and φ^{**} take on the same value on the the interval $\{0\} \cup [\overline{V}, +\infty]$. Therefore, since V_T takes value only on the set $\{0\} \cup [\overline{V}, +\infty]$, we have

$$E\left[\varphi(X) - \varphi(V_T)\right] = E\left[\varphi(X) - \varphi^{**}(V_T)\right] \geq E\left[\varphi^{**}(X) - \varphi^{**}(V_T)\right]$$

for any feasible X satisfying $E(M_T X) = V_0$. Now, from the first order conditions of Problem $\tilde{\mathcal{P}}$, we know that φ^{**} admits a supporting line at V_T with slope $-\lambda M_T$. Consequently, from the definition of supporting lines,

$$E\left[\varphi^{**}(X) - \varphi^{**}(V_T)\right] \geq -E\left[\lambda M_T\left(X - V_T\right)\right] = 0$$

thereby proving the optimality of V_T for Problem $\mathcal{P}$.

[23]When $-\lambda M_T = \varphi'(\overline{V})$ the first order conditions mandate that V_T could take any value in the interval $[0, \overline{V}]$ because φ^* has multiple supporting lines at $\varphi'(\overline{V})$. We selected the value 0 for V_T in the state $-\lambda M_T = \varphi'(\overline{V})$. This choice will have no impact on the final utility since the probability of the event $\{-\lambda M_T = \varphi'(\overline{V})\}$ is just 0 given the lognormal distribution of M_T.

Proof of Proposition 2. The valuation formula

$$V_t = E\left[\frac{M_T}{M_t} V_T \mid \mathcal{F}_t\right]$$

together with the dynamic equation for the state price density and the optimal firm value given in (10) give the expression for time t firm value, V_t, in Proposition 2. To see this, observe that conditional on $\mathcal{F}_t$, $\ln(M_T)$ is normally distributed with mean $\ln(M_t) - (r + \alpha^2/2)(T - t)$ and variance $\alpha^2(T - t)$. Substituting the expression for V_T in (10), and computing the expectation in the relevant regions yields the expression for V_t as a function of M_t. Application of Itô's Lemma to V_t results in equation (12) for time-t volatility.

Proof of Proposition 3. Expression (19) follows by computing the expectation in appropriate regions, using equation (3) and the valuation formula $C_t = E\left[\frac{M_T}{M_t} C_T \mid \mathcal{F}_t\right]$.

References

Aummann, R. and M. Perles (1965), A variational problem arising in economics, *Journal of mathematical analysis and applications* **11**, 488-503.

Baker, Malcolm, Richard Ruback, and Jeffrey Wurgler (2005), Behavioral Corporate Finance: A Survey, in Espen Eckbo (editor) The Handbook of Corporate Finance: Empirical Corporate Finance, edited by , Elsevier/North Holland, New York, forthcoming.

Black, Fisher, and Myron Scholes (1973), The pricing of options and corporate liabilities, *Journal of Political Economy* **81**, 637-654.

Carpenter, Jennifer (2000), Does option compensation increase managerial risk appetite?, *Journal of Finance* **55**, 2311-2331.

Cadenillas, Abel, Jaksa Cvitanic and Fernando Zapatero (2004), Leverage decision and manager compensation with choice of effort and volatility, *Journal of Financial Economics* **73**,71-92.

Collin-Dufresne, Pierre and Robert S. Goldstein (2001), Do credit spreads reflect stationary leverage ratios?, *Journal of Finance* **56**, 1929-1957.

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

Ekeland, Ivar and Thomas Turnbull (1983), Infinite-dimensional optimization and convexity, The University of Chicago Press, Chicago, Ill.

Eom, Young Ho, Jean Helwege, and Jing-Zhi Huang (2004), Structural models of corporate bond pricing: An empirical analysis, *Review of Financial Studies* **17**, 499-544 .

Fischer, Edwin O., Robert Heinkel, and Josef Zechner (1989), Dynamic capital structure choice - theory and tests, *Journal of Finance* **44**, 19-40.

Goldstein, Robert, Nengjiu Ju and Hayne Leland (2001), An EBIT-based model of dynamic capital structure, *Journal of Business* **74**, 483-512.

Graham, John (2000), How big are the tax benefits of debt?, *Journal of Finance* **55**, 1901-1941.

Hall B. J. and J. B. Liebman (1998), Are CEOs really paid like bureaucrats?, *Quarterly Journal of Economics* **113**, 653-691.

