Nengjiu Ju
Hong Kong University of Science and Technology

Hui Ou-Yang
Duke University

Capital Structure, Debt Maturity, and Stochastic Interest Rates*

This article develops a model in which optimal capital structure and debt maturity are jointly determined in a stochastic interest rate environment. The model yields leverage ratios that are consistent in spirit with empirical observations. The optimal maturity and credit spread of an optimally issued debt are found to be smaller than observed values. The long-run mean of the interest rate is found to be a key variable in determining optimal capital structure and debt maturity. Furthermore, the interest rate volatility and the correlation between the interest rate and the firm's asset value play important roles in determining debt maturity.

I. Introduction

The problem of optimal capital structure has long been an intriguing one among researchers. Brennan and Schwartz (1978) were perhaps the first to study this problem using the contingent-claims analysis approach of Black and Scholes (1973), Merton (1974), and Black and Cox (1976).[1] More recently, Leland (1994) introduced a model of optimal capital structure based on a perpetuity. Leland and Toft (1996) extended Leland (1994) to examine the effect of debt maturity on bond prices, credit spreads, and optimal leverage. The debt maturity in Leland and Toft is assumed rather than optimally determined.[2] Titman and Tsyplakov (2002) and Hennessy and Whited (2004) show that a dynamic trade-off model with features that are not typically included in previous capital structure models can explain many stylized facts.

* Contact the corresponding author, Hui Ou-Yang, at huiou@duke.edu. We are grateful to an anonymous referee for the detailed and thoughtful comments that have improved the article tremendously. We also thank Henry Cao, Hua He, Hayne Leland, Mark Rubinstein, and Matt Spiegel for their advice on an earlier version in which a static model was developed.

1. See also Kim (1978) for a mean-variance analysis of optimal capital structure.

2. See also Kane, Marcus, and McDonald 1985; Fischer, Heinkel, and Zechner 1989; Mello and Parsons 1992; Leland 1998; Fan and Sundaresan 2000; Goldstein, Ju, and Leland 2001; Morellec 2004; and Miao 2005.

(*Journal of Business*, 2006, vol. 79, no. 5)

0021-9398/2006/7905-0007$10.00

The traditional capital structure models, as represented by Brennan and Schwartz (1978), Leland (1994), and others, assume constant risk-free interest rates. Leland (1994) and Goldstein et al. (2001) show that the optimal capital structure is very sensitive to the changes in the level of the interest rate. In other words, the level of the interest rate is a key input in those models. In the absence of explicit modeling of the interest rate process, however, the traditional capital structure models do not provide any guidance on which interest rate, such as the spot rate, the yield-to-maturity (YTM) of a risk-free bond, or other interest-related variables, should be used in the determination of the optimal capital structure. Because their results depend critically on the interest rate, without knowing which interest rate to use, the traditional capital structure models cannot be employed directly to explain the empirical observations. Another restrictive assumption that limits the applicability of the traditional models is that debt maturity is exogenously specified. Because the leverage ratio and the credit spread depend crucially on debt maturity, it is of great importance to determine jointly the optimal capital structure and the optimal debt maturity. In the absence of an optimal maturity structure, the optimal capital structure, as obtained in the traditional models, is of limited empirical relevance.

In this article, we develop a dynamic model in which an optimal capital structure and an optimal debt maturity structure are jointly determined in a stochastic interest environment.[3] At time zero, the firm issues a T-year coupon bond. If the firm has not gone bankrupt in T years, the firm issues a new T-year coupon bond at time T. The optimal value of the new bond will depend on the firm value at time T. If at the end of the second T-year period the firm is still solvent, it issues another T-year coupon bond optimally. This process goes on indefinitely as long as the firm is solvent. The optimally levered firm value at time zero takes into account the tax benefits associated with all future leverages. To obtain the optimally levered firm value, we need to compute explicitly the total tax benefit, the total bankruptcy cost, and the total transactions cost of all future issues of debt. Generally, this is a difficult problem because the issuance of future debt depends on the fact that the firm has not gone bankrupt. We employ a scaling property and a fixed-point argument to obtain the present values of the total tax benefit, the total bankruptcy cost, and the total transactions cost explicitly. Valuation formulas are derived in closed form, and numerical solutions are used to obtain comparative statics.

In our model, the trade-off between tax shields and bankruptcy costs associated with debts yields an optimal capital structure. Firms may issue new debt as firm values change over time. The trade-off between the gains from dynamically adjusting the debt level and the transactions costs of doing so yields an optimal debt maturity structure. The optimal capital structure and

3. In the absence of optimal capital structure and optimal debt maturity, an active and growing body of work has studied the valuation of risky corporate bonds and other derivative instruments in a stochastic interest rate environment (e.g., Kim, Ramaswamy, and Sundaresan 1993; Jarrow and Turnbull 1995; Longstaff and Schwartz 1995; and Duffie and Singleton 1999).

the optimal maturity structure are interdependent. The interest rate is modeled as a Vasicek (1977) mean-reverting process. Modeling the interest rate as a mean-reverting process allows us to examine separately the impact of the long-run mean and the initial value of the interest rate.

When the interest rate is assumed to be a constant, the level of the constant interest rate plays an important role in determining the optimal leverage ratio and debt maturity. In our model with stochastic interest rates, however, the optimal leverage ratio and debt maturity are mainly determined by the long-run mean of the interest rate process. If the initial interest rate is below/above the long-run mean (an upward- or downward-sloping term structure), the firm optimally adjusts downward/upward the coupon and principal of the debt so that its market value (the present value of the coupon and principal) is independent of the initial interest rate level and is proportional to the prevailing asset value when the debt is issued. Consequently, the optimal leverage ratio and debt maturity are independent of the spot interest rate level. These results are quite different from those in a model with constant interest rates, in which the optimal leverage ratio, debt maturity, coupon rate, and principal all depend on the interest rate level.

We find that the long-run mean of the stochastic interest rate process is the most important variable in the determination of the optimal capital structure. This is intuitive because the long-run mean plays a key role in the determination of the tax shields and bankruptcy costs associated with all future debt issues. The reason is that while at any one point in time there is only one bond outstanding in our model, the capital structure and debt maturity are determined by taking into account all future debt issues, and the best information for the future interest rate is its long-run mean. Unlike the spot interest rate and a YTM, the long-run mean of a short-term interest rate process is not an empirical observable; one must estimate it using historical data as well as a specific model for the interest rate process. Therefore, it is essential to consider a stochastic interest rate process in the determination of a firm's capital structure.

The optimal leverage ratio of about 35% obtained from our dynamic model compares favorably with the historical average for a typical large publicly traded firm in the United States. By contrast, most static models typically predict very high leverage ratios.[4]

We find that the tax rate and the transactions cost are the two most important parameters in determining the optimal maturity because the optimal maturity is the result of the trade-off between the transactions cost of and the gain from adjusting the debt level in the future. On the one hand, the firm should issue short-maturity debt and therefore give itself the opportunity to issue new debt optimally, depending on the firm value when the old debt matures. On the other hand, if new debt is issued too often, transactions costs will become

4. Note that dynamic models with constant interest rates can also bring the predicted leverage ratios more in line with observed values (see, e.g., Goldstein et al. 2001).

too large. When the firm behaves optimally, in addition to an optimal capital structure, an optimal maturity structure emerges. Consistent with Childs, Mauer, and Ott (2003), a higher transaction cost yields a longer debt maturity because it is more costly to rebalance a firm's capital structure, and thus the firm does so less frequently. By contrast, it is more valuable to recapitalize the firm's capital structure if the corporate tax rate is higher. Therefore, a higher tax rate yields a shorter (optimal) debt maturity. The relationship among the tax rate, transaction cost, and optimal debt maturity is consistent with the predictions in Childs et al. (2003). However, the reasons behind the results of transactions cost and tax rate are not the same. Our explanation is the trade-off between tax benefits and the transactions costs of dynamically adjusting a firm's capital structure, whereas that in Childs et al. (2003) is based on minimizing the agency cost of the second-best exercising policy of the growth options.[5]

Another interesting result is that the optimal maturity is inversely related to the asset volatility level. This result is easy to understand in our framework because the flexibility to rebalance capital structure is an option. When the firm's asset return process is more volatile, the firm would like to issue debt more frequently to capture the tax benefit or reduce the bankruptcy cost. Due to a high volatility, the firm's asset value may change in greater magnitude. When its asset value goes up, the firm wants to issue more debt to capture the tax benefit. When its asset value goes down, the firm wants to issue less debt to reduce the bankruptcy cost. Consequently, the firm may want to issue debt with shorter maturity so that it can issue debt more frequently in response to the changes in its asset values. Similarly, we find that in addition to the long-run mean of the interest rate process, the volatility of the interest rate process and the correlation between the interest rate process and the firm's asset value process play nontrivial roles in the determination of the debt maturity structure.

The joint determination of the optimal capital structure and the optimal debt maturity leads to results that would not have been obtained in a model in which only the optimal capital structure is determined. For example, when the volatility of the asset return process or the interest rate process increases, one may expect that the credit spread of the firm's debt increases as well. When the firm can choose its debt maturity, however, it optimally issues debt with shorter maturity so that the credit spread of the firm's debt may not increase. The reason is that a debt with a shorter maturity has a lower default risk.

For further comparative statics, we calculate the optimal capital structure while fixing the maturity structure as well as the optimal maturity structure while fixing the capital structure. Some interesting results arise. For example, we find that the firm's optimal leverage ratio is not a monotone function of the debt maturity. Intuitively, one may believe that when the firm is constrained

5. There are no agency conflicts in our model.

to issue long-term debt, the firm may want to issue less debt because of a potentially higher probability of default risk. By contrast, however, if a firm is constrained to issue long-term debt, it means that the firm cannot issue debt as often as it desires. If a firm cannot issue debt more frequently, then it may want to issue a higher amount per debt to capture the tax benefit. The trade-off between the firm's desire to capture the tax benefit and its need to reduce bankruptcy costs leads to a nonmonotone relation between leverage ratio and debt maturity.

The remainder of this article is organized as follows. Section II describes the model. It summarizes the assumptions and introduces the variables and parameters to be used in the rest of the article. Section III derives various valuation expressions in closed form. Section IV presents numerical results. Section V summarizes and concludes. More technical materials are provided in the three appendices. Appendix A reviews the forward risk-neutral measure used to derive the valuation formulas in Section III. Complex derivations are deferred to appendix B. A formal proof of the optimality of our solution in Section III.C is provided in appendix C.

II. Model Description

ASSUMPTION 1. Financial markets are dynamically complete, and trading takes place continuously. Therefore, there exists an equivalent martingale measure (Harrison and Kreps 1979) or a risk-neutral measure (Cox and Ross 1976), Q, under which discounted price processes are martingales.

ASSUMPTION 2. The total before-tax value of the firm's unlevered assets under Q is described by a geometric Brownian motion process,

$$\frac{dV_t}{V_t} = (r_t - y)dt + \sigma_v dW_{vt}^Q, \tag{1}$$

where r_t is the interest rate at time t, y is a constant payout rate, σ_v denotes the constant volatility of asset returns, and W_{vt}^Q is a standard Wiener process defined on a complete probability space $(\Omega, \mathcal{P}, \mathcal{F})$.

This assumption implies that the after-tax value of the firm's unlevered assets, V_t^*, is simply $V_t^* = V_t(1 - \theta)$, where θ is the corporate tax rate. If a firm has debts in its capital structure, its after-tax levered value will be different from V_t^*.

ASSUMPTION 3. The interest rate under Q is modeled by a Vasicek (1977) process,

$$dr_t = \beta(\alpha - r_t)dt + \sigma_r dW_{rt}^Q. \tag{2}$$

The coefficients α, β, and σ_r are constants, and W_{rt}^Q is another standard Wiener process on the same probability space $(\Omega, \mathcal{P}, \mathcal{F})$. The instantaneous correlation between dW_{vt}^Q and dW_{rt}^Q is given by ρdt, with ρ being a constant.

As in the Vasicek model, the price of a zero-coupon bond at time t with a maturity of T is given by

$$\Lambda(r_t, t; T) = e^{A(t;T) - B(t;T)r_t}, \tag{3}$$

where

$$A(t;T) = \left(\frac{\sigma_r^2}{2\beta^2} - \alpha\right)(T-t) - \left(\frac{\sigma_r^2}{\beta^2} - \alpha\right)\frac{1 - e^{-\beta(T-t)}}{\beta} + \left(\frac{\sigma_r^2}{2\beta^2}\right)\frac{1 - e^{-2\beta(T-t)}}{2\beta}, \tag{4}$$

$$B(t;T) = \frac{1 - e^{-\beta(T-t)}}{\beta}. \tag{5}$$

Assumption 4. The firm issues coupon debt with a finite maturity date. Coupon payments are tax deductible at the corporate tax rate θ, and the full loss offset is assumed. Let (C, P, T) be the coupon, principal, and maturity of the debt, respectively. Assumption 2 implies that when the firm pays coupon C, it deducts taxes by the amount of θC.

