Modeling Expected Return on Defaultable Bonds

Fan Yu*
University of California, Irvine

First Draft: October 2001
Current Version: April 4, 2002

Forthcoming in the *Journal of Fixed Income*

---

*Assistant Professor of Finance, Graduate School of Management, University of California at Irvine, Irvine, CA 92697-3125, Phone: (949)824-8796, Fax: (949)725-2875, E-mail: fanyu@uci.edu. This paper can be downloaded at http://web.gsm.uci.edu/~fanyu. Helpful comments from Magnus Dahlquist, Darrell Duffie, Philippe Jorion, Haitao Li, and participants at the 12th Annual Derivative Securities Conference, the International Credit Risk Conference at HEC Montreal, and the Credit Risk Conference at the Stockholm Institute for Financial Research are gratefully acknowledged. All errors are my own.

Modeling Expected Return on Defaultable Bonds

Abstract

The existing literature on credit risk focuses on fitting bond prices and explaining yield spreads, while largely skirting the issue of expected return. The unique feature of credit risk, however, implies that the expected return on defaultable bonds is not synonymous with the (pre-default) price process as in the case of default-free bonds. In this paper, the expected return on defaultable bonds is examined within the framework of intensity-based credit risk models. It is shown that a defaultable bond's instantaneous expected return can be decomposed into three parts: a default-free component, a compensation for variations in default risk, and a compensation for investors' risk-aversion towards the default event. The methodology for estimating these components as well as the practical difficulties one might encounter in this estimation are discussed. Easily extended to include a non-default component, this decomposition can enrich our understanding of many empirical observations concerning credit risk.

1 Introduction

Recent research has greatly enhanced our understanding of default risk. For example, the popular reduced-form intensity approach, in the tradition of Jarrow and Turnbull (1995), Lando (1998), and Duffie and Singleton (1999), among others, has been used to model corporate-Treasury spreads (Duffee (1999)), LIBOR-swap spreads (Collin-Dufresne and Solnik (2001)), swap-Treasury spreads (Duffie and Singleton (1997), Liu, Longstaff and Mandell (2000)) and sovereign yield spreads (Duffie, Pedersen and Singleton (2001)). These studies focus on fitting and interpreting the time-series properties of the spreads, and on that measure they are quite successful. Others (Bakshi, Madan and Zhang (2001)) have taken to examine the out-of-sample pricing and hedging performance of such models, also with encouraging results.

Conspicuously missing in this literature is a discussion of the default risk premium, or more generally, the expected return on defaultable bonds. In the default-free term structure literature, it is recently realized that the risk premium specification in term structure models should be flexible enough to produce time-varying expected returns consistent with empirically observed bond excess returns (Duffee (2002)). Though much is still unknown about the empirical features of excess returns on defaultable bonds (perhaps due to the difficulty in measuring returns through default and/or restructuring), a general intensity-based default model might yield valuable guidances useful for subsequent empirical research on this topic.

It might be tempting to think that the issue of defaultable bond expected return has been rendered moot by the work of Duffie and Singleton (1999), which shows that under certain technical conditions, defaultable bonds are priced pre-default with an adjusted short rate that can be modeled as an affine function of diffusion state variables just as in the default-free case.[1] Indeed, for valuation purposes the default-free case and the defaultable case are methodologically indistinguishable—the tractability of affine models, for example, can be exploited in both cases. Unfortunately, this elegant result gives the (unintended) impression that the specification of default risk premium and defaultable bond expected return can also proceed as in default-free term structure models. The liberal use in the empirical literature of terms such as "price of default risk" and "credit premium," representing the

[1] For affine term structure models, see Duffie and Kan (1996). Dai and Singleton (2000) present a detailed classification and empirical study of affine models. Duarte (2000) and Duffee (2002) extend the standard affine framework using more general specifications of the market price of risk. In particular, Duffee (2002) derives general expressions of instantaneous expected return on zero-coupon bonds in affine models and their extensions.

drift adjustments on the affine diffusion state variables underlying the risk-neutral default intensity, is evidence of this confusion.

In a recent article, Jarrow, Lando and Yu (2001, JLY) consider the specification of default risk premium in reduced-form models. They show that there are generally two types of default risk premia, one due to systematic variations in the default intensity and the other a jump risk premium on the default event. When default risk is what they characterize as "conditionally diversifiable," the default event risk premium will disappear and one can specify the default risk premium using drift adjustments as in the default-free case. They propose to use the joint data on default rates and defaultable bond prices to estimate this default event risk premium.[2]

This paper clarifies the implication of JLY's risk premium specification for modeling the expected return on defaultable bonds. Despite the strong analogy between default-free and defaultable bond pricing, the similarity between the two sets of securities stops when it comes to expected return. Intuitively, to compute the expected return on defaultable bonds one has to know the product between the likelihood of default and how much of bond value is expected to be lost post-default, something appropriately termed the "mean-loss rate" under the physical measure. However, in Duffie and Singleton (1999), pre-default bond prices are dependent on the mean-loss rate under the risk-neutral measure. This subtle difference caused by the measure change is absent in the modeling of default-free expected returns.

Using a simple illustrative example from the affine class, Section 2 of this paper shows that the expected return on a defaultable bond can generally be decomposed into three parts. The first component is the expected return on an otherwise identical default-free bond. The second component is related to the parameters of the risk-neutral mean-loss rate and their risk adjustments under the physical measure. It can be understood as a compensation for variations in credit risk generated by the state variables. The third and heretofore ignored component is the difference between the risk-neutral and the physical mean-loss rates and can be thought of as a compensation demanded by investors for bearing the risk of the (systematic or non-diversifiable) default event. Incidentally, this result can also be interpreted along the lines of "survival bias" frequently mentioned in studies of stock returns (Brown, Goetzmann and Ross (1995), Li and Xu (2000)). If one uses only bonds not in default in the estimation of expected return, or if one computes expected

[2]Intuitively, this is not unlike the joint modeling of stock and option prices in order to identify the risk premium on jump factors commonly used to explain the volatility smirk.

return based on the pre-default pricing formula given by Duffie and Singleton (1999), then the expected return will be overstated. The extent of the overstatement is related to the physical mean-loss rate of the bond.

