

NBER WORKING PAPER SERIES

OPTIMAL DEBT AND PROFITABILITY IN THE TRADEOFF THEORY

Andrew B. Abel

Working Paper 21548

<http://www.nber.org/papers/w21548>

NATIONAL BUREAU OF ECONOMIC RESEARCH

1050 Massachusetts Avenue

Cambridge, MA 02138

September 2015

I thank Jeff Campbell, Vincent Glode, Joao Gomes, Christian Goulding, Richard Kihlstrom, Bob Mc-Donald, Adriano Rampini, Scott Richard, Michael Roberts, Ali Shourideh, Toni Whited and seminar participants at the Federal Reserve Bank of Chicago, Kellogg Finance, MIT Finance, Shanghai Advanced Institute of Finance, and Wharton Finance, and the Tepper/LAEF Conference on Macro-Finance and the Tel Aviv University Finance Conference for helpful comments and discussion. I also thank the editor, associate editor, and two referees for helpful comments. Part of this research was conducted while I was a Visiting Scholar at Chicago Booth and I thank colleagues there for their hospitality and support. The views expressed herein are those of the author and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2015 by Andrew B. Abel. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Optimal Debt and Profitability in the Tradeoff Theory
Andrew B. Abel
NBER Working Paper No. 21548
September 2015, Revised January 2017
JEL No. G3,G32

ABSTRACT

I develop a dynamic model of leverage with tax deductible interest and an endogenous cost of default. The interest rate includes a premium to compensate lenders for expected losses in default. A borrowing constraint is generated by lenders' unwillingness to lend an amount that would trigger immediate default. When the borrowing constraint is not binding, the tradeoff theory of debt holds: optimal debt equates the marginal tax shield and the marginal expected cost of default. Contrary to conventional interpretation, but consistent with empirical findings, increases in current or future profitability reduce the optimal leverage ratio when the tradeoff theory holds.

Andrew B. Abel
The Wharton School
University of Pennsylvania
2315 Steinberg Hall - Dietrich Hall
Philadelphia, PA 19104-6367
and NBER
abel@wharton.upenn.edu

The tradeoff theory of capital structure is the longest standing theory of capital structure¹ and underlies much of the large body of empirical work that studies capital structure. In the tradeoff theory, the optimal amount of debt equates the marginal benefit of a dollar of debt arising from the tax deductibility of interest payments with the marginal cost of a dollar of debt arising from increased exposure to default. This framework implies that changes in leverage over time, or variation in leverage across firms, can be attributed to differences in the marginal interest tax shield and/or differences in the marginal cost of default. The tradeoff theory has been conventionally interpreted² to imply that more profitable firms should have higher leverage ratios—a prediction that is contrary to the empirical fact³ that more profitable firms tend to have lower leverage ratios.⁴ Notable exceptions to this conventional interpretation are presented in quantitative models by Hennessy and Whited (2005) and Strebulaev (2007). In the stochastic dynamic model presented here, I analytically demonstrate an alternative explanation of the negative relationship between profitability and leverage that is found empirically.

In this paper, I develop and analyze a model of capital structure that incorporates an interest tax shield as well as the possibility of default. In some situations, the optimal amount of debt will be determined by the equality of the marginal benefit of debt arising from the interest tax shield and the marginal cost of debt associated with increased probability of default—that is, optimal debt will be characterized by the tradeoff theory. However, in other situations within the model, optimal debt will not be characterized by the equality of marginal benefit and marginal cost that epitomizes the tradeoff theory. Because the model allows for situations in which the tradeoff theory holds and situations in which it does not hold, the model has the potential to guide empirical tests by including both a null hypothesis in terms of the tradeoff theory and an alternative hypothesis that offers an explanation of leverage other than the tradeoff theory. In particular, I show that the model developed here accommodates situations in which higher profitability (either current profitability or expected future profitability) is associated with lower leverage, specifically, a lower value of

¹Robichek and Myers (1966), Kraus and Litzenberger (1973), and Scott (1976).

²Scott (1976), Fama and French (2002), Frank and Goyal (2007).

³Fama and French (2002).

⁴For 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."

market leverage, which is the ratio of debt to the market value of the firm. Furthermore, if the probability of default is nonzero, these situations arise when and only when the tradeoff theory is operative. That is, the empirical finding that more profitable firms tend to have lower leverage ratios, which has been viewed by others as evidence against the tradeoff theory, is viewed as evidence in favor of the tradeoff theory when seen through the lens of the model presented here.

As the tradeoff theory has developed over the past half century, it has become increasingly complex, especially in empirical structural models of the firm that are designed to capture realistic features of a firm's environment.⁵ The model I develop here is stripped of these complexities so that I can focus on its new features and implications in a framework that admits analytic results without relying on numerical solution. The model's biggest departure from standard models of debt concerns the maturity of the debt. Many standard models of debt⁶ assume that debt has infinite maturity and pays a fixed coupon over the infinite future, or until the firm defaults. Clearly, the assumption of infinite maturity is extreme, but it has been used productively over the years. I also make an extreme assumption about maturity, but in the opposite direction. I assume that debt must be repaid an instant after it is issued. The standard specification with infinite maturity can be viewed as the limiting case of long-term debt and my specification with zero maturity can be viewed as the limiting case of short-term debt, such as commercial paper, much of which has a maturity of only 1 to 4 days,⁷ as well as overnight repurchase agreements.

The specification of zero-maturity debt is motivated by two considerations. First, this specification makes salient the recurrent nature of the financing decision, in contrast to the once-and-for-all financing decision in many models of debt.⁸ At each instant of time, the firm decides whether to repay its debt or to default, and if it decides to repay its debt, it chooses the amount of debt to issue anew. Because I do not include any flotation, issuance, or adjustment costs, the amount of debt issued responds immediately and completely to changes in the firm's environment, and will not have the rich dynamics documented empirically and

⁵Hennessy and Whited (2007).

⁶Modigliani and Miller (1958), Leland (1994), and Gorbenko and Strebulaev (2010).

⁷In March 2016, the average daily issuance of nonfinancial AA commercial paper was 210 issues amounting to \$5.3 billion. 53% of these issues, accounting for 55% of the dollar volume, had a maturity of 1-4 days. Source: Board of Governors of the Federal Reserve System, Volume Statistics for Commercial Paper Issuance, Data as of April 12, 2016.

⁸Merton (1974), Leland (1994), and Gorbenko and Strebulaev (2010).

analyzed by Leary and Roberts (2005). Second, zero-maturity debt is always valued at par, which alleviates the need to calculate the value of debt that would arise with long-term debt. Therefore, the firm's decision about whether to default on its debt, which depends on a comparison of the total value of the firm and the value of the firm's debt, becomes transparent.

Because firms have the opportunity to default on their debt, rational lenders need to take account of the probability of default, as well as their losses in the event of default, in order to determine the appropriate interest rate on their loans to the firm. In this paper, risk-neutral lenders have the same information that the firm's shareholders have and they require a premium above the riskless rate in order to compensate for the expected losses in the event of default. In addition, if the amount of debt were sufficiently large, it would trigger immediate default. In the current framework with zero-maturity debt, lenders avoid being subject to immediate default by refusing to lend an amount greater than the contemporaneous value of the firm, which itself depends on the amount of debt issued.

An important component of the firm's financing decision is the stochastic process for the firm's pre-tax pre-interest cash flow, that is, EBIT (earnings before interest and taxes). I specify a continuous-time continuous-state process for EBIT. Instead of a diffusion process, as in Leland (1994) for example, I specify a Markov process in which EBIT remains unchanged for a random length of time and then a new value of EBIT arrives at dates governed by a Poisson process with arrival intensity λ . On these dates, the value of EBIT changes by a discrete amount, and a discrete decrease in the value of EBIT can lead the firm to default on its debt. In the event of default, a fraction $\alpha > 0$ of the firm's value disappears as a deadweight loss and the creditors take ownership of the remaining fraction $1 - \alpha$ of the firm. I do not consider re-negotiation between shareholders and creditors. Shareholders and creditors have common knowledge. If they were to re-negotiate when a new realization of EBIT would lead to default, the result of that re-negotiation would simply be a function of the amount of debt and the level of EBIT at that time. But if shareholders and lenders were to agree on terms that would apply when a new value of EBIT would lead to default, they could simply write those terms into the debt contract. But then the debt contract would become state-contingent in a way that would make the instrument not look like debt. Since I want to focus on the leverage ratio as measured using conventional debt, I do not

consider re-negotiation or the sort of "pre-negotiation" just described.⁹

For simplification, the new values of EBIT are i.i.d across arrivals. Nevertheless, EBIT displays persistence: EBIT remains constant during each regime, and regimes have a mean duration of $\frac{1}{\lambda}$, which could potentially be quite large. The unconditional correlation between two values of EBIT that occur at dates $x > 0$ units of time apart is $e^{-\lambda x}$, which is positive and declines monotonically in x . The advantage of using the i.i.d. specification is that it allows for persistence in EBIT as well as analytic tractability.¹⁰ In particular, as I will show, the optimal amount of debt is invariant to contemporaneous EBIT if the value of EBIT is high enough that tradeoff theory of debt is operative. As a result, the leverage ratio, which is the ratio of debt to the value of the firm, is a decreasing function of EBIT, since the value of the firm is increasing in EBIT. The stark result that optimal debt is invariant to EBIT when the tradeoff theory is operative is a consequence of the assumption that EBIT is i.i.d. across regimes. In Section 7, I extend the model to allow for persistence in EBIT across regimes. I show that optimal debt would no longer be invariant to EBIT when the tradeoff theory is operative. I present a simple example in which the conditional distribution of EBIT is uniform and I show that in this example, optimal debt and the leverage ratio are both decreasing in contemporaneous EBIT, thereby reinforcing the major result of the paper. I also show that the conditional probability of default is a decreasing function of EBIT, when and only when the tradeoff theory is operative.

When EBIT is i.i.d. across regimes, the firm's optimal capital structure depends on whether EBIT exceeds or falls short of an endogeneously determined critical value. For values of EBIT that exceed the critical value, the tradeoff theory is operative, provided that the firm faces a positive probability of default. When the tradeoff theory is operative, the optimal level of debt is invariant to EBIT; since an increase in EBIT increases the total value

⁹The debt contract that I analyze is not an optimal contract, but it is of the form commonly used in practice and in analytical models. Although it is not optimal, it does have the important feature that all participants in the debt instrument do so voluntarily and are not made worse off by participating. Biais, Mariotti, Plantin, and Rochet (2007) derive "performance sensitive" debt as an optimal contract in a dynamic framework.

¹⁰In their analysis of the impact of temporary shocks on optimal issuance of debt, Gorbenko and Strebulaev (2010) specify temporary shocks as Poisson arrivals from a given distribution; when they consider multiple shocks over time, they assume that they are i.i.d., as in the framework used here. The temporary shocks in Gorbenko and Strebulaev are drawn from a double exponential distribution; the shocks in the current model are drawn from a more general distribution on finite support for which the density is finite and non-decreasing and the restriction in equation (2) is satisfied.

of the firm, the optimal leverage ratio falls as profitability, measured by EBIT, increases. This negative relationship between profitability and the leverage ratio, which arises from the invariance of debt to EBIT, is reminiscent of Strebulaev’s (2007) finding that when adjustment costs prevent the level of debt from changing, an increase in profitability reduces the leverage ratio. However, the mechanisms that make debt invariant to profitability are different in the current model and in Strebulaev’s model. Specifically, the invariance of optimal debt arises in the current model in the complete absence of any flotation, issuance, or adjustment costs that prevent debt from adjusting in Strebulaev’s model.

For values of EBIT below the critical value, the borrowing constraint that prevents immediate default binds. In this situation, the tradeoff theory is not operative and the optimal leverage ratio is invariant to EBIT. To summarize, the model predicts a negative relationship between the optimal leverage ratio and contemporaneous EBIT (which is consistent with empirical findings) when the tradeoff theory is operative but not when the constraint on the firm’s borrowing is binding.

To examine the relationship between the optimal amount of debt and *expected future* profitability, I analyze the impact of a rightward translation of the unconditional distribution of profitability. This favorable shift of the unconditional distribution of profitability increases the continuation value of the firm, which the firm’s current shareholders stand to lose in the event of default; therefore, a favorable shift of the unconditional distribution increases the cost of default facing these shareholders.¹¹ In response to the increased cost of default, the firm adjusts its capital structure to reduce the probability of default. Whether this change in capital structure reduces the optimal amount of debt depends on the unconditional distribution of profitability. In the focal case of a uniform distribution, a rightward translation of the distribution reduces optimal debt when the tradeoff theory is operative.¹² Thus, whether one analyzes the relationship between the optimal leverage ratio and *contem-*

¹¹To the extent that creditors recover a portion of their loans in the event of default, the social cost of default is smaller than the loss to current shareholders. I assume that creditors receive a fraction $1 - \alpha$ of the continuation value of the firm in the event of default, so the social cost of default is a fraction α of the continuation value. Therefore, the social cost of default, as well as private cost to current shareholders, is increased by a favorable shift of the unconditional distribution of profitability.

¹²Appendix D in the online appendix shows that this finding holds more broadly within the class of truncated exponential distributions for density functions that do not slope upward too steeply. Only for sufficiently steeply upward-sloping density functions will optimal debt increase in response to a rightward translation of the distribution function when the tradeoff theory is operative.

poraneous EBIT (as described above) or between the optimal leverage ratio and *long-run average* EBIT, the model predicts a negative relationship between the optimal leverage ratio and EBIT if the tradeoff theory is operative, but not if the firm faces a binding constraint on its borrowing.

Section 1 presents the economic environment facing a firm, including the opportunity to borrow and the ability to default on outstanding debt. The availability and terms of loans to the firm depend on the valuation of the firm. Section 2 characterizes the optimal level of debt and the associated value of shareholders' equity, which together sum to the total valuation of the firm. An important component of this analysis is the critical value of current EBIT, which is the boundary between low values of EBIT for which the borrowing constraint binds and high values of EBIT for which the borrowing constraint does not bind. Depending on whether this critical value of EBIT is at the minimum value of the support of EBIT, the maximum value of the support of EBIT, or in between, the firm will find itself in one of three scenarios, which are characterized in Section 3. Although optimal behavior in two of the three scenarios can be derived very easily, optimal behavior in one of the scenarios (denoted as Scenario II) entails more extensive analysis, which is presented in Section 4. Of particular interest is subsection 4.1, which interprets the analytic results in terms of the tradeoff theory. Section 5 demonstrates that Scenario I, II, or III will prevail depending on whether the tax rate is low, intermediate, or high, and it provides explicit expressions for the values of the tax rate associated with the boundaries between adjacent scenarios. Section 6 analyzes the impact of a rightward translation of the unconditional distribution of EBIT on optimal debt, shareholder equity, the critical value of EBIT, and the probability of default. Section 7 extends the model to allow for persistence in EBIT across regimes. In the focal case of a uniform distribution, persistence strengthens the major result of the paper that the optimal leverage ratio is a decreasing function of EBIT if the tradeoff theory is operative. Concluding remarks are presented in Section 8. To avoid disruption in the narrative flow of the paper, the proofs of all lemmas, propositions, and corollaries are in Appendix A. Appendix B presents the details of the calculation of the value of the firm. An online appendix presents closed-form solutions for optimal debt and the value of the firm in the special case in which the unconditional distribution of EBIT is uniform and creditors recover zero in the event of default. It also contains a section analyzing the impact of rightward translation of the unconditional distribution of EBIT when the distribution is a

truncated exponential distribution.

1 The Firm's Economic Environment

Let $\phi(t)$ be EBIT, the pre-tax net cash flow from operations, before interest, at time t . The realizations of $\phi(t)$ are generated by an exogenous stochastic process and thus are independent of any actions taken by the firm. If EBIT at time t_0 is $\phi(t_0)$, then $\phi(t)$ remains equal to $\phi(t_0)$ until some random date $t_1 > t_0$. At time t_1 , a new value of EBIT, $\phi(t_1)$, is drawn from a distribution with c.d.f. $F(\phi)$, with $F(\Phi_L) = 0$, $F(\Phi_H) = 1$, and finite density $f(\phi) = F'(\phi) > 0$ everywhere on the support $[\Phi_L, \Phi_H]$, where $-\infty < \Phi_L < \Phi_H < \infty$. In addition, assume that $f'(\phi) \geq 0$ for all ϕ in the support of $F(\phi)$, which will ensure that the firm's objective function is concave in debt. The unconditional expected value of EBIT is assumed to be positive, that is, $E\{\phi\} = \int_{\Phi_L}^{\Phi_H} \phi dF(\phi) > 0$, which implies $\Phi_H > 0$ but does not place a restriction on the sign of Φ_L . The arrival date of a new value of EBIT, that is, the timing of t_1 , is governed by a Poisson process with an instantaneous probability λ of an arrival of a new value of $\phi(t)$. Realizations of new values of EBIT are i.i.d. across regimes.

I assume that the firm is managed on behalf of risk-neutral shareholders with constant rate of time preference equal to $\rho > 0$. The shareholders have "deep pockets" in the sense that they can infuse an unlimited amount of funds into the firm if its after-tax cash flow from operations net of interest is negative. The support of the distribution for EBIT may include negative values of EBIT, since Φ_L may be negative. If a negative realization of EBIT is sufficiently large in absolute value or sufficiently persistent, current shareholders might decide to abandon ownership of the firm to avoid a stream of future cash flows with a negative expected present value. In principle, the shareholders of a firm might abandon ownership of the firm either as a response to unfavorable current and prospective future EBIT or as a means to avoid repaying the firm's debt. My focus in this paper is on leverage and the potential default associated with it, so I will restrict the stochastic process for EBIT to be such that in the complete absence of borrowing, risk-neutral shareholders of the firm would never find it optimal to abandon the firm.¹³ To formalize this restriction, define $W(\phi(t))$ to be the expected present value of current and future cash flows, if the firm is

¹³In contrast to this assumption, Gorbenko and Strebulaev (2010) allow the unlevered value of the firm to become negative, in which case shareholders would abandon the firm.

an all-equity firm that never issues any debt, and continues to operate forever, regardless of the realization of $\phi(t)$. Therefore, $W(\phi(t)) \equiv E_t \left\{ \int_t^\infty (1 - \tau) \phi(s) e^{-\rho(s-t)} ds \right\}$, where $E_t \{ \}$ denotes the expectation conditional on information available at time t . The Poisson arrival process for new values of $\phi(s)$ implies that $E_t \{ \phi(s) \} = e^{-\lambda(s-t)} \phi(t) + (1 - e^{-\lambda(s-t)}) E \{ \phi \}$, for $s \geq t$, so it is straightforward to calculate the value of an all-equity firm that operates forever as

$$W(\phi(t)) = \frac{1 - \tau}{\rho + \lambda} \left[\phi(t) + \frac{\lambda}{\rho} E \{ \phi \} \right]. \quad (1)$$

Henceforth, I assume that

$$\Phi_L > -\frac{\lambda}{\rho} E \{ \phi \}, \quad (2)$$

so that $W(\phi(t)) > 0$ for all $\phi(t) \geq \Phi_L$ and hence an all-equity firm would never find it optimal to cease operation.¹⁴

The firm borrows by issuing bonds to risk-neutral lenders who have the same rate of time preference, ρ , as shareholders. The maturity of these bonds is vanishingly small. If the firm issues an amount D_t of bonds at time t , then it must repay these bonds at time $t + dt$, where dt is infinitesimal. Effectively these bonds have instantaneous maturity. The interest rate on these bonds, $r(t)$, is set at a level that compensates lenders for their expected losses in the event of default. If the firm defaults on its debt, the shareholders lose all ownership interest in the firm and receive zero. The creditors take ownership of the firm in default, but the act of default imposes a deadweight loss equal to a fraction α of the value of the firm, where $0 < \alpha \leq 1$. Therefore, in default at time $t + dt$, creditors receive $(1 - \alpha) V(\phi(t + dt))$, where $V(\phi(t + dt))$ is the value of the firm, which is described more formally in equation (7) below. Since creditors lose the value of the bonds, D_t , but recover $(1 - \alpha) V(\phi(t + dt))$ in default, the expected loss to creditors due to default in the interval from t to $t + dt$ is $\lambda \int_{V(\phi(t+dt)) < D_t} [D_t - (1 - \alpha) V(\phi(t + dt))] dF(\phi(t + dt))$. Therefore, the interest rate, $r(t)$, that compensates lenders for the expected loss in default is

$$r(t) = \rho + \lambda \int_{V(\phi) < D_t} \left[1 - (1 - \alpha) \frac{V(\phi)}{D_t} \right] dF(\phi) \geq \rho, \quad (3)$$

¹⁴Leland (1994, p. 1217) effectively makes the same assumption by specifying a stochastic process for the "asset value," which corresponds to $W(\phi(t)) \equiv E_t \left\{ \int_t^\infty (1 - \tau) \phi(s) e^{-\rho(s-t)} ds \right\}$, that is bounded away from zero.

where the first term, ρ , is the riskless rate in the absence of any chance of default, and the second term is the expected loss to creditors in default for each dollar of debt they hold.

In addition to requiring an interest rate that includes compensation for the expected loss in default, potential lenders to the firm will never lend an amount that is so large that the firm would default immediately upon receiving the funds from issuing the bond. This limit on the amount the firm can borrow is

$$D_t \leq V(\phi(t)). \quad (4)$$

I will refer to the limit on borrowing in equation (4) as an endogenous borrowing constraint, though it could also be described as a non-negativity constraint on the value of the firm's equity, or equivalently, as limited liability.

The taxable income of the firm, denoted $y(t)$, is operating profit, $\phi(t)$, minus interest payments, $r(t) D_t$, so using the interest rate in equation (3), taxable income is

$$y(t) = \phi(t) - \rho D_t - \lambda \int_{V(\phi) < D_t} [D_t - (1 - \alpha) V(\phi)] dF(\phi). \quad (5)$$

The tax rate on taxable income is constant and equal to τ , where $0 < \tau < 1$. If taxable income is negative, the firm receives a tax rebate of $-\tau y(t) > 0$.

I assume that shareholders do not retain any earnings in the firm, that is, the firm pays out all net cash flows as dividends, $dX(t)$, so

$$dX(t) = (1 - \tau) y(t) dt + dD_t, \quad (6)$$

where dD_t is the net increase in bonds at time t , which represents a net inflow of funds to the firm. The notation $dX(t)$ and dD_t allows for dividends and net increases in bonds to be discrete amounts, rather than flows, at a point of time. For instance, as I show later, the arrival of a new regime that increases $\phi(s)$ by a discrete amount may increase the amount of bonds by a discrete amount at a point of time, so that dD_t and $dX(t)$ are both positive.

The value of the firm at time t is

$$V(\phi(t)) = \max_{dD_s} E_t \left\{ \int_t^T e^{-\rho(s-t)} dX(s) \right\}, \quad (7)$$

where T is the date at which the shareholders decide to default on their debt, and hence is endogenous. $V(\phi(t))$ in equation (7) is the value of the firm at time t if it arrives at time t with zero bonds outstanding. Equivalently, $V(\phi(t))$ is the value of the firm at time t immediately after it has paid off (with interest) its outstanding debt issued at $t - dt$, and before it issues new debt at time t .

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

1. The firm arrives at time t with D_{t-} bonds outstanding.
2. The values of profitability, $\phi(t)$, and the value of the firm $V(\phi(t))$ 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 $V(\phi(t)) < D_{t-}$, the firm defaults on its bonds, creditors receive $(1 - \alpha)V(\phi(t))$ and shareholders receive nothing.
 - (b) If $V(\phi(t)) \geq D_{t-}$, the firm repays its bonds, and shareholders retain ownership.
4. If shareholders retain ownership, the firm issues bonds D_t subject to the borrowing constraint $D_t \leq V(\phi(t))$.
5. The firm receives $\phi(t)$, pays interest $r(t)D_t$, and pays taxes $\tau[\phi(t) - r(t)D_t]$.
6. The firm pays dividends $dX(t) = (1 - \tau)[\phi(t) - r(t)D_t] + dD_t$, where¹⁵ $dD_t = D_t - D_{t-}$.

2 Optimal Debt, Firm Value, and Shareholder Equity

Consider the firm at time t immediately after it has repaid its bonds D_{t-dt} . At this instant, the firm has no outstanding bonds and it decides how many bonds to issue at time t to

¹⁵As it turns out, the firm will change the amount of outstanding debt only at times of regime change, when $\phi(s)$ changes. At these times, the amount of debt will change by a discrete amount, dD_t .

achieve the maximum on the right hand side of equation (7). The firm's optimization problem can be represented by the following Bellman equation

$$V(\phi(t)) = \max_{D_s \leq V(\phi(s))} E_t \left\{ \int_t^{t_1} e^{-\rho(s-t)} dX(s) + e^{-\rho(t_1-t)} \max \left[V(\phi(t_1)) - D_{t_1^-}, 0 \right] \right\}, \quad (8)$$

where t_1 is the date of the first arrival of a new regime after time t . The value of the firm in equation (8) is the sum of two terms. The first term is the expected present value of dividends, dX , over the interval of time from t to the arrival of a new regime at time t_1 . The second term is the expected present value of the firm at time t_1 , after observing the new value of profitability, $\phi(t_1)$. When the firm observes $\phi(t_1)$, it can choose to become a firm with value $V(\phi(t_1))$, but to do so, it must repay its outstanding debt $D_{t_1^-}$, so the net value to the shareholders of continuing their ownership in the firm is $V(\phi(t_1)) - D_{t_1^-}$. Provided this net value of continued ownership of the firm is non-negative, shareholders will repay their outstanding bonds, and then issue a new amount of bonds, D_{t_1} . Otherwise, the firm will default on its bonds, and shareholders lose their ownership of the firm.

Other than deciding whether to default, the only decision that the firm makes is how many bonds to issue at each point in time. The structure of the stochastic process for $\phi(t)$ implies that the optimal value of D_t is simply a function of contemporaneous $\phi(t)$. I will use the notation $\widehat{D}(\phi(t))$ to denote the optimal value of D_t . Profitability is constant, and equal to $\phi(t)$, over the interval of time from t until t_1 , so after-tax income, $(1 - \tau)y(s)$, and the optimal amount of bonds outstanding, $\widehat{D}(\phi(s))$, are constant and equal to $(1 - \tau)y(t)$ and $\widehat{D}(\phi(t))$, respectively, for all s in $[t, t_1]$. With $\widehat{D}(\phi(s))$ constant over this interval of time, $dD_s = 0$ for all s in (t, t_1) , but $dD_t = \widehat{D}(\phi(t))$ for a firm at time t after it has repaid its debt $D_{t_1^-}$.

It is straightforward but tedious to calculate the value of the firm, so I have relegated the derivation to Appendix B, where I show that

$$V(\phi(t)) = \frac{1}{\rho + \lambda} \max_{D_t \leq V(\phi(t))} [(1 - \tau)\phi(t) + \lambda\bar{v} + A(D_t)], \quad (9)$$

where

$$A(D) \equiv \tau \left[\rho + \lambda \int_{V(\phi) < D} dF(\phi) \right] D - [\alpha + \tau(1 - \alpha)] \lambda \int_{V(\phi) < D} V(\phi) dF(\phi), \quad (10)$$

and $\bar{v} \equiv \int V(\phi) dF(\phi)$ is the unconditional mean of $V(\phi)$. To interpret the function $A(D)$ in equation (10), which arises naturally in the derivation of the value of the firm, use equation (3) to rewrite equation (10) as

$$A(D_t) = \tau r(t) D_t - \alpha \lambda \int_{V(\phi) < D_t} V(\phi) dF(\phi). \quad (11)$$

I will refer to $A(D)$ as the "tradeoff function" because it contains the elements of the tradeoff theory of debt. The first term on the right hand side of the expression for $A(D_t)$ in equation (11), $\tau r(t) D_t$, is the interest tax shield, which is simply the tax rate, τ , multiplied by interest payments, $r(t) D_t$. The second term is the expected deadweight loss associated with default, where the deadweight loss is $\alpha V(\phi)$ if default occurs when profitability equals ϕ .

