Macroeconomic Conditions and the Puzzles of Credit Spreads and Capital Structure

Hui Chen*

University of Chicago GSB

January 22, 2007

Abstract

This paper addresses two puzzles about corporate debt: the "credit spread puzzle" – why yield spreads between corporate bonds and treasuries are high and volatile – and the "under-leverage puzzle" – why firms use debt conservatively despite seemingly large tax benefits and low costs of financial distress. I propose a unified explanation for both puzzles: investors demand high risk premia for holding defaultable claims, including corporate bonds and levered firms, because (i) defaults tend to concentrate in bad times when marginal utility is high; (ii) default losses are also higher during such times. I study these comovements in a structural model, which endogenizes firms' financing and default decisions in an economy with business-cycle variation in expected growth rates and economic uncertainty. These dynamics coupled with recursive preferences generate countercyclical variation in risk prices, default probabilities, and default losses. The credit risk premia in my calibrated model are large enough to account for most of the high spreads and low leverage ratios. Relative to a standard structural model without business-cycle variation, the average spread between Baa and Aaa-rated bonds rises from 48 bp to around 100 bp, while the average optimal leverage ratio of a Baa-rated firm drops from 67% to 42%, both close to the U.S. data.

*I am grateful to the members of my dissertation committee: Monika Piazzesi (Chair), John Cochrane, Doug Diamond and Pietro Veronesi for constant support and many helpful discussions. I also thank Frederico Belo, George Constantinides, Gene Fama, Vito Gala, Raife Giovinazzo, Lars Hansen, Milt Harris, John Heaton, Steve Kaplan, Anil Kashyap, Robert Novy-Marx, Ioanid Rosu, Nick Roussanov, Morten Sorensen, Amir Sufi, and participants at the Chicago Finance Workshop and Finance Faculty lunch for comments. All errors are my own. Correspondence to hchen5@ChicagoGSB.edu.

1 Introduction

This paper addresses two puzzles about corporate debt. The first one is the "credit spread puzzle": yield spreads between corporate bonds and treasuries are high and volatile relative to the observed default probabilities and recovery rates. The second is the "underleverage puzzle": firms choose low leverage ratios despite facing seemingly large tax benefits of debt and small costs of financial distress.

To address these puzzles, I build a structural model that endogenizes firms' financing and default decisions over the business cycle. Aggregate consumption and firms' cash flows are exogenous, and their expected growth rates and volatility move over the cycle. Asset prices are determined by a representative household with recursive preferences. Firms choose their capital structure based on the trade-off between tax benefits of debt and deadweight losses of default. Examples of such deadweight losses include legal fees and losses made during asset liquidation. Ex ante, these losses are born by equity-holders, because they lower the value of bonds at issue. Due to lumpy adjustment costs, firms only change their capital structure infrequently. Corporate bond investors also suffer losses at default if they cannot recover the full amount of principal. The valuation of these default losses is key to solving the puzzles.

The main mechanism of the model is as follows. First, marginal utilities are high in recessions, which means that the default losses that occur during such times will affect investors more. Second, recessions are also times when cash flows are expected to grow slower and become more volatile. These factors, combined with higher risk prices at such times, imply lower continuation values for equity-holders, which make firms more likely to default in recessions. Third, since many firms are experiencing problems in recessions, asset liquidation can be particularly costly, which will result in higher default losses for bond and equity-holders. Taken together, the countercyclical variation in risk prices, default probabilities, and default losses raises the present value of expected default losses for bond and equity-holders, which leads to high credit spreads and low leverage ratios.

There are two types of shocks in the economy: small shocks that directly affect consumption levels, and large shocks that change the conditional moments of consumption and cash flow growth, which drive the business cycle in this model. I model large shocks with a continuous-time Markov chain, which not only helps me obtain closed form solutions for stock and bond prices (up to a system of nonlinear equations), but allows me to characterize firms' default policies analytically. Risk prices for small consumption shocks rise with the conditional volatility of consumption growth. Risk prices for large shocks will be zero with time-separable preferences, because they are uncorrelated with small consumption shocks. However, with recursive preferences, investors are concerned with news about future consumption. The arrival of a recession brings bad news of low expected growth rates, and investors will demand a high risk premium on securities that pay off poorly in such times. Risk prices for these shocks increase in the frequency, size, and persistence of the shocks, which change over the business cycle.

The calibration strategy is to match empirical moments of the exogenous fundamentals. I use data on aggregate consumption and corporate profits to calibrate consumption and systematic components of the cash flows of individual firms. The volatility of firm-specific shocks is calibrated to match the average default probabilities associated with a firm's credit rating. Next, I calibrate the preference parameters to match the moments of stocks and the riskfree rate. Finally, I estimate default losses from the data of recovery rates. Relative to a benchmark case where consumption and cash flow growth are *i.i.d.*, and default losses are constant, the average spread for a 10-year Baa-rated coupon bond rises from 57 bp to around 140 bp, while the spread between Baa and Aaa-rated bonds rises from 48 bp to around 100 bp. The average optimal leverage ratio of a Baa-rated firm drops from 67% to around 42%. These values are close to the U.S. data. There is also large variation in default probabilities and credit spreads. The volatility of the Baa-Aaa spread is about 35 bp, again close to the U.S. data.

Endogenizing firms' capital structure and default decisions has two advantages. First, the model is able to predict how default probabilities will depend on the business cycle while taking into account the endogenous adjustments in firms' capital structures. With infrequent adjustments in the capital structure, the model predicts that changes in the economic conditions can lead to large variation in the conditional default probabilities. Second, while default losses for bond-holders can be calculated from the observable recovery rates, default losses for equity-holders (deadweight losses) are not observable. However, there is a link between recovery rates and deadweight losses: recovery rates are determined by firm values at default net of deadweight losses. Since this model determines firm values at default endogenously, it provides a precise link between default losses for equity-holders and recovery rates.

Through this link, I estimate default losses as a function of the state of the economy using the simulated method of moments. The procedure matches the mean and volatility of recovery rates, as well as the correlations of recovery rates with macro variables, and it identifies countercyclical variation in default losses. The intuition is as follows. Although asset values are lower in recessions, they do not drop as much as do recovery rates. Moreover, firms tend to default at higher cash flow levels in recessions, which partially offsets the variation in asset values. Thus, default losses must be higher in recessions in order for the model to fit the recovery rates.

Figure 1 and 2 provide evidence on the business-cycle movements of default rates, credit spreads, and recovery rates. Panel A of Figure 1 plots the historical annual default

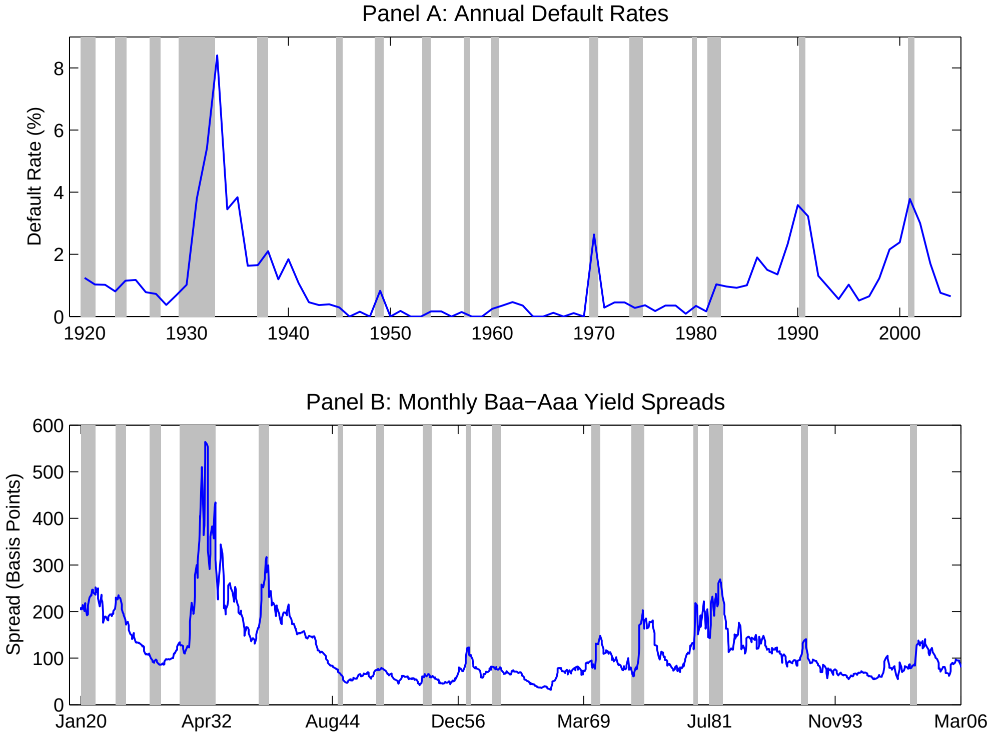

Figure 1: Annual Global Corporate Default Rates and Monthly Baa-Aaa Credit Spreads, 1920-2005. Shaded areas are NBER-dated recessions. For annual data, any calendar year with at least 5 months being in a recession as defined by NBER is treated as a recession year. Data source: Moody's.

rates from 1920 to 2005. There are several spikes in the default rates, all coinciding with an NBER recession. Panel B of Figure 1 plots the monthly Baa-Aaa spreads from 1920 to 2006. Credit spreads shoot up in almost every recession, including the ones during which default rates changed little.[1] These patterns suggest that understanding the high credit spreads in recessions is key to solving the credit spread puzzle. Business-cycle movements of the recovery rates are evident in Figure 2. Recovery rates during the three recessions in the sample, 1982, 1990 and 2001, were all significantly lower.[2] The difference in recovery rates between senior unsecured bonds and other bonds is negligible in bad times, but becomes significant in good times, suggesting that senior unsecured bonds are more affected by the cycle.

[1]The correlation between default rates and annual averages of monthly spreads is 0.65.

[2]Moody's calculate recovery rates as the weighted average of all corporate bond defaults, using closing bid prices on defaulted bonds observed roughly 30 days after the default date. For robustness, I also plot the value-weighted recovery rates from Altman and Pasternack (2006), who use the Altman Defaulted Bonds Data Set and measure recovery rates using closing bid prices as close to default date as possible. The results from these two methodologies are similar.

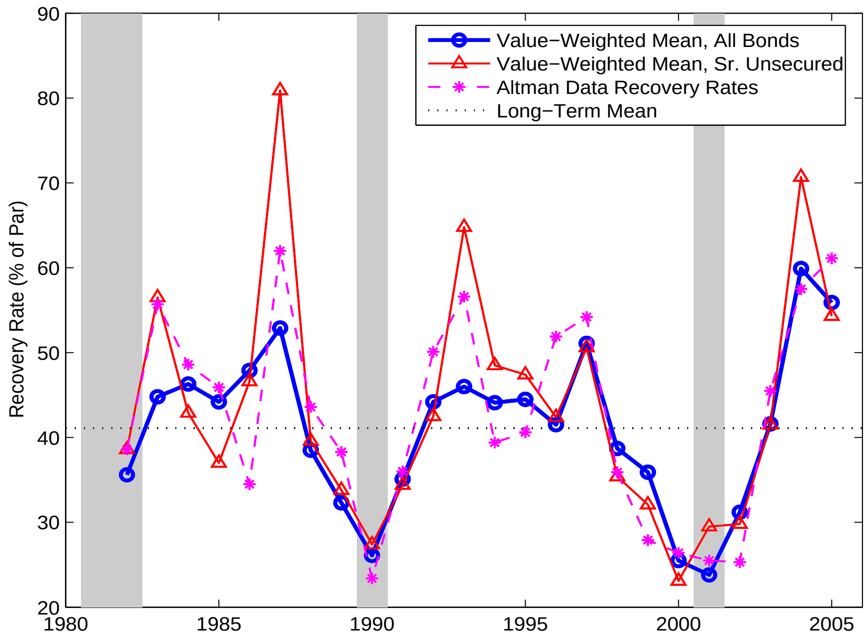

Figure 2: Annual Average Recovery Rates, 1982-2005. Value-weighted mean recovery rates for "All Bonds" and "Sr. Unsecured" are from Moody's. "Altman Data Recovery Rates" are from Altman and Pasternack (2006). Shaded areas are NBER-dated recessions.

Besides the business cycle, I also investigate the impact of risky tax benefits and costly equity issuance on the capital structure. Tax benefits are risky because firms lose part of their tax shield when they generate low cash flows for extended periods, which is more likely in bad times. Costs of (seasoned) equity issuance make leverage less attractive because they make it more costly for firms to issue equity to meet debt payments. I find considerable impact of the risky tax benefits on the capital structure, while the impact of equity issuance costs appears to be small.

The model has several additional implications. First, it predicts that firms are more likely to raise their debt levels in good times. Default probability will not rise as much following new debt issuance during such times, which reduces the effect of claim dilution on credit spreads. Second, I model default based on the dynamics of cash flows. With expected growth rates and risk premia changing over time, cash flows and market value of assets no longer have a one-to-one relation as in the earlier studies. As a result, both cash flows and market value of assets should be informative about default probabilities. For example, the model predicts that the optimal default boundaries based on cash flows are countercyclical. However, since the procyclical variation in price-dividend ratios still dominates, the resulting default boundaries based on asset value are procyclical.

Third, the model provides an explanation for default waves. The large shocks cause major changes in macroeconomic conditions, which can lead many firms to default simultaneously when the economy enters into a recession. Similarly, when the economy enters into an expansion, the model generates clustering of debt issuance, with many firms levering up simultaneously.

Related Literature

The credit spread puzzle refers to the finding of Huang and Huang (2003). They calibrated various structural models to match leverage ratios, default probabilities, and recovery rates, and found these models produce credit spreads well below historical averages. Miller (1977) highlights the challenge of the under-leverage puzzle: in expectation, default losses for firms seem disproportionately small compared to tax benefits of debt. For example, Graham (2000) estimates the capitalized tax benefits of debt to be as high as 5% of firm value, much larger than conventional estimates for the present values of default losses.

This paper builds on Chen, Collin-Dufresne, and Goldstein (2006) (CCDG), who find that strongly cyclical risk prices and default probabilities lead to high credit spreads. They focus on the credit spread puzzle, and treat firms' financing and default decisions as exogenous. This paper investigates how these decisions respond to the changes in macroeconomic conditions. It shows that these decisions have important implications for corporate bond pricing.

The connections between credit spreads and capital structure are also exploited by Almeida and Philippon (2006). They use a reduced-form approach, extracting risk-adjusted default probabilities from observed credit spreads to calculate expected default losses, and find the present value of expected default losses becomes much larger than traditional estimates. This paper goes further. It not only identifies the risks behind defaultable claims, but formally assesses the ability of a trade-off model to generate reasonable leverage ratios. Moreover, it identifies countercyclical default losses as a crucial ingredient for solving the under-leverage puzzle.

Countercyclical variation in default losses is consistent with Shleifer and Vishny (1992): liquidation of assets is more costly in bad times because the industry peers of the defaulted firm and other firms in the economy are likely experiencing similar problems. Acharya, Bharath, and Srinivasan (2006) find evidence that recovery rates are significantly lower when the industry of defaulted firm is in distress, and the relation is stronger for industries with non-redeployable assets. Altman, Brady, Resti, and Sironi (2005) also provide evidence that recovery rates are lower in recessions.

Lumpy capital structure adjustment is consistent with firms' financing behavior in reality. Welch (2004) documents that firms do not adjust their debt levels in response to changes in the market value of equity. Leary and Roberts (2005) find empirical evidence that such behaviors are likely due to adjustment costs. Strebulaev (2006) shows through simulation that a trade-off model with lumpy adjustment costs can replicate such effects. There is also evidence that such adjustment costs are asymmetric. For example, Gilson (1997) find that transaction costs for reducing debt are very high outside of Chapter 11.

The model's prediction of how default depends on market conditions echoes the findings of Pástor and Veronesi (2005) on IPO timing: just as new firms are more likely to exercise their options to go public in good times, existing firms are more likely to exercise their options to default (quit) in bad times. The model's prediction that both cash flows and market value of assets help predict default probabilities is consistent with the empirical finding of Davydenko (2005).

Theoretically, this model extends the literature on capital structure models, which include Leland (1994, 98), Leland and Toft (1996), Goldstein, Ju, and Leland (2001), Ju, Parrino, Poteshman, and Weisbach (2005), Hackbarth, Miao, and Morellec (2006), and earlier work of Brennan and Schwartz (1978), Kane, Marcus, and McDonald (1985), and Fischer, Heinkel, and Zechner (1989), etc. Business-cycle conditions have received limited attention among these models.[3] These models view default as an American option for equity-holders. Adding business cycles into these models increases the number of state variables, which brings the "curse of dimensionality". This paper provides a general solution to this problem by applying the option pricing technique for markov modulated processes (developed by Jobert and Rogers 2006). I approximate the dynamics of macroeconomic variables with a Markov chain, which helps reducing a high-dimensional free-boundary problem into a tractable system of ordinary differential equations.

This paper also contributes to the field of long-run risk models, led by Bansal and Yaron (2004), Bansal, Dittmar, and Lundblad (2005), Hansen, Heaton, and Li (2005), and others. Long-run risk models use predictable components in consumption growth to amplify the risk premia for financial claims, which helps generate high credit spreads and low leverage ratios in this models.[4] To get equilibrium pricing results, prior studies have mostly relied on the approximation method of Campbell (1993) or Hansen, Heaton, and Li (2005). Both approximations are exact when the intertemporal elasticity of substitution (IES) is equal to 1. Duffie, Schroder, and Skiadas (1997) derive close-form solutions for bond prices in continuous time in the case when the IES equals 1. This paper uses the Brownian motion–Markov chain setup to find closed form solutions for the prices of stocks,

[3]Hackbarth, Miao, and Morellec (2006) is a notable exception; they assume agents are risk-neutral, and study the effects of changes in the cash flow levels over the business cycle.

[4]An alternative way to generate big variation in risk premia is to use the habit formation model of Campbell and Cochrane (1999). Since the surplus-consumption ratio is a state variable that is driven by small consumption shocks, one cannot separately model the dynamics of this state variable with a Markov chain, which is key to tractability in this model.

bonds and other derivatives, which are exact even when the IES is not equal to 1. The Markov chain is flexible and can approximate rich dynamics of the economy. This method is useful more generally in models with recursive preferences.

The remainder of the paper is organized as follows. Section 2 presents a simple example to illustrate the main intuition. Section 3 specifies the model environment and firms' problems. Section 4 solves the static financing problem. Section 5 describes the calibration and the results. Section 6 solves the dynamic financing problem. Section 7 concludes.

2 Simple Two-Period Example

In this section, I present a simple two-period example to illustrate how the comovements among risk prices, default probabilities, and default losses lead to higher present value of expected default losses. Suppose the economy can either be in a good state (G) or bad state (B) at $t = 1$ with equal probability, as illustrated in Figure 3. The prices of one-period Arrow-Debreu securities that pay \$1 in one of the two states are Q_G and Q_B. Since marginal utility is high in the bad state, agents will pay more for consumption in that state: $Q_B > Q_G$.

There is a firm which issues one-period defaultable bonds with face value \$1 at $t = 0$. The probabilities of default in the two states, p_G and p_B, are different. Conditional on default, the losses in the two states are L_G and L_B.

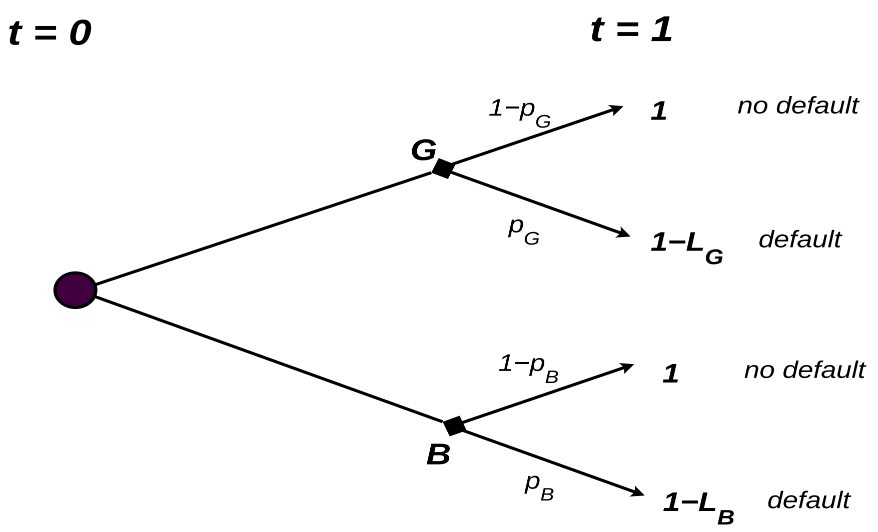

Figure 3: The payoff diagram of a defaultable zero-coupon bond in a two-period example.

The price of the zero-coupon bond at $t = 0$ is:

$$B = Q_G \left[(1 - p_G) \cdot 1 + p_G \cdot (1 - L_G)\right] + Q_B \left[(1 - p_B) \cdot 1 + p_B \cdot (1 - L_B)\right],$$

which can be rewritten as:

$$B = Q_G + Q_B - (Q_G p_G L_G + Q_B p_B L_B)\,.$$

This equation says that the price of a defaultable bond is equal to the price of a default-free bond minus the present value of expected losses at default.

In the benchmark case, the default probabilities and default losses are assumed to be the same across the two states, and are equal to their unconditional means: $\bar{p} = (p_G + p_B)/2$ and $\bar{L} = (L_G + L_B)/2$. Now, suppose that the average default probabilities and default losses are unchanged, but: (i) the bond is more likely to default in the bad state, $p_B > p_G$; (ii) the losses are higher in the bad state, $L_B > L_G$. Such "mean-preserving spreads" shift the credit losses to the state with a higher Arrow-Debreu price, which raise the present value of expected credit losses. As a result, the bond price at $t = 0$ is lower relative to the benchmark case. Moreover, the bigger the difference between the Arrow-Debreu prices Q_G and Q_B, the larger the above effects will be. The same logic applies when we calculate the present value of default losses for equity.

This simple example treats the Arrow-Debreu prices, default probabilities, and default losses as exogenous. In principle, firms could adjust their capital structure over the business cycle and avoid default in bad states. By endogenizing firms' financing decisions, this model takes such responses into account, and provides the link between default probabilities and business cycle variables. Moreover, the model derives the Arrow-Debreu prices from the representative household's marginal utilities, and estimates default losses from the data of recovery rates. I will check whether the comovements among these quantities are sufficient to solve the puzzles of credit spreads and leverage ratios.

3 The Economy

I study an economy with government, firms, and households. The government serves as a tax authority, levying taxes on corporate profit, dividend and interest income. Firms are financed by debt and equity, and generate infinite cash flow streams. Households are the owners and lenders of firms.

3.1 Preferences and Technology

There is a large number of identical infinitely lived households in the economy. The representative household has stochastic differential utility of Duffie and Epstein (1992b) and Duffie and Epstein (1992a), which is a continuous-time version of the recursive preferences of Kreps and Porteus (1978), Epstein and Zin (1989) and Weil (1990). I define the utility index at time t for a consumption process c as:

$$U_t = E_t\left(\int_t^\infty f\left(c_s, U_s\right) ds\right). \tag{1}$$

The function $f(c,v)$ is a normalized aggregator of consumption and continuation value in each period. It is defined as:

$$f(c,v) = \frac{\rho}{1-\frac{1}{\psi}} \frac{c^{1-\frac{1}{\psi}} - \left(\left(1-\gamma\right)v\right)^{\frac{1-1/\psi}{1-\gamma}}}{\left(\left(1-\gamma\right)v\right)^{\frac{1-1/\psi}{1-\gamma}-1}}. \tag{2}$$

where ρ is the rate of time preference, γ determines the coefficient of relative risk aversion for timeless gambles, and ψ determines the elasticity of intertemporal substitution for deterministic consumption paths.

Let J_t be the value function of the representative household at time t. Duffie and Epstein (1992b) and Duffie and Skiadas (1994) show that the stochastic discount factor in this economy is equal to:

$$m_t = e^{\int_0^t f_v(c_u, J_u) du} f_c\left(c_t, J_t\right). \tag{3}$$

There are two types of shocks in this economy: small shocks that directly affect output and nominal prices, and large but infrequent shocks that change expected growth rates and volatility. More specifically, a standard Brownian motion W_t^m provides systematic small shocks to the real economy. Large shocks come from the movements of a state variable s. I assume that s_t follows an n-state time-homogeneous Markov chain, and takes values in the set $\{1, \cdots, n\}$. The generator matrix for the Markov chain is $\mathbf{\Lambda} = [\lambda_{jk}]$ for $j, k \in \{1, \cdots, n\}$, which means that the probability of s_t changing from state j to k within time Δ is approximately $\lambda_{jk}\Delta$.

We can equivalently express this Markov chain as a sum of Poisson processes (see, e.g., Duffie 2001):

$$ds_t = \sum_{k \neq s_{t-}} \delta_k\left(s_{t-}\right) dN_t^{\left(s_{t-},k\right)}, \tag{4}$$

where

$$\delta_k(j) = k - j,$$

and $N^{(j,k)}$ $(j \neq k)$ are independent Poisson processes with intensity parameters λ_{jk}. The movements in the state variable are driven by these jumps.

Let Y_t denote the real aggregate output in the economy at time t, which evolves according to the following process:

$$\frac{dY_t}{Y_t} = \theta_m(s_t)\,dt + \sigma_m(s_t)\,dW_t^m. \tag{5}$$

The state variable s determines the conditional moments θ_m and σ_m, which represent the expected growth rate and volatility of aggregate output. Because s has n states, θ_m and σ_m can each take up to n different values.

In equilibrium, aggregate consumption equals aggregate output. We can solve for the value function J of the representative agent, and substitute J and Y into (3) to get the stochastic discount factor.

Proposition 1 *The real stochastic discount factor for this economy follows a Markov-modulated jump-diffusion:*

$$\frac{dm_t}{m_t} = -r(s_t)\,dt - \eta(s_t)\,dW_t^m + \sum_{s_t \neq s_{t-}} \left(e^{\kappa\left(s_{t-},s_t\right)} - 1\right) dM_t^{\left(s_{t-},s_t\right)}, \tag{6}$$

where r is the real riskfree rate; η is the risk price for systematic Brownian shocks W_t^m:

$$\eta(s) = \gamma\sigma_m(s); \tag{7}$$

$\kappa(j,k)$ determines the relative jump size of the discount factor when the Markov chain switches from state j to k; M_t is the vector of compensated processes,

$$dM_t^{(j,k)} = dN_t^{(j,k)} - \lambda_{jk}dt, \quad j \neq k, \tag{8}$$

where $N_t^{(j,k)}$ are the Poisson processes that move the state variable s_t as in equation (4). The expressions for r and κ are in Appendix A.

Proof. See Appendix A. ■

The stochastic discount factor is driven by the same set of shocks that drive aggregate output. Small systematic shocks affect marginal utility through today's consumption levels. The risk price for these shocks takes a familiar form (equation (7)), which says that the risk price rises with risk aversion and consumption volatility. Large shocks that change the state of the economy lead to jumps in the discount factor, even though consumption is perfectly smooth. The relative jump sizes $\kappa(j,k)$ are the risk prices for these large shocks.

With recursive preferences, investors care about the temporal distribution of risk, so that news about future consumption matters. The Markov chain that generates business-cycle variation in this economy brings such news. For example, investors will dislike news (large shocks) that lower the expected growth rates or raise the economic uncertainty, which means the stochastic discount factor will jump up when such news arrive. With a time-separable expected utility, investors would be indifferent to the temporal distribution of risk, and these large shocks would no longer be priced.