Heinkel, Robert and Neal M. Stoughton, The dynamics of portfolio management contracts, *Review of Finanancial Studies* **7** 351-387.

Hennessey, Chris (2004), Tobin's Q, debt overhang, and investment, *Journal of Finance* **59**, 1717-1742.

Hennessy Christopher A. and Toni M. Whited (2005), Debt dynamics, *Journal of Finance* **60**, 1129-1165.

Jarrow, R. and S. Turnbull (1995), Pricing derivatives on financial securities subject to credit risk, *Journal of Finance* **50**, 53-85.

Jensen, Michael C., and William H. Meckling (1976), Theory of the firm: Managerial behavior, agency costs and ownership structure, *Journal of Financial Economics* **3**, 305-360.

John, Teresa A. and Kose John (1993), Top-management compensation and capital structure, *Journal of Finance* **48**, 949-974.

Ju, Nengjiu, Robert Parrino, Allen M. Poteshman and Michael S. Weisbach (2003), Horses and rabbits? Optimal dynamic capital structure from shareholder and manager perspectives, University of Texas-Austin Working Paper.

Karatzas, I., J.P. Lehoczky and S.E. Shreve (1987), Optimal portfolio and consumption decisions for a 'small investor' on a finite horizon, *SIAM Journal of Control and Optimization* **25**, 1557-1586.

Kraus, Alan, and Robert H Litzenberger, A state-preference model of optimal financial leverage, *Journal of Finance* **28**, 911-922.

Leland, Hayne E. (1994), Debt value, bond covenants, and optimal capital structure, *Journal of Finance* **49**, 1213-1252.

Leland, Hayne E. (1998), Agency costs, risk management, and capital structure, *Journal of Finance* **53**, 1213-1243.

Leland, Hayne E. and Klaus Toft (1996), Optimal capital structure, endogenous bankruptcy and the term structure of credit spreads, *Journal of Finance* **51**, 987-1019.

Lewellen, Katharina (2003), Financing decisions when managers are risk averse, MIT Sloan School Working Paper.

Luenberger, David G. (1969), Optimization by vector space methods, John Wiley and Sons, Inc., New York, NY.

Mauer, David C. and Alexander J. Triantis (1994), Interactions of corporate financing and investment decisions: A dynamic framework, *Journal of Finance* **49**, 1253-1277.

Mello, Antonio S. and John E. Parsons (1992), Measuring the agency costs of debt, *Journal of Finance* **47**, 1887-1277.

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

Merton, Robert C. (1974), On the pricing of corporate debt: The risk structure of interest rates, *Journal of Finance* **29**, 449-469.

Morellec, Erwan (2004), Can managerial discretion explain observed leverage ratios?, *Review of Financial Studies* **17**, 257-294.

Moyen, Nathalie (2000), Dynamic investment decisions with a tax benefit and a default cost of debt, University of Colorado Working Paper.

Murphy, Kevin J. (1999), Executive compensation, in Orley Ashenfelter and David Card (eds.), Handbook of Labor Economics, Vol. 3, North Holland, 2485-2563.

Ortiz-Molina, Hernan (2005), Does capital structure matter in setting CEO pay?, University of British Columbia Working Paper.

Parrino, Robert, Allen M. Poteshman and Michael S. Weisbach (2004), Measuring investment distortions when risk-averse managers decide whether to undertake risky projects, University of Texas-Austin Working Paper.

Parrino, Robert and Michael S. Weisbach (1999), Measuring investment distortions arising from stockholder-bondholder conflicts, *Journal of Financial Economics* **53**, 3-42.

Ross, Stephen A. (2004), Compensation, incentives, and the duality of risk aversion and riskiness, *Journal of Finance* **59**, 207-225.

Titman, Sheridan and Sergey Tsyplakov (2003), A dynamic model of optimal capital structure, University of Texas-Austin Working Paper.