Equation (9) can be expressed in terms of flows at time t by multiplying both sides by $\rho + \lambda$ and then subtracting $\lambda V(\phi(t))$ from both sides to obtain the Hamilton-Jacobi-Bellman equation

$$\rho V(\phi(t)) = \max_{D_t \leq V(\phi(t))} [(1 - \tau) \phi(t) + \lambda (\bar{v} - V(\phi(t))) + A(D_t)]. \quad (12)$$

The left hand side of equation (12), $\rho V(\phi(t))$, is the required return on the firm, which is the product of the rate of return required by shareholders and the value of the firm. The expected return to the firm, on the right hand side, has three components: the after-tax profit from operations, $(1 - \tau) \phi(t)$; the expected change in the value of the firm if the regime changes, which is the product of, λ , the instantaneous probability of a regime change, and $\bar{v} - V(\phi(t))$, which is the expected change in the value of the firm when a new value of ϕ is drawn from its unconditional distribution; and $A(D_t)$, the contribution of bond financing consisting of the interest tax shield, $\tau r(t) D_t$, less the expected deadweight loss associated with default.

The optimal value of debt is the value of D_t that attains the maximum on the right hand side of equation (12). Since D_t enters the maximand on the right hand side of equation (12) only through $A(D)$, the optimal amount of debt is¹⁶

$$\hat{D}(\phi(t)) \equiv \arg \max_{D_t \leq V(\phi(t))} A(D_t). \quad (13)$$

¹⁶Equivalently, the optimal value of D_t attains the maximum on the right hand side of equation (9), which, since $\rho + \lambda > 0$, is also given by equation (13).

Since $A(D_t)$ is independent of $\phi(t)$, the optimal value of debt, $\widehat{D}(\phi(t))$, is independent of $\phi(t)$ if the borrowing constraint, $D_t \leq V(\phi(t))$, is not binding. However, if the borrowing constraint is binding, then $\widehat{D}(\phi(t)) = V(\phi(t))$, which is increasing in $\phi(t)$, since, as shown in Proposition 1 below, $V(\phi(t))$ is strictly increasing in $\phi(t)$.

To obtain the value function, substitute the optimal amount of debt, $\widehat{D}(\phi(t))$, from equation (13) into equation (9), which yields

$$V(\phi(t)) = \frac{(1 - \tau)\phi(t) + \lambda\bar{v} + A(\widehat{D}(\phi(t)))}{\rho + \lambda}. \quad (14)$$

Proposition 1 $V(\phi(t))$ is strictly increasing in $\phi(t)$. Moreover, for any $\phi_2 > \phi_1$ in the support $[\Phi_L, \Phi_H]$, $\frac{V(\phi_2) - V(\phi_1)}{\phi_2 - \phi_1} \geq \frac{1 - \tau}{\rho + \lambda} > 0$.

The proof of Proposition 1 is in Appendix A, but the logic of the proof is straightforward. Suppose that $\phi_2 > \phi_1$. If the firm were to maintain the same level of debt at $\phi(t) = \phi_2$ as at $\phi(t) = \phi_1$, which is $\widehat{D}(\phi_1)$, the value of the firm would increase by $\frac{1 - \tau}{\rho + \lambda}(\phi_2 - \phi_1)$, so that $\widehat{D}(\phi_1)$ would be feasible when $\phi(t) = \phi_2$. Allowing for the possibility that optimal debt changes when EBIT increases to ϕ_2 from ϕ_1 implies that the value of the firm may increase by even more, so $V(\phi(t))$ increases by at least $\frac{1 - \tau}{\rho + \lambda}(\phi_2 - \phi_1)$.

The following proposition is an immediate consequence of the definition of optimal debt, $\widehat{D}(\phi(t))$, in equation (13) and the fact that $V(\phi)$ is strictly increasing (Proposition 1).

Proposition 2 The optimal value of bonds, $\widehat{D}(\phi(t))$, is

1. non-decreasing in the state, $\phi(t)$, and
2. non-decreasing in time, t , outside of default.

Proposition 2 states that the optimal value of debt never falls below previous optimal values of debt as long as the firm does not default. This behavior is consistent with the time path of debt in Goldstein, Ju, and Leland (2001), in which it is *assumed* that, outside of default, the firm does not decrease its debt.¹⁷

¹⁷The second paragraph of Goldstein, Ju, and Leland (2001, p. 483) begins "Below, we consider only the option to increase future debt levels. While in theory management can both increase and decrease future debt levels, Gilson (1997) finds that transactions costs discourage debt reductions outside of Chapter 11."

Since shareholders owe a liability of $\widehat{D}(\phi(t))$, the value of shareholders' equity in the firm, $S(\phi(t))$, is

$$S(\phi(t)) = V(\phi(t)) - \widehat{D}(\phi(t)) \geq 0, \quad (15)$$

where the inequality is the borrowing constraint. Because shareholders receive an amount $\widehat{D}(\phi(t))$ immediately when they borrow, they choose D to maximize $V(\phi(t)) = S(\phi(t)) + D$ rather than to maximize shareholder equity exclusive of D , $S(\phi(t)) = V(\phi(t)) - D$.

The following proposition provides bounds on the optimal amount of bonds. Because the firm has the option to operate forever as an all-equity firm, $V(\phi(t))$ is at least as large as $W(\phi(t))$ in equation (1), and the assumption in equation (2) that $\Phi_L > -\frac{\lambda}{\rho} E\{\phi\}$ implies that $W(\phi(t)) > 0$. Therefore, even the lowest value of the firm, $V(\Phi_L)$, is at least as high as $W(\Phi_L) > 0$. It cannot be optimal to issue an amount of bonds smaller than $V(\Phi_L)$ because to do so would fail to take full advantage of the interest tax shield, without any exposure to the possibility of default. Therefore, the optimal amount of bonds is at least $V(\Phi_L)$. The borrowing constraint, $D \leq V(\phi(t))$, implies that even when the firm attains its highest possible value, the firm can never issue an amount of bonds greater than $V(\Phi_H)$.

Proposition 3 *If $0 < \tau < 1$, then $0 < V(\Phi_L) \leq \widehat{D}(\phi(t)) \leq V(\Phi_H)$.*

Corollary 4 *If $0 < \tau < 1$, then $\widehat{D}(\Phi_L) = V(\Phi_L)$, and $S(\Phi_L) = 0$.*

Corollary 4 follows directly from Proposition 3, which implies that $\widehat{D}(\Phi_L) \geq V(\Phi_L)$, and from the borrowing constraint $\widehat{D}(\Phi_L) \leq V(\Phi_L)$. Therefore, $\widehat{D}(\Phi_L) = V(\Phi_L)$. Hence, when $\phi(t)$ takes on its minimum value Φ_L , the optimal amount of debt equals $V(\Phi_L)$, which implies that shareholders' equity, $S(\Phi_L) = V(\Phi_L) - \widehat{D}(\Phi_L)$, is zero.

Proposition 5 *If (1) $0 < \tau < \frac{\alpha}{1+\alpha}$ and (2) $f(\phi) > 0$ is non-decreasing for all $\phi \in [\Phi_L, \Phi_H]$, then $V(\phi(t))$ is concave in $\phi(t)$ on the domain $[\Phi_L, \Phi_H]$.*

Proposition 5 provides sufficient conditions for the value function, $V(\phi(t))$, to be concave. One of the steps (formalized as Lemma 22 in Appendix A) in the proof of Proposition 5 is to prove that if $V(\phi(t))$ is concave, then the tradeoff function $A(D)$ is strictly concave in D for $V(\Phi_L) \leq D \leq V(\Phi_H)$, which is Corollary 6 below.¹⁸

¹⁸Corollary 6 provides sufficient conditions for $A(D)$ to be strictly concave in D . A necessary condition

Corollary 6 to Proposition 5. If (1) $0 < \tau < \frac{\alpha}{1+\alpha}$ and (2) $f(\phi) > 0$ is non-decreasing for all $\phi \in [\Phi_L, \Phi_H]$, then $A(D)$ is strictly concave in D for $V(\Phi_L) \leq D \leq V(\Phi_H)$.

Define D^* as the value of debt that maximizes the tradeoff function $A(D)$ over $D \in [V(\Phi_L), V(\Phi_H)]$, ignoring the borrowing constraint $D \leq V(\phi(t))$. Formally,

$$D^* \equiv \arg \max_{V(\Phi_L) \leq D \leq V(\Phi_H)} A(D). \quad (16)$$

The strict concavity of $A(D)$ in D implies that D^* is unique. Since, $A(D)$ is independent of $\phi(t)$, D^* is invariant to $\phi(t)$. If $D^* \leq V(\phi(t))$, the borrowing constraint does not bind, and optimal debt, $\widehat{D}(\phi(t))$, equals D^* . However, if $D^* > V(\phi(t))$, the firm cannot borrow as much as D^* and $\widehat{D}(\phi(t)) = V(\phi(t))$.

Define ϕ^* as the critical value of $\phi \in [\Phi_L, \Phi_H]$ above which $V(\phi) > \phi^*$ and below which $V(\phi) < \phi^*$. Formally,

$$\phi^* \equiv V^{-1}(D^*). \quad (17)$$

The following proposition characterizes optimal debt, $\widehat{D}(\phi(t))$, shareholder equity, $S(\phi(t))$, and total firm valuation, $V(\phi(t))$, first for $\phi(t) \geq \phi^*$ and then for $\phi(t) \leq \phi^*$.

Proposition 7 Define ϕ^* so that $V(\phi^*) = D^* \equiv \arg \max_{V(\Phi_L) \leq D \leq V(\Phi_H)} A(D)$. Then

- (1) for any $\phi(t) \geq \phi^*$,
 - (a) $\widehat{D}(\phi(t)) = D^*$,
 - (b) $S(\phi(t)) = \frac{1-\tau}{\rho+\lambda} [\phi(t) - \phi^*]$, and
 - (c) $V(\phi(t)) = \frac{1-\tau}{\rho+\lambda} [\phi(t) - \phi^*] + D^*$;
- (2) for any $\phi(t) \leq \phi^*$,
 - (a) $\widehat{D}(\phi(t)) \equiv V(\phi(t))$ is concave and strictly increasing in $\phi(t)$, and
 - (b) $S(\phi(t)) = 0$.

for this strict concavity is that default imposes a deadweight cost, that is, $\alpha > 0$. To see the necessity of this condition, set $\alpha = 0$ in the definition of $A(D)$ in equation (10) and differentiate twice with respect to D to obtain $A'(D) = \tau \left[\rho + \lambda \int_{V(\phi) < D} dF(\phi) \right]$ and $A''(D) = \tau \lambda V^{-1'}(D) f(V^{-1}(D)) > 0$ because $f(V^{-1}(D)) > 0$ and $V^{-1'}(D) > 0$. Therefore, if $\alpha = 0$, then $A(D)$ is strictly convex in D . Thus, $\alpha > 0$ is a necessary condition for strict concavity of $A(D)$ in D .

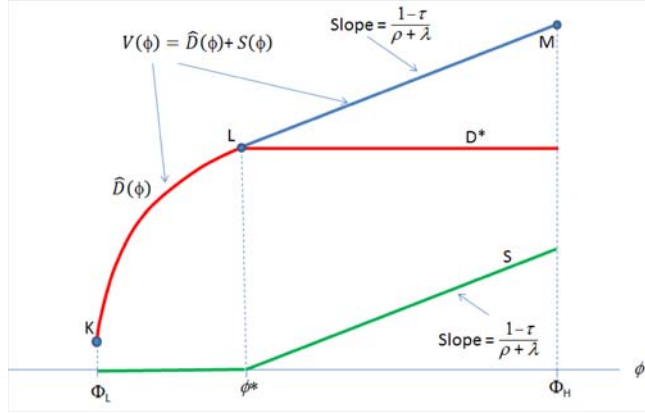


Figure 1: Optimal Debt, Equity and Firm Value

Figure 1 illustrates optimal debt, shareholders' equity, and the total value of the firm, each as a function of ϕ , for a situation in which $V(\Phi_L) < \phi^* < V(\Phi_H)$. For $\phi(t) \leq \phi^*$, the borrowing constraint, $D \leq V(\phi(t))$, is binding so shareholders' equity, $S(\phi(t)) = V(\phi(t)) - \widehat{D}(\phi(t))$, is identically zero. In Figure 1 optimal debt, $\widehat{D}(\phi(t))$, and the value of the firm, $V(\phi(t))$, are shown for $\phi(t) \leq \phi^*$ by the upward-sloping concave curve KL ; shareholders' equity is shown by the horizontal line segment on the horizontal axis for $\phi(t) \leq \phi^*$. For higher values of $\phi(t)$, specifically for $\phi(t) \geq \phi^*$, the borrowing constraint is not binding, and $\widehat{D}(\phi(t)) = D^*$, as shown by the horizontal line segment starting at point L and extending to the right. Shareholders' equity is $S(\phi(t)) = \frac{1-\tau}{\rho+\lambda} [\phi(t) - \phi^*]$, as shown by the line segment starting on the horizontal axis at $\phi = \phi^*$ and extending to the right with slope $\frac{1-\tau}{\rho+\lambda}$. The value of the firm is shown by the line segment LM with slope $\frac{1-\tau}{\rho+\lambda}$.

Corollary 8 Define the optimal leverage ratio as $L(\phi(t)) \equiv \frac{\widehat{D}(\phi(t))}{V(\phi(t))}$. If $\phi(t) \leq \phi^*$, then $L(\phi(t)) \equiv 1$; if $\phi(t) \geq \phi^*$, then $L(\phi(t)) = \frac{1}{\frac{1-\tau}{\rho+\lambda} \frac{\phi(t) - \phi^*}{D^*} + 1}$, which is strictly decreasing in $\phi(t)$.

The tradeoff theory is operative when optimal debt is determined by equating the marginal tax shield associated with interest deductibility and the marginal default cost. Formally, the tradeoff theory is operative when $A'(\widehat{D}(\phi(t))) = 0$ so that $\widehat{D}(\phi(t)) = D^*$. Proposition 7 implies that the tradeoff theory is operative when and only when $\phi(t) \geq \phi^*$. Therefore, Corollary 8 implies that when and only when the tradeoff theory is operative, the optimal leverage ratio is a decreasing function of contemporaneous profitability. This

result is remarkable because, as noted in the introduction to this paper, it is the opposite of the conventional interpretation of the tradeoff theory, which states that the leverage ratio is increasing in profitability. However, empirical studies find a negative relationship between the leverage ratio and productivity, consistent with Corollary 8 when the tradeoff theory is operative.

Proposition 2 states that outside of default, $\widehat{D}(\phi(t))$ is a weakly increasing function of the state $\phi(t)$ and of time t . The leverage ratio, however, is weakly decreasing in the state, $\phi(t)$, and is not monotonic in time t . To see that the leverage ratio can either increase or decrease over time, consider several consecutive regimes in which $\phi(t) \geq \phi^*$. The optimal value of bonds, $\widehat{D}(\phi(t))$, remains constant and equal to D^* over time since $\phi(t) \geq \phi^*$. When $\phi(t)$ increases from one of these regimes to the next regime, $V(\phi(t))$ increases and the leverage ratio falls; when $\phi(t)$ decreases from one of these regimes to the next regime, $V(\phi(t))$ falls and the leverage ratio increases.

Under the optimal debt policy, $\widehat{D}(\phi(t))$, the instantaneous probability of default is

$$P\left(\widehat{D}(\phi(t))\right) \equiv \lambda \int_{V(\phi) < \widehat{D}(\phi(t))} dF(\phi), \quad (18)$$

which is the probability that a new regime arrives that induces a value of the firm lower than the currently outstanding debt, $\widehat{D}(\phi(t))$.

Proposition 9 *The instantaneous probability of default $P\left(\widehat{D}(\phi(t))\right) = \lambda \min [F(\phi(t)), F(\phi^*)]$.*

In the framework analyzed here, default occurs only at times of regime change.¹⁹ If $\phi(t) < \phi^*$, then $\widehat{D}(\phi(t)) = V(\phi(t))$ so default will occur at the next regime change if the new value of profitability is lower than $\phi(t)$ so the new value of the firm is less than $V(\phi(t)) = \widehat{D}(\phi(t))$. Alternatively, if $\phi(t) \geq \phi^*$, then $\widehat{D}(\phi(t)) = D^* = V(\phi^*)$ so that default will occur at the next regime change if the new value of profitability is less than ϕ^* so the new value of the firm is less than $V(\phi^*) = \widehat{D}(\phi(t))$. Thus, regardless of whether $\phi(t)$ is above or below the critical value ϕ^* , default will occur at the next regime change if and only if the new value of profitability is less than $\min[\phi(t), \phi^*]$.

¹⁹The model in Gorbenco and Strebulaev (2010) has a similar feature. Specifically, in their model, default occurs only at the times at which a new value of the temporary shock has a Poisson arrival.

Proposition 10 *The instantaneous probability of default, $P\left(\widehat{D}(\phi(t))\right)$, and the interest rate, $r(t)$, are*

1. *strictly increasing in $\widehat{D}(\phi(t))$*
2. *strictly increasing in $\phi(t)$ if $\phi(t) < \phi^*$*
3. *invariant to $\phi(t)$ if $\phi(t) \geq \phi^*$.*

In principle, there are two endogenous channels that can increase the probability of default at the time, t' , of the next regime change: channel (1) is an increase in the current amount of bonds outstanding, $\widehat{D}(\phi(t))$; and channel (2) is a shift in the distribution of $\phi(t')$ that shifts the distribution of $V(\phi(t'))$ in an unfavorable way. Statement 1 of Proposition 10 captures channel (1). Channel (2) is not operative in the current specification because profitability, ϕ , is i.i.d. across regimes so the distribution of $V(\phi(t'))$ is invariant to $\phi(t)$.

If $\phi(t) < \phi^*$, an increase in current profitability, $\phi(t)$, *increases* the probability of default (Statement 2 of Proposition 10) because it leads the firm to increase its bonds, $\widehat{D}(\phi(t))$, (channel 1) and the increase in current profitability does not improve the distribution of $V(\phi(t'))$. If $\phi(t) \geq \phi^*$, an increase in $\phi(t)$ has no effect on the probability of default. These counter-intuitive results are based on the assumption that $\phi(t)$ is i.i.d. across regimes. In Section 7, I relax this i.i.d. assumption and allow profitability to be persistent across regimes. I show that an increase in $\phi(t)$ can reduce the probability of default, if the tradeoff theory is operative.

The following proposition describes aspects of optimal behavior when the borrowing constraint binds so $\widehat{D}(\phi(t)) = V(\phi(t))$.

Proposition 11 *If $\widehat{D}(\phi(t)) = V(\phi(t))$, then*

1. *after tax income, $(1 - \tau)y(t) \equiv (1 - \tau)\left[\phi(t) - r(t)\widehat{D}(\phi(t))\right] \leq 0$, with strict inequality if $\phi(t) < \Phi_H$*
2. *after-tax income $(1 - \tau)y(t)$ is strictly increasing in $\phi(t)$*
3. *if t is not a time of regime change, then dividends, $dX(t) = (1 - \tau)y(t) \leq 0$.*

Statement 1 of Proposition 11 implies that if the after-tax income of the firm is positive, then the borrowing constraint, $\widehat{D}(\phi(t)) \leq V(\phi(t))$, cannot bind. Statement 2 is that the after-tax income of the firm is strictly increasing in $\phi(t)$ if the borrowing constraint is binding, that is, if $\phi(t) < \phi^*$.²⁰ Statement 3 is that if the borrowing constraint is binding, and if there is not a change in regime, the shareholders of the firm will not receive positive dividends. Negative dividends in this context means that the shareholders inject funds into the firm to cover the negative flow of after-tax income because the binding borrowing constraint prevents the firm from issuing additional debt to cover the negative after-tax income.

3 Three Scenarios

D^* is defined in equation (16) as the value of D that maximizes $A(D)$ over the interval $V(\Phi_L) \leq D \leq V(\Phi_H)$. In this section, I define and analyze three scenarios that are defined by whether the lower bound $D \geq V(\Phi_L)$ strictly binds, the upper bound $D \leq V(\Phi_H)$ strictly binds, or neither bound strictly binds. Formally, Scenario I is the set of configurations of $\lambda, \rho, \alpha, \tau$, and $F(\phi)$ for which the lower bound on D strictly binds, Scenario II is the set of configurations of $\lambda, \rho, \alpha, \tau$, and $F(\phi)$ for which neither bound strictly binds, and Scenario III is the set of configurations of $\lambda, \rho, \alpha, \tau$, and $F(\phi)$ for which the upper bound on D strictly binds.

The definition of D^* in equation (16), $D^* \equiv \arg \max_{V(\Phi_L) \leq D \leq V(\Phi_H)} A(D)$, implies that

$$D^* = \arg \max_D A(D) + \omega_1 (D - V(\Phi_L)) + \omega_2 (V(\Phi_H) - D), \quad (19)$$

where $\omega_1 \geq 0$ is the multiplier on the constraint $V(\Phi_L) \leq D$ and $\omega_2 \geq 0$ is the multiplier on the constraint $D \leq V(\Phi_H)$. The first-order condition for optimal D is

$$A'(D^*) = \omega_2 - \omega_1 \quad (20)$$

²⁰If $\phi(t) \geq \phi^*$, after-tax income, $(1 - \tau) [\phi(t) - r(t) \widehat{D}(\phi(t))]$, is also strictly increasing in $\phi(t)$ because $r(t)$ and $\widehat{D}(\phi(t))$ are invariant to $\phi(t)$ when $\phi(t) \geq \phi^*$. Therefore, after-tax income is strictly increasing in $\phi(t)$, regardless of whether the borrowing constraint binds.

and the complementary slackness conditions are

$$\omega_1 (D^* - V(\Phi_L)) = 0 \quad (21)$$

and

$$\omega_2 (V(\Phi_H) - D^*) = 0. \quad (22)$$

Three distinct scenarios are defined according to whether the multipliers ω_1 and ω_2 are positive or zero:²¹

- Scenario I: $\omega_1 > 0$ and $\omega_2 = 0$
- Scenario II: $\omega_1 = \omega_2 = 0$
- Scenario III: $\omega_1 = 0$ and $\omega_2 > 0$

The remainder of this section will focus on Scenarios I and III. I will analyze Scenario II in more detail in Section 4.

3.1 Scenario I

In Scenario I, $\omega_1 > 0$ and $\omega_2 = 0$, so the first-order condition in equation (20) implies $A'(D^*) = -\omega_1 < 0$. Thus, the marginal interest tax shield, which provides an incentive to increase the amount of debt, is overwhelmed by the marginal expected cost of losing the continuation value of the firm in default. The complementary slackness condition in equation (21) implies $D^* = V(\Phi_L)$. Therefore, the optimal value of debt is $V(\Phi_L)$, which is the highest value of debt that the firm can issue without facing a positive probability of default.

Proposition 12 Define $D_0 \equiv \frac{1}{\rho+\lambda} \left[\Phi_L + \frac{\lambda}{\rho} E\{\phi\} \right]$. In Scenario I (where $\omega_1 > 0$ and $\omega_2 = 0$), for all $\phi(t) \in [\Phi_L, \Phi_H]$,

1. $\widehat{D}(\phi(t)) = D^* = D_0$, which is invariant to $\phi(t)$ and τ
2. $S(\phi(t)) = \frac{1-\tau}{\rho+\lambda} (\phi(t) - \Phi_L)$,

²¹Since $V(\Phi_L) \neq V(\Phi_H)$, ω_1 and ω_2 cannot both be positive.

$$3. V(\phi(t)) = D_0 + \frac{1-\tau}{\rho+\lambda} (\phi(t) - \Phi_L).$$

Proposition 12 states that in Scenario I, the optimal value of debt equals D_0 , which is invariant to the tax rate and invariant to the current value of EBIT. The invariance of optimal debt with respect to the tax rate reflects the fact that *at the margin* the tax shield is so weak that it is completely outweighed by the cost of default. Therefore, shareholders choose not to expose the firm to the probability of default. Nevertheless, because the tax rate τ is positive, the firm takes advantage of the tax shield on interest by choosing the maximal amount of debt that does not risk default. In fact, when $\phi(t) = \Phi_L$, the firm uses the tax shield to completely insulate its value from taxes. That is, when $\phi(t) = \Phi_L$, the ability to issue bonds and deduct interest payments increases the total value of the firm to $V(\Phi_L) = D_0 = \frac{1}{\rho+\lambda} [\Phi_L + \frac{\lambda}{\rho} E\{\phi\}]$ from $\frac{1-\tau}{\rho+\lambda} [\Phi_L + \frac{\lambda}{\rho} E\{\phi\}]$ in the absence of debt.

3.2 Scenario III

In Scenario III, $\omega_1 = 0$ and $\omega_2 > 0$, so the first-order condition in equation (20) implies that $A'(D^*) > 0$. The complementary slackness condition in equation (22) implies that $D^* = V(\Phi_H)$. Therefore, since $A(D)$ is strictly concave in D , $A'(D) > 0$ for all $D \in [V(\Phi_L), V(\Phi_H)]$. The marginal interest tax shield associated with an increase in debt overwhelms the increased exposure to default associated with an increase in debt. Therefore, the firm will issue as much debt as it can, driving shareholders' equity to zero, so that $\widehat{D}(\phi(t)) \equiv V(\phi(t))$ for all $\phi(t) \in [\Phi_L, \Phi_H]$.

Proposition 13 *In Scenario III (where $\omega_1 = 0$ and $\omega_2 > 0$), $V(\Phi_H) = \frac{1}{\rho+\lambda} [\Phi_H + (1-\alpha)\lambda\bar{v}]$.*

The expression for the value of the firm at $\phi(t) = \Phi_H$ depends on \bar{v} as well as on the exogenous parameters ρ , λ , Φ_H , and α . In the special case in which default destroys the entire value of the firm ($\alpha = 1$), Proposition 13 implies $V(\Phi_H) = \frac{1}{\rho+\lambda}\Phi_H$, which is simply the expected value of the flow of Φ_H from the current time until the next regime change, which triggers default because the new value of $V(\phi(t))$ will be less than $\widehat{D}(\Phi_H) = V(\Phi_H)$. In this special case, the firm value, $V(\Phi_H)$, is completely insulated from taxes.

4 Scenario II

Suppose that the firm is in Scenario II, so $\omega_1 = \omega_2 = 0$. Therefore, equation (20) implies that $A'(D^*) = 0$. Differentiating $A(D)$ in equation (10) with respect to D and evaluating $A'(D)$ at $D = D^*$ yields

$$A'(D^*) = \tau [\rho + \lambda F(V^{-1}(D^*))] - (1 - \tau) \alpha \lambda V^{-1'}(D^*) D^* f(V^{-1}(D^*)) = 0. \quad (23)$$

The following lemma provides a simple expression for $V^{-1'}(D^*)$, which appears in the second term in equation (23).