Assumption 5. Bankruptcy occurs when the firm's unlevered asset value falls to an exogenous default boundary $V_B(r_t, t; P, T)$, which is specified by

$$V_B(r_t, t; P, T) = P\Lambda(r_t, t; T)e^{y(T-t)}/(1-\theta). \tag{6}$$

If $V_t > V_B(r_t, t; P, T)$, the firm is solvent and makes the contractual coupon payment to its bondholders. In the event of default, $\phi \in (0,1)$ fraction of the firm's unlevered value is lost to the bankruptcy procedure, and the bondholders receive the after-tax value of the rest, $(1-\theta)(1-\phi)V_B(r_t, t; P, T)$.

This specification of the default boundary is for tractability. Note that $V_B(r_t, t; P, T)$ has the desired property that, at the maturity date, to avoid default the after-tax unlevered asset value has to be at least as large as the face value of the debt. For example, if the firm decides to retire the debt at T, $V_B(r_T, T; P, T)$ ensures that the firm has just enough after-tax assets to pay off the face value of the debt.[6] Furthermore, this assumption suggests that, given the drift of the firm's unlevered asset value process, the firm defaults (before maturity) at the first time its after-tax value declines below the present value of the principal, P, adjusted for the convenience yield, y, of the value process. In other words, the firm defaults as soon as its value declines to the point at which the expected after-tax firm value at maturity will be below the principal, P.

Assumption 6. The firm rebalances its capital structure every T years. At time zero, the firm issues a T-year coupon bond. If the firm has not gone

6. We have also considered the case where the factor $1/(1-\theta)$ is omitted in eq. (6). Although the results are similar, we feel that eq. (6) is a more reasonable specification because the asset process in eq. (1) is for before-tax value.

bankrupt in T years, it issues another T-year coupon bond at time T. This process continues indefinitely as long as the firm is solvent. Each bond is issued at (the same) $\lambda = 1/\Lambda(\alpha, 0; T)$ times of the value of a risk-free zero-coupon bond with the same face value.

We have two comments concerning assumption 6. First, the capital structures in Leland (1994) and Leland and Toft (1996) are static in the sense that as the firm's asset value evolves, the coupon and the face value of the debt remain the same. This is not optimal because the firm may want to change the debt structure when its asset value changes. When the firm rebalances its capital structure, assumption 6 allows the firm to scale up (down) the coupon and the face value if the firm's asset value has increased (decreased). Given this assumption, we show in Section III.C that the optimal initial debt value at each restructuring point is proportional to the prevailing asset value. For example, if the firm's asset value has doubled since the previous bond issue, the initial market value of the new bond should be twice that of the previous bond.

Second, the assumption that debt is issued at the same constant λ times of the price of a risk-free, zero-coupon bond with the same face value is a critical one for tractability. However, how the constant λ is chosen is not important for tractability, and different reasonable choices are possible. For example, if we require that the initial debt be priced at par, then we need to set $\lambda = 1/\Lambda(r_0, 0; T)$.[7] By contrast, if we set $\lambda = 1/\Lambda(\alpha, 0; T)$, where α is the long-run mean of the mean-reverting interest rate process, then the debt will be priced at par if the prevailing spot interest rate level is the long-run mean, α. With this choice of λ, half of the future debts will be priced above par (with the other half below par) over a long horizon. This latter choice of λ is state independent while the first one depends on the initial interest rate, r_0. We will choose $\lambda = 1/\Lambda(\alpha, 0; T)$ in our calculations.[8] The coupon rate is appropriately chosen to satisfy $D = \lambda P \Lambda(r_0, 0; T)$.

Assumption 7. The transactions cost of issuing and serving a bond is $\kappa \in (0, 1)$ fraction of the market value of the bond issued.

To examine the optimal maturity in a dynamic capital structure model, we need to introduce a transactions cost associated with issuing and serving the debt. The reason is that without it, the firm will rebalance its capital structure continuously.[9] The proportional cost structure is for tractability.

7. Since the debt is issued at λ times of the value of a riskless bond with the same face value, the debt value can be written as $D = \lambda P \Lambda(r_0, 0; T)$. If $\lambda = 1/\Lambda(r_0, 0; T)$, $D = P$ (initial debt priced at par). If $\lambda = 1/\Lambda(\alpha, 0; T)$, then the debt value will be issued at $D = P = \lambda P \Lambda(r_0, 0; T) = P\Lambda(r_0, 0; T)/\Lambda(\alpha, 0; T)$, which equals P (priced at par) when the initial spot rate $r_0 = \alpha$, the long-run mean.

8. While most debts are issued at par in practice, we do need to assume the same λ for all debts for tractability. Our choice of λ should minimize any effects of deviating from par pricing since half/half of the future debts over the long run will be issued above/below par.

9. Our assumption on the transactions cost is similar to those in Kane et al. (1985), Fischer et al. (1989), Goldstein et al. (2001), and Childs et al. (2003).

III. Model Derivation

The stochastic nature of the interest rate complicates the pricing of securities in this model tremendously. Nevertheless, the change of measure technique can be used to obtain valuation formulas in closed form.

A. *Preliminary Development*

Before we present the valuation formulas in the next two subsections, we provide some preliminary results. First, define the first passage time τ as

$$\tau = \min\left[t : V_t \leq V_B(r_t, t; P, T)\right], \tag{7}$$

which is the first time at which the asset value V_t hits $V_B(r_t, t; P, T)$ in some state $\omega \in \Omega$ under Q. Next, define

$$X_t = \log\left[V_t / V_B(r_t, t; P, T)\right]. \tag{8}$$

It is clear that equivalently, τ is the first passage time that X_t reaches the origin from above, starting at $X_0 = \log\left[V_0 / V_B(r_0, 0; P, T)\right] > 0$. Ito's lemma yields

$$dX_t = \left[\sigma_p^2(t; T)/2 - \sigma_v^2/2\right]dt + \sigma_v dW_{vt}^Q + \sigma_p(t; T)dW_{rt}^Q, \tag{9}$$

where

$$\sigma_p(t; T) = \sigma_r B(t; T). \tag{10}$$

Note that although the drifts of dV_t/V_t and $dV_B(r_t, t; P, T)/V_B(r_t, t; P, T)$ are stochastic, that of dX_t is time dependent and deterministic. We emphasize that the nonstochastic (time-dependent only) nature of the drift and diffusion of dX_t is one key feature that allows us to derive valuation formulas in closed form.[10]

In the next two subsections, we need the following two quantities:

$$G(t; T, X_0) = E_0^Q\left[\frac{e^{-\int_0^T r(u)du}}{\Lambda(r_0, 0; T)} 1(\tau < t)\right], \tag{11}$$

$$H(t; T, X_0) = E_0^Q\left[e^{-\sigma_v^2 t/2 + \sigma_v W_{vt}^Q} 1(\tau > t)\right], \tag{12}$$

where $1(\cdot)$ is the indicator function. Note that both $H(\cdot)$ and $G(\cdot)$ depend on T because $\sigma_p(t; T)$ in equation (10) does. The derivations that lead to the closed-form expressions for $H(\cdot)$ and $G(\cdot)$ are given in appendix B.

10. Another key aspect of closed-form derivations is the use of the change of measure technique introduced in app. A.

B. Valuation Formulas in Closed Form

Consider a bond that pays a coupon rate C, has a face value P, and matures at time T. The payment rate $d(s)$ to the bondholders at any time s ($s \leq T$) is equal to

$$\begin{aligned} d(s) = {} & C1(s \leq \tau) + P\delta(s - T)1(T \leq \tau) \\ & + (1-\theta)(1-\phi)V_B(r_s, s; P, T)1(\tau < T)\delta(s-\tau), \end{aligned} \tag{13}$$

where $\delta(\cdot)$ is the Dirac delta function. The factor $(1-\theta)$, in the last term in equation (13), reflects the fact that the firm must still pay taxes for the value of the remaining assets, $(1-\phi)V_B(r_s, s; P, T)$, if default occurs. Given the payment rate in (13), we have the debt value (with derivation in app. B):

$$\begin{aligned} D(T, X_0; r_0, P) = {} & C\int_0^T E_0^Q[e^{-\int_0^s r(u)du}1(s \leq \tau)]ds \\ & + P\Lambda(r_0, 0; T)[1 - G(T; T, X_0)] \\ & + (1-\phi)P\Lambda(r_0, 0; T)[G(T; T, X_0) + \hat{G}(T; T, X_0)], \end{aligned} \tag{14}$$

where

$$\hat{G}(T; T, X_0) = y\int_0^T e^{y(T-s)}G(s; T, X_0)ds. \tag{15}$$

When the coupon of C is paid, θC, where θ is the tax rate, is deducted from corporate taxes. Thus, the tax shield from the current issue of debt is θ times of the first term in (14). That is, the tax shield is given by

$$tb(T, X_0) = \theta C\int_0^T E_0^Q[e^{-\int_0^s r(u)du}1(s \leq \tau)]ds. \tag{16}$$

Given that

$$D(T, X_0; r_0, P) = \lambda P\Lambda(r_0, 0; T) = \lambda(1-\theta)V_0 e^{-X_0 - yT}, \tag{17}$$

from (14) we have

$$\begin{aligned} tb(T, X_0) = {} & \theta(1-\theta)V_0 e^{-X_0 - yT}[\lambda - 1 \\ & + \phi G(T; T, X_0) - (1-\phi)\hat{G}(T; T, X_0)]. \end{aligned} \tag{18}$$

When bankruptcy occurs, ϕ fraction of the asset value is lost to the bankruptcy procedure. It can be seen that the bankruptcy cost is given by[11]

$$bc(T, X_0) = \phi V_0 e^{-X_0 - yT}[G(T; T, X_0) + \hat{G}(T; T, X_0)]. \tag{19}$$

11. From eq. (13), $1/[(1-\phi)(1-\theta)]$ fraction of the last term in eq. (14) is the present value of the asset value when default occurs; ϕ fraction of this value is the default cost.

The (proportional) transactions cost of issuing and serving the debt is given by

$$tc(T, X_0) = \kappa D(T, X_0; r_0, P) = \kappa\lambda(1-\theta)V_0 e^{-X_0 - yT}, \qquad (20)$$

where κ is the transactions cost per dollar issued.

From (18–20), it is clear that, given both T and those parameters exogenous to the model, the tax shield, bankruptcy cost, and transactions cost depend only on one choice variable, X_0, which in turn determines P through $X_0 = \log\{V_0(1-\theta)/[P\Lambda(r_0, 0; T)e^{yT}]\}$.

C. Total Values in the Presence of Future Periods

Even though at any one point in time there is only one issue of debt outstanding, the total levered firm value will reflect the benefits and costs of all future issues of debt. Therefore, we need to find the total tax benefit, total bankruptcy cost, and total transactions cost from all future periods. Notice that the existence of future periods depends on the firm being solvent in all previous periods. This is a multiperiod, conditional, first-passage time problem. Generally, such problems are very difficult to solve. Fortunately, we are able to transform this difficult problem into a simple one-period, fixed-point one as we show now.

Consider the value of the tax shield at time T from the debt issued at time T. Similar to (18), the tax shield is given by

$$tb_2(T, X_T) = \theta(1-\theta)V_T e^{-X_T - yT}[\lambda - 1 + \phi G(T; T, X_T) - (1-\phi)\hat{G}(T; T, X_T)], \qquad (21)$$

where $X_T = \log\{V_T(1-\theta)/[P_T\Lambda(r_T, 0; T)e^{yT}]\}$, with P_T being the face value of the debt issued at time T. Note that (21) has exactly the same functional form as (18) except for the scaling factor, V_T versus V_0. Indeed, it is clear that the tax benefit of any debt issued in the future will have the same functional form as (18) but scale with the firm's asset value. Note that the scaling factor of the firm's asset value does not affect the optimal X. Hence, if the initial optimal capital structure corresponds with an optimal value X^* for X_0, X^* must also be the optimal value for X_T for the optimal capital structure at time T. In fact, X^* will be the optimal value for all future debt issues. Appendix C presents a detailed proof of this result.