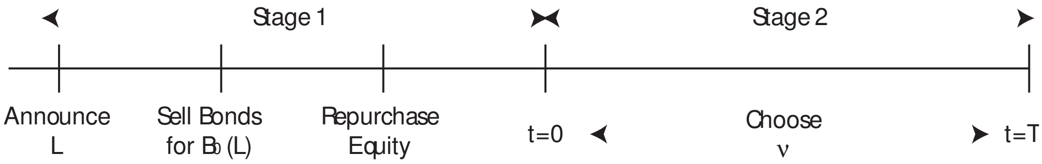

Figure 1: **Timeline describing the manager's decision problems.**

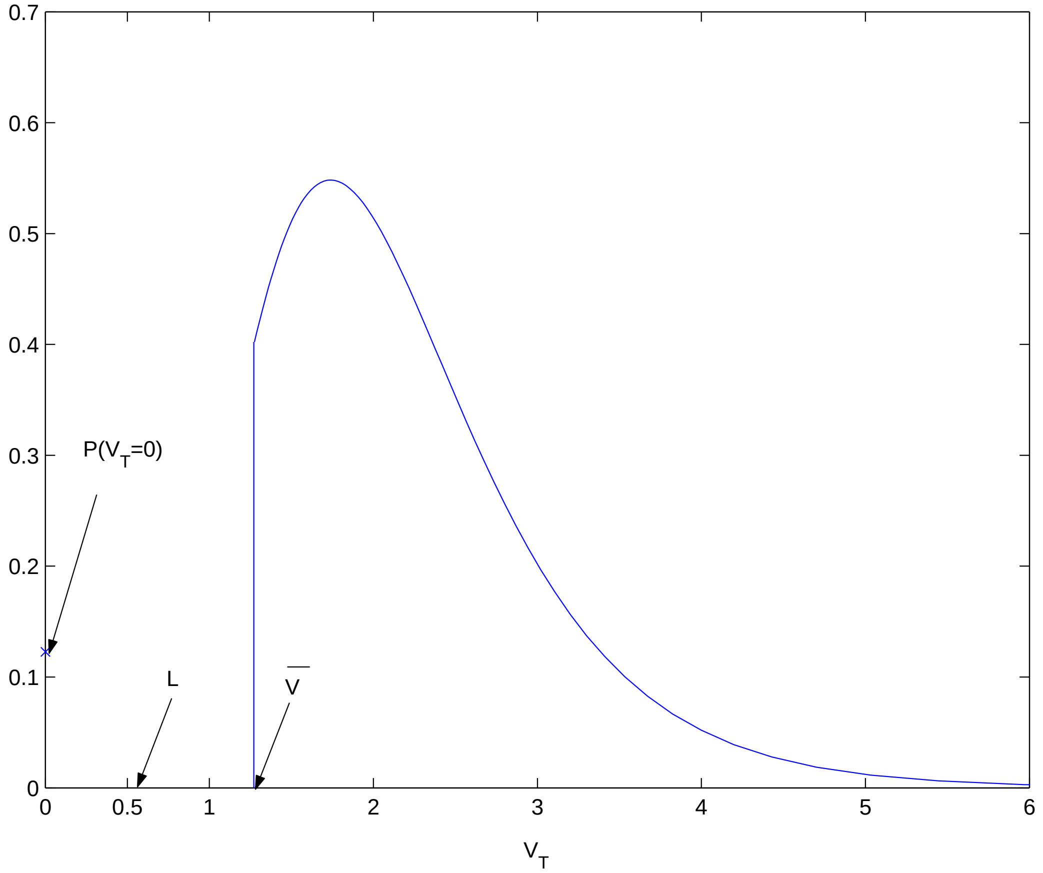

Figure 2: **Density function for V_T using the base case parameterization.**

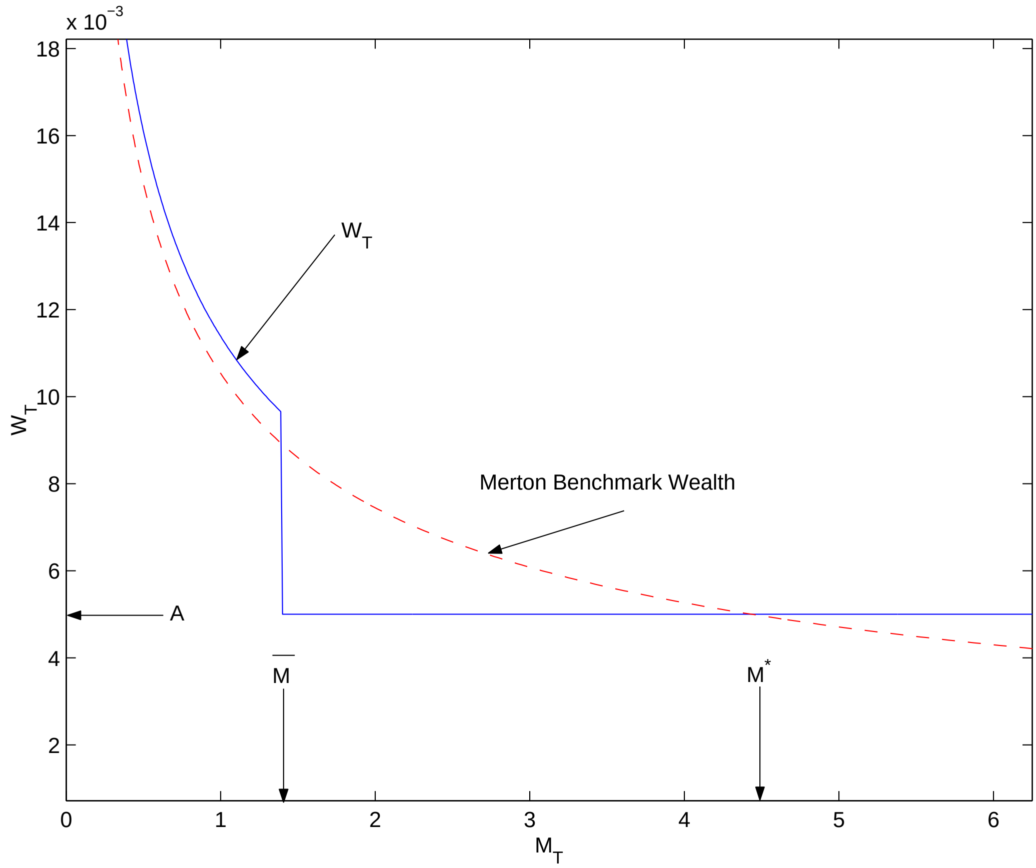