Lemma 14 *If $\omega_1 = \omega_2 = 0$, then $V'(\phi^*) = \frac{1-\tau}{\rho+\lambda}$ and $V^{-1'}(D^*) = \frac{\rho+\lambda}{1-\tau}$, where $\phi^* \equiv V^{-1}(D^*)$.*

Use Lemma 14 to rewrite equation (23) as

$$A'(D^*) = \tau [\rho + \lambda F(V^{-1}(D^*))] - (\rho + \lambda) \alpha \lambda D^* f(V^{-1}(D^*)) = 0. \quad (24)$$

Because $A(D)$ is strictly concave, there is at most one value of D^* that satisfies equation (24).²² For $\phi(t) \geq \phi^* \equiv V^{-1}(D^*)$, the borrowing constraint does not bind and the optimal amount of debt is D^* ; for $\phi(t) < \phi^* \equiv V^{-1}(D^*)$, the borrowing constraint binds and the optimal value of debt equals the value of the firm, $V(\phi(t))$.

4.1 The Tradeoff Theory

When the borrowing constraint is not binding, the optimal amount of debt is D^* , which satisfies the first-order condition, $A'(D^*) = 0$, in equation (24).²³ To interpret that first-

²²There will exist a value of D^* that satisfies equation (24) if $\tau_L \leq \tau \leq \tau_H$, where τ_L and τ_H are defined in Proposition 15.

²³Flotation costs for bonds can lead to a modification of the tradeoff theory. For instance, let $c(D)$ be the cost of issuing D units of bonds, and assume that $c(0) = 0$, $c'(D) \geq 0$, and $c''(D) \geq 0$ for $D \geq 0$. Subtracting these flotation costs from operating profits, the value of the firm in equation (9) can be written as $V(\phi(t)) = \frac{1}{\rho+\lambda} \max_{D_t \leq V(\phi(t))} [(1-\tau)\phi(t) + \lambda \bar{v} + \tilde{A}(D_t)]$, where $\tilde{A}(D_t) \equiv A(D_t) - c(D_t)$ is a "modified tradeoff function." Since $A(D_t)$ is strictly concave and $c(D_t)$ is convex, the modified tradeoff function, $\tilde{A}(D_t)$, is strictly concave. Define $\tilde{D}^* \equiv \max_{D_t \leq V(\Phi_H)} \tilde{A}(D_t)$. Suppose that $c'(V(\Phi_L)) < \tau\rho$, so that if $D_t \leq V(\Phi_L)$, then $\tilde{A}'(D_t) = A'(D_t) - c'(D_t) = \tau\rho - c'(D_t) > 0$, where the first equality follows from the

order condition in terms of the tradeoff theory, I rewrite it as

$$\tau (\rho + \lambda F (\phi^*)) = (\rho + \lambda) \alpha \lambda D^* f (\phi^*), \quad (25)$$

where $\phi^* \equiv V^{-1}(D^*)$. First consider the special case in which $\alpha = 1$ so that default completely destroys the firm. In this case, the interest rate in equation (3) is $\rho + \lambda F (\phi^*)$ so the left hand side of equation (25) is simply $\tau r (t)$, which is the marginal interest tax shield associated with an additional dollar of debt. To see why, suppose that the firm issues an additional dollar of debt at time t_0 , pays interest $\rho + \lambda F (\phi^*)$, then repays the dollar of debt at time $t_0 + dt$ with probability $e^{-\lambda F (\phi^*) dt}$, which is the probability that it is not optimal to default at $t_0 + dt$ or earlier. Therefore, *if the interest rate were to remain unchanged at $\rho + \lambda F (\phi^*)$* , the additional dollar of debt would increase the expected present value of the firm's after-tax cash flow over the next interval dt of time by $1 - (1 - \tau) (\rho + \lambda F (\phi^*)) dt - e^{-\rho dt} e^{-\lambda F (\phi^*) dt}$, which equals the left hand side of equation (25) for small dt . Thus, the left hand side of equation (25) is the marginal interest tax shield associated with an additional dollar of debt.

The right hand side of equation (25) is the marginal cost associated with the increased probability of default resulting from an additional dollar of debt. By increasing the probability of default, a one-dollar increase in debt increases the default premium, and hence increases the interest rate, $\rho + \lambda F (\phi^*)$, paid by the firm. To measure the increase in the probability of default, define $\psi (D)$ as the threshold value of ϕ that triggers default when outstanding debt is D . Formally, $D \equiv V (\psi (D))$ so $1 = V' (\psi (D)) \psi' (D)$. Since $V' (\phi^*) = \frac{1-\tau}{\rho+\lambda}$ (Lemma 14), $\psi' (D^*) = \frac{\rho+\lambda}{1-\tau}$. Therefore, a one-unit increase in D will increase the threshold level of ϕ at which default occurs by $\psi' (D^*) = \frac{\rho+\lambda}{1-\tau}$ and thus increase the probability of default by $\lambda f (\phi^*) \psi' (D^*) = \lambda f (\phi^*) \frac{\rho+\lambda}{1-\tau}$, which, in the case with $\alpha = 1$, increases the interest rate by $\lambda f (\phi^*) \frac{\rho+\lambda}{1-\tau}$, and increases the total flow of *after-tax* interest

definition of $\tilde{A} (D_t)$, the second equality follows from the definition of $A (D)$, which implies $A (D) \equiv \tau \rho D$ for $D \leq V (\Phi_L)$, and the final inequality follows from $c' (V (\Phi_L)) < \tau \rho$ and the convexity of $c (D)$. Therefore, $V (\Phi_L) < \tilde{D}^* \leq V (\Phi_H)$ and for any $\phi (t) \geq V^{-1} (\tilde{D}^*)$, the optimal amount of debt, $\hat{D} (\phi (t))$, equals \tilde{D}^* , and hence is invariant to $\phi (t)$. The optimal amount debt for these values of $\phi (t)$, satisfies $\tilde{A}' (\hat{D} (\phi (t))) = 0$, which can be viewed as a "modified tradeoff theory." Note that when $\tilde{A}' (D) = 0$, $A' (D) = c' (D) > 0$. In this case, the marginal interest tax shield is equated to the marginal expected cost of default, plus the marginal flotation cost, $c' (D)$. When this modified tradeoff theory holds, the optimal leverage ratio will be a declining function of $\phi (t)$, as in the text.

payments at time t_0 by $\lambda f(\phi^*)(\rho + \lambda)D^*$, which is the right hand side of equation (25) when $\alpha = 1$. Therefore, equation (25) represents the equality of the marginal interest tax shield and the marginal cost associated with increased exposure to default, which is the essence of the tradeoff theory.

In the more general case in which $\alpha \leq 1$, the interpretation of equation (25) in terms of the tradeoff theory is more nuanced. Define $R(D, \phi_0)$ as the flow of (pre-tax) interest payments at time t if the amount of outstanding debt is D and if the firm defaults if and only if $\phi < \phi_0$. Thus,

$$R(D, \phi_0) \equiv (\rho + \lambda F(\phi_0))D - \lambda(1 - \alpha) \int_{\Phi_L}^{\phi_0} V(\phi) dF(\phi), \quad (26)$$

so that $R(D_t, \psi(D_t)) = r(t)D_t$, where $r(t)$ is the interest rate in equation (3). If the value of ϕ that triggers default, ϕ_0 , were to remain unchanged, an increase in debt would increase interest payments by $\frac{\partial R(D^*, \phi_0)}{\partial D} = \rho + \lambda F(\phi_0)$. Therefore, when $D = D^*$ and $\phi_0 = \phi^*$, a one-dollar increase in D increases the tax shield associated with interest deductibility by

$$\tau \frac{\partial R(D^*, \phi^*)}{\partial D} = \tau(\rho + \lambda F(\phi^*)). \quad (27)$$

Therefore, the left hand side of equation (25) is the marginal tax shield.

Now consider the impact on interest payments of an increase in ϕ_0 , holding the amount of debt unchanged. Partially differentiating equation (26) with respect to ϕ_0 , evaluating this derivative at $\phi_0 = \phi^*$, and using $V(\phi^*) = D^*$, yields

$$\frac{\partial R(D^*, \phi^*)}{\partial \phi_0} = \alpha \lambda f(\phi^*) D^*. \quad (28)$$

Therefore, equation (28) implies that if ϕ_0 were to increase by $\psi'(D^*)$, as would be the case if D were to increase by one dollar, after-tax interest payments would increase by

$$(1 - \tau) \frac{\partial R(D^*, \phi^*)}{\partial \phi_0} \psi'(D^*) = (\rho + \lambda) \alpha \lambda f(\phi^*) D^*. \quad (29)$$

The left hand side of equation (29) is the marginal expected default cost, measured in flow terms, that reflects the increased probability of default when D increases by one dollar.

Specifically, a one-dollar increase in D increases the default threshold ϕ_0 by $\psi'(D^*) = \frac{\rho+\lambda}{1-\tau}$, thereby increasing total interest costs by $\frac{\partial R(D^*, \phi^*)}{\partial \phi_0} \psi'(D^*)$ and after-tax interest costs by $(1-\tau) \frac{\partial R(D^*, \phi^*)}{\partial \phi_0} \psi'(D^*)$, which is the left hand side of equation (29). The right hand side of equation (29) is identical to the right hand side of equation (25), so the right hand side of equation (25) is the marginal expected cost of default associated with a one-dollar increase in D . Thus, equation (25) is a statement of the tradeoff theory, equating the value of the interest tax shield associated with an additional dollar of debt, expressed per unit of time, and the marginal cost of increased exposure to default resulting from this increased debt, also expressed per unit of time.

4.2 Firm in Scenario II with $\phi(t) \leq \phi^*$

Suppose that the firm is in Scenario II and that $\phi(t) \leq \phi^*$. In this case, the firm issues as much debt as it can, driving the equity value to zero, so $V(\phi(t)) = \widehat{D}(\phi(t))$ and $S(\phi(t)) = 0$. Substitute $\widehat{D}(\phi(t))$ for $V(\phi(t))$ in equation (14), then multiply both sides of the resulting equation by $\rho + \lambda$ and subtract $A(\widehat{D}(\phi(t)))$ from both sides to obtain

$$(\rho + \lambda) \widehat{D}(\phi(t)) - A(\widehat{D}(\phi(t))) = (1 - \tau) \phi(t) + \lambda \bar{v}, \quad \text{if } \phi(t) \leq \phi^*. \quad (30)$$

Differentiate both sides of equation (30) with respect to $\phi(t)$ to obtain the ordinary differential equation (ODE)^{24,25}

$$\left(\rho + \lambda - A'(\widehat{D}(\phi(t))) \right) \widehat{D}'(\phi(t)) = 1 - \tau. \quad (31)$$

The boundary condition for the ODE in equation (31), which takes the form of a value-matching condition, is

$$\widehat{D}(\phi^*) = D^*. \quad (32)$$

²⁴This ODE is solved in the Online Appendix in Section E.

²⁵Differentiate equation (31) with respect to $\phi(t)$ to obtain $\left(\rho + \lambda - A'(\widehat{D}(\phi(t))) \right) \widehat{D}''(\phi(t)) - A''(\widehat{D}(\phi(t))) \left(\widehat{D}'(\phi(t)) \right)^2 = 0$. Differentiating $A(D)$ in equation (10) with respect to D implies that $A'(D) < \tau(\rho + \lambda) < \rho + \lambda$ so that $\widehat{D}''(\phi(t))$ has the same sign as $A''(\widehat{D}(\phi(t)))$, which is negative if $\tau < \frac{\alpha}{1+\alpha}$ and $f'(\phi) \geq 0$.

Evaluate equation (31) at $\phi(t) = \phi^*$ and use $A'(D^*) = 0$ and the boundary condition in equation (32) to obtain

$$\widehat{D}'(\phi^*) = \frac{1 - \tau}{\rho + \lambda}. \quad (33)$$

Since equation (33) was derived from the ODE that holds for $\phi \leq \phi^*$, the derivative in equation (33) is actually the left-hand derivative. Since $V(\phi(t)) \equiv \widehat{D}(\phi(t))$ for $\Phi_L \leq \phi(t) \leq \phi^*$, equation (33) implies that the left-hand derivative of $V(\phi(t))$ at $\phi(t) = \phi^*$ is $\frac{1-\tau}{\rho+\lambda}$. Proposition 7 implies that the right-hand derivative of $V(\phi(t))$ at $\phi(t) = \phi^*$ is also $\frac{1-\tau}{\rho+\lambda}$. Therefore, the right-hand and left-hand derivatives of $V(\phi(t))$ at $\phi(t) = \phi^*$ are equal to each other.

The value-matching condition in equation (32) is illustrated in Figure 1 at point L , where the curve through K and L , which represents $V(\phi)$ for $\phi \leq \phi^*$, meets the line segment LM , which represents $V(\phi)$ for $\phi \geq \phi^*$. As discussed, the left-hand and right-hand derivatives of $V(\phi)$ are equal to each other at $\phi = \phi^*$, so the meeting of the curve through K and L and the line segment LM is smooth, that is, differentiable, at point L .²⁶

5 Threshold Tax Rates

The tax rate, τ , determines whether the firm is in Scenario I, II, or III. For low values of $\tau > 0$, the marginal interest tax shield is so small that it is overwhelmed by the marginal expected default cost, and hence the firm is in Scenario I, where it does not expose itself to any chance of default. For high values of τ , the marginal interest tax shield is so strong that the firm always borrows as much as lenders are willing to lend, and hence the firm is in Scenario III. For intermediate values of τ , defined in the following proposition, the firm is in Scenario II, where the tradeoff theory is operative if $\phi(t) \geq \phi^*$, and the borrowing constraint binds if $\phi(t) < \phi^*$.

²⁶This smooth meeting of the curve through K and L and the line segment LM is superficially similar to the smooth-pasting condition that arises in optimal stopping problems with an underlying diffusion process. In those problems, the value-matching and smooth-pasting conditions are two separate boundary conditions that pin down two parameters in the solution. In the current framework, the fundamental stochastic variable has finite variation whereas in optimal stopping problems with a diffusion process, the stochastic variable has infinite variation. In the current framework, equality of the left-hand and right-hand derivatives of $V(\phi(t))$ at $\phi = \phi^*$ arises as a consequence of the value-matching condition. That equality does not impose any additional structure or restriction on the solution.

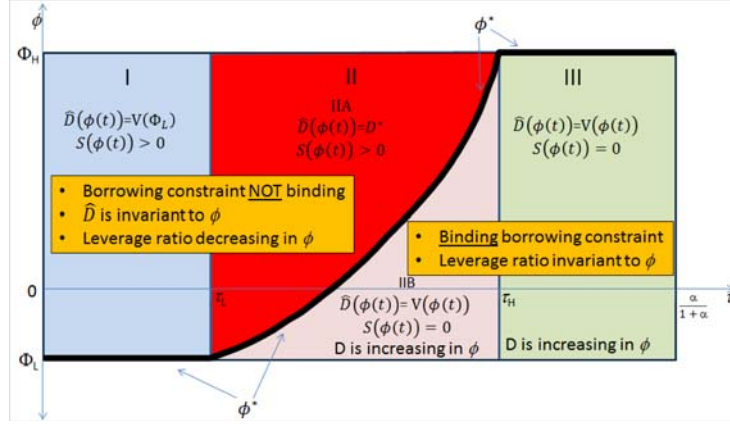


Figure 2: Three Scenarios

Proposition 15 Define $\tau_L \equiv \alpha \frac{\lambda}{\rho} f(\Phi_L) \left(\Phi_L + \frac{\lambda}{\rho} E\{\phi\} \right) > 0$ and $\tau_H \equiv \alpha \frac{\lambda}{\rho + \lambda} f(\Phi_H) (\Phi_H + (1 - \alpha) \lambda \bar{v}) > 0$. Suppose $\tau_L < \tau_H$.²⁷

- (1) If $\tau < \tau_L$, then the firm is in Scenario I and $\widehat{D}(\phi(t)) = \frac{1}{\rho + \lambda} \left[\Phi_L + \frac{\lambda}{\rho} E\{\phi\} \right]$ for all $\phi(t) \in [\Phi_L, \Phi_H]$.
- (2) If $\tau_L \leq \tau \leq \tau_H$, then the firm is in Scenario II and
 - (a) $\widehat{D}(\phi(t)) = D^*$ for $\phi(t) \geq \phi^*$
 - (b) $\widehat{D}(\phi(t)) = V(\phi(t))$ for $\phi(t) \leq \phi^*$.
- (3) If $\tau_H < \tau < \frac{\alpha}{1 + \alpha}$, then the firm is in Scenario III and $\widehat{D}(\phi(t)) = V(\phi(t))$ for all $\phi(t) \in [\Phi_L, \Phi_H]$.

Figure 2 illustrates the three scenarios and displays the behavior of debt and the leverage ratio in each scenario. The tax rate τ is measured along the horizontal axis and ϕ is measured

²⁷The interval $\tau_L \leq \tau \leq \tau_H$ is non-vacuous if $\frac{\tau_L}{\tau_H} \leq 1$. Using the definitions of τ_L and τ_H , $\frac{\tau_L}{\tau_H} = \frac{\frac{\lambda}{\rho} (\Phi_L + \frac{\lambda}{\rho} E\{\phi\})}{\frac{\lambda}{\rho + \lambda} (\Phi_H + (1 - \alpha) \lambda \bar{v})} \frac{f(\Phi_L)}{f(\Phi_H)} \leq \frac{\frac{\lambda}{\rho} (\Phi_L + \frac{\lambda}{\rho} E\{\phi\})}{\frac{\lambda}{\rho + \lambda} \Phi_H} \frac{f(\Phi_L)}{f(\Phi_H)}$, where the inequality follows from $(1 - \alpha) \lambda \bar{v} \geq 0$. Therefore, $\frac{\tau_L}{\tau_H} \leq (1 + \gamma) \frac{\Phi_L + \gamma E\{\phi\}}{\Phi_H} \frac{f(\Phi_L)}{f(\Phi_H)}$, where $\gamma \equiv \frac{\lambda}{\rho}$. As an example, if $f(\phi)$ is the density of a uniform distribution, then $\frac{f(\Phi_L)}{f(\Phi_H)} = 1$ and $E\{\phi\} = \frac{1}{2} (\Phi_L + \Phi_H)$, so $\frac{\tau_L}{\tau_H} \leq \frac{1}{2} (1 + \gamma) \frac{(2\Phi_L + \gamma(\Phi_L + \Phi_H))}{\Phi_H} = \frac{1}{2} (1 + \gamma) \left((2 + \gamma) \frac{\Phi_L}{\Phi_H} + \gamma \right)$. A sufficient condition for $\frac{\tau_L}{\tau_H} \leq 1$ in this example is $\frac{1}{2} (1 + \gamma) \left((2 + \gamma) \frac{\Phi_L}{\Phi_H} + \gamma \right) \leq 1$, which is satisfied if, for instance, $\Phi_L \leq 0$ and $(1 + \gamma) \gamma \leq 2$, which is satisfied for $0 \leq \gamma \leq 1$.

along the vertical axis. Scenario I prevails for $\tau < \tau_L$, Scenario II prevails for $\tau_L \leq \tau \leq \tau_H$, and Scenario III prevails for $\tau_H < \tau < \frac{\alpha}{1+\alpha}$. The ordinate of each point on the thick line that is horizontal at $\phi = \Phi_L$ in Scenario I, upward sloping in Scenario II, and horizontal at $\phi = \Phi_H$ in Scenario III is the value of ϕ^* for each value of τ . Everywhere above this line, that is, in Scenario I and in the upper portion of Scenario II, which is labelled IIA, $\phi(t) > \phi^*$ so the borrowing constraint is not binding and $\widehat{D}(\phi(t)) = D^*$. Since $\widehat{D}(\phi(t))$ is invariant to ϕ and $V(\phi(t))$ is strictly increasing in ϕ , the optimal leverage ratio, $L(\phi(t)) \equiv \frac{\widehat{D}(\phi(t))}{V(\phi(t))}$, is strictly decreasing in ϕ throughout Scenarios I and IIA. Everywhere below this line, that is, in the lower portion of Scenario II, which is labelled IIB, and in Scenario III, $\phi(t) < \phi^*$, so the borrowing constraint is binding and the leverage ratio is invariant to ϕ .

6 Effect of a Rightward Translation of $F(\phi)$

I have examined the relationship between optimal debt and *contemporaneous* profitability, $\phi(t)$. Now I turn attention to the relationship between optimal debt and the prospects for *future* profitability represented by the distribution function $F(\phi)$. Specifically, I examine an improvement in future prospects for profitability that is a rightward translation of the distribution $F(\phi)$ to a new distribution, indexed by $m \geq 0$, $G_m(\phi) \equiv F(\phi - m)$ on the support $[\Phi_L(m), \Phi_H(m)]$, where $\Phi_L(m) = \Phi_L + m$, $\Phi_H(m) = \Phi_H + m$ and $[\Phi_L, \Phi_H]$ is the support of the original distribution $F(\phi)$. For a given value of m , let $V(\phi(t); m)$ be the value of the firm when $\phi = \phi(t)$, let $\widehat{D}(\phi(t); m)$ be the optimal value of debt when $\phi = \phi(t)$, let $P(\phi(t); m)$ be the instantaneous probability of default when $\phi = \phi(t)$, let $A(D; m)$ be the tradeoff function $A(D)$, and let $\phi^*(m)$ and $D^*(m)$ be the values of ϕ^* and D^* , respectively.

The following proposition states that a rightward translation of the distribution $F(\phi)$ increases the value of the firm, $V(\phi(t); m)$, for any $\phi(t)$ in the intersection of the supports of the original distribution and the new distribution.

Proposition 16 *For $0 < m < \Phi_H - \Phi_L$ and any $\phi(t) \in [\Phi_L + m, \Phi_H]$, $V(\phi(t); m) > V(\phi(t); 0)$.*

6.1 Effect of a Rightward Translation of $F(\phi)$ in Scenario I

In Scenario I, the firm issues the amount of debt shown in Proposition 12, which is the highest amount of debt that it can issue without exposing itself to a chance of default.

Proposition 17 *In scenario I,*

1. $\frac{d\widehat{D}(\phi(t);m)}{dm} = \frac{1}{\rho} > 0$.
2. $\frac{dS(\phi(t);m)}{dm} = -\frac{1-\tau}{\rho+\lambda} < 0$.
3. $\frac{dV(\phi(t);m)}{dm} = \frac{\tau\rho+\lambda}{\rho(\rho+\lambda)} > 0$.
4. $\frac{dL(\phi(t);m)}{dm} > 0$.

A rightward translation of the distribution $F(\phi)$ by one unit increases EBIT by one unit in every state, which increases the expected present value of EBIT by $\frac{1}{\rho}$, since the firm will never default and hence will receive a stream of EBIT forever. Therefore, the optimal amount of debt increases by $\frac{1}{\rho}$ units (Statement 1). A rightward translation of $F(\phi)$ reduces the value of equity *for any given* $\phi(t)$ because the increase in the firm's debt outweighs the increase in total firm value resulting from improved future prospects of the firm (Statement 2).²⁸ That is, the decline in the value of equity for a given $\phi(t)$ is smaller than the increase in the optimal amount of debt, so that, consistent with Proposition 16, a rightward translation of $F(\phi)$ increases the total value of the firm for any given $\phi(t)$ (Statement 3). Finally, the increase in debt and the decrease in the value of equity increase the leverage ratio for any given $\phi(t)$ (Statement 4).

6.2 Effect of a Rightward Translation of $F(\phi)$ in Scenario II

In Scenario II, $\omega_1 = \omega_2 = 0$ so that $\frac{\partial A(D^*(m);m)}{\partial D} = 0$, where $D^*(m) = V(\phi^*(m), m)$. Therefore, the first-order condition in equation (24) can be written as

$$\frac{\partial A(D^*(m);m)}{\partial D} = \tau(\rho + \lambda F(\phi^*(m) - m)) - \alpha(\rho + \lambda)\lambda V(\phi^*(m);m) f(\phi^*(m) - m) = 0. \quad (34)$$

²⁸However, the entire distribution of $\phi(t)$ shifts to the right so the *unconditional* expected value of $S(\phi(t))$, which is $\frac{1-\tau}{\rho+\lambda}(E\{\phi\} - \Phi_L)$ in Scenario I, is unchanged since $\frac{d\Phi_L(m)}{dm} = \frac{dE\{\phi\}}{dm} = 1$.

Lemma 18 *If $0 < \tau < \frac{\alpha}{1+\alpha}$ and $f'(\phi) \geq 0$ for all ϕ , then $\phi^{*'}(m) < 1$.*

To understand why $\phi^{*'}(m) < 1$, suppose that $\phi^*(m) = \phi^*(0) + m$ so that a rightward translation of the distribution by m increases ϕ^* by m , which would leave the $F(\phi^*(m) - m)$ and $f(\phi^*(m) - m)$ unchanged but would increase $V(\phi^*(m); m)$ through the direct effect of an increase in ϕ on the firm's value and through the effect in Proposition 16. Therefore, the marginal interest tax shield in equation (34) would remain unchanged, but the marginal expected cost of default would increase, thereby reducing the reducing the optimal amount of debt and hence reducing the associated critical value $\phi^*(m)$ below $\phi^*(0) + m$.

To further explore the impact of a translation of the distribution $F(\phi - m)$, differentiate equation (34) with respect to m and use equation (34) to substitute $\tau \frac{\rho + \lambda F(\phi^*(m) - m)}{f(\phi^*(m) - m)}$ for $\alpha(\rho + \lambda)\lambda V(\phi^*(m), m)$ to obtain²⁹

$$\tau[\phi^{*'}(m) - 1]\chi(\phi^*(m) - m) = \alpha(\rho + \lambda)\lambda f(\phi^*(m) - m) \frac{dV(\phi^*(m); m)}{dm}, \quad (35)$$

where

$$\chi(\phi) \equiv \lambda f(\phi) - (\rho + \lambda F(\phi)) \frac{f'(\phi)}{f(\phi)}. \quad (36)$$

The following proposition uses equation (35) to examine the impacts on $P(\phi(t); m)$, $\phi^*(m)$, and $D^*(m)$ of a rightward translation of the distribution $F(\phi)$.

Proposition 19 *If $0 < \tau < \frac{\alpha}{1+\alpha}$ and (2) $f'(\phi) \geq 0$ for all ϕ , then*

1. $\frac{dP(\phi(t); m)}{dm} < 0$.
2. $\text{sign}\left(\frac{dV(\phi^*(m); m)}{dm}\right) = \text{sign}(D^{*'}(m)) = -\text{sign}(\chi(\phi^*(m) - m))$.
3. $\phi^{*'}(m) < 0$, if $\chi(\phi^*(m) - m) \geq 0$.

Since an increase in m does not increase $\phi^*(m)$ by an amount greater than or equal to the increase in m , it reduces $F(\phi^*(m) - m)$ and hence reduces the probability of default if $\widehat{D}(\phi(t); m) = D^*(m)$. Alternatively, if $\widehat{D}(\phi(t); m) < D^*(m)$, an increase in m also

²⁹ $\frac{dV(\phi^*(m); m)}{dm} = \frac{\partial V(\phi^*(m); m)}{\partial \phi} \phi^{*'}(m) + \frac{\partial V(\phi^*(m); m)}{\partial m}$ is the effect on $V(\phi^*)$ of a small increase in m . Proposition 16 implies that $\frac{\partial V(\phi^*(m); m)}{\partial m} > 0$, but does not imply that $\frac{dV(\phi^*(m); m)}{dm}$ is positive. Indeed, as shown in Statement 2 of Proposition 19, $\frac{dV(\phi^*(m); m)}{dm}$ will be negative when $\chi(\phi^*)$, defined in equation (36), is positive.