Consequently, the total optimal tax shield from the current and all future issues of debt also scales with the asset value when the debt is issued. Formally, if we let $TB(T, X^*)$ denote the total optimal tax benefit at time zero from the first issue of debt at time zero to all subsequent issues of debt, and if we let $TB_2(T, X^*)$ denote the total optimal tax benefit at time T from the debt issued at time T to all subsequent issues of debt, then it follows that $TB_2(T, X^*) = TB(T, X^*) * V_T/V_0$. However, the total tax benefit $TB(T, X^*)$ at time zero is the tax benefit $tb(T, X^*)$ from the debt issued at time zero, plus the present value of the total tax benefit $TB_2(T, X^*) = TB(T, X^*) * V_T/V_0$ at time T, conditional on no default occurred. Therefore, we have

$$\begin{aligned} TB(T, X^*) &= tb(T, X^*) + E_0^Q\left[\frac{V_T}{V_0} TB(T, X^*) 1(\tau > T) e^{-\int_0^T r(u)du}\right] = tb(T, X^*) \\ &\quad + TB(T, X^*) E_0^Q\left\{e^{\int_0^T [r(u) - y - \sigma_v^2/2]du + \sigma_v W_{vT}^Q} 1(\tau > T) e^{-\int_0^T r(u)du}\right\} \\ &= tb(T, X^*) + TB(T, X^*) e^{-yT} E_0^Q\left[e^{-\sigma_v^2 T/2 + \sigma_v W_{vT}^Q} 1(\tau > T)\right] \\ &= tb(T, X^*) + TB(T, X^*) e^{-yT} H(T; T, X^*), \end{aligned} \tag{22}$$

where

$$H(T; T, X^*) = E_0^Q[e^{-\sigma_v^2 T/2 + \sigma_v W_{vT}^Q} 1(\tau > T)].$$

Equation (22) indicates that the total optimal tax shield $TB(T, X^*)$ is a linear function of itself. Thus, we have reduced a complex multiperiod problem to a simple fixed-point one whose solution is given by

$$TB(T, X^*) = \frac{tb(T, X^*)}{1 - e^{-yT} H(T; T, X^*)}. \tag{23}$$

Similar arguments show that the total bankruptcy cost BC and transactions cost TC are, respectively, given by

$$BC(T, X^*) = \frac{bc(T, X^*)}{1 - e^{-yT} H(T; T, X^*)}, \tag{24}$$

$$TC(T, X^*) = \frac{tc(T, X^*)}{1 - e^{-yT} H(T; T, X^*)}. \tag{25}$$

Finally, the total levered firm value, $TV(T, X^*)$, equals the unlevered firm value, $V_0(1 - \theta)$, plus the value of tax shields, $TB(T, X^*)$, less the value of default and transaction costs, $BC(T, X^*) + TC(T, X^*)$:

$$TB(T, X^*) = V_0(1 - \theta) + TB(T, X^*) - [BC(T, X^*) + TC(T, X^*)]. \tag{26}$$

Note that given other parameters of the model, the total levered firm value $TV(T, X^*)$ is a function of only T and X^*. The firm chooses these two variables to maximize $TV(T, X^*)$. This is a simple unconstrained bivariate maximization problem, and a number of efficient library routines are well suited for this problem.[12]

We caution that without explicit solutions we cannot prove that local maximums do not exist, even though this is unlikely. The reason that an optimal

12. For example, “bconf” from the IMSL Math Library (2003) or “frprmn” from *Numerical Recipes* (Press et al. 2003).

capital structure exists is due to the trade-off between tax shields and the default costs of debt. The marginal tax benefit of debt declines while the marginal default cost of debt increases. Hence, too little or too much debt is unlikely to be optimal, and local maximums are not likely. Similarly, an optimal maturity structure exists because of the trade-off between the gains of tax shields of dynamically adjusting capital structure and the transaction costs of doing so. Again, adjusting too infrequently or too often is unlikely to be optimal, and local maximums are not likely. In our calculations, we have used wide ranges of T and X^* to locate the optimal (T, X^*).

D. Some General Properties of the Optimal Solution

Before we proceed to the numerical results in the next section, we make several general observations about the optimal solution. For notational purpose, let $t_n = nT$ denote the time of each restructuring point where $n = 0, 1, 2$, and so on.

First, note that since our choice of $\lambda = 1/\Lambda(\alpha, 0; T)$ does not depend on the initial interest level r_0, the optimal X^* does not depend on it either. The reason is that equations (19–21) and (24–27) make it clear that the optimal X^* depends on all model parameters and λ, but not on r_0. At first glance, this result may appear puzzling because the initial value at each restructuring point, $X_{t_n} = \log(V_{t_n}(1-\theta)/[P_{t_n}\Lambda(r_{t_n}, 0; T)e^{yT})]$, appears to depend on r_{t_n}. However, it is to be noted that it is the optimal face value of the bond that depends on r_{t_n} through $P^*_{t_n} = V_{t_n}(1-\theta)e^{-X^*-yT}/\Lambda(r_{t_n}, 0; T)$. Thus, the bond face value depends on the initial interest rate level in a particular way such that, given the assumptions of the model, X^* is independent of it. It is noted that, unlike in models with constant interest rates, the optimal coupon and face value at different restructuring points do not scale with the asset value because of their interest rate dependence.

Second, because X^* is the same, the probability of default is the same for all debt issues. The reason is that the drift and diffusion of X_t do not depend on the interest rate levels or the asset values (the two state variables).

Third, like in the Leland (1998)-type dynamic models with constant interest rates, the initial leverage ratio (market debt value over after-tax total levered firm value), D/TV, is the same at each restructuring point. The reason is that from equations (19–21) and (24–27) it is clear that the after-tax total levered firm value depends only on T and X^*, which are the same for all debt issues. The initial market value of any future debt also depends only on T and X^* because $D_{t_n} = \lambda(1-\theta)V_{t_n}e^{-X^*-yT}$. Therefore, the initial leverage ratio depends only on T and X^* as well and is the same for all debt issues.[13]

Finally, unlike in the Leland (1998)-type dynamic models with constant interest rates, the ratio of the coupon C to the face value P of each debt is

13. The asset value factor, V_{t_n}, cancels out in the ratio since the after-tax total levered firm value is also proportional to it.

different and depends on the interest rate level. Note that by the definition of X_t, the optimal face value is given by

$$P^*_{t_n} = V_{t_n}(1-\theta)e^{-X^*-yT}/\Lambda(r_{t_n}, 0; T), \quad (27)$$

where X^* is period independent. Thus $P^*_{t_n}$ is proportional to V_{t_n} and inversely proportional to $\Lambda(r_{t_n}, 0; T)$. Equations (16) and (18) and the similar equations at t_n imply that the optimal coupon, $C^*_{t_n}$, is given by

$$C^*_{t_n} = \frac{(1-\theta)V_{t_n}e^{-X^*-yT}[\lambda - 1 + \phi G(T; T, X^*) - (1-\phi)\hat{G}(T; T, X^*)]}{\int_0^T E^Q_{t_n}[e^{-\int_0^s r(t_n+u)du}1(s \leq \tau)]ds}. \quad (28)$$

Therefore, $C^*_{t_n}$ is also proportional to V_{t_n}. Although $C^*_{t_n}$ depends on r_{t_n}, it is not inversely proportional to $\Lambda(r_{t_n}, 0; T)$. It follows that $C^*_{t_n}/P^*_{t_n}$ is independent of V_{t_n} but not of r_{t_n}, and generally $C^*_{t_n}/P^*_{t_n} \neq C^*_{t_{n+1}}/P^*_{t_{n+1}}$ unless it happens that $r_{t_n} = r_{t_{n+1}}$.

IV. Numerical Results and Comparative Statics

In this section, we implement the dynamic model developed in the previous section. Although the valuation formulas are obtained in closed form, the joint determination of the optimal capital structure and optimal debt maturity structure needs to be performed numerically. In the numerical calculations, the base parameters are fixed as follows: the initial asset value, $V_0 = \$100$; the asset return volatility, $\sigma_v = 0.2$; the corporate tax rate, $\theta = 0.35$; the per dollar transactions cost of issuing debt, $\kappa = 2\%$;[14] the bankruptcy cost parameter, $\phi = 0.5$; the initial interest rate, $r_0 = 0.07$; the correlation coefficient between the firm's asset return and the interest rate, $\rho = 0$; and the payout rate, $y = 0.05$. The parameter values for the interest rate process are $\alpha = 0.0716$, $\beta = 0.261$, and $\sigma_r = 0.0224$.[15] The numerical results are reported in tables 1 and 2.

A. Optimal Capital Structure

Table 1 reports the comparative statics when the interest rate is constant, and table 2 reports the comparative statics when the interest rate is stochastic. Several features are notable.

First, the long-run mean α of the interest rate process, rather than the spot interest rate, has the greater impact on firms' optimal leverage ratios. The reason is that while at any time there is only one finite maturity debt outstanding, the optimal capital structure is based on the total tax shield and total

14. The transactions cost of issuing bonds is usually between 1% and 4%. Fischer et al. (1989) have used transactions costs ranging from 1% to 10%, while Kane et al. (1985) have considered 1% and 2% in their calculations.

15. These values are taken from the empirical estimates in Aït-Sahalia (1999). In particular, the long-run mean, α, is about 7%.

TABLE 1

(1)	T (Years) (2)	C ($) (3)	P ($) (4)	Credit Spread ($\times 10^4$) (5)	Leverage Ratio (%) (6)	TB ($) (7)	BC ($) (8)	TC ($) (9)	Debt Benefit (%) (10)	TV ($) (11)
Base	3.50	1.81	25.35	14.92	34.81	11.99	1.07	3.10	12.03	72.82
$\kappa = 0.015$	2.64	1.93	27.16	12.45	36.83	12.93	.98	3.23	13.43	73.73
$\kappa = 0.025$	4.42	1.72	24.01	17.44	33.28	11.27	1.16	2.96	10.99	72.14
$\theta = 0.2$	5.81	1.83	25.72	13.32	30.59	6.83	.76	2.01	5.07	84.05
$\theta = 0.5$	2.52	1.53	21.40	13.49	35.70	14.56	1.08	3.54	19.88	59.94
$\sigma_v = 0.15$	4.34	2.21	31.02	13.00	41.22	14.50	1.13	3.11	15.80	75.27
$\sigma_v = 0.25$	3.02	1.50	20.94	16.71	29.47	9.98	1.00	2.93	9.31	71.05
$r_0 = 0.05$	4.15	1.08	21.11	11.20	30.37	7.44	.72	2.22	6.92	69.50
$r_0 = 0.09$	3.16	2.63	28.66	18.38	37.42	16.83	1.41	3.81	17.85	76.60
$\phi = 0.4$	3.64	1.85	25.83	15.19	35.38	12.13	1.09	3.03	12.32	73.01
$\phi = 0.6$	3.39	1.78	24.97	14.67	34.35	11.86	1.05	3.14	11.80	72.67
$y = 0.04$	3.64	1.86	25.95	15.06	34.65	14.98	1.33	3.73	15.25	74.91
$y = 0.06$	3.37	1.77	24.76	14.62	34.67	9.97	.89	2.66	9.88	71.42

NOTE.—Column 1 denotes the changing parameter. The other parameters are fixed at their base values. Columns 2–11 represent the optimal maturity, optimal coupon rate, optimal principal, credit spread in basis points, optimal leverage ratio, total tax benefit, total bankruptcy cost, total transaction cost, debt benefit as a percentage of the after-tax unlevered firm value, $100(TB - BC - TC)/[V_0(1-\theta)]$, and optimal after-tax levered firm value, respectively.

default cost associated with both the existing debt and all future debts. It is the long-run mean rather than the spot interest rate level that has the greatest impact on the values of these future debts.

Our calculations indicate that the results with stochastic interest rates are similar to a model with a constant interest rate whose value is calibrated to the long-run mean of the stochastic interest rate process. Thus, our model provides a justification for the use of constant interest rates in capital structure models, provided that the long-run mean of the interest rate process instead of the spot interest rate or a YTM is used as the input for the constant interest rate. Without this insight, the traditional capital structure models could not be used to explain empirical findings. Because the long-run mean is not an empirical observable, one must estimate it using historical data as well as a specific model for the interest rate process. Hence, it is essential to consider a stochastic interest rate process in the determination of the capital structure.