Figure 3: **Optimal choice of V_T as a function of the pricing kernel M_T under the base case parameterization.**

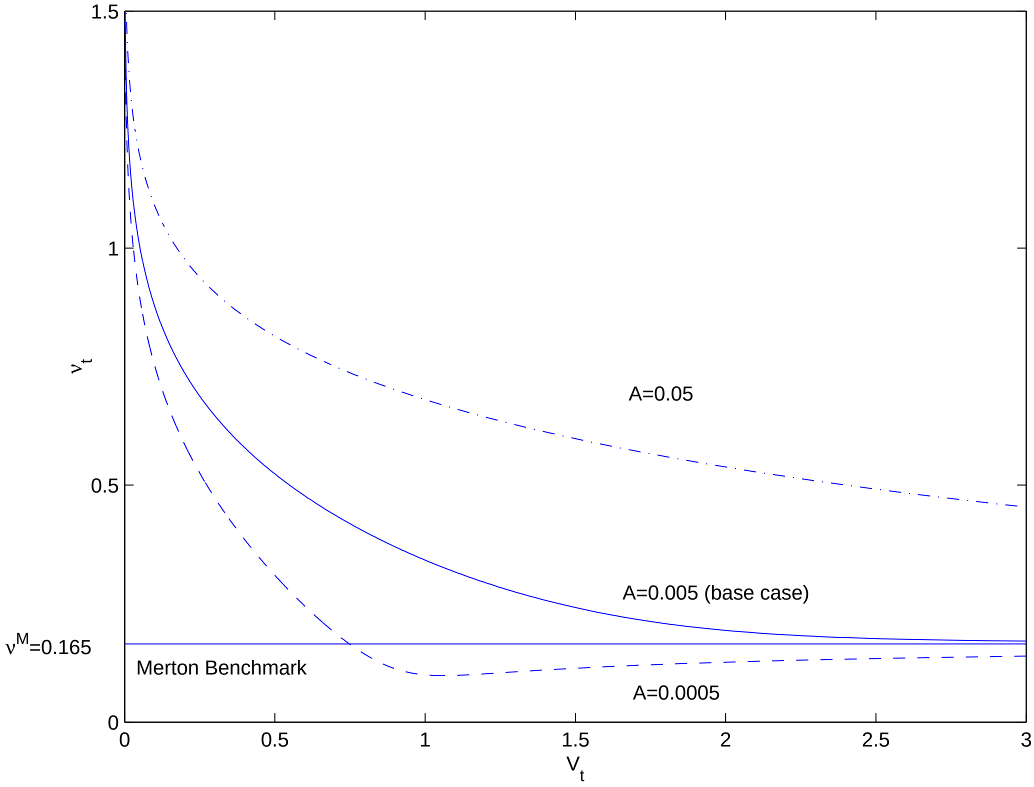

Figure 4: **Volatility choice as a function of the firm asset value in the base case.**