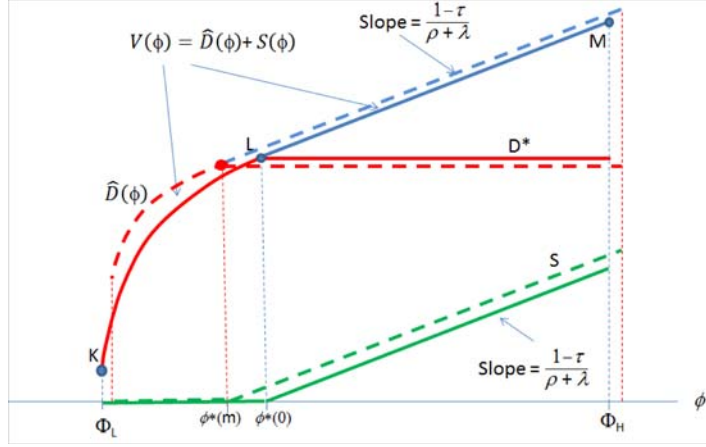


Figure 3: Rightward Translation of $F(\phi)$ in Scenario II

reduces the probability of default, which is $\lambda F(\phi(t) - m)$. Thus, for any $\phi(t)$, a rightward translation of the distribution of profitability reduces $P(\phi(t); m)$ (Statement 1 of Proposition 19). A rightward translation of the distribution, which is an increase in m , decreases $\rho + \lambda F(\phi; m)$, which enters the marginal interest tax shield, and, since $f'(\phi; m) \geq 0$, decreases $f(\phi; m)$, which enters the marginal expected cost of default. If $\chi(\phi) > 0$, then the percentage decrease in the marginal interest tax shield exceeds the percentage decrease in $f(\phi)$ when ϕ increases, and so the value of $D^*(m)$ falls. Since $D^*(m) = V(\phi^*(m); m)$, the fall in D^* implies that $V(\phi^*(m); m)$ falls (Statement 2 of Proposition 19). Furthermore, since an increase in m increases $V(\phi; m)$, the value of $\phi^*(m)$ falls to achieve the fall in $V(\phi^*(m); m)$ (Statement 3 of Proposition 19).

If the distribution $F(\phi)$ is uniform, so that $f'(\phi) \equiv 0$, it satisfies the condition in Proposition 19 that $f(\phi)$ is non-decreasing. With $F(\phi)$ being uniform, $\chi(\phi) = \lambda f(\phi) = \frac{\lambda}{\Phi_H - \Phi_L} > 0$. Therefore, Statements 2 and 3 of Proposition 19 imply that with a uniform distribution, a rightward translation reduces D^* , reduces ϕ^* , and reduces $V(\phi^*(m); m)$.

Figure 3 illustrates the impact of a rightward translation of $F(\phi)$ for the case in which $\chi(\phi^*(m) - m) > 0$, which includes the uniform distribution. In response to a rightward translation of $F(\phi)$, the value of ϕ^* falls from its initial value of $\phi^*(0)$ to its new value of $\phi^*(m)$ (Statement 3). In addition, the curves representing equity value, optimal debt, and the total value of the firm move from the solid curves to the dashed curves. As shown,

the rightward translation of $F(\phi)$ has no effect on equity for $\phi \leq \phi^*(m)$ and increases the equity value by a constant amount, $\frac{1-\tau}{\rho+\lambda}(\phi^*(0) - \phi^*(m)) > 0$ for $\phi > \phi^*(0)$. The total value of the firm increases for all $\phi(t)$ (Proposition 16), as shown by the shift from the solid curve through K , L , and M , to the dashed curve that lies above it. The response of optimal debt to a rightward translation of $F(\phi)$ depends on whether ϕ is high or low. For $\phi(t) \geq \phi^*(0)$, the optimal value of debt, which is invariant to $\phi(t)$ in this range, falls in response to a rightward translation of $F(\phi)$ (Statement 2). The continuation value of the firm increases, thereby increasing the cost of losing this future value by defaulting on debt. In response to this increased cost of default, the firm reduces its exposure to default by reducing the amount of debt it issues and reduces the critical value ϕ^* . For $\phi(t) \leq \phi^*(m)$, the optimal amount of debt increases in response to a rightward translation of $F(\phi)$ because the borrowing constraint is binding for low values of $\phi(t)$ and a rightward translation of $F(\phi)$ increases the total value of the firm, which allows the firm to borrow an increased amount. Finally, there is some $\tilde{\phi} \in (\phi^*(m), \phi^*(0))$, not labelled in Figure 3, such that a rightward translation of $F(\phi)$ increases optimal debt for $\phi(t) < \tilde{\phi}$ and decreases optimal debt for $\phi(t) > \tilde{\phi}$.³⁰

To understand the role of $\chi(\phi)$ in determining the impact of a rightward translation of $F(\phi)$,³¹ it is helpful to recall that $\frac{\partial A(D^*(m), m)}{\partial D}$ in the first-order condition in equation (34) equals the marginal interest tax shield minus the marginal expected default cost. *Holding $\phi^*(m)$ fixed*, an increase in m (1) *reduces* the marginal interest tax shield (by reducing $F(\phi^*(m) - m)$) and, (2) if $F(\phi)$ is uniform so that $f(\phi^*(m))$ is invariant to m , increases the marginal cost of default (by increasing $V(\phi^*(m), m)$). Therefore, an increase in m makes $\frac{\partial A(D^*(m), m)}{\partial D}$ negative. Because $A(D; m)$ is strictly concave in D , a reduction in D is needed to restore the first-order condition. In order to obtain the opposite effect, that is, in order for an increase in m to increase the optimal value of D^* , an increase in m must *reduce* the marginal default cost—and must do so by more than the reduction in the marginal

³⁰The text has shown that when $\chi(\phi^*) > 0$, $D^*(m) < D^*(0) = \widehat{D}(\phi^*(0); 0)$. Note that $D^*(m) = \widehat{D}(\phi^*(m); m) = V(\phi^*(m); m) > V(\phi^*(m); 0) = \widehat{D}(\phi^*(m); 0)$, where the final equality follows from the fact that the borrowing constraint binds for $\phi(t) = \phi^*(m)$ under the original distribution $F(\phi)$. Therefore, $\widehat{D}(\phi^*(m); 0) < D^*(m) < \widehat{D}(\phi^*(0); 0)$. Since $\widehat{D}(\phi; 0)$ is increasing in ϕ for $\phi < \phi^*(0)$, there is a unique $\tilde{\phi} \in (\phi^*(m), \phi^*(0))$ for which $\widehat{D}(\tilde{\phi}; 0) = D^*(m)$.

³¹Section D in the online appendix analyzes the case in which the unconditional distribution of EBIT is a truncated exponential distribution and distinguishes situations in which D^* falls in response to a rightward translation of $F(\phi)$ from situations in which D^* increases in response to a rightward translation of $F(\phi)$.

interest tax shield. Such a reduction in the marginal default cost requires that an increase in m reduces $f(\phi^*(m) - m)$, for given $\phi^*(m)$, by a sufficiently large amount, which will occur if $f'(\phi^*)$ is large enough to make $\chi(\phi^*) < 0$.

The tradeoff theory of capital structure is operative only in Scenario II and, indeed, only for $\phi(t) \geq \phi^*$ in Scenario II. When the tradeoff theory is operative, i.e., when $\phi(t) \geq \phi^*$, the optimal value of debt equals D^* . In the focal case of a uniform distribution, and more generally when $\chi(\phi^*) > 0$, Proposition 19 implies that a rightward translation of the distribution $F(\phi)$ reduces the optimal amount of debt and hence reduces the optimal leverage ratio. When $\phi(t) < \phi^*$, the tradeoff theory is not operative and $\widehat{D}(\phi(t); m) = V(\phi(t); m)$. Therefore, Proposition 16 implies that the optimal value of debt increases in response to a rightward translation of $F(\phi)$ in Scenario II when $\phi(t) < \phi^*$. Of course, whenever $\phi(t) \leq \phi^*$, the optimal leverage ratio equals one and thus is invariant to a translation of $F(\phi)$.

6.3 Effect of a Rightward Translation of $F(\phi)$ in Scenario III

In Scenario III, the borrowing constraint binds, that is, $\widehat{D}(\phi(t); m) = V(\phi(t); m)$ for all $\phi(t)$. Therefore, Proposition 16 implies that in Scenario III a rightward translation of $F(\phi)$ increases $V(\phi(t))$ and hence increases $\widehat{D}(\phi(t))$ for all $\phi(t)$. Of course, with $\widehat{D}(\phi(t); m) = V(\phi(t); m)$ for all $\phi(t)$, the optimal leverage ratio equals one for all $\phi(t)$ in Scenario III.

6.4 Summary of Effects of Increased Profitability

Table 1 summarizes the findings about the impact of increased profitability on optimal debt when $\chi(\phi) > 0$, which includes the case in which $F(\phi)$ is uniform. The top third of the table summarizes characteristics of the different scenarios. The tradeoff theory will be operative if and only if (1) the borrowing constraint (B.C., in the table) is not binding and (2) the firm faces a positive probability of default. Only Scenario IIA meets these two criteria, so the tradeoff theory is operative only in Scenario IIA. Among the scenarios in which the firm faces a positive probability of default (Scenarios IIA, IIB, and III), an increase in profitability—either an increase in current profitability, $\phi(t)$, or a rightward translation of the unconditional distribution $F(\phi)$ —can reduce the optimal leverage ratio only in Scenario IIA, that is, only when the tradeoff theory is operative. Thus, in the context of the model

Characteristics of Scenarios			
Scenario I	Scenario IIA	Scenario IIB	Scenario III
B.C.* not binding	B.C.* not binding	B.C.* binds	B.C.* binds
$\Pr\{\text{default}\}=0$	$\Pr\{\text{default}\}>0$	$\Pr\{\text{default}\}>0$	$\Pr\{\text{default}\}>0$
Not Tradeoff	Tradeoff Operative	Not Tradeoff	Not Tradeoff
Increase in $\phi(t)$			
Scenario I	Scenario IIA	Scenario IIB	Scenario III
\widehat{D} unchanged	\widehat{D} unchanged	$\widehat{D} \uparrow$	$\widehat{D} \uparrow$
$S \uparrow$	$S \uparrow$	$S \equiv 0$	$S \equiv 0$
$V \uparrow$	$V \uparrow$	$V \uparrow$	$V \uparrow$
$L \downarrow$	$L \downarrow$	$L \equiv 1$	$L \equiv 1$
Rightward Translation of $F(\phi)$			
Scenario I	Scenario IIA	Scenario IIB	Scenario III
$\widehat{D} \uparrow$	$\widehat{D} \downarrow$ if $\chi > 0$	$\widehat{D} \uparrow$	$\widehat{D} \uparrow$
$S \downarrow$	$S \uparrow$	$S \equiv 0$	$S \equiv 0$
$V \uparrow$	$V \uparrow$	$V \uparrow$	$V \uparrow$
$L \uparrow$	$L \downarrow$ if $\chi > 0$	$L \equiv 1$	$L \equiv 1$
*B.C. refers to the borrowing constraint $D(\phi) \leq \widehat{V}(\phi)$			

Table 1: Effects of Increased Profitability

presented here, in which the tradeoff theory can be operative or not, the empirical finding of a negative relationship between profitability and the leverage ratio is consistent only with the tradeoff theory being operative, if the probability of default is positive.

7 Persistence of Profitability Across Regimes

I derived the main result of this paper—that the leverage ratio is negatively related to profitability when the tradeoff theory is operative—by assuming that profitability is i.i.d. across regimes. Under this assumption, the tradeoff function $A(D)$ defined in equation (10) is independent of $\phi(t)$, which implies that D^* is independent of $\phi(t)$. And when the tradeoff theory is operative, optimal debt, $\widehat{D}(\phi(t))$, equals D^* so the optimal level of debt is invariant to $\phi(t)$. Therefore, the leverage ratio, $\frac{\widehat{D}(\phi(t))}{V(\phi(t))}$, is decreasing in $\phi(t)$ because the value of the firm is increasing in $\phi(t)$.

In this section, I show that with persistence across regimes, the optimal amount of debt

can be a decreasing function of profitability when the tradeoff theory is operative. To the extent that higher profitability reduces the amount of debt, it strengthens the major finding of this paper that when the tradeoff theory is operative, the leverage ratio is a decreasing function of profitability. And to the extent that higher profitability reduces the amount of debt, the probability of default is a decreasing function of profitability when the tradeoff theory is operative.

Suppose that the current time is t and let t' denote the time of the first regime change after t . Let $F(\phi(t'), \phi(t))$ be the distribution of $\phi(t')$ conditional on $\phi(t)$, and let $f(\phi(t'), \phi(t)) \equiv \frac{\partial F(\phi(t'), \phi(t))}{\partial \phi(t')}$ be the associated conditional density. I will model stable persistence in profitability across regimes by assuming that

$$0 \leq -\frac{\partial F(\phi(t'), \phi(t))}{\partial \phi(t)} \leq \frac{\partial F(\phi(t'), \phi(t))}{\partial \phi(t')}. \quad (37)$$

Persistence is captured by the first inequality in equation (37), which implies that if $\phi_2 > \phi_1$, then the distribution of $\phi(t')$ conditional on $\phi(t) = \phi_2$ first-order stochastically dominates the distribution of $\phi(t')$ conditional on $\phi(t) = \phi_1$. The second inequality captures stability. For instance, it implies that a one-unit increase in $\phi(t)$ increases each q -quantile of the conditional distribution of $\phi(t')$ by less than one unit.³²

Under the conditional distribution $F(\phi(t'), \phi(t))$, the tradeoff function $A(D)$ in equation (10) is

$$A(D, \phi(t)) \equiv \frac{\tau \left[\rho + \lambda \int_{V(\phi(t')) < D} f(\phi(t'), \phi(t)) d\phi(t') \right] D}{-\left[\alpha + \tau(1 - \alpha) \right] \lambda \int_{V(\phi(t')) < D} V(\phi) f(\phi(t'), \phi(t)) d\phi(t')} \quad (38)$$

and the value function is

$$V(\phi(t)) = \frac{1}{\rho + \lambda} \max_{D \leq V(\phi(t))} [(1 - \tau)\phi(t) + \bar{v} + A(D, \phi(t))]. \quad (39)$$

³²Formally, define $h(p, \phi(t))$ implicitly by $p \equiv F(h(p, \phi(t)), \phi(t))$. Stable positive first-order serial dependence is described by $0 \leq \frac{\partial h(p, \phi(t))}{\partial \phi(t)} < 1$. Implicitly differentiate $p \equiv F(h(p, \phi(t)), \phi(t))$ with respect to $\phi(t)$ for a given p to obtain $\frac{\partial h(p, \phi(t))}{\partial \phi(t)} = -\frac{\partial F(\phi(t'), \phi(t))}{\partial \phi(t)} \left[\frac{\partial F(\phi(t'), \phi(t))}{\partial \phi(t')} \right]^{-1}$. Therefore, equation (37) implies that $0 \leq \frac{\partial h(p, \phi(t))}{\partial \phi(t)} < 1$.

Proposition 20 *If $F(\phi(t'), \phi(t))$ is non-increasing in $\phi(t)$, the value function $V(\phi(t))$ is strictly increasing in $\phi(t)$.*

Because the conditional distribution of $\phi(t')$ depends on $\phi(t)$, $\widehat{D}(\phi(t)) \equiv \arg \max_{V(\Phi_L) \leq D \leq V(\Phi_H)} A(D, \phi(t))$ depends on $\phi(t)$. If $A(D, \phi(t))$ attains a local maximum at $D^*(\phi(t)) \in (\Phi_L, \Phi_H)$, then the associated first-order condition at $D = D^*$ is

$$\frac{\partial A(D^*, \phi(t))}{\partial D} = \tau [\rho + \lambda F(V^{-1}(D^*), \phi(t))] - (1 - \tau) \alpha \lambda V^{-\nu}(D^*) D^* f(V^{-1}(D^*), \phi(t)) = 0. \quad (40)$$

The first of the two terms on the right hand side of equation (40) is the marginal interest tax shield. Equation (37) implies that a higher value of $\phi(t)$ leads to a lower value of this marginal interest tax shield for any given value of D , thereby reducing the incentive to issue debt. The second term on the right hand of equation (40) is the marginal expected cost of default. A higher value of $\phi(t)$ could lead to a higher or a lower value of this cost by increasing or decreasing the density, $f(V^{-1}(D), \phi(t))$. If the impact on the marginal interest tax shield outweighs the impact of any change in the marginal expected cost of default, the value of D^* falls.

The instantaneous conditional probability of default at time t in equation (18) is

$$p(\phi(t)) \equiv \lambda F\left(V^{-1}\left(\widehat{D}(\phi(t))\right), \phi(t)\right), \quad (41)$$

which is the instantaneous conditional probability that a new regime will arrive with a value of $\phi(t')$ that is low enough that the new value of the firm, $V(\phi(t'))$, is less than outstanding bonds, $\widehat{D}(\phi(t))$. Therefore,

$$p'(\phi(t)) = \lambda f\left(V^{-1}\left(\widehat{D}(\phi(t))\right), \phi(t)\right) V^{-\nu}\left(\widehat{D}(\phi(t))\right) \widehat{D}'(\phi(t)) + \lambda \frac{\partial F\left(V^{-1}\left(\widehat{D}(\phi(t))\right), \phi(t)\right)}{\partial \phi(t)}. \quad (42)$$

The two terms on the right hand side of equation (42) correspond to channels (1) and (2), respectively, introduced in the discussion following Proposition 10. The first term captures the effect on the optimal amount of debt of an increase in $\phi(t)$, and the consequent effect on the probability of default. The second term captures the direct effect of a change in the conditional distribution on the probability of default at a given value of debt. If $\phi(t)$ is i.i.d.

across regimes, the second term is zero, and the sign of $p'(\phi(t))$ is the same as the sign of $\widehat{D}'(\phi(t))$: positive when the borrowing constraint is binding, that is, $\widehat{D}(\phi(t)) = V(\phi(t))$; and zero when the tradeoff theory is operative, so $\widehat{D}'(\phi(t)) = 0$.

The following proposition describes optimal debt and the probability of default in a simple example in which profitability is persistent across regimes.

Proposition 21 *Suppose that $F(\phi(t'), \phi(t))$ is uniform on $[g(\phi(t)) - d, g(\phi(t)) + d]$,³³ where $d > 0$ and $0 < g'(\phi(t)) < 1$.*

1. *If the tradeoff theory is operative in a neighborhood of $\phi(t)$, then*

$$(a) \widehat{D}'(\phi(t)) < 0$$

$$(b) L'(\phi(t)) < 0$$

$$(c) p'(\phi(t)) < 0.$$

2. *If the borrowing constraint is binding in a neighborhood of $\phi(t)$, then*

$$(a) \widehat{D}'(\phi(t)) > 0$$

$$(b) p'(\phi(t)) = \frac{1-g'(\phi(t))}{2d} > 0.$$

Proposition 21 implies that the introduction of stable persistence across regimes in the case of a uniform distribution has different effects depending on whether the tradeoff theory is operative or the borrowing constraint is binding. When the borrowing constraint is binding, an increase in current profitability increases the value of the firm and hence increases the optimal amount of debt, and increases the instantaneous conditional probability of default, which is the same counter-intuitive result that holds with profitability is i.i.d. across regimes.³⁴ However, when the tradeoff theory is operative, an increase in current profitability reduces the optimal amount of debt and hence reduces the leverage ratio, which reinforces the major result of this paper that profitability and leverage are negatively related when the tradeoff theory is operative. In addition, when the tradeoff theory is operative, an increase in current profitability reduces the instantaneous conditional probability of default.

³³To ensure that the support of the distribution of $\phi(t')$ remains within $[\Phi_L, \Phi_H]$, suppose that $\Phi_L + d \leq g(\Phi_L) < g(\Phi_H) \leq \Phi_H - d$.

³⁴These results in the case of binding borrowing constraints hold much more generally. When the borrowing constraint binds, $p(\phi(t)) \equiv \lambda F(V^{-1}(\widehat{D}(\phi(t))), \phi(t)) = \lambda F(V^{-1}(V(\phi(t))), \phi(t)) = \lambda F(\phi(t), \phi(t))$, which is (weakly) increasing in $\phi(t)$ under the general condition for stable persistence in equation (37).

8 Concluding Remarks

The model of debt choice developed and analyzed here is simple enough to be analytically tractable yet rich enough to include situations in which the tradeoff theory is operative as well as situations in which it is not operative. This model points to factors, such as the tax rate and the level of profitability, that determine whether or not the tradeoff theory is operative. To the extent that optimal behavior differs depending on whether or not the tradeoff theory is operative, one could potentially use such differences in behavior to empirically test the tradeoff theory.

The equality of the marginal tax shield resulting from interest deductibility and the marginal cost of increased exposure to default associated with increased debt is the defining feature of the tradeoff theory of debt. The mere presence of interest deductibility and deadweight costs of default is not sufficient to ensure that the tradeoff theory is operative. I have demonstrated three situations in which the tradeoff theory is not operative despite the presence of interest deductibility and deadweight default costs. First, if the tax rate is very low (Scenario I), the marginal benefit of the interest tax shield associated with an additional dollar of debt is completely overwhelmed by the marginal cost associated with increased exposure to default resulting from an additional dollar of debt. In this case, the firm will take advantage of the tax shield offered by interest deductibility, but will only borrow as much as it can without exposing itself to any possibility of default. Thus, the tradeoff theory is not operative for tax rates that are sufficiently low. Second, if the tax rate is sufficiently high (but not so high as to violate conditions for concavity of the value function), the marginal benefit of the interest tax shield associated with an additional dollar of debt completely overwhelms the marginal cost associated with increased exposure to default resulting from an additional dollar of debt. In this case, the firm borrows as much as lenders are willing to lend. The tradeoff theory is not operative because the borrowing constraint is strictly binding so the marginal benefit of the interest tax shield fails to equal the marginal cost of increased default probability. Third, even if the tax rate is neither too low nor too high, the tradeoff theory will fail to be operative if the current value of EBIT is lower than the critical value, denoted by ϕ^* in the model. In this situation, a low value of EBIT implies that the current value of firm is low, which implies that the constraint on how much the firm can borrow is strictly binding. In this situation, the marginal benefit of the interest

tax shield exceeds the marginal cost of increased default probability, so, again, the tradeoff theory is not operative. The only situation in the model in which the tradeoff theory is operative is when the tax rate is neither too low nor too high and the current value of EBIT is higher than the critical value ϕ^* . In this case, the optimal value of debt equates the marginal benefit of the interest tax shield and the marginal cost of increased exposure to default. In the baseline stochastic specification in which EBIT is i.i.d. across regimes, the optimal value of debt is invariant to the contemporaneous value of EBIT when the tradeoff theory is operative.

The relationship between profitability and optimal borrowing depends on whether the tradeoff theory is operative. Table 1 summarizes the impact on optimal borrowing of an increase in profitability. In this table are two measures of leverage—the amount of debt issued and the market leverage ratio—and two measures of profitability—current EBIT and the unconditional distribution from which new values of EBIT are drawn. Two major lessons emerge from this table. First, when the tradeoff theory is operative, the leverage ratio is a decreasing function of current EBIT; and provided that the unconditional density function of EBIT does not slope upward too steeply, a rightward translation of the unconditional distribution of EBIT reduces both the level of optimal debt and the optimal leverage ratio. This finding that an increase in profitability leads to a reduction in the leverage ratio is consistent with empirical analyses of this relationship. The second lesson is that provided the firm faces a positive probability of default, the optimal leverage ratio is negatively related to profitability only if the tradeoff theory is operative. So the empirical finding of a negative relationship between borrowing and profitability is not consistent with the alternatives to the tradeoff theory in this model.

I will conclude by describing four potential extensions of the analysis in this paper. First, the stochastic specification for EBIT can be generalized in various ways. For instance, the arrival rate of new regimes, λ , can be a random variable. In particular, the arrival of a new regime could be specified as a new realization of the vector (ϕ_t, λ_t) drawn from a distribution $F(\phi_t, \lambda_t)$. If the realizations of (ϕ_t, λ_t) are i.i.d. across regimes, the analysis of this paper is easily extended, and would allow an added dimension of cross-sectional heterogeneity across firms.³⁵ An alternative generalization of the stochastic structure would be to allow for persistence in EBIT across regimes, as in Section 7. The analysis in Section 7 suggests

³⁵I thank an anonymous referee for pointing out this extension.

that persistence across regimes strengthens the appeal of the tradeoff theory, at least in the example in which the conditional distribution of EBIT is uniform. In that case, when the tradeoff theory is operative, the optimal amount of bonds is decreasing in profitability, which reinforces the finding that the leverage ratio is decreasing in profitability, consistent with empirical findings. In addition, when the tradeoff theory is operative, the probability of default is decreasing in profitability.

Second, an important extension of the current model is to incorporate the capital investment decision. In ongoing work (Abel 2016), I develop and analyze a model in which a firm can accumulate (or decumulate) physical capital subject to convex costs of adjustment. In that framework, it continues to be the case that the optimal leverage ratio is a declining function of profitability when the tradeoff theory is operative.

A third type of extension is to include flotation costs for bonds. Footnote 23 outlines a simple extension with flotation costs specified as a convex function of the amount of bonds. In that case, the tradeoff theory is easily extended to a "modified tradeoff theory." Future research could analyze a broader variety of flotation costs, including perhaps a fixed cost component of these costs.

Fourth, the assumption of instantaneous maturity of bonds is analytically convenient and, in fact, seems appropriate for overnight repos and commercial paper. However, future research could examine bonds of longer maturity and address the optimal maturity of bonds.

References

- Abel, Andrew B.**, "Investment and Leverage," Working Paper, May 2016.
- Biais, Bruno, Thomas Mariotti, Guillaume Plantin and Jean-Charles Rochet**, "Dynamic Security Design: Convergence to Continuous Time and Asset Pricing Implications, *The Review of Economic Studies*, 74, 2 (April 2007), pp. 345-390.
- Fama, Eugene F. and Kenneth R. French**, "Testing Trade-Off and Pecking Order Predictions about Dividends and Debt," *The Review of Financial Studies*, 15, 1 (Spring 2002), 1-33.
- Frank, Murray Z. and Vidhan K. Goyal**, "Trade-off and Pecking Order Theories of Debt," in Espen Eckbo (editor) *Handbook of Corporate Finance: Empirical Corporate Finance*, Volume 2 (Elsevier Science/North-Holland), 135-202.
- Goldstein, Robert, Nengjiu Ju, and Hayne Leland**, "An EBIT-Based Model of Dynamic Capital Structure," *Journal of Business*, 74, 4 (October 2001), 483-512.
- Gorbenko, Alexander S. and Ilya A. Strebulaev**, "Temporary versus Permanent Shocks: Explaining Corporate Financial Policies," *The Review of Financial Studies*, 23, 7 (July 2010), 2591-2647.
- Hennessy, Christopher A. and Toni M. Whited**, "Debt Dynamics," *The Journal of Finance*, 60, 3, (June 2005), 1129-1165.
- Hennessy, Christopher A. and Toni M. Whited**, "How Costly is External Financing? Evidence from a Structural Estimation," *The Journal of Finance*, 62, 4 (August 2007), 1705-1745.
- Kraus, Alan and Robert H. Litzenberger**, "A State-Preference Model of Optimal Financial Leverage," *The Journal of Finance*, 28, 4 (September 1973), 911-922.
- Leary, Mark T. and Michael R. Roberts**, "Do Firms Rebalance Their Capital Structures?" *The Journal of Finance*, 60, 6 (December 2005), 2575-2619.
- Leland, Hayne E.**, "Corporate Debt Value, Bond Covenants, and Optimal Capital Structure," *The Journal of Finance*, 49, 4 (September 1994), 1213-1252.
- Merton, Robert C.** "On the Pricing of Corporate Debt: The Risk Structure of Interest Rates," *The Journal of Finance*, 29, 2 (May 1974), 449-470.
- 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.

