Investment with Leverage*

Andrew B. Abel
Wharton School of the University of Pennsylvania
National Bureau of Economic Research

October 30, 2016

Abstract

I examine the capital investment and leverage decisions of a firm. Optimal leverage depends on the interest tax shield and the cost of exposure to default, subject to an endogenous borrowing constraint. When the borrowing constraint is not binding, the tradeoff theory is operative; in that case, the market leverage ratio is a declining function of profitability, consistent with empirical findings. Bond financing increases q and investment, but given q, optimal investment and optimal leverage are independent. A novel expression for marginal q includes the expected present value of interest tax shields.

*I thank Amora Elsaify, Vito Gala, Itay Goldstein, Joao Gomes, Zhiguo He, Gregor Matvos, Stavros Panageas, Michael Roberts, Lawrence Summers, Fabrice Tourre, and Eric Zwick for helpful discussion. Part of this research was conducted while I was a Visiting Scholar at ChicagoBooth and I thank colleagues there for their hospitality and support. I also thank participants in seminars at the ChicagoBooth Finance Department and the Wharton Finance Department, and the Macro Finance Society.

How does a firm's financing decision interact with its capital investment decision? In a Modigliani-Miller environment, of course, a firm's debt-equity mix is irrelevant to the firm's value and to its capital investment. Therefore, to examine optimal leverage and its interaction with capital investment, I depart from the MM environment. Specifically, I introduce a tax on the firm's income, net of interest and capital investment costs, and a deadweight cost of default, which are the two main elements of the tradeoff theory of debt. I introduce these elements in the context of a fully-specified stochastic environment in which the firm borrows from lenders at an interest rate that compensates lenders for their expected losses in the event of default. Capital investment incurs convex costs of adjustment of the sort that underlie the q theory of investment. I specify the costs of adjustment to be linearly homogeneous in investment and the capital stock, so that average q and marginal q will be identically equal for a competitive firm with constant returns to scale. Therefore, the capital investment problem is similar to that underlying much of empirical literature on investment.

In the tradeoff theory of debt, the marginal benefit of an additional dollar of bonds is the interest tax shield associated with this additional dollar, and the marginal cost of an additional dollar of bonds is the increase in the expected costs associated with default. If the tradeoff theory is operative, the optimal amount of bonds equates the marginal benefit of the interest tax shield and the marginal cost of the additional exposure to default. Although the model in this paper includes these elements of the tradeoff theory, the tradeoff theory is not always operative because lenders will not lend an amount to a firm that exceeds the firm's total value. This feature is modeled as an endogenous limit on the total amount of bonds that a firm can issue, and I will describe this limit as an endogenous borrowing constraint. This limit might also be described as a non-negativity constraint on the firm's equity value, which equals the total value of the firm minus its outstanding debt. Equivalently, this limit on borrowing may be described as embodying limited liability. In situations in which this borrowing constraint strictly binds, the marginal benefit of debt *exceeds* the marginal cost of debt, so the tradeoff theory is not operative. Alternatively, in situations in which the borrowing constraint does not bind, the tradeoff theory is operative.

In addition to the features of the model described above, there are a few features that differ from standard theoretical models of debt issued by firms. The model is cast in continuous time, but instead of using a diffusion process to model underlying profitability, as in Leland (1994), I specify profitability to follow a regime-switching process. During a regime, profitability, denoted ϕ, remains fixed. A new regime arrives at a random date governed by a Poisson process. When a new regime arrives, the value of ϕ jumps to a new value, which can be higher or lower. A downward jump in ϕ may induce the firm to default, so default can occur along a path of realizations of ϕ. For all but Section 7 of the paper, I assume that the drawings of ϕ are i.i.d. across regimes, which simplifies the calculation of optimal debt. If the borrowing constraint were never binding at any ϕ, the assumption of i.i.d. realizations of ϕ across regimes would lead to a constant amount of the bond/capital ratio. However, the borrowing constraint may bind for intervals of ϕ, and the optimal bond/capital ratio will vary with ϕ within these intervals.

The other major departure from the existing literature concerns the maturity of the bonds issued by the firm. Unlike the bonds with infinite maturity in Leland (1994), I examine bonds with essentially zero maturity. If the firm does not default on these bonds, it repays them a tiny interval of time, dt, after issuing them. Of course, the firm can choose to issue new bonds and essentially roll over its bonds, provided lenders are willing to buy the new bonds. The instantaneous maturity of bonds provides two advantages as a modeling choice. First, the short maturity with the possibility of rollover emphasizes the dynamic nature of the decision to issue bonds. Firms need to revisit the financing decision, and this model makes the frequency of bond issuance arbitrarily high. The other advantage is that bonds with zero maturity always are valued at par, which alleviates the need to value bonds when the firm compares the value of its outstanding bonds to the total value of the firm as an ongoing enterprise. The firm makes this comparison at every point in time in deciding whether to repay its outstanding bonds or instead to default on them.

Another modeling simplification, which turns out to be purely expositional, is that I assume that in the event of default, the firm loses all value. That is, in default, the firm and its creditors all receive zero; equivalently, the deadweight cost of default is the total value of the firm. In Section 6, I extend the model to allow the deadweight cost of default to be a fraction α of the value of the firm, and let α be anywhere in $(0, 1]$. In that case, lenders recover a fraction $1 - \alpha$ of the firm's value at the time of default. I show that all of the major results in the paper continue *mutatis mutandis* to hold with $\alpha < 1$, and this is the sense in which the simplification of setting $\alpha = 1$ is purely expositional.

Five major findings emerge from the analysis in this paper. First, the market leverage ratio, which is the ratio of bonds to total firm value, is a decreasing function of contemporaneous profitabilty when and only when the tradeoff theory is operative. This finding is important because much of the existing theoretical literature on the tradeoff theory states that market leverage should be an increasing function of profitability, and yet empirical studies find a negative relationship between profitability and market leverage. The negative relationship derived here is consistent with empirical findings and extends the theoretical results in Abel (2015) to a more general stochastic framework that admits a richer variety of borrowing behavior.

Second, I show that the ability to issue bonds, and hence take advantage of the interest tax shield, increases the value of the firm and thus increases both average q and marginal q relative to the the case of an otherwise-identical all-equity firm that cannot issue bonds. Furthermore, because the investment-capital ratio is an increasing function of q, as is typical in neoclassical adjustment cost models of investment, the ability to issue bonds increases the optimal amount of capital investment.

Third, although bond finance increases the optimal amount of investment, there is an important sense in which the financing decision and the capital investment decision are independent of each other. Given the value function, which is an endogenously determined function, the optimal amount of bonds is independent of the amount of capital investment. And given the value of the firm, which is endogenous, optimal investment is independent of the financing decision. Equivalently, given the value of average q (which is identically equal to marginal q), optimal investment is independent of financing considerations. Therefore, a regression of the investment-capital ratio on q that correctly specifies the form of the inverse of the marginal adjustment cost function cannot detect the presence or absence of bonds; nor can it detect whether the endogenous borrowing constraint is binding.

Fourth, the paper derives a novel expression for marginal q. Marginal q can be calculated as the expected present value of the marginal profits accruing to the remaining undepreciated portion of a unit of capital over the indefinite future. With the commonly-used form of the adjustment cost function adopted here, the marginal profit of a unit of capital is typically calculated as the sum of (1) the marginal operating profit of capital, and (2) the reduction in the adjustment cost of a given rate of investment that is made possible by an additional unit of capital (which is the negative of the partial derivative of the adjustment cost function with respect to the capital stock). In this paper, I show that bond finance introduces a third, additive, term to the marginal profit of capital. This third term is the interest tax shield associated with the increase in the optimal amount of bonds resulting from a unit increase in the capital stock. This term is easily measured as the product of the tax rate and total interest payments divided by the capital stock.

Fifth, if capital expenditures, including adjustment costs, can be completely and immediately expensed, optimal investment is invariant to the tax rate for an all-equity firm. Under complete and immediate expensing, an increase in the tax rate increases optimal investment for a firm that issues bonds and takes advantage of the interest tax shield. Optimal investment for a firm that issues bonds will exceed optimal investment for an otherwise-identical firm that cannot issue bonds, provided that the tax rate is positive.

In Section 1, I specify the model of the firm including the stochastic environment governing operating profit, the adjustment cost function, and the opportunity to issue bonds at an interest rate that includes compensation to lenders for the ex ante cost of default. To solve the firm's decision problem, I use the Bellman equation in Section 2 to re-frame the firm's continuous-time decision problem as a pseudo-discrete-time problem. In Section 3, I analyze the optimal level of bonds and the optimal market leverage ratio, and in Section 4, I analyze the firm's optimal capital investment and introduce the novel formulation of marginal q in the presence of borrowing. Section 5 discusses the relationship between the financing decision and the capital investment decision, showing the extent to which these decisions are related to each other and also the sense in which, given the value function, they are independent of each other. The remaining two substantive sections are extensions to the basic model described above. In Section 6, I relax the assumption that the deadweight cost of default equals the total value of the firm. Instead, creditors recover a fraction $1-\alpha$ of the value of the firm in default, where the deadweight cost of default is a fraction α of the value of the firm. The major results of the paper continue to hold for any α in $(0,1]$. In Section 7, I explore the implications of positive serial correlation in the marginal operating profit of capital across successive regimes. Section 8 concludes. The appendices contain the proofs of all lemmas, propositions, and corollaries, as well as derivations that would interrupt the flow of the main text.

1 Model of the Firm

Consider a firm that is owned by risk-neutral shareholders who have a rate of time preference $\rho > 0$. Gross operating profit, that is, operating profit before taking account of capital expenditures and adjustment costs, at time t is $\phi_t K_t$, where K_t is the capital stock and ϕ_t is the profitability of capital at time t. This specification of profit can be derived easily, for example, for a competitive firm that uses capital and variable factors of production to produce output with a production function that has constant returns to scale. Profitability ϕ evolves over time according to a regime-switching process. A regime is a continuous interval of time during which profitablity ϕ remains constant. Regime changes, that is, changes in ϕ, arrive according to a Poisson process where λdt is the probability of a regime change over an infinitesimal interval of time dt.

I assume that the value of profitability, ϕ, is i.i.d. across regimes. Specifically, each new value of ϕ is an independent draw from an invariant unconditional distribution $F(\phi)$, which has finite support $[\phi_{\min}, \phi_{\max}]$, $\phi_{\max} > 0$. The distribution function $F(\phi)$ is differentiable with continuous density $f(\phi) \equiv F'(\phi) < \infty$ and has expected value $E\{\phi\} > 0$.[1] Despite the assumption that ϕ is i.i.d. across regimes, profitability displays persistence. The level of profitability, ϕ, remains constant during each regime, and regimes have a mean duration of $\frac{1}{\lambda}$, which could potentially be quite large. The unconditional correlation of ϕ_t and ϕ_{t+x}, for $x > 0$, is $e^{-\lambda x}$, which is positive and declines monotonically in x. In Section 7, I explore the implications of allowing for positive serial correlation in ϕ *across* regimes.

The capital stock changes over time as a result of capital investment by the firm and physical depreciation of capital. Specifically,

$$\dot{K}_t = I_t - \delta K_t, \tag{1}$$

where I_t is gross investment at time t, and δ is the constant rate of physical depreciation. The cost of undertaking gross investment at rate I_t consists of two components. The first component is the expenditure by the firm to purchase new uninstalled capital at a constant price $p_K > 0$ per unit of capital. I assume that the firm can sell capital, also at

[1] In addition, $\phi_{\min}$ can be negative, but, if it is negative, it must be small enough in absolute value so that an all-equity firm would never choose to cease operation. A sufficient condition on $\phi_{\min}$ is provided in equation (4).

a price p_K per unit of capital. Therefore, expenditure on new uninstalled capital is $p_K I_t$, which will be positive, negative, or zero, depending on whether the firm buys capital, sells capital, or undertakes zero gross investment. The second component of investment cost is an adjustment cost $c(\gamma_t) K_t$, where $\gamma_t \equiv \frac{I_t}{K_t}$ is the investment-capital ratio at time t, and $c(\gamma_t)$ is twice differentiable, strictly convex, and attains its minimum value of zero at $\gamma_t = 0$. An important feature of this formulation of the adjustment cost function, $c(\gamma_t) K_t$, is that it is linearly homogeneous in I_t and K_t. To ensure that the growth rate of the optimal capital stock, $\gamma_t - \delta$, is less than the discount rate, I assume that $\lim_{\gamma \nearrow \rho+\delta} c'(\gamma) = \infty$.[2] This assumption ensures that the value of the firm is finite. The total cost of investment comprises the purchase cost of uninstalled capital and the cost of adjustment,

$$p_K I_t + K_t c(\gamma_t) = [p_K \gamma_t + c(\gamma_t)] K_t. \tag{2}$$

The firm can borrow from risk-neutral lenders who have the same rate of time preference, ρ, as the shareholders. The firm borrows by issuing short-term bonds. Indeed, I assume that the maturity of bonds is so short as to be instantanenous. That is, the firm borrows an amount B_t at time t, pays interest $r_t B_t dt$, and then repays B_t to the lenders at time $t + dt$, where dt is an infinitesimal interval of time. Of course, at time $t + dt$ the firm can issue new debt, thereby rolling over its debt. If lenders knew with certainty that the firm would repay its bonds, then the interest rate on bonds, r_t, would equal the common rate of time preference, ρ. However, I allow for the possibility that the firm may default on its debt. Therefore, lenders require a default premium $r_t - \rho > 0$. I assume that lenders receive nothing in the event of default,[3] so the default premium is p_t, where $p_t dt$ is the probability of default over the infinitesimal interval of time from t to $t + dt$. Thus, the interest rate on the firm's bonds is $\rho + p_t$. In Section 6, I extend the model to allow creditors to receive a fraction of the firm's value in default, and I show that the results of this paper continue to

[2] This assumption rules out quadratic adjustment costs. A weaker assumption, which does not rule out quadratic adjustment costs, can be expressed in terms of $v(\phi)$ defined in equation (16): $(1-\tau)[p_K + c'(\rho + \delta - \varepsilon)] > v(\phi_{\max})$ for some $\varepsilon > 0$. If this condition is satisfied, the first-order condition in equation (24), along with the strict convexity of $c()$ and Proposition 1, which states that $v(\phi)$ is strictly increasing in ϕ, implies that optimal $\gamma_t < \rho + \delta$ for all ϕ_t.

[3] The assumption that lenders receive nothing in the event of default is extreme but it has two expositional advantages. First, it keeps the analysis easily tractable. Second, it rules out any collateral value for capital, and thus makes clear that the result (derived later) that optimal borrowing is proportional to the capital stock has nothing to do with any collateral value associated with capital.

hold *mutatis mutandi*.

I assume that the firm faces a constant tax rate τ on its taxable income. In calculating taxable income, the firm is permitted to deduct the total cost of investment in equation (2). That is, the firm is allowed to use immediate and full expensing of capital investment costs. Therefore, taxable income is gross operating profit, $\phi_t K_t$, less the total cost of investment, $[p_K \gamma_t + c(\gamma_t)] K_t$, and less interest payments, $(\rho + p_t) B_t$. If taxable income is negative, the firm receives a rebate from the government equal to the tax rate multiplied by the absolute value of taxable income.

Lenders recognize that if they are willing to buy the firm's bonds in an amount greater than the value of the firm, the firm would issue that large amount of bonds and then default immediately. To avoid that outcome, lenders impose an endogenous limit on the amount of bonds they are willing to buy. I will refer to this limit as a borrowing constraint, but it could also be described as a non-negativity constraint on equity value, or as limited liability. I will specify this borrowing constraint formally using the value function.

Define $V(\phi_t, K_t)$ as the value of a firm that arrives at time t with no bonds outstanding, so

$$V(\phi_t, K_t) = \max_{B_s \leq V(\phi_s, K_s), \gamma_s} E_t \left\{ \begin{array}{c} \int_t^T (1-\tau) [\phi_s K_s - [p_K \gamma_s + c(\gamma_s)] K_s - (\rho + p_s) B_s] e^{-\rho(s-t)} ds \\ + \int_t^T e^{-\rho(s-t)} dB_s \end{array} \right\}, \tag{3}$$

where T is the endogenous date at which the firm chooses to default and the constraint $B_s \leq V(\phi_s, K_s)$ formalizes the borrowing constraint just described. The first term on the right hand side of equation (3) is the expected present value of $(1-\tau)(\phi_s K_s - [p_K \gamma_s + c(\gamma_s)] K_s - (\rho + p_s) B_s)$, the after-tax cash flow from operations net of investment costs and interest costs from time t until the date of default, T. The second term is the expected present value of future cash flows to the firm generated by net issuances of bonds. In this term, dB_s is the net inflow of funds to the firm at date s when the firm changes the amount of its bonds outstanding. To ensure that $V(\phi_t, K_t) > 0$ for all

$\phi_t \in [\phi_{\min}, \phi_{\max}]$ and $K_t > 0$, I assume that[4]

$$\phi_{\min} > -E\{\phi\} \frac{\lambda}{\rho + \delta}. \tag{4}$$

The optimal net issuance of bonds at any date can either be a flow or a discrete change in the amount of bonds outstanding, depending on whether the regime changes at that date. To represent these differences, it is helpful to define t_j to be the date of the $j - th$ arrival of a new regime after the current date, t. When convenient, I use the notational convention $t_0 = t$ to denote the current date. With this notation, if the interval $[t, T)$ contains n regimes, the $j - th$ regime prevails during $[t_{j-1}, t_j)$, $j = 1, ..., n$, where $t_n = T$. During the $j - th$ regime, profitability, ϕ_s, is constant and equal to $\phi_{t_{j-1}}$.