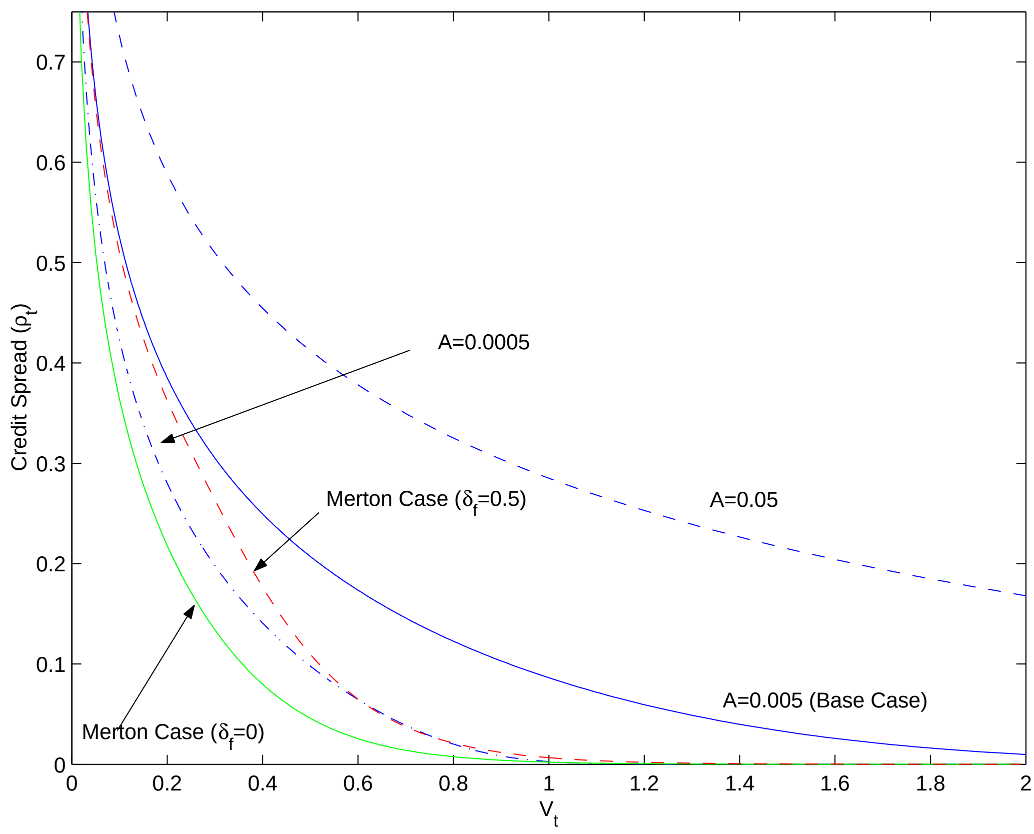

Figure 5: **Credit spreads as a function of the asset value in the base case.**

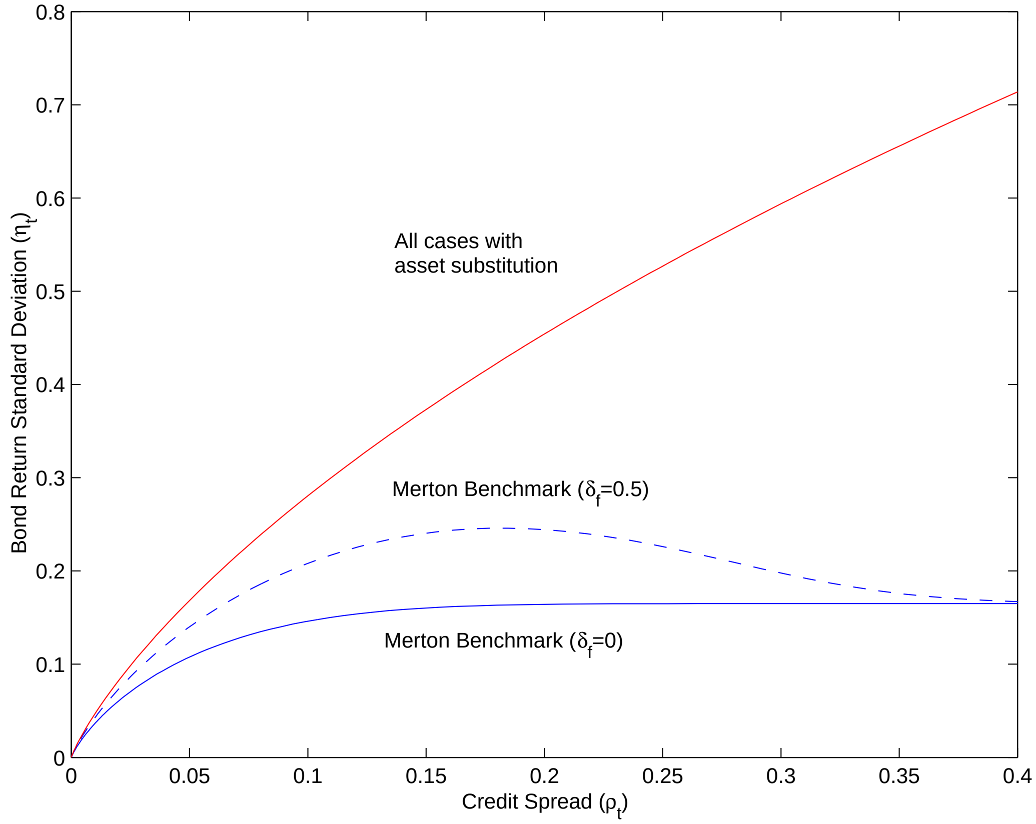

Figure 6: **Bond return volatility as a function of the bond yield in the base case.**