Robichek, Alexander A. and Stewart C. Myers, "Problems in the Theory of Optimal Capital Structure," *The Journal Financial and Quantitative Analysis*, 1, 2 (June 1966), pp. 1-35.

Scott, James H., Jr., "A Theory of Optimal Capital Structure," *The Bell Journal of Economics*, 7, 1 (Spring 1976), pp. 33-54.

Strebulaev, Ilya A., "Do Tests of Capital Structure Theory Mean What They Say?" *The Journal of Finance*, 62,4 (August 2007), 1747-1787.

A Appendix: Proofs

Proof. of Proposition 1. Assume that $\phi_2 > \phi_1$. If $\widehat{D}(\phi_1)$ is feasible when $\phi(t) = \phi_1$, then equation (9) implies that $V(\phi_2) \geq \frac{(1-\tau)\phi_2 + \lambda\bar{v} + A(\widehat{D}(\phi_1))}{\rho + \lambda}$ and $V(\phi_1) = \frac{(1-\tau)\phi_1 + \lambda\bar{v} + A(\widehat{D}(\phi_1))}{\rho + \lambda}$, so $V(\phi_2) - V(\phi_1) \geq \frac{(1-\tau)(\phi_2 - \phi_1)}{\rho + \lambda} > 0$. Note that $\widehat{D}(\phi_1) \leq V(\phi_1) < V(\phi_2)$ confirms that $\widehat{D}(\phi_1)$ is feasible when $\phi(t) = \phi_2$. Also $V(\phi_2) - V(\phi_1) \geq \frac{(1-\tau)(\phi_2 - \phi_1)}{\rho + \lambda}$ implies that $\frac{V(\phi_2) - V(\phi_1)}{\phi_2 - \phi_1} \geq \frac{1-\tau}{\rho + \lambda}$. ■

Proof. of Proposition 2. Assume that $\phi_2 > \phi_1$. Suppose, contrary to what is to be proved, that $\widehat{D}(\phi_2) \leq \widehat{D}(\phi_1)$. Therefore, $\widehat{D}(\phi_2) \leq \widehat{D}(\phi_1) \leq V(\phi_1) \leq V(\phi_2)$ so both $\widehat{D}(\phi_1)$ and $\widehat{D}(\phi_2)$ are feasible both when $\phi = \phi_1$ and when $\phi = \phi_2$. Since $\widehat{D}(\phi_1)$ is chosen by the firm when $\phi = \phi_1$, $A(\widehat{D}(\phi_1)) > A(\widehat{D}(\phi_2))$ because $A(D)$ is strictly concave. Since $\widehat{D}(\phi_2)$ is chosen by the firm when $\phi = \phi_2$, $A(\widehat{D}(\phi_2)) > A(\widehat{D}(\phi_1))$ because $A(D)$ is strictly concave, which is a contradiction. Therefore, $\widehat{D}(\phi_2) \geq \widehat{D}(\phi_1)$ (Statement 1). To prove that $\widehat{D}(\phi(t))$ is non-decreasing in time, it suffices to prove that for any $t_2 > t_1$ where t_1 and t_2 are in consecutive regimes, $\widehat{D}(\phi(t_2)) \geq \widehat{D}(\phi(t_1))$. Suppose, contrary to what is to be proved, that $\widehat{D}(\phi(t_2)) < \widehat{D}(\phi(t_1))$. Since, $\widehat{D}(\phi(t_2)) < \widehat{D}(\phi(t_1)) \leq V(\phi(t_1))$, the firm chooses $\widehat{D}(\phi(t_1))$ in preference to $\widehat{D}(\phi(t_2))$ when both are feasible. It would only choose $\widehat{D}(\phi(t_2))$ in preference to $\widehat{D}(\phi(t_1))$ when $\phi = \phi(t_2)$ if $\widehat{D}(\phi(t_1))$ were not feasible, that is, if $\widehat{D}(\phi(t_1)) > V(\phi(t_2))$. In this situation, the firm would default on its debt, $\widehat{D}(\phi(t_1))$, when the regime changes and ϕ changes from $\phi(t_1)$ to $\phi(t_2)$ (Statement 2). ■

Proof. of Proposition 3. To prove the first inequality, $0 < V(\Phi_L)$, observe that a feasible policy would be for the firm to never issue any bonds and to stay in operation forever, because the assumption $\Phi_L > -\frac{\lambda}{\rho}E\{\phi\}$ in equation (2) implies that $W(\Phi_L) > 0$. If $\phi(t) = \Phi_L$ and the firm followed this all-equity policy forever, its value would be $W(\Phi_L)$. Therefore, $V(\Phi_L) \geq W(\Phi_L) > 0$. To prove the second inequality, $V(\Phi_L) \leq \widehat{D}(\phi(t))$, observe that if the firm's outstanding bonds are $D_0 < V(\Phi_L)$, there no possibility that the firm will choose to default on the debt. In this case, $A(D_0) = \tau\rho D_0$, so that $(1-\tau)\phi(t) + \lambda\bar{v} + A(D_0) = (1-\tau)\phi(t) + \lambda\bar{v} + \tau\rho D_0$. If $D_0 < V(\Phi_L)$, then $(1-\tau)\phi(t) + \lambda\bar{v} + \tau\rho D_0 < (1-\tau)\phi(t) + \lambda\bar{v} + \tau\rho V(\Phi_L)$, so $D_0 < V(\Phi_L)$ cannot be optimal. Therefore, $V(\Phi_L) \leq \widehat{D}(\phi(t))$. The third inequality, $\widehat{D}(\phi(t)) \leq V(\Phi_H)$, is a direct consequence of the borrowing constraint, $D \leq V(\Phi_t)$, and the fact that $V(\phi)$ is strictly increasing (Proposition 1), so that $V(\phi(t)) \leq V(\Phi_H)$. ■

Proof. of Corollary 4. $V(\Phi_L) \geq \widehat{D}(\Phi_L) \geq V(\Phi_L)$, where the first inequality is the borrowing constraint that prevents the firm from borrowing an amount that leads to immediate default and the second inequality follows from Proposition 3. Since $S(\phi(t)) = V(\phi(t)) - \widehat{D}(\phi(t))$, $S(\Phi_L) = 0$. ■

Lemma 22 is stated (and proved) here, immediately preceding the proof of Proposition 5, rather than in the text, because its only role is to help prove Propositions 5 and 6.

Lemma 22 . *If (1) $f'(\phi(t)) \geq 0$ for all $\phi \in [\Phi_L, \Phi_H]$, (2) $\tau < \frac{\alpha}{1+\alpha}$, and (3) $V(\phi(t))$ is concave and strictly increasing, then $A(D_t)$ is strictly concave in D_t for $D_t \geq 0$.*

Proof. of Lemma 22. Assume that $V(\phi(t))$ is concave and strictly increasing. Twice differentiate $V^{-1}(V(\phi)) \equiv \phi$ with respect to ϕ to obtain $V^{-1'}(V(\phi))V'(\phi) \equiv 1$ and $V^{-1''}(V(\phi))[V'(\phi)]^2 + V^{-1'}(V(\phi))V''(\phi) = 0$. Since $V(\phi(t))$ is concave and strictly increasing, $V'(\phi) > 0$, $V^{-1'}(V(\phi)) > 0$, and $V''(\phi) \leq 0$, so that $V^{-1''}(D_t) \geq 0$. Twice partially differentiating $A(D_t) \equiv \tau \left[\rho + \lambda \int_{V(\phi) < D_t} dF(\phi) \right] D_t - [\alpha + \tau(1 - \alpha)] \lambda \int_{V(\phi) < D_t} V(\phi) dF(\phi)$ with respect to D_t yields $A'(D_t) = \tau \left[\rho + \lambda \int_{V(\phi) < D_t} dF(\phi) \right] - \alpha(1 - \tau) \lambda V^{-1'}(D_t) D_t f(V^{-1}(D_t))$ and
$$A''(D_t) = [(1 + \alpha)\tau - \alpha] \lambda V^{-1'}(D_t) f(V^{-1}(D_t)) - \alpha(1 - \tau) \lambda D_t \left(V^{-1''}(D_t) f(V^{-1}(D_t)) + f'(V^{-1}(D_t)) [V^{-1'}(D_t)]^2 \right) < 0$$
 for $D_t \geq 0$ since $\tau < \frac{\alpha}{1+\alpha}$ implies that the first term, $[(1 + \alpha)\tau - \alpha] \lambda V^{-1'}(D_t) f(V^{-1}(D_t))$, is negative, and $V^{-1''}(D_t) \geq 0$ and $f'(V^{-1}(D_t)) \geq 0$ imply that the second term, which follows (but does not include) a minus sign, is non-negative. Therefore, $A''(D_t) < 0$, so $A(D_t)$ is strictly concave in D_t for $D_t \geq 0$. ■

Proof. of Proposition 5. The proof begins by defining an operator T and proving that T maps concave, strictly increasing functions into concave, strictly increasing functions. Assume that (1) $f(\phi(t)) \geq 0$ is non-decreasing, (2) $\tau < \frac{\alpha}{1+\alpha}$, and (3) $V(\phi(t))$ is concave and strictly increasing. Lemma 22 implies that $A(D_t)$ is strictly concave in D_t for $D_t \geq 0$. Define the operator T by

$$(TV)(\phi(t)) = \max_{D \leq V(\phi(t))} \left[\frac{(1 - \tau)\phi(t) + \lambda \bar{v} + A(D)}{\rho + \lambda} \right], \quad (\text{A.1})$$

where $A(D(\phi(t)))$ defined in equation (10) depends on $V(\phi(t))$. Consider ϕ_1 and ϕ_2 and the respective corresponding optimal values of debt $\widehat{D}(\phi_1)$ and $\widehat{D}(\phi_2)$. Equation (14) implies $TV(\phi_i) = \frac{(1-\tau)\phi_i + \lambda \bar{v} + A(\widehat{D}(\phi_i))}{\rho + \lambda}$ and $\widehat{D}(\phi_i) \leq V(\phi_i)$, $i = 1, 2$. Now consider

$\phi(\gamma) = \gamma\phi_1 + (1 - \gamma)\phi_2$ and $D(\gamma) = \gamma\widehat{D}(\phi_1) + (1 - \gamma)\widehat{D}(\phi_2)$ for $0 \leq \gamma \leq 1$. Observe that $D(\gamma) = \gamma\widehat{D}(\phi_1) + (1 - \gamma)\widehat{D}(\phi_2) \leq \gamma V(\phi_1) + (1 - \gamma)V(\phi_2) \leq V(\phi(\gamma))$, so that $D(\gamma)$ is a feasible level of debt when $\phi = \phi(\gamma)$. Observe that $\gamma TV(\phi_1) + (1 - \gamma) TV(\phi_2) = \gamma \left[\frac{(1-\tau)\phi_1 + \lambda\bar{v} + A(\widehat{D}(\phi_1))}{\rho + \lambda} \right] + (1 - \gamma) \left[\frac{(1-\tau)\phi_2 + \lambda\bar{v} + A(\widehat{D}(\phi_2))}{\rho + \lambda} \right] \leq \frac{(1-\tau)\phi(\gamma) + \lambda\bar{v} + A(D(\gamma))}{\rho + \lambda} \leq \frac{(1-\tau)\phi(\gamma) + \lambda\bar{v} + \gamma A(\widehat{D}(\phi(\gamma)))}{\rho + \lambda} = TV(\phi(\gamma))$, where the first inequality follows from the concavity of $A(D)$ and the second inequality follows from the fact that $\widehat{D}(\phi(\gamma))$ maximizes $A(D)$ subject to $D \leq V(\phi(\gamma))$. Therefore, $TV(\phi(t))$ is concave. To prove that $TV(\phi)$ is strictly increasing, consider $\phi_2 > \phi_1$, and observe that $\widehat{D}(\phi_1) \leq V(\phi_1) < V(\phi_2)$ so that $\widehat{D}(\phi_1)$ is feasible when $\phi = \phi_2$. Then $TV(\phi_2) \geq \frac{(1-\tau)\phi_2 + \lambda\bar{v} + A(\widehat{D}(\phi_1))}{\rho + \lambda}$ and $TV(\phi_1) = \frac{(1-\tau)\phi_1 + \lambda\bar{v} + A(\widehat{D}(\phi_1))}{\rho + \lambda}$, so $TV(\phi_2) - TV(\phi_1) \geq \frac{(1-\tau)(\phi_2 - \phi_1)}{\rho + \lambda} > 0$. Therefore, $TV(\phi)$ is strictly increasing. Therefore, the operator T maps concave, strictly increasing functions into concave, strictly increasing functions. It remains to show that T is a contraction so that there is a unique concave, strictly increasing function that is a fixed point of T .

Monotonicity: I will show that if $V_2(\phi(t)) \geq V_1(\phi(t))$, then $TV_2(\phi(t)) \geq TV_1(\phi(t))$. Suppose that $V_2(\phi(t)) \geq V_1(\phi(t))$. Use equation (A.1)

$$TV_i(\phi(t)) = \frac{(1 - \tau)\phi(t) + \lambda\bar{v}_i + A_i(\widehat{D}_i(\phi(t)))}{\rho + \lambda}, \quad (\text{A.2})$$

where $\widehat{D}_i(\phi(t))$ is the optimal value of debt when the value function is $V_i(\phi)$ and $\phi = \phi(t)$. Use the definition of $A(D)$ in equation (10) to rewrite equation (A.2) as

$$TV_i(\phi(t)) = \frac{(1 - \tau)\phi(t) + \int g_i(\phi, D) dF(\phi)}{\rho + \lambda}, \quad (\text{A.3})$$

where

$$g_i(\phi, D) \equiv \begin{cases} \tau\rho D + \tau\lambda D + (1 - \alpha)(1 - \tau)\lambda V_i(\phi), & \text{if } V_i(\phi) < D \\ \tau\rho D + \lambda V_i(\phi), & \text{if } V_i(\phi) \geq D \end{cases}. \quad (\text{A.4})$$

The definition of $g_i(\phi, D)$ in equation (A.4) implies

$$\begin{aligned}
& (1 - \alpha)(1 - \tau)\lambda(V_2(\phi) - V_1(\phi)) \geq 0, & \text{if } V_2(\phi) < D \\
g_2(\phi, D) - g_1(\phi, D) = & \lambda V_2(\phi) - [\tau\lambda D + (1 - \alpha)(1 - \tau)\lambda V_1(\phi)], & \text{if } V_1(\phi) < D \leq V_2(\phi) . \\
& \lambda(V_2(\phi) - V_1(\phi)) \geq 0, & \text{if } V_1(\phi) \geq D
\end{aligned} \tag{A.5}$$

When $V_1(\phi) < D \leq V_2(\phi)$, $g_2(\phi, D) - g_1(\phi, D) = \lambda V_2(\phi) - [\tau\lambda D + (1 - \alpha)(1 - \tau)\lambda V_1(\phi)] \geq \lambda V_2(\phi) - [\tau + (1 - \alpha)(1 - \tau)]\lambda D \geq 0$, so that equation (A.5) implies that $g_2(\phi, D) \geq g_1(\phi, D)$. Since $\widehat{D}_1(\phi(t)) \leq V_1(\phi(t)) \leq V_2(\phi(t))$, $\widehat{D}_1(\phi(t))$ is feasible under $V_2(\phi(t))$. Therefore,

$$TV_2(\phi(t)) \geq \frac{(1 - \tau)\phi(t) + \int g_2(\phi, \widehat{D}_1(\phi(t))) dF(\phi)}{\rho + \lambda} \tag{A.6}$$

and since $g_2(\phi, D) \geq g_1(\phi, D)$,

$$TV_2(\phi(t)) \geq \frac{(1 - \tau)\phi(t) + \int g_1(\phi, \widehat{D}_1(\phi(t))) dF(\phi)}{\rho + \lambda}. \tag{A.7}$$

The right hand side of equation (A.7) equals $TV_1(\phi(t))$, so $TV_2(\phi(t)) \geq TV_1(\phi(t))$. Therefore, the operator T satisfies the monotonicity property of the Blackwell conditions.

Discounting: For $a \geq 0$,

$$[T(V + a)](\phi(t)) = \frac{(1 - \tau)\phi(t) + \int g_a(\phi, \widehat{D}_a(\phi(t))) dF(\phi)}{\rho + \lambda}, \tag{A.8}$$

where $\widehat{D}_a(\phi(t))$ is the optimal amount of debt under the value function $V(\phi(t)) + a$, and, following equation (A.4),

$$g_a(\phi, D) \equiv \begin{cases} \tau\rho D + \tau\lambda D + (1 - \alpha)(1 - \tau)\lambda(V(\phi) + a), & \text{if } V(\phi) + a < D \\ \tau\rho D + \lambda(V(\phi) + a), & \text{if } V(\phi) + a \geq D \end{cases}. \tag{A.9}$$

Since $\widehat{D}_a(\phi(t))$ is feasible under $V(\phi) + a$, it follows that $\widehat{D}_a(\phi(t)) \leq V(\phi(t)) + a$, which implies $\widehat{D}_a(\phi(t)) - a \leq V(\phi(t))$, so that $\widehat{D}_a(\phi(t)) - a$ is feasible under $V(\phi(t))$ and

hence

$$TV(\phi(t)) \geq \frac{(1-\tau)\phi(t) + \int g_0\left(\phi, \widehat{D}_a(\phi(t) - a)\right) dF(\phi)}{\rho + \lambda}. \quad (\text{A.10})$$

Subtracting equation (A.10) from equation (A.8) yields

$$[T(V+a)](\phi(t)) - TV(\phi(t)) \leq \frac{\int \left[g_a\left(\phi, \widehat{D}_a(\phi(t))\right) - g_0\left(\phi, \widehat{D}_a(\phi(t) - a)\right) \right] dF(\phi)}{\rho + \lambda}. \quad (\text{A.11})$$

The definition of $g_a(\phi, D)$ in equation (A.9) implies

$$g_a\left(\phi, \widehat{D}_a(\phi(t))\right) - g_0\left(\phi, \widehat{D}_a(\phi(t) - a)\right) = \begin{cases} [\tau\rho + (1-\alpha(1-\tau))\lambda]a, & \text{if } V(\phi) + a < \widehat{D}_a(\phi(t)) \\ [\tau\rho + \lambda]a, & \text{if } V(\phi) + a \geq \widehat{D}_a(\phi(t)) \end{cases}. \quad (\text{A.12})$$

Equations (A.11) and (A.12) imply that

$$[T(V+a)](\phi(t)) - TV(\phi(t)) \leq \frac{\tau\rho + \lambda}{\rho + \lambda}a, \quad (\text{A.13})$$

which can be written as

$$[T(V+a)](\phi(t)) \leq TV(\phi(t)) + \beta a, \quad (\text{A.14})$$

where $\beta \equiv \frac{\tau\rho + \lambda}{\rho + \lambda} < 1$. Hence, the operator T satisfies the discounting property of the Blackwell conditions.

Therefore, the operator T takes concave, strictly increasing functions on the domain $[\Phi_L, \Phi_H]$ into concave, strictly increasing functions on the domain $[\Phi_L, \Phi_H]$, and it has a unique fixed point $V(\phi(t))$. Therefore, $V(\phi(t))$ is concave on the domain $[\Phi_L, \Phi_H]$. ■

Proof. of Corollary 6. Lemma 22 and Proposition 5 immediately imply that $A(D)$ is strictly concave in D for $V(\Phi_L) \leq D \leq V(\Phi_H)$ under the conditions in this corollary. ■

Proof. of Proposition 7. If $\phi(t) \geq \phi^*$, then $V(\phi(t)) \geq V(\phi^*) = D^*$ so the constraint $D \leq V(\phi(t))$ will not bind and hence the optimal value of debt is $\widehat{D}(\phi(t)) = D^*$. Therefore, for any $\phi(t) \geq \phi^*$, the value of $V(\phi(t))$ defined in equation (14) is $V(\phi(t)) = \frac{(1-\tau)\phi(t) + \lambda\bar{v} + A(D^*)}{\rho + \lambda}$, so

$$V(\phi_2) - V(\phi_1) = \frac{1-\tau}{\rho + \lambda}(\phi_2 - \phi_1) \text{ for } \phi_1 \geq \phi^* \text{ and } \phi_2 \geq \phi^*. \quad (\text{A.15})$$

Setting ϕ_1 in equation (A.15) equal to ϕ^* , and using $V(\phi^*) = D^*$ yields

$$V(\phi(t)) = \frac{1-\tau}{\rho+\lambda} [\phi(t) - \phi^*] + D^*, \quad \text{for } \phi(t) \geq \phi^*. \quad (\text{A.16})$$

Since $\widehat{D}(\phi(t)) = D^*$ when $\phi(t) \geq \phi^*$, equations (A.16) and (15) imply

$$S(\phi(t)) = \frac{1-\tau}{\rho+\lambda} [\phi(t) - \phi^*], \quad \text{for } \phi(t) \geq \phi^*. \quad (\text{A.17})$$

For values of $\phi(t) < \phi^*$, setting $D = D^*$ would violate the constraint $D \leq V(\phi(t))$ because $V(\phi(t)) \leq V(\phi^*) = D^*$, so the optimal value of D , $\widehat{D}(\phi(t))$, will be less than D^* . Since $A(D)$ is strictly concave in D (Corollary 6), the constraint $V(\phi(t)) \geq D$ will strictly bind for any $\phi(t) < \phi^*$ so that

$$V(\phi) \equiv \widehat{D}(\phi), \quad \text{for } \phi(t) \leq \phi^*,$$

and

$$S(\phi(t)) = 0, \quad \text{for } \phi(t) \leq \phi^*. \quad (\text{A.18})$$

■

Proof. of Corollary 8. For $\phi(t) \geq \phi^*$, use Proposition 7 to substitute D^* for $\widehat{D}(\phi(t))$ and $\frac{1-\tau}{\rho+\lambda} [\phi(t) - \phi^*] + D^*$ for $V(\phi(t))$ in the definition of the leverage ratio, $L(\phi(t)) \equiv \frac{\widehat{D}(\phi(t))}{V(\phi(t))}$, and divide numerator and denominator by D^* to obtain $L(\phi(t)) = \frac{1}{\frac{1-\tau}{\rho+\lambda} \frac{\phi(t) - \phi^*}{D^*} + 1}$. For $\phi(t) \leq \phi^*$,

Proposition 7 implies $\widehat{D}(\phi(t)) = V(\phi(t))$, so the leverage ratio equals one. ■

Proof. of Proposition 9. If $\phi(t) < \phi^*$, then $\widehat{D}(\phi(t)) = V(\phi(t))$ so $V(\phi) < \widehat{D}(\phi(t))$ if and only if $V(\phi) < V(\phi(t))$, i.e., if and only if $\phi < \phi(t)$, since $V(\phi)$ is strictly increasing. Therefore, $P(\phi(t)) = \lambda F(\phi(t)) < \lambda F(\phi^*)$. If $\phi(t) \geq \phi^*$, then $\widehat{D}(\phi(t)) = D^*$, so $V(\phi) < \widehat{D}(\phi(t)) = D^* = V(\phi^*)$ if and only if $V(\phi) < V(\phi^*)$, i.e., if and only if $\phi < \phi^*$, since $V(\phi)$ is strictly increasing. Therefore, $P(\phi(t)) = \lambda F(\phi^*) \leq \lambda F(\phi(t))$. Therefore, $P(\phi(t)) = \lambda \min [F(\phi(t)), F(\phi^*)]$. ■

Proof. of Proposition 10. From equation (18), $P(\widehat{D}(\phi(t))) \equiv \lambda \int_{V(\phi) < \widehat{D}(\phi(t))} dF(\phi) = \lambda \int_{\phi < V^{-1}(\widehat{D}(\phi(t)))} dF(\phi)$, so $\frac{dP}{d\widehat{D}(\phi(t))} = V^{-1'}(\widehat{D}(\phi(t))) f(V^{-1}(\widehat{D}(\phi(t)))) > 0$. If $\phi(t) < \phi^*$, then $\frac{d\widehat{D}(\phi(t))}{d\phi(t)} > 0$, so $\frac{dP}{d\phi(t)} = \frac{dP}{d\widehat{D}(\phi(t))} \frac{d\widehat{D}(\phi(t))}{d\phi(t)} > 0$. If $\phi(t) \geq \phi^*$, then $\frac{d\widehat{D}(\phi(t))}{d\phi(t)} \equiv 0$, so $\frac{dP}{d\phi(t)} =$

$\frac{dF}{d\widehat{D}(\phi(t))} \frac{d\widehat{D}(\phi(t))}{d\phi(t)} \equiv 0$. From equation (3), $r(t) = \rho + \lambda \int_{V(\phi) < \widehat{D}(\phi(t))} \left[1 - (1 - \alpha) \frac{V(\phi)}{\widehat{D}(\phi(t))} \right] dF(\phi) = \rho + \lambda \int_{\phi < V^{-1}(\widehat{D}(\phi(t)))} \left[1 - (1 - \alpha) \frac{V(\phi)}{\widehat{D}(\phi(t))} \right] dF(\phi)$, so $\frac{dr(t)}{d\widehat{D}(\phi(t))} = \lambda V^{-1'}(\widehat{D}(\phi(t))) \alpha f(V^{-1}(\widehat{D}(\phi(t))))$
 $+ \lambda \int_{\phi < V^{-1}(\widehat{D}(\phi(t)))} \left[(1 - \alpha) \frac{V(\phi)}{[\widehat{D}(\phi(t))]^2} \right] dF(\phi) > 0$. If $\phi(t) < \phi^*$, then $\frac{d\widehat{D}(\phi(t))}{d\phi(t)} > 0$, so $\frac{dr(t)}{d\phi(t)} = \frac{dr(t)}{d\widehat{D}(\phi(t))} \frac{d\widehat{D}(\phi(t))}{d\phi(t)} > 0$. If $\phi(t) \geq \phi^*$, then $\frac{d\widehat{D}(\phi(t))}{d\phi(t)} \equiv 0$, so $\frac{dr(t)}{d\phi(t)} = \frac{dr(t)}{d\widehat{D}(\phi(t))} \frac{d\widehat{D}(\phi(t))}{d\phi(t)} \equiv 0$. ■

Proof. of Proposition 11. Suppose that $\widehat{D}(\phi(t)) = V(\phi(t))$. Substitute $V(\phi(t))$ for $\widehat{D}(\phi(t))$ in the definition of $A(D_t)$ in equation (11) and then substitute the resulting equation into equation (12) to obtain

$$\rho V(\phi(t)) = (1 - \tau)\phi(t) + \lambda(\bar{v} - V(\phi(t))) + \tau r(t)V(\phi(t)) - \alpha\lambda \int_{V(\phi) < V(\phi(t))} V(\phi) dF(\phi). \quad (\text{A.19})$$

Now substitute $V(\phi(t))$ for D_t in equation (3) to obtain

$$r(t)V(\phi(t)) = \rho V(\phi(t)) + \lambda \int_{V(\phi) < V(\phi(t))} [V(\phi(t)) - (1 - \alpha)V(\phi)] dF(\phi). \quad (\text{A.20})$$

Combine equations (A.19) and (A.20) to eliminate $\rho V(\phi(t))$ and obtain

$$0 = (1 - \tau)\phi(t) + \lambda(\bar{v} - V(\phi(t))) - (1 - \tau)r(t)V(\phi(t)) + \lambda \int_{V(\phi) < V(\phi(t))} [V(\phi(t)) - V(\phi)] dF(\phi). \quad (\text{A.21})$$

Substitute $\int [V(\phi) - V(\phi(t))] dF(\phi)$ for $\bar{v} - V(\phi(t))$ in equation (A.21) and rearrange to obtain

$$(1 - \tau)[\phi(t) - r(t)V(\phi(t))] = -\lambda \int_{V(\phi) \geq V(\phi(t))} [V(\phi) - V(\phi(t))] dF(\phi) \leq 0. \quad (\text{A.22})$$

When $\widehat{D}(\phi(t)) = V(\phi(t))$, $(1 - \tau)y(t) = (1 - \tau)[\phi(t) - r(t)V(\phi(t))]$, so equation (A.22) implies that

$$(1 - \tau)y(t) = -\lambda \int_{V(\phi) \geq V(\phi(t))} [V(\phi) - V(\phi(t))] dF(\phi) \leq 0, \quad (\text{A.23})$$

with strict inequality if $\phi(t) < \Phi_H$, which proves Statement 1.