Finally, since credit spreads are based on nominal yields and taxes are collected on nominal cash flows, I specify a stochastic consumption price index to get nominal prices and quantities. The price index follows the diffusion

$$\frac{dP_t}{P_t} = \pi dt + \sigma_{P,1} dW_t^m + \sigma_{P,2} dW_t^P, \tag{9}$$

where W_t^P is another independent Brownian motion that generates additional shocks to nominal prices. For simplicity, the expected inflation rate π and volatility $(\sigma_{P,1}, \sigma_{P,2})$ are constant. Then, the nominal stochastic discount factor is:

$$n_t = \frac{m_t}{P_t}. \tag{10}$$

Applying Ito's formula to n_t, we get the nominal interest rate:

$$r^n(s_t) = r(s_t) + \pi - \sigma_{P,1}\eta(s_t) - \sigma_P^2. \tag{11}$$

3.2 Firms

The technology of firm i is a machine that produces a perpetual stream of real cash flows. The cash flow net of investments at time t is Y_t^i. Since operating expenses such as wages are not included in a firm's earnings, but are still part of aggregate output, the Y_t^i's across firms do not add up to the aggregate real output Y_t. The dynamics of Y_t^i is governed by the following process:

$$\frac{dY_t^i}{Y_t^i} = \theta^i(s_t)\, dt + \sigma_m^i(s_t)\, dW_t^m + \sigma_f^i dW_t^i, \tag{12}$$

where θ^i and σ_m^i are firm i's mean growth rate and systematic volatility, W_t^i is a standard Brownian motion independent of W_t^m, which generates idiosyncratic shocks specific to firm i. Finally, σ_f^i is firm i's idiosyncratic volatility, which is constant over time.

In principle, the expected growth rates and systematic volatility of cash flows can differ across firms. For computational reasons, however, it is important to keep number of states in the Markov chain low. I therefore assume that they are perfectly correlated with the aggregate expected growth rate and volatility:

$$\begin{aligned}\theta^i(s) &= a^i(\theta_m(s) - \overline{\theta}_m) + \overline{\theta}^i_m\,, \\ \sigma^i_m(s) &= b^i(\sigma_m(s) - \overline{\sigma}_m) + \overline{\sigma}^i_m\,,\end{aligned}$$

where $\overline{\theta}_m$ and $\overline{\sigma}_m$ are the average growth rate and volatility of aggregate output, $\overline{\theta}^i_m$ and $\overline{\sigma}^i_m$ are the average growth rate and systematic volatility of firm i. The coefficients a^i and b^i determine the sensitivity of firm-level expected growth rate and volatility are to changes in the aggregate values.

Firms issue bonds and pay taxes on a nominal basis. The nominal cash flow of firm i is denoted $X^i_t = Y^i_t P_t$. An application of the Ito's formula gives:

$$\frac{dX^i_t}{X^i_t} = \theta^i_X(s_t)\,dt + \sigma^i_{X,m}(s_t)\,dW^m_t + \sigma_{P,2}dW^P_t + \sigma^i_f dW^i_t, \tag{13}$$

where

$$\begin{aligned}\theta^i_X(s_t) &= \theta^i(s_t) + \pi + \sigma^i_m(s_t)\,\sigma_{P,1}\,, \\ \sigma^i_{X,m}(s_t) &= \sigma^i_m(s_t) + \sigma_{P,1}\,.\end{aligned}$$

Valuation of Unlevered Firms and Default-free Bonds

If a firm never takes on any leverage, its value (before taxes) is simply the expected value of future cash flows discounted with the stochastic discount factor. Equivalently, the value is the expected value of cash flows discounted with riskfree rates under the risk-neutral probability measure $\mathcal{Q}$. Technical details for the change of measure are in Appendix B.

The risk-neutral measure adjusts for risks by changing the distributions of shocks. Under $\mathcal{Q}$, the expected growth rate of firm i's nominal cash flows becomes:

$$\tilde{\theta}^i_X(s_t) = \theta^i_X(s_t) - \sigma^i_{X,m}(s_t)\,(\eta(s_t) + \sigma_{P,1}) - \sigma^2_{P,2}, \tag{14}$$

where θ^i_X is the expected growth rate under the physical measure $\mathcal{P}$. If cash flows are positively correlated with marginal utility, the adjustment lowers the expected growth rate of cash flows under $\mathcal{Q}$.

In addition, the generator matrix for the Markov chain becomes $\tilde{\mathbf{\Lambda}} = \left[\tilde{\lambda}_{jk}\right]$, where the transition intensities are adjusted by the corresponding jump sizes of the stochastic discount factor (see equation (6)):

$$\tilde{\lambda}_{jk} = e^{\kappa(j,k)}\lambda_{jk}\,, \quad j \neq k \tag{15a}$$

$$\tilde{\lambda}_{jj} = -\sum_{k\neq j}\tilde{\lambda}_{jk}\,. \tag{15b}$$

Bad news about future cash flows are particularly "painful" if they occur at the same time when the economy enters into a recession (marginal utility jumps up). The risk-neutral measure adjusts for such risks by increasing the probability that the economy will enter into a bad state, and reducing the probability that it will leave a bad state for a good one. For example, if marginal utility jumps up when the economy changes from state i to j, $\kappa(j,k) > 0$, then the jump intensity associated with this change of state will be higher under the risk-neutral measure.

Next, the value of an unlevered firm is the expected value of its future nominal cash flows discounted with the nominal interest rates. The following proposition gives the pricing formula.

Proposition 2 *Suppose firm i's cash flows evolve according to (13) and it never levers up. If its current cash flow is X^i, and the economy is in state s, then the value of the firm (before taxes) is:*

$$V^i\left(X^i, s\right) = X^i v^i\left(s\right). \tag{16}$$

Let $\mathbf{v}^i = \left[v^i\left(1\right), ..., v^i\left(n\right)\right]'$, then

$$\mathbf{v}^i = \left(\mathbf{r}^n - \tilde{\theta}^i_X - \tilde{\mathbf{\Lambda}}\right)^{-1}\mathbf{1}, \tag{17}$$

where $\mathbf{r}^n \triangleq \mathbf{diag}\left(\left[r^n\left(1\right), ..., r^n\left(n\right)\right]'\right)$, $\tilde{\theta}^i_X \triangleq \mathbf{diag}\left(\left[\tilde{\theta}^i_X\left(1\right), ..., \tilde{\theta}^i_X\left(n\right)\right]'\right)$, with $\tilde{\theta}^i_X\left(s\right)$ defined in (14), $\mathbf{1}$ is an $n \times 1$ vector of ones, and $\tilde{\mathbf{\Lambda}}$ is the generator of the Markov chain under the risk-neutral measure defined by (15a-15b).

Proof. See Appendix C. ■

The value of the firm is given by the Gordon growth formula. Without large shocks, the ratio of value to cash flows, v, is equal to $1/(r^n - \tilde{\theta})$, where $\tilde{\theta}$ is the expected growth rate of cash flows under the risk-neutral measure. Proposition 2 extends the Gordon formula to the more general case with large shocks. The new feature is that the expected growth rate is now adjusted by $\tilde{\mathbf{\Lambda}}$, the risk-neutral Markov chain generator, which accounts for possible changes of the state in the future.

Bad times come with higher risk prices, higher cash flow volatility and lower expected growth rate. According to equation (14), all these lead to a lower risk-neutral growth rate, which overcomes the lower real interest rate in such times, and yields lower ratios of value to cash flows. Moreover, since the adjustments in the transition probabilities increases the duration of bad times, they lead to even lower asset values in bad times.

A default-free consol bond is a cash flow stream with expected growth rate and volatility equal to 0. Thus, we can determine its value as a special case of Proposition 2.

Corollary 1 *In state s, the value of a default-free nominal consol bond with coupon rate C (before taxes) is:*

$$B(C, s) = Cb(s), \tag{18}$$

where

$$\mathbf{b} = [b(1), \cdots, b(n)]' = \left(\mathbf{r}^n - \tilde{\mathbf{\Lambda}}\right)^{-1} \mathbf{1}, \tag{19}$$

and $\mathbf{r}^n$, $\tilde{\mathbf{\Lambda}}$ and $\mathbf{1}$ are defined in Proposition 2.

3.3 Financing Decisions

The setup of firms' financing problems closely follows that of Goldstein, Ju, and Leland (2001). Firms make financing and default decisions. Their objective is to maximize equity-holders' value. Because interest expenses are tax deductible, firms lever up with debt to exploit the tax shield. As they take on more and more debt, the probability of financial distress rises, which raises the expected default losses. Thus, firms will lever up to a point when the marginal benefit of debt is zero.

Firms have access to two types of external financing: debt and equity, and they are initially financed entirely by equity. I assume that firms do not hold cash reserves. In each period, a levered firm first uses its cash flow net of investments to make interest payments on its debt, then pay taxes, and finally distributes the rest to equity-holders as dividend. The firm faces a "liquidity crunch" whenever its internally generated cash flows fall short of the interest expenses. To finance its debt payments, the firm can issue additional equity. If the "liquidity crunch" becomes too severe and equity-holders are no longer willing to contribute more capital, the firm defaults.

Debt is in the form of a consol bond, i.e., a perpetuity with constant coupon rate C. This is a standard assumption in the literature (see, e.g., Fischer, Heinkel, and Zechner (1989), Leland (1994), Duffie and Lando (2001), Goldstein, Ju, and Leland (2001)), which helps maintain a time-homogeneous setting for the model. One interpretation for this assumption is that firms commit to a constant financing plan, rolling over debt perpetually. All bonds have a pari passu covenant, which requires newly issued bonds have equal seniority as any old issues. This assumption helps to simplify the seniority structure of outstanding debt.

Bond and equity issues are costly. For equity, these costs are a constant fraction e of the proceeds from issuance. For debt, these costs are "quasi-fixed", i.e., they are a fraction q of the amount of debt outstanding after issuance (not the amount newly issued). The idea behind behind this assumption is that debt issuance incurs two types of costs: underwriting costs, which are proportional to the value of new issues, and costs of negotiating with the firm's existing debt-holders (to get the permission to issue additional pari passu debt), which are proportional to the value of old issues. These adjustment costs help the model match the lumpiness of debt issues in the data.[5]

Default losses are proportional to the value of a firm's unlevered assets at the time of default. This assumption is standard in the literature. These costs are likely to be higher in bad times, when the demand for both physical and intangible assets is low, making liquidation more costly. I therefore allow the fractional default losses $\alpha(s)$ to depend on the state of the economy s.

The tax environment consists of a constant tax rate τ_i for personal interest income, and τ_d for dividend income. A firm's taxable income is equal to cash flow (EBIT) minus interest expenses. Positive taxable income is taxed at rate τ_c^+, while negative taxable income is taxed at a lower rate τ_c^-. The assumption of two different corporate tax rates is a crude way to model "partial loss offset". The US tax laws allow firms to carry net operating losses backward and forward for a limited number of years, which means a firm can lose part of the tax shield when earnings are low.[6] Since cash flows are more likely to be low in bad times, so will tax benefits, which increases the riskiness of tax benefits.

I study firms' financing decisions in two settings: a static setting where firms only issue debt once at time 0 and makes no adjustment later on, and a dynamic setting where firms can make subsequent adjustments to their debt levels.

Static Financing Decisions

The static financing problem is to choose an amount of debt and a default policy that maximize the value of equity *right before issuance*, E_U, which is equal to the expected present value of the firm's cash flow stream, plus the tax benefits of debt, minus default losses and debt/equity issuance costs:

$$\max_{\{C,\mathcal{T}_D\}} E_U\left(C, \mathcal{T}_D, \chi_0\right), \tag{20}$$

where C is the coupon rate of perpetual debt issued at time 0, $\mathcal{T}_D$ is a stopping time that

[5]Technically, this assumption together with the "pari passu covenant" helps relax the requirement in Goldstein, Ju, and Leland (2001) that a firm retires all its outstanding debt before issuing new debt.

[6]A more realistic way to model "partial loss offset" will be to assume τ_c^- decreases with the net losses, since firms lose their tax shield only when they accumulate net losses for an extended period of time.

determines the default policy, and χ_0 contains all the state variables at time 0.

Dynamic Financing Decisions

The dynamic problem allows firms to issue additional debt after time 0, which I refer to as "upward restructuring". Now, in addition to the initial coupon rate and default policy, a firm also needs to decide when to increase its debt level, and by how much. Thus, the firm's problem becomes:

$$\max_{\{C, \mathcal{T}_D, \{\mathcal{T}_U\}, \{C_{\mathcal{T}_U}\}\}} E_U\left(C, \mathcal{T}_D, \{\mathcal{T}_U\}, \{C_{\mathcal{T}_U}\}, \chi_0\right), \tag{21}$$

where $\{\mathcal{T}_U\}$ is a series of stopping times that determines the firm's restructuring policy, and $\{C_{\mathcal{T}_U}\}$ are the new coupon rates at each restructuring point.

4 Static Financing Decisions

The static financing problem is solved in three steps. The first step computes debt and equity values for a fixed amount of debt outstanding and a fixed set of default boundaries. The second step determines the optimal default boundaries for a fixed amount of debt outstanding. The third step determines the optimal amount of debt by maximizing the value of equity before debt issuance.

There is no need to distinguish between firms yet, so I will temporarily drop the superscript i for cash flow X_t. For a fixed amount of debt, the default policy is an optimal stopping problem. This policy is characterized by a set of default boundaries, X_D^k for state k, $k = 1, \cdots, n$. A firm defaults if its cash flows fall below the boundary X_D^k while the economy is in state k. Although these boundaries are endogenous, I can always re-order the macroeconomic states such that:

$$X_D^1 \le X_D^2 \le \cdots \le X_D^n \le C. \tag{22}$$

The last inequality follows from the optimality of default. It is never optimal to default when the value of equity is above zero, which will be the case if cash flows at default are higher than interest payments.

The default boundaries and coupon rate divide the relevant range for cash flows into $n+1$ regions: $\mathcal{D}_k \triangleq [X_D^k, X_D^{k+1})$ for $k < n$, $\mathcal{D}_n \triangleq [X_D^n, C)$ and $\mathcal{D}_{n+1} \triangleq [C, +\infty)$. In regions $\mathcal{D}_1$ through $\mathcal{D}_{n-1}$, firms face immediate default threats. For example, suppose the economy is currently in state 1, the state with the lowest default boundary. If a firm's cash flow is in region $\mathcal{D}_{n-1}$, then it is below the default boundary in state n, but above the boundary for the current state. The firm will not default now, but if a big shock suddenly changes the state from 1 to n, thus raising the default boundary above current cash flow, it will default immediately. In region $\mathcal{D}_n$, there is no immediate danger of default, but firms face a liquidity crunch because they are short of internal cash flows to cover interest payments. Finally, $\mathcal{D}_{n+1}$ is the "normal" region (without default threats or liquidity problems).

4.1 Debt and Equity Value

Debt and equity are contingent claims based on a firm's cash flows as well as the state of the economy. They belong to a general class of perpetual securities $J(X_t, s_t)$, paying a dividend $F(X_t, s_t)$ for as long as the firm is solvent, and a default payment $H(X_{\mathcal{T}_D}, s_{\mathcal{T}_D})$ when default occurs at time $\mathcal{T}_D$. What distinguishes one security is the dividend stream and the default payment. I define $\mathbf{J}(X)$ as an n-dimensional vector of the security J's values in the n states.

For debt, the "dividend" is the after-tax coupon rate. With strict priority, the default payment is equal to the residual value of the firm at default:

$$V_B(X,s) = \left(1-\tau_c^+\right)(1-\tau_d)(1-\alpha(s))V(X,s), \tag{23}$$

which is the value of the unlevered firm V,[7] net of taxes (τ_c^+, τ_d) and default losses $\alpha(s)$. Thus, the dividend and default payment for debt are:

$$F(X,s) = (1-\tau_i)\,C\,, \tag{24a}$$
$$H(X,s) = V_B(X,s)\,. \tag{24b}$$

For equity, the dividend is positive when cash flows exceed interest expenses. If cash flows are less than interest, the firm faces a liquidity crunch. On such occasions, as long as the present value of future dividend income exceeds their debt obligations, equity-holders will contribute additional capital through costly equity issuance. The issuance costs are a fraction e of the proceeds. If the firm defaults, default payment to equity-holders is zero. So, the dividend and default payment for equity are:

$$F(X,s) = \begin{cases} (1-\tau_d)\left(1-\tau_c^+\right)(X-C) & X \geq C \\ -\left(1-\tau_c^-\right)(C-X)/(1-e) & X < C \end{cases}, \tag{25a}$$
$$H(X,s) = 0. \tag{25b}$$

[7] In principle, debt-holders should be able to takeover the residual assets and lever up optimally. I use the simplifying assumption to avoid the fix-point problem, which leads to a small downward bias on default losses when the model is calibrated to match recovery rates.

Let $D(X,s)$ and $E(X,s)$ be the value of debt and equity in state s. The following two propositions summarize the valuation of debt and equity given the default policy.

Proposition 3 *Suppose a firm has a consol bond outstanding with coupon rate C and a default policy characterized by a set of default boundaries $(X_D^1, \cdots, X_D^n)$, which satisfy the ordering of (22). Then, the value of debt is:*

$$\mathbf{D}(X;C) = \sum_{j=1}^{2k} w_{k,j}^D \mathbf{g}_{k,j} X^{\beta_{k,j}} + \xi_k^D X + \zeta_k^D, \quad X \in \mathcal{D}_k, \quad k = 1, \cdots, n+1. \tag{26}$$

The coefficients $\mathbf{g}$, β, (w^D, ξ^D, ζ^D) are given in Appendix D.

Proof. See Appendix D. ■

This proposition specifies the value of debt in each of the $n+1$ regions $\mathcal{D}_k$. In the first $n-1$ regions, the firm will already be in default for some of the states, and the value of debt corresponding to those states will be 0. In the last region $\mathcal{D}_{n+1}$, the firm is alive in all n states. Given the amount of debt outstanding, as X increases, the firm gets further away from bankruptcy. In the limit, the firm is free of default risk. Thus, the value of the corporate consol is bounded from above by that of a default-free consol:

$$\lim_{X \uparrow +\infty} \mathbf{D}(X;C) = (1-\tau_i)\, C\mathbf{b},$$

where $\mathbf{b}$ is the value of a default-free consol with unit coupon rate as given in Corollary 1. This intuition suggests that the coefficients $w_{n+1,j}^D$ associated with those exponents $\beta_{n+1,j}$ that are positive will be zero, ξ_{n+1}^D will be zero, and ζ_{n+1}^D will be equal to $(1-\tau_i)C\mathbf{b}$.

The values of all perpetual securities $J(X,s)$ described earlier can be written in the same form as debt, and they share the same coefficients $\mathbf{g}$ and β. However, the coefficients w^D, ξ^D and ζ^D are specific to debt. They are determined by the dividend rate, the default payment, and a set of conditions that ensure that the value of the claim is continuous and smooth across adjacent regions.

Proposition 4 *For a given coupon rate C, the value of equity can be decomposed into two parts: the value of future positive dividend payments, and the costs of equity contribution to cover future shortfalls in cash for debt payments.*

$$E(X,s;C) = (1-\tau_d)\left(1-\tau_c^+\right) E^+(X,s;C) - \frac{1-\tau_c^-}{1-e} E^-(X,s;C), \tag{27}$$

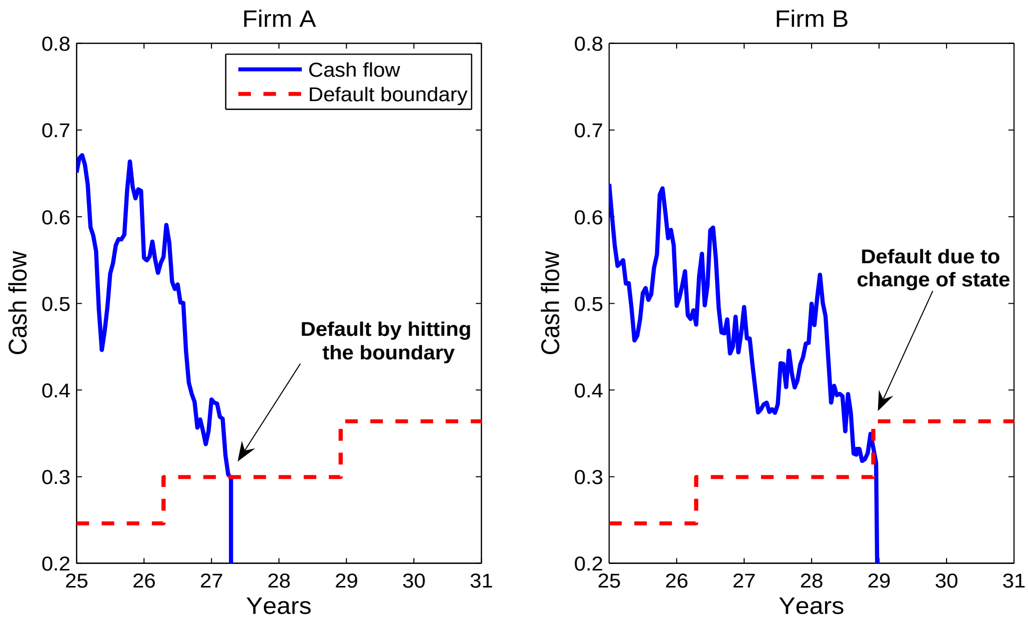

Figure 4: Illustration of Two Types of Defaults

where

$$\mathbf{E}^{+}(X;C)=\sum_{j=1}^{2k}w_{k,j}^{E^{+}}\mathbf{g}_{k,j}X^{\beta_{k,j}}+\xi_{k}^{E^{+}}X+\zeta_{k}^{E^{+}},\quad X\in\mathcal{D}_{k}\,,\quad k=1,\cdots,n+1. \tag{28}$$

and

$$\mathbf{E}^{-}(X;C)=\sum_{j=1}^{2k}w_{k,j}^{E^{-}}\mathbf{g}_{k,j}X^{\beta_{k,j}}+\xi_{k}^{E^{-}}X+\zeta_{k}^{E^{-}},\quad X\in\mathcal{D}_{k}\,,\quad k=1,\cdots,n+1. \tag{29}$$

The coefficients $\mathbf{g}$ *and* β *are given in Proposition 3, while* $(w^{E^+},\xi^{E^+},\zeta^{E^+})$ *and* $(w^{E^-},\xi^{E^-},\zeta^{E^-})$ *are given in Appendix D.*

Proof. See Appendix D. ∎

When cash flows are sufficiently large, partial loss offset becomes irrelevant, and the firm no longer needs to issue equity to finance debt payments. In the limit, the value of equity should be equal to the value of future cash flows net of the value of the default-free debt and taxes:

$$\lim_{X\uparrow+\infty}\mathbf{E}(X;C)=(1-\tau_d)\left(1-\tau_c^{+}\right)(X\mathbf{v}-C\mathbf{b})\,,$$

where $\mathbf{v}$ is the value-cash flow ratio given in Proposition 2. This intuition implies the following: in the region $\mathcal{D}_{n+1}$, all the coefficients $w_{n+1,j}^{E^+}, w_{n+1,j}^{E^-}$ associated with positive exponents $\beta_{n+1,j}$ are equal to zero, and so are $\xi_{n+1}^{E^-}$ and $\zeta_{n+1}^{E^-}$, while $\xi_{n+1}^{E^+}=\mathbf{v}$, $\zeta_{n+1}^{E^+}=-C\mathbf{b}$.

For any default policy (a set of default boundaries), we are interested in the conditional probability that a firm will default within a given amount of time. In other words, we are interested in the distribution of the stopping time $\mathcal{T}_D$, the first time that cash flow X is below one of the n default boundaries while the economy is in the corresponding state:

$$\mathcal{T}_D \triangleq \inf \left\{ u > 0 \mid X_{t+u} \leq X_D^k, s_{t+u} = k \ for \ any \ k \ between \ 1 \ and \ n \right\}.$$

In Appendix E, I provide an algorithm to evaluate the distribution of stopping time $\mathcal{T}_D$.

Default can be triggered by small shocks or large shocks. For example, the economy could remain in state i while X_t keeps decreasing until it reaches X_D^i. Alternatively, X_t could already be below X_D^i, but the economy is currently in state j with $j < i$. Then a large shock that changes the economy from state j to i will cause the firm to default immediately. Figure 4 illustrates these two types of defaults. Firm A and B have the same cash flow processes and default boundaries, but they experience different idiosyncratic shocks. Firm A defaults shortly after year 27, as a series of small shocks drive its cash flow below the default boundary. Firm B's cash flows stay above the default boundary until the end of year 29, when a big shock causes the default boundary to jump above the firm's cash flow level, which leads to default.

The second type of default is especially interesting because it suggests that those firms with cash flows between two default boundaries can default at the same time when the boundary jumps up. Hackbarth et. al. (2006) point out that this mechanism can be used to explain default waves. Their model predicts that default waves occur when aggregate cash flow levels jump down, while in this model default waves occur when expected growth rates, volatility, and risk prices change.

4.2 Optimal Default Boundaries and Capital Structure

The optimal default boundaries satisfy the smooth-pasting conditions for equity:

$$\frac{\partial}{\partial X} E\left(X, k; C\right)\bigg|_{X=X_D^k} = 0, \quad k = 1, ..., n. \tag{30}$$

Given the pricing formula for equity in Proposition 4, the n smooth-pasting conditions translate into a system of nonlinear equations (see Appendix H).

The optimal amount of debt to issue at time 0 is determined by the coupon rate that maximizes the value of equity right before issuing debt. This value is equal to the sum of equity and debt right after issuance minus debt issuance costs, which are a fraction q of debt value. Thus, the value of equity right before debt issuance is:

$$E_U(X, s; C) = E(X, s; C) + (1 - q) D(X, s; C), \tag{31}$$

and the optimal coupon rate is:

$$C^*(X, s) = \arg\max_C E_U(X, s; C). \tag{32}$$

5 The Puzzles of Credit Spreads and Leverage Ratio

I first calibrate the process for aggregate output to the consumption data. Next, I calibrate preferences so that the model can match the key moments of the asset market. Then, I calibrate the cash flow processes, default probability, and recovery rates to the data for firms with different credit ratings. Using these parameters, I calculate the optimal leverage ratios and credit spreads in the model.

While the model provides close-form solutions for the credit spreads of consols, these numbers are not directly comparable with those of finite maturity coupon bonds. A main reason is that all the cash flows of a consol are subject to personal taxes, while the principal payment of a finite maturity coupon bond is not. Thus, I also compute the credit spreads of hypothetical 10-year coupon bonds, which have exactly the same default probabilities and recovery rates as firms with the same credit ratings.

For target credit spreads, I use the estimates of Duffee (1998). In his sample, the average credit spread of a Baa-rated medium-maturity (close to 10 years) bond in the industrial sector is 148 bp, while the average Baa-Aaa spread is 101 bp. The advantage of Duffee's estimates is that they are based on corporate bonds without option-like features. His sample covers the period 1985-1995, a period when the Baa-Aaa spread is relatively low and smooth. Huang and Huang (2003) estimate credit spreads over the sample period 1973-1993. Their estimates are higher (194 bp for Baa, 131 bp for Baa-Aaa) because of the embedded call options and the inclusion of two recessions with high spreads. I calculate the volatility of Baa-Aaa spreads using the Moody's data, which is 40 bp.

5.1 Calibration

I calibrate the Markov chain that controls the conditional moments of consumption growth to be consistent with the consumption model of Bansal and Yaron (2004), which are in turn calibrated to the annual consumption data from 1929 to 1998. Appendix K provides the details of the calibration. For numerical reasons, I choose a small number of states ($n = 9$) for the Markov chain. Simulations show that the Markov chain captures

Table 1: ASSET PRICING IMPLICATIONS OF THE MARKOV CHAIN MODEL

	Data		Model	
Variable	Estimate	SE	$\gamma = 7.5$	$\gamma = 10$
$E(r_m - r_f)$	6.33	(2.15)	5.16	6.44
$E(r_f)$	0.86	(0.42)	1.42	1.18
$\sigma(r_m)$	19.42	(3.07)	14.77	14.11
$\sigma(r_f)$	0.97	(0.28)	1.00	0.93
$E(SR)$	0.33	-	0.35	0.46
$E(P/D)$	26.56	(2.53)	27.84	21.17
$\sigma(log(P/D))$	0.29	(0.04)	0.13	0.11

NOTE: The statistics of the data are from BY (2004) (Table IV). The variables r_m and r_f are returns on the market portfolio and risk-free rate; SR is the Sharpe ratio; P/D is the price-dividend ratio for the market portfolio. Two additional preference parameters are $\psi = 1.5$, and $\rho = 0.009$. All values are annualized when applicable.

the main properties of consumption reasonably well. Some of the median values from simulations (with corresponding sample estimates reported in parentheses) are: average annual growth rate 1.81% (1.80%), volatility 2.64% (2.93%), first order autocorrelation 0.42 (0.49), second order autocorrelation 0.18 (0.15).