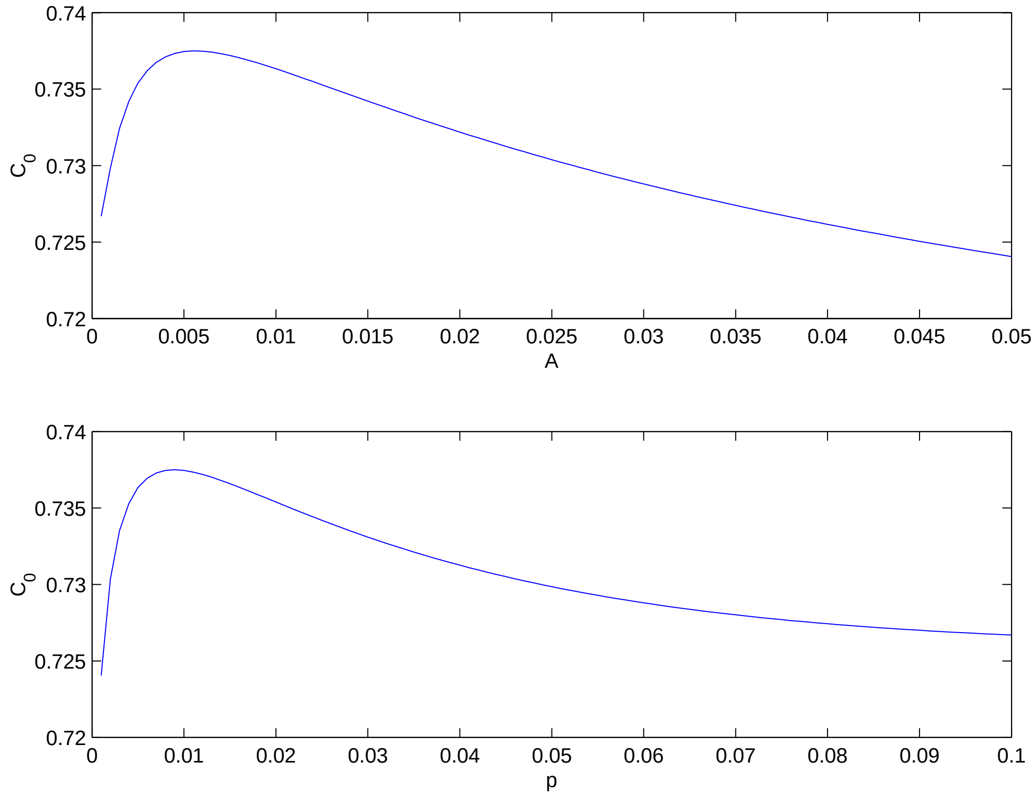

Figure 7: **Firm value as functions of the manager's cash payment A (top panel) and stock holdings p (bottom panel).** All other parameters are as in the base case.