Differentiate equation (A.23) with respect to $\phi(t)$ to obtain

$$\frac{d[(1-\tau)y(t)]}{d\phi(t)} = \lambda V'(\phi(t)) \int_{V(\phi) \geq V(\phi(t))} dF(\phi) > 0, \quad (\text{A.24})$$

which proves Statement 2. Finally, to prove Statement 3, if t is not a time of regime change, then $dD_t = 0$, so equation (6) implies $dX(t) = (1-\tau)y(t)dt$, which is non-positive from Statement 1. ■

Proof. of Proposition 12. Assume that the firm is in Scenario I so that $\omega_1 > 0$. The complementary slackness condition in equation (21) implies $D^* = V(\Phi_L)$. Therefore, $\phi^* = V^{-1}(D^*) = V^{-1}(V(\Phi_L)) = \Phi_L$. Proposition 7 implies

$$V(\phi(t)) = \frac{1-\tau}{\rho+\lambda} [\phi(t) - \Phi_L] + V(\Phi_L). \quad (\text{A.25})$$

Taking the unconditional expectation of both sides of equation (A.25) yields

$$\bar{v} = \frac{1-\tau}{\rho+\lambda} [E\{\phi\} - \Phi_L] + V(\Phi_L), \quad (\text{A.26})$$

where $\bar{v} \equiv E\{V(\phi)\}$. Evaluate $A(D)$ in equation (10) at $D = V(\Phi_L)$ to obtain $A(V(\Phi_L)) = \tau\rho V(\Phi_L)$ so that equation (14) implies

$$V(\Phi_L) = \frac{(1-\tau)\Phi_L + \lambda\bar{v} + \tau\rho V(\Phi_L)}{\rho+\lambda}. \quad (\text{A.27})$$

Equations (A.26) and (A.27) are two linear equations in $V(\Phi_L)$ and \bar{v} , which can be solved to obtain $V(\Phi_L) = \frac{1}{\rho+\lambda} \left[\Phi_L + \frac{\lambda}{\rho} E\{\phi\} \right]$. Therefore, $\widehat{D}(\phi(t)) = D^* = V(\Phi_L) = \frac{1}{\rho+\lambda} \left[\Phi_L + \frac{\lambda}{\rho} E\{\phi\} \right]$. Since $\phi^* = \Phi_L$, Proposition 7 implies that $S(\phi(t)) = \frac{1-\tau}{\rho+\lambda} [\phi(t) - \Phi_L]$ and $V(\phi(t)) = D^* + S(\phi(t))$. ■

Proof. of Proposition 13. Since $\omega_2 > 0$, equation (22) implies $D^* = V(\Phi_H)$ and hence $\widehat{D}(\Phi_H) = V(\Phi_H)$. Evaluating equation (14) at $\phi(t) = \Phi_H$, and setting $\widehat{D}(\Phi_H) = V(\Phi_H)$ yields $V(\Phi_H) = \frac{(1-\tau)\Phi_H + \lambda\bar{v} + A(V(\Phi_H))}{\rho+\lambda}$. The definition of $A(D)$ in equation (10) implies $A(V(\Phi_H)) \equiv \tau(\rho+\lambda)V(\Phi_H) - [\alpha + \tau(1-\alpha)]\lambda\bar{v}$. Substituting this expression for $A(V(\Phi_H))$ into the preceding equation, $V(\Phi_H) = \frac{(1-\tau)\Phi_H + \lambda\bar{v} + A(V(\Phi_H))}{\rho+\lambda}$, yields $V(\Phi_H) = \frac{(1-\tau)\Phi_H + \lambda\bar{v} + \tau(\rho+\lambda)V(\Phi_H) - [\alpha + \tau(1-\alpha)]\lambda\bar{v}}{\rho+\lambda}$. Multiply both sides by $\rho+\lambda$ and then subtract $\tau(\rho+\lambda)V(\Phi_H)$

from both sides to obtain $(1 - \tau)(\rho + \lambda)V(\Phi_H) = (1 - \tau)\Phi_H + \lambda(1 - \tau)(1 - \alpha)\bar{v}$, which implies $(\rho + \lambda)V(\Phi_H) = \Phi_H + \lambda(1 - \alpha)\bar{v}$, or equivalently, $V(\Phi_H) = \frac{\Phi_H + \lambda(1 - \alpha)\bar{v}}{\rho + \lambda}$. ■

Proof. of Lemma 14. Since, $\omega_1 = \omega_1 = 0$, equation (20) implies $A'(D^*) = 0$. Since $V(\Phi_L) \leq D^* \leq V(\Phi_H)$ and $V(\phi)$ is strictly increasing in ϕ , there is a unique $\phi^* \in [\Phi_L, \Phi_H]$ such that $V(\phi^*) = D^*$. Therefore, $\widehat{D}(\phi^*) = D^*$. Differentiating $V(\phi(t))$ in equation (14) with respect to $\phi(t)$ yields $V'(\phi(t)) = \frac{1}{\rho + \lambda} \left[1 - \tau + A'(\widehat{D}(\phi(t)))\widehat{D}'(\phi(t)) \right]$. Evaluating this derivative at $\phi(t) = \phi^*$, where $\widehat{D}(\phi(t)) = \widehat{D}(\phi^*) = D^*$, and using $A'(D^*) = 0$, yields $V'(\phi^*) = \frac{1 - \tau}{\rho + \lambda}$. Therefore, $V^{-1'}(D^*) = \frac{\rho + \lambda}{1 - \tau}$. ■

Proof. of Proposition 15. The proof proceeds by presenting, for each scenario, the value of D^* that satisfies the first-order condition in equation (20) and the values of ω_1 and ω_2 that, along with D^* , lead to satisfaction of the complementary slackness conditions in equations (21) and (22). Because the maximand in equation (19) is strictly concave in D , a value of D that satisfies equations (20), (21), and (22) is the unique value of D^* .

(I) Assume that $\tau < \tau_L$. Suppose that $D^* = V(\Phi_L)$, $\omega_1 > 0$, and $\omega_2 = 0$. Since $D^* = V(\Phi_L)$, the complementary slackness condition in equation (21) is satisfied and since $\omega_2 = 0$, the complementary slackness condition in equation (22) is satisfied. Since $\omega_2 - \omega_1 < 0$, the first-order condition in equation (20) implies that $A'(D^*) < 0$. Thus, it suffices to show that $A'(V(\Phi_L)) < 0$. Differentiating $A(D)$ in equation (10) with respect to D and evaluating the derivative at $D = D^* = V(\Phi_L)$, and using $\phi^* = V^{-1}(D^*) = V^{-1}(V(\Phi_L)) = \Phi_L$, yields

$$A'(D^*) = \tau\rho - \alpha(1 - \tau)\lambda V^{-1'}(V(\Phi_L))f(\Phi_L)V(\Phi_L). \quad (\text{A.28})$$

Since $\omega_1 > 0$ and $\omega_2 = 0$, the firm is in Scenario I. Therefore, Statement 3 of Proposition 12 implies $V(\Phi_L) = D^* = \frac{1}{\rho + \lambda} \left[\Phi_L + \frac{\lambda}{\rho} E\{\phi\} \right]$ and also implies that $V'(\phi(t)) = \frac{1 - \tau}{\rho + \lambda}$ for all $\phi(t)$, so that $V^{-1'}(V(\Phi_L)) = \frac{\rho + \lambda}{1 - \tau}$. Therefore, equation (A.28) can be written as

$$A'(D^*) = \tau\rho - \alpha \left[\Phi_L + \frac{\lambda}{\rho} E\{\phi\} \right] \lambda f(\Phi_L). \quad (\text{A.29})$$

Use the definition $\tau_L \equiv \alpha \frac{\lambda}{\rho} f(\Phi_L) \left(\Phi_L + \frac{\lambda}{\rho} E\{\phi\} \right)$ to rewrite equation (A.29) as

$$A'(D^*) = (\tau - \tau_L)\rho < 0, \quad (\text{A.30})$$

where the inequality follows from $\rho > 0$ and the assumption that $\tau < \tau_L$. Therefore, $D^* = V(\Phi_L)$, $\omega_1 > 0$, and $\omega_2 = 0$ satisfy the first-order condition in equation (20) and the complementary slackness conditions in equations (21) and (22), and the firm is in Scenario I.

(II) Assume that $\tau_L \leq \tau \leq \tau_H$. Suppose that $\omega_1 = \omega_2 = 0$. Therefore, the complementary slackness conditions in equations (21) and (22) are satisfied. Since $\omega_2 - \omega_1 = 0$, the first-order condition in equation (20) implies that $A'(D^*) = 0$. Thus, it suffices to show that $A'(D) = 0$ for some $D \in [V(\Phi_L), V(\Phi_H)]$. The proof will proceed by showing that if $D^* = V(\Phi_L)$ so that $\phi^* = \Phi_L$, then $A'(D^*) \geq 0$ and that if $D^* = V(\Phi_H)$ so that $\phi^* = \Phi_H$, then $A'(D^*) \leq 0$, so there is a $D^* \in [V(\Phi_L), V(\Phi_H)]$ that satisfies $A'(D^*) = 0$.

Differentiate $A(D)$ in equation (10) with respect to D and use Lemma 14 to obtain

$$A'(D^*) = \tau(\rho + \lambda F(V^{-1}(D^*))) - (\rho + \lambda)\alpha\lambda D^* f(V^{-1}(D^*)). \quad (\text{A.31})$$

Consider the possibility that $D^* = V(\Phi_L)$, which implies $\phi^* = \Phi_L$. Evaluate $A(D)$ in equation (10) at $D = V(\Phi_L)$ to obtain $A(V(\Phi_L)) = \tau\rho V(\Phi_L)$, so equation (14) evaluated at $\phi(t) = \Phi_L$ and $\widehat{D}(\phi(t)) = D^* = V(\Phi_L)$ yields $V(\Phi_L) = \frac{(1-\tau)\Phi_L + \lambda\bar{v} + \tau\rho V(\Phi_L)}{\rho + \lambda}$, which implies

$$((1-\tau)\rho + \lambda)V(\Phi_L) = (1-\tau)\Phi_L + \lambda\bar{v}. \quad (\text{A.32})$$

Taking the expectation of both sides of $V(\phi(t)) = \frac{1-\tau}{\rho + \lambda}[\phi(t) - \phi^*] + D^*$ from Proposition 7 yields

$$\bar{v} = \frac{1-\tau}{\rho + \lambda}[E\{\phi\} - \Phi_L] + V(\Phi_L), \quad \text{when } D^* = V(\Phi_L) \text{ and } \phi^* = \Phi_L. \quad (\text{A.33})$$

Substituting equation (A.33) into equation (A.32) and rearranging yields

$$V(\Phi_L) = \frac{1}{\rho + \lambda}\Phi_L + \frac{\lambda/\rho}{\rho + \lambda}E\{\phi\}. \quad (\text{A.34})$$

Now evaluate $A'(D^*)$ in equation (A.31) at $D^* = V(\Phi_L)$ and use the expression for $V(\Phi_L)$ in equation (A.34) to obtain

$$A'(V(\Phi_L)) = \tau\rho - \alpha\lambda \left[\Phi_L + \frac{\lambda}{\rho}E\{\phi\} \right] f(\Phi_L). \quad (\text{A.35})$$

Use the definition $\tau_L \equiv \alpha \frac{\lambda}{\rho} f(\Phi_L) \left(\Phi_L + \frac{\lambda}{\rho} E\{\phi\} \right)$ to rewrite equation (A.35) as

$$A'(V(\Phi_L)) = (\tau - \tau_L) \rho \geq 0, \quad (\text{A.36})$$

where the inequality follows from $\rho > 0$ and the assumption that $\tau \geq \tau_L$.

Evaluate $A'(D)$ at $D = V(\Phi_H)$ to obtain

$$A'(V(\Phi_H)) = (\tau - \alpha \lambda V(\Phi_H) f(\Phi_H)) (\rho + \lambda). \quad (\text{A.37})$$

Evaluate $A(D)$ in equation (10) at $D = V(\Phi_H)$ to obtain $A(V(\Phi_H)) = \tau(\rho + \lambda)V(\Phi_H) - [\alpha + \tau(1 - \alpha)]\lambda\bar{v}$, so equation (14) evaluated at $\phi(t) = \Phi_H$ and $\widehat{D}(\phi(t)) = D^* = V(\Phi_H)$ yields $V(\Phi_H) = \frac{(1-\tau)\Phi_H + \lambda\bar{v} + \tau(\rho + \lambda)V(\Phi_H) - [\alpha + \tau(1-\alpha)]\lambda\bar{v}}{\rho + \lambda}$, which implies

$$V(\Phi_H) = \frac{1}{\rho + \lambda} [\Phi_H + (1 - \alpha)\lambda\bar{v}]. \quad (\text{A.38})$$

Substitute equation (A.38) into equation (A.37) and use the definition $\tau_H \equiv \alpha \frac{\lambda}{\rho + \lambda} f(\Phi_H) (\Phi_H + (1 - \alpha)\lambda\bar{v})$ to obtain

$$A'(V(\Phi_H)) = (\tau - \tau_H) (\rho + \lambda) \leq 0, \quad (\text{A.39})$$

where the inequality follows from $\rho + \lambda > 0$ and the assumption $\tau \leq \tau_H$. Therefore, there is a $D^* \in [V(\Phi_L), V(\Phi_H)]$ for which $A'(D^*) = 0$ when $\omega_1 = \omega_2 = 0$. Hence, the firm is in Scenario II. Proposition 7 immediately implies that $\widehat{D}(\phi(t)) = D^*$ for $\phi(t) \geq \phi^*$ and $\widehat{D}(\phi(t)) = V(\phi(t))$ for $\phi(t) \leq \phi^*$.

(III) Assume that $\frac{\alpha}{1+\alpha} > \tau > \tau_H$. Suppose that $D^* = V(\Phi_H)$, $\omega_1 = 0$, and $\omega_2 > 0$, so that the firm is in Scenario III. Since $\omega_1 = 0$, the complementary slackness condition in equation (21) is satisfied. Since $D^* = V(\Phi_H)$, the complementary slackness condition in equation (22) is satisfied. Since $\omega_2 - \omega_1 > 0$, the first-order condition in equation (20) implies that $A'(D^*) > 0$. Thus, it suffices to show that $A'(V(\Phi_H)) > 0$. Differentiating $A(D)$ in equation (10) with respect to D and evaluating the derivative at $D = D^* = V(\Phi_H)$, and using $\phi^* = V^{-1}(D^*) = V^{-1}(V(\Phi_H)) = \Phi_H$, yields

$$A'(V(\Phi_H)) = \tau(\rho + \lambda) - \alpha(1 - \tau)\lambda V^{-1'}(V(\Phi_H)) f(\Phi_H) V(\Phi_H). \quad (\text{A.40})$$

Proposition 1 implies that $V'(V(\Phi_H)) \geq \frac{1-\tau}{\rho+\lambda}$, so $V^{-1}(V(\Phi_H)) < \frac{\rho+\lambda}{1-\tau}$ and equation (A.40) implies

$$A'(V(\Phi_H)) \geq \tau(\rho + \lambda) - \alpha(\rho + \lambda)\lambda f(\Phi_H)V(\Phi_H). \quad (\text{A.41})$$

Use $V(\Phi_H) = \frac{1}{\rho+\lambda}[\Phi_H + (1-\alpha)\lambda\bar{v}]$ from Proposition 13 and the definition $\tau_H \equiv \alpha\frac{\lambda}{\rho+\lambda}f(\Phi_H)(\Phi_H + (1-\alpha)\lambda\bar{v})$ to rewrite equation (A.41) as

$$A'(V(\Phi_H)) \geq (\tau - \tau_H)(\rho + \lambda) > 0, \quad (\text{A.42})$$

where the second inequality follows from $\rho + \lambda > 0$ and $\tau > \tau_H$. Therefore, $D^* = V(\Phi_H)$, $\omega_1 = 0$, and $\omega_2 > 0$ satisfy the first-order condition in equation (20) and the complementary slackness conditions in equations (21) and (22), and the firm is in Scenario III. ■

Proof. of Proposition 16. Let $y(\phi(t); 0)$ be the optimal taxable income under the original distribution ($m = 0$) and recall from Statement 2 of Proposition 11 and footnote 20 that $y(\phi(t); 0)$ is strictly increasing in $\phi(t)$. Consider the following feasible financing plan under $G_m(\phi)$: (1) if $\phi(t) \in [\Phi_L(m), \Phi_H(0)]$, set $D(\phi(t); m) = \widehat{D}(\phi(t); 0)$, where $\widehat{D}(\phi(t); 0)$ is the optimal amount of debt under $G_0(\phi)$, and pay the same default premium that would be paid under $G_0(\phi)$, so that taxable income $y(\phi(t); m) = y(\phi(t); 0)$. It will be verified below that the default premium under the new distribution, $G_m(\phi)$, is no higher than under the original distribution, $G_0(\phi)$. (2) if $\phi(t) \in (\Phi_H(0), \Phi_H(m)]$, set $D(\phi(t); m) = \widehat{D}(\Phi_H(0); 0)$ and pay the same default premium that would be paid under $G_0(\phi)$ when $\phi(t) = \Phi_H(0)$, so that $y(\phi(t); m) = y(\Phi_H(0); m) + \phi(t) - \Phi_H(0) = y(\Phi_H(0); 0) + \phi(t) - \Phi_H(0) > y(\Phi_H(0); 0)$ (3) at any future date of regime change, t_j , (a) if $\phi(t_j) \in [\Phi_L(m), \Phi_H(0)]$, default if and only if the firm would optimally default under $G_0(\phi)$ when $\phi = \phi(t_j)$ and (b) if $\phi(t_j) \in (\Phi_H(0), \Phi_H(m)]$, do not default. For $\phi(t_0) \in [\Phi_L(m), \Phi_H(0)]$, the firm will have the same cash flow $G_m(\phi)$ as it would under the optimal policy under $G_0(\phi)$. However, with the financing plan under $G_m(\phi)$, the continuation value under $G_m(\phi)$ will exceed the continuation value under $G_0(\phi)$ by $\int_{\Phi_H(0)}^{\Phi_H(m)} (1-\tau)y(\phi(t); m)dF(\phi - m) - \int_{\Phi_L(0)}^{\Phi_L(m)} (1-\tau)y(\phi(t); 0)dF(\phi) > (1-\tau)y(\Phi_H(0); 0) \int_{\Phi_H(0)}^{\Phi_H(m)} dF(\phi - m) - (1-\tau)y(\Phi_L(m); 0) \int_{\Phi_L(0)}^{\Phi_L(m)} dF(\phi) = (1-\tau)y(\Phi_H(0); 0) \int_{\Phi_H(0)-m}^{\Phi_H(0)} dF(\phi) - (1-\tau)y(\Phi_L(m); 0) \int_{\Phi_L(0)}^{\Phi_L(0)+m} dF(\phi) > 0$, where the first inequality follows from $y(\phi(t); m) > y(\Phi_H(0); 0)$ if $\phi(t) > \Phi_H(0)$, as shown above, and from the fact that $y(\phi(t); 0)$ is strictly increasing in $\phi(t)$ as shown earlier in this proof;

the second inequality follows from $y(\Phi_H(0); 0) > y(\Phi_L(m); 0)$ and $\int_{\Phi_H(0)-m}^{\Phi_H(0)} dF(\phi) > \int_{\Phi_L(0)}^{\Phi_L(0)+m} dF(\phi)$ since $f(\phi) \equiv F'(\phi)$ is non-decreasing. Therefore, the continuation value under $G_m(\phi)$ exceeds the continuation value under $G_0(\phi)$, and hence $V(\phi(t); m) > V(\phi(t); 0)$ for $m > 0$. Finally, because the continuation value of the firm for any given $\phi(t)$ is higher under $G_m(\phi)$ than under $G_0(\phi)$, the default premium under $G_m(\phi)$ is no higher than under $G_0(\phi)$. ■

Proof. of Proposition 17. From Proposition 12, $\widehat{D}(\phi(t); m) = \frac{1}{\rho+\lambda} \left[\Phi_L(m) + \frac{\lambda}{\rho} E\{\phi; m\} \right]$ in Scenario I, so $\frac{d\widehat{D}(\phi(t); m)}{dm} = \frac{1}{\rho+\lambda} \left[\frac{d\Phi_L(m)}{dm} + \frac{\lambda}{\rho} \frac{dE\{\phi; m\}}{dm} \right]$, where $E\{\phi; m\} = \int_{\Phi_L(m)}^{\Phi_H(m)} \phi dF(\phi - m)$. Since $\frac{d\Phi_L(m)}{dm} = \frac{dE\{\phi; m\}}{dm} = 1$, $\frac{d\widehat{D}(\phi(t); m)}{dm} = \frac{1}{\rho+\lambda} \left(1 + \frac{\lambda}{\rho} \right) = \frac{1}{\rho}$ (Statement 1). From Proposition 12, $S(\phi(t); m) = \frac{1-\tau}{\rho+\lambda} (\phi(t) - \Phi_L(m))$ in Scenario I, so $\frac{dS(\phi(t); m)}{dm} = -\frac{1-\tau}{\rho+\lambda} \frac{d\Phi_L(m)}{dm}$. Since $\frac{d\Phi_L(m)}{dm} = 1$, $\frac{dS(\phi(t); m)}{dm} = -\frac{1-\tau}{\rho+\lambda}$ (Statement 2). Since $V(\phi(t); m) = \widehat{D}(\phi(t); m) + S(\phi(t); m)$, $\frac{dV(\phi(t); m)}{dm} = \frac{d\widehat{D}(\phi(t); m)}{dm} + \frac{dS(\phi(t); m)}{dm} = \frac{1}{\rho} - \frac{1-\tau}{\rho+\lambda} = \frac{1}{\rho} \frac{\rho\tau+\lambda}{\rho+\lambda} > 0$ (Statement 3). Since $L(\phi(t); m) \equiv \frac{\widehat{D}(\phi(t); m)}{V(\phi(t); m)}$, $\frac{dL(\phi(t); m)}{dm} = \frac{1}{V(\phi(t); m)} \frac{1}{\rho} \left[1 - L(\phi(t); m) \frac{\tau\rho+\lambda}{\rho+\lambda} \right] > 0$, where the inequality follows from $L(\phi(t); m) \leq 1$ and $\tau < 1$ (Statement 4). ■

Proof. of Lemma 18. Suppose that the distribution shifts to the right by $d > 0$ and consider whether $\widehat{\phi} \equiv \phi^*(m) + d$ satisfies the first-order condition in equation (34) under the new distribution. That is, consider whether $\Psi \equiv \tau \left(\rho + \lambda F \left(\widehat{\phi} - (m + d) \right) \right) - \alpha \lambda (\rho + \lambda) V \left(\widehat{\phi}, m + d \right) f \left(\widehat{\phi} - (m + d) \right)$ equals zero. The definition of $\widehat{\phi}$ implies that $\widehat{\phi} - (m + d) = \phi^*(m) - m$, so $\Psi = \tau (\rho + \lambda F(\phi^*(m) - m)) - \alpha \lambda (\rho + \lambda) V(\phi^*(m) + d, m + d) f(\phi^*(m) - m)$. Use equation (34) to replace $\tau (\rho + \lambda F(\phi^*(m) - m))$ by $\alpha \lambda (\rho + \lambda) V(\phi^*(m), m) f(\phi^*(m) - m)$ in Ψ to obtain $\Psi = \alpha \lambda (\rho + \lambda) f(\phi^*(m) - m) [V(\phi^*(m), m) - V(\phi^*(m) + d, m + d)]$. Since $d > 0$, we have $V(\phi^*(m) + d, m + d) > V(\phi^*(m), m + d) > V(\phi^*(m), m)$ where the first inequality follows from Proposition 1 and the second inequality follows from Proposition 16. Since $\alpha \lambda f(\phi^*(m) - m) > 0$ and $V(\phi^*(m) + d, m + d) > V(\phi^*(m), m)$, we have $\Psi < 0$, that is, $\frac{\partial A(V(\widehat{\phi}, m + d); m + d)}{\partial D} < 0$ and since (from Corollary 6) $A(D; m)$ is strictly concave in D , the value of D^* is less than $V(\widehat{\phi}, m + d)$ so $\phi^*(m + d) < \widehat{\phi} \equiv \phi^*(m) + d$. Therefore, for any $d > 0$, $\frac{\phi^*(m + d) - \phi^*(m)}{d} < 1$, so $\phi^{*'}(m) < 1$. ■

Proof. of Proposition 19. Proposition 9 implies that

$$P(\phi(t); m) = \lambda \min [F(\phi(t) - m), F(\phi^*(m) - m)]. \quad \frac{dF(\phi^*(m) - m)}{dm} = f(\phi^*(m) - m) (\phi^{*'}(m) - 1)$$

< 0 , where the strict inequality follows from $f(\phi^*(m) - m) > 0$ and Lemma 18. Also, $\frac{dF(\phi(t)-m)}{dm} = -f(\phi(t) - m) < 0$. Therefore, $\frac{dP(\phi(t);m)}{dm} < 0$ (Statement 1).

Since $D^*(m) \equiv V(\phi^*(m); m)$, $\text{sign}(D^{*'}(m)) = \text{sign}\left(\frac{dV(\phi^*(m);m)}{dm}\right)$. Since $\tau > 0$, $\alpha\lambda(\rho + \lambda)f(\phi^*(0)) > 0$, and $\phi^{*'}(m) < 1$ (Lemma 18), equation (35) implies $\text{sign}\left(\frac{dV(\phi^*(m);m)}{dm}\right) = -\text{sign}(\chi(\phi^*(m) - m))$. Therefore, $\text{sign}\left(\frac{dV(\phi^*(m);m)}{dm}\right) = \text{sign}(D^{*'}(m)) = -\text{sign}(\chi(\phi^*(m) - m))$ (Statement 2).