When the interest rate is assumed to be a constant, its level does have a significant impact on a firm's capital structure decisions. For example, when the constant interest rate is 5%, the optimal levered firm value, net debt benefit, and the market value of debt (leverage times levered firm value) are $69.5, 6.92%, and $21.11, respectively, whereas the corresponding values are $76.6, 17.85%, and $28.66, respectively, when the constant interest rate is 9%. There are two reasons that the interest rate level is important in a constant interest rate model. First, the usual explanation is that at a higher interest rate, the risk-neutral drift of the asset value process is higher, and the risk-neutral default probability becomes lower for the same coupon and principal. Therefore, the

TABLE 2

(1)	T (Years) (2)	C ($) (3)	P ($) (4)	Credit Spread ($\times 10^4$) (5)	Leverage Ratio (%) (6)	TB ($) (7)	BC ($) (8)	TC ($) (9)	Debt Benefit (%) (10)	TV ($) (11)
A. Changing one parameter value at a time:										
Base	3.20	1.86	25.59	14.15	35.22	12.35	1.03	3.40	12.17	72.91
$\kappa = 0.015$	2.48	1.99	27.42	12.08	37.21	13.33	.96	3.47	13.69	73.90
$\kappa = 0.025$	3.93	1.76	24.20	16.09	33.67	11.58	1.10	3.33	11.02	72.16
$\theta = 0.2$	5.06	1.86	25.80	11.94	30.85	6.98	.70	2.29	4.99	84.00
$\theta = 0.5$	2.37	1.57	21.60	13.11	36.01	15.01	1.07	3.80	20.29	60.14
$\sigma_v = 0.15$	3.68	2.26	31.20	11.67	41.62	14.92	1.03	3.65	15.75	75.23
$\sigma_v = 0.25$	2.85	1.54	21.18	16.23	29.85	10.30	.99	3.14	9.50	71.17
$r_0 = 0.05$	3.20	1.82	24.50	14.15	35.22	12.35	1.03	3.40	12.17	72.91
$r_0 = 0.09$	3.20	1.90	26.73	14.15	35.22	12.35	1.03	3.40	12.17	72.91
$\phi = 0.4$	3.30	1.89	26.07	14.28	35.79	12.49	1.05	3.35	12.45	73.09
$\phi = 0.6$	3.13	1.83	25.21	14.00	34.77	12.22	1.02	3.44	11.95	72.77
$y = 0.04$	3.29	1.90	26.16	14.10	35.00	15.43	1.27	4.15	15.40	75.01
$y = 0.06$	3.12	1.82	25.03	14.01	35.13	10.26	.86	2.90	10.01	71.50
$\rho = -0.3$	3.53	1.90	26.11	14.78	35.77	12.54	1.09	3.17	12.73	73.28
$\rho = 0.3$	2.99	1.83	25.12	13.77	34.72	12.16	.99	3.56	11.70	72.60
$\alpha = 0.04$	4.35	.77	19.93	8.31	27.16	5.24	.48	1.87	4.44	67.88
$\alpha = 0.10$	2.80	2.96	28.33	18.58	38.35	19.36	1.48	4.46	20.64	78.42
$\beta = 0.15$	3.12	1.86	25.64	13.80	35.33	12.37	1.01	3.49	12.10	72.86
$\beta = 0.35$	3.26	1.86	25.57	14.38	35.16	12.33	1.05	3.34	12.22	72.94
$\sigma_r = 0.01$	3.40	1.87	25.56	14.96	35.11	12.36	1.09	3.21	12.41	73.07
$\sigma_r = 0.04$	2.87	1.84	25.61	12.74	35.41	12.27	.94	3.77	11.62	72.55
B. Holding r_0 at 4% and 6-year risk-free yield at 7%:										
$\alpha = 0.1028$	2.78	2.97	26.90	19.01	38.55	20.07	1.52	4.55	21.53	79.00
$\beta = 3.3352$	3.46	1.86	25.39	15.19	35.06	12.36	1.10	3.16	12.46	73.10
$\alpha = 0.0415$	4.26	.86	21.95	8.62	27.73	5.55	.51	1.95	4.75	68.09
$\sigma_r = 0.0859$	2.23	1.79	26.63	9.67	35.59	11.76	.72	4.74	9.70	71.31

NOTE.—Column 1 denotes the changing parameter. The other parameters are fixed at their base values. Columns 2–11 represent the optimal maturity, optimal coupon rate, optimal principal, credit spread in basis points, optimal leverage ratio, total tax benefit, total bankruptcy cost, total transaction cost, debt benefit as a percentage of the after-tax unlevered firm value, $100(TB - BC - TC)/[V_0(1-\theta)]$, and optimal after-tax levered firm value, respectively. In panels B and C, α, β, or σ_r is chosen so that given the initial interest rate, 4% or 10%, the 6-year risk-free yield is 7%.

firm optimally levers more if the interest rate is higher, 30.37% versus 37.42%, when the constant interest rate is 5% versus 9%. An alternative explanation for this result is that at a higher interest rate, to reduce the impact of more discounting to obtain the present value of tax shields, the firm optimally takes on more debt in earlier periods, resulting in a higher initial leverage ratio.[16]

By contrast, when the interest rate is stochastic, the firm makes capital structure decisions based mainly on the long-run mean of the interest rate process. Given that all debts are issued at $\lambda = 1/\Lambda(\alpha, 0; T)$ times of the value of a risk-free, zero-coupon debt with the same face value (assumption 6), the optimal X^* is independent of the initial (spot) interest rate level. Consequently, the levered firm value, net debt benefit, and optimal leverage ratio are all independent of the initial interest rate. Furthermore, the initial leverage ratio

16. We are grateful to the anonymous referee for this explanation.

at each restructuring point is the same, although the prevailing interest rate at each restructuring point is different.[17]

As discussed in Section III.D, we note that the optimal coupon and principal of the debt in our model do depend on the initial interest rate level, although the debt value (the present value of the coupon and principal) and initial leverage ratio do not. Thus, the firm's optimal strategy is to issue debt at each restructuring point whose market value is proportional to the prevailing asset value but independent of the interest rate level. This can be achieved by issuing debt with higher/lower coupon and principal if the interest rate has increased/decreased but by keeping the present value of coupon and principal independent of the interest rate level.

To examine further how the changes in the term structure affect the capital structure decisions, we fix the initial interest rate at 4% in panel B and 10% in panel C in table 2 but change α, β, or σ_r of the interest rate process so that the 6-year risk-free yield is kept at 7%. The corresponding long-run means, α, in the last four entries in table 2 are 10.28%, 7.16%, 4.15%, and 7.16%, respectively. Thus, relative to the 6-year, risk-free yield or the long-run mean, the two entries in panel B correspond with an upward sloping term structure, and those in panel C correspond with a downward sloping one. The results are consistent with the previous observation that the most important interest rate parameter is the long-run mean of the interest rate process. For example, the first entry in panel B with a long-run mean of 10.28% levers more aggressively than the first entry in panel C with a long-run mean of 4.15%. The reason is that the long-run mean is more important than the initial interest rate in determining capital structures. A model with constant interest rates would have predicted that the firm would lever more aggressively at 10% than at 4% of the interest rate.

Second, the leverage ratios obtained from our model (around 35%) compare more favorably with the historical average of about 30% for a typical large publicly traded firm in the United States than the predictions of most static optimal capital structure models. The reason for the lower leverage ratio in our dynamic model is that, in a static model, a firm cannot adjust its debt level in the future and, therefore, issues debt more aggressively. By contrast, a firm with an option to restructure in the future issues debt less aggressively, for it can adjust its capital structure when the firm's asset value changes.[18] Moreover, the cost of default for the firm includes not only the cost of bankruptcy but also the loss of the option value of adjusting the debt level in the future. Hence, default is more costly in the dynamic setting, further decreasing the initial optimal leverage ratio.

Third, the comparative static results with respect to the correlation coef-

17. This result is the same as in some dynamic capital structure models with constant interest rates (see, e.g., Goldstein et al. 2001; Ju et al. 2005). In those models, however, the leverage ratio is the same at each restructuring point, but it is different when the interest rate is different.

18. This was first noted in Goldstein et al. (2001) in a different setting with constant interest rates. A similar implication is also noted in Titman and Tsyplakov (2002).

ficient between the interest rate process and the firm's asset value process in table 2 reveal that the correlation coefficient has little impact on a firm's capital structure decisions. The reason is that the covariance $\sigma_v \rho_i \sigma_i$ between the asset value return and the interest rate process is much smaller than the typical values of the variance of the asset value returns, σ_v^2.[19] Thus, σ_v has a first-order effect on the pricing of securities.

Finally, in the last six rows of panel A in table 2, we examine the comparative statics on the term structure parameters α, β, and σ_r. When α changes, the long-run mean of the interest rate also changes. Note that a firm optimally issues more debt when the long-run mean of the interest rate process is higher, similar to the effects of a higher interest rate level in a model with constant interest rates.[20] The two rows with $\beta = 0.15$ and $\beta = 0.35$ consider the effect of the speed of the mean-reverting parameter β. Given that the initial interest rate of 7% is close to the long-run mean of 7.16%, the mean-reversion parameter β has a small impact on model outputs. The last two rows of panel A in table 2 indicate that although the interest rate volatility increases four times from 0.01 to 0.04, the increase in the leverage ratio is quite small. The reason is that the return of the asset value is much more volatile than the interest rate process, and thus the added volatility in the interest rate process has little impact on the first-passage time to the default boundary.

In sum, our calculations suggest that the long-run mean is the most important variable of a stochastic interest rate process in determining the optimal capital structure.

B. *Optimal Maturity Structure*

The shareholders' strategic consideration of dynamic adjustments of the debt level yields an optimal maturity structure. The trade-off in this case is between the transactions costs of issuing debt and the gains from adjusting debt level dynamically. On the one hand, the firm should issue short-maturity debt and therefore afford itself the opportunity to issue new debt optimally, depending on the firm value when the old debt matures. On the other hand, if new debts are issued too often, transactions costs will become too large. When the firm behaves optimally, in addition to an optimal capital structure, an optimal maturity structure emerges.

The optimal maturities are reported in column 2 in table 1 and in table 2. The tables show that the two most important parameters are the tax rate and the transactions cost because the optimal maturity is the result of the trade-off between the transactions cost of and the gain from adjusting the debt level in the future.

Barclay and Smith (1995) find that during the period of 1974–92, firms in their sample have 51.7% of their debts due in more than 3 years. Because on average, the debt would have existed for half of its lifetime at any point in

19. See, e.g., eq. (B7).
20. Compare the two rows with $r_0 = 5\%$ and $r_0 = 9\%$ in table 1.

time, the median maturity at issue appears to be a little over 6 years. Stohs and Mauer (1996) find that the median time to maturity is 3.38 years for their sample. Therefore, the mean maturity of all debts at issue appears to be between 6 and 7 years, consistent with Barclay and Smith (1995). Most of the optimal maturities in the two tables are between 3 and 4 years, shorter than the empirical values. While our theoretical values are smaller than 6 years, it is to be noted that the optimal maturities in the two tables with a 20% tax rate are close to 6 years. Even though 35% is the top corporate tax rate, the effective corporate tax rate for most firms is likely to be lower due to investment credit, loss-carry forward, and personal tax effects.

The comparative static results with respect to the firm's asset volatility indicate that the optimal debt maturity is inversely related to it.[21] This is due to an optionlike effect. The higher the asset volatility, the greater the value of the option to adjust the capital structure in the future. The gains from dynamically adjusting the debt level are reduced if the firm's asset value is less volatile. In this case the firm recapitalizes its capital structure less frequently. This appears to be consistent with the empirical evidence. For example, in the sample of Barclay and Smith (1995), 36.6% of equally weighted debts mature in more than 5 years, but the percentage increases to 45.9% with value-weighted debts. Therefore, larger firms tend to have longer maturity debts. Because larger and more mature firms are less volatile than smaller and less mature ones, it follows that less volatile firms have longer maturity debts. We also note a similar inverse relation between optimal maturity and the interest rate process volatility, because a higher interest rate volatility implies more variability in the risk-neural firm asset value process. For example, the optimal maturity is 3.4 years when $\sigma_r = 0.01$, while it is 2.87 years when $\sigma_r = 0.04$. Similarly, the correlation between the interest rate process and the asset value process has a similar impact on debt maturity structure.

The comparative static results with respect to the long-run mean α reveal that the optimal maturity is inversely related to α. When the long-run mean is low, the future growth rate in the firm's asset value, as given in equation (1), is likely low as well. With a small asset growth rate, the potential benefit of adjusting the debt level is small. Therefore, the firm optimally issues longer maturity debts when the long-run mean is smaller.