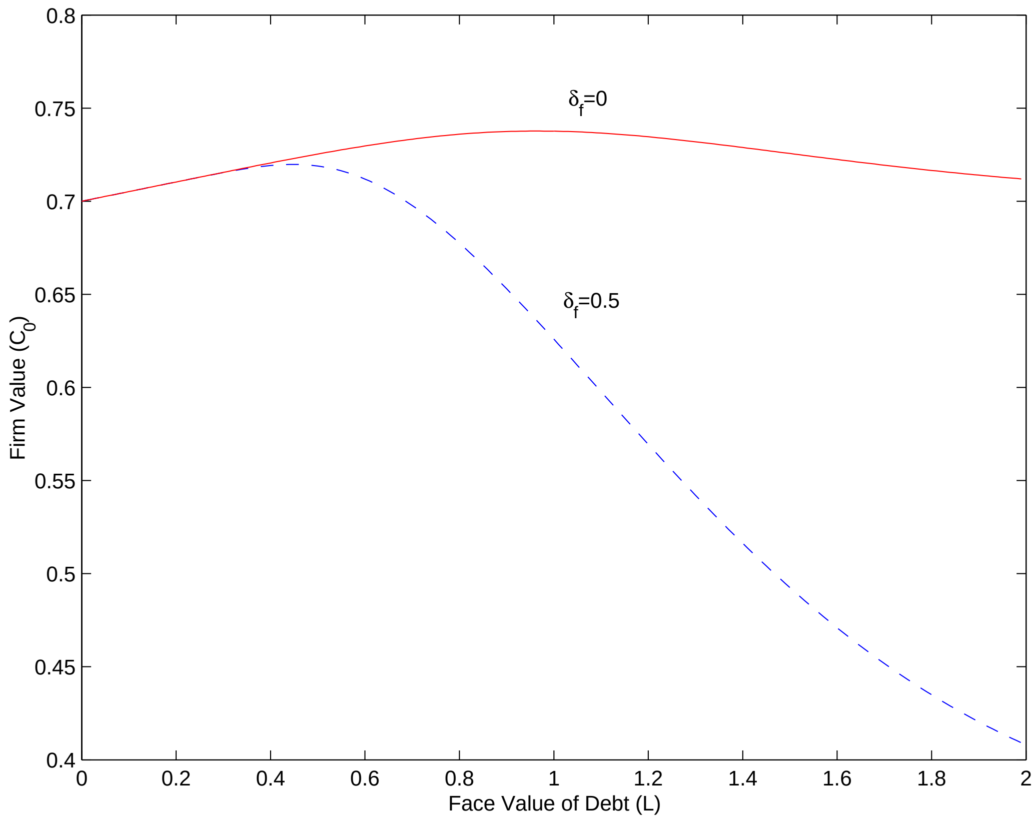

Figure 8: **Firm value as a function of the face value of debt in the Merton benchmark model.**