The remainder of this paper is organized as follows. Section 3 discusses the empirical strategy for estimating defaultable bond expected return, with an emphasis on the difficulties one might encounter in the estimation. The conditional default rate example in JLY is reconsidered with a simple but more realistic control for liquidity and tax effects. Section 4 entertains the possibility that the expected return decomposition can help us understand several empirical "irregularities" concerning credit risk. Examples are the finding of a negative credit premium in Liu, Longstaff and Mandell (2000), and the contention that bankruptcy risk is non-systematic in Dichev (1998). Section 5 concludes.

2 Decomposition of Expected Return

This section presents a simple calculation of the expected return on defaultable bonds. To illustrate and emphasize the conceptual distinction from default-free bonds, a standard affine model of the default-free term structure is first considered. Although the derivation below specializes to the CIR specification, the decomposition result can be broadly extended to the entire affine class.

2.1 Default-Free Case

The benchmark case for comparison purposes is a one-factor CIR model of the default-free term structure. In this model, the dynamics of the short rate r_t under the physical measure P and the risk-neutral measure Q are assumed to be:

$$\begin{aligned} dr_t &= \kappa(\theta - r_t)\,dt + \sigma\sqrt{r_t}dW_t, \text{ under } P, &(1)\\ dr_t &= (\kappa+\nu)\left(\frac{\kappa\theta}{\kappa+\nu} - r_t\right)dt + \sigma\sqrt{r_t}d\widetilde{W}_t, \text{ under } Q, &(2)\end{aligned}$$

where κ, θ, ν are constants and W_t and $\widetilde{W}_t$ are P- and Q-Brownian motion, respectively.

In this case, it is well-known that zero-coupon bond prices can be expressed as an exponentially affine function of the short rate:

$$p(t,T) = H_1 \exp(-H_2 r_t), \tag{3}$$

where

$$H_1 = \left(\frac{2\gamma e^{(\gamma+\kappa+\nu)(T-t)/2}}{(\gamma+\kappa+\nu)\left(e^{\gamma(T-t)}-1\right)+2\gamma} \right)^{2\kappa\theta/\sigma^2}, \tag{4}$$

$$H_2 = \frac{2\left(e^{\gamma(T-t)}-1\right)}{(\gamma+\kappa+\nu)\left(e^{\gamma(T-t)}-1\right)+2\gamma}, \tag{5}$$

$$\gamma = \sqrt{(\kappa+\nu)^2+2\sigma^2}. \tag{6}$$

It is then straightforward to apply Ito's lemma to equation (3), yielding the following dynamics for $p(t,T)$:

$$\frac{dp(t,T)}{p(t,T)} = (r_t + H_2 \nu r_t)\, dt - H_2 \sigma \sqrt{r_t} dW_t. \tag{7}$$

Thus the instantaneous expected return is

$$r_t + H_2 \nu r_t, \tag{8}$$

which is an affine function of r_t.

This example shows that the expected return depends on both the risk-neutral dynamics of r_t (through H_2) and the parameter ν which determines the physical dynamics of r_t. For this reason ν is often referred to as the market price of interest rate risk.

2.2 Defaultable Case

Now consider the computation of the expected return on defaultable bonds. For simplicity, the one-factor CIR model for the default-free term structure is maintained. Consideration for default risk is achieved by constructing a risk-neutral mean-loss rate $\widetilde{s}_t$ that follows a square-root diffusion:

$$d\widetilde{s}_t = \left(\kappa' + \nu'\right) \left(\frac{\kappa'\theta'}{\kappa'+\nu'} - \widetilde{s}_t \right) dt + \sigma' \sqrt{\widetilde{s}_t} d\widetilde{W}'_t, \tag{9}$$

where κ', θ', ν' are constants and $\widetilde{W}'_t$ is a Brownian motion independent of $\widetilde{W}_t$ under Q. Similarly, it is assumed that the P-dynamics of $\widetilde{s}_t$ is given by

$$d\widetilde{s}_t = \kappa' \left(\theta' - \widetilde{s}_t\right) dt + \sigma' \sqrt{\widetilde{s}_t} dW'_t, \tag{10}$$

where W'_t is a Brownian motion independent of W_t under P. This specification of the mean-loss rate ignores the empirical observation that yield spreads change systematically with the default-free term structure.[3] However, it suffices for this simple illustration.

[3] For a more general specification that allows such dependence, see Duffee (1999).

As shown by Duffie and Singleton (1999), under certain technical assumptions defaultable bonds can be valued like default-free bonds—one only needs to modify the short rate used for discounting from r_t to $R_t = r_t + \widetilde{s}_t$.[4] The risk-neutral mean-loss rate $\widetilde{s}_t$ is equal to $\widetilde{\lambda}_t (1 - \delta_t)$ where $\widetilde{\lambda}_t$ is the risk-neutral default intensity and δ_t is the recovery rate as a fraction of pre-default market value of the bond. In this setting, once we make the assumption that R_t is an affine function of diffusion state variables, defaultable bond pricing collapses to the default-free case.