Table 1 reports the pricing implications of the Markov chain model. With $\gamma = 7.5$, the model generates an equity premium that is slightly lower than the data, but has a better match of the Sharpe ratio and the price-dividend ratio. Changing γ to 10 raises the equity premium, but also raises the Sharpe ratio and lowers the price-dividend ratio. In both cases, the model generates a low volatility of the log price-dividend ratio, and requires a tiny subjective discount factor to keep the risk-free rate down. Moreover, the model predicts that short term interest rates are higher in good times, and that the real yield curve is downward sloping on average. This result is consistent with the findings of Piazzesi and Schneider (2006). I use $\gamma = 7.5$ as the benchmark case in this paper.

There are 5 parameters associated with a firm's cash flow process (see equation (12)). I assume that the long-run average growth rate of cash flows for all firms are the same as that of aggregate consumption. For a Baa-rated firm, I set the multipliers a_i and b_i to 3 and 4.5, and the average systematic volatility $\overline{\sigma}^i_m$ to 0.141, so that the cash flows fit the moments of the real growth rates of corporate profits for nonfinancial firms as reported by NIPA. Finally, I calibrate the idiosyncratic volatility σ^i_f to match the 10-year default probability of Baa-rated firms (4.9%). It can be the case that a typical Baa-rated firm has less volatile cash flows than an average nonfinancial firm. In that case, we will need higher idiosyncratic volatility to match the average 10-year default rates.

It is difficult to calibrate the cash flow process for an Aaa-rated firm directly to the

Table 2: PARAMETERS OF THE MODEL

Inflation, Taxes, and Issuance Costs

π	σ_P	$\rho_{P,m}$	τ_c^+	τ_c^-	τ_d	τ_i	q	e
0.036	0.014	-0.12	0.35	0.20	0.12	0.296	0.01	0.05

Cash Flow Process

	$\overline{\theta}_m^i$	$\overline{\sigma}_m^i$	σ_f^i	a_i	b_i
Baa	0.018	0.141	0.169	3.0	4.5
Aaa	0.018	0.093	0.117	2.0	3.0

NOTE: Variables are annualized, when applicable. Inflation data are from NIPA. Tax rates (except τ_c^-) are from Graham (2000). Issuance costs are from Altinkihc and Hansen (2000).

data, because there are few Aaa-rated nonfinancial firms available. Instead, I adopt a somewhat arbitrary method: scaling down a_i, b_i, $\overline{\sigma}_m^i$ and σ_f^i from their Baa values by the same proportion to match the 10-year default rate for Aaa-rated firms (0.6%). This assumption makes the cash flows of Aaa firms less volatile, but still have the same correlation with consumption as Baa firms.

Miller (1977) points out that tax benefits of debt at the corporate level is partially offset by individual tax disadvantages of interest income. Under certain simplifying assumptions, Miller gives the condition for tax benefits to be positive:

$$\tau_c > \frac{\tau_i - \tau_d}{1 - \tau_d}.$$

In the Miller equilibrium, $\tau_d = 0$, and $\tau_c = \tau_i$, so that the tax benefits zero. While this is an extreme case, it shows that the optimal leverage ratio will depend on the tax rates. To address this concern, I use the tax rate estimates of Graham (2000), which take into account the sheltering of capital income at the personal level. There is no good guidance on how large τ_c^- should be. I set it to .2, which is small enough to violate the above condition.

The inflation statistics are based on the price index for nondurables and services from NIPA. The costs of debt and equity issuance are based on the estimates of Altinkihc and Hansen (2000) on the underwriting fees for straight bond and seasoned equity offerings. Table 2 summarizes the calibrated parameters.

The Cyclicality of Recovery Rates and default losses

There are direct and indirect costs for a firm when it is in financial distress. Examples of direct costs include litigation expenses and loss due to fire sales of assets. Examples

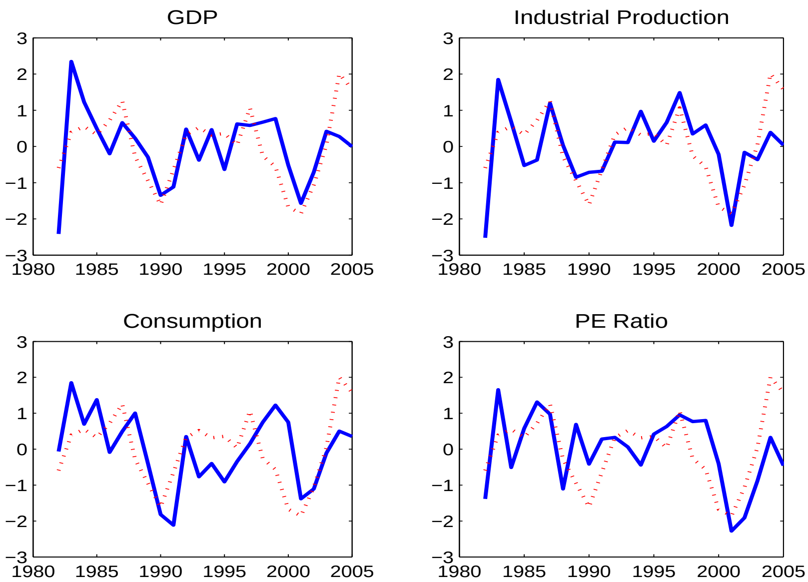

Figure 5: Recovery Rates and Macroeconomic Variables, 1982-2005. All the series are normalized to have mean 0 and standard deviation 1. The dotted line is the normalized recovery rate. GDP, IP and consumption data are from NIPA. Consumption is the sum of nondurables and services deflated with a chain-weighted price indice. Price-Earnings ratios are from Robert Shiller's web site. All macro variables are annual growth rates.

of indirect costs include loss of customers, human capital, and growth options, etc. With business-cycle variation, not only the average levels, but the distribution of default losses over different states of the economy matters.

Shleifer and Vishny (1992) argue that liquidation of assets will be particularly costly in recessions when many firms are in distress. This suggests that default losses are countercyclical. However, default losses are difficult to measure, partly because it is hard to distinguish between costs of financial and losses due to economic distress. With most defaults happening in bad times, it is even harder to measure the variation in default losses over time. A common practice in structural models is to assume that default losses are a constant fraction α of the market value of firms' assets at default. In fact, many studies set this fraction to the estimates by Andrade and Kaplan (1998), which suggest a number between $10 \sim 20\%$.

However, this approach is problematic for several reasons. The estimates of Andrade and Kaplan (1998) are relative to the *pre*-distress value of a firm, which are likely to be significantly larger than the firm value at default. Thus, default losses as a fraction of the firm value at default could well exceed 20%. Moreover, it is unclear how well these estimates represent the default losses of a typical firm. On the one hand, Leland (1998) argues that firms choosing to undergo highly leveraged buyouts might have lower default

Table 3: EXPLAINING AGGREGATE DEFAULT RATES

Dependent Variable – Default Rate (DR)							
Intercept	2.01	2.50	1.74	2.75	3.62	3.59	3.74
	(0.38)	(0.64)	(0.27)	(0.54)	(0.40)	(0.31)	(0.33)
ΔIP	-0.14					-0.09	-0.10
	(0.07)					(0.06)	(0.05)
ΔGDP		-0.28					
		(0.15)					
ΔPE			-0.03				
			(0.01)				
g				-0.55	-1.82	-1.70	-1.77
				(0.17)	(0.38)	(0.28)	(0.31)
g^2					0.33	0.34	0.37
					(0.12)	(0.09)	(0.10)
r_f							-0.24
							(0.23)
R^2	0.28	0.29	0.15	0.32	0.50	0.60	0.61
Adj R^2	0.25	0.26	0.11	0.29	0.45	0.54	0.53

NOTE: DR - Default Rate, ΔIP - real industrial production growth, ΔGDP - real GDP growth, ΔPE - growth rate of Price/Earnings ratio, g - real consumption growth, r_f - real riskfree rate. Numbers in brackets are standard errors computed with GMM based on Newey-West with lag 3. All variables are annualized, from 1982 to 2005 (24 observations). GDP, IP, consumption and CPI series are from NIPA. PE ratios are from Robert Shiller's web site. Riskfree rates are the 1-month T-bill rates from Ken French's web site. Default Rates are from Moody's.

losses than others. On the other hand, the distress periods of many firms in their sample coincide with the 1990-91 recession. If default losses are higher in bad times, then the estimates of Andrade and Kaplan might be higher than average.

I use a different approach to identify the cyclical variation in default losses. Unlike default losses, recovery rates for corporate bonds are straightforward to measure, and have a relatively long time series (Moody's average recovery rates series starts in 1982). If we know when a firm will default (the default boundary), we can compute the recovery rate by deducting default losses and taxes from the value of assets at the default boundary. Thus, using the endogenously determined default boundaries, we can identify the variation in default losses across different states of the economy from the variation of recovery rates.

I first provide more evidence for the cyclical variation in recovery rates. Figure 2 shows that recovery rates are lower during recessions. Figure 5 shows that recovery rates covary with several macroeconomic variables: GDP, industrial production, consumption, and price-earnings ratio.

Table 4: EXPLAINING AGGREGATE RECOVERY RATES

	Dependent Variable – Recovery Rate (RR)							
Intercept	52.96	37.05	32.92	39.60	33.58	28.03	31.68	31.79
	(2.69)	(2.61)	(4.09)	(2.27)	(3.47)	(2.47)	(2.90)	(2.37)
DR	-7.36							-6.95
	(1.14)							(1.83)
ΔIP		1.43						
		(0.44)						
ΔGDP			2.55					
			(0.95)					
ΔPE				0.33		0.31	0.35	0.35
				(0.07)		(0.08)	(0.09)	(0.07)
g					3.63	12.86	11.04	10.99
					(1.42)	(3.36)	(2.52)	(2.83)
g^2						-2.83	-2.29	-2.28
						(0.85)	(0.67)	(0.74)
r_f							-5.39	-5.55
							(3.54)	(2.38)
R^2	0.60	0.32	0.27	0.23	0.16	0.42	0.50	0.74
Adj R^2	0.58	0.29	0.24	0.20	0.12	0.33	0.40	0.67

NOTE: DR - Default Rate, RR - Recovery Rate, ΔIP - real industrial production growth, ΔGDP - real GDP growth, ΔPE - growth rate of Price/Earnings ratio, g - real consumption growth, r_f - real riskfree rate. Numbers in brackets are standard errors computed with GMM based on Newey-West with lag 3. All variables are annualized, from 1982 to 2005 (24 observations). GDP, IP, consumption and CPI series are from NIPA. PE ratios are from Robert Shiller's web site. Riskfree rates are the 1-month T-bill rates from Ken French's web site. Default rates and recovery rates are from Moody's.

We can use regressions to formally assess the relationship between default rates and macro variables. Altman et al. (2005) find that the levels and changes in default rates have strong explanatory power for recovery rates, while macro variables appear to explain little. However, default rates are themselves strongly affected by macroeconomic conditions: as shown in Table 3, the growth rates of industrial production, GDP, price-earnings ratio, and consumption all show significant explanatory power. For example, consumption growth and squared consumption growth alone can explain nearly half of the variation in default rates. The signs of the coefficients are as expected, with lower growth rates in industrial production, GDP, price-earnings ratio and consumption all leading to higher default rates. The squared consumption growth term captures the nonlinear relationship between default rates and consumption growth: default rates rise more rapidly when consumption growth becomes negative.

In Table 4, the univariate regression of recovery rates on default rates confirms the finding of Altman et al. (2005). However, a regression with only macro variables (PE, g and g^2) can explain 42% of the variation in recovery rates. This number increases to 50% where the riskfree rate is included. Default rates appear to contain information about recovery rates that is not captured by the macro variables. In a two-stage regression (Table 4, last column), the residuals from the regression of default rates on the other macro variables still have significant explanatory power for recovery rates, suggesting that other factors, such as the supply and demand of defaulted securities as identified by Altman et al. (2005), could also affect recovery rates.

In light of the regression results, I model default losses as a function of the expected growth rate $\theta_m(s)$ and volatility $\sigma_m(s)$ of aggregate consumption:

$$\alpha(s) = a_0 + a_1\theta_m(s) + a_2\theta_m^2(s) + a_3\sigma_m(s). \tag{33}$$

A quadratic term is included to help the model capture the long tail in default losses. I estimate the 4 coefficients for Baa and Aaa firms separately using the simulated method of moments. The target moments are: the mean and volatility of recovery rate, plus the correlations between recovery rate and default rate, price-dividend ratio, realized consumption growth. They are the same for Baa and Aaa firms.

The average recovery rate for all corporate bonds between 1982-2005 is \$41.1 per \$100 par, with a standard deviation of \$9.4. These numbers do not apply to debt instruments such as bank loans or mortgages, which likely have higher and more stable recovery rates. For example, Moody's report that the value-weighted average recovery rate of senior secured bank loans is \$64.2. According to the Flow of Funds data, bank loans account for a relatively small fraction of debt instruments ($10 \sim 20\%$). To be conservative, I assume that around 70% of debt instruments have recovery rates similar to corporate bonds, and the rest similar to bank loans, which leads to the estimates of mean and volatility of recovery rates for all debt instruments. The target moments and resulting estimates of the coefficients in equation (33) are given in Table 5.

5.2 Credit Spreads and Leverage Ratios

To illustrate the difficulty for standard structural models to generate reasonable credit spreads and leverage ratios, I first study the benchmark case of this model by shutting down the business-cycle variation in aggregate consumption and cash flows. I set all variables to their unconditional means, with two exceptions: the default loss coefficient α, and total volatility of cash flow, σ. I use these two variables to match the average recovery rate and 10-year default probability of a Baa and Aaa-rated firm.

Table 5: Estimating Default Losses

Panel A: Moments for Recovery Rates

Mean:	48%
Volatility:	7%
Correlation with default rates:	−0.77
Correlation with consumption growth:	0.40
Correlation with changes in price-earnings ratio:	0.48

Panel B: SMM Estimates

Baa:	$\alpha(s) = -0.04 - 12.88 \times \theta_m(s) + 209.02 \times \theta_m^2(s) + 10.29 \times \sigma_m(s)$
Aaa:	$\alpha(s) = -0.15 - 9.61 \times \theta_m(s) + 109.26 \times \theta_m^2(s) + 19.80 \times \sigma_m(s)$

Figure 6 reports results for a wide range of recovery rates, from \$40 to \$60. Credit spreads are rather insensitive to changes in recovery rates, while the optimal leverage ratio rises with the recovery rate. The latter is intuitive: as recovery rates rise, default losses drop, making firms take on more debt. Higher leverage raises the probability of default, which cancels out the effect of higher recovery rates on bond prices, thus leaving the credit spread flat. The expected excess returns for levered firms appear to be high, which is because these firms are highly-levered, making their dividend processes volatile. The rise in expected excess return with recovery rates is again due to rising leverage.

For Baa firms, with a relative risk aversion of 7.5, and a recovery rate of 48%, the model generates a credit spread of 57 bp for a 10-year coupon bond, far short of the average spread in the data (148 bp). The model predicts a leverage ratio of 67%, significantly higher than the average leverage of 42% for Baa firms, or 35% for all nonfinancial firms (according to the Flow of Funds Accounts data). The interest coverage, measured as the ratio of cash flow to interest expenses, is 0.7, much lower than the number in the data (around 3). These discrepancies highlight the dual puzzles of credit spreads and leverage ratio. The puzzles get worse as recovery rate rises. With a recovery rate of \$60, 10-year credit spread drops to 50 bp, while leverage ratio rises to 76%.

One can not resolve the puzzles simply by raising the risk aversion. While a higher risk aversion does push up the credit spreads, it increases the equity premium dramatically. Moreover, a higher risk aversion actually increases the leverage ratio. It does increase the expected costs of financial distress, which leads to lower optimal coupon rate lower and higher interest coverage. However, a drop in debt value comes with a bigger drop in equity value, resulting in a higher leverage ratio.

Table 6 compares the results after introducing variation in macroeconomic conditions with those of the benchmark case. I first consider the case that leaves out partial loss offset and equity issuance costs. A firm can lever up in any state. Rather than reporting

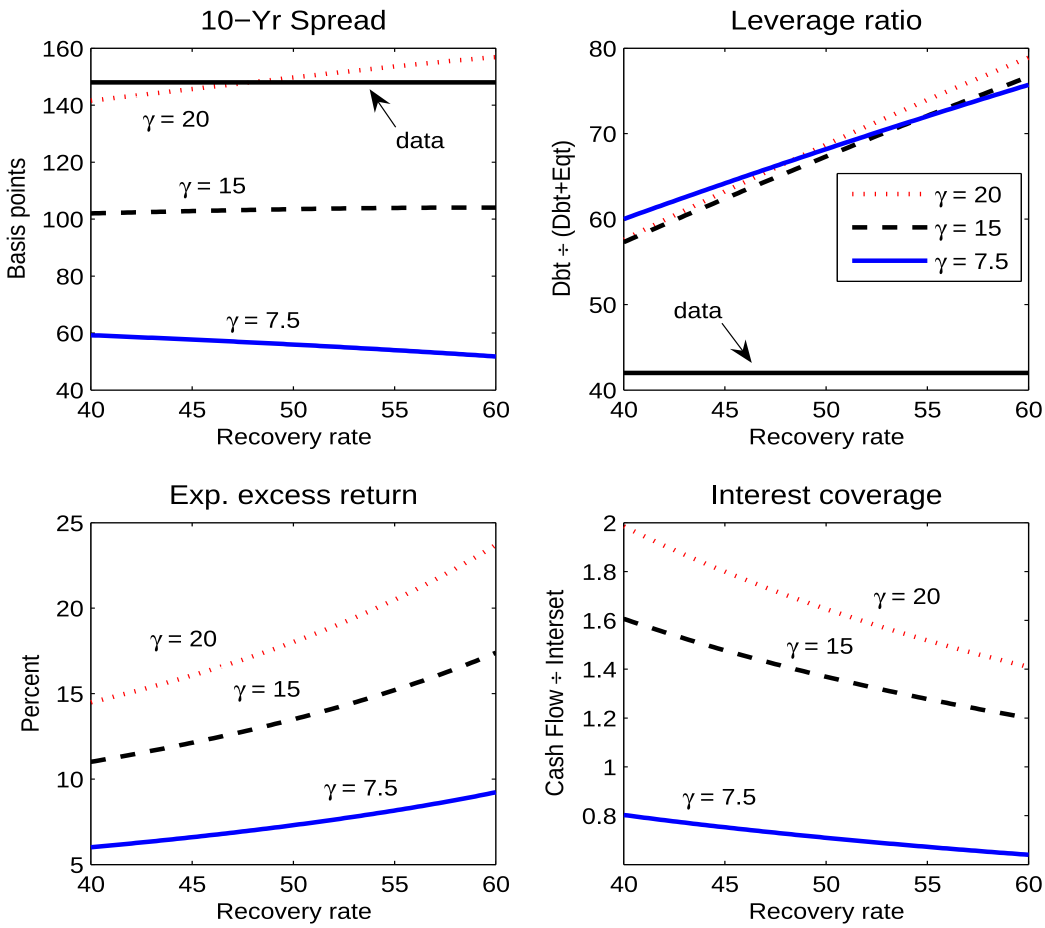

Figure 6: Benchmark Case: No Variation in Macroeconomic Conditions. All variables are set to their unconditional averages, except for α and σ, which are calibrated to match the recovery rate and 10-year default probability for Baa-rated firms (4.9%).

the results for all nine states, the table reports the average values across all states for each variable, along with their standard deviations. From the first few columns, we can see that the model matches the 10-year default probability, and the mean and volatility of recovery rates quite well.

The model has some success in addressing the two puzzles. It raises the average credit spread of a 10-year Baa-rated bond from 57 to 141 bp, while the average credit spread between Baa and Aaa-rated bonds is 98 bp. The market leverage drops from 67% in the benchmark model to 50%. The levered firm has an expected excess return of 9.3%, which is a little high. The value of the net tax benefits, which is the percentage increase in the value of a firm when it takes on optimal leverage, is about 5.3%, which is much lower than the 10.8% in the benchmark. Finally, since Aaa-rated firms have safer cash flows in this model, they have much higher leverage ratios and net tax benefits.

Table 6: Results For The Static Model

Panel A: Benchmark Case

	Def10	Rec	VolRec	Spr10	Lev	IntCov	TaxBen	Sprd	ERx
Baa	4.9%	48.0%	-	56.5	66.7%	0.7	10.8%	79.1	7.0%
Aaa	0.6%	48.0%	-	7.1	91.2%	0.1	20.0%	4.2	8.3%

Panel B: Model with Business-cycle Variation

	Def10	Rec	VolRec	Spr10	Lev	IntCov	TaxBen	Sprd	ERx
Baa	4.9%	47.3%	6.9%	141.3	50.4%	1.7	5.3%	262.8	9.3%
	(1.2%)	(1.7%)	(0.3%)	(9.7)	(3.5%)	(0.4)	(0.4%)	(50.5)	(1.6%)
Aaa	0.6%	47.7%	7.0%	43.4	52.2%	1.3	6.9%	81.4	6.6%
	(0.1%)	(1.4%)	(0.2%)	(1.1)	(1.8%)	(0.2)	(0.3%)	(17.1)	(1.2%)

Note: Def10 - 10-year cumulative default probability; Rec - average recovery rate for firm's debt; VolRec - volatility of recovery rates; Spr10 - average credit spread for a 10-year coupon bond; Lev - market leverage; IntCov - Interest Coverage (Cash Flow/Coupon); TaxBen - Net tax benefits as measured by percentage increases in firm value; sprd - average credit spread of consol bond; ERx - exp. excess return on equity.

The standard deviations for credit spreads reported in Table 6 do not measure the volatility of credit spreads for firms with certain credit ratings. They measure the deviation in credit spreads across optimally levered firms in different states. Because of the lumpy adjustment costs, a firm does not always adjust its capital structure immediately following a large shock. The firm's leverage ratio and credit spread will change, but not necessarily the credit rating, because rating agencies assign ratings through the cycle. Thus, the lumpiness of a firm's capital structure can lead to high volatilities in credit spreads.[8] I calculate the volatility of credit spreads across different states for the same bond issued in the normal state (with medium expected growth rate and medium volatility). For a 10-year Baa-rated bond, the volatility is 35.2 bp (40 bp in the data).

To see how the optimal leverage ratios, default boundaries, recovery rates, and default losses vary over the business cycle, I simulate the state of the economy for 100 years, and plot the corresponding values of the above variables in Figure 7. Recessions, marked with shades in the plots, are periods when the expected growth rates are negative. The darkness of the shade represents the severity of a recession. Those recessions with high (low) volatility are the most (least) severe, and are marked with the darkest (lightest)

[8] David (2006) argues that the time-varying leverage ratios can also lead to higher average credit spreads over time, because credit spreads are convex functions of the solvency ratio (inverse of leverage ratio). However, CCDG (2006) show that the bias due to convexity is small once the model is calibrated to match historical default rates, recovery rates, and Sharpe ratios.

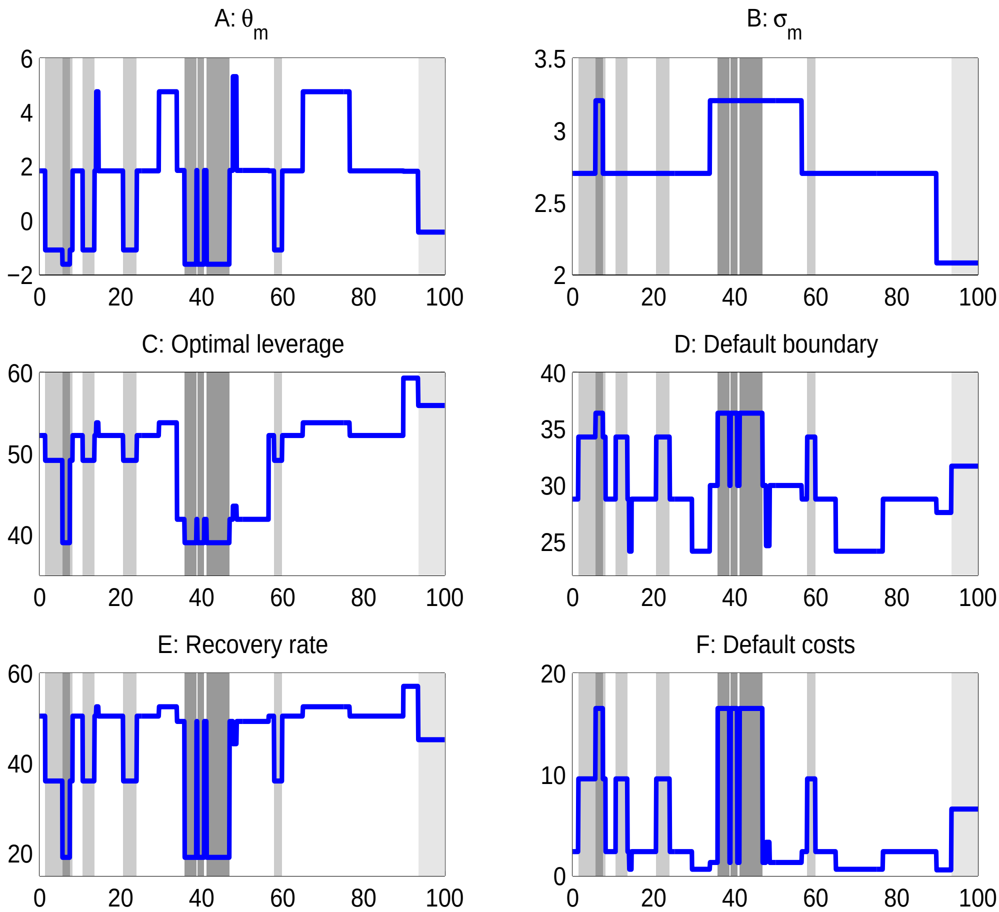

Figure 7: Conditional moments of aggregate consumption, optimal leverage ratios, default boundaries, recovery rates, and default losses in a simulation over 100 years. Default boundaries are relative to initial cash flow level. Default losses are relative to *pre*-distress firm value. All variables are in percentages.

shades. The optimal leverage ratios are lower in recessions, and they appear to be more sensitive to the movements in volatility than in expected growth rates. The default boundaries are higher in recessions, and they appear to be more sensitive to changes in the expected growth rates. Recovery rates are lower in recessions, especially when the volatility is high. Default losses, specified as percentages of *pre*-distress firm value, are rather low outside of recessions. They rise significantly in recessions, up to 16% in a most severe recession, which is still below the upper bound estimated by Andrade and Kaplan (1998).