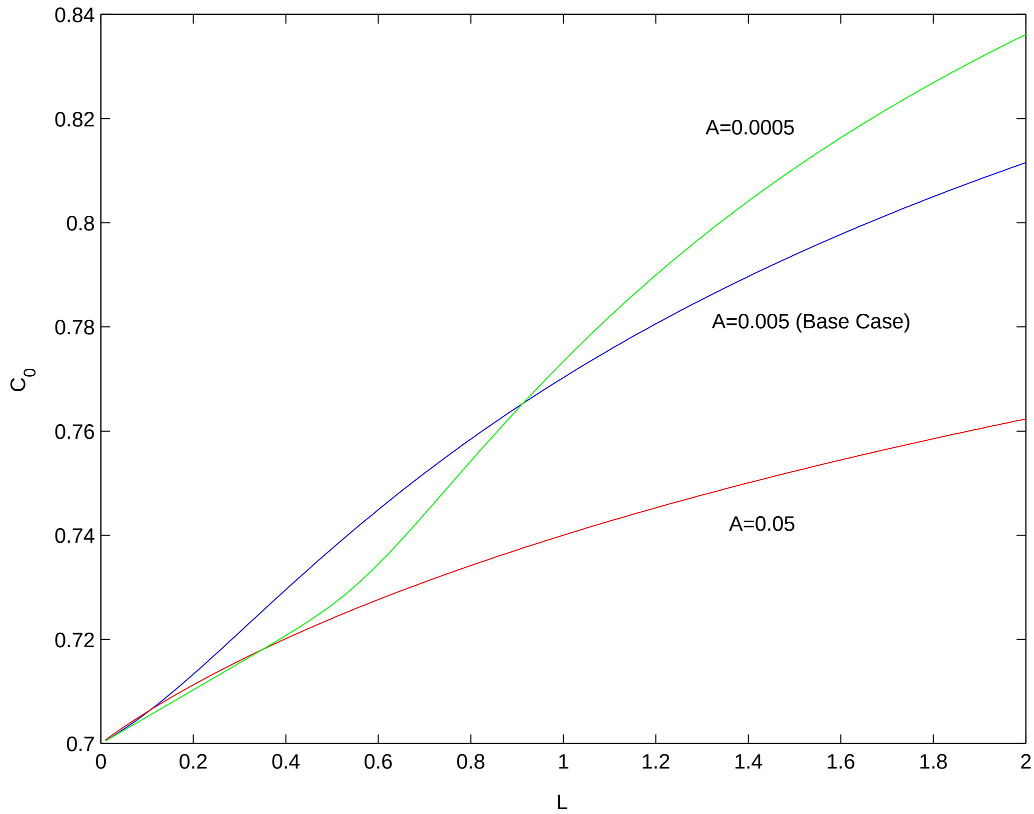

Figure 9: **Firm value as a function of the face value of debt for various levels of the cash component of compensation.**

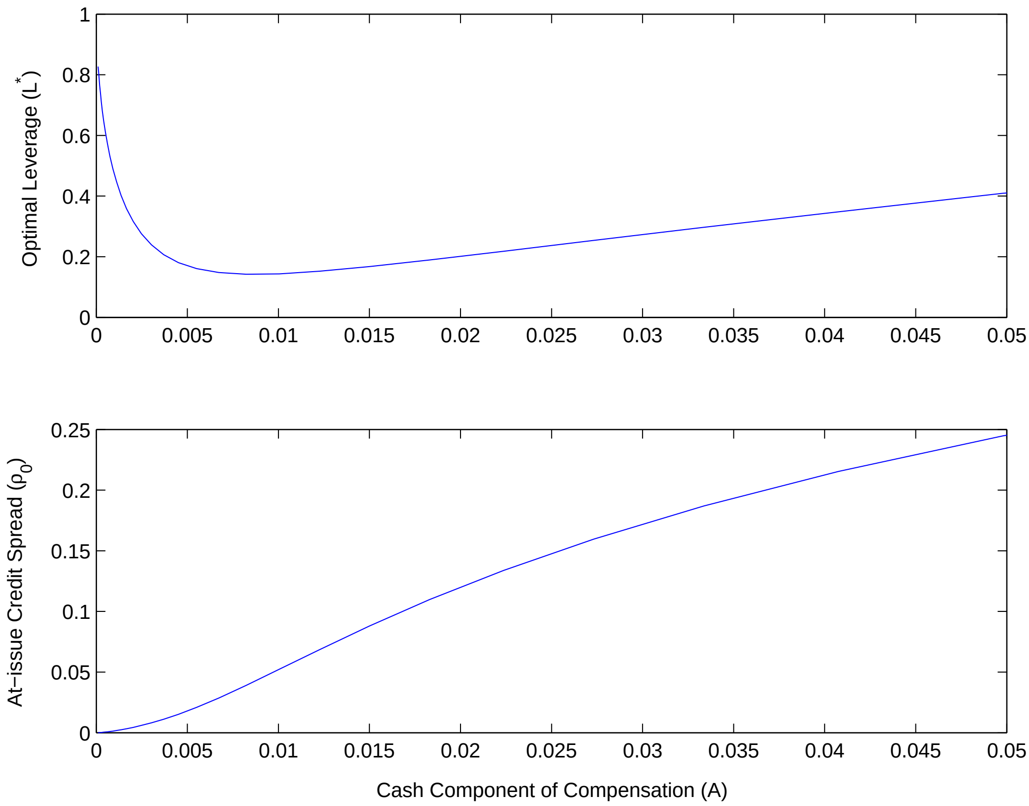

Figure 10: **The optimal choice of leverage (top panel) and the credit spread at issuance (bottom panel) as functions of the cash component of the manager's compensation.**