Define $b_t \equiv \frac{B_t}{K_t}$ as the ratio of the firm's bonds to its capital stock at time t. Because of the Markovian nature of the stochastic environment and the linear homogeneity of the firm's optimization problem, the optimal values of b_s and γ_s and the value of the default probability p_s depend only on the contemporaneous value of profitability, ϕ_s, and hence are constant over time within a regime. I will use $b(\phi_s)$, $\gamma(\phi_s)$, and $p(\phi_s)$ to denote the values of b_s, γ_s, and p_s, respectively, that prevail when profitability is ϕ_s and the firm chooses γ_s and b_s optimally. Using this notation, the optimal value of outstanding bonds is

$$B_s = b(\phi_s) K_s, \text{ for } s \geq t = t_0. \tag{5}$$

Therefore,

$$dB_t = B_t = b(\phi_t) K_t, \text{ for } t = t_0 \tag{6}$$

$$dB_{t_j} = \left[b\left(\phi_{t_j}\right) - b\left(\phi_{t_{j-1}}\right) \right] K_{t_j}, \text{ for } j = 1, 2, 3... \tag{7}$$

and

$$dB_s = \left(\gamma_{t_j} - \delta\right) b\left(\phi_{t_j}\right) K_s ds, \text{ for } s \in (t_j, t_{j+1}), \ j = 0, 1, 2, ... \tag{8}$$

Equation (6), which is the issuance of bonds at the current time t_0, reflects the assumption

[4] Since it is feasible for the firm to set $B_s = 0 = \gamma_s$ and $dB_s = 0$ for all $s \geq t$, and thereby never default, $V(\phi_t, K_t) \geq (1 - \tau) E_t \left\{ \int_t^\infty \phi_s K_s e^{-\rho(s-t)} ds \right\}$. With $\gamma_s = 0$ for all $s \geq t$, $K_s = e^{-\delta(s-t))} K_t$ and it is straightforward to show that $V(\phi_t, K_t) \geq \frac{1-\tau}{\rho+\delta+\lambda} \left[\phi_t + E\{\phi\} \frac{\lambda}{\rho+\delta} \right] K_t > 0$, where the final inequality follows from equation (4) and $K_t > 0$.

that the firm has no bonds outstanding when time t_0 arrives. Therefore, the firm's net issuance of bonds equals its gross issuance of bonds at time t_0, B_{t_0}, because it does not have any outstanding bonds to repay. Equation (7) shows that on dates t_j at which at which subsequent new regimes arrive, the amount of bonds can change by a discrete amount because $b(\phi)$ can change by a discrete amount when a new value of ϕ arrives. Equation (8) shows the flow of new bonds issued when the regime is unchanged so that the ratio $b(\phi_s)$ is constant. To the extent that the size of the capital stock, K_s, changes during a regime, it changes continuously over time and the amount of bonds changes continuously in proportion to K_s to maintain $\frac{B_s}{K_s}$ unchanged during the regime.

The value function in equation (3) can be written as

$$V(\phi_t, K_t) = \max_{B_s \leq V(\phi_s, K_s), \gamma_s} E_t \left\{ \int_t^T e^{-\rho(s-t)} dD_s \right\}, \tag{9}$$

where

$$dD_s \equiv (1-\tau)\left[\phi_s K_s - \left[p_K \gamma_s + c(\gamma_s)\right] K_s - (\rho + p_s) B_s\right] ds + dB_s \tag{10}$$

is amount of dividends at time s, if the firm does not retain any earnings. The amount of dividends can be negative at a point of time, if either taxable income, $\left[\phi_s K_s - \left[p_K \gamma_s + c(\gamma_s)\right] K_s - (\rho + p_s) B_s\right] ds$, is negative or if the firm is reducing its outstanding debt so $dB_s < 0$. When dividends are negative, existing shareholders inject funds into the firm. I assume that these shareholders have deep pockets, that is, they have unlimited capacity to inject funds into the firm as they see fit.

1.1 The Timing of Events

In this subsection, I summarize the timing of events. To describe this timing around time t, it is useful to define $B_{t^-} \equiv \lim_{dt \searrow 0} B_{t-dt}$ as the amount of bonds issued by the firm an instant before time t.

1. The firm arrives at time t with B_{t^-} bonds outstanding.

2. At time t, the value of profitability, ϕ_t, and hence the value of the firm $V(\phi_t, K_t)$ are realized and observed by the firm and by lenders.

3. The firm decides whether to repay or default on its outstanding bonds.

(a) If $V(\phi_t, K_t) < B_{t^-}$, the firm defaults on its bonds; creditors and shareholders receive nothing.

(b) If $V(\phi_t, K_t) \geq B_{t^-}$, the firm repays its bonds, and the shareholders retain ownership.

4. If shareholders retain ownership, the firm issues bonds, B_t, subject to the borrowing constraint $B_t \leq V(\phi_t, K_t)$.

5. The firm receives $\phi_t K_t$, undertakes capital investment $I_t = \gamma_t K_t$, pays interest $r_t B_t$, and pays taxes $\tau(\phi_t K_t - [p_K \gamma_t + c(\gamma_t)] K_t - r_t B_t)$.

6. The firm pays dividends $dD_t = (1-\tau)(\phi_t K_t - [p_K \gamma_t + c(\gamma_t)] K_t - r_t B_t) + dB_t$, where $dB_t = B_t - B_{t^-}$.

2 Bellman Equation

To solve for the optimal amounts of bonds and investment, it is convenient to express the firm's continuous-time problem as a pseudo-discrete-time problem. I use the adjective "pseudo" because the intervals of the time are of random length; the intervals of time correspond to regimes, during which ϕ_s, and hence γ_s, b_s, and p_s, are constant. Thus, for instance, if the current regime prevails continuously until time t_1, when a new regime arrives, the capital stock at time $s \in [t, t_1]$ is $e^{(\gamma_t - \delta)(s-t)} K_t$ and taxable income at time $s \in [t, t_1)$ is $(\phi_t - [p_K \gamma_t + c(\gamma_t)] - (\rho + p_t) b_t) e^{(\gamma_t - \delta)(s-t)} K_t$. Therefore, the present value of taxable income from time t to time t_1 is

$$\int_t^{t_1} \left(\begin{array}{c} \phi_s K_s - [p_K \gamma_s + c(\gamma_s)] K_s \\ -(\rho + p_s) B_s \end{array} \right) e^{-\rho(s-t)} ds = \left(\begin{array}{c} \phi_t - [p_K \gamma_t + c(\gamma_t)] \\ -(\rho + p_t) b_t \end{array} \right) H(t, t_1) K_t, \quad (11)$$

where

$$H(t, t_1) \equiv \int_t^{t_1} e^{-(\rho + \delta - \gamma_t)(s-t)} ds = \frac{1 - e^{-(\rho + \delta - \gamma_t)(t_1 - t)}}{\rho + \delta - \gamma_t}. \quad (12)$$

The present value of the net inflow of funds raised by issuing bonds during the interval of time $[t, t_1)$, $\int_t^{t_1} e^{-\rho(s-t)} dB_s$, is, using equations (6) and (8) and $K_s = e^{(\gamma_t - \delta)(s-t)} K_t$,

$$\int_t^{t_1} e^{-\rho(s-t)} dB_s = [1 + (\gamma_t - \delta) H(t, t_1)] b_t K_t. \tag{13}$$

These funds consist of a discrete inflow of funds at time t when the firm issues $B_t = b(\phi_t) K_t$, followed by a net flow of funds as the firm adjusts its outstanding level of bonds continuously over the time interval (t, t_1) to maintain $B_s = b(\phi_t) K_s$.

Use equations (11) and (13) to rewrite the value function in equation (3) as the following Bellman equation

$$V(\phi_t, K_t) = \max_{b_t \leq \frac{V(\phi_t, K_t)}{K_t}, \gamma_t} E_t \left\{ \begin{array}{c} (1-\tau)\{\phi_t - [p_K \gamma_t + c(\gamma_t)] - (\rho + p_t) b_t\} H(t, t_1) K_t \\ + [1 + (\gamma_t - \delta) H(t, t_1)] b_t K_t \\ + e^{-\rho(t_1 - t)} \max\left[V\left(\phi_{t_1}, K_{t_1}\right) - B_{t_1^-}, 0\right] \end{array} \right\}, \tag{14}$$

where $B_{t_1^-}$ is the amount of bonds outstanding when a new regime arrives at time t_1. The first two lines on the right hand side of equation (14) are the expected present value of net inflows of cash over the duration of the current regime from time t to time t_1. In particular, as explained above, the first line is the expected present value of the after-tax operating profit less total investment costs and less interest payments over the interval of time after t until the next regime change. The second line is the expected present value of funds obtained by net bond issuance over the interval of time until the next regime change, which occurs at time t_1. The third line on the right hand side of equation (14) is the expected present value of the continuation value of the firm when the next regime arrives. If the firm chooses to repay its outstanding debt, $B_{t_1^-}$, at time t_1, the continuation value of the firm will be $V\left(\phi_{t_1}, K_{t_1}\right) - B_{t_1^-}$, which is the value of the firm if it had no outstanding debt less the value of its outstanding debt. Provided that $V\left(\phi_{t_1}, K_{t_1}\right) - B_{t_1^-} \geq 0$, the firm will repay its outstanding debt and continue operation. However, if $V\left(\phi_{t_1}, K_{t_1}\right) - B_{t_1^-} < 0$, then the firm will default and its continuation value would be zero.

The value function is a function of two state variables, ϕ_t and K_t. It turns out that the value function is linearly homogeneous in K_t and can be written as[5]

$$V(\phi_t, K_t) = v(\phi_t) K_t. \tag{15}$$

Straightforward, but tedious, calculation, which is relegated to Appendix B, shows that

$$v(\phi_t) = \max_{b_t \leq v(\phi_t), \gamma_t} \frac{(1-\tau)[\phi_t - [p_K \gamma_t + c(\gamma_t)]] + \lambda \overline{v} + A(b_t)}{\rho + \lambda + \delta - \gamma_t}, \tag{16}$$

where

$$\overline{v} \equiv \int v(\phi) \, dF(\phi) \tag{17}$$

is the unconditional expectation of $v(\phi)$ and

$$A(b_t) \equiv \tau \left[\rho + \lambda \int_{v(\phi)<b_t} dF(\phi)\right] b_t - \lambda \int_{v(\phi)<b_t} v(\phi) \, dF(\phi). \tag{18}$$

The function $A(b)$ defined in equation (18) contains the essential elements of the tradeoff theory of debt, and I will refer to $A(b_t)$ as the "tradeoff function." The first term on the right hand side of the tradeoff function in equation (18), $\tau\left[\rho + \lambda \int_{v(\phi)<b_t} dF(\phi)\right] b_t$, is the tax shield associated with the deductibility of interest payments. This interest tax shield is the product of the tax rate, τ, and interest payments by the firm, which are the product of the interest rate, $\rho + \lambda \int_{v(\phi)<b_t} dF(\phi)$, and the amount of debt b_t.[6] The second term, $\lambda \int_{v(\phi)<b_t} v(\phi) \, dF(\phi)$, is the expected value of the deadweight loss arising from default, which occurs when the new realization of ϕ in the next regime leads to a value of the firm that is smaller than the outstanding debt, b_t. Both terms are increasing in b_t. An increase in b_t increases the interest tax shield by increasing interest payments, both by increasing the interest rate and by increasing the amount of debt on which interest is paid. An increase in b_t also increases the expected cost of default by enlarging the set of values of ϕ in the next regime that would lead to default. In general, an increase in b_t can increase or decrease $A(b_t)$ depending on whether the resulting increase in the interest tax shield is larger or smaller

[5] The first two lines in the expectation on the right hand side of equation (14) are proportional to K_t, and, since $V\left(\phi_{t_1}, K_{t_1}\right)$ and $B_{t_1^-} = b\left(\phi_{t_1^-}\right) K_{t_1}$ are both proportional to K_{t_1}, the third line is proportional to $K_{t_1} = e^{(\gamma_t - \delta)(s-t)} K_t$.

[6] More precisely, the amount of debt is $b_t K_t$, so the interest tax shield in the first term on the right hand side of equation (18) is the interest tax shield per unit of capital.

than the resulting increase in the expected cost of default.

The following proposition is not surprising, but since subsequent discussion and proofs use the fact that $v(\phi)$ is strictly increasing, and hence invertible, it is useful to formalize and prove this property.

Proposition 1 *$v(\phi_t)$ is strictly increasing in ϕ_t.*

3 Optimal Leverage

The optimal value of b_t attains the maximum on the right hand side of equation (16) subject to the borrowing constraint $b_t \leq v(\phi_t)$. Since b_t enters the maximand in equation (16) only through $\frac{A(b_t)}{\rho+\lambda+\delta-\gamma_t}$, and since $\rho + \lambda + \delta - \gamma_t > 0$, the optimal value of b_t maximizes $A(b_t)$ over the domain $[0, v(\phi_t)]$. Without further restrictions on the distribution function $F(\phi)$, I cannot rule out the possibility that $A(b_t)$ has multiple local maxima. To resolve any multiplicities that may arise, I specify the optimal value of b_t as the smallest value of $b_t \in [0, v(\phi_t)]$ at which $A(b_t)$ is maximized. I use the notation $b(\phi_t)$ denote the optimal value of b_t so

$$b(\phi_t) = \min\left\{\arg\max_{b\in[0,v(\phi_t)]} A(b)\right\}. \tag{19}$$

Even without further restrictions on $F(\phi)$, the definition of the tradeoff function $A(b_t)$ in equation (18) and the fact that $v(\phi) > 0$ imply the following lemma, which presents properties of $A(b_t)$ that are useful in characterizing $b(\phi_t)$.

Lemma 2 *Assume that $\tau\rho > 0$, and define $A(b_t)$ as in equation (18). Then*

1. *$A(b_t) = \tau\rho b_t \geq 0$ for $0 \leq b_t \leq v(\phi_{\min})$,*
2. *$\arg\max_{b\in[0,v(\phi_t)]} A(b) \geq v(\phi_{\min}) > 0$.*

The properties of $A(b)$ in Lemma 2 are illustrated in Figure 1. The top panel of Figure 1 shows $A(b)$ for $b \in [0, v(\phi_{\max})]$. Inspection of the definition of $A(b_t)$ in equation (18) reveals that although $A(b)$ depends on the distribution of ϕ, it is independent of current profitability, ϕ_t. Therefore, the graph of $A(b)$ in the top panel of Figure 1 is identical for all values of ϕ_t in the support of $F(\phi)$. Statement 1 of Lemma 2 implies that $A(0) = 0$,

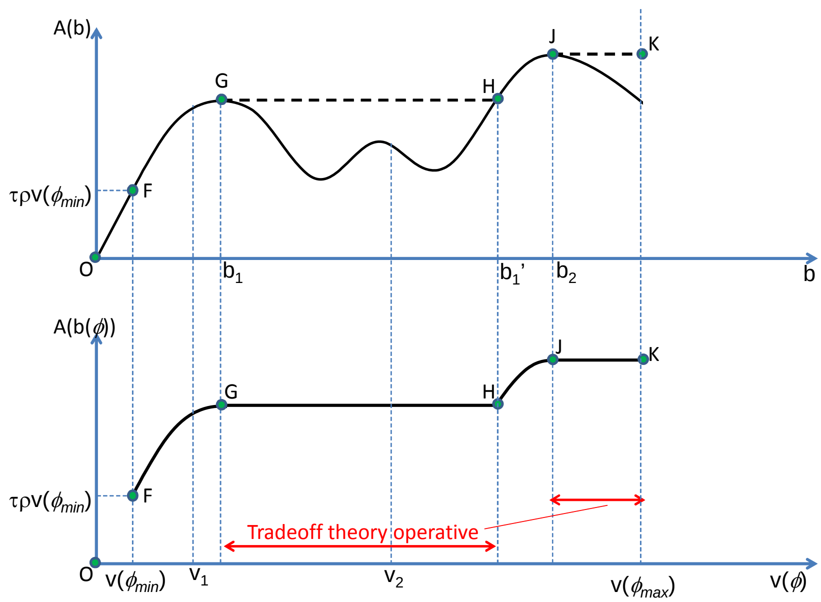

Figure 1:

represented by the origin O; $A(v(\phi_{\min})) = \tau\rho v(\phi_{\min})$, represented by point F; and the segment of $A(b)$ from point O to point F is linear with slope $\tau\rho > 0$. Statement 2 implies that the value of $b \in [0, v(\phi_t)]$ that maximizes $A(b)$ cannot be smaller than $v(\phi_{\min})$, and this result is illustrated by the ordinate at point F, which is higher than the ordinates of all points to the left of point F.