The countercyclical default boundaries shown in Figure 7 implies that equity-holders will voluntarily default earlier (at higher cash flow levels) in recessions. This feature, combined with the fact that low expected growth rates and high uncertainty in bad times

Table 7: Comparative Statics For The Static Model – Baa Firms

	Def10	Rec	Spr10	Lev	IntCov	TaxBen	Sprd	ERx
Risky Tax	3.9%	48.8%	122.4	45.9%	1.96	4.7%	238.1	8.8%
Benefits	(0.8%)	(1.6%)	(6.2)	(2.9%)	(0.41)	(0.3%)	(42.0)	(1.6%)
Costly Eqt	4.7%	47.7%	136.5	49.1%	1.79	5.1%	255.9	9.2%
Issuance	(1.1%)	(1.6%)	(8.6)	(3.3%)	(0.38)	(0.4%)	(47.8)	(1.6%)
Combine	3.8%	49.2%	119.5	44.9%	2.02	4.6%	233.1	8.7%
Above Two	(0.8%)	(1.6%)	(5.7)	(2.8%)	(0.42)	(0.3%)	(40.5)	(1.6%)
Constant	13.4%	50.6%	281.8	65.6%	1.19	7.9%	321.9	12.1%
Deflt Costs	(0.7%)	(1.1%)	(28.0)	(0.5%)	(0.16)	(0.2%)	(42.6)	(2.4%)

Note: Def10 - 10-year cumulative default probability; Rec - average recovery rate for firm's debt; Spr10 - average credit spread for a 10-year coupon bond; Lev - market leverage; IntCov - Interest Coverage (Cash Flow/Coupon); TaxBen - Net tax benefits as measured by percentage increases in firm value; ERx - exp. excess return on equity.

make it more likely for a firm to enter into distress (by reaching low cash flow levels), results in high default probabilities in recessions.

Why do firms choose higher default boundaries in bad times? As pointed out by Geske (1977), equity-holders of a levered firm hold a perpetual compound option. At every point, they can either retain the option by making debt payments, or forfeit the firm's future cash flows to debt-holders in exchange for the waiver of their debt obligation. In bad times, higher risk premia lower the present value of future cash flows. Expected growth rates are also lower in bad times, which not only lower the present value of future cash flows directly, but also raise the probability of default and the probability that the firm loses part of its tax shield. I refer to the first channel as the "discount rate effect", the second the "cash flow effect". Both effects reduce the continuation value for equity-holders, which makes them default earlier. Finally, volatility is likely to be high in recessions, which makes the option to wait more valuable. I call this channel the "volatility effect". In the model, the discount rate and cash flow effects dominate the volatility effect, thus making firms default earlier in recessions.

Table 7 reports the results of four comparative static exercises. Case I considers the risk aspect of tax benefits by modeling partial loss offset (setting $\tau_c^- = 0.2$). Case II considers costly equity issuance, while Case III combines Cases I and II. All the other variables are unchanged. With partial loss offset, the optimal leverage drops significantly, to 45.9%. Notice that the model predicts a 10-year default probability of only 3.9%. If we recalibrate the model to match the default rate of 4.9%, the leverage ratio will drop further. In contrast, the effect of costly equity issuance appears to be small.

Case IV sets the default losses α to their unconditional mean. This case ignores partial loss offset and costs of equity issuance. The average leverage ratio jumps back to 65.6%, almost the same as the benchmark case. Moreover, the 10-year default probability is way too high. If we recalibrate the model to bring down the default probability, the leverage ratio will become even higher. The contrast between Case IV and the results in Table 6 Panel B clearly demonstrates the central role that countercyclical default losses play in explaining the leverage puzzle. When α is constant, procyclical variation in asset value leads to lower default losses in bad times, which actually generates a negative risky premium on the defaultable claim for equity-holders. This is why the leverage ratio becomes so much higher.

6 Dynamic Financing Decisions

The static model offers most of the intuition for the impact of macroeconomic conditions on the capital structure and credit spreads. The obvious limitation of the model is that firms cannot adjust their debt levels. This restriction is quite unrealistic, and it could introduce a bias in leverage decisions. Firms might take on "too much" debt initially because they cannot issue more debt later on. The opposite could also be true: firms might take on "too little" debt because they cannot reduce leverage when they get into trouble. Another undesirable feature is that the model generates nonstationary leverage ratios, which drop to zero as the firm size increases.

In this section, I address some of these concerns by adding the option of debt restructuring into firms' problems. For simplicity, I assume costs for downward restructuring are too high, so that firms only restructure upward. Gilson (1997) find that financially distressed firms remain highly leveraged because transaction costs discourage debt reduction outside of Chapter 11. Goldstein, Ju, and Leland (2001) argue that the value of the downward restructuring option is much smaller than that of the upward restructuring option for a healthy firm. Still, we need to be aware that excluding downward restructuring will lead to a downward bias in the optimal leverage ratio. Finally, given the limited effects of partial loss offset and equity issuance costs in the static model, I exclude these two features in the dynamic model.

The costs of debt issuance are the sum of two parts: the underwriting costs for the new debt, and the costs of negotiation to get the current creditors' approval to issue additional pari passu debt. With diffused debt ownership, the larger the amount of debt outstanding, the more costly the negotiation process. For simplicity, I assume both types of costs are 1%, which means the total debt issuance costs are 1% of the value of all debt after restructuring. These "quasi-fixed" issuance costs help generate stickiness in the capital structure.

Consider a firm's problem at time 0 as defined in (21). After the firm chooses a coupon rate C, it needs a default policy and a restructuring policy. The former is described by a set of default boundaries $(X_D^1, ..., X_D^n)$, and the latter is described by a set of upward restructuring boundaries $(X_U^1, ..., X_U^n)$. The firm will issue more debt when its cash flow is above a restructuring boundary while the economy is in the corresponding state.

As in the static model, I assume the ordering of the default boundaries is:

$$X_D^1 \leq X_D^2 \leq \cdots \leq X_D^n.$$

However, there is no guarantee that the restructuring boundaries will have the same order. To accommodate arbitrary orderings, I define a function $u(\cdot)$ that maps the order of restructuring boundaries across states into the indices for the states. For example, $u(i)$ denotes the state with the i^{th} lowest restructuring boundary. Then, by definition,

$$X_U^{u(1)} \leq X_U^{u(2)} \leq \cdots \leq X_U^{u(n)}.$$

For the parameters considered in this paper, the default and restructuring boundaries are sufficiently apart such that $X_D^n < X_U^{u(1)}$.

Next, I define the default regions $\mathcal{D}_k \triangleq \left[X_D^k, X_D^{k+1}\right)$ for $k < n$, and region $\mathcal{D}_n \triangleq \left[X_D^n, X_U^{u(1)}\right]$ where neither default or restructure will occur immediately due to a change of state. In addition, there are restructuring regions $\mathcal{D}_{n+k} \triangleq \left(X_U^{u(k)}, X_U^{u(k+1)}\right]$ for $k < n$.

After adding upward restructuring, for any corporate security $J(X_t, s_t)$, we have to specify not only the dividend $F\left(X_t, s_t\right)$ (before default and upward restructuring) and default payment $H\left(X_{\mathcal{T}_D}, s_{\mathcal{T}_D}\right)$, but also the restructuring payment at all potential restructuring points, $K\left(X_{\mathcal{T}_U}, s_{\mathcal{T}_U}\right)$.

Scaling Property

Since firms have infinite horizon, they face essentially the same problem at each restructuring point. The cash flow level will be higher than at the previous restructuring point, and the economy might be in a different state. Thus, the optimal capital structure problem is recursive and can be formulated into a dynamic programming problem. A scaling property, which is also used by Leland (1998), Goldstein, Ju, and Leland (2001) and Hackbarth, Miao, and Morellec (2006), can further simplify the problem into a static one. I state it formally in the following lemma.

Lemma 1 *If the state of the economy at the new restructuring point is the same as the state at the previous one, then the new coupon rate and default/restructuring boundaries will scale up by the same proportion from their previous values as do the cash flows. Moreover, for $k = 1, \cdots, n$, the value of total debt outstanding and the value of equity at the two restructuring points satisfy:*

$$D\left(X_U^k, k; C(X_U^k, k)\right) = \frac{X_U^k}{X_0} D\left(X_0, k; C(X_0, k)\right), \tag{34}$$

$$E\left(X_U^k, k; C(X_U^k, k)\right) = \frac{X_U^k}{X_0} E\left(X_0, k; C(X_0, k)\right). \tag{35}$$

Proof. See Appendix F. ■

The intuition of the scaling property is the following. If the state is the same, the firm essentially faces an identical problem, except that the current cash flow level is higher than at the previous restructuring point. This is due to the log-normality of cash flows, and the costs of debt issuance being proportional to the total amount of debt after restructuring. Thus, it will be optimal to to scale up the coupon rate, default boundaries, and restructuring boundaries by the same proportion as cash flows, which leaves the conditional probability of default and restructuring unchanged.

The scaling only holds after conditioning on the state, which introduces history dependence in the following sense. When the firm raises debt at time 0, the coupon rate will change with the initial state s_0, which in turn affects the default and restructuring boundaries. Suppose the firm reaches the restructuring boundary X_U^k at time t with $s_t = k$, then the scaling factor is X_U^k / X_0. When we apply this scaling factor to get the new debt level, default and restructuring boundaries, etc., we cannot scale up the time 0 values of these variables in state s_0, but their "shadow values" at time 0 in state k.

6.1 Debt and Equity

For debt, the dividend rate before default and restructuring is the same as in the static model, and so is the payment at default. After restructuring, the outstanding debt from previous issues gets diluted by new pari passu debt. Suppose at time 0, the state is i and cash flow equals X_0. Let the optimal coupon rate as a function of cash flow and state be $C(X, s)$. Then, if restructuring occurs in state j, the value of old debt becomes:

$$D\left(X_U^j, j; C\left(X_0, i\right)\right) = \frac{C(X_0, i)}{C(X_U^j, j)} D\left(X_U^j, j; C(X_U^j, j)\right) = \frac{C\left(X_0, i\right)}{C\left(X_0, j\right)} D\left(X_0, j; C\left(X_0, i\right)\right).$$

The first equality follows from the pari passu covenant, and the second equality is the result of the scaling property.

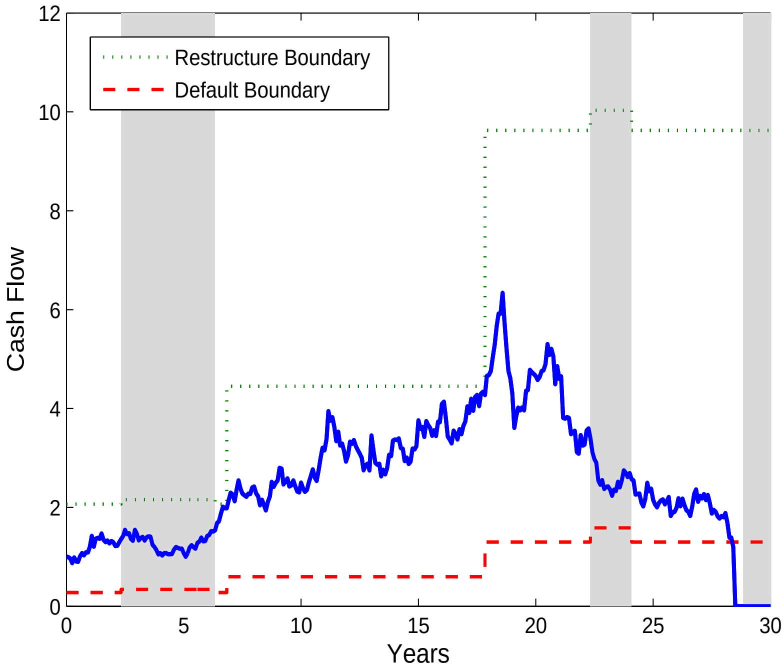

Figure 8: Cash Flow Sample Path for A Firm in the Dynamic Model. Light shades denote the state of low growth and median uncertainty.

When two consecutive restructurings occur in the same state, the value of old debt after restructuring will be exactly the same as its value at the previous restructuring point. The optimal coupon rate is higher in good times. Thus, if the economy is in a better state than at the previous restructuring point, the firm will increase its debt level disproportionately more, causing a bigger reduction in the value of old debt. If restructuring takes place in a worse state, there will be less dilution in old debt.

The following proposition gives the value of debt and equity for given coupon rate and default/restructuring boundaries.

Proposition 5 *Suppose a firm has a consol bond outstanding with coupon rate C; it follows a default policy characterized by a set of default boundaries $(X_D^1, \cdots, X_D^n)$, and a restructuring policy characterized by (i) a set of restructuring boundaries $(X_U^1, \cdots, X_U^n)$, and (ii) scaling of coupon rates and default/restructuring boundaries at each restructuring point. Then, the value of debt is:*

$$\mathbf{D}(X; C) = \sum_j \bar{w}_{k,j}^D \bar{\mathbf{g}}_{k,j} X^{\bar{\beta}_{k,j}} + \bar{\xi}_k^D X + \bar{\zeta}_k^D, \quad X \in \mathcal{D}_k, \quad k = 1, \cdots, 2n-1. \tag{36}$$

The value of equity is:

$$\mathbf{E}(X;C) = \sum_j \bar{w}^E_{k,j} \bar{\mathbf{g}}_{k,j} X^{\bar{\beta}_{k,j}} + \bar{\xi}^E_k X + \bar{\zeta}^E_k, \quad X \in \mathcal{D}_k, \quad k = 1, \cdots, 2n-1. \tag{37}$$

The coefficients $\bar{\mathbf{g}}$, $\bar{\beta}$, $(\bar{w}^D, \bar{\xi}^D, \bar{\zeta}^D)$ *and* $(\bar{w}^E, \bar{\xi}^E, \bar{\zeta}^E)$ *are given in Appendix G.*

Proof. See Appendix G. ■

Given the current debt level, the conditions for optimal default boundaries again translate into a set of nonlinear equations. As in the static model, these equations are easy to evaluate. The conditions for the restructuring boundaries based on the current debt level require a numerical evaluation of the smooth pasting conditions (see Appendix H). Finally, we can choose optimal debt level C to maximize the total value of the firm before levering up. The scaling property simplifies the problem into a static one, from which we can solve for the complete set of coupon rates $\{C_1, ..., C_n\}$ at time 0.

Figure 8 plots a sample path of cash flows for a firm in the dynamic problem. The firm enjoys strong growth in early periods. Cash flows rise and hit the restructuring boundaries twice, leading the firm to raise more debt. When the firm restructures, both the default and restructuring boundaries scale up proportionally. There are two additional cases where the default and restructuring boundaries change, which occur when the economy changes from the normal state into a state of low growth. The firm's luck reverses after 20 years, and its cash flow keeps declining until the default boundary is hit, and the firm is in default.

The graph shows that, like the default boundaries, restructuring boundaries are also countercyclical. The reasons that firms are less willing to restructure in bad times are similar to those that make firms default earlier in bad times. Restructuring is also an option to equity-holders. By exercising this option, a firm gets a bigger tax shield, but at the expense of the lump-sum debt issuance costs. The discount rate, cash flow, and volatility effects are still at work. The only difference is that earlier exercise implies lower boundaries for restructuring, but higher boundaries for default.

Table 8 reports the results for the dynamic model. In order to match the 10-year default probability of a Baa firm, the dynamic model requires a higher idiosyncratic volatility ($\sigma^i_f = 0.206$). Parameters for the default loss function (equation (33)) are also recalibrated. All the other cash flow parameters remain the same as in the static model. As expected, the optimal leverage ratio is significantly lower than in the static model (Table 6). On average, the optimal leverage is 41.8%, compared to 50.4% in the static model. The interest coverage rises to 2.1, and net tax benefits rise to 6.9%. Given the results from the static model, we expect that the optimal leverage ratio will drop further once the effect of partial loss offset and equity issuance costs are taken into account. The

Table 8: Results Of The Dynamic Model

	Def10	Rec	Spr10	Lev	IntCov	TaxBen	Sprd	ERx
Baa	4.9%	47.5%	139.8	41.8%	2.1	6.9%	261.5	7.7%
	(0.9%)	(1.5%)	(8.5)	(2.4%)	(0.3)	(0.4%)	(30.0)	(1.4%)
Aaa	0.6%	48.0%	42.4	45.5%	1.4	7.6%	78.8	5.9%
	(0.1%)	(1.2%)	(0.9)	(1.6%)	(0.1)	(0.3%)	(14.9)	(1.1%)

Note: Def10 - 10-year cumulative default probability; Rec - average recovery rate for firm's debt; Spr10 - average credit spread for a 10-year coupon bond; Lev - market leverage; IntCov - Interest Coverage (Cash Flow/Coupon); TaxBen - Net tax benefits as measured by percentage increases in firm value; ERx - exp. excess return on equity.

credit spreads of the consol and 10-year coupon bond in the dynamic model do not differ much from their values in the static model. This is because both models have similar default probabilities and recovery rates.

Collin-Dufresne and Goldstein (2001) argue that claim dilution due to firms issuing additional equal priority debt can raise credit spreads ex ante. Such effects appear to be small in this model. The impact of new debt on default probability is smaller in good times. Thus, the fact that firms are more likely to issue new debt in good times limits the effect of dilution.

Finally, to illustrate the countercyclical default rates and the clustering of defaults, I simulate 1000 identical firms over 50 years, and record the timing of defaults. These firms experience the same aggregate shocks, but have different outcome due to the idiosyncratic shocks. Figure 9 plots the default counts and corresponding annual default rates for a typical simulation. During this simulation, the economy experiences 3 states – (high growth, median uncertainty), (low growth, median uncertainty), and (low growth, high uncertainty) (at the end of 50 years). Most of the defaults occur in the latter two states where growth rate is low. The simulation nicely replicates the countercyclical default rates in the data, and we see the dramatic increase in default rate when the economy moves into the "worst state" – low growth and high uncertainty. In the graph of default counts, the two highest spikes occur right at the time when the economy moves from a high growth state into a low growth one. These are examples of default clustering: firms default at the same time due to the sudden increase of default boundary.

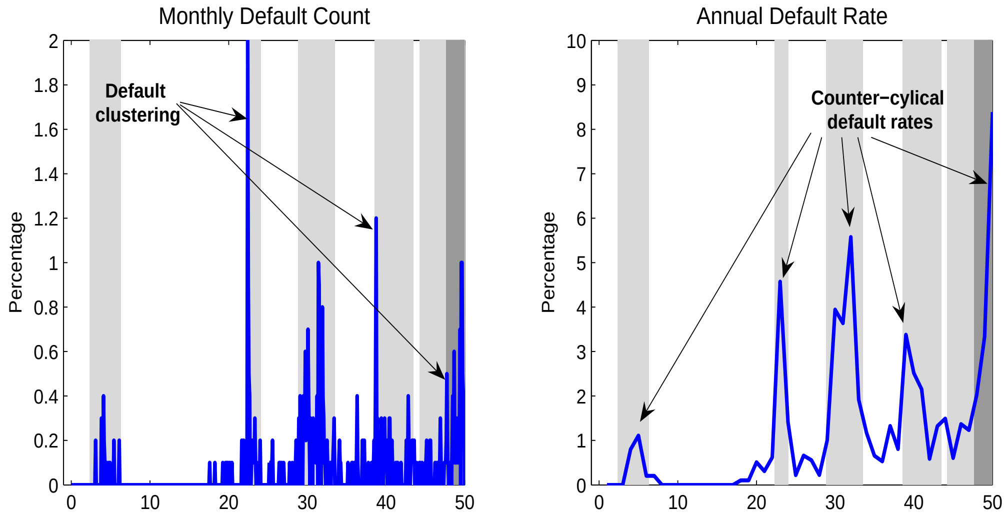

Figure 9: Simulated Default and the Annualized Default Rates. Areas with no shades are periods where the economy is in the state of high growth and median uncertainty. Light shades denote the state of low growth and median uncertainty. Dark shades denote the state of low growth and high uncertainty.

7 Concluding Remarks

Since defaults tend to concentrate in bad times when marginal utility is high, and default losses are particularly high during such times, investors will demand high risk premia for holding defaultable claims, including corporate bonds and levered firms. In this paper, I formally study these comovements in a structural model, and show that the risk premia are large enough to account for the credit spread puzzle and under-leverage puzzle.

I consider a basic trade-off model of capital structure. While the model abstracts from many realistic features, such as cash reserves, endogenous investments, different equity regimes (positive, zero or negative distributions, as in Hennessy and Whited 2005), agency costs, or strategic debt services (e.g., Anderson and Sundaresan (1996), Mella-Barral and Perraudin (1997)), it highlights the effects of macroeconomic conditions and risk premia on firms' financing decisions in a clean way. It will be interesting to see how macroeconomic conditions interact with the different features above in affecting the capital structure.

The dynamic model rules out downward restructuring. Such a restriction makes firms less capable of avoiding default. This could bias the leverage ratio downward ex ante, although it is unlikely to change the result that firms take on less debt in a dynamic model. The net effect depends on exactly how restructuring is done. In practice, firms do restructure their liabilities downward when they are in distress, and they do so through renegotiating debt, cutting investments, and selling assets. This model can be extended to incorporate "mechanical" asset sales (as in Strebulaev 2006). For more realistic features, we will need to endogenize investments and cash flows.

The model makes a prediction about debt capacity in the cross section. Since the covariation between default probabilities and macroeconomic conditions is key to low leverage, the model predicts that firms with more cyclical cash flows will have less debt. Firms and industries with more cyclical recovery rates should also take on less leverage.

The paper's finding on the connections between macroeconomic conditions and credit risk premia has important consequences for risk management. Cyclical risk factors can result in large swings in financial institutions' exposure to credit risk. As Allen and Saunders (2004) point out, under the new Basel Capital Accord, these risk factors affect bank capital requirements and lending capacity, which can exacerbate business-cycle fluctuations. This mechanism resembles the "financial accelerator" of Bernanke, Gertler, and Gilchrist (1996), but it is through a different channel. It will be very interesting to investigate this channel in a general equilibrium setting.

Finally, the model provides an important role for macro variables in the determination of credit spreads, which is consistent with the empirical findings of Collin-Dufresne, Goldstein, and Martin (2001) and Elton, Gruber, Agrawal, and Mann (2001) that market wide factors have additional explanatory power for credit spread variation. There is a large body of research that use default spreads (levels or changes) to predict returns for stocks and bonds (Cochrane (2006) surveys these studies). Unlike stocks, bond prices are less exposed to small cash flow shocks. Moreover, this model suggests that credit spreads are especially sensitive to risk prices in bad states. These features could make changes in credit spreads better proxies for the variation in risk prices than other variables such as price-dividend ratios.

Appendix

A Proof of Proposition 1

Proof. Before proceeding, I comment on the stochastic differential utility defined in (1). First, as Duffie and Epstein (1992b) point out, because the information structure of the economy is generated by both Brownian motions and Poisson jumps, (1) characterizes recursive utility for only a subclass of aggregators, which include the Kreps-Porteus specification in this paper (equation (2)). Second, the infinite-horizon specification used in this paper is defined as a pointwise limit of the finite horizon utilities. I assume that the technical conditions for this limit to be well defined (Duffie, Epstein and Skiadas (1992)) are satisfied.

To get the stochastic discount factor, we first need to solve for the value function of the representative household. In equilibrium, the representative household consumes aggregate output, which is given in (5). Thus, I directly define the value function of the representative agent as:

$$J\left(Y_t, s_t\right) = E_t\left[\int_0^\infty f\left(Y_{t+s}, J_{t+s}\right) ds\right]. \tag{A.1}$$

The Hamilton-Jacoby-Bellman equation in state i is:

$$0 = f\left(Y, J\left(Y,i\right)\right) + J_c\left(Y,i\right) Y\theta_m\left(i\right) + \frac{1}{2}J_{cc}\left(Y,i\right)Y^2\sigma_m^2\left(i\right) + \sum_{j\neq i}\lambda_{ij}\left(J\left(Y,j\right) - J\left(Y,i\right)\right). \tag{A.2}$$

There are n such differential equations for the n states. Thus, by using a Markov chain to model the expected growth rate and volatility, we replace a high-dimensional partial differential equation with a system of ordinary differential equations. As long as the number of states for the Markov chain is not too large, the ODE system will be relatively easy to handle.

I conjecture that the solution for J is:

$$J\left(Y,s\right) = \frac{\left(h\left(s\right)Y\right)^{1-\gamma}}{1-\gamma}, \tag{A.3}$$

where h is a function of the state variable s. Substituting J into the differential equations above, we get a system of nonlinear equations for h:

$$\begin{aligned} 0 \;=\;& \rho\frac{1-\gamma}{1-\delta}h(i)^{\delta-\gamma} + \left[\left(1-\gamma\right)\theta_m\left(i\right) - \frac{1}{2}\gamma\left(1-\gamma\right)\sigma_m^2\left(i\right) - \rho\frac{1-\gamma}{1-\delta}\right]h\left(i\right)^{1-\gamma} \\ & + \sum_{j\neq i}\lambda_{ij}\left(h\left(j\right)^{1-\gamma} - h\left(i\right)^{1-\gamma}\right), \qquad i = 1,\cdots,n \end{aligned} \tag{A.4}$$

where $\delta = 1/\psi$, the inverse of the intertemporal elasticity of substitution. These equations can be solved quickly using a nonlinear equation solver, even in the case when the number of states is fairly large, say 50.

Plugging J and Y into (3) gives:

$$m_t = \exp\left(\int_0^t \frac{\rho(1-\gamma)}{1-\delta}\left[\left(\frac{\delta-\gamma}{1-\gamma}\right)h(s_u)^{\delta-1}-1\right]du\right)\rho h(s_t)^{\delta-\gamma}Y_t^{-\gamma}. \tag{A.5}$$

Applying Ito's formula with jumps (see, e.g., Duffie 2001, Appendix F) to m, we get:

$$\frac{dm_t}{m_t} = -r(s_t)\,dt - \eta(s_t)\,dB_t + \sum_{s_t \neq s_{t-}}\left(e^{\kappa(s_{t-},s_t)}-1\right)dM_t^{(s_{t-},s_t)}, \tag{A.6}$$

where

$$\begin{aligned} r(i) &= -\frac{\rho(1-\gamma)}{1-\delta}\left[\left(\frac{\delta-\gamma}{1-\gamma}\right)h(i)^{\delta-1}-1\right]+\gamma\theta_m(i) \\ &\quad -\frac{1}{2}\gamma(1+\gamma)\sigma_m^2(i) - \sum_{j\neq i}\lambda_{ij}\left(e^{\kappa(i,j)}-1\right), && \text{(A.7a)} \\ \eta(i) &= \gamma\sigma_m(i), && \text{(A.7b)} \\ \kappa(i,j) &= (\delta-\gamma)\log\left(\frac{h(j)}{h(i)}\right). && \text{(A.7c)} \end{aligned}$$

Consider two special cases. In the first case, $\delta = 1/\psi = \gamma$. In this case, the normalized aggregator reduces to the standard CRRA utility. According to (A.7c), the stochastic discount factor does not jump in this case, so that large shocks are no longer priced. The risk free rate in this case simplifies to:

$$r(i) = \rho + \gamma\theta_m(i) - \frac{1}{2}\gamma(1+\gamma)\sigma_m^2(i).$$

Moreover, the nonlinear equation (A.4) simplifies to a linear equation, and $h(i)$ can be solved analytically.