2.2.1 The Naive Derivation

In the above example, it can be shown that the defaultable zero-coupon bond price is given by

$$v(t, T) = H_1 H_1^{'} \exp\left(-H_2 r_t - H_2^{'} \widetilde{s}_t\right), \tag{11}$$

where H_1 and H_2 are defined by equations (4)-(6) and $H_1^{'}$ and $H_2^{'}$ are the corresponding primed version. Similar to the derivation of equation (7), it is tempting to conclude that the instantaneous expected return on the defaultable bond is

$$r_t + \widetilde{s}_t + H_2 \nu r_t + H_2^{'} \nu' \widetilde{s}_t, \tag{12}$$

which follows from the dynamics of $v(t, T)$:

$$\frac{dv(t, T)}{v(t, T)} = \left(r_t + \widetilde{s}_t + H_2 \nu r_t + H_2^{'} \nu' \widetilde{s}_t\right) dt - H_2 \sigma \sqrt{r_t} dW_t - H_2^{'} \sigma' \sqrt{\widetilde{s}_t} dW_t'. \tag{13}$$

The above derivation of expected return is similar to that of LLM, who use a multifactor Vasicek specification instead. However, it is not difficult to see that the expression in equation (12) is incomplete. Take the simple example of the market prices of risk being zero, i.e. $\nu = 0$ and $\nu' = 0$. In this case the defaultable bond is priced under risk-neutrality and the instantaneous expected return ought to be just the risk-free rate r_t. Instead, equation (12) gives $r_t + \widetilde{s}_t$. Hence, an important piece is missing from the previous derivation. As shown below, this missing piece is related to the expected return conditional on a default event.

2.2.2 The Correct Derivation

The intuition why equations (12)-(13) can go wrong is simple once it is realized that although Duffie and Singleton's methodology can reduce the pricing of defaultable bonds to that of

[4] For simplicity, other contributors of the spread, such as liquidity, taxes, and asymmetric information, are ignored. Considerations for these factors can be added later on.

default-free bonds, it is only applicable prior to the default event. In deriving expected returns on defaultable bonds, however, one cannot avoid the question of what happens to bond value in default and how likely default will occur. The following presents a somewhat heuristic derivation of the instantaneous expected return on a defaultable bond.

To start, it is assumed that under the physical measure P, the default intensity is given by λ_t.[5] Over the interval $(t, t+\Delta t)$, the likelihood of default is given by $\lambda_t \Delta t$. Thus the return over the interval is

$$\begin{aligned} & \left.\frac{\Delta v(t,T)}{v(t,T)}\right|_{\text{no default}} (1-\lambda_t \Delta t) + \frac{\delta_t v(t-,T) - v(t-,T)}{v(t-,T)} \lambda_t \Delta t \\ \approx \; & \left.\frac{\Delta v(t,T)}{v(t,T)}\right|_{\text{no default}} - (1-\delta_t)\lambda_t \Delta t, \end{aligned} \tag{14}$$

where the first term becomes equation (13) in the limit as $\Delta t \to 0$. Hence the instantaneous expected return is slightly modified from equation (12) as

$$r_t + H_2 \nu r_t + H_2^{'} \nu' \widetilde{s}_t + \widetilde{s}_t - s_t, \tag{15}$$

where $s_t = \lambda_t (1-\delta_t)$ is the mean-loss rate under the physical measure.

The difference between the two equations is essentially a consequence of survival bias—expected returns are bound to be overstated when they are estimated using only the survived sub-sample. The extent of the survival bias is related to the physical default intensity λ_t and not the risk-neutral intensity $\widetilde{\lambda}_t$. While the issue of survival bias is largely an afterthought in the literature on stock returns,[6] it is of paramount importance here since the focus is precisely on the risk and return associated with default and bankruptcy.

2.2.3 Interpretation of JLY's Default Risk Premium Specification

The instantaneous expected return on defaultable bonds, given in equation (15), can be seen to consist of three components. The first component is the return on an otherwise identical default-free bond (the first two terms).[7] The second is an affine function of state variables (third term). The third is the difference between the risk-neutral and physical expected loss rate.

These components can help us understand the default risk premium specification in the recent work of JLY. They consider an intensity-based model of multiple issuers whose

[5] The existence of a point process intensity under equivalent changes of measure is shown by Artzner and Delbaen (1995).

[6] The debate on the importance of the survival bias for equity returns can be found in Brown, Goetzmann and Ross (1995) and Li and Xu (2000).

[7] This term contains the usual interest rate risk premiums.

defaults are independent conditional on a set of state variables. By using a diversification argument similar to that of the original APT, they are able to show asymptotic equivalence between risk-neutral and physical default intensities. The implication for equation (15) is that the third component will disappear because $\lambda_t = \widetilde{\lambda}_t$. In such cases, investors do not demand any compensation for bearing the risk of the default event. Instead, they are compensated because of variation in default risk, which enters equation (15) as the second component. It is in this sense that the parameter ν' can be considered a "default risk premium."

In a situation where the risk inherent in the actual default event cannot be diversified away, one has to consider the third component explicitly. This is made possible because of Girsanov-type results showing that under an equivalent change of measure, point process intensities are linked by a strictly positive predictable process μ_t through the relation $\widetilde{\lambda}_t = \mu_t \lambda_t$. Hence the compensation for the default event becomes $(\mu_t - 1)\, s_t$. It is in this sense that the parameter μ_t is labeled a "default event risk premium." Since investors' risk-aversion toward the default event implies a μ_t greater than one, ignoring the third component will lead to an underestimation of the expected return on defaultable bonds.

2.2.4 Non-Default Sources of Expected Return

Duffie and Singleton (1999) show that liquidity and other non-default factors can be absorbed into the risk-neutral mean-loss rate in pricing defaultable bonds. One simply re-interprets $\widetilde{s}_t$ to include an additive term l_t, possibly state dependent, which captures the effect of non-default factors. The mean-loss rate can still be modeled by an affine function of state variables, preserving the original structure and results.

To re-interpret equation (15), it is noted that the non-default factors can affect expected returns in two ways. First, there is a compensation if they vary with the state variables. This is similar to the compensation for changes in default risk. Second, the modified risk-neutral mean-loss rate now contains non-default factors, while the physical mean-loss rate does not. This difference represents a direct compensation for the non-default factors.