The bottom panel of Figure 1 helps take account of the borrowing constraint, $b_t \leq v(\phi_t)$, in the determination of the optimal amount of bonds. The horizontal axis in the bottom panel is $v(\phi_t)$, which is a strictly increasing function of ϕ_t. This axis is aligned directly below the horizontal axis in the top panel so that, for instance, the level of b at point G in the top panel is equal to the level of $v(\phi_t)$ at point G in the bottom panel. I will illustrate how to use Figure 1 to calculate optimal b at two different values of ϕ_t. First, suppose that that $v(\phi_t) = v_1$, as shown on the horizontal axis in the bottom panel. Then the optimal value of b must be less than or equal to v_1, that is, it must lie in $[0, v_1]$. As shown in the top panel, the highest value of $A(b)$ for b in the interval $[0, v_1]$ is attained at $b = v_1$, so the borrowing constraint is binding. Indeed, the borrowing constraint will bind for any value of $v(\phi_t)$ corresponding to the segment FG along the curve representing $A(b)$.

For the second illustrative example, suppose that $v(\phi_t) = v_2$ on the horizontal axis in the bottom panel, so the optimal value of b must lie in $[0, v_2]$. As shown in the top panel, the highest value of $A(b)$ for b in $[0, v_2]$ is attained at point G, where $b = b_1$, as shown on the horizontal axis in the top panel. In fact, for any value of $v(\phi_t)$ in the interval $[b_1, b_1']$, the highest value of $A(b)$ for b in $[0, v(\phi_t)]$ is attained at point G, where $b = b_1$. That is, the optimal amount of bonds is invariant to ϕ_t for any $v(\phi_t) \in [b_1, b_1']$. Moreover, the borrowing constraint is not binding for any $v(\phi_t) \in [b_1, b_1')$ so that at the optimal amount of bonds, $b(\phi_t)$, $A'(b_t) = A'(b(\phi_t)) = 0$; therefore, as I discuss in Subsection 3.1, the tradeoff theory is operative for any $v(\phi_t) \in [b_1, b_1')$.

Proposition 3 *The optimal value of b_t represented by $b(\phi_t)$ in equation (19) has the following properties:*

1. $b(\phi_t) \geq v(\phi_{\min}) > 0$*;*
2. $A(b(\phi_t)) > 0$*;*
3. $b(\phi_t)$ *is weakly increasing in* ϕ_t*; and*
4. $A(b(\phi_t))$ *is weakly increasing in* ϕ_t*.*

Statement 1 of Proposition 3 is that the optimal amount of bonds is strictly positive. The firm will always borrow at least much as the minimum possible value of the firm, $v(\phi_{\min}) K_t$. If the firm were to borrow an amount smaller than $v(\phi_{\min}) K_t$, it would be in a position from which it could increase its interest tax shield by increasing its borrowing without exposing itself or its lenders to any probability of default. Statement 2 of Proposition 3 is that $A(b(\phi_t)) > 0$ so that the opportunity to issue bonds increases the value of the firm relative to a situation in which the firm could not issue bonds (since $A(0) = 0$ from Statement 1 of Lemma 2). Statement 3 of Proposition 3 states that $b(\phi_t)$ is weakly increasing in ϕ_t. Therefore, the amount of bonds outstanding at time t, $B_t = b(\phi_t) K_t$, is a weakly increasing function of contemporaneous profitability, ϕ_t, and proportional to the contemporaneous capital stock, K_t. Finally, Statement 4 is that $A(b(\phi_t))$ is weakly increasing in ϕ_t.

Proposition 3 describes the behavior of optimal bonds, $b(\phi_t) K_t$, as a function of ϕ_t. This proposition can be used to describe the behavior of the optimal amount of borrowing, $b(\phi_t) K_t$, as a function of time. The capital stock, K_t, is a continuous function of time and can increase, decrease, or remain unchanged over time, depending on whether net investment, $\gamma_t - \delta$, is positive, negative, or zero. The value of $b(\phi_t)$ changes only when the value of ϕ_t changes, that is, only when the regime changes. Although $b(\phi_t)$ can increase over time, it will never decrease over time for an ongoing firm; to do so would imply that the previous level of debt $b(\phi_{t^-}) K_t$, is no longer feasible, that is, $b(\phi_{t^-}) > v(\phi_t)$, so the firm would default. The facts that (1) $B_t = b(\phi_t) K_t$, (2) K_t is a continuous function of time; and (3) $b(\phi_t)$ cannot fall over time imply, as stated in Corollary 4 below, that optimal B_t will never fall by a discrete amount at a point of time.[7] The only reason that shareholders might inject a discrete amount of funds at a point of time would be to pay for a discrete reduction in the amount of bonds outstanding. Since optimal B_t never falls by a discrete amount at a point of time, shareholders will never inject a discrete amount of funds at a point of time.

Corollary 4 *An ongoing firm will never decrease its amount of debt by a discrete amount at a point of time. Therefore, shareholders will never inject a discrete amount of funds into an ongoing firm at a point of time.*

3.1 The Tradeoff Theory of Debt

The tradeoff theory of debt states that a firm's optimal level of debt reflects a tradeoff between the increased interest tax shield associated with an additional dollar of debt and the additional exposure to default and its consequent losses associated with an additional dollar of debt. Formally, the tradeoff theory is operative when $A'(b(\phi_t)) = 0$. Differentiating $A(b)$ in equation (18) with respect to b yields

$$A'(b) = \tau \left[\rho + \lambda F\left(v^{-1}(b)\right)\right] - (1-\tau)\lambda b v^{-1\prime}(b) f\left(v^{-1}(b)\right). \tag{20}$$

To calculate $v^{-1\prime}(b)$ in the second term on the right hand side of equation (20) when $A'(b) = 0$, first substitute the optimal values of γ_t and b_t, which are $\gamma(\phi_t)$ and $b(\phi_t)$, respectively, into the right side of equation (16), differentiate with respect to ϕ_t and apply

[7] Goldstein, Ju, and Leland (2001) assume that, except for default, the firm never reduces its debt. In the current paper, optimal debt will fall continuously over time during a regime if $\gamma_t < \delta$ so that K_t falls continuously over time. However, optimal debt $B_t = b(\phi_t) K_t$ will never fall by a discrete amount at a point of time, except for default.

the envelope theorem to obtain[8]

$$v'(\phi_t) = \frac{1-\tau}{\rho+\lambda+\delta-\gamma(\phi_t)}, \text{ when } A'(b(\phi_t)) = 0. \tag{21}$$

Substituting $\frac{\rho+\lambda+\delta-\gamma(\phi_t)}{1-\tau} = \frac{\rho+\lambda+\delta-\gamma\left(v^{-1}(b)\right)}{1-\tau}$ for $v^{-1\prime}(b)$ on the right hand side of equation (20) when $A'(b) = 0$ yields

$$A'(b) = \tau\left[\rho+\lambda F\left(v^{-1}(b)\right)\right] - \left[\rho+\lambda+\delta-\gamma\left(v^{-1}(b)\right)\right]\lambda b f\left(v^{-1}(b)\right) = 0. \tag{22}$$

The first term on the right hand side of equation (22) is the interest tax shield associated with an additional dollar of bonds. It is the product of the tax rate and the interest rate, $\rho+\lambda F\left(v^{-1}(b)\right)$. The second term reflects the cost associated with increased exposure to default associated with an additional dollar of bonds. When the marginal benefit of the increased interest tax shield equals the marginal cost associated with increased exposure to default, then $A'(b(\phi_t)) = 0$, and the tradeoff theory of debt is operative. The tradeoff theory is operative when the borrowing constraint, $b(\phi_t) \leq v(\phi_t)$ is not binding. When the borrowing constraint is binding, $A'(b(\phi_t)) > 0$, and the tradeoff theory is not operative.

Definition 5 *The tradeoff theory is strictly operative at ϕ_t if and only if $b(\phi_t) < v(\phi_t)$.*[9]

Proposition 6 *Let $\Phi \equiv [\phi_L, \phi_H]$ be a non-degenerate interval in the support of $F(\phi)$. For all $\phi_t \in \Phi$,*

1. *$b(\phi_t)$ is invariant to ϕ_t, if $b(\phi_t) < v(\phi_t)$ for all $\phi_t \in \Phi$.*

   (tradeoff theory strictly operative)

2. *$b(\phi_t)$ is strictly increasing in ϕ_t, if $b(\phi_t) = v(\phi_t)$ for all $\phi_t \in \Phi$.*

   (tradeoff theory not strictly operative)

[8] $v(\phi_t)$ is not differentiable at values of ϕ_t for which optimal amount of bonds, $b(\phi_t)$, is discontinuous, as at the value of ϕ corresponding to point H in Figure 1. But $v(\phi_t)$ is differentiable at values of ϕ_t for which $A'(b(\phi_t)) = 0$.

[9] If $A'(\phi_t) = 0$ and $b(\phi_t) = v(\phi_t)$, for instance, as at points G and J in Figure 1, the tradeoff theory is operative, but is not strictly operative in the sense of Definition 5.

It is possible for the tradeoff theory to be strictly operative over two disjoint intervals of ϕ_t separated by a nontrivial interval in which the tradeoff theory is not strictly operative. In such a situation, $b(\phi_t)$ is invariant to ϕ_t *within* each interval in which the tradeoff theory is strictly operative, but $b(\phi_t)$ will be higher in the interval with the higher values of ϕ_t than in the interval with the lower values of ϕ_t. For instance, in Figure 1, the tradeoff theory is strictly operative for all ϕ_t such that $v(\phi_t) \in [b_1, b_1']$, corresponding to the segment GH, and for all ϕ_t such that $v(\phi_t) \in [b_2, v(\phi_{\max})]$, corresponding to the segment JK, and the value of $b(\phi_t)$ is higher in the second interval than in the first interval.

3.2 Leverage Ratios

Two common measures of leverage are the book leverage ratio, which is the ratio of the firm's outstanding debt to the book value of its assets, and the market leverage ratio, which is the ratio of the value of the firm's outstanding debt to its total market value. In the model analyzed here, the book leverage ratio is the ratio of bonds outstanding, B_t, to the replacement cost of the firm's capital, $p_K K_t$, which is simply $\frac{b(\phi_t)}{p_K}$ because $b_t \equiv \frac{B_t}{K_t}$. Since p_K is constant in this model, book leverage, $\frac{b(\phi_t)}{p_K}$, inherits the properties of $b(\phi_t)$ in Propositions 3 and 6. Therefore, book leverage is positive for all ϕ_t. It is invariant to ϕ_t when the tradeoff theory is strictly operative; when the tradeoff theory is not operative, the book leverage ratio is strictly increasing in ϕ_t .

The market leverage ratio is $l(\phi_t) \equiv \frac{B_t}{V(\phi_t, K_t)} = \frac{b(\phi_t)}{v(\phi_t)} \leq 1$, where the weak inequality is simply a restatement of the borrowing constraint.

Corollary 7 *to Proposition 6. The market leverage ratio,* $l(\phi_t) \equiv \frac{b(\phi_t)}{v(\phi_t)}$, *is*

1. *strictly decreasing in* ϕ_t, *if* $b(\phi_t) < v(\phi_t)$ *in a neighborhood around* ϕ_t*;*

   (tradeoff theory is strictly operative)

2. *identically equal to one, if* $b(\phi_t) = v(\phi_t)$*.*

   (tradeoff theory is not strictly operative)

Corollary 7 states that the market leverage ratio is a decreasing function of profitability, ϕ_t, if and only if the tradeoff theory is strictly operative. This result is remarkable because the conventional wisdom is that theoretical models of the tradeoff theory predict that the market leverage ratio is an increasing function of profitability, but empirical studies find a negative relation between market leverage and profitability.[10] Thus, when the tradeoff theory is strictly operative in the model presented here, the relation between market leverage and profitability is consistent with the negative empirical relationship between market leverage and profitability.

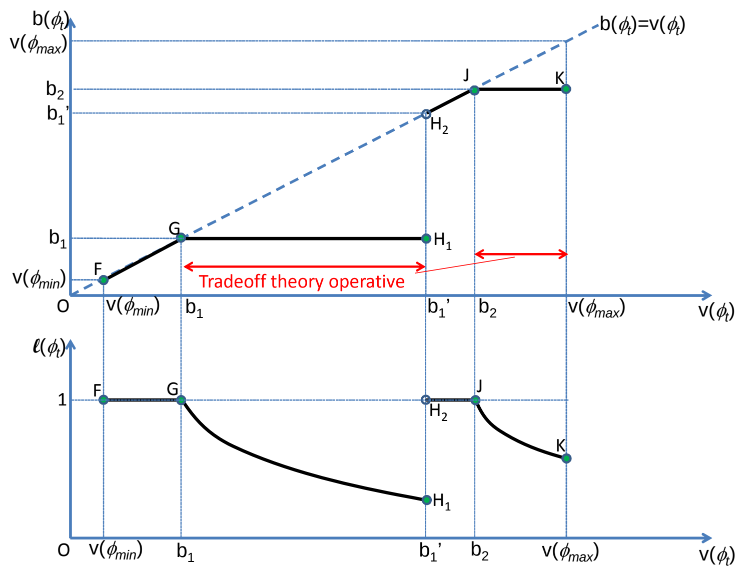

Figure 2:

Figure 2 illustrates the behavior of optimal bonds (in the top panel) and optimal market leverage (in the bottom panel) as functions of profitability. Rather than measuring profitability, ϕ_t, itself on the horizontal axis, Figure 2 measures $v(\phi_t)$, which is a monotonic transformation of ϕ_t, on the horizontal axis. The values of $v(\phi_{\min})$, b_1, b_1', b_2, and $v(\phi_{\max})$ in Figure 2 are the same as in Figure 1, and the points labelled O, F, G, J, and K correspond to the points in Figure 2 with those labels.[11] Figure 2 partitions the set of possible

[10] For instance, Myers (1993, p. 6) states "The most telling evidence against the static tradeoff theory is the strong inverse correlation between profitability and financial leverage.... Yet the static tradeoff story would predict just the opposite relationship. Higher profits mean more dollars for debt service and more taxable income to shield. They should mean higher target debt ratios."

[11] Since $b(\phi_t)$ is discontinuous at $v(\phi_t) = b_1'$, the points H_1 and H_2 in Figure 2 correspond to point H in Figure 1.

values of $v(\phi_t)$, $[v(\phi_{\min}), v(\phi_{\max})]$, into four intervals:[12] (1) If $v(\phi_t) \in [v(\phi_{\min}), b_1]$, then $\arg\max_{b\in[0,v(\phi_t)]} A(b) = v(\phi_t)$, so the optimal value of b_t, $b(\phi_t)$, equals $v(\phi_t)$, as shown in the top panel of Figure 2. Therefore, the tradeoff theory is not strictly operative in this interval. For values of ϕ_t that are low enough to lead to $v(\phi_t)$ in the interval $[v(\phi_{\min}), b_1]$, the firm borrows as much as lenders will lend. Since lenders will not lend an amount greater than $v(\phi_t) K_t$, the firm issues bonds in the amount $v(\phi_t) K_t$, so that $b(\phi_t) = v(\phi_t)$. Since the amount of bonds outstanding equals the value of the firm, the market leverage ratio is equal to one for all ϕ_t in this interval, as shown in the bottom panel of Figure 2. (2) If $v(\phi_t) \in (b_1, b_1']$, then $\min\left\{\arg\max_{b\in[0,v(\phi_t)]} A(b)\right\} = b_1$, so the optimal value of $b(\phi_t)$ equals $b_1 < v(\phi_t)$ and the tradeoff theory is strictly operative. Therefore, the optimal amount of bonds is invariant to ϕ_t for values of ϕ_t for which $v(\phi_t) \in (b_1, b_1']$, as shown in the top panel, and the market leverage ratio is less than one and is a decreasing function of ϕ_t, as shown in the bottom panel. (3) If $v(\phi_t) \in (b_1', b_2]$, then $\arg\max_{b\in[0,v(\phi_t)]} A(b) = v(\phi_t)$, so that $b(\phi_t) = v(\phi_t)$, as in the top panel. As in case (1), the tradeoff theory is not strictly operative and the optimal market leverage ratio equals one throughout this interval, as shown in the bottom panel. (4) If $v(\phi_t) \in (b_2, v(\phi_{\max})]$, then $\arg\max_{b\in[0,v(\phi_t)]} A(b) = b_2 < v(\phi_t)$ and the tradeoff theory is strictly operative. As in case (2), the optimal amount of bonds is invariant to ϕ_t in this interval (top panel), and the optimal market leverage ratio is a decreasing function of ϕ_t (bottom panel).

Figure 2 illustrates that the optimal value of $b(\phi_t)$ can be a discontinuous function of the state, ϕ_t. In particular, $b(\phi_t)$ has a discontinuity at $\phi_t = v^{-1}(b_1')$, equivalently, $v(\phi_t) = b_1'$, which corresponds to points H_1 and H_2 in Figure 2.[13]

Not only is $b(\phi_t)$ potentially discontinuous in the state, ϕ_t, it is a discontinuous function of time. Specifically, the optimal value of debt changes by a discrete amount when ϕ_t changes, that is, when the regime changes, as indicated by equation (7). However, according to

[12] Formally, and more generally, define D as the set of all $b_j \in [0, v(\phi_{\max})]$ such that $A(b)$ attains a local maximum at b_j and $A(b_j) > A(b)$ for all $b < b_j$. Assume that there are n such b_j and arrange them such that b_j is increasing in j. Define b_j' as the largest value in (b_j, b_{j+1}) such that $A(b_j') = A(b_j)$. If $\phi_t \leq v^{-1}(b_1)$, then $b(\phi_t) = v(\phi_t)$. If $\phi_t \in \left[v^{-1}(b_j), v^{-1}(b_j')\right]$, then $b(\phi_t) = b_j$, $j = 1, ..., n-1$. If $\phi_t \in \left[v^{-1}(b_j'), v^{-1}(b_{j+1})\right]$, then $b(\phi_t) = v(\phi_t)$, $j = 1, ..., n-1$. If $\phi_t \geq v^{-1}(b_n)$, then $b(\phi_t) = b_n$.