C. *Credit Spreads*

To compute the credit spread, let D_d (y_d) and D_f (y_f) be the prices (yields) of the defaultable and risk-free debts with maturity T, respectively. Solve y_d from $D_d = C\int_0^T e^{-y_d ds} + Pe^{-y_d T} = C/y_d(1 - e^{-y_d T}) + Pe^{-y_d T}$. Similarly, solve y_f from $D_f = C\int_0^T e^{-y_f ds} + Pe^{-y_f T} = C/y_f(1 - e^{-y_f T}) + Pe^{-y_f T}$. The credit spread (in basis points) is computed as $10{,}000(y_d - y_f)$.

Table 1 and table 2 indicate that with our model inputs, when firms are

21. In Childs et al. (2003) the optimal maturity is increasing in the assets-in-place volatility.

optimally levered, the credit spreads are around 14 basis points. These values are lower than the historical credit spread. Leland (1994) reports that the average historical credit spread is about 52 basis points after the impact of call provisions is eliminated. There are two main reasons that the credit spreads in our model are small. First, debt maturities in the two tables are generally quite short. Shorter maturity debts normally have lower default risks and thus lower credit spreads. Second, it is well known that structural models like ours generally yield credit spreads that are too small, indicating that factors not considered in a typical contingent-claim structural model play important roles in determining the credit spread.[22] For example, Tang and Yan (2004) demonstrate that macroeconomic conditions can affect credit spreads significantly. Childs et al. (2003), by contrast, show that incorporating agency conflicts between bondholders and shareholders can generate credit spreads that are more consistent with observations.

While our calculations show that the model predictions of credit spreads are small, comparative static calculations do indicate that the various underlying variables affect the credit spreads in expected ways. For example, with the coupon and principal fixed at their optimal base case values, that is, 1.86 and 25.59, respectively, increasing the debt maturity to 6 years increases the credit spread to 47.27 basis points from 14.15, while it increases to 65.63 basis points if the asset volatility increases to 0.25. When the fraction of the asset value lost to default increases to 0.6, the credit spread increases to 16.87. When the correlation coefficient between the interest rate process and the asset value process increases from 0 to 0.3, the credit spread increases to 19.44.

D. Optimal Capital Structure with Fixed Maturity or Optimal Maturity with Fixed Debt Capacity

So far we have discussed the model implications assuming that the firm optimizes both its capital structure and its debt maturity. With a dynamic framework, our model is flexible enough to allow for the determination of the optimal capital structure with a fixed debt maturity or for the optimal debt maturity with a fixed amount of debt to issue. The resulting model is still dynamic in that the firm optimally rebalances its capital structure with an exogenously given frequency (debt maturity) or optimally chooses its debt maturity in meeting its requirement of issuing a certain amount of debt. The results are reported in table 3. Since the initial interest rate of 7% is close to the long-run mean of 7.16%, the results for constant and stochastic interest rates are similar, and we will focus only on panel C and panel D. When the maturity is shorter, the default risk is smaller for a given level of debt, and thus the firm can optimally issue more debt. This results in a higher leverage ratio. However, since the default risk is quite small even at high leverage ratios when the debt maturity is short, the credit spread is quite low. When the firm is constrained to issue longer maturity debts, as expected, the credit

22. Jones, Mason, and Rosenfeld 1984; Tang and Yan 2004.

TABLE 3

(1)	$D \mid T$ (\$ | Years) (2)	C (\$) (3)	P (\$) (4)	Credit Spread ($\times 10^4$) (5)	Leverage Ratio (%) (6)	TB (\$) (7)	BC (\$) (8)	TC (\$) (9)	Debt Benefit (%) (10)	TV (\$) (11)
A. Constant interest rate and changing maturity:										
$T = 2.0$	28.44	2.01	28.44	8.00	39.46	13.63	.67	5.90	10.87	72.07
$T = 4.0$	24.70	1.77	24.70	16.99	33.93	11.63	1.18	2.66	11.99	72.79
$T = 6.0$	23.15	1.68	23.15	24.55	31.95	10.72	1.52	1.74	11.49	72.47
$T = 8.0$	22.61	1.65	22.61	31.55	31.33	10.29	1.80	1.33	11.01	72.16
$T = 10.0$	22.67	1.67	22.67	38.29	31.51	10.14	2.08	1.11	10.70	71.95
$T = 12.0$	23.16	1.73	23.16	44.97	32.24	10.17	2.34	.99	10.53	71.84
B. Constant interest rate and changing debt capacity:										
$D = 10.0$	9.35	.70	10.0	.87	14.50	4.50	.02	.54	6.07	68.95
$D = 15.0$	7.32	1.06	15.0	3.63	21.20	6.88	.15	.98	8.85	70.75
$D = 20.0$	5.27	1.42	20.0	8.16	27.71	9.34	.46	1.71	11.03	72.17
$D = 25.0$	3.59	1.79	25.0	14.40	34.33	11.82	1.02	2.98	12.03	72.82
$D = 30.0$	2.39	2.17	30.0	23.67	41.61	14.23	2.01	5.12	10.93	72.10
$D = 35.0$	1.57	2.59	35.0	40.50	50.72	16.45	3.91	8.54	6.16	69.00

C. Stochastic interest rate and changing maturity:										
$T = 2.0$	28.47	2.05	28.40	8.30	39.37	13.89	.69	5.90	11.22	72.30
$T = 4.0$	24.47	1.78	24.37	17.60	33.60	11.66	1.20	2.64	12.02	72.82
$T = 6.0$	22.62	1.65	22.51	25.53	31.28	10.54	1.53	1.70	11.24	72.31
$T = 8.0$	21.76	1.60	21.64	32.85	30.28	9.93	1.81	1.28	10.53	71.84
$T = 10.0$	21.49	1.59	21.37	39.84	30.06	9.61	2.05	1.05	10.01	71.50
$T = 12.0$	21.66	1.61	21.53	46.66	30.38	9.49	2.29	.92	9.66	71.28
D. Stochastic interest rate and changing debt capacity:										
$D = 10.0$	8.32	.70	9.95	.85	14.51	4.55	.02	.59	6.06	68.94
$D = 15.0$	6.45	1.06	14.93	3.24	21.20	6.96	.13	1.08	8.83	70.74
$D = 20.0$	4.75	1.43	19.91	7.19	27.71	9.47	.41	1.88	11.05	72.18
$D = 25.0$	3.36	1.81	24.91	13.13	34.29	12.00	.93	3.17	12.15	72.90
$D = 30.0$	2.31	2.20	29.92	22.45	41.50	14.49	1.91	5.29	11.21	72.29
$D = 35.0$	1.55	2.64	34.93	39.34	50.50	16.77	3.80	8.67	6.62	69.30

NOTE.—Column 1 denotes the changing debt maturity or debt capacity. Column 2 denotes the optimal debt capacity given maturity or the optimal maturity given debt capacity. Columns 3–11 represent the coupon rate, principal, credit spread in basis points, leverage ratio, total tax benefit, total bankruptcy cost, total transaction cost, debt benefit as a percentage of the after-tax unlevered firm value, $100(TB - BC - TC)/[V_0(1-\theta)]$, and the after-tax levered firm value, respectively. In panels A and B, the interest rate is kept at a constant 7%, while in panels C and D, the spot rate r_0 is kept at 7%, and the interest rate process parameters, α, β, σ_r, and ρ, are kept at their base values.

spread widens. For example, the credit spread of the optimally issued debt increases from 8.3 basis points when debt maturity is fixed at 2 years to 46.7 basis points when debt maturity is fixed at 12 years.

Note that the optimal leverage ratio is not monotonically decreasing in debt maturity. When the debt maturity is short, the firm can afford to optimally issue more debt because the default risk is small. By contrast, if the firm is constrained to rebalance less frequently (longer debt maturity), then the gain of dynamically adjusting capital structure is small. As a result, the firm issues debt more aggressively. The reasons for issuing debt more aggressively at shorter and longer maturity are different. Neither too short nor too long maturity may be optimal because rebalancing too often increases transaction costs and because rebalancing too infrequently reduces the tax benefits.

Panel D in table 3 examines a firm's debt maturity choice given that it needs to issue a certain amount of debt. When the firm is constrained to have low leverage, the firm optimally issues longer maturity debt because the benefit of rebalancing is limited. By contrast, when the firm is constrained to issue higher amounts of debt, it optimally chooses short maturity to reduce the risk of default. Constraining the firm to too low leverage or too high leverage is not optimal for the firm value. Given the choice, the firm will optimally lever moderately.

In figures 1–4, we continue to exploit the relationships among firm value, leverage, and maturity. Figure 1 plots the optimal leverage ratio as a function of debt maturity. As we discussed above, a firm optimally levers more aggressively when the debt maturity is low or high. Figures 2–4 plot the levered firm value as a function of the leverage ratio when the debt maturity is fixed at 2, 6, and 10 years, respectively. Initially, the levered firm value increases as the leverage increases. However, after a point, the marginal benefit of leveraging becomes smaller than the marginal cost. The peak in each figure denotes the optimal levered firm value versus the optimal leverage ratio.

V. Concluding Remarks

The existing models of the optimal capital structure consider neither stochastic interest rates nor optimal maturity structure. This article develops a model of optimal capital structure and optimal maturity structure with a Vasicek (1977) interest rate process. Valuation formulas are obtained in closed form. A novel fixed-point argument is used to obtain the total tax shield, default cost, and transactions cost for a dynamic model with potentially an infinite number of debt issues.

The trade-off between the bankruptcy costs and the tax shields of debts yields an optimal capital structure. The trade-off between the gains from adjusting the capital structure periodically and the costs of doing so yields an optimal maturity structure. In our model the optimal capital structure and optimal maturity structure are interdependent and determined jointly.

When the interest rate is a constant, the level of the interest rate affects

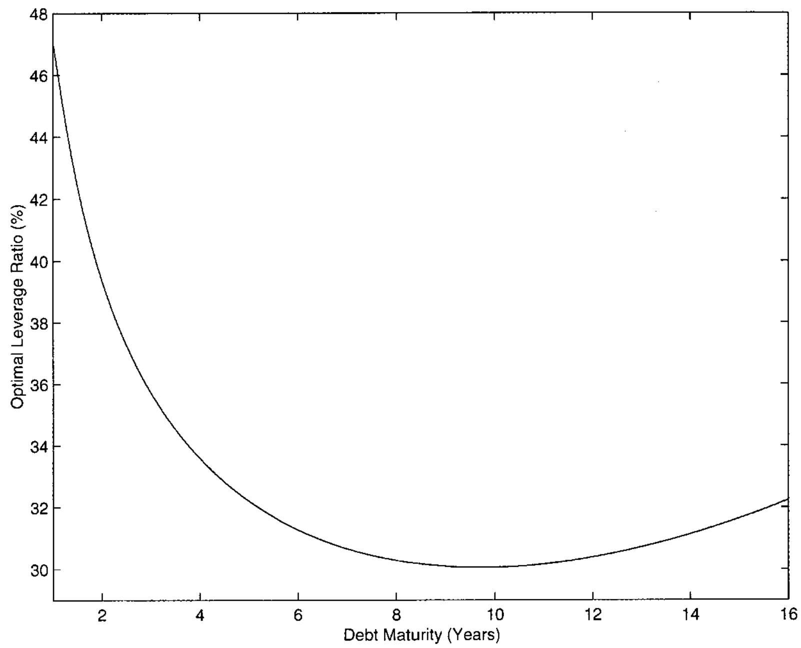

FIG. 1.—Optimal leverage ration as a function of debt maturity. The short-term interest rate is assumed to follow a Vasicek process. Its parameter values are $\alpha =$ 0.0716, $\beta = 0.261$, and $\sigma_r = 0.0224$. The correlation between the interest rate process and the firm asset value procedure is assumed to be zero.

both the optimal leverage ratio and the optimal maturity. When the interest rate follows a mean-reverting stochastic process, the initial level as well as the long-run mean of the interest rate process affect the pricing of both the coupon and the principal of the debt. The coupon and principal are determined in such a way that the market value of the debt (the present value of both the coupon payments and the principal) is independent of the spot interest rate level. In particular, when the interest rate is higher/lower, coupon and principal are also higher/lower, but their present values remain the same. Consequently, the optimal leverage ratio and optimal maturity are independent of the spot interest rate level.