Statement 2 implies that if $\chi(\phi^*(m) - m) \geq 0$, then $\frac{dV(\phi^*(m);m)}{dm} \leq 0$. Since $\frac{dV(\phi^*(m);m)}{dm} = \frac{\partial V(\phi^*(m);m)}{\partial \phi} \phi^{*'}(m) + \frac{\partial V(\phi^*(m);m)}{\partial m}$, it follows that if $\chi(\phi^*(m) - m) \geq 0$, then $\frac{\partial V(\phi^*(m);m)}{\partial \phi} \phi^{*'}(m) = \frac{dV(\phi^*(m);m)}{dm} - \frac{\partial V(\phi^*(m);m)}{\partial m} \leq -\frac{\partial V(\phi^*(m);m)}{\partial m} < 0$, where the final inequality follows from Proposition 16. Finally, since $\frac{\partial V(\phi^*(m);m)}{\partial \phi} > 0$ (Proposition 1), $\phi^{*'}(m) \leq -\frac{\frac{\partial V(\phi^*(m);m)}{\partial m}}{\frac{\partial V(\phi^*(m);m)}{\partial \phi}} < 0$ (Statement 3). ■

Proof. of Proposition 20. Define the operator T as

$TV(\phi(t)) = \max_{D \leq V(\phi(t))} \frac{(1-\tau)\phi(t) + \lambda \int V(\phi(t))f(\phi(t'), \phi(t))d\phi(t') + A(D, \phi(t))}{\rho + \lambda}$. 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 satisfies monotonicity and discounting.

Mapping non-decreasing functions into strictly increasing functions: Assume that $V(\phi)$ is non-decreasing in ϕ . Therefore, for $\phi_2 > \phi_1$, $V(\phi_2) \geq V(\phi_1)$ so $\widehat{D}(\phi_1) \leq V(\phi_1) \leq V(\phi_2)$ and hence $\widehat{D}(\phi_1)$ is feasible when $\phi(t) = \phi_2$.

Substituting the definition of $A(D, \phi(t))$ from equation (38) into the expression for $TV(\phi(t))$ and evaluating the resulting expression at $D = \widehat{D}(\phi(t))$ yields $TV(\phi(t)) = \frac{(1-\tau)\phi(t) + \lambda \int g(\phi(t'), \widehat{D}(\phi(t)))f(\phi(t'), \phi(t))d\phi(t')}{\rho + \lambda}$, where

$$g(\phi, D) \equiv \begin{cases} \tau\rho D + \lambda\tau D + (1-\tau)(1-\alpha)\lambda V(\phi), & \text{if } V(\phi) < D \\ \tau\rho D + \lambda V(\phi), & \text{if } V(\phi) \geq D \end{cases}. \quad (\text{A.43})$$

Note that $g(\phi, D)$ is non-decreasing in ϕ .³⁶ Therefore, $TV(\phi_2) \geq \frac{(1-\tau)\phi_2 + \lambda \int g(\phi(t'), \widehat{D}(\phi_1))f(\phi(t'), \phi_2)d\phi(t')}{\rho + \lambda} > \frac{(1-\tau)\phi_1 + \lambda \int g(\phi(t'), \widehat{D}(\phi_1))f(\phi(t'), \phi_2)d\phi(t')}{\rho + \lambda} \geq \frac{(1-\tau)\phi_1 + \lambda \int g(\phi(t'), \widehat{D}(\phi_1))f(\phi(t'), \phi_1)d\phi(t')}{\rho + \lambda} = TV(\phi_1)$, where the first inequality follows from the fact that $\widehat{D}(\phi_1)$ is feasible when $\phi(t) = \phi_2$; the second inequality follows from $\phi_2 > \phi_1$; and

³⁶To show that $g(\phi, D)$ is non-decreasing in ϕ , it suffices to show that $\tau\rho D + \lambda\tau D + (1-\tau)(1-\alpha)\lambda D \leq \tau\rho D + \lambda D$, or equivalently that $0 \leq (1-\tau)\lambda D - (1-\tau)(1-\alpha)\lambda D = \alpha(1-\tau)\lambda D$, which follows from $\alpha > 0$, $\tau < 1$, $\lambda > 0$, and $D \geq 0$.

the third inequality follows from the facts that $g(\phi, D)$ is non-decreasing in ϕ and that $F(\phi(t'), \phi_2)$ first-order stochastically dominates $F(\phi(t'), \phi_1)$. Therefore T maps non-decreasing functions into strictly increasing functions.

Monotonicity: Suppose that $V_2(\phi(t)) \geq V_1(\phi(t))$. Define $g_i(\phi, D)$ to be the function $g(\phi, D)$ defined in equation (A.43) where $V(\phi)$ is replaced by $V_i(\phi)$, $i = 1, 2$. Therefore, $TV_2(\phi(t)) \geq \frac{(1-\tau)\phi(t) + \int g_2(\phi(t'), \widehat{D}_1(\phi(t))) f(\phi(t'), \phi(t)) d\phi(t')}{\rho + \lambda}$ and $TV_1(\phi(t)) = \frac{(1-\tau)\phi(t) + \int g_1(\phi(t'), \widehat{D}_1(\phi(t))) f(\phi(t'), \phi(t)) d\phi(t')}{\rho + \lambda}$, where $\widehat{D}_1(\phi(t))$ attains the maximum of $\frac{(1-\tau)\phi(t) + \int g_1(\phi(t'), D) f(\phi(t'), \phi(t)) d\phi(t')}{\rho + \lambda}$, subject to $D \leq V_1(\phi(t))$ and $\widehat{D}_1(\phi(t))$ is feasible under $V_2(\phi(t))$ because $\widehat{D}_1(\phi(t)) \leq V_1(\phi(t)) \leq V_2(\phi(t))$. Therefore, $TV_2(\phi(t)) - TV_1(\phi(t)) \geq \frac{\int [g_2(\phi(t'), \widehat{D}_1(\phi(t))) - g_1(\phi(t'), \widehat{D}_1(\phi(t)))] f(\phi(t'), \phi(t)) d\phi(t')}{\rho + \lambda}$. To prove that $TV_2(\phi(t)) - TV_1(\phi(t)) \geq 0$, it suffices to prove that $g_2(\phi(t'), \widehat{D}_1(\phi(t))) - g_1(\phi(t'), \widehat{D}_1(\phi(t))) \geq 0$ for all $\phi(t)$ in the support of F . The definition of $g(\phi, D)$ implies

$$\begin{aligned} g_2(\phi, D) &= (1-\tau)(1-\alpha)\lambda[V_2(\phi) - V_1(\phi)] \geq 0, & \text{if } V_2(\phi) < D \\ -g_1(\phi, D) &= \lambda V_2(\phi) - [\lambda\tau D + (1-\tau)(1-\alpha)\lambda V_1(\phi)], & \text{if } V_1(\phi) < D \leq V_2(\phi) \\ &= \lambda[V_2(\phi) - V_1(\phi)] \geq 0, & \text{if } V_1(\phi) \geq D \end{aligned}$$

It remains to prove that $g_2(\phi, D) - g_1(\phi, D) \geq 0$ when $V_1(\phi) < D \leq V_2(\phi)$. Observe that when $V_1(\phi) < D \leq V_2(\phi)$, $g_2(\phi, D) - g_1(\phi, D) = \lambda V_2(\phi) - [\lambda\tau D + (1-\tau)(1-\alpha)\lambda V_1(\phi)] > \lambda V_2(\phi) - [\lambda\tau D + (1-\tau)(1-\alpha)\lambda D] = \lambda(V_2(\phi) - [1 - (1-\tau)\alpha]D) \geq 0$. Therefore, the operator T is monotonic.

Discounting: Define $g^{(a)}(\phi, D)$ to be the function $g(\phi, D)$ defined in equation (A.43) where $V(\phi)$ is replaced by $V(\phi) + a$, so that

$$g^{(0)}(\phi, \widehat{D}^{(a)}(\phi(t)) - a) \equiv \begin{cases} \begin{bmatrix} \tau\rho(\widehat{D}^{(a)}(\phi(t)) - a) \\ +\lambda\tau(\widehat{D}^{(a)}(\phi(t)) - a) \\ +(1-\tau)(1-\alpha)\lambda V(\phi) \end{bmatrix}, & \text{if } V(\phi) < \widehat{D}^{(a)}(\phi(t)) - a \\ \begin{bmatrix} \tau\rho(\widehat{D}^{(a)}(\phi(t)) - a) \\ +\lambda V(\phi) \end{bmatrix}, & \text{if } V(\phi) \geq \widehat{D}^{(a)}(\phi(t)) - a \end{cases}$$

and

$$g^{(a)}\left(\phi, \widehat{D}^{(a)}(\phi(t))\right) \equiv \begin{cases} \left[\begin{array}{l} \tau\rho\widehat{D}^{(a)}(\phi(t)) + \lambda\tau\widehat{D}^{(a)}(\phi(t)) \\ + (1-\tau)(1-\alpha)\lambda(V(\phi) + a) \end{array} \right], & \text{if } V(\phi) + a < \widehat{D}^{(a)}(\phi(t)) \\ \tau\rho\widehat{D}^{(a)}(\phi(t)) + \lambda(V(\phi) + a), & \text{if } V(\phi) + a \geq \widehat{D}^{(a)}(\phi(t)) \end{cases}.$$

For $a \geq 0$, $[T(V+a)](\phi(t)) = \frac{(1-\tau)\phi(t) + \int g^{(a)}(\phi(t'), \widehat{D}^{(a)}(\phi(t)))f(\phi(t'), \phi(t))d\phi(t')}{\rho+\lambda}$. Since $\widehat{D}^{(a)}(\phi(t)) \leq V(\phi(t)) + a$, $\widehat{D}^{(a)}(\phi(t)) - a \leq V(\phi(t))$ and is feasible under $V(\phi(t))$. Therefore, $TV(\phi(t)) \geq \frac{(1-\tau)\phi(t) + \int g^{(0)}(\phi(t'), \widehat{D}^{(a)}(\phi(t)) - a)f(\phi(t'), \phi(t))d\phi(t')}{\rho+\lambda}$. Therefore, $[T(v+a)](\phi(t)) - Tv(\phi(t)) \leq \frac{\int [g^{(a)}(\phi(t'), \widehat{D}^{(a)}(\phi(t))) - g^{(0)}(\phi(t'), \widehat{D}^{(a)}(\phi(t)) - a)]f(\phi(t'), \phi(t))d\phi(t')}{\rho+\lambda}$, where

$$\begin{aligned} g^{(a)}\left(\phi, \widehat{D}^{(a)}(\phi(t))\right) &= \tau\rho a + [1 - \alpha(1 - \tau)]\lambda a, & \text{if } V(\phi) + a < \widehat{D}^{(a)}(\phi(t)) \\ -g^{(0)}\left(\phi, \widehat{D}^{(a)}(\phi(t)) - a\right) &= \tau\rho a + \lambda a, & \text{if } V(\phi) + a \geq \widehat{D}^{(a)}(\phi(t)) \end{aligned}.$$

Therefore, $[T(V+a)](\phi(t)) - TV(\phi(t)) \leq \beta a$, where $\beta \equiv \frac{\tau\rho+\lambda}{\rho+\lambda} < 1$. Therefore, T satisfies the discounting property and hence T is a contraction. Therefore, $V(\phi)$ is strictly increasing. ■

Proof. of Proposition 21. The uniform distribution in this proposition is $F(\phi(t'), \phi(t)) = \frac{\phi(t') - g(\phi(t)) + d}{2d}$. First consider the case in which the tradeoff theory is operative in a neighborhood of $\phi(t)$, which implies that $\widehat{D}(\phi(t))$ satisfies the first-order condition in equation (40). This first-order condition can be written as $\tau\left[\rho + \lambda\frac{V^{-1}(D) - g(\phi(t)) + d}{2d}\right] - (1-\tau)\alpha\lambda V^{-1}(D)D\frac{1}{2d} = 0$, the second-order condition that must hold for the tradeoff theory to be operative is that the left hand side of this first-order condition is decreasing in D . An increase in $\phi(t)$ reduces the first term on the left hand side, which is the marginal interest tax shield. Therefore, the value of D must fall to satisfy the first-order condition. Hence, $\widehat{D}'(\phi(t)) < 0$ (Statement 1a). Since $\widehat{D}'(\phi(t)) < 0$ and from Proposition 20 $V(\phi(t))$ is strictly increasing in $\phi(t)$, the leverage ratio, $L(\phi(t)) \equiv \frac{\widehat{D}(\phi(t))}{V(\phi(t))}$ is strictly decreasing in $\phi(t)$ (Statement 1b). With $F(\phi(t'), \phi(t)) = \frac{\phi(t') - g(\phi(t)) + d}{2d}$, the conditional probability of default in equation (41) is $p(\phi(t)) = \frac{V^{-1}(\widehat{D}(\phi(t))) - g(\phi(t)) + d}{2d}$ regardless of whether the tradeoff theory is operative or the borrowing constraint is binding. If the tradeoff theory is operative, $\widehat{D}'(\phi(t)) < 0$ so an increase in $\phi(t)$ reduces $V^{-1}(\widehat{D}(\phi(t)))$. An increase in $\phi(t)$ increases $g(\phi(t))$. Therefore,

because $V^{-1}(\widehat{D}(\phi(t)))$ decreases and $g(\phi(t))$ increases, the probability of default, $p(\phi(t))$ falls. (Statement 1c).

Now consider the case in which the borrowing constraint is binding in a neighborhood of $\phi(t)$. Therefore, $\widehat{D}(\phi(t)) \equiv V(\phi(t))$, and hence Proposition 20 implies that $\widehat{D}'(\phi(t)) > 0$ (Statement 2a). Since $\widehat{D}(\phi(t)) \equiv V(\phi(t))$, we have $V^{-1}(\widehat{D}(\phi(t))) = \phi(t)$, so the expression for $p(\phi(t))$ can be written as $p(\phi(t)) = \frac{\phi(t) - g(\phi(t)) + d}{2d}$. Therefore, $p'(\phi(t)) = \frac{1 - g'(\phi(t))}{2d} > 0$, where the inequality follows from the assumption $0 < g'(\phi(t)) < 1$ (Statement 2b). ■

B Calculation of the Value of the Firm

This Appendix derives the expression for the value of the firm in equation (9). First, substitute equation (6) into equation (8) to obtain

$$V(\phi(t)) = \max_{D_s \leq V(\phi(s))} E_t \left\{ \int_t^{t_1} (1 - \tau) y(s) e^{-\rho(s-t)} ds + \int_t^{t_1} e^{-\rho(s-t)} dD_s \right. \\ \left. + e^{-\rho(t_1-t)} \max[V(\phi(t_1)) - D_{t_1^-}, 0] \right\}. \quad (\text{B.1})$$

Set $y(s) = y(t)$ for all s in $[t, t_1)$, $dD_t = D_t$, $dD_s = 0$ for all s in (t, t_1) and $D_{t_1^-} = D_t$ to obtain

$$V(\phi(t)) = \max_{D_t \leq V(\phi(t))} E_t \left\{ (1 - \tau) y(t) \int_t^{t_1} e^{-\rho(s-t)} ds + D_t + e^{-\rho(t_1-t)} \max[V(\phi(t_1)) - D_t, 0] \right\}. \quad (\text{B.2})$$

The density of t_1 is $\lambda e^{-\lambda(t_1-t)}$ so

$$E_t \{ e^{-\rho(t_1-t)} ds \} = \frac{\lambda}{\rho + \lambda} \quad (\text{B.3})$$

and

$$E_t \left\{ \int_t^{t_1} e^{-\rho(s-t)} ds \right\} = \frac{1}{\rho + \lambda}. \quad (\text{B.4})$$

Use the expectations in equations (B.3) and (B.4) to calculate the expectation in equation

(B.2) and rearrange to obtain

$$V(\phi(t)) = \frac{1}{\rho + \lambda} \max_{D_t \leq V(\phi(t))} E_t \left\{ (1 - \tau) y_t + (\rho + \lambda) D_t + \lambda \int_{V(\phi) \geq D_t} [V(\phi) - D_t] dF(\phi) \right\}. \quad (\text{B.5})$$

Substitute the expression for taxable income, $y(t)$, from equation (5) into equation (B.5) to obtain

$$V(\phi(t)) = \frac{1}{\rho + \lambda} \max_{D_t \leq V(\phi(t))} E_t \left\{ \begin{aligned} (1 - \tau) \left[\phi(t) - \rho D_t - \lambda \int_{V(\phi) < D_t} [D_t - (1 - \alpha) V(\phi)] dF(\phi) \right] \\ + (\rho + \lambda) D_t + \lambda \int_{V(\phi) \geq D_t} [V(\phi) - D_t] dF(\phi) \end{aligned} \right\}. \quad (\text{B.6})$$

Collect terms in D_t and separately collect terms in $V(\phi)$ to obtain

$$V(\phi(t)) = \frac{1}{\rho + \lambda} \max_{D_t \leq V(\phi(t))} E_t \left\{ \begin{aligned} & (1 - \tau) \phi(t) \\ & \tau \rho D_t + \lambda D_t - \lambda (1 - \tau) D_t \int_{V(\phi) < D_t} dF(\phi) - \lambda D_t \int_{V(\phi) \geq D_t} dF(\phi) \\ & + \lambda \int_{V(\phi) \geq D_t} V(\phi) dF(\phi) + \lambda (1 - \tau) (1 - \alpha) \int_{V(\phi) < D_t} V(\phi) dF(\phi) \end{aligned} \right\}. \quad (\text{B.7})$$

Use $1 - \int_{V(\phi) \geq D_t} dF(\phi) = \int_{V(\phi) < D_t} dF(\phi)$ and $\int_{V(\phi) \geq D_t} V(\phi) dF(\phi) = \int V(\phi) dF(\phi) - \int_{V(\phi) < D_t} V(\phi) dF(\phi)$ and rearrange to rewrite equation (B.7) as

$$V(\phi(t)) = \frac{1}{\rho + \lambda} \max_{D_t \leq V(\phi(t))} E_t \left\{ \begin{aligned} & (1 - \tau) \phi(t) \\ & \tau \left[\rho + \lambda \int_{V(\phi) < D_t} dF(\phi) \right] D_t \\ & + \lambda \int V(\phi) dF(\phi) - [1 - (1 - \tau)(1 - \alpha)] \lambda \int_{V(\phi) < D_t} V(\phi) dF(\phi) \end{aligned} \right\}. \quad (\text{B.8})$$

Use the definitions $A(D) \equiv \tau \left[\rho + \lambda \int_{V(\phi) < D} dF(\phi) \right] D - [\alpha + \tau(1 - \alpha)] \lambda \int_{V(\phi) < D} V(\phi) dF(\phi)$ and $\bar{v} \equiv \int V(\phi) dF(\phi)$ to rewrite equation (B.8) as

$$V(\phi(t)) = \max_{D_t \leq V(\phi(t))} \frac{(1 - \tau) \phi(t) + \lambda \bar{v} + A(D_t)}{\rho + \lambda}. \quad (\text{B.9})$$

Online Appendix

C Closed-Form Solutions: $\alpha = 1$ and $F(\phi)$ is Uniform on $[\Phi_L, \Phi_H]$

In this online appendix, I derive closed-form solutions for optimal debt and shareholders' equity in the special case in which (1) $\alpha = 1$, so that default completely destroys the value of the firm; and (2) the unconditional distribution $F(\phi)$ is uniform. The assumption that $\alpha = 1$ simplifies the solution of the shareholders' problem in Scenario II by permitting the calculation of ϕ^* , $\widehat{D}(\phi(t))$, and $V(\phi(t))$ for $\phi(t) \geq \phi^*$ without having to simultaneously calculate $V(\phi(t))$ for values of $\phi(t) < \phi^*$.³⁷

It will be convenient to represent the uniform distribution on $[\Phi_L, \Phi_H]$ by the mean $\mu \equiv \frac{1}{2}(\Phi_L + \Phi_H)$ and by $\delta \equiv \Phi_H - \Phi_L$. Therefore, $f(x) = \frac{1}{\delta}$, $F(x) = \frac{x - \mu + \frac{1}{2}\delta}{\delta}$, and hence $x = \mu - \frac{1}{2}\delta + \delta F(x)$, for $x \in [\Phi_L, \Phi_H]$.

Proposition 12 contains closed-form solutions for $\widehat{D}(\phi(t))$, $S(\phi(t))$, and $V(\phi(t))$ in Scenario I for the more general case in which $0 < \alpha \leq 1$ and $f'(\phi) \geq 0$, which includes the uniform distribution as a special case. This Appendix presents closed-form solutions that apply in Scenario II and Scenario III in the special case with $\alpha = 1$ and uniform $F(\phi)$.

C.1 Closed-Form Solutions for Scenario II

Consider Scenario II. Evaluate the value function in equation (14) at $\phi(t) = \phi^*$, use $V(\phi^*) = D^* = \widehat{D}(\phi^*)$, the definition of $A(D)$ in equation (10) with $\alpha = 1$, and multiply both sides by $\rho + \lambda$ to obtain

$$(\rho + \lambda) D^* = (1 - \tau) \phi^* + \tau(\rho + \lambda F^*) D^* + \lambda \int_{\phi^*}^{\Phi_H} V(\phi) dF(\phi), \quad (\text{C.1})$$

where $F^* = F(\phi^*)$. Proposition 7 along with the uniform distribution $dF(\phi) = \frac{1}{\delta}$ implies

$$\int_{\phi^*}^{\Phi_H} V(\phi) dF(\phi) = (1 - F^*) D^* + \frac{1 - \tau}{\rho + \lambda} \frac{1}{\delta} \int_{\phi^*}^{\Phi_H} (\phi - \phi^*) d\phi. \quad (\text{C.2})$$

³⁷In the more general case with $\alpha \neq 0$, $\widehat{V}(\phi(t))$ for $\phi \geq \phi^*$ depends on $\alpha \widehat{V}(\phi(t))$ for $\phi(t) < \phi^*$.

Evaluate the integral in equation (C.2) and use $\frac{\Phi_H - \phi^*}{\delta} = 1 - F^*$ to obtain

$$\int_{\phi^*}^{\Phi_H} V(\phi) dF(\phi) = (1 - F^*) D^* + \frac{1}{2} \delta \frac{1 - \tau}{\rho + \lambda} (1 - F^*)^2. \quad (\text{C.3})$$

Substitute equation (C.3) into equation (C.1) and rearrange terms to obtain

$$(\rho + \lambda F^*) D^* = \phi^* + \frac{1}{2} \delta \frac{\lambda}{\rho + \lambda} (1 - F^*)^2. \quad (\text{C.4})$$

The first-order condition in equation (25) when $\alpha = 1$ and $f(\phi) = \frac{1}{\delta}$ is

$$\tau(\rho + \lambda F^*) = (\rho + \lambda) D^* \frac{\lambda}{\delta}. \quad (\text{C.5})$$

Use equation (C.5) to eliminate D^* from equation (C.4) and rearrange to obtain

$$\tau \frac{\delta}{\lambda} (\rho + \lambda F^*)^2 = (\rho + \lambda) \phi^* + \frac{1}{2} \delta \lambda (1 - F^*)^2. \quad (\text{C.6})$$

Use the definition $\tau_L \equiv \alpha \frac{\lambda}{\rho} f(\Phi_L) \left(\Phi_L + \frac{\lambda}{\rho} E\{\phi\} \right)$ from Proposition 15, and set $\alpha = 1$, $f(\Phi_L) = \frac{1}{\delta}$, $E\{\phi\} = \mu$, and use $\phi^* = \Phi_L + \delta F^*$ to obtain

$$\tau_L = \frac{\lambda}{\rho} \frac{1}{\delta} \left(\phi^* - \delta F^* + \frac{\lambda}{\rho} \mu \right). \quad (\text{C.7})$$

Use equation (C.7) to eliminate ϕ^* from equation (C.6) and rearrange to obtain

$$\tau(\rho + \lambda F^*)^2 - (\rho + \lambda) \rho \tau_L - (\rho + \lambda) \lambda F^* + (\rho + \lambda) \frac{\lambda}{\delta} \frac{\lambda}{\rho} \mu - \frac{1}{2} \lambda^2 (1 - F^*)^2 = 0. \quad (\text{C.8})$$

Now collect terms in F^{*2} , F^* , and a constant to obtain

$$\begin{aligned} & \frac{1}{2} (2\tau - 1) \lambda^2 F^{*2} + (2\tau - 1) \rho \lambda F^* + (\tau - \tau_L) \rho^2 \\ & - \lambda \rho \left(\tau_L - \frac{\lambda}{\rho} \frac{1}{\delta} \left[\left(1 + \frac{\lambda}{\rho} \right) \mu - \frac{1}{2} \delta \right] \right) = 0. \end{aligned} \quad (\text{C.9})$$

Now use the fact that $\tau_L \equiv \alpha \frac{\lambda}{\rho} f(\Phi_L) \left(\Phi_L + \frac{\lambda}{\rho} E\{\phi\} \right) = \frac{\lambda}{\rho} \frac{1}{\delta} \left(\mu - \frac{1}{2} \delta + \frac{\lambda}{\rho} \mu \right) = \frac{\lambda}{\rho} \frac{1}{\delta} \left(\left(1 + \frac{\lambda}{\rho} \right) \mu - \frac{1}{2} \delta \right)$ to set the final term on the left hand side of equation (C.9) equal to zero and then multiply

both sides of the resulting equation by $\frac{1}{\rho^2} \frac{2}{2\tau-1}$ to obtain

$$\left(\frac{\lambda}{\rho} F^*\right)^2 + 2\frac{\lambda}{\rho} F^* + \frac{2(\tau - \tau_L)}{2\tau - 1} = 0. \quad (\text{C.10})$$

Equation (C.10) is a quadratic equation in $\frac{\lambda}{\rho} F^*$ and the positive root of this equation (recall that $\tau_L < \tau < \frac{\alpha}{1+\alpha} < \frac{1}{2}$) is

$$\frac{\lambda}{\rho} F^* = -1 + \sqrt{\frac{1 - 2\tau_L}{1 - 2\tau}}. \quad (\text{C.11})$$

Use the fact that $\phi^* = \delta F(\phi^*) + \mu - \frac{1}{2}\delta = \delta F(\phi^*) + \Phi_L$ to obtain

$$\phi^* = \mu - \frac{1}{2}\delta - \delta\frac{\rho}{\lambda} + \delta\frac{\rho}{\lambda}\sqrt{\frac{1 - 2\tau_L}{1 - 2\tau}} = \Phi_L + \delta\frac{\rho}{\lambda}\left(-1 + \sqrt{\frac{1 - 2\tau_L}{1 - 2\tau}}\right). \quad (\text{C.12})$$

To calculate the optimal level of debt, D^* , use equation (C.5) along with equation (C.11) to obtain

$$D^* = \tau\frac{\delta}{\lambda}\frac{\rho + \lambda F(\phi^*)}{\rho + \lambda} = \frac{\tau\delta}{\lambda}\frac{\rho}{\rho + \lambda}\sqrt{\frac{1 - 2\tau_L}{1 - 2\tau}}. \quad (\text{C.13})$$

Next calculate the optimal amount of debt when $\phi(t) \leq \phi^*$. It will be helpful to define Γ and use the expression in equation (C.3) to obtain

$$\Gamma \equiv \lambda \int_{\phi^*}^{\Phi_H} V(\phi) dF(\phi) = (1 - F^*) D^* + \frac{1}{2}\delta\frac{1 - \tau}{\rho + \lambda} (1 - F^*)^2. \quad (\text{C.14})$$