[13] Optimal debt $b(\phi_t)$ is a discontinuous function of the state, ϕ_t, for $\phi_t = \widehat{\phi}$ if for small $\varepsilon > 0$, the borrowing constraint is not binding for $\phi_t \in \left(\widehat{\phi} - \varepsilon, \widehat{\phi}\right)$ and the borrowing constraint is binding for $\phi_t \in \left(\widehat{\phi}, \widehat{\phi} + \varepsilon\right)$. That is, $b(\phi_t)$ is discontinuous at $\phi_t = \widehat{\phi}$ if the tradeoff theory is strictly operative for ϕ_t slightly smaller than $\widehat{\phi}$ but is not operative for ϕ_t slightly greater than $\widehat{\phi}$.

Corollary 4, the optimal amount of bonds will never jump downward, because any regime change that reduces the optimal amount of bonds, $b(\phi_t)$, by a discrete amount will lead to immediate default.

4 Optimal Investment

Having analyzed the optimal amount of bonds in Section 3, I turn in this section to the optimal rate of investment. The firm chooses optimal values of b_t and γ_t at time t to attain the maximum value of the right hand side of the expression for $v(\phi_t)$ equation (16). As discussed in Section 3, the maximand on the right hand side of equation (16) can maximized with respect to b_t simply by choosing b_t to maximize the tradeoff function $A(b_t)$, regardless of the value of γ_t. Substituting the optimal value of b_t, which is denoted $b(\phi_t)$, into $A(b_t)$ in equation (16) yields

$$v(\phi_t) = \max_{\gamma_t} \frac{(1-\tau)\left[\phi_t - \left[p_K\gamma_t + c(\gamma_t)\right]\right] + \lambda\overline{v} + A(b(\phi_t))}{\rho + \lambda + \delta - \gamma_t}. \tag{23}$$

The optimal value of the investment-capital ratio, γ_t, maximizes the expression on the right hand side of equation (23). Differentiating this expression with respect to γ_t and setting the derivative equal to zero yields an expression that is the fundamental equation of the q theory of investment,

$$(1-\tau)\left[p_K + c'(\gamma_t)\right] = v(\phi_t). \tag{24}$$

To interpret the first-order condition in equation (24), it is helpful to define tax-adjusted Tobin's q as the market value of the firm, $V(\phi_t, K_t) = v(\phi_t) K_t$ from equation (15), divided by the tax-adjusted replacement cost of the capital stock, $(1-\tau) p_K K_t$, so tax-adjusted Tobin's q is

$$q(\phi_t, \tau) \equiv \frac{V(\phi_t, K_t)}{(1-\tau) p_K K_t} = \frac{v(\phi_t)}{(1-\tau) p_K}. \tag{25}$$

Using this expression for $q(\phi_t, \tau)$, the first-order condition for the optimal investment-capital ratio γ_t in equation (24) can be written as

$$\gamma(\phi_t, \tau) = c'^{-1}\left(p_K\left[q(\phi_t, \tau) - 1\right]\right). \tag{26}$$

Up to this point, I have suppressed the dependence on τ of the optimal value of $\gamma(\phi_t)$, but I introduce it in the notation here to prepare for the discussion of the impact of taxes in Section 5. Since $c(\gamma_t)$ is strictly convex, $c'^{-1}()$ is strictly increasing, so the first-order condition in equation (26) shows that the optimal investment-capital ratio, represented by $\gamma(\phi_t)$, is a strictly increasing function of tax-adjusted Tobin's q.

The first-order condition in equation (24) can also be written as a function of $v_t = v(\phi_t)$ as

$$\gamma_t = c'^{-1}\left(\frac{v_t}{1-\tau} - p_K\right). \tag{27}$$

This alternative presentation of the first-order condition for optimal γ_t will be useful in Section 7, where I analyze the first-order condition for optimal bonds under serially correlated drawings of profitability across regimes .

Tax-adjusted *marginal* q is the value to the firm of an additional unit of capital, $\frac{\partial V(\phi_t,K_t,)}{\partial K_t}$, divided by the tax-adjusted cost of purchasing an additional unit of uninstalled capital, $(1-\tau)p_K$. Equation (15) immediately implies that $\frac{\partial V(\phi_t,K_t,)}{\partial K_t} = v(\phi_t)$, so that tax-adjusted marginal q equals tax-adjusted Tobin's (average) q. An alternative derivation of marginal q calculates $\frac{\partial V(\phi_t,K_t,)}{\partial K_t}$ as the expected present value of the net marginal contribution to profit accruing over the indefinite future to the undepreciated portion of a unit of capital installed today. The net marginal contribution to profit accruing to an additional unit of capital at time s consists of three components: (1) $(1-\tau)\phi_s$, which is the marginal contribution to after-tax operating profit, $(1-\tau)\phi_s K_s$, of a unit of capital at time s; (2) $(1-\tau)[\gamma_s c'(\gamma_s) - c(\gamma_s)]$, which is the reduction in the cost of adjustment associated with a given rate of investment, I_s, that is attributable to an additional unit of installed capital. This component is calculated as (the negative of) the partial derivative of after-tax total investment cost in equation (2) at time s with respect to K_s; and (3) $\tau[\rho + \lambda F(v^{-1}(b_s))]b_s$, which is the increase in the interest tax shield at time s when a one-unit increase in K_s leads to an increase in bonds of b_s and thus an increase of $\tau[\rho + \lambda F(v^{-1}(b_s))]b_s$ in the interest tax shield. This increase in borrowing associated with a one-unit increase in K_s is not a reflection of any collateral value of capital that might secure additional borrowing. In the current model, lenders receive nothing in default, so capital has no collateral value. The increase in borrowing associated with an increase in capital arises because additional capital increases the continuation value of the firm and thus increases the amount of bonds that the firm would be willing to repay.

Define M_s as the total contribution of an additional unit of capital at time s to net after-tax total profit, which is the sum of the three components just described. Therefore,

$$M_s \equiv (1-\tau)\left[\phi_s + \gamma_s c'(\gamma_s) - c(\gamma_s)\right] + \tau\left[\rho + \lambda F\left(v^{-1}(b_s)\right)\right] b_s. \tag{28}$$

Define $\chi(s)$ to be an indicator function that indicates whether firm has reached time s without defaulting. Specifically,

$$\chi(s) \equiv \begin{array}{l} 0, \text{ if the firm defaults at time } s \text{ or earlier} \\ 1, \text{ if the firm does not default before or at time } s \end{array}. \tag{29}$$

The marginal valuation of a unit of installed capital, $\frac{\partial V(\phi_t, K_t)}{\partial K_t}$, can be calculated as the expected present value of $e^{-\delta(s-t)}M_s$ over the period of time until the firm defaults. That is,

$$\frac{\partial V(\phi_t, K_t)}{\partial K_t} = E_t\left\{\int_t^{\infty} \chi(s)\, M_s e^{-(\rho+\delta)(s-t)} ds\right\}. \tag{30}$$

Proposition 8 $E_t\left\{\int_t^{\infty} \chi(s)\, M_s e^{-(\rho+\delta)(s-t)} ds\right\} = v(\phi_t)$, *where* $M_s \equiv (1-\tau)\left[\phi_s + \gamma_s c'(\gamma_s) - c(\gamma_s)\right] + \tau\left[\rho + \lambda F\left(v^{-1}(b_s)\right)\right] b_s$.

Proposition 8, along with equation (30), implies that marginal q and average q are equal to each other. Of course, that equality is a direct consequence of the fact that the value function in equation (15) is proportional to K_t. The contribution of Proposition 8 is that it provides an alternative method of calculating q. This alternative method is similar to that developed in Abel and Blanchard (1986) and used in Gilchrist and Himmelberg (1995), Bontempi et. al. (2004), and Chirinko and Schaller (2011), with an important extension: the new version of M_s presented here includes the increase in the interest tax shield made possible by an additional unit of capital. This component of M_s is readily observable and, notably, is invariant to whether the borrowing constraint is binding.[14]

[14]The inclusion of the interest tax shield in the expression for the marginal value of capital may appear to resemble Philippon's (2009) "bond market's q" which uses bond yields to calculate a measure of q. However, Philippon uses a Modigliani-Miller framework with no taxes and no deadweight cost of default and the amount of bonds is simply determined by assumption (p. 1018). The framework of the current paper departs from Modigliani-Miller by having a positive tax rate and a positive deadweight cost of default. In this non-MM framework, firms choose the amount of bonds optimally to maximize firm value.

5 Interaction between Investment and Financing Decisions

In a Modigliani-Miller environment, the mix of debt and equity is irrelevant to the firm. Specifically, the value of the firm and the optimal amount of capital investment are both unaffected by the mix of debt and equity. The environment in this paper departs from Modigliani-Miller because of the presence of taxes on the firm and deadweight costs associated with default. With immediate expensing of the total costs of investment, and a positive tax rate on after-tax income of the firm, the opportunity to issue bonds increases the value of the firm and increases the optimal rate of investment for any given value of the capital stock. Therefore, unlike in a Modigliani-Miller environment, capital investment is affected by the availability of bond financing. That is, optimal investment is not independent of the financing decision. However, there is an important sense in which the optimal amount of investment is independent of the financing decision: Given the value of tax-adjusted Tobin's q, the optimal investment-capital ratio is completely determined by the first-order condition in equation (26), without any need to take account of whether the borrowing constraint is binding or indeed whether or how many bonds are issued by the firm.

In this section, I compare the optimal investment of a firm that can issue bonds with the optimal investment of an otherwise-identical all-equity firm that cannot issue bonds. Let $q^E(\phi_t,\tau)$ be the value of $q(\phi_t,\tau)$ in equation (25) of an all-equity firm, and let $\gamma^E(\phi_t,\tau)$ be the value of $\gamma(\phi_t,\tau)$ in equation (26) for such a firm. Substituting $\gamma^E(\phi_s,\tau)$ for γ_s and setting $B_s = dB_s = 0$ for all s in equation (3) yields the value of the all-equity firm facing a permanent tax rate τ. Then using equation (25) yields

$$q^E(\phi_t,\tau) = \frac{1}{p_K} E_t\left\{\int_t^T \left(\phi_s - \left[p_K\gamma^E(\phi_s,\tau) + c\left(\gamma^E(\phi_s,\tau)\right)\right]\right) K_s e^{-\rho(s-t)} ds\right\}. \tag{31}$$

The following Proposition shows the role of bond financing in determining Tobin's q and its relation to the tax rate τ.

Proposition 9 *For any* $\tau \in [0,1)$,

1. $q^E(\phi_t,\tau)$ *and* $\gamma^E(\phi_t,\tau)$ *are invariant to* τ;

   (neutrality of taxes for all-equity firm)

2. $q(\phi_t, 0) = q^E(\phi_t, \tau)$ *and* $\gamma(\phi_t, 0) = \gamma^E(\phi_t, \tau)$;

   (irrelevance of bond financing if $\tau = 0$*)*

3. $q(\phi_t, \tau)$ *and* $\gamma(\phi_t, \tau)$ *are increasing in* τ; *and*

   (interest tax shield is increasing in τ*)*

4. $q(\phi_t, \tau) > q^E(\phi_t, \tau)$ *and* $\gamma(\phi_t, \tau) > \gamma^E(\phi_t, \tau)$ *for* $\tau > 0$.

   (bonds increase q *and* γ *relative to all-equity financing, if* $\tau > 0$*)*

Each of the numbered statements in Proposition 9 is actually a pair of statements: one statement about tax-adjusted q and a corresponding statement about the optimal value of the investment-capital ratio, γ. These statements are paired together because the first-order condition for the optimal investment-capital ratio in equation (26) indicates that the optimal value of γ is an increasing function of tax-adjusted Tobin's q. That function is invariant to the tax rate τ and, moreover, is invariant to whether the firm issues bonds optimally or is prohibited from issuing bonds.

Statement 1 is simply a restatement of the classic result, in Hall and Jorgenson (1971), for instance, that optimal investment is invariant to the tax rate for a firm that is financed entirely by equity and is allowed to expense its investment expenditures completely and immediately. Inspection of equation (31) shows that the tax rate τ does not enter the expression for tax-adjusted q, so that $q^E(\phi_t, \tau)$ does not depend on τ. An additional implication is that the optimal value of the investment-capital ratio, $\gamma^E(\phi_t, \tau)$, is invariant to τ.

Statement 2 is that the opportunity to issue bonds will not affect tax-adjusted Tobin's q or the optimal investment-capital ratio if the tax rate is zero. In that case, bonds do not offer an interest tax shield and there is no reason for the firm to issue bonds.

Statement 3 is that an increase in the tax rate τ increases tax-adjusted Tobin's q and hence increases the optimal rate of investment for a firm that can take advantage of the interest tax shield associated with bonds and can fully expense its capital expenditures. Finally, Statement 4, which is an implication of Statements 1, 2, and 3, is that the opportunity to issue bonds increases the values of tax-adjusted Tobin's q and the optimal investment-capital ratio. One way to see this implication (though it is not how I proved Statement 4) is to observe that the interest tax shield, $\tau\left[\rho + \lambda F\left(v^{-1}(b_s)\right)\right] b_s$, increases the contribution, M_s, to after-tax total profit accruing to an additional unit of capital in equation (28), and hence increases tax-adjusted q and optimal γ.

6 Lenders Receive Partial Recovery in Default

So far I have assumed that all shareholders and lenders to the firm receive zero when the firm defaults. Put differently, I have assumed that default imposes a deadweight cost equal to the entire value of the firm. In this section, I extend the analysis to allow the deadweight loss to be a fraction α, $0 < \alpha \leq 1$, of the value of the firm. Specifically, in the event of default, the firm's creditors take ownership of the firm, but a fraction α of the firm's capital stock disappears; the rest of the firm's economic environment is unaffected by default. If default occurs at time t_1 when $b(\phi_t) K_{t_1}$ bonds are outstanding, lenders take ownership of the firm with diminished capital stock, $(1-\alpha) K_{t_1}$, and thus recover an amount $(1-\alpha) v\left(\phi_{t_1};\alpha\right) K_{t_1}$ to partially offset the loss of $b(\phi_t) K_{t_1}$. In previous sections, where the firm and its creditors all receive zero in default, the implicit assumption was $\alpha = 1$. In this section and the following section, I consider the more general case in which α can be anywhere in $(0, 1]$. I will economize a bit on notation by suppressing the dependence on α in the notation for the optimal values of b_t and γ_t, which I will continue to denote as $b(\phi_t)$ and $\gamma(\phi_t)$ for the remainder of the paper.

In default, creditors suffer a loss of $b(\phi_t) K_{t_1}$ that is partially offset by taking ownership of the firm with remaining capital stock $(1-\alpha) K_{t_1}$, which has value $(1-\alpha) v\left(\phi_{t_1};\alpha\right) K_{t_1}$. Thus, the net loss suffered by lenders in default is $\left[b(\phi_t) - (1-\alpha) v\left(\phi_{t_1};\alpha\right)\right] K_{t_1}$, and the expected loss to lenders from default over the short interval of time from t to $t + dt$, is $\eta(b(\phi_t);\alpha) K_t dt$, where

$$\eta(b_t;\alpha) \equiv \lambda \int_{v(\phi;\alpha)<b_t} \left[b_t - (1-\alpha) v(\phi;\alpha)\right] dF(\phi) \geq 0. \tag{32}$$

For $0 \leq b_t \leq v(\phi_{\min};\alpha)$, there is zero probability of default and hence $\eta(b_t;\alpha) = 0$. For $b_t \geq v(\phi_{\min};\alpha)$, $\eta(b_t;\alpha)$ is strictly increasing in b_t. If $\alpha = 1$, then all of the firm's value is lost in the event of default, so $\eta(\phi_t;1) = \lambda b_t \int_{v(\phi;1)<b_t} dF(\phi)$, which is the instantaneous probability of default, denoted as p_t earlier the paper, multiplied by b_t.