Another special case is when $\psi = 1$. As $\psi \to 1$, the aggregator takes the form

$$f(c,v) = \rho(1-\gamma)v\left[\log(c) - \frac{1}{1-\gamma}\log((1-\gamma)v)\right]. \tag{A.8}$$

In this case, the value function is still given by (A.3), but the system of nonlinear equations for h becomes:

$$\begin{aligned} 0 &= -\rho(1-\gamma)h(i)^{1-\gamma}\log(h(s)) + (1-\gamma)\left(\theta(i)-\frac{1}{2}\gamma\sigma^2(i)\right)h(i)^{1-\gamma} \\ &\quad +\sum_{j\neq i}\lambda_{ij}\left(h(j)^{1-\gamma}-h(i)^{1-\gamma}\right), \qquad i = 1,\cdots,n. && \text{(A.9)} \end{aligned}$$

Similarly, the risk free rate becomes:

$$r(i) = \rho + \rho(1-\gamma)\log(h(i)) + \gamma\theta(i) - \frac{1}{2}\gamma(1+\gamma)\sigma^2(i) - \sum_{j\neq i}\lambda_{ij}\left(e^{\kappa(i,j)}-1\right), \tag{A.10}$$

and the relative jump size of the discount factor becomes:

$$\kappa(i,j) = (1-\gamma)\log\left(\frac{h(j)}{h(i)}\right). \tag{A.11}$$

There is an important distinction between the stochastic discount factor of the stochastic differential utility in this paper and that of the discrete time Epstein-Zin preferences in Bansal and Yaron (2004) for the case $\psi = 1$. In the BY model, shocks to expected growth rate and volatility are not priced when $\psi = 1$, which is not the case in this model. See Chen (2006) for more details of the derivations above and a comparison of the asset pricing implications of this model with that of Bansal and Yaron (2004). ■

B The Risk-neutral Measure

Let $(\mathbf{\Omega}, \mathfrak{F}, \mathcal{P})$ be the probability space on which the Brownian motions and Poisson processes in the model are defined. Let the corresponding information filtration be $(\mathfrak{F}_t)$. Applying Ito's formula with jumps to (10), we get the dynamics of the nominal stochastic discount factor n_t,

$$\frac{dn_t}{n_t} = -r^n(s_t)\,dt - \eta^m(s_t)\,dW_t^m - \eta^P dW_t^P + \sum_{s_t \neq s_{t-}} \left(e^{\kappa\left(s_{t-}, s_t\right)} - 1\right) dM_t^{\left(s_{t-}, s_t\right)}, \tag{B.1}$$

where the nominal risk-free rate is

$$r^n(s_t) = r(s_t) + \pi - \sigma_{P,1}\eta(s_t) - \sigma_P^2, \tag{B.2}$$

and the risk prices for the two Brownian motions are

$$\eta^m(s_t) = \eta(s_t) + \sigma_{P,1}, \tag{B.3}$$

$$\eta^P = \sigma_{P,2}. \tag{B.4}$$

We can define the risk-neutral measure $\mathcal{Q}$ associated with the nominal stochastic discount factor n_t (equation (B.1)) by specifying the density process ξ_t,

$$\xi_t = E_t\left[\frac{dQ}{dP}\right],$$

which evolves according to the following process:

$$\frac{d\xi_t}{\xi_t} = -\eta^m(s_t)\,dW_t^m - \eta^P dW_t^P + \sum_{s_t \neq s_{t-}} \left(e^{\kappa\left(s_{t-}, s_t\right)} - 1\right) dM_t^{\left(s_{t-}, s_t\right)}. \tag{B.5}$$

Applying the Girsanov theorem, we get the new standard Brownian motions under $\mathcal{Q}$, $\tilde{W}^m$

and $\tilde{W}^P$, which solve:

$$d\tilde{W}_t^m = dW_t^m + \eta^m(s_t)\,dt, \tag{B.6}$$
$$d\tilde{W}_t^P = dW_t^P + \eta^P dt. \tag{B.7}$$

The Girsanov theorem for point processes (see Elliott (1982)) gives the new jump intensity of the Poisson process under $\mathcal{Q}$:

$$\tilde{\lambda}_{jk} = E\left[e^{\kappa(j,k)}\right]\lambda_{jk} = e^{\kappa(j,k)}\lambda_{jk}, \quad j \neq k \tag{B.8}$$

which adjusts the intensity of the Poisson processes under measure $\mathcal{P}$ by the expected jump size of the density ξ_t. Finally, the diagonal elements of the generator has to be reset to make each row sum up to zero,

$$\tilde{\lambda}_{jj} = -\sum_{k\neq j}\tilde{\lambda}_{jk}. \tag{B.9}$$

These two equations characterize the new generator matrix $\tilde{\mathbf{\Lambda}}$ under $\mathcal{Q}$.

C Proof of Proposition 2

Proof. I compute the value of a cash flow stream by solving a system of ordinary differential equations. Veronesi (2000) provides an alternative proof, which exploits the right-continuity of the continuous-time Markov chain and obtains the same pricing formula with a limit argument.

Under the risk-neutral measure $\mathcal{Q}$, the nominal cash flow process for firm i is:

$$\frac{dX_t^i}{X_t^i} = \tilde{\theta}_X^i(s_{t-})\,dt + \sigma_{X,m}^i(s_{t-})\,d\tilde{W}_t^m + \sigma_{P,2}d\tilde{W}_t^P + \sigma_f^i dW_t^i,$$

where $\tilde{\theta}_X^i$ is the risk-neutral growth rate,

$$\tilde{\theta}_X^i(s_t) = \theta_X^i(s_t) - \sigma_{X,m}^i(s_{t-})\,\eta^m(s_{t-}) - \sigma_{P,2}\eta^P.$$

The total value of firm i's cash-flows before taxes is:

$$V^i\left(X_t^i, s_t\right) = \mathbb{E}_t^Q\left[\int_t^\infty \exp\left(-\int_t^\tau r^n(s_u)\,du\right)X_\tau^i d\tau\right]. \tag{C.1}$$

Define the log nominal cash flow $x_t^i \triangleq \log(X_t^i)$, total volatility

$$\sigma_X^i(s_t) \triangleq \sqrt{\left(\sigma_{X,m}^i(s_t)\right)^2 + \sigma_{P,2}^2 + \left(\sigma_f^i\right)^2}, \tag{C.2}$$

and a new Brownian motion that aggregates all the shocks for firm i,

$$d\tilde{W}_t^i \triangleq \frac{\sigma_{X,m}^i(s_t)}{\sigma_X^i(s_t)} d\tilde{W}_t^m + \frac{\sigma_{P,2}}{\sigma_X^i(s_t)} d\tilde{W}_t^P + \frac{\sigma_f^i}{\sigma_X^i(s_t)} dW_t^i. \tag{C.3}$$

Then, the risk-neutral dynamics of the log of firm i's cash flow can be written as:

$$dx_t^i = \left(\tilde{\theta_X^i}(s_t) - \frac{1}{2}\sigma_X^i(s_t)^2\right) dt + \sigma_X^i(s_t)\, d\tilde{W}_t^i. \tag{C.4}$$

Let $\mathbf{V}^i(x) = \left[V^i(x,1), ..., V^i(x,n)\right]'$ be the vector of firm i's asset values in n states. The Feynman-Kac formula implies that $\mathbf{V}^i$ satisfies the following system of ODEs:

$$\mathbf{r}^n \mathbf{V}^i = \left(\tilde{\theta}_X^i - \frac{1}{2}\boldsymbol{\Sigma}_X^i\right)\mathbf{V}_x^i + \frac{1}{2}\boldsymbol{\Sigma}_X^i \mathbf{V}_{xx} + \tilde{\boldsymbol{\Lambda}}\mathbf{V}^i + e^x \cdot \mathbf{1}, \tag{C.5}$$

where $\mathbf{r}^n \triangleq \mathbf{diag}\left([r^n(1), \cdots, r^n(n)]'\right)$, $\tilde{\theta}_X^i \triangleq \mathbf{diag}\left(\left[\tilde{\theta}_X^i(1), \cdots, \tilde{\theta}_X^i(n)\right]'\right)$, $\mathbf{1}$ is an $n \times 1$ vector of ones, and $\boldsymbol{\Sigma}_X^i \triangleq \mathbf{diag}\left(\left[\sigma_X^i(1)^2, \cdots, \sigma_X^i(n)^2\right]'\right)$. The set of boundary conditions are:

$$\lim_{x \downarrow -\infty} \mathbf{V}^i(x) = \mathbf{0}. \tag{C.6}$$

In fact, given the log-linear process for X^i, V^i must be linear in X^i, which will also satisfy the boundary conditions. Thus, I directly search for solution of the type:

$$\mathbf{V}^i(x) = e^x \cdot \mathbf{v}^i,$$

where $\mathbf{v}^i$ is an $n \times 1$ vector of constants. Plugging this guess into the ODE system gives:

$$\mathbf{r}^n \mathbf{v}^i = \left(\tilde{\theta}_X^i - \frac{1}{2}\boldsymbol{\Sigma}_X^i\right)\mathbf{v}^i + \frac{1}{2}\boldsymbol{\Sigma}_X^i \mathbf{v}^i + \tilde{\boldsymbol{\Lambda}}\mathbf{v}^i + \mathbf{1},$$

or

$$\left(\mathbf{r}^n - \tilde{\theta}_X^i - \tilde{\boldsymbol{\Lambda}}\right)\mathbf{v}^i = \mathbf{1}.$$

Thus,

$$\mathbf{v}^i = \left(\mathbf{r}^n - \tilde{\theta}_X^i - \tilde{\boldsymbol{\Lambda}}\right)^{-1}\mathbf{1}. \tag{C.7}$$

■

D Proof of Proposition 3 and 4

Proof. To simplify notation, I temporarily drop the superscripts that denote the cash flows of different firms. Start with a perpetual security $J(x_t, s_t)$, which pays a dividend rate $F(x_t, s_t)$ for as long as the firm is solvent, and a default payment $H(x_\tau, s_\tau)$ when default occurs at time τ. Let $\mathbf{F}(x)$ be an $n \times 1$ vector of dividend rate across n states, and $\mathbf{H}(x)$ an $n \times 1$ vector of the default payments. I also define an $n \times n$ diagonal matrix $\mathcal{A}$. Its ith diagonal element $\mathcal{A}^i$ is the infinitesimal generator for any C^2 function $\phi(x)$ in state i, where x is the log nominal cash flow specified in (C.2):

$$\mathcal{A}^i \phi(x) \triangleq \left(\tilde{\theta}_X(i) - \frac{1}{2}\sigma_X^2(i) \right) \frac{\partial}{\partial x}\phi(x) + \frac{1}{2}\sigma_X^2(i) \frac{\partial^2}{\partial x^2}\phi(x). \tag{D.1}$$

When cash flow X is in the region $\mathcal{D}_k = [X_D^k, X_D^{k+1})$ (for $k < n$), the firm will already be in default in all states $s > k$. Thus, the security will only be " alive" in the first k states. Define a set $\mathcal{I}_k \triangleq \{1, ..., k\}$ and its complement $\mathcal{I}_k^c \triangleq \{k+1, ..., n\}$. When $X \in \mathcal{D}_k$, the claims that are not in default yet are $\mathbf{J}_{[\mathcal{I}_k]}$, which satisfy the following system of ordinary differential equations:

$$\mathcal{A}_{[\mathcal{I}_k,\mathcal{I}_k]}\mathbf{J}_{[\mathcal{I}_k]} + \mathbf{F}_{[\mathcal{I}_k]} + \tilde{\mathbf{\Lambda}}_{[\mathcal{I}_k,\mathcal{I}_k]}\mathbf{J}_{[\mathcal{I}_k]} + \tilde{\mathbf{\Lambda}}_{[\mathcal{I}_k,\mathcal{I}_k^c]}\mathbf{H}_{[\mathcal{I}_k^c]} = \mathbf{r}^n_{[\mathcal{I}_k,\mathcal{I}_k]}\mathbf{J}_{[\mathcal{I}_k]}. \tag{D.2}$$

This equation states that, under the risk-neutral measure, the instantaneous expected return of a claim in any state should be equal to the riskfree rate in the corresponding state. A sudden change of the state can lead to abrupt changes in the value of the claim. It could also lead to immediate default, in which case the default payment is realized. These explain the last two terms on the LHS of the equation.

In regions $\mathcal{D}_n$ and $\mathcal{D}_{n+1}$, the firm is alive in all states. Sudden change of the state will not cause default. Thus, the ODE becomes:

$$\mathcal{A}\mathbf{J} + \mathbf{F} + \tilde{\mathbf{\Lambda}}\mathbf{J} = \mathbf{r}^n\mathbf{J}. \tag{D.3}$$

The homogeneous equation in region $\mathcal{D}_k$ can be written as:

$$\mathcal{A}_{[\mathcal{I}_k,\mathcal{I}_k]}\mathbf{J}_{[\mathcal{I}_k]} + \left(\tilde{\mathbf{\Lambda}}_{[\mathcal{I}_k,\mathcal{I}_k]} - \mathbf{r}^n_{[\mathcal{I}_k,\mathcal{I}_k]} \right) \mathbf{J}_{[\mathcal{I}_k]} = \mathbf{0}, \tag{D.4}$$

which is a quadratic eigenvalue problem (see Kennedy and Williams (1990)). Jobert and Rogers (2006) show its solution takes the following form:

$$\mathbf{J}(x)_{[\mathcal{I}_k]} = \sum_{j=1}^{2k} w_{k,j}\mathbf{g}_{k,j} \exp(\beta_{k,j} x), \tag{D.5}$$

where $\mathbf{g}_{k,j}$ and $\beta_{k,j}$ are solutions to the following standard eigenvalue problem:

$$\begin{bmatrix} 0 & \mathbf{I} \\ -\left(2\mathbf{\Sigma}_X^{-1}\left(\tilde{\mathbf{\Lambda}} - \mathbf{r}^n\right)\right)_{[\mathcal{I}_k,\mathcal{I}_k]} & -\left(2\mathbf{\Sigma}_X^{-1}\tilde{\theta}_X - \mathbf{I}\right)_{[\mathcal{I}_k,\mathcal{I}_k]} \end{bmatrix} \begin{bmatrix} \mathbf{g}_k \\ \mathbf{h}_k \end{bmatrix} = \beta_k \begin{bmatrix} \mathbf{g}_k \\ \mathbf{h}_k \end{bmatrix}, \tag{D.6}$$

where $\mathbf{I}$ is an $n \times n$ identity matrix, $\mathbf{r}^n$, $\tilde{\theta}_X$ and $\mathbf{\Sigma}_X$ are defined in (C.5). The coefficients $w_{k,j}$ will be different for different securities. Barlow, Rogers, and Williams (1980) show that there are exactly n eigenvalues with negative real parts, and n with positive real parts.

The remaining tasks are to find a particular solution for the inhomogeneous equation, and solve for the coefficients $w_{k,j}$ through the boundary conditions. Both the inhomogeneous equation and the boundary conditions will depend on the specific type of security under consideration.

D.1 Debt

Let $D(x,s)$ be the total value of corporate debt outstanding when the firm has log cash flow x and the economy is in state s. As shown earlier,

$$\begin{aligned} F(X,s) &= (1-\tau_i)\,C, && \text{(D.7)} \\ H(X,s) &= V_B(X,s). && \text{(D.8)} \end{aligned}$$

Plug these values into equation (D.2). When $X \in \mathcal{D}_k$ $(k<n)$, for those states $i \in \mathcal{I}_k$, the total value of debt satisfies:

$$\begin{aligned} r^n(i)D(x,i) &= \mathcal{A}^i D(x,i) + \tilde{\lambda}_{i,1}D(x,1) + \cdots + \tilde{\lambda}_{i,k}D(x,k) \\ &\quad + \tilde{\lambda}_{i,k+1}V_B(x,k+1) + \cdots + \tilde{\lambda}_{i,n}V_B(x,n) + (1-\tau_i)\,C. \qquad \text{(D.9)} \end{aligned}$$

Stacking up $D(X,s)$ in a vector, $\mathbf{D}(X) = [D(X,1),\cdots,D(X,n)]'$. We first get the following solution to the homogeneous equations,

$$\mathbf{D}(x)_{[\mathcal{I}_k]} = \sum_{j=1}^{2k} w_{k,j}^D \mathbf{g}_{k,j}\exp(\beta_{k,j}x), \quad k<n, \qquad \text{(D.10)}$$

where $\mathbf{g}_{k,j}$ and $\beta_{k,j}$ are characterized in the eigenvalue problem (D.6).

The inhomogeneous equation has the additional term that is linear in e^x:

$$\begin{aligned} & \tilde{\lambda}_{i,k+1}V_B(x,k+1) + \cdots + \tilde{\lambda}_{i,n}V_B(x,n) + (1-\tau_i)\,C \\ = \; & \left(1-\tau_c^+\right)(1-\tau_d)\sum_{j=k+1}^{n}\tilde{\lambda}_{ij}\,(1-\alpha(j))\,v(j)\,e^x + (1-\tau_i)\,C. \end{aligned}$$

Thus, I will seek a particular solution that is linear in e^x. For $X \in \mathcal{D}_k$ $(k<n)$,

$$D(x,i) = \xi_k^D(i)\,e^x + \zeta_k^D(i). \qquad \text{(D.11)}$$

The coefficients $\xi_k^D(i)$ and $\zeta_k^D(i)$ will be zero for $i \in \mathcal{I}_k^c$, because the firm is already in default in those states. Substituting the guess into (D.9) gives:

$$\begin{aligned} r^n(i)\left(\xi_k^D(i)\,e^x + \zeta_k^D(i)\right) &= \xi_k^D(i)\,\tilde{\theta}_X(i)\,e^x + \sum_{j=1}^{k}\tilde{\lambda}_{ij}\left(\xi_k^D(j)\,e^x + \zeta_k^D(j)\right) \\ &\quad + \left(1-\tau_c^+\right)(1-\tau_d)\sum_{j=k+1}^{n}\tilde{\lambda}_{ij}\,(1-\alpha(j))\,v(j)\,e^x + (1-\tau_i)\,C. \end{aligned}$$

After collecting terms, we get:

$$\begin{aligned}
r^n(i)\zeta_k^D(i) &= \sum_{j=1}^{k}\tilde{\lambda}_{ij}\zeta_k^D(j) + (1-\tau_i)\,C, \\
r^n(i)\xi_k^D(i) &= \xi_k^D(i)\,\tilde{\theta}_X(i) + \sum_{j=1}^{k}\tilde{\lambda}_{ij}\xi_k^D(j) + \left(1-\tau_c^+\right)(1-\tau_d)\sum_{j=k+1}^{n}\tilde{\lambda}_{ij}\left(1-\alpha(j)\right)v(j).
\end{aligned}$$

Thus,

$$\begin{aligned}
\xi_k^D(\mathcal{I}_k) &= \left(1-\tau_c^+\right)(1-\tau_d)\left(\mathbf{r}^n - \tilde{\mathbf{\Lambda}} - \tilde{\theta}_X\right)^{-1}_{[\mathcal{I}_k,\mathcal{I}_k]}\left(\tilde{\mathbf{\Lambda}}_{[\mathcal{I}_k,\mathcal{I}_k^c]}\left(\alpha\odot v\right)_{[\mathcal{I}_k^c]}\right), &\text{(D.12)}\\
\zeta_k^D(\mathcal{I}_k) &= (1-\tau_i)\,C\left(\mathbf{r}^n - \tilde{\mathbf{\Lambda}}\right)^{-1}_{[\mathcal{I}_k,\mathcal{I}_k]}\mathbf{1}_k, &\text{(D.13)}
\end{aligned}$$

where the symbol $\odot$ denotes element-by-element multiplication; $\xi_k^D(\mathcal{I}_k^c)$ and $\zeta_k^D(\mathcal{I}_k^c)$ are zero.

In the region $\mathcal{D}_n \cup \mathcal{D}_{n+1} = [X_D^n, +\infty)$, a sudden change of state will not lead to immediate default. Thus, the extra term in the inhomogeneous equation no longer depends on x. The total value of debt satisfies

$$r^n(i)D(x,i) = \mathcal{A}^i D(x,i) + \tilde{\lambda}_{i,1}D(x,1) + \cdots + \tilde{\lambda}_{i,n}D(x,n) + (1-\tau_i)\,C. \quad \text{(D.14)}$$

Now the solution to the homogeneous equation is:

$$\mathbf{D}(x) = \sum_{j=1}^{n} w_{n,j}^D\mathbf{g}_{n,j}\exp\left(\beta_{n,j}x\right), \quad \text{(D.15)}$$

and a particular solution in this region is:

$$D(x,i) = \zeta_n^D(i), \quad \text{(D.16)}$$

where

$$\zeta_n^D = (1-\tau_i)\,C\mathbf{b}. \quad \text{(D.17)}$$

To summarize, and rewrite the value of debt in terms of cash flows,

$$\begin{aligned}
\mathbf{D}(X)_{[\mathcal{I}_k]} &= \sum_{j=1}^{2k} w_{k,j}^D\mathbf{g}_{k,j}X^{\beta_{k,j}} + \xi_k^D(\mathcal{I}_k)X + \zeta_k^D(\mathcal{I}_k), \quad X\in\mathcal{D}_k, \quad k<n, &\text{(D.18)}\\
\mathbf{D}(X) &= \sum_{j=1}^{2n} w_{n,j}^D\mathbf{g}_{n,j}X^{\beta_{n,j}} + (1-\tau_i)\,C\mathbf{b}, \quad X\in\mathcal{D}_n\cup\mathcal{D}_{n+1}. &\text{(D.19)}
\end{aligned}$$

Next, I specify the boundary conditions that determine the coefficients $w_{k,j}^D$.

Debt value should be finite as x goes to infinity. To exclude any explosive terms, we need to set $w_{n,j}^D$ associated with all the $\beta_{n,j}$ with positive real parts (n of them) to zero. Then, the value of debt as cash flow gets large approaches that of a perpetuity without default risk:

$$\lim_{X \to +\infty} \mathbf{D}(X) = (1-\tau_i) C\mathbf{b}. \tag{D.20}$$

Another set of boundary conditions specify the value of debt at the n different default boundaries:

$$D\left(X_D^i, i\right) = V_B\left(X_D^i, i\right), \quad i = 1, \cdots, n. \tag{D.21}$$

Because the payoff function F and terminal payoff H are bounded and piecewise-continuous in X, while the discount rate r is constant in each state, an application of Theorem 4.9 (Karatzas and Shreve 1991, page 271) shows that $D(X, s)$ must be piecewise C^2 with respect to X over the region where it is defined, $[X_D^s, +\infty)$. Thus, for any $i \in \mathcal{I}_{n-1}$, we need to ensure that $D(X, i)$ is C^0 and C^1 at the boundaries $X_D^{i+1}, \cdots, X_D^n$:

$$\begin{aligned}
\lim_{X \uparrow X_D^k} D(X, i) &= \lim_{X \downarrow X_D^k} D(X, i), \quad k = i+1, ..., n \\
\lim_{X \uparrow X_D^k} D_X(X, i) &= \lim_{X \downarrow X_D^k} D_X(X, i), \quad k = i+1, ..., n
\end{aligned}$$

There are $2n^2$ of unknown coefficients for $\left\{w_{k,j}^D\right\}$. The continuity of D and its derivatives at the different default boundaries also give us $2n^2$ conditions. So we can solve for $\left\{w_{k,j}^D\right\}$ from a system of linear equations.

D.2 Equity

The dividend rate for equity naturally suggests a decomposition of equity into two parts, corresponding to the positive and negative part of the payoff,

$$E(x, i) = (1-\tau_d)\left(1-\tau_c^+\right) E^+(x, i) - \frac{1-\tau_c^-}{1-e} E^-(x, i), \tag{D.22}$$

First consider E^+. This claim's "dividend rate" and payment upon default are:

$$\begin{aligned}
F(X, s) &= \max(X - C, 0), && \text{(D.23)} \\
H(X, s) &= 0. && \text{(D.24)}
\end{aligned}$$

By definition, cash flow falls short of the interest expense in all regions except $\mathcal{D}_{n+1}$. So, for $X \in \mathcal{D}_k$ $(k \leq n)$, E^+ satisfies:

$$r^n(i) E^+(x, i) = \mathcal{A}^i E^+(x, i) + \tilde{\lambda}_{i,1} E^+(x, 1) + \cdots + \tilde{\lambda}_{i,k} E^+(x, k). \tag{D.25}$$

When $X \in \mathcal{D}_{n+1}$, E^+ satisfies:

$$r^n(i) E^+(x, i) = \mathcal{A}^i E^+(x, i) + \tilde{\lambda}_{i,1} E^+(x, 1) + \cdots + \tilde{\lambda}_{i,n} E^+(x, n) + (e^x - C). \tag{D.26}$$

Define $\mathbf{E}^{+}(x) = \left[\mathbf{E}^{+}(x,1), \cdots, \mathbf{E}^{+}(x,n)\right]'$. Again, the solution to the homogeneous equation in $\mathcal{D}_k$ $(k \leq n)$ is

$$\mathbf{E}^{+}(x)_{[\mathcal{I}_k]} = \sum_{j=1}^{2k} w_{k,j}^{E^+} \mathbf{g}_{k,j} \exp\left(\beta_{k,j} x\right).$$

In $\mathcal{D}_{n+1}$, the homogeneous equation is identical to that in the region $\mathcal{D}_n$. Thus, the solution shares the same $\mathbf{g}$ and β:

$$\mathbf{E}^{+}(x) = \sum_{j=1}^{2n} w_{n+1,j}^{E^+} \mathbf{g}_{n,j} \exp\left(\beta_{n,j} x\right).$$

As for the inhomogeneous equation in region $\mathcal{D}_{n+1}$, a natural guess of a particular solution is:

$$E^{+}(x,i) = \xi_{n+1}^{E^+}(i)\, e^{x} + \zeta_{n+1}^{E^+}(i).$$

Plug the guess into (D.26),

$$r^{n}(i)\left(\xi_{n+1}^{E^+}(i)\, e^{x} + \zeta_{n+1}^{E^+}(i)\right) = \xi_{n+1}^{E^+}(i)\, \tilde{\theta}_X(i)\, e^{x} + \sum_{j=1}^{n} \tilde{\lambda}_{ij}\left(\xi_{n+1}^{E^+}(j)\, e^{x} + \zeta_{n+1}^{E^+}(j)\right) + \left(e^{x} - C\right).$$

Collecting terms leads to:

$$\begin{aligned}
r^{n}(i)\zeta_{n+1}^{E^+}(i) &= \sum_{j=1}^{k} \tilde{\lambda}_{ij}\zeta_{n+1}^{E^+}(j) - C, \\
r^{n}(i)\xi_{n+1}^{E^+}(i) &= \tilde{\theta}(i)\, \xi_{n+1}^{E^+}(i) + \sum_{j=1}^{n} \tilde{\lambda}_{j,k}\xi_{n+1}^{E^+}(j) + 1.
\end{aligned}$$

Thus,

$$\xi_{n+1}^{E^+} = \left(\mathbf{r}^{n} - \tilde{\theta}_X - \tilde{\mathbf{\Lambda}}\right)^{-1} \mathbf{1}_n = \mathbf{v} \tag{D.27}$$

$$\zeta_{n+1}^{E^+} = -C\left(\mathbf{r}^{n} - \tilde{\mathbf{\Lambda}}\right)^{-1} \mathbf{1}_n = -C\mathbf{b} \tag{D.28}$$

The boundary conditions for E^{+} are similar to those for debt. As X becomes large, a firm becomes essentially free of default risk, which makes the claim $\mathbf{E}^{+}$ equivalent to the difference between a claim on the cash flow stream and a riskfree perpetuity.

$$\lim_{X\to+\infty} \mathbf{E}^{+}(X) = \left(X\mathbf{v} - C\mathbf{b}\right).$$

To satisfy this boundary condition, we need to set $w_{n+1,j}^{E^+}$ associated with all the $\beta_{n,j}$ with positive real parts.(n of them) to zero in region $\mathcal{D}_{n+1}$.

The rest of the boundary conditions:

$$E^{+}\left(X_D^{i}, i\right) = 0, \quad i = 1, \cdots, n.$$

Also need $E^{+}(X,i)$ to be C^0 and C^1 at $X_D^{i+1},\cdots,X_D^n$ and C for $i=1,...,n-1$,

$$\begin{aligned}
\lim_{X\uparrow X_D^k} E^{+}(X,i) &= \lim_{X\downarrow X_D^k} E^{+}(X,i), \quad k=i+1,...,n \\
\lim_{X\uparrow X_D^k} E_X^{+}(X,i) &= \lim_{X\downarrow X_D^k} E_X^{+}(X,i), \quad k=i+1,...,n \\
\lim_{X\uparrow C} E^{+}(X,i) &= \lim_{X\downarrow C} E^{+}(X,i) \\
\lim_{X\uparrow C} E_X^{+}(X,i) &= \lim_{X\downarrow C} E_X^{+}(X,i)
\end{aligned}$$

This time, there are $2n^2+2n$ of unknown coefficients for $\left\{w_{k,j}^{E^+}\right\}$. There are the same number of continuity conditions above, so we can solve for $\left\{w_{k,j}^{E^+}\right\}$ from a system of linear equations.