3 Estimation of Default Event Risk Premium

As is evident from the previous section, the issue of expected return is reduced to a standard calculation using estimated model parameters (say within the affine class) when default event risk is not priced. Since this part of the problem is well understood from the term structure literature, the estimation of defaultable bond expected return boils down to the estimation of the default event risk premium. This section focuses on the practical issues associated with this estimation.

3.1 Methodology

Consider the yield spread at time t on a zero-coupon bond with maturity T. It is noted that

$$\begin{aligned}\text{Spread}\,(t,T) &= -\frac{1}{T-t}\ln\frac{v\,(t,T)}{p\,(t,T)} \\ &= -\frac{1}{T-t}\left(\ln\widetilde{E}_t\left(e^{-\int_t^T (r_u+\widetilde{s}_u)du}\right)-\ln\widetilde{E}_t\left(e^{-\int_t^T r_u du}\right)\right) \\ &= -\frac{1}{T-t}\left(\ln\widetilde{E}_t\left(e^{-\int_t^T (r_u+\mu_u s_u+l_u)du}\right)-\ln\widetilde{E}_t\left(e^{-\int_t^T r_u du}\right)\right), \qquad (16)\end{aligned}$$

where $\widetilde{s}$, s, μ, and l are respectively the risk-neutral mean-loss rate, the physical mean-loss rate, the default event risk premium, and the non-default component of the mean-loss rate, and $\widetilde{E}_t(\cdot)$ denotes the expectation under the risk-neutral measure. This expression shows that any empirical feature of the yield spread must be attributable to one of four sources: 1) expected default and recovery through s; 2) risk premium for default event risk through μ; 3) risk premium for variations in the spread through the drift adjustments on $\widetilde{s}$ when changing to the risk-neutral measure; and 4) non-default reasons such as tax and liquidity through l. Abstracting away from non-default sources for the time being, the spread $\widetilde{s}$ and the risk premium for variations in the spread can be identified via standard affine term structure estimation methods. If the physical mean-loss rate s can be separately estimated, the default event risk premium μ can be inferred as $\mu = \widetilde{s}/s$.

3.2 Risk-Neutral Mean-Loss Rate

For the first step of this procedure, the risk-neutral mean-loss rate $\widetilde{s}$ would have to be estimated. As a more specific and workable approach, one can model the risk-neutral mean-loss rate as an affine function of default-free short rate factors, a “spread index” factor, plus a firm-specific latent factor. The inclusion of short rate factors is to account for the well-documented relationship between yield spread and the default-free term structure (see Duffee (1998)). The spread index can for example be the difference between Moody’s Baa and Aaa yield indices, which is meant to capture a major part of the systematic variation in the cross-section of yield spread changes (see Collin-Dufresne, Goldstein and Martin (2001)). The short rate factors and the spread index factor can be jointly estimated in the first stage, and given their estimated values, in the second stage the firm-specific factor can be inverted from bond prices.

The short rate factors and the spread index factor should capture enough of the systematic variation in the spreads such that there should be no significant remaining risk premium for the latent firm-specific factor. The presence of a firm-specific default risk premium suggests that the inferred firm-specific factor probably subsumes other common factors contributing to changes in the yield spread. This in no way affects the subsequent estimation of the default event risk premium. The JLY framework of diversifiable default risk, for example, can be easily extended to include a firm-specific component in the default intensity, encompassing specifications such as that used by Duffee (1999).

3.3 Physical Mean-Loss Rate

Next, one has to estimate the physical mean-loss rate. To simplify the task, one can assume that the recovery given default is a constant equal to some measure of a historical average such as those compiled by Moody's. The physical mean-loss rate is then equivalent to the physical default intensity. Still, firm-specific estimates of a time-varying physical default intensity are difficult to obtain. Since default is a rare event, such estimates would probably have to be based on a representative firm analysis using financial statement variables and equity returns. This approach has been applied to bankruptcy prediction (see Altman (1968, 1993), Shumway (2001)), but it has not been tested for its ability to produce accurate default probabilities. An alternative would be to use the contingent claims analysis to calculate default probabilities based on the time-varying capital structure of a firm (Merton (1974)). This approach is used by KMV to construct their EDF (expected default frequency) measure and presumably is also the approach used by Moody's to rate the bond issues. However, these are proprietary models employing proprietary data, making them infeasible for academic research.[8]

If one gives up the hope of having firm-specific and time-varying default intensities, then Moody's historical default rates by credit rating class can be used. These are estimated by following a cohort of firms through a given horizon and then averaging the default rates across all possible starting years in the sample. Alternatively one can use the average one-year transition matrix to compute default rates with various horizons. Because of the

[8] Vassalou and Xing (2001) use Merton's (1974) model to compute default probabilities. They show that this measure of default risk works better than accounting-based measures or default spreads in explaining the cross-section of stock returns. Thus it appears that although the structural approach runs into difficulties as a valuation model for corporate debt (see Jones, Mason and Rosenfeld (1984)), it may be well-suited for constructing physical default probabilities.

time-aggregation in this procedure, the ability to estimate a random process for μ is lost. As a result, one may simply assume that μ is equal to a constant and proceed to calculate the implied physical default rates by integrating over the risk-neutral default intensity divided by μ. By minimizing the distance between these implied values and those from Moody's, one obtains an estimate of μ. This can be used for example to test the diversification condition (whether μ is equal to one) hypothesized by JLY.