For any $\alpha \in (0, 1]$, risk-neutral lenders are compensated for the expected loss in the event of default, so that total interest payments exceed $\rho b_t K_t$ by $\eta(b_t;\alpha) K_t$. Hence,

$$\text{interest payments at time } t = (\rho b_t + \eta(b_t;\alpha)) K_t. \tag{33}$$

With the extension to situations with $\alpha < 1$, the value of the firm, $V(\phi_t, K_t; \alpha)$, continues to be proportional to the capital stock and can be written as $v(\phi_t;\alpha) K_t$, as in equation (15). An expression for $v(\phi_t;\alpha)$ can be derived by substituting the interest payments from equation (33) for the interest payments $(\rho + p_t) b_t K_t$ in the Bellman equation in (14) and using the definition of $\eta(b_t;\alpha)$ in equation (32), as well as $K_{t_1} = e^{(\gamma_t - \delta)(t_1 - t)} K_t$, and $B_{t_1^-} = b_t K_{t_1}$. These calculations are performed in Appendix C, where it is shown that

$$v(\phi_t;\alpha) = \frac{(1-\tau)\{\phi_t - [p_K \gamma(\phi_t) + c(\gamma(\phi_t))]\} + \lambda \overline{v}(\alpha) + A(b(\phi_t);\alpha)}{\rho + \delta + \lambda - \gamma(\phi_t)}, \tag{34}$$

where

$$A(b;\alpha) \equiv \tau(\rho b + \eta(b;\alpha)) - \alpha\lambda \int_{v(\phi;\alpha)<b} v(\phi;\alpha)\, dF(\phi) \tag{35}$$

and

$$\overline{v}(\alpha) \equiv \int v(\phi;\alpha)\, dF(\phi). \tag{36}$$

The expression for $v(\phi_t;\alpha)$ in equation (34) is the same as the expression for $v(\phi_t)$ in equation (16), except that $A(b;\alpha)$ in equation (34) replaces $A(b)$ in equation (16) and $\overline{v}(\alpha)$ replaces $\overline{v}$. Notice that $A(b;1)$ equals $A(b)$ defined in equation (18), as it should, since when $\alpha = 1$, the deadweight loss equals the entire value of the firm and lenders receive nothing in default, as assumed in earlier sections. Inspection of the definitions of $\eta(b;\alpha)$ and $A(b;\alpha)$ in equations (32) and (35), respectively, reveals that for any value of α in $[0,1]$, $A(b;\alpha) = \tau\rho b$ if $0 \leq b \leq v(\phi_{\min})$, so that $A(b;\alpha)$ has the properties listed in Lemma 2. Therefore, Proposition 3, Corollary 4, Proposition 6, Corollary 7, and Proposition 9 continue to hold.

When the tradeoff theory is strictly operative, $A'(b;\alpha) = 0$. Allowing $\alpha < 1$ introduces a simple modification of the expression for $A'(b) = 0$ in equation (20). Substituting the definition of $\eta(b_t;\alpha)$ in equation (32) into equation (35) and differentiating the resulting expression for $A(b;\alpha)$ with respect to b yields[15]

$$A'(b;\alpha) = \tau\left[\rho + \lambda F\left(v^{-1}(b;\alpha)\right)\right] - \alpha(1-\tau)\lambda b v^{-1\prime}(b;\alpha) f\left(v^{-1}(b;\alpha)\right), \tag{37}$$

which differs from the expression for $A'(b)$ in equation (20) only by the factor α in the second term. That is, when creditors recover a fraction $1-\alpha$ of the firm's value in default, the marginal cost of increased debt associated with increased exposure to default is only α times as large as in the baseline case in which all of the firm's value disappears in default.

The prospect that lenders partially recover some of their bonds in default leads to a modification of Proposition 8, which is Proposition 10 below. Rewrite the value function $v(\phi_t;\alpha)$ in equation (34) by multiplying both sides by $\rho+\delta+\lambda-\gamma(\phi_t)$, adding $\gamma(\phi_t)v(\phi_t;\alpha)$ to both sides, using the first-order condition in equation (24) to replace $\gamma(\phi_t)v(\phi_t;\alpha)$ by $(1-\tau)[p_K\gamma(\phi_t)+\gamma(\phi_t)c'(\gamma(\phi_t))]$, and finally dividing both sides by $\rho+\delta+\lambda$ to obtain

$$v(\phi_t;\alpha) = \frac{(1-\tau)[\phi_t + \gamma(\phi_t)c'(\gamma(\phi_t)) - c(\gamma(\phi_t))] + \lambda\overline{v}(\alpha) + A(b(\phi_t);\alpha)}{\rho+\delta+\lambda}. \tag{38}$$

Define $M(\phi_t;\alpha)$ to be the marginal contribution to net after-tax cash flow of an additional unit of capital, so that

$$M(\phi_t;\alpha) \equiv (1-\tau)[\phi_t + \gamma(\phi_t)c'(\gamma(\phi_t)) - c(\gamma(\phi_t))] + \tau(\rho b(\phi_t) + \eta(\phi_t;\alpha)), \tag{39}$$

which is equivalent to M_t defined in equation (28), with the interest payments per unit of capital $\rho b(\phi_t)+\eta(\phi_t;\alpha)$ replacing $[\rho+\lambda F(v^{-1}(b_t))]\,b_t$ in equation (28).

Proposition 10 *$E_t\left\{\int_t^\infty (1-\alpha)^{n(s)} M(\phi_s;\alpha)e^{-(\rho+\delta)(s-t)}ds\right\} = v(\phi_t;\alpha)$, where $M(\phi_s;\alpha)$ is the marginal contribution to after-tax total profits defined in equation (39) and $n(s)$ is the number of defaults before or at time s.*

Proposition 10 implies that the marginal value of a unit of installed capital is the expected

[15] To differentiate $A(b;\alpha)$ with respect to b, it is helpful to rewrite equation (35) using the expression for $\eta(b;\alpha)$ from equation (32), as $A(b;\alpha) = \tau\rho b + \lambda\int_{\phi_{\min}}^{v^{-1}(b;\alpha)}[\tau b_t - (\alpha+\tau(1-\alpha))v(\phi;\alpha)]\,dF(\phi)$, which requires $v^{-1}(b;\alpha)$ to be strictly increasing in b, or equivalently $v(\phi;\alpha)$ to be strictly increasing in ϕ. The proof of Proposition 1 can be easily extended to prove that $v(\phi;\alpha)$ is strictly increasing in ϕ even when $\alpha>0$ is less than one. Alternatively, a special case of Proposition 11 in Section 7 proves that $v(\phi;\alpha)$ in the current section is strictly increasing in ϕ if $(1-\tau)\rho > \gamma_t - \delta$ for all possible optimal values of γ_t.

present value of $(1-\alpha)^{n(s)} M(\phi_s;\alpha) e^{-\delta(s-t)}$, the contribution to after-tax total profits accruing to the remaining undepreciated portion of a unit of capital installed at time t. This interpretation incorporates a broad view of depreciation that comprises two components: the continuous depreciation of capital at rate $\delta > 0$, and the disappearance of a fraction $\alpha > 0$ of the capital stock in a discrete lump every time the firm defaults. In the polar case of $\alpha = 1$, which is analyzed in previous sections, the factor $(1-\alpha)^{n(s)}$ equals one when $n(s) = 0$ and equals zero when $n(s) > 0$; in terms of $\chi(s)$ defined in equation (29), $(1-\alpha)^{n(s)} = \chi(s)$. As in the baseline case in which $\alpha = 1$, the marginal contribution to after-tax cash flow is the sum of three terms: (1) $(1-\tau)\phi_t$, the marginal contribution to operating profit; (2) $(1-\tau)[\gamma(\phi_t) c'(\gamma(\phi_t)) - c(\gamma(\phi_t))]$, the reduction in the adjustment cost of a given rate of investment; and (3) $\tau(\rho b(\phi_t) + \eta(\phi_t;\alpha))$, the increase in the interest tax shield associated with additional borrowing when the capital stock increases by one unit. This increase in the interest tax shield is simply the tax rate, τ, multiplied total interest payments and divided by the capital stock. Proposition 10 makes clear that the value of an additional unit of installed capital is the expected present value of total profits acrruing to the undepreciated portion of a unit of capital not just until the time of the first default, but indeed over the infinite future, even when the capital is owned by the firm's creditors, and the creditors of those creditors, and so on.

7 Correlated Profitability Across Regimes

The optimal amount of bonds is invariant to profitability, ϕ_t, when the tradeoff theory is operative, as discussed in Section 3. More precisely, if the tradeoff theory is strictly operative for all ϕ_t in a continuous interval $[\phi_L, \phi_H]$, then the optimal value of b_t, $b(\phi_t)$, is invariant to $\phi_t \in [\phi_L, \phi_H]$. This invariance of $b(\phi_t)$ implies that the market leverage ratio, $l(\phi_t;\alpha) \equiv \frac{b(\phi_t)}{v(\phi_t;\alpha)}$, is a decreasing function of ϕ_t when the tradeoff theory is strictly operative. However, this invariance depends on the assumption that realizations of ϕ are i.i.d. across regimes. In this section, I allow for first-order persistence in the realizations of ϕ across regimes and present an example in which positive persistence makes $b(\phi_t)$ a decreasing function of ϕ_t when the tradeoff theory is strictly operative, which reinforces the result that the optimal market leverage ratio, $l(\phi_t;\alpha) \equiv \frac{b(\phi_t)}{v(\phi_t;\alpha)}$, is a decreasing function of ϕ_t when the tradeoff theory is strictly operative.

Suppose that the current regime, in which profitability is ϕ_t, prevails from the current time t until time t_1, which is the time of the next regime change. Let $F\left(\phi_{t_1}|\phi_t\right)$ be the c.d.f. of profitability in the next regime, ϕ_{t_1}, conditional on current profitability, ϕ_t. Suppose that $F\left(\phi_{t_1}|\phi_t\right)$ is non-increasing in ϕ_t. That is, if $\phi_{t,2} > \phi_{t,1}$, then $F\left(\phi_{t_1}|\phi_{t,2}\right) \leq F\left(\phi_{t_1}|\phi_{t,1}\right)$ so that $F\left(\phi_{t_1}|\phi_{t,2}\right)$ first-order stochastically dominates $F\left(\phi_{t_1}|\phi_{t,1}\right)$. A regime with high profitability ϕ_t is more likely to be followed by a regime with high profitability.

When the tradeoff theory of debt is strictly operative, the constraint $b\left(\phi_t\right) \leq v\left(\phi_t;\alpha\right)$ is not binding so the optimal value of debt, $b\left(\phi_t\right)$, is a local maximum of $A\left(b;\alpha|\phi_t\right)$, where the notation indicates the dependence on ϕ_t of the distribution of the next realization of profitability, $F\left(\phi_{t_1}|\phi_t\right)$, in the appropriately-modified definition of $A\left(b;\alpha\right)$ in equation (35). Because $b\left(\phi_t\right)$ is a local maximum of $A\left(b;\alpha|\phi_t\right)$, the first-order condition $A'\left(b;\alpha|\phi_t\right) = 0$ and the second-order condition $A''\left(b;\alpha|\phi_t\right) \leq 0$ both hold, when the first and second derivatives of $A\left(b;\alpha|\phi\right)$ are evaluated at $b = b\left(\phi_t\right)$. With the specification of persistence presented here, the first-order condition for the optimal amount of debt in the current regime when the tradeoff theory is strictly operative is, from equation (37),

$$A'\left(b;\alpha|\phi_t\right) = \begin{array}{c} \tau\left[\rho+\lambda F\left(v^{-1}\left(b;\alpha\right)|\phi_t\right)\right] \\ -\alpha\left[\rho+\lambda+\delta-c'^{-1}\left(\frac{b}{1-\tau}-p_K\right)\right]\lambda b f\left(v^{-1}\left(b;\alpha\right)|\phi_t\right) \end{array} = 0, \tag{40}$$

where $f\left(\phi_{t_1}|\phi_t\right)$ is the density function associated with $F\left(\phi_{t_1}|\phi_t\right)$, and I have used equation (21) to replace $v^{-1\prime}\left(b;\alpha\right)$ by $\frac{\rho+\lambda+\delta-\gamma\left(v^{-1}(b;\alpha)\right)}{1-\tau}$ and equation (27) to replace $\gamma\left(v^{-1}\left(b;\alpha\right)\right)$ with $c'^{-1}\left(\frac{b}{1-\tau}-p_K\right)$.

As discussed in Section 3, the first of the two terms in equation (40), $\tau\left[\rho+\lambda F\left(v^{-1}\left(b;\alpha\right)|\phi_t\right)\right]$, is the interest tax shield associated with an additional unit of debt and the second term is the increase in the expected default cost associated with an additional unit of debt. The formulation of persistence specified in this section implies that a higher value of ϕ_t decreases $F\left(v^{-1}\left(b;\alpha\right)|\phi_t\right)$, and hence decreases the interest tax shield associated with an additional dollar of debt. That is, an increase in ϕ_t makes the distribution of ϕ_{t_1} more favorable, thereby reducing the probability of default at time t_1 and hence reducing $\rho+\lambda F\left(v^{-1}\left(b;\alpha\right)|\phi_t\right)$. The second of the two terms in equation (40), $\alpha\left[\rho+\lambda+\delta-c'^{-1}\left(\frac{b}{1-\tau}-p_K\right)\right]\lambda b f\left(v^{-1}\left(b;\alpha\right)|\phi_t\right)$, is the increase in the expected cost of default associated with an additional unit of debt. At this level of generality, an increase in ϕ_t could increase, decrease, or leave unchanged the marginal default cost because $f\left(v^{-1}(b;\alpha)\,|\phi_t\right)$ could increase, decrease, or remain unchanged in response to an increase in ϕ_t. If the impact on $f\left(v^{-1}(b;\alpha)\,|\phi_t\right)$ is small enough in absolute value, the decrease in the marginal interest tax shield dominates any change in the marginal default cost and would reduce the optimal value of bonds if the tradeoff theory is operative.

As an example, suppose that $F\left(\phi_{t_1}|\phi_t\right)$ is uniform on the interval $[m_{t_1} - d, m_{t_1} + d]$, where $\phi_{\min} \leq m_{t_1} - d < m_{t_1} + d \leq \phi_{\max}$. The conditional mean, m_{t_1}, depends on the realization of ϕ_t. To model positive serial dependence, assume $m_{t_1} = g(\phi_t)$ where $g'(\phi_t) \geq 0$. To ensure that the support of the distribution of ϕ_{t_1} remains within $[\phi_{\min}, \phi_{\max}]$, suppose that $\phi_{\min} + d \leq g(\phi_{\min}) < g(\phi_{\max}) \leq \phi_{\max} - d$. Under this specification, an increase in ϕ_t simply translates the c.d.f. $F\left(\phi_{t_1}|\phi_t\right)$ to the right but leaves $f\left(v^{-1}(b;\alpha)\,|\phi_t\right)$ unchanged for b such that $v^{-1}(b;\alpha)$ is in the supports of both the original distribution and the new, translated, distribution. If the tradeoff theory is operative, this change decreases $A'(b;\alpha|\phi_t)$ in equation (40) so that $A'(b;\alpha|\phi_t) < 0$. To restore the first-order condition, the value of b must fall since $A''(b;\alpha|\phi_t) \leq 0$. In this simple example, a higher value of ϕ_t leads to a lower value of $b(\phi_t)$ and hence reduces the optimal value of bonds, when the tradeoff theory is operative. That is, $b(\phi_t)$ is decreasing in ϕ_t in this simple example.

To analyze the relation between the market leverage ratio, $l(\phi_t;\alpha) \equiv \frac{b(\phi_t)}{v(\phi_t;\alpha)}$, and ϕ_t, I use the relation between $v(\phi_t;\alpha)$ and ϕ_t stated in the following proposition.

Proposition 11 *Suppose that $(1-\tau)\rho > \gamma(\phi_t) - \delta$ for all $\phi \in [\phi_{\min}, \phi_{\max}]$ and that $F\left(\phi|\phi_{t,2}\right)$ first-order stochastically dominates $F\left(\phi|\phi_{t,1}\right)$ if $\phi_{t,2} > \phi_{t,1}$. Then $v(\phi_t;\alpha)$ is strictly increasing.*

The condition that $(1-\tau)\rho > \gamma(\phi_t) - \delta$ for all $\phi \in [\phi_{\min}, \phi_{\max}]$ means that the after-tax riskless rate exceeds the growth rate of the capital stock for all possible realizations of ϕ. It is stronger than $\rho > \gamma(\phi_t) - \delta$, which is used in previous sections, and is introduced here to facilitate the use of a contraction to prove Proposition 11.[16,17]

[16] If the operator T defined in the proof of Proposition 11 ignored the borrowing constraint $b_t \leq v(\phi_t;\alpha)$, then the more familiar condition $\gamma(\phi_t) - \delta < \rho$ would be sufficient to prove that T satisfies the discounting property of a contraction. However, T takes account of the borrowing constraint, and the borrowing constraint depends on the value function itself, so a higher value of the firm increases the feasible level of borrowing, which potentially increases the interest tax shield. Therefore, the stronger condition $\gamma(\phi_t) - \delta < (1-\tau)\rho$ is used to ensure that T satisfies the discounting property.

[17] The condition $(1-\tau)\rho > \gamma(\phi_t) - \delta$ will be satisfied if $\lim_{\gamma \nearrow (1-\tau)\rho+\delta} c'(\gamma) = \infty$. Alternatively, this condition will be satisfied if $(1-\tau)\left[p_K + c'\left((1-\tau)\rho + \delta - \varepsilon\right)\right] > v(\phi_{\max})$ for some $\varepsilon > 0$.

In the simple example with a uniform distribution of ϕ, $b(\phi_t)$ is decreasing in ϕ_t when the tradeoff theory is operative and $v(\phi_t;\alpha)$ is increasing in ϕ_t so the market leverage ratio, $l(\phi_t;\alpha) \equiv \frac{b(\phi_t)}{v(\phi_t;\alpha)}$, is decreasing in ϕ_t. This finding is consistent with the finding in Section 3 that when the tradeoff theory is operative, $l(\phi_t;\alpha)$ is decreasing under the assumption of i.i.d. realizations of ϕ_t across regimes. This finding is also consistent with the empirical relation between market leverage and profitability.