Next, consider E^-. This claim's "dividend rate" and payment upon default are:

$$F(X,s) = \max(C-X,0), \tag{D.29}$$

$$H(X,s) = 0. \tag{D.30}$$

For $X\in\mathcal{D}_k$ $(k\le n)$, the cash flow falls short of the interest expense. So, E^- satisfies

$$r^n(i)E^-(x,i) = \mathcal{A}^i E^-(x,i) + \tilde{\lambda}_{i,1}E^-(x,1) + \cdots + \tilde{\lambda}_{i,k}E^-(x,k) + (C-e^x). \tag{D.31}$$

For $X\in\mathcal{D}_{n+1}$, the cash flow exceeds the interest expense. So, E^- satisfies

$$r^n(i)E^-(x,i) = \mathcal{A}^i E^-(x,i) + \tilde{\lambda}_{i,1}E^-(x,1) + \cdots + \tilde{\lambda}_{i,n}E^-(x,n). \tag{D.32}$$

In this region, no particular solution is required.

The solutions to the homogeneous equations are:

$$\begin{aligned}
\mathbf{E}^-(x)_{[\mathcal{I}_k]} &= \sum_{j=1}^{2k} w_{k,j}^{E^-}\mathbf{g}_{k,j}\exp(\beta_{k,j}x), \quad for\ X\in\mathcal{D}_k\ (k\le n) \\
\mathbf{E}^-(x) &= \sum_{j=1}^{2n} w_{n+1,j}^{E^-}\mathbf{g}_{n,j}\exp(\beta_{n,j}x), \quad for\ X\in\mathcal{D}_{n+1}
\end{aligned}$$

Consider the inhomogeneous equation in the region $\mathcal{D}_k(k\le n)$. Try the following particular solution,

$$E^-(x,i) = \xi_k^{E^-}(i)\,e^x + \zeta_k^{E^-}(i).$$

The coefficients $\xi_k^{E^-}(i)$ and $\zeta_k^{E^-}(i)$ will be zero for $i\in\mathcal{I}_k^c$, because the firm is already in default in those states. Then,

$$r^n(i)\left(\xi_k^{E^-}(i)\,e^x + \zeta_k^{E^-}(i)\right) = \xi_k^{E^-}(i)\,\tilde{\theta}_X(i)\,e^x + \sum_{j=1}^{k}\tilde{\lambda}_{ij}\left(\xi_k^{E^-}(j)\,e^x + \zeta_k^{E^-}(j)\right) + (C-e^x)$$

Collecting terms, we get, for $i \in \mathcal{I}_k$,

$$\begin{aligned} r^n(i)\zeta_k^{E^-}(i) &= \sum_{j=1}^{k} \tilde{\lambda}_{ij}\zeta_k^{E^-}(j) + C, \\ r^n(i)\xi_k^{E^-}(i) &= \xi_k^{E^-}(i)\,\tilde{\theta}(i) + \sum_{j=1}^{k} \tilde{\lambda}_{ij}\xi_k^{E^-}(j) - 1. \end{aligned}$$

Thus,

$$\xi_k^{E^-}(\mathcal{I}_k) = -\left(\mathbf{r}^n - \tilde{\theta}_X - \tilde{\mathbf{\Lambda}}\right)^{-1}_{[\mathcal{I}_k,\mathcal{I}_k]} \mathbf{1}_k, \tag{D.33}$$

$$\zeta_k^{E^-}(\mathcal{I}_k) = C\left(\mathbf{r}^n - \tilde{\mathbf{\Lambda}}\right)^{-1}_{[\mathcal{I}_k,\mathcal{I}_k]} \mathbf{1}_k, \tag{D.34}$$

while $\xi_k^{E^-}(\mathcal{I}_k^c)$ and $\zeta_k^{E^-}(\mathcal{I}_k^c)$ are equal to zero.

The first set of boundary conditions specify that $\mathbf{E}^-$ should approach zero as X becomes large. This requires that, like E^+, the coefficients $w_{n+1,j}^{E^-}$ associated with all the $\beta_{n,j}$ with positive real parts.(n of them) must equal zero. The rest of boundary conditions are identical to those for E^+.

In summary,

$$E(x,i) = (1-\tau_d)\left(1-\tau_c^+\right)E^+(x,i) - \frac{1-\tau_c^-}{1-e}E^-(x,i), \tag{D.35}$$

where

$$\mathbf{E}^+(X)_{[\mathcal{I}_k]} = \sum_{j=1}^{2k} w_{k,j}^{E^+}\mathbf{g}_{k,j}X^{\beta_{k,j}}, \quad X \in \mathcal{D}_k, \quad k \leq n, \tag{D.36}$$

$$\mathbf{E}^+(X) = \sum_{j=1}^{n} w_{n+1,j}^{E^+}\mathbf{g}_{n,j}X^{\beta_{n,j}} + X\mathbf{v} - C\mathbf{b}, \quad X \in \mathcal{D}_{n+1}. \tag{D.37}$$

and

$$\mathbf{E}^-(X)_{[\mathcal{I}_k]} = \sum_{j=1}^{2k} w_{k,j}^{E^-}\mathbf{g}_{k,j}X^{\beta_{k,j}} + \xi_k^{E^-}(\mathcal{I}_k)\,X + \zeta_k^{E^-}(\mathcal{I}_k), \quad X \in \mathcal{D}_k, \quad k \leq n, \tag{D.38}$$

$$\mathbf{E}^-(X) = \sum_{j=1}^{n} w_{n+1,j}^{E^-}\mathbf{g}_{n,j}X^{\beta_{n,j}}, \quad X \in \mathcal{D}_{n+1}. \tag{D.39}$$

■

E Algorithm to Compute the First-Hitting Density of a Markov Modulated Process[9]

In this section, I provide an algorithm to compute the density of the stopping time $\mathcal{T}_D$. Consider the cash flow process

$$\frac{dX_t}{X_t} = \theta\left(s_t\right)dt + \sigma\left(s_t\right)dW_t. \tag{E.1}$$

There are n default boundaries $X_D^1, \cdots, X_D^n$ for the n states. The conditional stopping time $\mathcal{T}_D$ is defined as

$$\mathcal{T}_D\left(X_t, s_t\right) \triangleq \inf\left\{u \geq 0, X_{t+u} \leq X_D^i, s_{t+u} = i \ for \ \forall i \in \mathcal{I}_n \mid X_t, s_t\right\}. \tag{E.2}$$

If both the drift θ and diffusion σ are constant, then X_t becomes a geometric Brownian motion. The density of the first hitting time for a geometric Brownian motion is well-known (see, e.g., Borodin and Salminen 2002). Suppose X_t is a GBM with drift θ, diffusion σ, and starting value X. The probability that X will hit m ($m < X$) before t is:

$$F\left(X, m, t; \theta, \sigma\right) = N\left[h_1\left(t\right)\right] + \left(\frac{X}{m}\right)^{-2a} N\left[h_2\left(t\right)\right], \tag{E.3}$$

where

$$h_1\left(t\right) = \frac{-\ln\left(X/m\right) - a\sigma^2 t}{\sigma\sqrt{t}}, \quad h_2\left(t\right) = \frac{-\ln\left(X/m\right) + a\sigma^2 t}{\sigma\sqrt{t}}, \quad a = \frac{\theta}{\sigma^2} - \frac{1}{2}. \tag{E.4}$$

Now consider a time interval Δ. When Δ is sufficiently small, the probability that the Markov chain will change its state before Δ will be small. This is especially the case when the Markov chain is highly persistent. Thus, I can approximate cash flows X_t with $\tilde{X}_t$, whose drift and diffusion are controled by a discrete time Markov chain – state changes only occur at times $\Delta, 2\Delta, 3\Delta, \cdots$. Suppose the initial state is $s_0 = i$. Between 0 and Δ, θ and σ will remain equal to $\theta\left(i\right)$ and $\sigma\left(i\right)$. At time Δ, the probability of a change from state i to j is given by p_{ij}^{Δ}, where

$$\begin{aligned} p_{ij}^{\Delta} &= \lambda_{ij}\Delta, \quad i \neq j \\ p_{ii}^{\Delta} &= 1 - \sum_{j \neq i} \lambda_{ij}\Delta. \end{aligned}$$

Next, I characterize the density of the corresponding stopping time $\tilde{\mathcal{T}}_D$,

$$\tilde{\mathcal{T}}_D\left(\tilde{X}_t, \tilde{s}_t\right) \triangleq \inf\left\{u \geq 0, \ \tilde{X}_{t+u} \leq X_D^i, \tilde{s}_{t+u} = i \ for \ \forall i \in \mathcal{I}_n \mid \tilde{X}_t, \tilde{s}_t\right\}. \tag{E.5}$$

Let $V\left(X, i, t\right)$ be the probability that the firm will default before time T, conditional on the

[9] I thank Robert Novy-Marx for helpful discussions on this problem.

cash flow X and state i at time t.

$$V(X,i,t)=\Pr\left(\tilde{\mathcal{T}}_D\left(\tilde{X}_t=X,\tilde{s}_t=i\right)\leq T-t\right).$$

The firm could default at t:

$$V(X,i,t)=1 \quad X\leq X_D^i. \tag{E.6}$$

If $X>X_D^i$, one of the following two scenarios has to be true: $\tilde{X}$ can either hit X_D^s before $t+\Delta$, or reach a different value y without hitting the boundary. Then,

$$\begin{aligned} V(X,i,t) &= F\left(X,X_D^s,\Delta;\theta(s),\sigma(s)\right) \\ &\quad +\sum_{j=1}^{n} p_{ij}^{\Delta}\int_Y V(Y,j,t+\Delta)\,\mathbb{P}\left(\tilde{X}_{t+\Delta}\in dY,\tilde{\mathcal{T}}_D\geq\Delta|\tilde{X}_t=X,\tilde{s}_t=i\right), \end{aligned} \tag{E.7}$$

where $\mathbb{P}\left(\tilde{X}_{t+\Delta}\in dY,\tilde{\mathcal{T}}_D\geq\Delta|\tilde{X}_t=X,\tilde{s}_t=i\right)$ is the transition probability that $\tilde{X}$ reaches Y at $t+\Delta$ without hitting the boundary between t and $t+\Delta$. Since the state does not change between t and $t+\Delta$, it is the same as the transition density of a geometric Brownian motion travelling from X to Y within Δ without hitting the lower bound X_D^i. The following lemma characterizes the value of this probability.

Lemma 2 *Let X be a geometric Brownian motion with drift θ, diffusion σ and initial value x. Let $\mathcal{T}$ be the stopping time for X hitting z $(z<x)$. Then, for $y>z$,*

$$\begin{aligned} &\mathbb{P}\left(X_T\in dy,\mathcal{T}>T|\,X_0=x\right) \\ &= \left\{\frac{1}{\sqrt{2\pi\sigma^2T}}\exp\left(-\frac{\left(\ln\left(\frac{y}{x}\right)-\left(\theta-\frac{\sigma^2}{2}\right)T\right)^2}{2\sigma^2T}\right)\right. \\ &\quad \left.-\frac{-\ln\left(\frac{z}{x}\right)}{2\sigma^2\pi}\left[\int_0^T\frac{s^{-3/2}}{\sqrt{T-s}}\exp\left(-\frac{\left(\ln\left(\frac{z}{x}\right)-\left(\theta-\frac{\sigma^2}{2}\right)s\right)^2}{2\sigma^2 s}-\frac{\left(\ln\left(\frac{y}{z}\right)\right)^2}{2\sigma^2(T-s)}\right)ds\right]\right\}\frac{1}{y}\,dy. \end{aligned} \tag{E.8}$$

Proof. Let W_t be a standard Brownian motion, then X_T is distributed as $X_0\exp\left[\sigma\left(\left(\frac{\theta}{\sigma}-\frac{\sigma}{2}\right)T+x_T\right)\right]$.

$$\mathbb{P}\left(X_T\in dy,\mathcal{T}>T|\,X_0=x\right)=\mathbb{P}\left(X_T\in dy|\,X_0=x\right)-\mathbb{P}\left(X_T\in dy,\mathcal{T}\leq T|\,X_0=x\right),$$

where the first term is easy to compute:

$$\mathbb{P}\left(X_T\in dy|\,X_0=x\right)=\frac{1}{\sqrt{2\pi\sigma^2T}}\exp\left(-\frac{\left(\ln\left(\frac{y}{x}\right)-\left(\theta-\frac{\sigma^2}{2}\right)T\right)^2}{2\sigma^2T}\right)\frac{1}{y}\,dy.$$

Define $\tilde{W}_t=\left(\frac{\mu}{\sigma}-\frac{\sigma}{2}\right)t+W_t$, which is a Brownian motion with a drift under $\mathbb{P}$. Then X_t is equal to $X_0 \exp\left(\sigma \tilde{W}_t\right)$. Let $\tilde{\tau}$ be the stopping time for $\tilde{W}_t$ hitting lower bound m. Define a new measure $\tilde{\mathbb{P}}$ with the Radon-Nykodym derivative Z_t,

$$\begin{aligned} Z_t &= \exp\left(-\int_0^t \left(\frac{\theta}{\sigma} - \frac{\sigma}{2}\right) dW_s - \frac{1}{2}\left(\frac{\theta}{\sigma} - \frac{\sigma}{2}\right)^2 t\right) \\ &= \exp\left(-\left(\frac{\theta}{\sigma} - \frac{\sigma}{2}\right) \tilde{W}_t + \frac{1}{2}\left(\frac{\theta}{\sigma} - \frac{\sigma}{2}\right)^2 t\right), \end{aligned}$$

then $\tilde{W}_t$ will be a standard Brownian motion under $\tilde{\mathbb{P}}$ (see, e.g., Karatzas and Shreve 1991). Then, using the first hitting density and transition density of a standard Brownian motion, and the fact that Z_t^{-1} is a martingale under $\tilde{\mathbb{P}}$, we have:

$$\begin{aligned} \mathbb{P}\left(\tilde{W}_T \in db, \tilde{\tau} \le T\right) &= \mathbb{E}\left[1_{\{\tilde{x}_T \in db, \tilde{\tau} \le T\}}\right] = \tilde{\mathbb{E}}\left[1_{\{\tilde{x}_T \in db, \tilde{\tau} \le T\}} Z_T^{-1}\right] \\ &= \tilde{\mathbb{E}}\left[1_{\{\tilde{x}_T \in db, \tilde{\tau} \le T\}} \tilde{\mathbb{E}}\left[Z_T^{-1} \middle| \mathfrak{F}_{\tilde{\tau}\wedge T}\right]\right] = \tilde{\mathbb{E}}\left[1_{\{\tilde{x}_T \in db, \tilde{\tau} \le T\}} Z_{\tilde{\tau}\wedge T}^{-1}\right] \\ &= \tilde{\mathbb{E}}\left[1_{\{\tilde{x}_T \in db, \tilde{\tau} \le T\}} \exp\left(\left(\frac{\theta}{\sigma} - \frac{\sigma}{2}\right) m - \left(\frac{\theta}{\sigma} - \frac{\sigma}{2}\right)^2 \frac{\tilde{\tau}}{2}\right)\right] \\ &= \int_0^T \exp\left(\left(\frac{\theta}{\sigma} - \frac{\sigma}{2}\right) m - \left(\frac{\theta}{\sigma} - \frac{\sigma}{2}\right)^2 \frac{s}{2}\right) \tilde{\mathbb{P}}\left(\tilde{\tau} \in ds\right) \tilde{\mathbb{P}}\left(\tilde{x}_T \in db | \tilde{\tau} = s\right) \\ &= \int_0^T \exp\left(\left(\frac{\theta}{\sigma} - \frac{\sigma}{2}\right) m - \left(\frac{\theta}{\sigma} - \frac{\sigma}{2}\right)^2 \frac{s}{2}\right) \frac{-m e^{-\frac{m^2}{2s}}}{s^{3/2}\sqrt{2\pi}} \frac{e^{-\frac{(b-m)^2}{2(T-s)}}}{\sqrt{2\pi\left(T-s\right)}} \, ds \, db \\ &= \frac{-m}{2\pi}\left[\int_0^T \frac{s^{-3/2}}{\sqrt{T-s}} \exp\left(-\frac{\left(m - \left(\frac{\theta}{\sigma} - \frac{\sigma}{2}\right)s\right)^2}{2s} - \frac{\left(b-m\right)^2}{2\left(T-s\right)}\right) ds\right] db. \end{aligned}$$

Then, it follows from a change of variables that:

$$\begin{aligned} &\mathbb{P}\left(X_T \in dy, \mathcal{T} \le T \middle| X_0 = x\right) \\ =& \frac{-\ln\left(\frac{z}{x}\right)}{2\sigma^2\pi}\left[\int_0^T \frac{s^{-3/2}}{\sqrt{T-s}} \exp\left(-\frac{\left(\ln\left(\frac{z}{x}\right) - \left(\theta - \frac{\sigma^2}{2}\right)s\right)^2}{2\sigma^2 s} - \frac{\left(\ln\left(\frac{y}{z}\right)\right)^2}{2\sigma^2\left(T-s\right)}\right) ds\right] \frac{1}{y} \, dy. \end{aligned}$$

Putting two pieces together leads to the result. ■

Finally, the terminal conditions for V at time T are:

$$V\left(X, i, T\right) = \begin{cases} 0, & X > X_D^i \\ 1, & X \le X_D^i \end{cases}, \tag{E.9}$$

and we can solve equations (E.6) and (E.7) backward to get $V\left(X, s, 0\right)$.

F Proof of Lemma 1

Proof. Outline of the proof: I first show that, conditional on the state, the value of debt, equity and the boundary values, are all homogeneous of degree 1 in (X, C); then, I show that, again conditional on the state, the optimal C is proportional to X.

The value of debt is:

$$\begin{aligned} D\left(X_0, s_0, C\right) &= \mathbb{E}\left[\left.\int_0^{\mathcal{T}_D \wedge \mathcal{T}_U} \frac{n_t}{n_0}\left(1-\tau_i\right) C dt \right| \mathfrak{F}_0\right] \\ &+\mathbb{E}\left[\left. 1_{\{\mathcal{T}_U > \mathcal{T}_D\}} \frac{n_{\mathcal{T}_D}}{n_0}\left(1-\tau_{eff}\right) V_B\left(X_{\mathcal{T}_D}, s_{\mathcal{T}_D}\right) \right| \mathfrak{F}_0\right] \\ &+\mathbb{E}\left[\left. 1_{\{\mathcal{T}_U < \mathcal{T}_D\}} \frac{n_{\mathcal{T}_U}}{n_0} \frac{C}{C\left(X_{\mathcal{T}_U}, s_{\mathcal{T}_U}\right)} D\left(X_{\mathcal{T}_U}, s_{\mathcal{T}_U}; C\left(X_{\mathcal{T}_U}, s_{\mathcal{T}_U}\right)\right) \right| \mathfrak{F}_0\right] \end{aligned} \tag{F.1}$$

When scaling (X_0, C) to (aX_0, aC), if the firm also scales up the default and restructuring boundaries by a, the distributions of $\mathcal{T}_D$ and $\mathcal{T}_U$ will be unchanged. Under this condition, we have:

$$D\left(aX_0, s_0, aC\right) = aD\left(X_0, s_0, C\right). \tag{F.2}$$

The value of equity *after restructuring* is given by the following Bellman equation:

$$\begin{aligned} E\left(X_0, s_0, C\right) &= \max_{\mathcal{T}_D, \mathcal{T}_U, C\left(X_{\mathcal{T}_U}, s_{\mathcal{T}_U}\right)} \left\{ \mathbb{E}\left[\left.\int_0^{\mathcal{T}_D \wedge \mathcal{T}_U} \frac{n_t}{n_0}\left(1-\tau_{eff}\right)\left(X_t - C\right) dt \right| \mathfrak{F}_0\right] \right. \\ &+\mathbb{E}\left[\left. 1_{\{\mathcal{T}_U < \mathcal{T}_D\}} \frac{n_{\mathcal{T}_U}}{n_0}\left(1-\tau_{eff}\right) \left[\begin{array}{c} \left(1-q\right) D\left(X_{\mathcal{T}_U}, s_{\mathcal{T}_U}; C\left(X_{\mathcal{T}_U}, s_{\mathcal{T}_U}\right)\right) \\ -\frac{C}{C\left(X_0, s_{\mathcal{T}_U}\right)} D\left(X_0, s_{\mathcal{T}_U}; C\left(X_0, s_{\mathcal{T}_U}\right)\right) \end{array} \right] \right| \mathfrak{F}_0\right] \\ &\left. +\mathbb{E}\left[\left. 1_{\{\mathcal{T}_U < \mathcal{T}_D\}} \frac{n_{\mathcal{T}_U}}{n_0} E\left(X_{\mathcal{T}_U}, s_{\mathcal{T}_U}, C\left(X_{\mathcal{T}_U}, s_{\mathcal{T}_U}\right)\right) \right| \mathfrak{F}_0\right] \right\}. \end{aligned} \tag{F.3}$$

Suppose the optimal stopping times and coupon rates are $\left(\mathcal{T}_D^*, \mathcal{T}_U^*, C^*\left(X_{\mathcal{T}_U}, s_{\mathcal{T}_U}\right)\right)$. When changing (X_0, C) to (aX_0, aC), it is feasible for the firm to scale up future coupons and boundaries by a, which will again leave the distributions of $\mathcal{T}_D^*$ and $\mathcal{T}_U^*$ unchanged. Then,

$$\begin{aligned} & E\left(aX_0, s_0, aC\right) \geq \mathbb{E}\left[\left.\int_0^{\mathcal{T}_D^* \wedge \mathcal{T}_U^*} \frac{n_t}{n_0}\left(1-\tau_{eff}\right)\left(X_t - aC\right) dt \right| \mathfrak{F}_0\right] \\ & +\mathbb{E}\left[\left. 1_{\{\mathcal{T}_U^* < \mathcal{T}_D^*\}} \frac{n_{\mathcal{T}_U^*}}{n_0}\left(1-\tau_{eff}\right) \left[\begin{array}{c} \left(1-q\right) D\left(aX_{\mathcal{T}_U}, s_{\mathcal{T}_U}; aC\left(X_{\mathcal{T}_U}, s_{\mathcal{T}_U}\right)\right) \\ -\frac{aC}{aC\left(X_0, s_{\mathcal{T}_U}\right)} D\left(aX_0, s_{\mathcal{T}_U^*}; aC\left(X_0, s_{\mathcal{T}_U^*}\right)\right) \end{array} \right] \right| \mathfrak{F}_0\right] \\ & +\mathbb{E}\left[\left. 1_{\{\mathcal{T}_U^* < \mathcal{T}_D^*\}} \frac{n_{\mathcal{T}_U^*}}{n_0} E\left(aX_{\mathcal{T}_U^*}, s_{\mathcal{T}_U^*}, aC\left(X_{\mathcal{T}_U^*}, s_{\mathcal{T}_U^*}\right)\right) \right| \mathfrak{F}_0\right] \\ & \geq aE\left(X_0, s, C\right). \end{aligned} \tag{F.4}$$

Applying the same argument to the case when scaling from (aX_0, aC) to (X_0, C) leads to:

$$E\left(X_0, s, C\right) \geq \frac{1}{a} E\left(aX_0, s, aC\right). \tag{F.5}$$

Thus,

$$E(aX_0, s, aC) = aE(X_0, s, C). \tag{F.6}$$

Since the optimum is obtained by scaling future coupon rates and default boundaries by a, these choices must be optimal. It is also straightforward to directly check that optimal boundaries are homogeneous of degree 1 in (X, C). Suppose the default boundary for state s is x. According to the smooth-pasting conditions, x must satisfy:

$$\left.\frac{\partial}{\partial X} E(X, s, C)\right|_{X=x} = 0. \tag{F.7}$$

We also have

$$\left.\frac{\partial}{\partial X} E(X, s; aC)\right|_{X=ax} = a \frac{\partial}{\partial X} E\left(\frac{X}{a}, s; C\right)\Big|_{X=ax} = 0, \tag{F.8}$$

suggesting that the optimal default boundary after scaling is indeed ax. Essentially the same argument shows that the optimal restructuring boundaries should scale up by a as well.

Finally, we need to show that the optimal initial coupon rate C is indeed proportional to X. The initial coupon rate is chosen to maximize the value of equity before issuing debt,

$$E_U(X, s) = \max_C \{(1-q) D(X, s, C) + E(X, s, C)\}. \tag{F.9}$$

Since both $D(X, s, C)$ and $E(X, s, C)$ are homogeneous of degree 1 in (X, C), we can repeat the "sandwich" argument above to show that the optimal C must be proportional to X. ■

G Proof of Proposition 5

Proof. After adding the option of upward restructuring, for any corporate perpetual security $J(x_t, s_t)$, we need to specify restructuring payment $K(x_{\mathcal{T}_U}, s_{\mathcal{T}_U})$, in addition to dividend rate $F(x_t, s_t)$ and default payment $H(x_{\mathcal{T}_D}, s_{\mathcal{T}_D})$.

Now we have the following boundaries, $\left(X_D^1, \cdots, X_D^n, X_U^{u_1}, \cdots, X_U^{u_n}\right)$. Denote the regions with

$$\mathcal{D}_1, \cdots \mathcal{D}_n, \mathcal{D}_{n+1}, \cdots, \mathcal{D}_{2n-1},$$

where

$$\begin{aligned}
\mathcal{D}_k &= [X_D^k, X_D^{k+1}), \quad k = 1, \cdots, n-1 \\
\mathcal{D}_n &= \left[X_D^n, X_U^{u_1}\right], \\
\mathcal{D}_{n+k} &= (X_U^{u(k)}, X_U^{u(k+1)}], \quad k = 1, \cdots, n-1.
\end{aligned}$$

I use index set $\mathcal{I}_{n+k} \triangleq \{u(k+1), \cdots, u(n)\}$ to denote states where the firm has not yet restructured, with its compliment $\mathcal{I}_{n+k}^c \triangleq \{u(1), \cdots, u(k)\}$ denoting the states where restructuring has occurred.

In regions $\mathcal{D}_k$ $(k < n)$, the equation governing J is identical to those in the static case (see equation (D.2)). The same is true in $\mathcal{D}_n$, where the firm will neither default nor restructure because of a sudden change of state. Thus, I will focus on the restructuring regions. In $\mathcal{D}_{n+k}$ $(k < n)$, the firm has not restructured yet for any of the states in $\mathcal{I}_{n+k}$, thus:

$$\mathcal{A}_{[\mathcal{I}_{n+k},\mathcal{I}_{n+k}]}\mathbf{J}_{[\mathcal{I}_{n+k}]} + \mathbf{F}_{[\mathcal{I}_{n+k}]} + \tilde{\mathbf{\Lambda}}_{[\mathcal{I}_{n+k},\mathcal{I}_{n+k}]}\mathbf{J}_{[\mathcal{I}_{n+k}]} + \tilde{\mathbf{\Lambda}}_{[\mathcal{I}_{n+k},\mathcal{I}^c_{n+k}]}\mathbf{K}_{[\mathcal{I}^c_{n+k}]} = \mathbf{r}^n_{[\mathcal{I}_{n+k},\mathcal{I}_{n+k}]}\mathbf{J}_{[\mathcal{I}_{n+k}]}. \tag{G.1}$$

The homogeneous equation in region $\mathcal{D}_k$ $(k \leq n)$ is the same as in the static model, and will have the same solution. The homogeneous equation in region $\mathcal{D}_{n+k}$ $(k \leq n-1)$ can be written as:

$$\mathcal{A}_{[\mathcal{I}_{n+k},\mathcal{I}_{n+k}]}\mathbf{J}_{[\mathcal{I}_{n+k}]} + \left(\tilde{\mathbf{\Lambda}}_{[\mathcal{I}_{n+k},\mathcal{I}_{n+k}]} - \mathbf{r}^n_{[\mathcal{I}_{n+k},\mathcal{I}_{n+k}]}\right)\mathbf{J}_{[\mathcal{I}_{n+k}]} = \mathbf{0}. \tag{G.2}$$

Its solution takes the following form:

$$\mathbf{J}\left(x\right)_{[\mathcal{I}_{n+k}]} = \sum_{j=1}^{2(n-k)} w_{n+k,j}\bar{\mathbf{g}}_{n+k,j}\exp\left(\bar{\beta}_{n+k,j}x\right), \tag{G.3}$$

where $\bar{\mathbf{g}}_{n+k,j}$ and $\bar{\beta}_{n+k,j}$ are solutions to the following standard eigenvalue problem:

$$\begin{bmatrix} 0 & \mathbf{I} \\ -\left(2\mathbf{\Sigma}_X^{-1}\left(\tilde{\mathbf{\Lambda}} - \mathbf{r}^n\right)\right)_{[\mathcal{I}_{n+k},\mathcal{I}_{n+k}]} & -\left(2\mathbf{\Sigma}_X^{-1}\tilde{\theta}_X - \mathbf{I}\right)_{[\mathcal{I}_{n+k},\mathcal{I}_{n+k}]} \end{bmatrix} \begin{bmatrix} \bar{\mathbf{g}}_{n+k} \\ \bar{\mathbf{h}}_{n+k} \end{bmatrix} = \bar{\beta}_{n+k} \begin{bmatrix} \bar{\mathbf{g}}_{n+k} \\ \bar{\mathbf{h}}_{n+k} \end{bmatrix}, \tag{G.4}$$

where $\mathbf{I}$ is an $n \times n$ identity matrix, $\mathbf{r}^n$, $\tilde{\theta}_X$ and $\mathbf{\Sigma}_X$ are defined in (C.5). As in the static case, the coefficients $w_{n+k,j}$ will be different for different securities.