[Figure 1 about here]

Caution must be exercised, however, in interpreting the result obtained by assuming a constant event risk premium. If the realistic situation dictates a time-varying one, and especially if it is characterized by occasional episodes of high values during, say, market crashes, such a simplified approach may lead to erroneous or incomplete conclusions. To get a glimpse of this issue, let us consider the relation between expected default and the yield spread. Figure 1 shows the speculative-grade 12-month trailing default rate as given by Moody's and the speculative-grade spread against Treasuries as given by Datastream. The speculative-grade 12-month trailing default rate is calculated by identifying all speculative-grade issuers and then computing the fraction of issuers that have gone into default in the following year. It is clear from Figure 1 that the two series follow the same trend over the sample period. It also appears that the spread leads the default rate by about one year despite the use of a trailing measure. This is because the spread is calculated using bonds whose maturities are most likely beyond one year. If one were able to use an expected default rate over an appropriate horizon, say the average duration of bonds in the speculative-grade index, such lead-lag effects will probably be reduced.

Notable in Figure 1 are episodes in which dramatic changes in the spread are not accompanied by concurrent or lagged responses in the default rate. The latter half of 1998 seems to be a good example. This episode has been commonly attributed to "low liquidity and shifting investor preferences" (Keenan (2000)). If there is no change in the physical mean-loss rate, in our framework yield spreads can change for three possible reasons—respectively, changes in drift adjustments on the state variables, changes in the event risk premium, and changes in non-default factors such as liquidity. A priori there is no reason to rule out any of these possibilities. Since all three may be intricately linked in practice, the identification of the event risk premium becomes even more daunting.

3.4 Non-Default Factors

The estimation of the event risk premium also requires a "clean" risk-neutral mean-loss rate that is uncontaminated by taxes, liquidity and other non-default factors.[9] Although a fully specified model of non-default sources of the spread remains elusive, ad hoc adjustments have been suggested. For example, JLY subtract the average value of the Aaa-implied intensity from the intensities implied from other ratings. They argue that this is a crude way of eliminating the effect of non-default factors. This section first focuses on the role of non-default factors in explaining the short-end spread, then revisits the JLY example with another ad hoc but potentially more realistic adjustment for the effect of non-default factors. It is shown that this produces implied default rates that better conform to Moody's estimates. The purpose of this example is to illustrate the importance of some form of adjustment for non-default factors before one can embark on a more rigorous analysis of the default event risk premium.

3.4.1 Short-End Spreads

Consider the yield spread on a zero-coupon bond as given in equation (16). Assuming continuity of the processes, it is easy to see that

$$\text{Spread}\,(t, t+) = \widetilde{s}_t. \tag{17}$$

On the other hand, the probability of survival until time T is

$$\text{SP}\,(t, T) = E_t\left(e^{-\int_t^T \lambda_u du}\right), \tag{18}$$

where λ_u is the physical default intensity and $E_t\left(\cdot\right)$ denotes the expectation under the physical measure. One can then calculate the instantaneous conditional default rate as

$$\text{CDR}\,(t, T) = -\frac{\partial \ln \text{SP}\,(t, T)}{\partial T}. \tag{19}$$

This is the default probability per unit time conditional on no default within the interval (t, T). In the limit as $T \to t+$, the conditional default rate becomes λ_t. Assuming diversifiable default risk as in JLY, one has $\lambda_t = \widetilde{\lambda}_t$ and the following:

$$\text{CDR}\,(t, t+) = \frac{\text{Spread}\,(t, t+)}{1 - \delta_t}, \tag{20}$$

[9] See Duffie and Lando (2001), Schultz (2001), Elton et al. (2001), and Huang and Huang (2000) for discussions of these issues.

which results from the definition of $\widetilde{s}_t$. Therefore, under diversifiable default risk and if there were no non-default factors, there is a simple relation that ties down default rates and spreads in the short-end.[10]

[Figure 2 about here]

The data, however, suggest something much more complex. Figure 2 presents two plots. The first plot is the term structure of conditional default rates assuming that the issuer is endowed with a given initial credit rating. This is based on Moody's one-year transition matrix averaged between 1980 and 1999.[11] The second plot is the term structure of yield spreads computed from the empirical estimates provided in Duffee (1999). Although the two plots come from time periods that are not entirely comparable, the stark contrast in the short-end probably transcends such differences. Take the Aa curves for example, the short-end spread is roughly 60 bps (basis points). With a recovery rate on senior unsecured debt estimated by Moody's to be about 44 percent, this translates into a default intensity of about 100 bps. In comparison, the conditional default rate measured at the end of year one is only 4 bps. This is the most extreme among the three cases presented. Even in the mildest case of Baa-rated issuers, a short-end spread of 100 bps yields an intensity of about 180 bps, much in excess of the corresponding short-end default rate of 20 bps. The relation in equation (20) is therefore grossly violated.

Without the presence of non-default related factors, to reconcile the magnitude of the short-end spread with such a small default rate would require an incredibly large μ_t—much too large, in fact, to be compatible with their apparent agreement in the mid- to long-maturity range.[12] To explain this observation, it seems that one is forced to appeal to non-default related causes such as liquidity and taxes, whose effect on required yields may be declining as a function of maturity. These considerations may also affect expected returns because in equation (15), the risk-neutral "expected loss" $\widetilde{s}_t$ now includes the effect of liquidity and taxes, while the physical expected loss s_t does not.

[10] In deriving this result, it is assumed that the intensities are driven by diffusion state variables. However, the result generalizes to jump state variables as long as the rate of jumps is finite.

[11] With the one-year transition matrix, only yearly estimates of survival probabilities can be computed. Therefore, standing in year 0, it is assumed that the probability of surviving until year n is p_n. The conditional default rate in year n is then computed as $(p_n - p_{n+1})/p_n$. This is of course a discrete-time approximation to the instantaneous conditional default rate defined above.

[12] The same is true for ν', the drift adjustment of the risk-neutral expected loss rate. A numerical example illustrating this point is given by JLY.