For $\phi(t) \leq \phi^*$, the borrowing constraint is binding so $V(\phi(t)) \equiv \widehat{D}(\phi(t))$. Therefore, set $V(\phi(t)) = \widehat{D}(\phi(t))$ and $\alpha = 1$ in the definition of $A(D)$ in equation (10) to obtain

$$A\left(\widehat{D}(\phi(t))\right) = \tau[\rho + \lambda F(\phi(t))]\widehat{D}(\phi(t)) - \lambda \int_{\Phi_L}^{\phi(t)} V(\phi) dF(\phi). \quad (\text{C.15})$$

Substitute equation (C.15) into equation (14), use $\widehat{D}(\phi(t)) \equiv V(\phi(t))$ and $F(\phi(t)) = \frac{\phi(t) - (\mu - \frac{1}{2}\delta)}{\delta}$ to obtain

$$(a - b\phi(t))\widehat{D}(\phi(t)) = (1 - \tau)\phi(t) + \frac{\lambda}{\delta} \int_{\phi(t)}^{\phi^*} \widehat{D}(x) dx + \Gamma, \quad (\text{C.16})$$

where

$$a \equiv (1 - \tau) \rho + \lambda + \frac{\tau \lambda}{\delta} \left(\mu - \frac{1}{2} \delta \right) > 0 \quad (\text{C.17})$$

and

$$b \equiv \frac{\tau \lambda}{\delta} > 0, \quad (\text{C.18})$$

so that

$$a - b\phi(t) = (1 - \tau) \rho + \lambda [1 - \tau F(\phi(t))] > 0. \quad (\text{C.19})$$

In Appendix C.3, I show that the functional equation in equation (C.16) is satisfied by

$$\widehat{D}(\phi(t)) = (a - b\phi(t))^{\frac{1-\tau}{\tau}} C + \tau \frac{1}{b}, \quad (\text{C.20})$$

where C is a constant of integration that is determined by the boundary condition in equation (32), so that

$$\widehat{D}(\phi(t)) = \frac{\delta}{\lambda} \left[1 - \left(\frac{a - b\phi^*}{\rho + \lambda} \right) \left(\frac{a - b\phi(t)}{a - b\phi^*} \right)^{\frac{1-\tau}{\tau}} \right], \text{ for } \phi(t) \leq \phi^*. \quad (\text{C.21})$$

Differentiating equation (C.21) with respect to $\phi(t)$ and using $\frac{b}{\tau} = \frac{\lambda}{\delta}$ from equation (C.18) yields

$$\widehat{D}'(\phi(t)) = \frac{1 - \tau}{\rho + \lambda} \left(\frac{a - b\phi(t)}{a - b\phi^*} \right)^{\frac{1}{\tau} - 2} \geq \frac{1 - \tau}{\rho + \lambda}, \text{ for } \phi(t) \leq \phi^*, \quad (\text{C.22})$$

where the inequality follows from $\tau < \frac{\alpha}{1+\alpha} \leq \frac{1}{2}$ and the inequality is strict for $\phi(t) < \phi^*$. Since $\frac{1-\tau}{\rho+\lambda} > 0$, $\widehat{D}'(\phi(t)) > 0$ for $\phi(t) < \phi^*$. Therefore, consistent with Proposition 7, for low values of $\phi(t)$, that is, for $\phi(t) \leq \phi^*$, the optimal amount of debt is strictly increasing in $\phi(t)$. Evaluating the derivative in equation (C.22) at $\phi = \phi^*$ yields $\widehat{D}'(\phi^*) = \frac{1-\tau}{\rho+\lambda}$, which is consistent with equation (33).³⁸

³⁸As a check, note that when $\tau = \tau_L$, so that $\phi^* = \Phi_L$ and $F(\phi^*) = 0$, equation (C.21) implies that

$$\widehat{D}(\Phi_L) = \frac{\delta}{\lambda} \left[1 - \left(\frac{a - b\Phi_L}{\rho + \lambda} \right) \right]. \quad (\text{C.23})$$

Equation (C.19) implies that $a - b\Phi_L = (1 - \tau_L) \rho + \lambda$ when $\tau = \tau_L$ so equation (C.23) can be rewritten as

$$\widehat{D}(\Phi_L) = \frac{\delta}{\lambda} \frac{\rho}{\rho + \lambda} \tau_L. \quad (\text{C.24})$$

C.2 Closed-Form Solutions for Scenario III

In Scenario III, the borrowing constraint binds for all values of EBIT so that $\widehat{D}(\phi(t)) \equiv V(\phi(t))$ for all $\phi(t) \in [\Phi_L, \Phi_H]$ and optimal debt is given by equation (C.20). However, the constant of integration, C , is determined by a different boundary condition than in Scenario II. To obtain an appropriate boundary condition, set $\alpha = 1$ and evaluate $A(D)$ in equation (10) at $D = \widehat{D}(\Phi_H) \equiv V(\Phi_H)$ to obtain

$$A\left(\widehat{D}(\Phi_H)\right) \equiv \tau(\rho + \lambda)\widehat{D}(\Phi_H) - \lambda\bar{v}. \quad (\text{C.26})$$

Substitute equation (C.26) into equation (14) and use $\widehat{D}(\phi(t)) \equiv V(\phi(t))$ to obtain

$$\widehat{D}(\Phi_H) = V(\Phi_H) = \frac{1}{\rho + \lambda}\Phi_H. \quad (\text{C.27})$$

As shown in Appendix C.3, the boundary condition in equation (C.27), along with equation (C.20), implies that³⁹

$$\widehat{D}(\phi(t)) = \frac{\delta}{\lambda} \left[1 - (1 - \tau_H) \left(\frac{a - b\phi(t)}{a - b\Phi_H} \right)^{\frac{1-\tau}{\tau}} \right]. \quad (\text{C.28})$$

C.3 Appendix: Solving the Functional Equation in (C.16)

The functional equation in equation (C.16) holds for $\phi(t) \leq \phi^*$. To solve the functional equation in equation (C.16), differentiate this equation with respect to $\phi(t)$ and use $\frac{b}{\tau} = \frac{\lambda}{\delta}$ from equation (C.18) to write the resulting differential equation in canonical form as

$$\widehat{D}'(\phi(t)) + \frac{1-\tau}{\tau} \frac{b}{a - b\phi(t)} \widehat{D}(\phi(t)) = \frac{1-\tau}{a - b\phi(t)}. \quad (\text{C.29})$$

Substitute τ_L from equation (C.7) into equation (C.24) recognizing that $\phi^* = \Phi_L$ and $F(\phi^*) = 0$ to obtain

$$\widehat{D}(\Phi_L) = \frac{1}{\rho + \lambda} \left(\frac{\lambda}{\rho} \mu + \Phi_L \right), \quad (\text{C.25})$$

which is identical to the optimal value of debt in Regime I shown in Proposition 12.

³⁹As a check, note that when $\tau = \tau_H$, so that $\phi^* = \Phi_H$ and $F(\phi^*) = 1$, equation (C.28) and the definition of τ_H in Proposition 15 with $\alpha = 1$ and $f(\Phi_H) = \frac{1}{\delta}$ imply that $\widehat{D}(\Phi_H) = \frac{\delta}{\lambda} \tau_H = \frac{1}{\rho + \lambda} \Phi_H$, which is identical to equation (C.27).

Multiply both sides of equation (C.29) by the integrating factor $(a - b\phi(t))^{-\frac{1-\tau}{\tau}}$ to obtain

$$(a - b\phi(t))^{-\frac{1-\tau}{\tau}} \widehat{D}'(\phi(t)) + \frac{1-\tau}{\tau} b (a - b\phi(t))^{-\frac{1}{\tau}} \widehat{D}(\phi(t)) = (1-\tau) (a - b\phi(t))^{-\frac{1}{\tau}}. \quad (\text{C.30})$$

Integrate both sides of equation (C.30) and then divide both sides of the resulting equation by the integrating factor $(a - b\phi(t))^{-\frac{1-\tau}{\tau}}$ to obtain

$$\widehat{D}(\phi(t)) = (a - b\phi(t))^{\frac{1-\tau}{\tau}} C + \tau \frac{1}{b}, \quad (\text{C.31})$$

where C is a constant of integration. To determine the value of C , substitute the solution for $D(\phi(t))$ from equation (C.31) into the functional equation (C.16) to obtain

$$(a - b\phi(t)) \left[(a - b\phi(t))^{\frac{1-\tau}{\tau}} C + \tau \frac{1}{b} \right] = (1-\tau) \phi(t) + \frac{\lambda}{\delta} \int_{\phi(t)}^{\phi^*} \left[(a - bx)^{\frac{1-\tau}{\tau}} C + \tau \frac{1}{b} \right] dx + \Gamma. \quad (\text{C.32})$$

Evaluating equation (C.32) at $\phi(t) = \phi^*$ and simplifying yields

$$C = \left[\phi^* + \Gamma - \tau \frac{a}{b} \right] (a - b\phi^*)^{-\frac{1}{\tau}}. \quad (\text{C.33})$$

Finally, substitute equation (C.33) into equation (C.31) to get an expression for the optimal amount of debt

$$\widehat{D}(\phi(t)) = \tau \frac{1}{b} + J \left(\frac{a - b\phi(t)}{a - b\phi^*} \right)^{\frac{1-\tau}{\tau}}, \quad (\text{C.34})$$

where

$$J \equiv \frac{(1-\tau) \phi^* + \Gamma}{a - b\phi^*} - \tau \frac{1}{b}. \quad (\text{C.35})$$

To calculate J in Scenario II, evaluate equation (C.34) at $\phi(t) = \phi^*$ and use the value-matching condition in equation (32) to obtain

$$D^* = \tau \frac{1}{b} + J, \quad \text{in Scenario II.} \quad (\text{C.36})$$

Use the expression for D^* in equation (C.13) and the definition $b \equiv \frac{\tau\lambda}{\delta}$ to obtain

$$D^* = \tau \frac{\rho + \lambda F(\phi^*)}{\rho + \lambda} \frac{\tau}{b}, \quad \text{in Scenario II.} \quad (\text{C.37})$$

Equating the right hand sides of equations (C.36) and (C.37) yields

$$J = \left[\tau \frac{\rho + \lambda F(\phi^*)}{\rho + \lambda} - 1 \right] \frac{\tau}{b} < 0, \quad \text{in Scenario II,} \quad (\text{C.38})$$

where the inequality follows from $\lambda > 0$, $0 < \tau < 1$, and $b > 0$. Use equation (C.19) to obtain

$$\tau \frac{\rho + \lambda F(\phi^*)}{\rho + \lambda} - 1 = -\frac{a - b\phi^*}{\rho + \lambda}, \quad \text{in Scenario II.} \quad (\text{C.39})$$

Substitute equation (C.39) into equation (C.38) to obtain

$$J = -\frac{a - b\phi^*}{\rho + \lambda} \frac{\tau}{b}, \quad \text{in Scenario II.} \quad (\text{C.40})$$

Substituting equation (C.40) into equation (C.34) yields

$$\widehat{D}(\phi(t)) = \frac{\tau}{b} \left[1 - \left(\frac{a - b\phi^*}{\rho + \lambda} \right) \left(\frac{a - b\phi(t)}{a - b\phi^*} \right)^{\frac{1-\tau}{\tau}} \right], \quad \text{in Scenario II.} \quad (\text{C.41})$$

which is the same as equation (C.21) since $\frac{\tau}{b} = \frac{\delta}{\lambda}$.

To calculate J in Scenario III, use $\phi^* = \Phi_H$, so that $\Gamma = 0$ and equation (C.19) implies $a - b\phi^* = (1 - \tau)(\rho + \lambda)$, to rewrite equation (C.35) as

$$J = \frac{\Phi_H}{\rho + \lambda} - \frac{\tau}{b}, \quad \text{in Scenario III.} \quad (\text{C.42})$$

Substituting equation (C.42) into equation (C.34) and using $\frac{\tau}{b} = \frac{\delta}{\lambda}$ yields

$$\widehat{D}(\phi(t)) = \frac{\delta}{\lambda} \left[1 - \left(1 - \frac{\lambda}{\rho + \lambda} \frac{1}{\delta} \Phi_H \right) \left(\frac{a - b\phi(t)}{a - b\Phi_H} \right)^{\frac{1-\tau}{\tau}} \right], \quad \text{in Scenario III.} \quad (\text{C.43})$$

Now use the definition of τ_H in Proposition 15 with $\alpha = 1$ and $f(\Phi_H) = \frac{1}{\delta}$ to obtain $\tau_H = \frac{\lambda}{\delta} \frac{\Phi_H}{\rho + \lambda}$ so

$$\widehat{D}(\phi(t)) = \frac{\delta}{\lambda} \left[1 - (1 - \tau_H) \left(\frac{a - b\phi(t)}{a - b\Phi_H} \right)^{\frac{1-\tau}{\tau}} \right], \quad \text{in Scenario III,} \quad (\text{C.44})$$

which is identical to equation (C.28).

C.4 Appendix: Verification that Equation (C.34) Satisfies Functional Equation (C.16)

Use equation (C.34) to express the right hand side of equation (C.16) as

$$(1 - \tau) \phi(t) + \frac{\lambda}{\delta} \int_{\phi(t)}^{\phi^*} \widehat{D}(x) dx + \Gamma = (1 - \tau) \phi(t) + \frac{\lambda}{\delta} \int_{\phi(t)}^{\phi^*} \left[\frac{\tau}{b} + J \left(\frac{a - bx}{a - b\phi^*} \right)^{\frac{1-\tau}{\tau}} \right] dx + \Gamma. \quad (\text{C.45})$$

Perform the integration on the right hand side of equation (C.45) to obtain

$$(1 - \tau) \phi(t) + \frac{\lambda}{\delta} \int_{\phi(t)}^{\phi^*} \widehat{D}(x) dx + \Gamma = \left[\begin{array}{l} (1 - \tau) \phi(t) + \frac{\lambda \tau}{\delta b} (\phi^* - \phi(t)) \\ + \frac{\lambda}{\delta} J \tau \left(-\frac{a - b\phi^*}{b} \right) \left(\frac{a - b\phi(t)}{a - b\phi^*} \right)^{\frac{1}{\tau}} |_{\phi(t)}^{\phi^*} + \Gamma \end{array} \right], \quad (\text{C.46})$$

which can be written as

$$(1 - \tau) \phi(t) + \frac{\lambda}{\delta} \int_{\phi(t)}^{\phi^*} \widehat{D}(x) dx + \Gamma = \left[\begin{array}{l} (1 - \tau) \phi(t) + \frac{\lambda \tau}{\delta b} (\phi^* - \phi(t)) \\ + \frac{\lambda}{\delta} J \tau \left(-\frac{a - b\phi^*}{b} \right) \left[1 - \left(\frac{a - b\phi(t)}{a - b\phi^*} \right)^{\frac{1}{\tau}} \right] + \Gamma \end{array} \right]. \quad (\text{C.47})$$

Use the definition $b \equiv \frac{\tau \lambda}{\delta}$ to obtain

$$(1 - \tau) \phi(t) + \frac{\lambda}{\delta} \int_{\phi(t)}^{\phi^*} \widehat{D}(x) dx + \Gamma = -\tau \phi(t) + \phi^* - J(a - b\phi^*) \left[1 - \left(\frac{a - b\phi(t)}{a - b\phi^*} \right)^{\frac{1}{\tau}} \right] + \Gamma. \quad (\text{C.48})$$

Use the definition $J \equiv \frac{(1-\tau)\phi^* + \Gamma}{a - b\phi^*} - \frac{\tau}{b}$ in equation (C.35), which implies $(a - b\phi^*) J = \phi^* + \Gamma - \tau \frac{a}{b}$, to obtain

$$(1 - \tau) \phi(t) + \frac{\lambda}{\delta} \int_{\phi(t)}^{\phi^*} \widehat{D}(x) dx + \Gamma = -\tau \phi(t) + \frac{a}{b} \tau + J(a - b\phi^*) \left(\frac{a - b\phi(t)}{a - b\phi^*} \right)^{\frac{1}{\tau}}. \quad (\text{C.49})$$

Multiply and divide the final term on the right hand side of equation (C.49) by $a - b\phi(t)$ and simplify to obtain

$$(1 - \tau) \phi(t) + \frac{\lambda}{\delta} \int_{\phi(t)}^{\phi^*} \widehat{D}(x) dx + \Gamma = (a - b\phi(t)) \left[\frac{\tau}{b} + J \left(\frac{a - b\phi(t)}{a - b\phi^*} \right)^{\frac{1}{\tau} - 1} \right]. \quad (\text{C.50})$$

Use the solution for the level of debt in equation (C.34) on the right hand side of equation (C.50) to obtain

$$(1 - \tau) \phi(t) + \frac{\lambda}{\delta} \int_{\phi(t)}^{\phi^*} \widehat{D}(x) dx + \Gamma = (a - b\phi(t)) \widehat{D}(\phi(t)), \quad (\text{C.51})$$

which is the same as equation (C.16). In Scenario II, the solution for the optimal level of debt in equation (C.34) becomes equation (C.21) as demonstrated in the derivation of equation (C.41) and in Scenario III, the solution for the optimal level of debt in equation (C.34) becomes equation (C.28) as demonstrated in the derivation of equation (C.44).

D Appendix: Translation of Truncated Exponential Distribution

To distinguish situations in which D^* falls in response to a rightward translation of $F(\phi)$ from situations in which D^* increases in response to a rightward translation of $F(\phi)$, consider the truncated exponential distribution $F(\phi) = \frac{\exp[\eta(\phi - \Phi_L)] - 1}{\exp[\eta(\Phi_H - \Phi_L)] - 1}$ which has density $f(\phi) = \frac{\eta \exp[\eta(\phi - \Phi_L)]}{\exp[\eta(\Phi_H - \Phi_L)] - 1}$ on the support $[\Phi_L, \Phi_H]$. Assume that $\eta \geq 0$,⁴⁰ so that $f'(\phi) = \eta f(\phi) \geq 0$, which ensures that $V(\phi)$ is concave. With this distribution function, the function $\chi(\phi)$ defined in equation (36) becomes

$$\chi(\phi) = \left(\frac{\lambda}{\exp[\eta(\Phi_H - \Phi_L)] - 1} - \rho \right) \eta, \quad (\text{D.1})$$

which is invariant to ϕ . The following corollary provides a condition on $\eta(\Phi_H - \Phi_L)$ that determines the sign of $\chi(\phi)$ and hence determines the sign of $D^{*'}(\phi)$.

⁴⁰The uniform distribution is the limiting case as η approaches zero, as may be seen by using L'Hopital's Rule to obtain $\lim_{\eta \rightarrow 0} F(\phi) = \lim_{\eta \rightarrow 0} \frac{(\phi - \Phi_L) \exp[\eta(\phi - \Phi_L)]}{(\Phi_H - \Phi_L) \exp[\eta(\Phi_H - \Phi_L)]} = \frac{\phi - \Phi_L}{\Phi_H - \Phi_L}$.

Corollary 23 *If (1) $0 < \tau < \frac{\alpha}{1+\alpha}$ and (2) the unconditional distribution of ϕ is the truncated exponential distribution $F(\phi) = \frac{\exp[\eta(\phi-\Phi_L)]-1}{\exp[\eta(\Phi_H-\Phi_L)]-1}$ for $\phi \in [\Phi_L, \Phi_H]$ where $\eta \geq 0$, then $D^{*'}(m) \lesseqgtr 0$ as $\eta(\Phi_H - \Phi_L) \lesseqgtr \ln\left(1 + \frac{\lambda}{\rho}\right)$.*

Proof. of Corollary 23. Since $\eta \geq 0$, equation (D.1) implies that $\chi(\phi) \gtrless 0$ as $\frac{\lambda}{\rho} \gtrless \exp[\eta(\Phi_H - \Phi_L)] - 1$, so that $\chi(\phi) \gtrless 0$ as $\eta(\Phi_H - \Phi_L) \lesseqgtr \ln\left(1 + \frac{\lambda}{\rho}\right)$. Therefore, Statement 2 of Proposition 19 implies that $D^{*'}(m) \lesseqgtr 0$ as $\eta(\Phi_H - \Phi_L) \lesseqgtr \ln\left(1 + \frac{\lambda}{\rho}\right)$. ■

Since $\ln\left(1 + \frac{\lambda}{\rho}\right) > 0$, Corollary 23 indicates that a rightward translation of the truncated exponential distribution will reduce D^* for small $\eta > 0$, as well as for the uniform distribution ($\eta = 0$). In order to get the opposite result, that is, to get D^* to increase in response to a rightward translation of $F(\phi)$, η must be sufficiently large, that is, the density function must have a sufficiently steep upward slope.

E Appendix: Solution to ODE in equation (31)

Consider $\phi(t) \leq \phi^*$ so $\widehat{D}(\phi(t)) \equiv V(\phi(t))$ and evaluate $A(D)$ in equation (10) at $D = \widehat{D}(\phi(t))$ for $\phi(t) \leq \phi^*$ using $V^{-1}(\widehat{D}(\phi(t))) = V^{-1}(V(\phi(t))) = \phi(t)$ to obtain

$$A(\widehat{D}(\phi(t))) = \tau[\rho + \lambda F(\phi(t))] \widehat{D}(\phi(t)) - [\alpha + \tau(1 - \alpha)] \lambda \int_{\Phi_L}^{\phi(t)} V(\phi) dF(\phi) \quad (\text{E.1})$$

Substitute equation (E.1) into equation (30) and use $\bar{v} = \int_{\Phi_L}^{\phi(t)} V(\phi) dF(\phi) + \int_{\phi(t)}^{\phi^*} V(\phi) dF(\phi) + \int_{\phi^*}^{\Phi_H} V(\phi) dF(\phi)$ to obtain

$$\begin{aligned} [(1 - \tau)\rho + \lambda(1 - \tau F(\phi(t)))] \widehat{D}(\phi(t)) &= (1 - \tau)\phi(t) \\ &+ (1 - \alpha)\lambda(1 - \tau) \int_{\Phi_L}^{\phi(t)} V(\phi) dF(\phi) \\ &+ \lambda \int_{\phi(t)}^{\phi^*} V(\phi) dF(\phi) + \Gamma, \end{aligned} \quad (\text{E.2})$$

where

$$\Gamma \equiv \lambda \int_{\phi^*}^{\Phi_H} V(\phi) dF(\phi). \quad (\text{E.3})$$

Equation (E.2) is a functional equation in $\widehat{D}(\phi(t)) \equiv V(\phi(t))$ for $\phi(t) \leq \phi^*$. A strategy for

solving this functional equation is to differentiate both sides with respect to $\phi(t)$ and use the fact that $V(\phi(t)) \equiv \widehat{D}(\phi(t))$ for $\phi(t) \leq \phi^*$ to obtain the following first-order linear ordinary differential equation (ODE)⁴¹

$$[(1-\tau)\rho + \lambda(1-\tau F(\phi(t)))] \widehat{D}'(\phi(t)) + (1-\tau)\alpha\lambda f(\phi(t)) \widehat{D}(\phi(t)) = 1-\tau. \quad (\text{E.4})$$

The ODE in equation (E.4) holds for $\phi(t) \leq \phi^*$. To solve this ODE, rewrite this equation in canonical form as

$$\widehat{D}'(\phi(t)) + \frac{\alpha(1-\tau)\lambda f(\phi(t))}{(1-\tau)\rho + \lambda(1-\tau F(\phi(t)))} \widehat{D}(\phi(t)) = \frac{1-\tau}{(1-\tau)\rho + \lambda(1-\tau F(\phi(t)))}. \quad (\text{E.5})$$

Next multiply both sides of equation (E.5) by the integrating factor $\exp\left(\int \frac{\alpha(1-\tau)\lambda f(\phi(t))}{(1-\tau)\rho + \lambda(1-\tau F(\phi(t)))} d\phi(t)\right) = [(1-\tau)\rho + \lambda(1-\tau F(\phi(t)))]^{-\frac{\alpha(1-\tau)}{\tau}}$ to obtain

$$\left[\begin{array}{l} [(1-\tau)\rho + \lambda(1-\tau F(\phi(t)))]^{-\frac{\alpha(1-\tau)}{\tau}} D'(\phi(t)) + \\ [(1-\tau)\rho + \lambda(1-\tau F(\phi(t)))]^{-\frac{\alpha(1-\tau)}{\tau}-1} \\ \times \alpha(1-\tau)\lambda f(\phi(t)) D(\phi(t)) \end{array} \right] = (1-\tau) [(1-\tau)\rho + \lambda(1-\tau F(\phi(t)))]^{-\frac{\alpha(1-\tau)}{\tau}-1}. \quad (\text{E.6})$$

Integrating both sides of equation (E.6) yields

$$\left(\begin{array}{l} [(1-\tau)\rho + \lambda(1-\tau F(\phi(t)))]^{-\frac{\alpha(1-\tau)}{\tau}} \\ \times D(\phi(t)) \end{array} \right) = \left[\begin{array}{l} C + \\ (1-\tau) \int [(1-\tau)\rho + \lambda(1-\tau F(\phi(t)))]^{-\frac{\alpha(1-\tau)}{\tau}-1} d\phi(t) \end{array} \right], \quad (\text{E.7})$$

where C is a constant of integration determined by the boundary condition $\widehat{D}(\phi^*) = D^*$ in equation (32). Divide both sides of equation (E.7) by

$[(1-\tau)\rho + \lambda(1-\tau F(\phi(t)))]^{-\frac{\alpha(1-\tau)}{\tau}}$ to obtain

⁴¹An alternative derivation of equation (E.4) evaluates $A'(D)$ at $D = \widehat{D}(\phi(t))$ to obtain $A'(\widehat{D}(\phi(t))) = \tau[\rho + \lambda F(\phi(t))] - (1-\tau)\alpha\lambda V^{-1'}(\widehat{D}(\phi(t))) \widehat{D}(\phi(t)) f(\phi(t))$, so that $A'(\widehat{D}(\phi(t))) \widehat{D}'(\phi(t)) = \tau[\rho + \lambda F(\phi(t))] \widehat{D}'(\phi(t)) - (1-\tau)\alpha\lambda \widehat{D}(\phi(t)) f(\phi(t))$ because $V^{-1}(\widehat{D}(\phi(t))) = \phi(t)$, which implies $V^{-1'}(\widehat{D}(\phi(t))) \widehat{D}'(\phi(t)) = 1$. Substitute this expression for $A'(\widehat{D}(\phi(t))) \widehat{D}'(\phi(t))$ into the ODE in equation (31) to obtain $(\rho + \lambda) \widehat{D}'(\phi(t)) - \tau[\rho + \lambda F(\phi(t))] \widehat{D}'(\phi(t)) + (1-\tau)\alpha\lambda \widehat{D}(\phi(t)) f(\phi(t)) = 1-\tau$, which is equivalent to equation (E.4).

$$\begin{aligned}
D(\phi(t)) &= [(1-\tau)\rho + \lambda(1-\tau F(\phi(t)))]^{\frac{\alpha(1-\tau)}{\tau}} \\
&\times \left[C + (1-\tau) \int [(1-\tau)\rho + \lambda(1-\tau F(\phi(t)))]^{-\frac{\alpha(1-\tau)}{\tau}-1} d\phi(t) \right].
\end{aligned} \tag{E.8}$$