8 Conclusion

The interest tax shield provided by tax deductibility of interest payments provides an incentive for firms to issue bonds. This incentive is tempered by the increased exposure to default that would accompany an increase in bonds outstanding. In addition, the ability to issue bonds is limited by an endogenous borrowing constraint that recognizes that potential creditors will not buy such a large amount of bonds that a firm would choose to default immediately after issuing the bonds. The interest tax shield and the increased exposure to default are the elements of the tradeoff theory of debt. The tradeoff theory is operative if the marginal interest tax shield associated with an additional dollar of bonds equals the marginal cost of the associated increased exposure to default. This equality of marginal benefit and marginal cost can fail to hold if the endogeneous borrowing constraint is binding. When the borrowing constraint binds, the marginal benefit associated with the increased interest tax shield exceeds the marginal cost of increased exposure to default, and the firm would increase its debt if it could. Thus, whether or not the tradeoff theory of debt is operative depends endogenously on whether the borrowing constraint is binding. When the borrowing constraint is not binding, the tradeoff theory is operative. I have shown that when the tradeoff theory is operative, the optimal market leverage ratio is a decreasing function of profitability. This negative relation between market leverage and profitability is consistent with empirical findings, yet has proved to be a challenge for many theoretical analyses of the tradeoff theory.

Capital investment decisions reflect the equality of the marginal cost of investment and tax-adjusted marginal q, which is the value of an additional unit of capital divided by the after-tax cost of purchasing capital. Under the commonly-used assumptions that the operating profit of the firm is proportional to the capital stock and the adjustment cost function is linearly homogeneous in investment and the capital stock, the value of the firm is proportional to the capital stock. Therefore, the marginal value of a unit of installed capital is identically equal to the average value of each unit of installed capital, and hence average q and marginal q are identically equal to each other under these assumptions. The ability to issue bonds and take advantage of the interest tax shield increases the value of the firm for a given capital stock, and thus increases (marginal and average) q and hence increases investment. This increase in the value of the firm becomes evident when computing the marginal value of a unit of capital as the expected present value of the increase in after-tax total profits accruing to the undepreciated remains of a unit of capital over the indefinite future. Specifically, the increase in these total profits accruing to one unit of capital at any point in time is the sum of (1) the after-tax marginal operating profit of capital; (2) the reduction in the after-tax adjustment cost associated with a given rate of investment that is made possible by an additional unit of capital; and (3) the interest tax shield divided by the capital stock. The third term, the interest tax shield, has been overlooked in previous calculations of marginal q. It is easily calculated as the product of the tax rate and total interest payments by the firm divided by the firm's capital stock.

Because the ability to issue bonds increases the value of firm, and hence increases both q and investment, it is evident that there is an interaction between financing and capital investment. However, even though the firm is outside the Modigliani-Miller framework, there is an interesting sense in which the capital investment decision and the financing decision are independent. Specifically, if the value function is taken as given, the optimal investment-capital ratio does not depend on the amount of bonds, and the optimal amount of bonds does not depend on the rate of investment.

For an all-equity firm that can never issue bonds, tax-adjusted q is invariant to the tax rate if the firm can immediately and fully expense its total cost of investment. Consequently, optimal investment is invariant to the tax rate for such a firm. Allowing a firm to issue bonds and take advantage of the interest tax shield increases tax-adjusted marginal q and hence increases investment. Furthermore, the higher is the tax rate, the higher are tax-adjusted q and optimal investment, when the firm can immediately and completely expense its total cost of investment.

References

Abel, Andrew B., "Optimal Debt and Profitability in the Tradeoff Theory," National Bureau of Economic Research Working Paper No. 21548, September 2015.

Abel, Andrew B. and Olivier J. Blanchard, "The Expected Present Value of Profits and the Cyclical Variability of Investment," Econometrica, 54, 2 (March 1986), 249-273.

Bontempi, Elena, Alessandra Del Boca, Alessandra Franzosi, Marzio Galeotti and Paola Rota, "Capital Heterogeneity: Does It Matter? Fundamental Q and Investment on a Panel of Italian Firms," The RAND Journal of Economics, 35, 4 (Winter, 2004), pp. 674-690.

Chirinko, Robert S. and Huntley Schaller, "Fundamentals, Misvaluation, and Investment," Journal of Money, Credit, and Banking, 43, 7 (October 2011), 1423-1442.

Gilchrist, Simon and Charles Himmelberg, "Evidence on the Role of Cash Flow for Investment," Journal of Monetary Economics, 36 (1995), 541-572.

Goldstein, Robert, Nengjiu Ju, and Hayne Leland, "An EBIT-Based Model of Dynamic Capital Structure," Journal of Business, 74, 4 (October 2001), 483-512.

Leland, Hayne E., "Corporate Debt Value, Bond Covenants, and Optimal Capital Structure," The Journal of Finance, 49, 4 (September 1994), 1213-1252.

Modigliani, Franco and Merton H. Miller, "The Cost of Capital, Corporation Finance and the Theory of Investment," American Economic Review, 48, 3 (June 1958), 261-297.

Myers, Stewart C., "Still Searching for Optimal Capital Structure," Journal of Applied Corporate Finance, 6, 1 (Spring 1993), 4-14.

Philippon, Thomas, "The Bond Market's q," The Quarterly Journal of Economics, 124, 3 (August 2009), pp. 1011-1056.

A Appendix: Proofs

Proof. of Proposition 1. Consider $\phi_2 > \phi_1$. Observe that $v(\phi_2) = \max_{b_t,\gamma_t} \frac{(1-\tau)[\phi_2-[p_K\gamma_t+c(\gamma_t)]]+\lambda\overline{v}+A(b_t)}{\rho+\lambda+\delta-\gamma_t} \geq \frac{(1-\tau)[\phi_2-[p_K\gamma(\phi_1)+c(\gamma(\phi_1))]]+\lambda\overline{v}+A(b(\phi_1))}{\rho+\lambda+\delta-\gamma(\phi_1)} > \frac{(1-\tau)[\phi_1-[p_K\gamma(\phi_1)+c(\gamma(\phi_1))]]+\lambda\overline{v}+A(b(\phi_1))}{\rho+\lambda+\delta-\gamma(\phi_1)} = v(\phi_1)$, where the first inequality follows from the maximization operator and the second inequality follows because $\frac{1-\tau}{\rho+\lambda+\delta-\gamma(\phi_t)}$ is positive, since $\tau < 1$ and the assumption $\lim_{\gamma\nearrow\rho+\delta} c'(\gamma-\varepsilon) = \infty$ for some $\varepsilon > 0$ implies that $\gamma(\phi_t) < \rho+\delta$. To complete the proof, one needs to verify that $b(\phi_1)$ is feasible when $\phi_t = \phi_2$. It suffices to prove that $b(\phi_1) \leq v(\phi_2)$, which follows from $b(\phi_1) \leq v(\phi_1) < v(\phi_2)$. ■

Proof. of Lemma 2. If $b_t \leq v(\phi_{\min})$, then $F(v^{-1}(b_t)) = 0$ and $\int_{v(\phi)<b_t} v(\phi)\,dF(\phi) = 0$, so the definition of $A(b_t)$ implies $A(b_t) \equiv \tau\rho b_t$ (Statement 1). Suppose that, contrary to what is to be proved, $\arg\max_{b\in[0,v(\phi_t)]} A(b) < v(\phi_{\min})$, and let $\widehat{b} < v(\phi_{\min})$ denote this value of $\arg\max_{b\in[0,v(\phi_t)]} A(b)$. Statement 1 implies that $A\left(\widehat{b}\right) = \tau\rho\widehat{b} < \tau\rho v(\phi_{\min}) = A(v(\phi_{\min}))$, which implies that $\widehat{b} \neq \arg\max_{b\in[0,v(\phi_t)]} A(b)$. Finally, as shown in footnote 4, the condition in equation (4) implies that $V(\phi_t, K_t) > 0$ for all $\phi_t \in [\phi_{\min}, \phi_{\max}]$ when $K_t > 0$, so that $v(\phi_{\min}) > 0$ (Statement 2). ■

Proof. of Proposition 3. Statement 2 in Lemma 2 and the expression for $b(\phi_t)$ in equation (19) directly imply that $b(\phi_t) \geq v(\phi_{\min}) > 0$ (Statement 1).

Consider $\phi_1 < \phi_2$, which implies $v(\phi_1) < v(\phi_2)$ and hence $[0, v(\phi_1)] \subset [0, v(\phi_2)]$. Therefore, $\max_{b\in[0,v(\phi_2)]} A(b) \geq \max_{b\in[0,v(\phi_1)]} A(b)$. If $\max_{b\in[0,v(\phi_2)]} A(b) > \max_{b\in[0,v(\phi_1)]} A(b)$, then $b(\phi_2) = \min\left\{\arg\max_{b\in[0,v(\phi_2)]} A(b)\right\} > v(\phi_1) \geq b(\phi_1)$. Alternatively, if $\max_{b\in[0,v(\phi_2)]} A(b) = \max_{b\in[0,v(\phi_1)]} A(b)$, then $\min\left\{\arg\max_{b\in[0,v(\phi_2)]} A(b)\right\} \in [0, v(\phi_1)]$, so $b(\phi_2) = \min\left\{\arg\max_{b\in[0,v(\phi_2)]} A(b)\right\} = \min\left\{\arg\max_{b\in[0,v(\phi_1)]} A(b)\right\} = b(\phi_1)$. Therefore, $b(\phi_t)$ is non-decreasing in ϕ_t (Statement 3). Statement 2 of Lemma 2 and the expression for $b(\phi_t)$ in equation (19) imply that $A(b(\phi_t)) > 0$ (Statement 2). Finally, the expression for $b(\phi_t)$ in equation (19) implies that $A(b(\phi_t)) = \max_{b\in[0,v(\phi_t)]} A(b)$, so the statement earlier in this proof that $\max_{b\in[0,v(\phi_2)]} A(b) \geq \max_{b\in[0,v(\phi_1)]} A(b)$ implies that $A(b(\phi_2)) \geq A(b(\phi_1))$ (Statement 4). ■

Proof. Corollary 4 (by contradiction) The optimal amount of bonds changes by a discrete amount at a point of time only when the regime changes so that profitability, ϕ, changes. Assume that the firm reduces its outstanding debt by a discrete amount at time t, from $b(\phi_{t^-}) K_t$ to $b(\phi_{t^+}) K_t$, when a regime changes causes profitability to change from ϕ_{t^-}

to ϕ_{t^+}. Therefore, $b(\phi_{t^+}) < b(\phi_{t^-})$. If $v(\phi_{t^+}) \geq b(\phi_{t^-})$, then $b(\phi_{t^+}) \geq b(\phi_{t^-})$, which is a contradiction. Therefore, $v(\phi_{t^+}) < b(\phi_{t^-})$. Since, $v(\phi_{t^+}) < b(\phi_{t^-})$, the firm would choose to default at time t^+, and the firm would cease to be ongoing at that time. Therefore, $dB_t < 0$ cannot be a discrete amount. The definition of dividends in equation (10) implies that when shareholders inject a discrete amount of funds into the firm at a point of time, $dD_t < 0$ is a discrete amount rather than a flow, and hence $dB_t < 0$ is a discrete amount, which contradicts the result that $dB_t < 0$ cannot be a discrete amount. ■

Proof. of Proposition 6. Assume that $b(\phi_t) < v(\phi_t)$ for all $\phi_t \in \Phi$. Define $b^* \equiv b(\phi_2) \leq v(\phi_2)$. If $v(\phi_1) \geq b^*$, then $b(\phi_t) = b^*$ for all $\phi_t \in \Phi$. If $v(\phi_1) < b^*$, then there exists a $\phi^* \in \Phi$ such that $v(\phi^*) = b^*$ and $b(\phi^*) = b^*$. Therefore, $v(\phi^*) = b(\phi^*)$, which contradicts the assumption that $b(\phi_t) < v(\phi_t)$ for all $\phi_t \in \Phi$, and completes the proof of Statement 1. Statement 2 follows immediately from Proposition 1 and the assumption that $b(\phi_t) = v(\phi_t)$ for all $\phi_t \in \Phi$. ■

Proof. of Corollary 7. Assume that $b(\phi_t) < v(\phi_t)$ in a neighborhood around ϕ_t. Proposition 6 implies that $b(\phi_t)$ is invariant to ϕ_t and Proposition 1 implies that $v(\phi_t)$ is strictly increasing in ϕ_t, so the market leverage ratio, $l(\phi_t) \equiv \frac{b(\phi_t)}{v(\phi_t)}$, is strictly decreasing in ϕ_t, which proves Statement 1. Assume that $b(\phi_t) = v(\phi_t)$, which immediately implies $l(\phi_t) \equiv \frac{b(\phi_t)}{v(\phi_t)} = 1$ and thus proves Statement 2. ■

Proof. of Proposition 8. Replace γ_t and b_t in equation (16) by their optimal values, $\gamma(\phi_t)$ and $b(\phi_t)$, respectively, multiply both sides of the equation by $\rho + \lambda + \delta - \gamma(\phi_t)$, then add $\gamma(\phi_t)\, v(\phi_t)$ to both sides of the resulting equation and use the first-order condition for investment in equation (24) to obtain

$$\begin{aligned}(\rho + \lambda + \delta)\, v(\phi_t) &= (1-\tau)\left[\phi_t + \gamma(\phi_t)\, c'(\gamma(\phi_t)) - c(\gamma(\phi_t))\right] \\ &\quad + \lambda \overline{v} + A(b(\phi_t))\,.\end{aligned} \tag{41}$$

Use the definition of $A(b_t)$ given in equation (18) to rewrite equation (41) and rearrange to obtain

$$v(\phi_t) = \frac{1}{\rho + \lambda + \delta} \left\{ \begin{array}{c} (1-\tau)\left[\phi_t + \gamma(\phi_t)\, c'(\gamma(\phi_t)) - c(\gamma(\phi_t))\right] \\ +\tau\left[\rho + \lambda F\left(v^{-1}(b(\phi_t))\right)\right] b(\phi_t) + \lambda \int_{v(\phi) \geq b(\phi_t)} v(\phi)\, dF(\phi) \end{array} \right\}. \tag{42}$$

Use the definition of M_s from equation (28) to rewrite equation (42) as

$$v(\phi_t) = \frac{1}{\rho+\lambda+\delta}\left[M_t + \lambda \int_{v(\phi)\geq b(\phi_t)} v(\phi)\, dF(\phi)\right]. \tag{43}$$

A solution of the recursive equation in equation (43) is

$$v(\phi_t) = E_t\left\{\sum_{j=0}^{\infty} \chi(t_j)\left(\frac{\lambda}{\rho+\lambda+\delta}\right)^j \frac{M_{t_j}}{\rho+\lambda+\delta}\right\}, \tag{44}$$

where $\chi(t_j)$ is the indicator function defined in equation (29) and $t_0 = t$.

Since the conditional density of t_1 as of time t is $\lambda e^{-\lambda(t_1-t)}$,

$$E_t\left\{e^{-(\rho+\delta)(t_1-t)}\right\} = \frac{\lambda}{\rho+\lambda+\delta}, \tag{45}$$

and since $M_s = M_{t_j}$ for $t_j \leq s < t_{j+1}$,

$$E_t\left\{\int_{t_j}^{t_{j+1}} e^{-\delta(s-t)} M_s e^{-\rho(s-t_j)} ds\right\} = \frac{M_{t_j}}{\rho+\lambda+\delta}. \tag{46}$$

Equation (45) implies that

$$E_t\left\{e^{-(\rho+\delta)(t_j-t)}\right\} = \left(\frac{\lambda}{\rho+\lambda+\delta}\right)^j. \tag{47}$$

Use equations (46) and (47) and the assumption that ϕ_t is i.i.d. across regimes to rewrite equation (44) as

$$v(\phi_t) = E_t\left\{\sum_{j=0}^{\infty} \chi(t_j)\, e^{-(\rho+\delta)(t_j-t)} \int_{t_j}^{t_{j+1}} e^{-\delta(s-t)} M_s e^{-\rho(s-t_j)} ds\right\}. \tag{48}$$

Since $\chi(s) = \chi(t_j)$ for $t_j \leq s < t_{j+1}$, equation (48) implies

$$v(\phi_t) = E_t\left\{\int_t^{\infty} \chi(s)\, M_s e^{-(\rho+\delta)(s-t)} ds\right\}. \tag{49}$$

■

Proof. of Proposition 9. Proof of Statement 1: Let

$V^E(\phi_t, K_t; \tau_i) = \max_{\gamma_s} (1-\tau_i) E_t \left\{\int_t^\infty (\phi_s - [p_K\gamma_s + c(\gamma_s)]) K_s e^{-\rho(s-t)} ds\right\} > 0$ be the value of an all-equity firm facing a constant tax rate τ_i. Suppose that $0 \leq \tau_1 < \tau_2 < 1$. $V^E(\phi_t, K_t; \tau_2) \geq (1-\tau_2) E_t \left\{\int_t^T \left(\phi_s - \left[p_K\gamma^E(\phi_s;\tau_1) + c\left(\gamma^E(\phi_s;\tau_1)\right)\right]\right) K_{1,s} e^{-\rho(s-t)} ds\right\} = (1-\tau_2)\frac{V^E(\phi_t,K_t;\tau_1)}{1-\tau_1}$, so $\frac{V^E(\phi_t,K_t;\tau_2)}{1-\tau_2} \geq \frac{V^E(\phi_t,K_t;\tau_1)}{1-\tau_1}$. Similarly,

$V^E(\phi_t, K_t; \tau_1) \geq (1-\tau_1) E_t \left\{\int_t^T \left(\phi_s - \left[p_K\gamma^E(\phi_s;\tau_2) + c\left(\gamma^E(\phi_s;\tau_2)\right)\right]\right) K_{2,s} e^{-\rho(s-t)} ds\right\}$ $= (1-\tau_1)\frac{V^E(\phi_t,K_t;\tau_2)}{1-\tau_2}$, so $\frac{V^E(\phi_t,K_t;\tau_1)}{1-\tau_1} \geq \frac{V^E(\phi_t,K_t;\tau_2)}{1-\tau_2}$. Therefore, $\frac{V^E(\phi_t,K_t;\tau_2)}{1-\tau_2} \geq \frac{V^E(\phi_t,K_t;\tau_1)}{1-\tau_1}$ $\geq \frac{V^E(\phi_t,K_t;\tau_2)}{1-\tau_2}$, which implies $q^E(\phi_t,\tau_2) \equiv \frac{V^E(\phi_t,K_t;\tau_2)}{(1-\tau_2)p_K} = \frac{V^E(\phi_t,K_t;\tau_1)}{(1-\tau_1)p_K} \equiv q^E(\phi_t,\tau_1)$. Then $q^E(\phi_t,\tau_2) = q^E(\phi_t,\tau_1)$ implies $\gamma^E(\phi_t,\tau_2) = \gamma^E(\phi_t,\tau_1)$.