While the long-run mean is shown to be important in determining the optimal capital structure, numerical results show that the correlation between the stochastic interest rate and the return of the firm's asset has little impact. The leverage ratios obtained from our model are consistent in spirit with the empirical observations.

The optimal maturity and the credit spread of optimally issued debt in our model appear to be small compared with observed values, indicating that factors not considered play important roles in determining them. The trade-

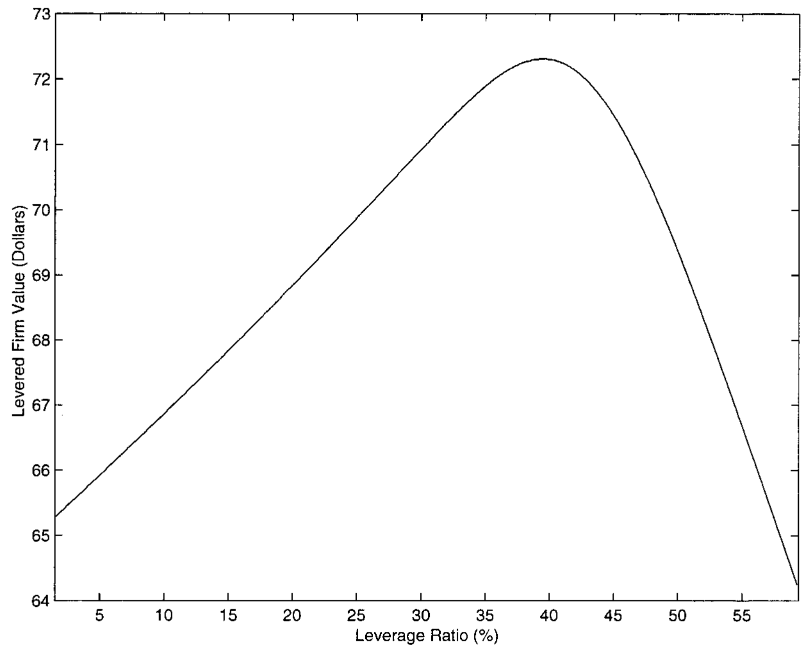

FIG. 2.—Levered firm value as a function of leverage ratio. Debt maturity is fixed at $T = 2.0$ years. The short-term interest rate is assumed to follow a Vasicek process. Its parameter values are $\alpha = 0.0716$, $\beta = 0.261$, and $\sigma_r = 0.0224$. The correlation between the interest rate process and the firm asset value process is assumed to be zero.

off between the transactions costs of issuing bonds and the gains from adjusting bonds dynamically may not be the only factor determining the optimal maturity structure. It is well known that contingent-claim-based structural models tend to generate credit spreads that are too small. Models that incorporate macroeconomic variables and agency conflicts seem to be able to yield credit spreads that are closer to observed values (Childs et al. 2003; Tang and Yan 2004).

For tractability, the default boundary $V_B(y_t, t; P, T)$ is exogenously specified. Extending our model to allow for an endogenous default boundary, in the spirit of Leland (1994) and Leland and Toft (1996), would be an important but challenging topic for future research. In addition, Darrough and Stoughton (1986) develop a capital structure model in which moral hazard and adverse selection problems exist and financing decisions affect investments. Titman and Tsyplakov (2002) also allow a firm's financing decisions to affect its investments. Incorporating these features would be another fruitful but difficult topic for future research.

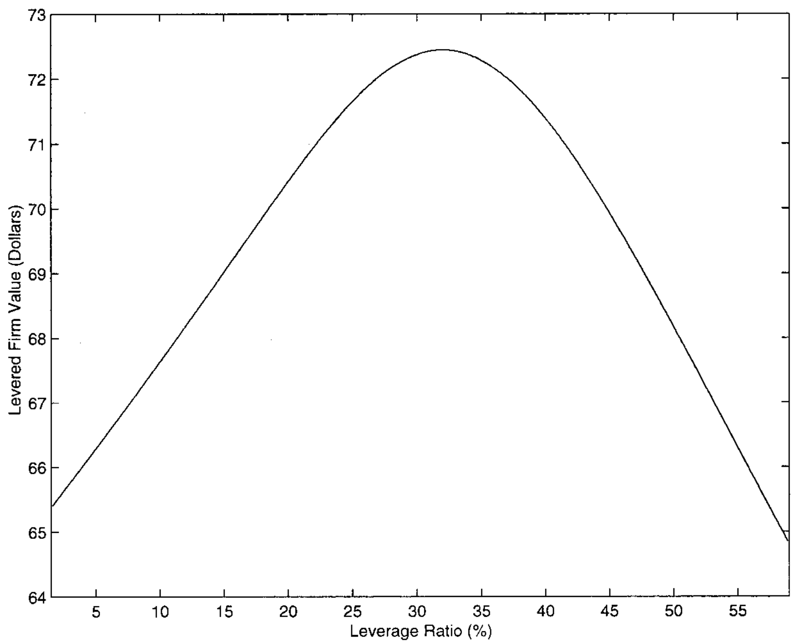

FIG. 3.—Levered firm value as a function of leverage ratio. Debt maturity is fixed at $T = 6.0$ years. The short-term interest rate is assumed to follow a Vasicek process. Its parameter values are $\alpha = 0.0716$, $\beta = 0.261$, and $\sigma_r = 0.0224$. The correlation between the interest rate process and the firm asset value process is assumed to be zero.

Appendix A

The Forward Risk-Neutral Measure

In this appendix, we use the Girsanov theorem to derive the T-forward risk-neutral measure in a multidimensional setting. Without loss of generality, we assume a probability space Q generated by two standard Wiener processes:

$$W_t^Q = \begin{pmatrix} W_{1t}^Q \\ W_{2t}^Q \end{pmatrix}, \tag{A1}$$

with correlation matrix

$$\tilde{\rho}(t) = \begin{pmatrix} 1 & \rho(t) \\ \rho(t) & 1 \end{pmatrix}. \tag{A2}$$

In the following, Q should be interpreted as the risk-neutral probability measure and r_t as the risk-free interest rate given by

$$dr_t = \mu(r,t)dt + \sigma(r,t)dW_{2t}^Q. \tag{A3}$$

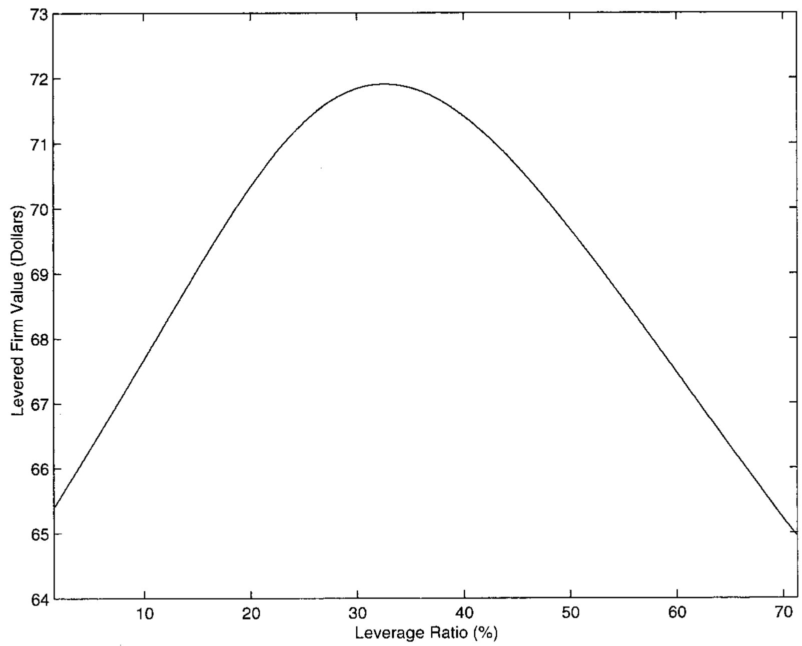

FIG. 4.—Levered firm value as a function of leverage ratio. Debt maturity is fixed at $T = 10.0$ years. The short-term interest rate is assumed to follow a Vasicek process. Its parameter values are $\alpha = 0.0716$, $\beta = 0.261$, and $\sigma_r = 0.0224$. The correlation between the interest rate process and the firm asset value process is assumed to be zero.

We leave other random variables generated by W_{1t}^Q and W_{2t}^Q unspecified.

Suppose we want to compute the following expectation

$$h = E_0^Q\left[\frac{e^{-\int_0^T r(u)du}}{\Lambda(r_0, 0; T)} Z(\{\cdots\}, T)\right], \tag{A4}$$

where $\Lambda(r_0, 0; T) = E_0^Q\left\{\exp\left[-\int_0^T r(u)du\right]\right\}$ is the discount bond price at $t = 0$ with maturity T, and the curly brackets $\{\cdots\}$ indicate that the random variable $Z(\{\cdots\}, T)$ may depend on the sample path in probability space Q from 0 to T. Define

$$\xi_T = \frac{e^{-\int_0^T r(u)du}}{\Lambda(r_0, 0; T)}. \tag{A5}$$

We then have

$$h = E_0^Q[\xi_T Z(\{\cdots\}, T)]. \tag{A6}$$

It is clear that ξ_T is strictly positive and $E_0^Q(\xi_T) = 1$. Therefore, ξ_T can be used as a

Radon-Nikodym derivative to define a new probability measure R_T equivalent to the original measure Q. That is,

$$E_0^{R_T}(1_{\{A\}}) = E_0^Q(\xi_T 1_{\{A\}}) \tag{A7}$$

for any event A. Under the new (forward) risk-neutral measure R_T,

$$h = E_0^{R_T}[Z(\{\cdots\}, T)]. \tag{A8}$$

To find the Wiener processes under R_T, define the likelihood ratio

$$\xi_t = E_t^Q(\xi_T) = \frac{e^{-\int_0^t r(s)ds}\Lambda(r_t, t; T)}{\Lambda(r_0, 0; T)}. \tag{A9}$$

Ito's lemma implies that

$$\begin{aligned} d\log\xi_t &= -rdt + \frac{d\Lambda(r_t,t;T)}{\Lambda(r_t,t;T)} - \frac{1}{2}\left[\frac{d\Lambda(r_t,t;T)}{\Lambda(r_t,t;T)}\right] \\ &= [-r\Lambda + \Lambda_t + u(r,t)\Lambda_r + (1/2)\sigma^2(r,t)\Lambda_{rr}]dt/\Lambda \\ &\quad + \frac{\sigma(r,t)\Lambda_r}{\Lambda}dW_2^Q(t) - \frac{1}{2}\left[\frac{\sigma(r,t)\Lambda_r}{\Lambda}\right]dt. \end{aligned} \tag{A10}$$

The term inside the square bracket on the second line of (A10) is the fundamental partial differential equation satisfied by the discounted bond price Λ and equals zero. Therefore (with $\xi_0 = 1$), we have

$$\xi_t = e^{-(1/2)\int_0^t \varphi^T(s)\varphi(s)ds + \varphi^T(s)dW_t^Q}, \tag{A11}$$

where T denotes the transpose and

$$\varphi(t) = \begin{bmatrix} 0 \\ \dfrac{\sigma(r,t)\Lambda_r}{\Lambda} \end{bmatrix}, dW_t^Q = \begin{pmatrix} dW_{1t}^Q \\ dW_{2t}^Q \end{pmatrix}. \tag{A12}$$

Thus, using (A11) we can rewrite (A4) as

$$h = E_0^Q\left\{\exp\left[\int_0^t -\varphi^T(s)\varphi(s)ds/2 + \varphi^T(s)dW_t^Q\right]Z(\{\cdots\}, T)\right\} = E_0^{R_T}[Z(\{\cdots\}, T)]. \tag{A13}$$

Now the multidimensional Girsanov theorem means that $W_t^{R_T}$ defined by[23]

$$W_t^{R_T} = \begin{pmatrix} W_{1t}^{R_T} \\ W_{2t}^{R_T} \end{pmatrix} = W_t^Q - \int_0^t \tilde{\rho}(s)\varphi(s)ds \tag{A14}$$

are two standard Wiener processes under R_T with the correlation matrix $\tilde{\rho}(t)$. In differential form, we have

$$dW_t^{R_T} = dW_t^Q - \tilde{\rho}(t)\varphi(t)dt. \tag{A15}$$

23. See, e.g., Duffie (2001) for a review.

Appendix B

Derivation of $G(t;T, X_0)$, $H(t;T, X_0)$ and (14)

Recall that

$$G(t;T,X_0) = E_0^Q\left[\frac{e^{-\int_0^T r(u)du}}{\Lambda(r_0,0;T)}1(\tau<t)\right], \tag{B1}$$

$$H(t;T,X_0) = E_0^Q[e^{-\sigma_v^2 t/2+\sigma_v W_{vt}^Q}1(\tau>t)], \tag{B2}$$

where $1(\cdot)$ is the indicator function, and where $\tau = \min\{t : X_t \leq 0\}$ is the first passage time that X_t, defined in (9), crosses the origin from above ($X_0 > 0$).