G.1 Debt

Let $D\left(x, s; C\right)$ be the value of corporate debt payments before restructuring occurs. These payments include the coupon payments, and the recovery value at default, if default occurs before restructuring. The intermediate cash flows before default and restructuring are the same as in the static model, and so are the payments at default.

$$F\left(X, s\right) = \left(1 - \tau_i\right)C, \tag{G.5}$$

$$H\left(X, s\right) = V_B\left(X, s\right). \tag{G.6}$$

After restructuring, the outstanding debt from previous issues gets diluted by new issues. Suppose at time 0, the state is i and cash flow equals X_0. Let the corresponding optimal coupon rate be $C\left(X_0, i\right)$. Next, consider a restructuring that occurs in state j when cash flow is $X_{\mathcal{T}_\mathcal{U}}$. Notice that $X_{\mathcal{T}_\mathcal{U}}$ does not have to be equal to X_U^j because restructuring can also be triggered by a change of state, as in the case of default. The value of old debt at a restructuring point in state j is:

$$K\left(X_{\mathcal{T}_{\mathcal{U}}},j\right)=D\left(X_{\mathcal{T}_{\mathcal{U}}},j;C(X_0,i)\right) = \frac{C(X_0,i)}{C(X_{\mathcal{T}_{\mathcal{U}}},j)}D\left(X_{\mathcal{T}_{\mathcal{U}}},j;C(X_{\mathcal{T}_{\mathcal{U}}},j)\right) = \frac{C(X_0,i)}{C\left(X_0,j\right)}D\left(X_0,j;C(X_0,i)\right), \tag{G.7}$$

where the second equality is due to newly issued debt being pari passu, and the third equality follows from the scaling property. Thus, the value of old issues does not depend on the cash flow level at the restructuring point. What matters is the state where restructuring occurs.

When $X\in\mathcal{D}_k$ $(k<n)$, for $i\in\mathcal{I}_k$,

$$\begin{aligned} r^n(i)D\left(x,i\right) &= \mathcal{A}^iD\left(x,i\right)+\tilde{\lambda}_{i,1}D\left(x,1\right)+\cdots+\tilde{\lambda}_{i,k}D\left(x,k\right) \\ &\quad +\tilde{\lambda}_{i,k+1}V_B\left(x,k+1\right)+\cdots+\tilde{\lambda}_{i,n}V_B\left(x,n\right)+\left(1-\tau_i\right)C. \end{aligned} \tag{G.8}$$

The particular solution is the same as in the static case:

$$D\left(x,i\right)=\bar{\xi}_k^D\left(i\right)e^x+\bar{\zeta}_k^D\left(i\right), \tag{G.9}$$

where $\bar{\xi}_k^D$ and $\bar{\zeta}_k^D$ are the same as in the static case.

When $X\in\mathcal{D}_n$, for $i\in\mathcal{I}_n$,

$$r^n(i)D\left(x,i\right)=\mathcal{A}^iD\left(x,i\right)+\tilde{\lambda}_{i,1}D\left(x,1\right)+\cdots+\tilde{\lambda}_{i,n}D\left(x,n\right)+\left(1-\tau_i\right)C.$$

Again, the particular solution is the same as in the static case:

$$D\left(x,i\right)=\bar{\zeta}_n^D\left(i\right), \tag{G.10}$$

where $\bar{\zeta}_n^D$ is the same as in the static case.

When $X\in\mathcal{D}_{n+k}$ $(k<n)$, for $i\in\mathcal{I}_{n+k}$,

$$\begin{aligned} r^n(i)D\left(x,i\right) &= \mathcal{A}^iD\left(x,i\right)+\tilde{\lambda}_{i,u(1)}K\left(x,u\left(1\right)\right)+\cdots+\tilde{\lambda}_{i,u(k)}K\left(x,u\left(k\right)\right) \\ &\quad +\tilde{\lambda}_{i,u(k+1)}D\left(x,u\left(k+1\right)\right)+\cdots+\tilde{\lambda}_{i,u(n)}D\left(x,u\left(n\right)\right)+\left(1-\tau_i\right)C. \end{aligned} \tag{G.11}$$

Here, the values $K\left(x,\cdot\right)$ depends on the initial value of debt. We will need to solve for that recursively. Assume these values are known for now. Guess that a particular solution is:

$$D\left(x,i\right)=\bar{\zeta}_{n+k}^D\left(i\right). \tag{G.12}$$

Then,

$$\begin{aligned} r^n(i)\bar{\zeta}_{n+k}^D\left(i\right) &= \tilde{\lambda}_{i,u(1)}\frac{C}{C\left(X_0,u\left(1\right)\right)}D\left(X_0,u\left(1\right)\right)+\cdots+\tilde{\lambda}_{i,u(k)}\frac{C}{C\left(X_0,u\left(k\right)\right)}D\left(X_0,u\left(k\right)\right) \\ &\quad +\tilde{\lambda}_{i,u(k+1)}\bar{\zeta}_{n+k}^D\left(u\left(k+1\right)\right)+\cdots+\tilde{\lambda}_{i,u(n)}\bar{\zeta}_{n+k}^D\left(u\left(n\right)\right)+\left(1-\tau_i\right)C, \end{aligned}$$

which implies:

$$\bar{\zeta}^D_{n+k}(\mathcal{I}_{n+k}) = C\left(\mathbf{r}^n - \tilde{\mathbf{\Lambda}}\right)^{-1}_{[\mathcal{I}_{n+k},\mathcal{I}_{n+k}]}\left[(1-\tau_i)\,\mathbf{1}_k + \tilde{\mathbf{\Lambda}}_{[\mathcal{I}_{n+k},\mathcal{I}^c_{n+k}]}\left[\mathbf{D}\left(X_0\right)\oslash\mathbf{C}\left(X_0\right)\right]_{[\mathcal{I}^c_{n+k}]}\right], \tag{G.13}$$

where $\oslash$ denotes element-by-element division, $\mathbf{D}(X_0) = \left[D\left(X_0,1\right),\cdots,D\left(X_0,n\right)\right]'$, and $\mathbf{C}(X_0) = \left[C\left(X_0,1\right),\cdots,C\left(X_0,n\right)\right]'$. Since $\bar{\zeta}^D_{n+k}$ depends on the initial debt value, we have to solve for the value of debt in the dynamic case recursively.

The boundary conditions are as follows. First, as in the static model, there are n conditions specifying the value of debt at the n different default boundaries:

$$D\left(X^i_D,i\right) = V_B\left(X^i_D,i\right),\quad i=1,\cdots,n. \tag{G.14}$$

Another n conditions specify the value of debt at the restructuring boundaries:

$$D\left(X^{u(i)}_U,u\left(i\right)\right) = \frac{C}{C\left(X_0,u\left(i\right)\right)}D\left(X_0,u\left(i\right)\right),\quad i=1,\cdots,n. \tag{G.15}$$

Moreover, we need to ensure that $D\left(X,i\right)$ is C^0 and C^1 at all the boundaries for which neither default or restructure has occurred.

$$\begin{aligned}\lim_{X\uparrow X^k_D} D\left(X,i\right) &= \lim_{X\downarrow X^k_D} D\left(X,i\right),\quad k=i+1,\cdots,n\\ \lim_{X\uparrow X^k_D} D_X\left(X,i\right) &= \lim_{X\downarrow X^k_D} D_X\left(X,i\right),\quad k=i+1,\cdots,n\end{aligned}$$

and

$$\begin{aligned}\lim_{X\uparrow X^{u(k)}_U} D\left(X,u\left(i\right)\right) &= \lim_{X\downarrow X^{u(k)}_U} D\left(X,u\left(i\right)\right),\quad k=1,\cdots,i-1\\ \lim_{X\uparrow X^{u(k)}_U} D_X\left(X,u\left(i\right)\right) &= \lim_{X\downarrow X^{u(k)}_U} D_X\left(X,u\left(i\right)\right),\quad k=1,\cdots,i-1\end{aligned}$$

There are $2n^2$ of unknown coefficients for $\left\{w^D\right\}$ $\left(2\left(1+\cdots+n+\cdots+1\right)\right)$. The boundary conditions combined to give $2n^2$ conditions, so we can solve for $\left\{w^D\right\}$ from a system of linear equations.

Thus, in summary,

$$\begin{aligned}\mathbf{D}\left(X\right)_{[\mathcal{I}_k]} &= \sum_{j=1}^{2k}\bar{w}^D_{k,j}\bar{\mathbf{g}}_{k,j}X^{\bar{\beta}_{k,j}} + \bar{\xi}^D_k(\mathcal{I}_k)X + \bar{\zeta}^D_k(\mathcal{I}_k),\quad X\in\mathcal{D}_k,\quad k=1,\cdots,n-1.\\ \mathbf{D}\left(X\right)_{[\mathcal{I}_n]} &= \sum_{j=1}^{2n}\bar{w}^D_{n,j}\bar{\mathbf{g}}_{n,j}X^{\bar{\beta}_{n,j}} + \bar{\zeta}^D_n(\mathcal{I}_n),\quad X\in\mathcal{D}_n \qquad\qquad (G.16)\\ \mathbf{D}\left(X\right)_{[\mathcal{I}_{n+k}]} &= \sum_{j=1}^{2(n-k)}\bar{w}^D_{n+k,j}\bar{\mathbf{g}}_{n+k,j}X^{\bar{\beta}_{n+k,j}} + \bar{\zeta}^D_{n+k}(\mathcal{I}_{n+k}),\quad X\in\mathcal{D}_{n+k},\quad k=1,\cdots,n-1.\end{aligned}$$

To compute the value of debt, start with initial guesses of debt value in all the states. Then recalculate $D(X_0, s)$ for all the states based on the formula above, and replace the initial guesses with these new debt values. Iterate till convergence.

G.2 Equity

Consider the value of equity after restructuring. Without partial loss offset and equity issuance costs, I define the "effective tax rate" for equity-holders as $\tau_{eff} = 1 - (1-\tau_d)(1-\tau_c)$. Applying the scaling property,

$$
\begin{aligned}
& E(X_0, s_0, C) = \mathbb{E}\left[\int_0^{\mathcal{T}_D \wedge \mathcal{T}_U} \frac{n_t}{n_0}(1-\tau_{eff})(X_t - C)\,dt \middle| \mathfrak{F}_0\right] \\
& + \mathbb{E}\left[1_{\{\mathcal{T}_U < \mathcal{T}_D\}} \frac{n_{\mathcal{T}_U}}{n_0} D\left(X_0, s_{\mathcal{T}_U}; C\left(X_0, s_{\mathcal{T}_U}\right)\right)\left[(1-q)\frac{X_U^{s_{\mathcal{T}_U}}}{X_0} - \frac{C}{C\left(X_0, s_{\mathcal{T}_U}\right)}\right] \middle| \mathfrak{F}_0\right] \\
& + \mathbb{E}\left[1_{\{\mathcal{T}_U < \mathcal{T}_D\}} \frac{n_{\mathcal{T}_U}}{n_0} \frac{X_U^{s_{\mathcal{T}_U}}}{X_0} E\left(X_0, s_{\mathcal{T}_U}, C\left(X_0, s_{\mathcal{T}_U}\right)\right) \middle| \mathfrak{F}_0\right]. \qquad \text{(G.17)}
\end{aligned}
$$

The first term in the equation specifies the value of dividend payments until default or restructuring. The second and third term specifies that, at a restructuring point, equity-holders receive the proceeds from new debt issuance (net of issuance costs), plus the scaled-up equity claim after restructuring. Thus, the dividend rate, default payment and restructuring payment for equity are:

$$
\begin{aligned}
F(X, s) &= (1-\tau_{eff})(X - C), && \text{(G.18)} \\
H(X, s) &= 0, && \text{(G.19)} \\
K(X, s) &= D(X_0, s; C(X_0, s))\left((1-q)\frac{X}{X_0} - \frac{C(X_0, s_0)}{C(X_0, s)}\right) + \frac{X}{X_0} E(X_0, s, C(X_0, s)) \\
&= k^0(s) + k^1(s)X. && \text{(G.20)}
\end{aligned}
$$

When $X \in \mathcal{D}_k$ $(k \le n)$, for $i \in \mathcal{I}_k$,

$$
r^n(i)E(x, i) = \mathcal{A}^i E(x, i) + \tilde{\lambda}_{i,1} E(x, 1) + \cdots + \tilde{\lambda}_{i,k} E(x, k) + (1-\tau_{eff})(e^x - C). \qquad \text{(G.21)}
$$

The particular solution is the same as in the static case:

$$
E(x, i) = \bar{\xi}_k^E(i)\, e^x + \bar{\zeta}_k^E(i). \qquad \text{(G.22)}
$$

Plug the guess into the ODE above,

$$
\begin{aligned}
r^n(i)\left(\bar{\xi}_k^E(i)\, e^x + \bar{\zeta}_k^E(i)\right) &= \bar{\xi}_k^E(i)\, \tilde{\theta}(i)\, e^x + \sum_{j=1}^{k} \tilde{\lambda}_{ij}\left(\bar{\xi}_k^E(i)\exp(x) + \bar{\zeta}_k^E(i)\right) \\
&+ (1-\tau_{eff})(e^x - C).
\end{aligned}
$$

Collecting terms leads to:

$$\begin{aligned} r^n(i)\bar{\zeta}_k^E(i) &= \sum_{j=1}^{k} \tilde{\lambda}_{ij}\bar{\zeta}_k^E(i) - (1-\tau_{eff})\,C. \\ r^n(i)\bar{\xi}_k^E(i) &= \tilde{\theta}(i)\,\bar{\xi}_k^E(i) + \sum_{j=1}^{k} \tilde{\lambda}_{j,k}\bar{\xi}_k^E(i) + (1-\tau_{eff}). \end{aligned}$$

Thus,

$$\bar{\xi}_k^E(\mathcal{I}_k) = (1-\tau_{eff})\left(\mathbf{r}^n - \tilde{\theta}_X - \tilde{\mathbf{\Lambda}}\right)^{-1}_{[\mathcal{I}_k,\mathcal{I}_k]}\mathbf{1}_k, \tag{G.23}$$

$$\bar{\zeta}_k^E(\mathcal{I}_k) = -(1-\tau_{eff})\,C\left(\mathbf{r}^n - \tilde{\mathbf{\Lambda}}\right)^{-1}_{[\mathcal{I}_k,\mathcal{I}_k]}\mathbf{1}_k. \tag{G.24}$$

When $X \in \mathcal{D}_{n+k}$ $(k < n)$, for $i \in \mathcal{I}_{n+k}$,

$$\begin{aligned} r^n(i)E(x,i) &= \mathcal{A}^i E(x,i) + \tilde{\lambda}_{i,u(1)}K(x,u(1)) + \cdots + \tilde{\lambda}_{i,u(k)}K(x,u(k)) \\ &\quad + \tilde{\lambda}_{i,u(k+1)}E(x,u(k+1)) + \cdots + \tilde{\lambda}_{i,u(n)}E(x,u(n)) + (1-\tau_{eff})(e^x - C). \end{aligned}$$

The particular solution is the same as in the static case:

$$E(x,i) = \bar{\xi}_{n+k}^E(i)\,e^x + \bar{\zeta}_{n+k}^E(i). \tag{G.25}$$

Plug the guess into the ODE above,

$$\begin{aligned} r^n(i)\left(\bar{\xi}_{n+k}^E(i)\,e^x + \bar{\zeta}_{n+k}^E(i)\right) &= \bar{\xi}_{n+k}^E(i)\,\tilde{\theta}(i)\,e^x + \sum_{j=1}^{k}\tilde{\lambda}_{i,u(j)}\left(k^0(u(j)) + k^1(u(j))\,e^x\right) \\ &\quad + \sum_{j=k+1}^{n}\tilde{\lambda}_{i,u(j)}\left(\bar{\xi}_{n+k}^E(u(j))\,e^x + \bar{\zeta}_{n+k}^E(u(j))\right) + (1-\tau_{eff})(e^x - C). \end{aligned}$$

Collecting terms leads to:

$$\begin{aligned} r^n(i)\bar{\zeta}_{n+k}^E(j) &= \sum_{j=k+1}^{n}\tilde{\lambda}_{i,u(j)}\bar{\zeta}_{n+k}^E(u(j)) + \sum_{j=1}^{k}\tilde{\lambda}_{i,u(j)}k^0(u(j)) - (1-\tau_{eff})\,C. \\ r^n(i)\bar{\xi}_{n+k}^E(i) &= \tilde{\theta}(i)\,\bar{\xi}_{n+k}^E(i) + \sum_{j=k+1}^{n}\tilde{\lambda}_{i,u(j)}\bar{\xi}_{n+k}^E(j) + \sum_{j=1}^{k}\tilde{\lambda}_{i,u(j)}k^1(u(j)) + (1-\tau_{eff}). \end{aligned}$$

Thus,

$$\bar{\xi}_{n+k}^E(\mathcal{I}_{n+k}) = \left(\mathbf{r}^n - \tilde{\theta}_X - \tilde{\mathbf{\Lambda}}\right)^{-1}_{[\mathcal{I}_{n+k},\mathcal{I}_{n+k}]}\left[(1-\tau_{eff})\,\mathbf{1}_k + \tilde{\mathbf{\Lambda}}_{[\mathcal{I}_{n+k},\mathcal{I}^c_{n+k}]}\mathbf{k}^1_{[\mathcal{I}^c_{n+k}]}\right], \tag{G.26}$$

$$\bar{\zeta}_{n+k}^E(\mathcal{I}_{n+k}) = \left(\mathbf{r}^n - \tilde{\mathbf{\Lambda}}\right)^{-1}_{[\mathcal{I}_k,\mathcal{I}_k]}\left[\tilde{\mathbf{\Lambda}}_{[\mathcal{I}_{n+k},\mathcal{I}^c_{n+k}]}\mathbf{k}^0_{[\mathcal{I}^c_{n+k}]} - (1-\tau_{eff})\,C\mathbf{1}_k\right]. \tag{G.27}$$

The boundary conditions for E are similar to those for debt. First, as in the static model, there are n conditions specifying the value of debt at the n different default boundaries:

$$E\left(X_D^i, i\right) = 0, \quad i = 1, \cdots, n. \tag{G.28}$$

Another n conditions specify the value of debt at the restructuring boundaries, for $i = 1, \cdots, n$,

$$\begin{aligned} E\left(X_U^{u(i)}, u\left(i\right)\right) &= D\left(X_0, u\left(i\right); C\left(X_0, u\left(i\right)\right)\right)\left(\left(1-q\right)\frac{X_U^{u(i)}}{X_0} - \frac{C\left(X_0, s_0\right)}{C\left(X_0, u\left(i\right)\right)}\right) \\ &\quad + \frac{X_U^{u(i)}}{X_0} E\left(X_0, u\left(i\right); C\left(X_0, u\left(i\right)\right)\right). \end{aligned} \tag{G.29}$$

Finally, we need to ensure that $E\left(X, i\right)$ is C^0 and C^1 , which lead to an identical set of conditions as for D. These boundary conditions help determine the coefficients $\left\{w_{k,j}^E\right\}$.

Define $\mathbf{E}(X) = \left[E\left(X, 1\right), \cdots, E\left(X, n\right)\right]'$. Then, in summary,

$$\begin{aligned} \mathbf{E}\left(X\right)_{[\mathcal{I}_k]} &= \sum_{j=1}^{2k} \bar{w}_{k,j}^E \bar{\mathbf{g}}_{k,j} X^{\bar{\beta}_{k,j}} + \bar{\xi}_k^E(\mathcal{I}_k) X + \bar{\zeta}_k^E(\mathcal{I}_k), \quad X \in \mathcal{D}_k, \quad k = 1, \cdots, n. \\ \mathbf{E}\left(X\right)_{[\mathcal{I}_{n+k}]} &= \sum_{j=1}^{2(n-k)} \bar{w}_{n+k,j}^E \bar{\mathbf{g}}_{n+k,j} X^{\bar{\beta}_{n+k,j}} + \bar{\xi}_{n+k}^E(\mathcal{I}_{n+k}) X + \bar{\zeta}_{n+k}^E(\mathcal{I}_{n+k}), \quad X \in \mathcal{D}_{n+k}, \quad k = 1, \cdots, n-1. \end{aligned}$$

■

H Smooth-Pasting Conditions

H.1 Static Model

Since formulae for equity values are available in close-form, we can evaluate the smooth-pasting conditions directly.

$$E\left(X, i\right) = \left(1 - \tau_d\right)\left(1 - \tau_c^+\right) E^+\left(X, i\right) + \frac{1 - \tau_c^-}{1 - e} E^-\left(X, i\right),$$

thus the n smooth pasting conditions are:

$$\begin{aligned} \left.\frac{\partial}{\partial X} E\left(X, k\right)\right|_{X \downarrow X_D^k} &= \left(1 - \tau_d\right)\left(1 - \tau_c^+\right) \left.\frac{\partial}{\partial X} E^+\left(X, k\right)\right|_{X \downarrow X_D^k} + \frac{1 - \tau_c^-}{1 - e} \left.\frac{\partial}{\partial X} E^-\left(X, k\right)\right|_{X \downarrow X_D^k} \\ &= 0, \quad k = 1, \cdots, n \end{aligned}$$

$$
\begin{aligned}
\left.\frac{\partial}{\partial X}E^{+}(X,k)\right|_{X\downarrow X_D^k} &= \sum_{j=1}^{2k} w_{k,j}^{E^+}\mathbf{g}_{k,j}(k)\,\beta_{k,j}\left(X_D^k\right)^{\beta_{k,j}-1},\\
\left.\frac{\partial}{\partial X}E^{-}(X,k)\right|_{X\downarrow X_D^k} &= \sum_{j=1}^{2k} w_{k,j}^{E^-}\mathbf{g}_{k,j}(k)\,\beta_{k,j}\left(X_D^k\right)^{\beta_{k,j}-1}+\xi_k^{E^-}(k).
\end{aligned}
$$

Thus,

$$
\begin{aligned}
0 &= (1-\tau_d)\left(1-\tau_c^+\right)\left(\sum_{j=1}^{2k} w_{k,j}^{E^+}\mathbf{g}_{k,j}(k)\,\beta_{k,j}\left(X_D^k\right)^{\beta_{k,j}-1}\right)\\
&\quad+\frac{1-\tau_c^-}{1-e}\left(\sum_{j=1}^{2k} w_{k,j}^{E^-}\mathbf{g}_{k,j}(k)\,\beta_{k,j}\left(X_D^k\right)^{\beta_{k,j}-1}+\xi_k^{E^-}(k)\right)
\end{aligned}
$$

or

$$
\sum_{j=1}^{2k}\left(\frac{\left(1-\tau_c^+\right)(1-\tau_d)(1-e)}{1-\tau_c^-}w_{k,j}^{E^+}+w_{k,j}^{E^-}\right)\mathbf{g}_{k,j}(k)\,\beta_{k,j}\left(X_D^k\right)^{\beta_{k,j}-1}+\xi_k^{E^-}(k)=0,\quad k=1,\cdots,n. \tag{H.1}
$$

Be aware that w also depend on the n default boundaries. So, we need to solve these n nonlinear equations simultaneously to get the optimal default boundaries.

H.2 Dynamic Model

For default boundaries, for $k=1,\cdots,n$,

$$
\left.\frac{\partial}{\partial X}E(X,k,C)\right|_{X\downarrow X_D^k}=\sum_{j=1}^{2k}\bar{w}_{k,j}^{E}\bar{\mathbf{g}}_{k,j}(k)\,\bar{\beta}_{k,j}\left(X_D^k\right)^{\bar{\beta}_{k,j}-1}+\bar{\xi}_k^{E}(k)=0. \tag{H.2}
$$

For restructuring boundaries, for $k=1,\cdots,n$,

$$
\begin{aligned}
&\left.\frac{\partial}{\partial X}E(X,u(k),C)\right|_{X\uparrow X_U^{u(k)}}\\
&= \sum_{j=1}^{2(n-k+1)} \bar{w}_{n+k-1,j}^{E}\bar{\mathbf{g}}_{n+k-1,j}(u(k))\,\bar{\beta}_{n+k-1,j}\left(X_U^{u(k)}\right)^{\bar{\beta}_{n+k-1,j}-1}+\bar{\xi}_{n+k-1}^{E}(u(k))\\
&= \frac{\partial}{\partial X_U^{u(k)}}\left[\begin{array}{c} D\left(X_0,u(k),C\left(X_0,u(k)\right)\right)\left((1-q)\frac{X_U^{u(k)}}{X_0}-\frac{C(X_0,s_0)}{C(X_0,u(k))}\right)\\ +\frac{X_U^{u(k)}}{X_0}E\left(X_0,u(k),C\left(X_0,u(k)\right)\right)\end{array}\right].
\end{aligned} \tag{H.3}
$$

The partial derivative following the last equality needs to be evaluated numerically.

I Returns on Equity and Debt

To compare the pricing implications of the Markov chain model with that of Bansal and Yaron (2004), I consider a real dividend stream, which is a levered up version of aggregate consumption (as in BY).

$$\frac{dD_t}{D_t} = \theta_D(s_t)\,dt + \sigma_{D,m}(s_t)\,dW_t^m, \tag{I.1}$$

with

$$\begin{aligned} \theta_D(s) &= \overline{\theta}_m + \phi\left(\theta_m(s_t) - \overline{\theta}_m\right), & \text{(I.2)} \\ \sigma_{D,m}(s) &= \varphi_d \sigma_m(s_t). & \text{(I.3)} \end{aligned}$$

Thus, the dividend stream has the same expected growth rate as aggregate consumption. Denote the (real) value of the stock as S, which will be a function of current dividend and state, $S(D_t, s_t)$. Its value can be derived using the same method in Proposition 2, but we have to use the real stochastic discount factor m_t instead of n_t. Ignoring taxes,

$$S(D, s) = Dv^D(s), \tag{I.4}$$

where $v^D(s)$ is the pricie-dividend ratio in state s, which is given in a vector,

$$\mathbf{v}^D = \left(\mathbf{r} - \tilde{\theta}_D - \tilde{\mathbf{\Lambda}}\right)^{-1}\mathbf{1}, \tag{I.5}$$

where $\tilde{\mathbf{\Lambda}}$ is again the generator matrix under risk-neutral measure $\mathcal{Q}$, $\mathbf{r} \triangleq \mathbf{diag}([r(1), \cdots, r(n)])$, $\tilde{\theta}_D \triangleq \mathbf{diag}\left(\left[\tilde{\theta}_D(1), \cdots, \tilde{\theta}_D(n)\right]\right)$, with the risk-neutral growth rates defined as:

$$\tilde{\theta}_D(s) = \theta_D(s) - \sigma_{D,m}(s)\,\eta(s). \tag{I.6}$$

Then, in state i,

$$\frac{dS}{S} = o(dt) + \sigma_{D,m}(i)\,dW_t^m + \sum_{j \neq i}\left(\frac{v^D(j)}{v^D(i)} - 1\right)dN^{(i,j)}. \tag{I.7}$$

The risk premium for S, μ_S, is determined by its covariance with the discount factor. Thus, in state i,

$$\begin{aligned} \mu_S(i) &= -\frac{1}{dt}cov_t\left(\frac{dS}{S}, \frac{dm}{m}\right) \\ &= \sigma_{D,m}(i)\,\eta(i) - \sum_{j \neq i}\lambda_{ij}\left(\frac{v^D(j)}{v^D(i)} - 1\right)\left(e^{\kappa(i,j)} - 1\right). & \text{(I.8)} \end{aligned}$$

The first term for the risk premium comes from the risk exposure to small shocks. If the small shocks tend to move the stock price and stochastic discount factor in opposite directions (e.g, price drops as marginal utility rises), then the stock is risky and demand a positive premium.