3.4.2 Adjustment for Non-Default Factors

To account for the effect of non-default factors, one needs to consider an adjustment to the default intensity. In the absence of any guiding empirical work, it is noted that a constant adjustment to the intensity leads to a constant non-default component of the spread, while adjusting all bond prices downward by a constant fraction leads to a spread adjustment that approaches zero as maturity increases. Specifically, it decreases as $\frac{1}{T-t}$. The realistic situation perhaps lies somewhere between these two cases. For simplicity, a case of diversifiable default risk is assumed so that the example below is not complicated by a nontrivial μ_t process. For computations of default rates and spreads at time t, let the adjusted default intensity be defined by

$$\lambda_u' = \lambda_u - \frac{\alpha}{u-t+\beta},\ u \geq t, \tag{21}$$

where α and β are positive constants. It is straightforward to show that various adjusted quantities of interest (primed variables) are

$$\begin{aligned} \mathrm{Spread}'(t,T) &= \mathrm{Spread}(t,T) - \frac{\alpha(1-\delta)}{T-t}\ln\left(\frac{T-t+\beta}{\beta}\right), \\ \mathrm{SP}'(t,T) &= \mathrm{SP}(t,T)\left(\frac{T-t+\beta}{\beta}\right)^{\alpha}, \\ \mathrm{CDR}'(t,T) &= \mathrm{CDR}(t,T) - \frac{\alpha}{T-t+\beta}, \end{aligned} \tag{22}$$

where a constant recovery rate δ is assumed. Hence after adjusting for non-default, "true" spreads are narrower, "true" survival probabilities are larger, and "true" conditional default rates are lower. Focusing on the conditional default rate above, one can see that the parameters α and β control its level at the short-end as well as the speed at which the adjustment term decays with maturity. Clearly, these parameters can be chosen in some suitable fashion to produce a good match of spreads and conditional default rates in the short-end in relation to equation (20).[13] An example is given next.

[13] JLY have argued that by adjusting for the effect of state taxes and liquidity as in Elton et al. (2001) and others, short-end spreads can be boiled down to physical mean-loss rates alone. Ericsson and Renault (1999) construct a structural model of liquidity and credit risk. One of their findings is that the liquidity component of the spread is a decreasing function of bond maturity and that the price discount approaches a constant as maturity increases. These results support the interpretation of the ad hoc adjustment given above as due to non-default factors. Of course, the simple adjustment here ignores the interaction between credit and liquidity risk which is the focus of Ericsson and Renault (1999). However, this can be incorporated into the general intensity-based framework by assuming state-dependent liquidity adjustments.

3.4.3 Term Structure of Implied Conditional Default Rates

Incidentally, the above non-default adjustment can shed light on another interesting observation concerning the shape of the term structure of implied conditional default rates. JLY, for example, have relied on their notion of asymptotic equivalence between the risk-neutral and the physical default intensity to argue that the term structure of physical default rates can be computed with the risk-neutral intensity extracted from corporate bond prices. However, they show that the result of such an exercise does not resemble that produced using Moody's one-year rating transition matrix, which is based on historical default data. Specifically, for investment-grade issuers the former is high and flat because of the fast mean-reversion of the estimated risk-neutral intensity under the physical measure, while the latter starts out low but gradually increases with maturity. One can experiment with changing the initial value of the risk-neutral default intensity, but this does little to reconcile the differences because as Duffee (1999) shows, the intensity is bounded below by a positive constant and the process mean-reverts rather quickly under the physical measure.

[Figure 3 about here]

This observation is reproduced in Figure 3 with a number of plots for the term structure of conditional default rates. The "unadjusted" series is the result of applying the JLY methodology to Duffee's intensity estimates for Baa-rated issuers. The "actual" series is the Baa default rate computed from Moody's average one-year transition matrix between 1980 and 1999. Under JLY's diversifiable default risk hypothesis, these two curves ought to be identical. In reality, however, they are indistinguishable only in the mid- to long-maturity range but differ significantly in the short-end.

Incorporating the non-default adjustment of the previous subsection, the parameters α and β are chosen to minimize the deviation of the non-default adjusted conditional default rate of equation (22) from the "actual" in the short-end.[14] Evidently, such a simple model of liquidity, with only two parameters, does a good job in isolating the effect of liquidity and other confounding factors.[15] As expected, the difference between the two curves is minimal in the short-end. In the long-end, their difference could be due to a nontrivial μ_t, but is likely to be statistically insignificant given the precision of Duffee's estimates. As a result,

[14] Specifically, the sum of squared difference between actual and adjusted one-year conditional default rate over $T-t=0,1,\ldots,4$ is minimized. The long-end is left out of the optimization since liquidity likely has a lesser role there.

[15] Ericsson and Renault (1999) show that the downward sloping term structure of liquidity spreads can explain the apparent failure of structural models to produce a flat term structure of yield spreads for investment-grade issuers. Theirs is another look at essentially the same problem from a different angle.

the overall shape of the term structure of “implied” physical default rates is much closer to the one obtained from actual default data. What was previously perceived as a slowly mean-reverting intensity process necessary to generate the gently upward sloping default rates may simply be an artifact of the declining liquidity premium with maturity.

4 Understanding Empirical “Irregularities”

This section entertains the idea that some of the credit risk “irregularities” uncovered by recent empirical studies can be explained by the notion of default event risk and its associated risk premium. Two examples are considered. The first is the finding by Liu, Longstaff and Mandell (2000) that the excess return on LIBOR bonds as implied by swap-Treasury spreads is negative in over half of their sample period. The second is the finding by Dichev (1998) that firms with higher bankruptcy risk do not earn higher subsequent returns on their stocks.