Proof of Statement 2: If $\tau = 0$, then the definition of $A(b_t)$ in equation (18) implies that $A(b_t) = -\lambda\int_{v(\phi)<b_t} v(\phi)\, dF(\phi) \leq 0$, so that $\max_{b_t \leq v(\phi_t)} A(b_t) = 0$. Setting $A(b_t) = 0$ in equation (16) yields $v(\phi_t) = \max_{\gamma_t} \frac{(1-\tau)[\phi_t - [p_K\gamma_t + c(\gamma_t)]] + \lambda\overline{v}}{\rho+\lambda+\delta-\gamma_t}$, which is identical to $v(\phi_t)$ for an all-equity firm. Therefore, $q(\phi_t, 0) = q^E(\phi_t, 0) = q^E(\phi_t, \tau)$, where the second equality follows from Statement 1 of this proposition. Therefore, $\gamma(\phi_t, 0) = \gamma^E(\phi_t, 0) = \gamma^E(\phi_t, \tau)$.

Proof of Statement 3: As will be verified below, when the tax rate is $\tau_2 > \tau_1$, it is feasible to set $\gamma_s = \gamma(\phi_s;\tau_1)$ and $b_s = \frac{1-\tau_2}{1-\tau_1} b(\phi_s;\tau_1)$ so the value of the firm given in equation (3) can be written as[18] $V(\phi_t, K_t; \tau_2) \geq$

$$E_t \left\{ \begin{array}{c} \int_t^T (1-\tau_2) \left(\begin{array}{c} \phi_s - [p_K\gamma(\phi_s;\tau_1) + c(\gamma(\phi_s;\tau_1))] \\ -\left(\rho + P\left(\frac{1-\tau_2}{1-\tau_1} b(\phi_s;\tau_1);\tau_2\right)\right) \frac{1-\tau_2}{1-\tau_1} b(\phi_s;\tau_1) \end{array} \right) K_{1,s} e^{-\rho(s-t)} ds \\ + \int_t^T e^{-\rho(s-t)} \frac{1-\tau_2}{1-\tau_1} dB_{1,s} \end{array} \right\}, \text{where}$$

$K_{i,s}$ is the capital stock, and $B_{i,s}$ is the value of bonds outstanding, if the permanent tax rate is τ_i; $P(b;\tau_i) \equiv \lambda\int_{v(\phi;\tau_i)<b} dF(\phi)$ is the probability that the firm will default over the next infinitesimal interval of time, if it issues bonds $bK_{i,s}$ and the tax rate is τ_i.[19] This inequality can be rewritten as

$V(\phi_t, K_t; \tau_2) \geq$

[18] The value of T is the date of default under the tax rate τ_1. As will be shown, $\frac{V(\phi_t,K_t;\tau_2)}{(1-\tau_2)p_K K_t} > \frac{V(\phi_t,K_t;\tau_1)}{(1-\tau_1)p_K K_t}$, which implies $v(\phi_t;\tau_2) > \frac{1-\tau_2}{1-\tau_1} v(\phi_t;\tau_1)$. Therefore, $v(\phi_t;\tau_2) \geq \frac{1-\tau_2}{1-\tau_1} b(\phi_t;\tau_1)$ whenever $v(\phi_t;\tau_1) \geq b(\phi_t;\tau_1)$. Hence, the default date, $T_2 \geq T_1$, where T_i is the default date under τ_i.

[19] $P(b;\tau_i)$ is the default probability for an arbitrary level of b when $\tau = \tau_i$. The probability $p(\phi_t)$ that is the default premium in the interest rate is the probability that prevails when the firm issues the optimal amount of bonds, $b(\phi_t) K_t$. That is, when the tax rate is τ_i, $p(\phi_t) = \lambda\int_{v(\phi;\tau_i)<b(\phi_t)} dF(\phi) = P(b(\phi_t);\tau_i)$.

$$\frac{1-\tau_2}{1-\tau_1}E_t\left\{\begin{array}{c}\int_t^T(1-\tau_1)\left(\begin{array}{c}\phi_s-\left[p_K\gamma(\phi_s;\tau_1)+c(\gamma(\phi_s;\tau_1))\right]\\-\left(\rho+P\left(\frac{1-\tau_2}{1-\tau_1}b(\phi_s;\tau_1);\tau_2\right)\right)\frac{1-\tau_2}{1-\tau_1}b(\phi_s;\tau_1)\end{array}\right)K_{1,s}e^{-\rho(s-t)}ds\\+\int_t^T e^{-\rho(s-t)}dB_{1,s}\end{array}\right\}.$$

If, as will be shown,

$\int_t^T\left(\rho+P\left(\frac{1-\tau_2}{1-\tau_1}b(\phi_s;\tau_1);\tau_2\right)\right)\frac{1-\tau_2}{1-\tau_1}b(\phi_s;\tau_1)K_{1,s}e^{-\rho(s-t)}ds <$
$\int_t^T(\rho+P(b(\phi_s;\tau_1);\tau_1))b(\phi_s;\tau_1)K_{1,s}e^{-\rho(s-t)}ds$, then $V(\phi_t,K_t;\tau_2) >$

$$\frac{1-\tau_2}{1-\tau_1}E_t\left\{\begin{array}{c}\int_t^T(1-\tau_1)\left(\begin{array}{c}\phi_s-\left[p_K\gamma(\phi_s;\tau_1)+c(\gamma(\phi_s;\tau_1))\right]\\-(\rho+P(b(\phi_s;\tau_1);\tau_1))b(\phi_s;\tau_1)\end{array}\right)K_{1,s}e^{-\rho(s-t)}ds\\+\int_t^T e^{-\rho(s-t)}dB_{1,s}\end{array}\right\}=\frac{1-\tau_2}{1-\tau_1}V(\phi_t,K_t;\tau_1).$$

Therefore, $q(\phi_t;\tau_2)\equiv\frac{V(\phi_t,K_t;\tau_2)}{(1-\tau_2)p_KK_t}>\frac{V(\phi_t,K_t;\tau_1)}{(1-\tau_1)p_KK_t}\equiv q(\phi_t;\tau_1)$, and hence $\gamma(\phi_t;\tau_2)>\gamma(\phi_t;\tau_1)$. It remains to be shown that

$\int_t^T\left(\rho+P\left(\frac{1-\tau_2}{1-\tau_1}b(\phi_s;\tau_1);\tau_2\right)\right)\frac{1-\tau_2}{1-\tau_1}b(\phi_s;\tau_1)K_{1,s}e^{-\rho(s-t)}ds <$
$\int_t^T(\rho+P(b(\phi_s;\tau_1);\tau_1))b(\phi_s;\tau_1)K_{1,s}e^{-\rho(s-t)}ds$. Since $\frac{1-\tau_2}{1-\tau_1}<1$ and $b(\phi_s;\tau_1)K_{1,s}>0$, it suffices to show that $P\left(\frac{1-\tau_2}{1-\tau_1}b(\phi_s;\tau_1);\tau_2\right)<P(b(\phi_s;\tau_1);\tau_1)$, which is equivalent to $\int_{v(\phi,\tau_2)<\frac{1-\tau_2}{1-\tau_1}b(\phi_s;\tau_1)}dF(\phi)<\int_{v(\phi,\tau_1)<b(\phi_s;\tau_1)}dF(\phi)$ and is satisfied because $V(\phi_t,K_t;\tau_2)>\frac{1-\tau_2}{1-\tau_1}V(\phi_t,K_t;\tau_1)$. Finally, note that $\frac{1-\tau_2}{1-\tau_1}b(\phi_s;\tau_1)$ is feasible with $\tau=\tau_2$ because $\frac{1-\tau_2}{1-\tau_1}b(\phi_s;\tau_1)\leq\frac{1-\tau_2}{1-\tau_1}v(\phi_s;\tau_1)<v(\phi_s;\tau_2)$.

Proof of Statement 4: Statements 1, 2, and 3 directly imply Statement 4. ■

Proof. of Proposition 10. Use the definition of $A(b;\alpha)$ from equation (35) and the definition of $M(\phi_t;\alpha)$ from equation (39), along with $\overline{v}\equiv\int v(\phi)\,dF(\phi)$, to rewrite equation (38) as

$$v(\phi_t;\alpha)=\frac{M(\phi_t;\alpha)+(1-\alpha)\lambda\int_{v(\phi;\alpha)<b(\phi_t)}v(\phi;\alpha)\,dF(\phi)+\lambda\int_{v(\phi;\alpha)\geq b(\phi_t)}v(\phi;\alpha)\,dF(\phi)}{\rho+\delta+\lambda}, \tag{50}$$

which implies

$$v\left(\phi_{t_j};\alpha\right)=\frac{M\left(\phi_{t_j};\alpha\right)+\lambda E_t\left\{(1-\alpha)^{[n(t_{j+1})-n(t_j)]}v(\phi;\alpha)\right\}}{\rho+\delta+\lambda}, \tag{51}$$

where $n(t_j)$ is the number of defaults before or at time t_j.

A solution of the recursive equation in equation (51) is

$$v(\phi_t;\alpha) = E_t\left\{\sum_{j=0}^{\infty}(1-\alpha)^{n(t_j)}\left(\frac{\lambda}{\rho+\delta+\lambda}\right)^j \frac{M\left(\phi_{t_j};\alpha\right)}{\rho+\delta+\lambda}\right\}. \tag{52}$$

Since the conditional density of t_1 as of time t is $\lambda e^{-\lambda(t_1-t)}$,

$$E_t\left\{e^{-(\rho+\delta)(t_1-t)}\right\} = \frac{\lambda}{\rho+\delta+\lambda}, \tag{53}$$

and since $M(\phi_s;\alpha) = M\left(\phi_{t_j};\alpha\right)$ for $t_j \leq s < t_{j+1}$,

$$E_t\left\{\int_{t_j}^{t_{j+1}} e^{-\delta(s-t)} M(\phi_s;\alpha)\, e^{-\rho(s-t_j)} ds\right\} = \frac{M\left(\phi_{t_j};\alpha\right)}{\rho+\delta+\lambda}. \tag{54}$$

Equation (53) implies that

$$E_t\left\{e^{-(\rho+\delta)(t_j-t)}\right\} = \left(\frac{\lambda}{\rho+\delta+\lambda}\right)^j. \tag{55}$$

Use equations (54) and (55) to rewrite equation (52) as

$$v(\phi_t;\alpha) = E_t\left\{\sum_{j=0}^{\infty}(1-\alpha)^{n(t_j)} e^{-(\rho+\delta)(t_j-t)} \int_{t_j}^{t_{j+1}} e^{-\delta(s-t_j)} M(\phi_s;\alpha)\, e^{-\rho(s-t_j)} ds\right\}. \tag{56}$$

Since $n(s) = n(t_j)$ for $t_j \leq s < t_{j+1}$, equation (56) implies

$$v(\phi_t;\alpha) = E_t\left\{\int_t^{\infty}(1-\alpha)^{n(s)} M(\phi_s;\alpha)\, e^{-(\rho+\delta)(s-t)} ds\right\}. \tag{57}$$

■

Proof. of Proposition 11.

Define the operator T as

$Tv(\phi_t;\alpha) = \max_{b\leq v(\phi_t;\alpha),\gamma} \frac{(1-\tau)[\phi_t-[p_K\gamma+c(\gamma)]]+\lambda\int v(\phi;\alpha)dF(\phi|\phi_t)+A(b;\alpha|\phi_t)}{\rho+\lambda+\delta-\gamma}$. First, I will prove that T maps non-decreasing functions into strictly increasing functions. Then I will prove that T is a contraction by proving that it satisifies monotonicity and discounting.

Mapping non-decreasing functions into strictly increasing functions: Assume that $v(\phi;\alpha)$ is non-decreasing in ϕ. Therefore, for $\phi_{t,2} > \phi_{t,1}$, $v\left(\phi_{t,2};\alpha\right) \geq v\left(\phi_{t,1};\alpha\right)$ so $b\left(\phi_{t,1}\right) \leq v\left(\phi_{t,1};\alpha\right) \leq v\left(\phi_{t,2};\alpha\right)$ and hence $b\left(\phi_{t,1}\right)$ is feasible when $\phi_t = \phi_{t,2}$.

Substituting the definition of $\eta(\phi_t;\alpha)$ from equation (32) into the definition of $A(b;\alpha|\phi_t)$ in equation (35), evaluating the resulting equation at $b = b(\phi_t) = \arg\max_{b \leq v(\phi_t)} A\left(b;\alpha|\phi_{t,2}\right)$ and then substituting the resulting equation into equation (34) yields

$$Tv(\phi_t;\alpha) = \frac{(1-\tau)\{\phi_t - [p_K\gamma(\phi_t) + c(\gamma(\phi_t))]\} + \int g(\phi, b(\phi_t);\alpha)\, dF(\phi)}{\rho + \delta + \lambda - \gamma(\phi_t)}$$

where

$$g(\phi, b;\alpha) \equiv \begin{array}{ll} \tau\rho b + \lambda\tau b + (1-\tau)(1-\alpha)\lambda v(\phi;\alpha), & \text{if } v(\phi;\alpha) < b \\ \tau\rho b + \lambda v(\phi;\alpha), & \text{if } v(\phi;\alpha) \geq b \end{array} \tag{58}$$

and $g(\phi, b;\alpha)$ is non-decreasing in ϕ.[20]

Therefore, $Tv\left(\phi_{t,2};\alpha\right) \geq \frac{(1-\tau)\left[\phi_{t,2} - \left[p_K\gamma\left(\phi_{t,1}\right) + c\left(\gamma\left(\phi_{t,1}\right)\right)\right]\right] + \int g\left(\phi, b\left(\phi_{t,1}\right);\alpha\right) dF\left(\phi|\phi_{t,2}\right)}{\rho+\lambda+\delta-\gamma\left(\phi_{t,1}\right)} > \frac{(1-\tau)\left[\phi_{t,1} - \left[p_K\gamma\left(\phi_{t,1}\right) + c\left(\gamma\left(\phi_{t,1}\right)\right)\right]\right] + \int g\left(\phi, b\left(\phi_{t,1}\right);\alpha\right) dF\left(\phi|\phi_{t,2}\right)}{\rho+\lambda+\delta-\gamma\left(\phi_{t,1}\right)} \geq \frac{(1-\tau)\left[\phi_{t,1} - \left[p_K\gamma\left(\phi_{t,1}\right) + c\left(\gamma\left(\phi_{t,1}\right)\right)\right]\right] + \int g\left(\phi, b\left(\phi_{t,1}\right);\alpha\right) dF\left(\phi|\phi_{t,1}\right)}{\rho+\lambda+\delta-\gamma\left(\phi_{t,1}\right)} = Tv\left(\phi_{t,1}\right)$, where the first inequality from the fact that $\gamma\left(\phi_{t,1}\right)$ and $b\left(\phi_{t,1}\right)$ are feasible when $\phi_t = \phi_{t,2}$; the second (strict) inequality follows from $\phi_{t,2} > \phi_{t,1}$; and the third inequality follows from the facts that $g(\phi, b;\alpha)$ is non-decreasing in ϕ and $F\left(\phi|\phi_{t,2}\right)$ first-order stochastically dominates $F\left(\phi|\phi_{t,1}\right)$. Therefore T maps non-decreasing functions into strictly increasing functions.