A. Derivation of $G(t;T, X_0)$

Note that (B1) has the form of (A4). Therefore, we can define a new probability measure R_T under which

$$G(t;T,X_0) = E_0^{R_T}[1(\tau<t)] \tag{B3}$$

and is the distribution function of τ under R_T. With the interest rate given by the Vasicek process in (2), the corresponding $\varphi(t)$ in (A12) becomes $\varphi^T(t) = [0, -\sigma_r B(t;T)] = [0, -\sigma_p(t;T)]$, and the new standard Wiener processes in (A15) are given by

$$dW_{vt}^{R_T} = dW_{vt}^Q + \rho\sigma_p(t;T)dt, \tag{B4}$$

$$dW_{rt}^{R_T} = dW_{rt}^Q + \sigma_p(t;T)dt. \tag{B5}$$

Rewriting (9) using the new $dW_{vt}^{R_T}$ and $dW_{rt}^{R_T}$, we have

$$\begin{aligned} dX_t &= -[\sigma_v^2 + \sigma_p^2(t;T) + 2\rho\sigma_p(t;T)\sigma_v]/2 + \sigma_v dW_{vt}^{R_T} + \sigma_p(t;T)dW_{rt}^{R_T} \\ &= -\sigma^2(t;T)dt/2 + \sigma(t;T)dW_t^{R_T}, \end{aligned} \tag{B6}$$

where

$$\sigma(t;T) = \sqrt{\sigma_v^2 + \sigma_p^2(t;T) + 2\rho\sigma_v\sigma_p(t;T)} \tag{B7}$$

and $W_t^{R_T}$ is a one-dimensional standard Wiener process.[24]

For the process X_t defined in (B6), the distribution function in (B3) is given by[25]

$$G(t;T,X_0) = N\left[\frac{-X_0-\mu_g(t;T)}{\sqrt{\Sigma(t;T)}}\right] + e^{-2\mu_g(t;T)X_0/\Sigma(t;T)}N\left[\frac{-X_0+\mu_g(t;T)}{\sqrt{\Sigma(t;T)}}\right], \tag{B8}$$

24. No confusion should occur with the generic two-dimensional $W_t^{R_T}$ in (A14).

25. For a Brownian motion with constant drift μ and diffusion σ, the distribution function of its first-passage time to reach the origin from above ($X_0 > 0$) is given by $\Pr(\tau < t) = N[(-X_0 + \mu t)/(\sigma\sqrt{t})] + \exp(-2\mu X_0/\sigma^2)N[(-X_0 + \mu t)/(\sigma\sqrt{t})]$. See, e.g., eq. 4 in Leland and Toft (1996) or eq. 34b in Ingersoll (1987, 353). Note that there is a typo in eq. 34b of Ingersoll, where the minus sign in the second term should be the plus sign. The constant drift and diffusion solution provides a guidance for the solution in (B8) when the drift and diffusion are deterministic functions of time.

where

$$\Sigma(t;T) = \int_0^t \sigma^2(s;T)ds = \int_0^t [\sigma_v^2 + \sigma_p^2(t;T) + 2\rho\sigma_p(t;T)\sigma_v]ds = \sigma_v^2 t + \frac{\sigma_r^2}{\beta^2}[t + e^{-2\beta(T-t)}B_2(t) - 2e^{-\beta(T-t)}B_1(t)] + \frac{\sigma_r}{\beta}[t - e^{-\beta(T-t)}B_1(t)], \quad \text{(B9)}$$

$$\mu_g(t;T) = -\int_0^t \sigma^2(s;T)ds/2 = -\Sigma(t;T)/2, \quad \text{(B10)}$$

$B_1(t) = (1 - e^{-\beta t})/\beta$, and $B_2(t) = (1 - e^{-2\beta t})/(2\beta)$.

Proof. Given X_t in (B6), it is well known that the distribution function $G(t;T, X_0)$ satisfies the Kolmogorov backward (not the forward) equation,

$$-\frac{1}{2}\sigma^2(t;T)\frac{\partial G}{\partial X_0} + \frac{1}{2}\sigma^2(t;T)\frac{\partial^2 G}{\partial X_0^2} - \frac{\partial G}{\partial t} = 0, \quad \text{(B11)}$$

with boundary conditions $G(0;T,X_0) = 0$ for $X_0 > 0$ and $G(t;T,0) = 1$.

It is easy to check that (B8) satisfies the two boundary conditions. So we need to verify that (B8) satisfies (B11). To this end, let $n(x) = n(-x) = e^{-x^2/2}/\sqrt{2\pi}$ be the standard normal density function. It is easy to see that $n'(x) = -xn(x)$. Note that given the definition of μ_g and Σ, $-2\mu_g/\Sigma = 1$. Straightforward calculations yield

$$\frac{\partial G}{\partial X_0} = -n\left(\frac{X_0 - \Sigma/2}{\sqrt{\Sigma}}\right)\frac{1}{\sqrt{\Sigma}} - n\left(\frac{X_0 + \Sigma/2}{\sqrt{\Sigma}}\right)\frac{e^{X_0}}{\sqrt{\Sigma}} + e^{X_0}N\left(\frac{-X_0 - \Sigma/2}{\sqrt{\Sigma}}\right), \quad \text{(B12)}$$

$$\frac{\partial^2 G}{\partial X_0^2} = n\left(\frac{X_0 - \Sigma/2}{\sqrt{\Sigma}}\right)\frac{X_0 - \Sigma/2}{\Sigma^{3/2}} + n\left(\frac{X_0 + \Sigma/2}{\sqrt{\Sigma}}\right)\frac{(X_0 + \Sigma/2)e^{X_0}}{\Sigma^{3/2}} - 2n\left(\frac{X_0 + \Sigma/2}{\sqrt{\Sigma}}\right)\frac{e^{X_0}}{\sqrt{\Sigma}} + e^{X_0}N\left(\frac{-X_0 - \Sigma/2}{\sqrt{\Sigma}}\right), \quad \text{(B13)}$$

$$\frac{\partial G}{\partial t} = n\left(\frac{X_0 - \Sigma/2}{\sqrt{\Sigma}}\right)\left[\frac{X_0\sigma^2(t;T)}{2\Sigma^{3/2}} + \frac{\sigma^2(t;T)}{4\Sigma^{1/2}}\right] + e^{X_0}n\left(\frac{X_0 - \Sigma/2}{\sqrt{\Sigma}}\right)\left[\frac{X_0\sigma^2(t;T)}{2\Sigma^{3/2}} - \frac{\sigma^2(t;T)}{4\Sigma^{1/2}}\right]. \quad \text{(B14)}$$

Plugging (B12–B14) into (B11), we see that G satisfies (B11). Q.E.D.

B. Derivation of $H(t;T, X_0)$

To obtain $H(t;T, X_0)$ in closed form, we note that (B2) is already in the standard form of the Girsanov theorem. The corresponding $\varphi^T(t)$ is given by $\varphi^T(t) = (\sigma_v, 0)$. We can define another equivalent measure (to Q), R_T', such that $W_{vt}^{R_T'}$ and $W_{rt}^{R_T'}$, defined by

$$dW_{vt}^{R_T'} = dW_{vt}^{Q} - \sigma_v dt, \quad \text{(B15)}$$

$$dW_{rt}^{R_T'} = dW_{rt}^{Q} - \rho\sigma_v dt, \quad \text{(B16)}$$

are standard Wiener processes under R'_T, with the instantaneous correlation between $dW_{vt}^{R'_T}$ and $dW_{rt}^{R'_T}$ being given by ρdt. It can be seen that under R'_T,

$$dX_t = \sigma^2(t;T)dt/2 + \sigma(t;T)dW_t^{R'_T}, \tag{B17}$$

where $W_t^{R'_T}$ is a new standard Wiener process under R'_T.

Using the Girsanov theorem and the newly defined R'_T, we have

$$H(t;T,X_0) = E_0^{R'_T}[1(\tau > t)] = 1 - E_0^{R'_T}[1(\tau < t)]. \tag{B18}$$

Thus, $H(t;T,\ X_0)$ is the complementary distribution function of τ under R'_T. With X_t defined by (B17), $H(t;T,\ X_0)$ is given by

$$\begin{aligned} H(t;T,X_0) &= 1 - N\left[\frac{-X_0 - \mu_h(t;T)}{\sqrt{\Sigma(t;T)}}\right] - e^{-2\mu_h(t;T)X_0/\Sigma(t;T)}N\left[\frac{-X_0 + \mu_h(t;T)}{\sqrt{\Sigma(t;T)}}\right], \\ &= N\left[\frac{X_0 + \mu_h(t;T)}{\sqrt{\Sigma(t;T)}}\right] - e^{-2\mu_h(t;T)X_0/\Sigma(t;T)}N\left[\frac{-X_0 + \mu_h(t;T)}{\sqrt{\Sigma(t;T)}}\right], \end{aligned} \tag{B19}$$

where

$$\mu_h(t;T) = \int_0^t \sigma^2(s;T)/2 = \Sigma(t;T)/2. \tag{B20}$$

Proof. Let $\hat{H}(t;T,X_0) = E_0^{R'_T}[1(\tau < t)]$ be the distribution function of τ under R'_T. Comparing (B17) with (B6) reveals that dX_t under R'_T differs from dX_t under R_T only by the sign of its drift function. Therefore, by changing $\mu_g(t;T)$ to $-\mu_g(t;T)$ in (B8), it becomes $\hat{H}(t;T,X_0)$. That is, $\hat{H}(t;T,X_0)$ is given by

$$\hat{H}(t;T,X_0) = N\left[\frac{-X_0 - \mu_h(t;T)}{\sqrt{\Sigma(t;T)}}\right] + e^{-2\mu_h(t;T)X_0/\Sigma(t;T)}N\left[\frac{-X_0 + \mu_h(t;T)}{\sqrt{\Sigma(t;T)}}\right], \tag{B21}$$

where $\mu_h(t;T) = -\mu_g(t;T) = \Sigma(t;T)/2$. Now $1 - \hat{H}(t;T,X_0)$ yields (B19). Q.E.D.

C. Derivation of (14)

First, we note that for an arbitrary random variable $\tilde{x}$, the expectation, $E[\delta(\tilde{x} - x)]$, yields its density at x. To see this, let $f(t)$ be the density function of $\tilde{x}$. It follows then that $E[\delta(\tilde{x} - x)] = \int_{-\infty}^{\infty}\delta(t - x)f(t)dt = f(x)$. Second, we note that $(1 - \theta)V_B(r_s,\ s;\ P,\ T) = E_s^Q[e^{-\int_s^T r(u)du}]Pe^{y(T-s)}$.

Let PV_3 denote the present value of the third (last) term in (13). Using the risk-neutral valuation we have

$$\begin{aligned} PV_3 &= (1-\phi)P\int_0^T E_0^Q\{e^{-\int_0^s r(u)du}E_s^Q[e^{-\int_s^T r(u)du}]e^{y(T-s)}\delta(s-\tau)\}ds \\ &= (1-\phi)P\int_0^T E_0^Q[e^{-\int_0^T r(u)du}e^{y(T-s)}\delta(s-\tau)]ds, \\ &= (1-\phi)P\Lambda(r_0,0;T)\int_0^T E_0^Q\left[\frac{e^{-\int_0^T r(u)du}}{\Lambda(r_0,0;T)}\delta(s-\tau)\right]e^{y(T-s)}ds, \end{aligned} \tag{B22}$$

where the second equality follows from the law of iterated expectations, and thus the expectation operator E_s^Q drops out. Note that the third equality has the form of (A4). Using the new probability measure R_T, defined in Sec. B of this appendix, we can rewrite (B22) as

$$\begin{aligned} PV_3 &= (1-\phi)P\Lambda(r_0,0;T)\int_0^T E_0^{R_T}[\delta(s-\tau)]e^{y(T-s)}ds \\ &= (1-\phi)P\Lambda(r_0,0;T)\int_0^T g(s;T,X_0)e^{y(T-s)}ds \\ &= (1-\phi)P\Lambda(r_0,0;T)\int_0^T e^{y(T-s)}dG(s;T,X_0) \qquad \text{(B23)} \\ &= (1-\phi)P\Lambda(r_0,0;T)[G(T;T,X_0)+\hat{G}(T;T,X_0)], \end{aligned}$$

where $\hat{G}(T;T,X_0)$ is defined in (15), $g(s;T,X_0) = E_0^{R_T}[\delta(s-\tau)]$ is the density of τ under R_T, and $G(s;T,X_0)$, defined in (B3), is its distribution function.