4.1 Credit Premium in the Swap Market

In a recent study of the swap market, Liu, Longstaff and Mandell (2000, LLM) compute the expected return on bonds issued by the fictitious “LIBOR” issuer using the intensity implied from swap-Treasury spreads. They produce the puzzling observation that the part of this expected return attributable to credit risk—what they call “credit premium”—is negative in over half of their sample period. An explanation of this puzzle is presented below using the framework of Section 2.[16]

The interest rate swap market differs from the U.S. corporate bond market in the source of its credit risk. Corporate bonds are issued by companies and hence reflect their individual credit worthiness. In comparison, the interest rate swap contract itself contains minimal credit risk from the counterparties due to the contract design.[17] Instead, the credit risk reflected in the swap rates comes from the three-month LIBOR rate that determines the floating side of the swap, which is agreed upon by a consortium of large international banks located in London. This has led JLY to suspect that the default event underlying the swap-Treasury spread is systematic rather than diversifiable, hence commanding its own risk premium.

[16] LLM use a four-factor Vasicek model in order to capture features of the swap-Treasury spread. In the illustration to follow we use a significantly simpler setup which incidentally does not allow the possibility of a negative credit premium. However, the basic intuition easily migrates to the more complex model.

[17] For example, see Duffie and Huang (1996), Jarrow and Turnbull (1997), Li (1998), and Collin-Dufresne and Solnik (2001).

Consider the instantaneous expected return on a default-free and a defaultable zero-coupon bond. In Section 2, they are given respectively by equations (8) and (15). Taking the difference, the risk premium is given by

$$H_2^{'}\nu'\widetilde{s}_t + \widetilde{s}_t - s_t. \tag{23}$$

This shows that the risk premium contains two parts. The first part is due to changes in the spread and the second part is due to the systematic nature of the default event. Since LLM have only considered the first part, it is informative to take a detailed look at the entire expression. Using the relation $\widetilde{s}_t = \mu_t s_t + l_t$ where l_t represents the effect of liquidity differential between swap contracts and Treasury bond benchmarks used in computing the spread, we see that the risk premium is

$$H_2^{'}\nu'\mu_t s_t + (\mu_t - 1)\, s_t + H_2^{'}\nu' l_t + l_t. \tag{24}$$

To interpret this expression, the first term is a compensation for changes in default risk. The second term is a default event risk premium. The third and fourth terms are liquidity premiums, with the fourth term representing a direct compensation and the third a compensation for changes in liquidity risk. One would naturally consider the sum of the first two terms as defining the credit premium.

LLM, however, define their credit premium as the sum of the first and the third terms above. They argue that this can turn negative primarily because l_t may be negative (meaning that swap contracts have higher liquidity than on-the-run Treasury bonds). Consequently, the "liquidity adjusted" credit premium (essentially just the first term) will be higher and in fact can become positive in their four-factor Vasicek model.

While the logic of this argument is correct, we see that it ignores an important piece of the puzzle.[18] With a non-trivial μ_t (greater than one on average) affecting the swap market, even LLM's liquidity adjusted definition will be understating the true credit premium. As argued above, this is quite plausible given the mechanism for determining the underlying LIBOR rate. Therefore, the credit premium can be positive even in the absence of the alleged swap-Treasury liquidity differential.

[18] On a minor note, the argument for a negative liquidity premium suggests that the default part of the spread should be measured with respect to a fictitious Treasury bond with the same liquidity as a swap contract. This seems to conflict with LLM's earlier choice of the specialness premium in the Treasury market as the liquidity component of the swap spread process.

4.2 Default Risk and Stock Returns

A recent paper by Dichev (1998) examines the question whether bankruptcy risk is systematic. The approach is to regress realized monthly stock returns, which are proxies for expected return and systematic risk, onto proxies of bankruptcy risk, one of which is Altman (1968, 1993)'s Z-score. The paper finds that the regression coefficient of bankruptcy risk is not significantly different from zero, and is even negative in some sample periods, suggesting that firms with higher bankruptcy risk earn lower than average expected returns. The author concludes that default risk is not systematic.

The expected return decomposition of equation (24) suggests additional complexity in answering the above question.[19] After all, the default likelihood under the physical measure, which Altman's Z-score purports to measure, is only one of several determinants of corporate bond (as well as common stock) expected return. It is therefore not difficult to give a number of conjectures that might explain Dichev's empirical finding.

First, it is noted that Dichev's data include delisting returns. To the extent that the monthly returns of every firm in the sample, up to and including its potential bankruptcy, are correctly accounted for by this procedure, equation (24) would be the appropriate expression for assessing expected return. Shumway (1997) and Shumway and Warther (1999), however, show that this is not the case. Not including or not adequately accounting for the large negative returns on defaulted firms in the analysis would amount to having an expected return equal to just the first two terms in equation (23). Intuitively, when the return on bankrupt firms is overlooked, realized return will tend to be overstated. This "survival bias" is more severe for speculative-grade issuers since their default probabilities are larger than those of investment-grade issuers. Unfortunately, it is easy to see that this line of thinking cannot explain why firms with high bankruptcy risk tend to have below-average returns.

Doubts about the data aside, a plausible explanation for Dichev's findings is that speculative-grade default and recovery rates may be less sensitive to the state variable than their investment-grade counterparts. This will generate a difference in the market price of risk parameter ν' which might offset a positive relation between bankruptcy risk and expected return. For example, consider the distinction between an oil company heavily invested in oil exploration and, say, GM. The oil company's fortune depends a great deal on the success of its small number of current projects, which are highly risky but idiosyn-

[19] To be sure, equation (23) addresses the expected return on defaultable bonds, while Dichev (1998) studies the relation between bankruptcy risk and stock return. However, it is not unreasonable to assume that the insights from studying the first problem will have some bearing on the second.

cratic. On the other hand, GM is a much safer bet, but automobile sales are typically highly pro-cyclical. Note, however, that the above argument applies in a one-factor setting. With multiple factors driving changes in the default intensity over the business cycle, various factor exposures can also produce offsetting effects on expected return when moving from high to low credit quality, thus explaining the lack of correlation between bankruptcy probability and expected return.