Monotonicity: Suppose that $v_2(\phi_t;\alpha) \geq v_1(\phi_t;\alpha)$. Therefore, $Tv_2(\phi_t;\alpha) \geq \frac{(1-\tau)\left[\phi_t - \left[p_K\gamma^{(1)}(\phi_t) + c\left(\gamma^{(1)}(\phi_t)\right)\right]\right] + \lambda\int v_2(\phi;\alpha) dF(\phi|\phi_t) + A_2\left(b^{(1)}(\phi_t);\alpha|\phi_t\right)}{\rho+\lambda+\delta-\gamma^{(1)}(\phi_t)} = \frac{(1-\tau)\left[\phi_t - \left[p_K\gamma^{(1)}(\phi_t) + c\left(\gamma^{(1)}(\phi_t)\right)\right]\right] + \int g_2\left(\phi, b^{(1)}(\phi_t);\alpha\right) dF(\phi|\phi_t)}{\rho+\lambda+\delta-\gamma^{(1)}(\phi_t)}$ and $Tv_1(\phi_t;\alpha) = \frac{(1-\tau)\left[\phi_t - \left[p_K\gamma^{(1)}(\phi_t) + c\left(\gamma^{(1)}(\phi_t)\right)\right]\right] + \int g_1\left(\phi, b^{(1)}(\phi_t);\alpha\right) dF(\phi|\phi_t)}{\rho+\lambda+\delta-\gamma^{(1)}(\phi_t)}$, where $\gamma^{(1)}(\phi_t)$ and $b^{(1)}(\phi_t)$ attain the maximum of $\frac{(1-\tau)[\phi_t - [p_K\gamma + c(\gamma)]] + \int g_1(\phi,b;\alpha) dF(\phi|\phi_t)}{\rho+\lambda+\delta-\gamma}$. Therefore, $Tv_2(\phi_t;\alpha) - Tv_1(\phi_t;\alpha) \geq \frac{\int\left[g_2\left(\phi, b^{(1)}(\phi_t);\alpha\right) - g_1\left(\phi, b^{(1)}(\phi_t);\alpha\right)\right] dF(\phi|\phi_t)}{\rho+\lambda+\delta-\gamma^{(1)}(\phi_t)}$. To prove that $Tv_2(\phi_t;\alpha) - Tv_1(\phi_t;\alpha) \geq 0$, it suffices

[20] To show that $g(\phi, b;\alpha)$ is non-decreasing in ϕ, it suffices to show that $\tau\rho b + \lambda\tau b + (1-\tau)(1-\alpha)\lambda b \leq \tau\rho b + \lambda b$, or equivalently that $0 \leq (1-\tau)\lambda b - (1-\tau)(1-\alpha)\lambda b = \alpha(1-\tau)\lambda b$, which is satisfied because $\alpha > 0$, $\tau < 1$, $\lambda > 0$, and $b \geq 0$.

to prove that $g_2\left(\phi, b^{(1)}(\phi_t);\alpha\right) - g_1\left(\phi, b^{(1)}(\phi_t);\alpha\right) \geq 0$ for all ϕ in the support of ϕ. The definition of $g(\phi, b; \alpha)$ implies

$$\begin{array}{c} g_2(\phi, b; \alpha) \\ -g_1(\phi, b; \alpha) \end{array} = \begin{array}{ll} (1-\tau)(1-\alpha)\lambda[v_2(\phi;\alpha) - v_1(\phi;\alpha)] \geq 0, & \text{if } v_2(\phi;\alpha) < b \\ \lambda v_2(\phi;\alpha) - [\lambda\tau b + (1-\tau)(1-\alpha)\lambda v_1(\phi;\alpha)], & \text{if } v_1(\phi;\alpha) < b \leq v_2(\phi;\alpha) \\ \lambda[v_2(\phi;\alpha) - v_1(\phi;\alpha)] \geq 0, & \text{if } v_1(\phi;\alpha) \geq b \end{array}.$$

To show that $g_2(\phi, b;\alpha) - g_1(\phi, b;\alpha) \geq 0$ when $v_1(\phi;\alpha) < b \leq v_2(\phi;\alpha)$, observe that when $v_1(\phi;\alpha) < b \leq v_2(\phi;\alpha)$, $g_2(\phi, b;\alpha) - g_1(\phi, b;\alpha) = \lambda v_2(\phi;\alpha) - [\lambda\tau b + (1-\tau)(1-\alpha)\lambda v_1(\phi;\alpha)]$ $> \lambda v_2(\phi;\alpha) - [\lambda\tau b + (1-\tau)(1-\alpha)\lambda b] = \lambda(v_2(\phi;\alpha) - [\tau b + (1-\tau)b - \alpha(1-\tau)b]) = \lambda(v_2(\phi;\alpha) - [1-\alpha(1-\tau)]b) \geq 0$. Thus, the operator T is monotonic.

Discounting: Define $g^{(a)}\left(\phi, b^{(a)}(\phi_t);\alpha\right)$ to be the function $g(\phi, b;\alpha)$ defined in equation (58) with $v(\phi;\alpha)$ replaced by $v(\phi;\alpha) + a$, so that

$$g^{(0)}\left(\phi, b^{(a)}(\phi_t) - a;\alpha\right) \equiv \begin{array}{ll} \left[\begin{array}{c} \tau\rho\left(b^{(a)}(\phi_t) - a\right) \\ +\lambda\tau\left(b^{(a)}(\phi_t) - a\right) \\ +(1-\tau)(1-\alpha)\lambda v(\phi;\alpha) \end{array}\right], & \text{if } v(\phi;\alpha) < b^{(a)}(\phi_t) - a \\ \left[\begin{array}{c} \tau\rho\left(b^{(a)}(\phi_t) - a\right) \\ +\lambda v(\phi;\alpha) \end{array}\right], & \text{if } v(\phi;\alpha) \geq b^{(a)}(\phi_t) - a \end{array}$$

and

$$g^{(a)}\left(\phi, b^{(a)}(\phi_t);\alpha\right) \equiv \begin{array}{ll} \left[\begin{array}{c} \tau\rho b^{(a)}(\phi_t) + \lambda\tau b^{(a)}(\phi_t) \\ +(1-\tau)(1-\alpha)\lambda(v(\phi;\alpha) + a) \end{array}\right], & \text{if } v(\phi;\alpha) + a < b^{(a)}(\phi_t) \\ \tau\rho b^{(a)}(\phi_t) + \lambda(v(\phi;\alpha) + a), & \text{if } v(\phi;\alpha) + a \geq b^{(a)}(\phi_t) \end{array}.$$

For $a \geq 0$, $[T(v+a)](\phi_t) = \frac{(1-\tau)\left[\phi_t - \left[p_K\gamma^{(a)}(\phi_t) + c\left(\gamma^{(a)}(\phi_t)\right)\right]\right] + \int g^{(a)}\left(\phi, b^{(a)}(\phi_t);\alpha\right)dF(\phi|\phi_t)}{\rho+\lambda+\delta-\gamma^{(a)}(\phi_t)}$. Since $b^{(a)}(\phi_t) - a$ is feasible under $v(\phi)$, $Tv(\phi_t) \geq \frac{(1-\tau)\left[\phi_t - \left[p_K\gamma^{(a)}(\phi_t) + c\left(\gamma^{(a)}(\phi_t)\right)\right]\right] + \int g^{(0)}\left(\phi, b^{(a)}(\phi_t) - a;\alpha\right)dF(\phi|\phi_t)}{\rho+\lambda+\delta-\gamma^{(a)}(\phi_t)}$. Therefore, $[T(v+a)](\phi_t) - Tv(\phi_t) \leq \frac{\int\left[g^{(a)}\left(\phi, b^{(a)}(\phi_t);\alpha\right) - g^{(0)}\left(\phi, b^{(a)}(\phi_t) - a;\alpha\right)\right]dF(\phi|\phi_t)}{\rho+\lambda+\delta-\gamma^{(a)}(\phi_t)}$, where

$$\begin{array}{c} g^{(a)}\left(\phi, b^{(a)}(\phi_t);\alpha\right) \\ -g^{(0)}\left(\phi, b^{(a)}(\phi_t) - a;\alpha\right) \end{array} = \begin{array}{ll} \tau\rho a + [1-\alpha(1-\tau)]\lambda a, & \text{if } v(\phi;\alpha) + a < b^{(a)}(\phi_t) \\ \tau\rho a + \lambda a, & \text{if } v(\phi;\alpha) + a \geq b^{(a)}(\phi_t) \end{array}.$$

Therefore, $[T(v+a)](\phi_t) - Tv(\phi_t) \leq \beta a$, where $\beta \equiv \frac{\tau\rho+\lambda}{\rho+\lambda+\delta-\gamma^{(a)}(\phi_t)} = \frac{\rho+\lambda-(1-\tau)\rho}{\rho+\lambda-\left(\gamma^{(a)}(\phi_t)-\delta\right)} < 1$, since $(1-\tau)\rho > \left(\gamma^{(a)}(\phi_t) - \delta\right)$. Therefore, T satisfies the discounting property and hence T is a contraction. Therefore, $v(\phi)$ is strictly increasing. ■

B Calculation of Value Function in Equation (14)

Since K_t grows at the constant rate $\gamma_t - \delta$ from t to t_1, $K_{t_1} = e^{(\gamma_t - \delta)(t_t - t)} K_t$ and hence

$$V\left(\phi_{t_1}, K_{t_1}\right) = v\left(\phi_{t_1}\right) K_{t_1} = v\left(\phi_{t_1}\right) e^{(\gamma_t - \delta)(t_t - t)} K_t. \tag{59}$$

Substituting equation (59) into the Bellman equation in equation (14), using the independence of profitability ϕ across regimes, and combining several of the terms in b_t yields

$$V\left(\phi_t, K_t\right) = \max_{b_t \leq v(\phi_t), \gamma_t} K_t E_t \left\{ \begin{array}{c} b_t \\ + \left[\begin{array}{c} (1-\tau)\left[\phi_t - \left[p_K \gamma_t + c\left(\gamma_t\right)\right]\right] \\ + \left[\gamma_t - \delta - (1-\tau)\left(\rho + p_t\right)\right] b_t \end{array} \right] H\left(t, t_1\right) \\ + e^{-(\rho+\delta-\gamma_t)(t_1 - t)} \max\left[\left(v\left(\phi_{t_1}\right) - b_t\right), 0\right] \end{array} \right\}. \tag{60}$$

To calculate the expectation on the right hand side of equation (60), first calculate $E_t\left\{e^{-(\rho+\delta-\gamma_t)(t_1-t)}\right\}$ and $E_t\left\{H\left(t, t_1\right)\right\}$. The random variable inside these expectations is the date t_1 at which the next new regime arrives. The density of t_1 is $\lambda e^{-\lambda(t_1 - t)}$ so[21]

$$E_t\left\{e^{-(\rho+\delta-\gamma_t)(t_1-t)}\right\} = \frac{\lambda}{\rho + \delta + \lambda - \gamma_t} \tag{61}$$

and

$$E_t\left\{H\left(t, t_1\right)\right\} = \frac{1}{\rho + \delta + \lambda - \gamma_t}. \tag{62}$$

Now substitute equations (61) and (62) into the Bellman equation in equation (60), and rearrange, to obtain

$$V\left(\phi_t, K_t\right) = \max_{b_t \leq v(\phi_t), \gamma_t} \frac{K_t}{\rho + \lambda + \delta - \gamma_t} \left\{ \begin{array}{c} \left[\begin{array}{c} (1-\tau)\left[\phi_t - \left[p_K \gamma_t + c\left(\gamma_t\right)\right]\right] \\ + \left[\tau\rho + \lambda - (1-\tau) p_t\right] b_t \end{array} \right] \\ + \lambda E_t\left\{\max\left[\left(v\left(\phi_{t_1}\right) - b_t\right), 0\right]\right\} \end{array} \right\}. \tag{63}$$

Finally, use $p_t = \lambda F\left(v^{-1}\left(b_t\right)\right)$ and $E_t\left\{\max\left[\left(v\left(\phi_{t_1}\right) - b_t\right), 0\right]\right\} = \int_{v(\phi) \geq b_t} \left(v(\phi) - b_t\right) dF(\phi)$

[21] Equation (61) is derived as $E_t\left\{e^{-(\rho+\delta-\gamma_t)(t_1-t)}\right\} = \int_t^\infty \lambda e^{-\lambda(t_1-t)} e^{-(\rho+\delta-\gamma_t)(t_1-t)} dt_1 = \frac{\lambda}{\rho+\delta+\lambda-\gamma_t}$. Equation (62) is derived as $E_t\left\{H\left(t, t_1\right)\right\} = \int_t^\infty \lambda e^{-\lambda(t_1-t)} \left\{\frac{1 - e^{-(\rho+\delta-\gamma_t)(t_1-t)}}{\rho+\delta-\gamma_t}\right\} dt_1 = \frac{1}{\rho+\delta-\gamma_t}\left[1 - \frac{\lambda}{\rho+\delta+\lambda-\gamma_t}\right] = \frac{1}{\rho+\delta+\lambda-\gamma_t}$.

in equation (63) and rearrange to obtain

$$V(\phi_t, K_t) = \max_{b_t \le v(\phi_t), \gamma_t} \frac{K_t}{\rho + \lambda + \delta - \gamma_t} \left\{ \begin{array}{c} (1-\tau)[\phi_t - [p_K \gamma_t + c(\gamma_t)]] \\ +\lambda \overline{v} \\ +A(b_t) \end{array} \right\} \quad (64)$$

where $\overline{v} = \int v(\phi)\, dF(\phi)$ and

$$A(b_t) \equiv \tau\left[\rho + \lambda F\left(v^{-1}(b_t)\right)\right] b_t - \lambda \int_{v(\phi) < b_t} v(\phi)\, dF(\phi). \quad (65)$$

Therefore, $v(\phi_t)$ defined implicitly in equation (15) is

$$v(\phi_t) = \max_{b_t \le v(\phi_t), \gamma_t} \frac{1}{\rho + \lambda + \delta - \gamma_t} \left\{ \begin{array}{c} (1-\tau)[\phi_t - [p_K \gamma_t + c(\gamma_t)]] \\ +\lambda \overline{v} \\ +A(b_t) \end{array} \right\}. \quad (66)$$

C Derivation of Value Function When Lenders Recover Some of Firm Value in Default

Substitute the interest payments from equation (33) for the interest payments $(\rho + p_t) b_t K_t$ in the Bellman equation in (14) and use $K_{t_1} = e^{(\gamma_t - \delta)(t_1 - t)} K_t$ and $B_{t_1^-} = b_t K_{t_1}$ to obtain

$$V(\phi_t, K_t; \alpha) = \max_{b_t \le \frac{V(\phi_t, K_t; \alpha)}{K_t}, \gamma_t} E_t \left\{ \begin{array}{c} (1-\tau)\{\phi_t - [p_K \gamma_t + c(\gamma_t)] - (\rho b_t + \eta(b_t; \alpha))\} H(t, t_1) \\ +[1 + (\gamma_t - \delta) H(t, t_1)] b_t \\ +e^{-(\rho + \delta - \gamma_t)(t_1 - t)} \max\left[v\left(\phi_{t_1}; \alpha\right) - b_t, 0\right] \end{array} \right\} K_t. \quad (67)$$

Substitute the optimal values $\gamma(\phi_t)$ and $b(\phi_t)$ for γ_t and b_t, respectively, and use the expressions for $E_t\{H(t, t_1)\}$ and $E_t\left\{e^{-(\rho+\delta-\gamma_t)(t_1-t)}\right\}$ in equations (62) and (61) respectively, to obtain

$$V(\phi_t, K_t; \alpha) = v(\phi_t; \alpha) K_t$$

where

$$v(\phi_t;\alpha) = \frac{1}{\rho+\delta+\lambda-\gamma_t} E_t \left\{ \begin{array}{c} (1-\tau)\{\phi_t - [p_K \gamma(\phi_t) + c(\gamma(\phi_t))] - (\rho b(\phi_t) + \eta(b_t;\alpha))\} \\ +(\rho+\lambda) b(\phi_t) \\ +\lambda \max\left[v\left(\phi_{t_1};\alpha\right) - b(\phi_t), 0\right] \end{array} \right\}. \tag{68}$$

Now substitute $\int_{v(\phi;\alpha)\geq b(\phi_t)} [v(\phi;\alpha) - b(\phi_t)]\, dF(\phi)$ for $E_t\left\{\max\left[v\left(\phi_{t_1};\alpha\right) - b(\phi_t), 0\right]\right\}$, use the definition $\overline{v}(\alpha) = \int v(\phi;\alpha)\, dF(\phi)$ in equation (36) and rearrange to obtain

$$v(\phi_t;\alpha) = \frac{1}{\rho+\delta+\lambda-\gamma_t} E_t \left\{ \begin{array}{c} (1-\tau)\{\phi_t - [p_K \gamma(\phi_t) + c(\gamma(\phi_t))]\} \\ +\tau[\rho b(\phi_t) + \eta(b_t;\alpha)] - \eta(b_t;\alpha) + \lambda\overline{v}(\alpha) \\ -\lambda \int_{v(\phi;\alpha)<b(\phi_t)} v(\phi)\, dF(\phi) + \lambda \int_{v(\phi;\alpha)<b(\phi_t)} b(\phi_t)\, dF(\phi) \end{array} \right\}. \tag{69}$$

Then use the definition of $\eta(b_t;\alpha)$ in equation (32) to obtain

$$v(\phi_t;\alpha) = \frac{(1-\tau)\{\phi_t - [p_K\gamma(\phi_t) + c(\gamma(\phi_t))]\} + \lambda\overline{v}(\alpha) + A(b;\alpha)}{\rho+\delta+\lambda-\gamma_t}, \tag{70}$$

where

$$A(b;\alpha) \equiv \tau(\rho b + \eta(\phi_t;\alpha)) - \alpha\lambda \int_{v(\phi;\alpha)<b} v(\phi;\alpha)\, dF(\phi). \tag{71}$$