To prove (14), using the risk-neutral pricing, we have the debt value given by

$$\begin{aligned} D(T,X_0;r_0,P) &= \int_0^T E_0^Q[e^{-\int_0^s r(u)du}d(s)]ds \\ &= C\int_0^T E_0^Q[e^{-\int_0^s r(u)du}1(s\le\tau)]ds + PE_0^Q[e^{-\int_0^T r(u)du}1(T\le\tau)] + PV_3 \\ &= C\int_0^T E_0^Q[e^{-\int_0^s r(u)du}1(s\le\tau)]ds + P\Lambda(r_0,0;T)E_0^{R_T}[1(T\le\tau)] + PV_3 \\ &= C\int_0^T E_0^Q[e^{-\int_0^s r(u)du}1(s\le\tau)]ds + P\Lambda(r_0,0;T)[1-G(T;T,X_0)] \\ &\quad + (1-\phi)P\Lambda(r_0,0;T)[G(T;T,X_0)+\hat{G}(T;T,X_0)]. \end{aligned}$$

(B24)

Appendix C

Further Discussion of the Scaling Property

In this appendix, we provide a detailed discussion of the scaling property used in Section III.C to calculate the total levered firm value with potentially an infinite number of periods. A key result is that the optimal solution X^* remains the same for all the bonds optimally issued by the firm.

Let τ_n be the first passage time when $V_t = V_B(y_t, t; P, T)$ in period n. Note that from

(18–20), the net benefit of debt (tax benefit of debt less the after-tax cost of default and transaction) from the debt issued at time zero can be rewritten as[26]

$$NB_0(T, X_0) \equiv tb(T, X_0) - [bc(T, X_0) + tc(T, X_0)](1-\theta) = V_0 f(T, X_{t_0}), \quad \text{(C1)}$$

where $f(T, X_{t_0})$ is given by

$$f(T, X_{t_0}) = e^{-X_{t_0} - yT}\theta((1-\theta)[(\lambda - 1) + \phi G(T; T, X_{t_0}) - (1-\phi)\hat{G}(T; T, X_{t_0})]$$
$$-(1-\theta)\{\phi[G(T; T, X_{t_0}) + \hat{G}(T; T, X_{t_0})] - \kappa\lambda(1-\theta)\}). \quad \text{(C2)}$$

Similarly, the net benefit of debt in period n is given by

$$NB_{t_{n-1}} = V_{t_{n-1}} f(T, X_{t_{n-1}}). \quad \text{(C3)}$$

Therefore, the total net benefit at $t_0 = 0, F_0$, is given by

$$F_0(T, X_0, X_T, \ldots) = \sum_{n=1}^{\infty} E_0^Q[e^{-\int_0^{t_{n-1}} r(s)ds} V_{t_{n-1}} f(T, X_{t_{n-1}}) \Pi_{i=1}^{n-1} 1(\tau_i > T)]$$
$$= V_0 \sum_{n=1}^{\infty} e^{-yt_{n-1}} E_0^Q[e^{-\sigma_v^2 t_{n-1}/2 + \sigma_v W_{vt_{n-1}}^Q} f(T, X_{t_{n-1}}) \Pi_{i=1}^{n-1} 1(\tau_i > T)], \quad \text{(C4)}$$

where $\Pi_{i=1}^{n-1} 1(\tau_i > T) = 1(\tau_1 > T)1(\tau_2 > T)\ldots 1(\tau_{n-1} > T)$ is the product of indicator functions and $\Pi_{i=1}^{n-1} 1(\tau_i > T) = 1$ if $n = 1$. Notice that the relation,

$$V_{t_n} = V_0 \exp\left\{\int_0^{t_n} [r(s) - y - \sigma_v^2/2]ds + \sigma_v W_{vT_n}^Q\right\}, \quad \text{(C5)}$$

has been used to obtain the second equality, and, as a result, the second equality no longer depends directly on the stochastic interest rate. In (C4), $X_{t_{n-1}}$ is the choice variable of the firm in period n.

Now consider the total net benefit at $t_1 = T, F_T$. It is easy to see that it is given by

$$F_T(T, X_0', X_T', \ldots) = V_T \sum_{m=1}^{\infty} e^{-yt_{m-1}} E_0^Q[e^{-\sigma_v^2 t_{m-1}/2 + \sigma_v W_{vt_{m-1}}^Q} f(T, X_{t_{m-1}}') \Pi_{i=1}^{m-1} 1(\tau_i > T)], \quad \text{(C6)}$$

where $X_{t_{m-1}}'$ is the choice variable of the firm in period m.[27]

Note that, besides the factors V_0 and V_T, (C4) and (C6) have the same functional form of the choice variables. That is, if we let

$$F_0(T, X_0, X_T, \ldots) = V_0 U(T, X_{t_i}|_{i=0,1,\ldots,\infty}), \quad \text{(C7)}$$

then we have

$$F_T(T, X_0', X_T', \ldots) = V_T U(T, X_{t_i}'|_{i=0,1,\ldots,\infty}) \quad \text{(C8)}$$

with the same function $U(\cdot)$. Therefore, if the optimal value of X_0 in (C4), which is the choice variable at time zero, is X^*, X^* must also be the optimal value for X_0' in (C6), which is the choice variable at time T. Note that at time zero the firm's objective is to maximize the total after-tax firm value given by $V_0(1-\theta) + F_0 = V_0[(1-\theta) +$

26. Remember that $t_n = nT$, where $n = 0, 1, 2, \ldots$, denotes the time of each restructuring point.

27. For notational purpose, we have renamed T as time zero. In other words, the first period starts at T when we compute the total net benefit of debts at T.

$U(T, X_{t_i}|_{i=0,1,\ldots,\infty})$]. Equivalently, the firm maximizes $U(T, X_{t_i}|_{i=0,1,\ldots,\infty})$, which is independent of the firm's initial value V_0. Likewise, the firm maximizes $U(T, X'_{t_i}|_{i=0,1,\ldots,\infty})$ at time T. Suppose that at time zero the optimal solution for the first period is given by X_0^*. At time T, the optimal solution for the first period, which is the second period viewed at time zero, must be the same as X_0^* or $X_0'^* = X_0^*$, because $U(T, X_{t_i}|_{i=0,1,\ldots,\infty})$ and $U(T, X'_{t_i}|_{i=0,1,\ldots,\infty})$ have the same form. In other words, the optimal solutions for X_0 and X_T are the same in the first two periods. Similarly, it can be seen that the optimal value for X_{t_i} in all periods must be the same.

Denote the (same) optimal solution by X^*.[28] Equations (C7) and (C8) and the proceeding discussion show that at the optimal $X_t = X^*$, the total net benefit at time zero, $F_0(T, X^*)$, and that at time T, $F_T(T, X^*)$, are related by $F_T(T, X^*) = F_0(T, X^*) * V_T/V_0$. This relation is used in Section III.C to obtain the optimal total net benefit in closed form.

References

Aït-Sahalia, Yacine. 1999. Transition densities for interest rate and other nonlinear diffusions. *Journal of Finance* 54:1361–95.

Barclay, Michael, and Clifford Smith. 1995. The maturity structure of corporate debt. *Journal of Finance* 50:609–31.

Black, Fischer, and John Cox. 1976. Valuing corporate securities: Some effects of bond indenture provisions. *Journal of Finance* 31:351–67.

Black, Fischer, and Myron Scholes. 1973. The pricing of options and corporate liabilities. *Journal of Political Economy* 81:637–54.

Brennan, Michael, and Eduardo Schwartz. 1978. Corporate income taxes, valuation, and the problem of optimal capital structure. *Journal of Business* 51:103–14.

Childs, Paul, David Mauer, and Steven Ott. 2003. Interaction of corporate financing and investment decision: The effects of agency conflicts. *Journal of Financial Economics* 76:667–90.

Cox, John, and Stephen Ross. 1976. The valuation of options for alternative stochastic processes. *Journal of Financial Economics* 3:145–66.

Darrough, Masako, and Neal Stoughton. 1986. Moral hazard and adverse selection: The question of financial structure. *Journal of Finance* 41:501–13.

Duffie, Darrell. 2001. *Dynamic asset pricing theory*. Princeton, NJ: Princeton University Press.

Duffie, Darrell, and Kenneth Singleton. 1999. Modelling term structures of defaultable bond. *Review of Financial Studies* 12:687–720.

Fan, Hua, and Suresh Sundaresan. 2000. Debt valuation, renegotiation, and optimal dividend policy. *Review of Financial Studies* 13:1057–99.

Fischer, Edwin, Robert Heinkel, and Josef Zechner. 1989. Dynamic capital structure choice: Theory and tests. *Journal of Finance* 44:19–40.

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

Harrison, Michael, and David Kreps. 1979. Martingales and arbitrage in multi-period securities markets. *Journal of Economic Theory* 20:381–408.

Hennessy, Christopher, and Toni Whited. 2004. Debt dynamics. *Journal of Finance* 60:1129–65.

IMSL Math Library. 2003. Visual numerics incorporated, http://www.vni.com.

Ingersoll, John. 1987. *Theory of financial decision making*. Totowa, NJ: Rowman & Littlefield.

Jarrow, Robert, and Stuart Turnbull. 1995. Pricing options on financial securities subject to default risk. *Journal of Finance* 50:53–86.

Jones, Philip, Scott Mason, and Eric Rosenfeld. 1984. Contingent-claims analysis of corporate capital structures: An empirical investigation. *Journal of Finance* 39:611–27.

Ju, Nengjiu, Robert Parrino, Allen Poteshman, and Michael Weisbach. 2005. Horses and rabbits?

28. A key point is that at times 0, T, $2T$, … , the firm faces an infinite time horizon, and the U functions take the same form. Hence, the optimal solution for the first period viewed at different times must be the same.

Trade-off theory and optimal capital structure. *Journal of Financial and Quantitative Analysis* 40:259–81.

Kane, Alex, Alan Marcus, and Robert McDonald. 1985. Debt policy and the rate of return premium to leverage. *Journal of Financial and Quantitative Analysis* 20:479–99.

Kim, Han. 1978. A mean-variance theory of optimal capital structure and corporate debt capacity. *Journal of Finance* 33:45–64.

Kim, In Joon, Krishna Ramaswamy, and Suresh Sundaresan. 1993. Valuation of corporate fix-income securities. *Financial Management* 22:117–31.

Leland, Hayne. 1994. Risky debt, bond covenants and optimal capital structure. *Journal of Finance* 49:1213–52.

———. 1998. Agency costs, risk management, and capital structure. *Journal of Finance* 53: 1213–43.

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

Longstaff, Francis, and Eduardo Schwartz. 1995. A simple approach to valuing risky fixed and floating rate debt. *Journal of Finance* 50:789–819.

Mello, Antonio, and John Parsons. 1992. Measuring the agency cost of debt. *Journal of Finance* 47:1887–1904.

Merton, Robert. 1974. On the pricing of corporate debt: The risk structure of interest rates. *Journal of Finance* 29:449–69.

Miao, Jianjun. 2005. Optimal capital structure and industry dynamics in stationary equilibrium. *Journal of Finance* 60:2621–59.

Morellec, Erwan. 2004. Can managerial discretion explain observed leverage ratios? *Review of Financial Studies* 17:257–94.

Press, William, Brian Flannery, Saul Teukolsky, and William Vetterling. 2003. *Numerical recipes*. New York: Cambridge University Press.

Stohs, Mark, and David Mauer. 1996. The determinants of corporate debt maturity structure. *Journal of Business* 69:279–312.

Tang, Yongjun, and Hong Yan. 2004. Macroeconomic conditions and credit spread dynamics: A theoretical exploration. Working paper, University of Texas at Austin, Department of Finance.

Titman, Sheridan, and Sergey Tsyplakov. 2002. A dynamic model of optimal capital structure. Working paper, University of Texas at Austin and University of South Carolina, Department of Finance.

Vasicek, Oldrich. 1977. An equilibrium characterization of the term structure. *Journal of Financial Economics* 5:177–88.