The same intuition applies to the second term of the risk premium, which comes from the exposure to large shocks.

The total volatility of return consists of two parts, volatility due to Brownian motion and jumps:

$$\sigma_R(i) = \sqrt{\sigma_{D,m}^2(i) + \sum_{j \neq i} \lambda_{ij} \left(\frac{v^D(j)}{v^D(i)} - 1 \right)^2}. \tag{I.9}$$

Next, I calculate the risk premium for the equity and debt of levered firms. Levered equity can be regarded as a portfolio of two claims, $E^+(X,s)$ and $E^-(X,s)$. From equation (27), we have:

$$\frac{dE}{E} = (1-\tau_d)\left(1-\tau_c^+\right) \frac{E^+}{E} \frac{dE^+}{E^+} - \frac{1-\tau_c^-}{1-e} \frac{E^-}{E} \frac{dE^-}{E^-}. \tag{I.10}$$

Denote the risk premium for E^+ as μ_{E^+}. Applying Ito's formula with jumps, we have:

$$\begin{aligned}
\mu_{E^+}(X,s) &= -\frac{1}{dt} cov_t \left(\frac{dE^+(X,s)}{E^+(X,s)}, \frac{dn_t}{n_t} \right) \\
&= \frac{\sum_{j=1}^n w_{n+1,j}^{E^+} g_{n,j}(s) \beta_{n,j} X^{\beta_{n,j}} + Xv(s)}{\sum_{j=1}^n w_{n+1,j}^{E^+} g_{n,j}(s) X^{\beta_{n,j}} + Xv(s) - Cb(s)} \left(\sigma_{X,m}(s) \eta^m(s) + \sigma_{P,2} \eta^P \right) \\
&\quad - \sum_{s' \neq s} \lambda_{ss'} \left(\frac{E^+(X,s')}{E^+(X,s)} - 1 \right) \left(e^{\kappa(s,s')} - 1 \right).
\end{aligned} \tag{I.11}$$

Similarly, let the instantaneous expected excess return of E^- be μ_{E^-}. Then,

$$\begin{aligned}
\mu_{E^-}(X,s) &= \frac{\sum_{j=1}^n w_{n+1,j}^{E^-} g_{n,j}(s) \beta_{n,j} X^{\beta_{n,j}}}{\sum_{j=1}^n w_{n+1,j}^{E^-} g_{n,j}(s) X^{\beta_{n,j}}} \left(\sigma_{X,m}(s) \eta^m(s) + \sigma_{P,2} \eta^P \right) \\
&\quad - \sum_{s' \neq s} \lambda_{ss'} \left(\frac{E^-(X,s')}{E^-(X,s)} - 1 \right) \left(e^{\kappa(s,s')} - 1 \right).
\end{aligned} \tag{I.12}$$

Thus, the conditional risk premium for levered equity is:

$$\mu_E(X,s) = (1-\tau_d)\left(1-\tau_c^+\right) \frac{E^+(X,s)}{E(X,s)} \mu_{E^+}(X,s) - \frac{1-\tau_c^-}{1-e} \frac{E^-(X,s)}{E(X,s)} \mu_{E^-}(X,s). \tag{I.13}$$

Similarly, the conditional risk premium for the corporate bond is

$$\begin{aligned}
\mu_D(X,s) &= \frac{\sum_j w_{k,j}^D g_{k,j}(s) \beta_{k,j} X^{\beta_{k,j}}}{\sum_j w_{k,j}^D g_{k,j}(s) X^{\beta_{k,j}} + (1-\tau_i) Cb(s)} \left(\sigma_{X,m}(s) \eta^m(s) + \sigma_{P,2} \eta^P \right) \\
&\quad - \sum_{s' \neq s} \lambda_{ss'} \left(\frac{D(X,s')}{D(X,s)} - 1 \right) \left(e^{\kappa(s,s')} - 1 \right).
\end{aligned} \tag{I.14}$$

Finally, define the total value of a levered firm as the sum of its debt and equity,

$$A(X,s) = D(X,s) + E(X,s), \tag{I.15}$$

then:

$$\mu_A(X,s)=\frac{D(X,s)}{A(X,s)}\mu_D(X,s)+\frac{E(X,s)}{A(X,s)}\mu_E(X,s). \tag{I.16}$$

J Comments on the Numerical Procedure

We need to solve for the coefficients $w_{k,j}^{()}$ numerically. For debt, $\mathbf{w}^D$ is $n^2\times 1$. It is determined by a system of linear equations, arising from the boundary conditions. The boundary conditions can be translated into three groups of linear equations in $\mathbf{w}^D$. The first n equations specify the value of debt at the n default boundaries. The the next $n(n-1)$ equations specify the continuity of D and its first derivative at the boundaries where it is defined. More specifically, $D(X,i)$ is defined over the region $\left(X_D^i,+\infty\right)$. So, it must be continuous and have continuous first derivative at $X_D^{i+1},X_D^{i+2},\cdots,X_D^n$. The total number of equations here is:

$$2\left[(n-1)+(n-2)+\cdots+1\right]=n(n-1).$$

For equity, $\mathbf{w}^{E^+}$ and $\mathbf{w}^{E^-}$ are both $\left(n^2+2n\right)\times 1$. The extra coefficients are due to the additional region $[X_D^n,C)$, where there are $2n$ terms. The boundary conditions are largely identical to those for debt. The only difference comes from the continuity of E and its first derivative at $X=C$, which applies to all $E(X,i)$, hence adding $2n$ equations.

I first stack up all the coefficients $\mathbf{w}$ into a vector, in the order of regions ($\mathcal{D}_1,\mathcal{D}_2,\cdots\mathcal{D}_n$ and $\mathcal{D}_{n+1}$). Then, the boundary conditions can be summarized by a system of the form:

$$\mathbf{Aw}=\mathbf{a},$$

where $\mathbf{A}$ is either $n^2\times n^2$ (for claims whose payoffs do not depend on whether $X>C$) or $\left(n^2+2n\right)\times\left(n^2+2n\right)$ (for equity).

To illustrate the structure of this system, consider the case when $n=2$. In this case we have two default boundaries, X_D^1 and X_D^2. The first 2 boundary conditions for debt are:

$$\begin{aligned}
\sum_{j=1}^{2} w_{1,j}^D g_{1,j}(1)\left(X_D^1\right)^{\beta_{1,j}}+\xi_1(1)X_D^1+\zeta_1(1) &= \left(1-\tau_c^+\right)\alpha(1)X_D^1 v(1),\\
\sum_{j=1}^{2} w_{2,j}^D g_{2,j}(2)\left(X_D^2\right)^{\beta_{2,j}}+(1-\tau_i)Cb(2) &= \left(1-\tau_c^+\right)\alpha(2)X_D^2 v(2).
\end{aligned}$$

The next 2 boundary conditions ensure the continuity of $D(X,1)$ and its first derivative at X_D^2:

$$\begin{aligned}
\sum_{j=1}^{2} w_{1,j}^D g_{1,j}(1)\left(X_D^2\right)^{\beta_{1,j}}+\xi_1(1)X_D^2+\zeta_1(1) &= \sum_{j=1}^{2} w_{2,j}^D g_{2,j}(1)\left(X_D^2\right)^{\beta_{2,j}}+(1-\tau_i)Cb(1),\\
\sum_{j=1}^{2} w_{1,j}^D g_{1,j}(1)\beta_{1,j}\left(X_D^2\right)^{\beta_{1,j}-1}+\xi_1(1) &= \sum_{j=1}^{2} w_{2,j}^D g_{2,j}(1)\beta_{2,j}\left(X_D^2\right)^{\beta_{2,j}-1}.
\end{aligned}$$

Thus, define $\mathbf{w}, \mathbf{a}, \mathbf{A}$ as

$$\mathbf{w} = \begin{bmatrix} w_{1,1}^D \\ w_{1,2}^D \\ w_{2,1}^D \\ w_{2,2}^D \end{bmatrix}, \quad \mathbf{a} = \begin{bmatrix} (1-\tau_c^+)\,\alpha(1)\,X_D^1 v(1) - \left(\xi_1(1)\,X_D^1 + \zeta_1(1)\right) \\ (1-\tau_c^+)\,\alpha(2)\,X_D^2 v(2) - (1-\tau_i)\,Cb(2) \\ (1-\tau_i)\,Cb(1) - \left(\xi_1(1)\,X_D^2 + \zeta_1(1)\right) \\ -\xi_1(1)\,X_D^2 \end{bmatrix}, \tag{J.1}$$

and

$$\mathbf{A} = \begin{bmatrix} g_{1,1}(1)\left(X_D^1\right)^{\beta_{1,1}} & g_{1,2}(1)\left(X_D^1\right)^{\beta_{1,2}} & 0 & 0 \\ 0 & 0 & g_{2,1}(2)\left(X_D^2\right)^{\beta_{2,1}} & g_{2,2}(2)\left(X_D^2\right)^{\beta_{2,2}} \\ g_{1,1}(1)\left(X_D^2\right)^{\beta_{1,1}} & g_{1,2}(1)\left(X_D^2\right)^{\beta_{1,2}} & -g_{2,1}(1)\left(X_D^2\right)^{\beta_{2,1}} & -g_{2,2}(1)\left(X_D^2\right)^{\beta_{2,2}} \\ g_{1,1}(1)\,\beta_{1,1}\left(X_D^2\right)^{\beta_{1,1}} & g_{1,2}(1)\,\beta_{1,2}\left(X_D^2\right)^{\beta_{1,2}} & -g_{2,1}(1)\,\beta_{2,1}\left(X_D^2\right)^{\beta_{2,1}} & -g_{2,2}(1)\,\beta_{2,2}\left(X_D^2\right)^{\beta_{2,2}} \end{bmatrix}. \tag{J.2}$$

As n increases, the dimension of the system quickly explodes. In practice, the matrix $\mathbf{A}$ is often ill-conditioned, especially when n gets large. Thus, it will be quite helpful if we can reduce the dimension of the system. This can be achieved by exploring the structure more carefully. In addition, proper scaling of the value of initial cash flow X_0 also helps with the ill-conditioning problem.

K Details of the Calibration

The Markov chain for the expected growth rate and volatility of aggregate consumption is calibrated using a two-step procedure. Start with the discrete-time system of consumption dynamics of Bansal and Yaron (2004) (BY):

$$\begin{aligned} g_{t+1} &= \mu_c + x_t + \sqrt{v_t}\eta_{t+1}, \\ x_{t+1} &= \kappa_x x_t + \sigma_x \sqrt{v_t} e_{t+1}, \\ v_{t+1} &= \bar{v} + \kappa_v\left(v_t - \bar{v}\right) + \sigma_v w_{t+1}, \end{aligned}$$

where g is log consumption growth. I use the parameters from BY, which are at monthly frequency and calibrated to the annual consumption data from 1929 to 1998.

An important restriction of this system is that shocks to consumption, η_{t+1}, and shocks to the conditional moments, e_{t+1}, w_{t+1}, are mutually independent. This restriction allows me to approximate the dynamics of (x, v) with a Markov chain. I first obtain a discrete-time Markov chain using the quadrature method of Tauchen and Hussey (1991). For numerical reasons, I choose a small number of states ($n = 9$) for the Markov chain, with three different values for v, and three values for x for each v.

To build the grid for the Markov chain, I first pick 3 grid points for v corresponding to the

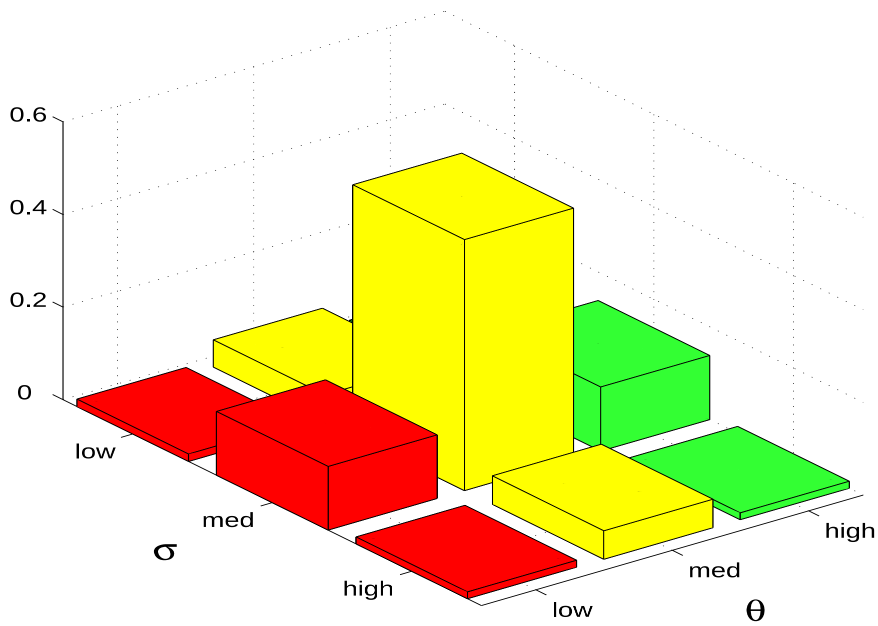

Figure 10: Stationary Distribution of the Markov Chain

Gauss-Hermite quadrature nodes for the unconditional distribution of v:

$$v \sim N\left(\bar{v}, \frac{\sigma_v^2}{1-\kappa_v^2}\right).$$

Next, fixing v, the distribution of x is again normal. So, for each chosen v, I pick 3 values for x using the quadrature nodes for the distribution of x with given v. These grid points lead to 9 states for (θ_m, σ_m).

Given the grid, I calculate the transition probabilities of the Markov chain at quarterly frequency using the method of Tauchen and Hussey (1991), which is an approximation of the conditional density of the continuous-state process. Finally, I convert the grid for (x, v) into a grid for (θ_m, σ_m) as in equation (5), and transform the discrete-time transition matrix into the generator of a continuous-time Markov chain by assuming that the probability of more than one change of state is close to zero within a quarter. The formulas for this transformation are in Appendix A of Jarrow, Lando, and Turnbull (1997). With just 9 states, the grid points are relatively far away from each other. I compute the discrete time Markov chain at the quarterly frequency so that the transition probabilities are not too small, and the assumption of no more than one jump within the period is reasonable. Under my calibration, the economy spends about 54% of the time in the "center" state (with median expected growth rate and volatility). More information about the stationary distribution of the Markov chain is in Figure 10.

Table 9 Panel A shows the parameters for the discrete time consumption model of BY; Panel

Table 9: MARKOV CHAIN APPROXIMATION OF THE BY MODEL

Panel A: Paramters for the BY Model

μ_c	μ_d	ϕ	σ_d	κ_x	σ_x	κ_v	$\bar{v}$	σ_v
0.0015	0.0015	3	4.5	0.979	0.044	0.987	6.08×10^{-5}	0.23×10^{-5}

Panel B: Properties of Annualized Time-Averaged Growth Rates

	Data		BY			Markov Chain		
Variable	Estimate	SE	5%	50%	95%	5%	50%	95%
$\mu(g)$	1.80	-	0.59	1.79	2.99	0.74	1.81	2.90
$\sigma(g)$	2.93	(0.69)	2.26	2.79	3.44	2.19	2.64	3.12
AC(1)	0.49	(0.14)	0.25	0.46	0.63	0.23	0.42	0.58
AC(2)	0.15	(0.22)	-0.04	0.22	0.45	-0.05	0.18	0.39
AC(5)	-0.08	(0.10)	-0.17	0.06	0.30	-0.17	0.05	0.28
AC(10)	0.05	(0.09)	-0.24	-0.03	0.20	-0.24	-0.02	0.19
VR(2)	1.61	(0.34)	1.25	1.46	1.63	1.23	1.42	1.58
VR(5)	2.01	(1.23)	1.36	2.13	2.91	1.33	2.01	2.69
VR(10)	1.57	(2.07)	1.20	2.47	4.21	1.15	2.33	3.85

NOTE: Parameters in Panel A are from the discrete time model of BY (Table IV). In Panel B, the statistics of the data are from BY (2004) (Table I), based on annual observations from 1929 to 1998. The statistics for the two models are based on 5,000 simulations, each with 70 years of data. The simulations are done at high frequency and then aggregated to get annual growth rates. The symbols $\mu(g)$ and $\sigma(g)$ are mean and standard deviation of growth rates; $AC(j)$ is the jth autocorrelation; $VR(j)$ is the jth variance ratio.

B compares the statistical properties of consumption growth rates in the data with those of the simulated data from the BY model and the Markov chain model. With just 9 states, the Markov chain approximation does a good in matching the mean, volatility, autocorrelation and variance ratio of consumption growth in the BY model. The noticeable differences are that the Markov chain appears to generate a distribution of volatility and variance ratios with lighter right tail, which is likely due to the non-extreme grid points.

References

Acharya, V. V., S. T. Bharath, and A. Srinivasan, 2006, "Does Industry-Wide Distress Affect Defaulted Firms? - Evidence from Creditor Recoveries," *forthcoming, Journal of Financial Economics.*

Allen, L., and A. Saunders, 2004, "Incorporating Systemic Influences Into Risk Measurements: A Survey of the Literature," *Journal of Financial Services Research*, 26, 161–191.

Almeida, H., and T. Philippon, 2006, "The Risk-Adjusted Cost of Financial Distress," *Working Paper, New York University.*

Altinkihc, O., and R. S. Hansen, 2000, "Are There Economies of Scale in Underwriting Fees? Evidence of Rising External Financing Costs," *Review of Financial Studies*, 13, 191–218.

Altman, E., B. Brady, A. Resti, and A. Sironi, 2005, "The Link Between Default and Recovery Rates: Theory, Empirical Evidence and Implications," *Journal of Business*, 78, 2203–2227.

Altman, E., and B. Pasternack, 2006, "Defaults and Returns in the High Yield Bond Market: The Year 2005 in Review and Market Outlook," *Journal of Applied Research in Accounting and Finance*, 1, 3–29.

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

Andrade, G., and S. N. Kaplan, 1998, "How Costly is Financial (Not Economic) Distress? Evidence from Highly Leveraged Transactions That Became Distressed," *Journal of Finance*, 53, 1443–1493.

Bansal, R., R. F. Dittmar, and C. Lundblad, 2005, "Consumption, Dividends, and the Cross-Section of Equity Returns," *Journal of Finance*, 60, 1639–1672.

Bansal, R., and A. Yaron, 2004, "Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles," *Journal of Finance*, 59, 1481 – 1509.

Barlow, M. T., L. C. G. Rogers, and D. Williams, 1980, *Wiener-Hopf Factorization for Matrices, in Séminaire de Probabilités XIV, Lecture Notes in Math. 784* . pp. 324–331, Springer-Verlag, Berlin.

Bernanke, B. S., M. L. Gertler, and S. Gilchrist, 1996, "The Financial Accelerator and the Flight to Quality," *Review of Economics and Statistics*, 78, 1–15.

Borodin, A. N., and P. Salminen, 2002, *Handbook of Brownian Motion: Facts and Formulae*, Springer Verlag.

Brennan, M. J., and E. S. Schwartz, 1978, "Corporate Income Taxes, Valuation, and the Problem of Optimal Capital Structure," *Journal of Business*, 51, 103–114.

Campbell, J. Y., 1993, “Intertemporal Asset Pricing Without Consumption Data,” *American Economic Review*, 83, 487–512.

Campbell, J. Y., and J. H. Cochrane, 1999, “By Force of Habit: A Consumption Based Explanation of Aggregate Stock Market Behavior,” *Journal of Political Economy*, 107, 205–251.

Chen, H., 2006, “Asset Pricing with Recursive Preferences: Analytical Results,” *Working Paper, University of Chicago.*

Chen, L., P. Collin-Dufresne, and R. S. Goldstein, 2006, “On the Relation Between the Credit Spread Puzzle and the Equity Premium Puzzle,” *Working Paper, Michigan State University, U.C. Berkeley, and University of Minnesota.*

Cochrane, J. H., 2006, “Financial Markets and the Real Economy,” *Working Paper, University of Chicago.*

Collin-Dufresne, P., and R. S. Goldstein, 2001, “Do Credit Spreads Reflect Stationary Leverage Ratios?,” *Journal of Finance*, 56, 1929–1957.

Collin-Dufresne, P., R. S. Goldstein, and J. S. Martin, 2001, “The Determinants of Credit Spread Changes,” *Journal of Finance*, 56, 2177–2207.

David, A., 2006, “Inflation Uncertainty, Asset Valuations, and the Credit Spreads Puzzle,” *Working paper, University of Calgary.*

Davydenko, S. A., 2005, “When Do Firms Default? A Study of the Default Boundary,” *Working Paper, University of Toronto.*

Duffee, G. R., 1998, “The Relation Between Treasury Yields and Corporate Bond Yield Spreads,” *Journal of Finance*, 53, 2225–2241.

Duffie, D., 2001, *Dynamic Asset Pricing Theory*, Princeton University Press, Princeton, 3rd edn.

Duffie, D., and L. G. Epstein, 1992a, “Asset Pricing with Stochastic Differential Utility,” *Review of Financial Studies*, 5, 411–436.

Duffie, D., and L. G. Epstein, 1992b, “Stochastic Differential Utility,” *Econometrica*, 60, 353–394.

Duffie, D., and D. Lando, 2001, “Term Structures of Credit Spreads with Incomplete Accounting Information,” *Econometrica*, 69, 633–664.

Duffie, D., M. Schroder, and C. Skiadas, 1997, “A Term Structure Model with Preferences for the Timing of Resolution of Uncertainty,” *Economic Theory*, 9, 3–22.

Duffie, D., and C. Skiadas, 1994, “Continuous-Time Security Pricing: A Utility Gradient Approach,” *Journal of Mathematical Economics*, 23, 107–132.

Elton, E. J., M. J. Gruber, D. Agrawal, and C. Mann, 2001, "Explaining the Rate Spread on Corporate Bonds," *Journal of Finance*, 56, 247–277.

Epstein, L., and S. Zin, 1989, "Substitution, Risk Aversion, and the Temporal Behavior of Consumption Growth and Asset Returns I: A Theoretical Framework," *Econometrica*, 57, 937–969.

Fischer, E. O., R. Heinkel, and J. Zechner, 1989, "Dynamic Capital Structure Choice: Theory and Tests," *Journal of Finance*, 44, 19–40.

Geske, R., 1977, "The Valuation of Corporate Liabilities as Compound Options," *Journal of Financial and Quantitative Analysis*, 12, 541–552.

Gilson, S. C., 1997, "Transactions Costs and Capital Structure Choice: Evidence from Financially Distressed Firms," *Journal of Finance*, 52, 161–196.

Goldstein, R., N. Ju, and H. Leland, 2001, "An EBIT-Based Model of Dynamic Capital Structure," *Journal of Business*, 74, 483–512.

Graham, J. R., 2000, "How Big Are the Tax Benefits of Debt?," *Journal of Finance*, 55, 1901–1941.

Hackbarth, D., J. Miao, and E. Morellec, 2006, "Capital Structure, Credit Risk, and Macroeconomic Conditions," *forthcoming in Journal of Financial Economics.*

Hansen, L. P., J. C. Heaton, and N. Li, 2005, "Consumption Strikes Back? Measuring Long Run Risk," *Working Paper, University of Chicago.*

Hennessy, C. A., and T. M. Whited, 2005, "Debt Dynamics," *Journal of Finance*, 60, 1129–1165.

Huang, J., and M. Huang, 2003, "How Much of the Corporate-Treasury Yield Spread is Due to Credit Risk?," *Working paper, Penn State University and Stanford University.*

Jarrow, R. A., D. Lando, and S. M. Turnbull, 1997, "A Markov Model for the Term Structure of Credit Risk Spreads," *Review of Financial Studies*, 10, 481–523.

Jobert, A., and L. C. G. Rogers, 2006, "Option Pricing with Markov-Modulated Dynamics," *SIAM Journal on Control and Optimization*, 44, 2063–2078.

Ju, N., R. Parrino, A. Poteshman, and M. Weisbach, 2005, "Horses and Rabbits? Trade-Off Theory and Optimal Capital Structure," *Journal of Financial and Quantitative Analysis*, 40, 259–281.

Kane, A., A. J. Marcus, and R. L. McDonald, 1985, "Debt Policy and the Rate of Return Premium to Leverage," *Journal of Financial and Quantitative Analysis*, 20, 479–499.

Karatzas, I., and S. E. Shreve, 1991, *Brownian Motion and Stochastic Calculus*, Springer-Verlag, New York, 2nd edn.

Kennedy, J. E., and D. Williams, 1990, “Probabilistic Factorization of a Quadratic Matrix Polynomial,” *Math. Proc. Cambridge Philos. Soc.*, 107, 591–600.

Kreps, D. M., and E. L. Porteus, 1978, “Temporal Resolution of Uncertainty and Dynamic Choice Theory,” *Econometrica*, 46, 185–200.

Leary, M. T., and M. R. Roberts, 2005, “Do Firms Rebalance Their Capital Structures?,” *Journal of Finance*, 60, 2575–2619.

Leland, H., 1994, “Corporate Debt Value, bond covenants, and Optimal Capital Structure,” *Journal of Finance*, 49, 1213–1252.

Leland, H., 1998, “Agency Costs, Risk Management, and Capital Structure,” *Journal of Finance*, 53, 1213–1243.

Leland, H., and K. B. Toft, 1996, “Optimal Capital Structure, Endogenous Bankruptcy, and the Term Structure of Credit Spreads,” *Journal of Finance*, 51, 987–1019.

Mella-Barral, P., and W. Perraudin, 1997, “Strategic Debt Service,” *Journal of Finance*, 52, 531–556.

Miller, M. H., 1977, “Debt and Taxes,” *Journal of Finance*, 32, 261–275.

Pástor, L., and P. Veronesi, 2005, “Rational IPO Waves,” *Journal of Finance*, 60, 1713–1757.

Piazzesi, M., and M. Schneider, 2006, “Equilibrium Yield Curves,” *NBER Macroeconomics Annual.*

Shleifer, A., and R. W. Vishny, 1992, “Liquidation Values and Debt Capacity: A Market Equilibrium Approach,” *Journal of Finance*, 47, 1343–1366.

Strebulaev, I. A., 2006, “Do Tests of Capital Structure Theory Mean What They Say?,” *forthcoming, Journal of Finance.*

Tauchen, G., and R. Hussey, 1991, “Quadrature Based Methods for Obtaining Approximate Solutions to Nonlinear Asset Pricing Models.,” *Econometrica*, 59, 371–396.

Veronesi, P., 2000, “How Does Information Quality Affect Stock Returns?,” *Journal of Finance*, 55, 807–837.

Weil, P., 1990, “Non-Expected Utility in Macroeconomics,” *Quarterly Journal of Economics*, 105, 29–42.

Welch, I., 2004, “Capital Structure and Stock Returns,” *Journal of Political Economy*, 112, 106–131.