A second explanation is that speculative-grade issuers may command a smaller event risk premium than investment-grade issuers. This has a similar offsetting effect on expected returns. This explanation is plausible if, for example, low quality or "junk" issues tend to be small and reside in large portfolios while high quality issues tend to be large and held by only a small number of investors (which implies that their portfolios are rather concentrated). Alternatively, investors in junk bonds may be used to the notion of default risk and as such, they may be risk-neutral with respect to the default events. In comparison, defaults by investment-grade bonds are rare and to that extent investors in high quality debt may well demand a compensation for a "surprising" default event.

Without actual empirical work, these "explanations" are merely speculations. Nevertheless, they show that Dichev's conclusion is limited by his definition of "systematic default risk." In particular, before one launches any non-risk based explanation of Dichev's results, it seems necessary to achieve a more detailed understanding of bankruptcy risk than is currently available. This can be done by utilizing the structure of expected return analyzed in this paper.

5 Conclusion

This paper examines the expected return on defaultable bonds. With the proliferation of standard term structure modeling techniques in the credit risk literature, one has to exercise caution in transplanting certain procedures for default-free bonds to defaultable bonds. This paper emphasizes expected return calculation as an example.

Using a simple affine framework, this paper shows that the expected return on defaultable bonds can be decomposed into a default-free component, a compensation for variations in default risk, and a compensation for bearing the default event. The last component has heretofore been ignored in the empirical credit risk literature. It originates from a "survival bias" that arises, for example, whenever one drops defaulted firms in computing realized returns on a bond portfolio.

The expected return decomposition is then used in an attempt to resolve several empirical "puzzles" in the credit risk literature. First, the default event risk premium can help to explain the negative "credit premium" identified by Liu, Longstaff and Mandell (2000) in their study of the LIBOR swap-Treasury spread. Second, the default event risk premium may be one of the reasons why Dichev (1998) finds an insignificant relation between stock return and default risk.

This paper also presents a general methodology for estimating the default event risk premium. However, many difficulties remain for this estimation, including the specification and identification of stochastic processes describing the physical default intensity as well as the non-default part of the yield spread. These difficulties must be overcome before one can test the diversification hypothesis posited by JLY.

References

Altman, E. I., 1968, "Financial Ratios, Discriminant Analysis, and the Prediction of Corporate Bankruptcy," *Journal of Finance*, 23, 589-609.

Altman, E. I., 1993, *Corporate Financial Distress and Bankruptcy: A Complete Guide to Predicting and Avoiding Distress and Profiting from Bankruptcy*, John Wiley & Sons, New York.

Artzner, P., and F. Delbaen, 1995, "Default Risk Insurance and Incomplete Markets," *Mathematical Finance*, 5, 187-195.

Bakshi, G., D. Madan, and F. Zhang, 2001, "Investigating the Sources of Default Risk: Lessons from Empirically Evaluating Credit Risk Models," Working Paper, University of Maryland.

Brown, S., W. Goetzmann, and S. A. Ross, 1995, "Survival," *Journal of Finance*, 50, 853-873.

Collin-Dufresne, P., and B. Solnik, 2001, "On the Term Structure of Default Premia in the Swap and LIBOR Markets," *Journal of Finance*, 56, 1095-2015.

Collin-Dufresne, P., R. Goldstein, and S. Martin, 2001, "The Determinants of yield spread Changes," *Journal of Finance*, 56, 2177-2207.

Dai, Q., and K. Singleton, 2000, "Specification Analysis of Affine Term Structure Models," *Journal of Finance*, 55, 1943-1978.

Dichev, I., 1998, "Is the Risk of Bankruptcy a Systematic Risk?" *Journal of Finance*, 53, 1131-1148.

Duarte, J., 2000, "The Relevance of the Price of Risk in Affine Term Structure Models," Working Paper, University of Chicago.

Duffee, G. R., 1998, "The Relation between Treasury Yields and Corporate Bond Yield Spreads," *Journal of Finance*, 53, 2225-2242.

Duffee, G. R., 1999, "Estimating the Price of Default Risk," *Review of Financial Studies*, 12, 197-226.

Duffee, G. R., 2002, "Term Premia and Interest Rate Forecasts in Affine Models," *Journal of Finance*, 57, 405-443.

Duffie, J. D., and M. Huang, 1996, "Swap Rate and Credit Quality," *Journal of Finance*, 51, 921-949.

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

Duffie, J. D., and R. Kan, 1996, "A Yield-Factor Model of Interest Rates," *Mathematical Finance,* 6, 379-406.

Duffie, J. D., L. Pedersen, and K. J. Singleton, 2001, "Modeling Sovereign Yield Spreads: A Case Study of Russian Debt," Working Paper, Stanford University.

Duffie, J. D., and K. J. Singleton, 1997, "An Econometric Model of the Term Structure of Interest Rate Swap Yields," *Journal of Finance* 52, 1287-1321.

Duffie, J. D., and K. J. Singleton, 1999, "Modeling Term Structures of Defaultable Bonds," *Review of Financial Studies,* 12, 687-720.

Elton, E., M. Gruber, D. Agrawal, and C. Mann, 2001, "Explaining the Rate Spread on Corporate Bonds," *Journal of Finance* 56, 247-278.

Ericsson, J., and O. Renault, 1999, "Credit and Liquidity Risk," Working Paper, McGill University.

Huang, J., and M. Huang, 2000, "How Much of the Corporate-Treasury Yield Spread is Due to Credit Risk?" Working Paper, Pennsylvania State University.

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

Jarrow, R. A., D. Lando, and F. Yu, 2001, "Default Risk and Diversification: Theory and Applications," Working Paper, UC-Irvine.

Jarrow, R. A., and S. M. Turnbull, 1995, "Pricing Derivatives on Financial Securities Subject to Credit Risk," *Journal of Finance,* 50, 53-85.

Jarrow, R. A., and S. M. Turnbull, 1997, "When Swaps Are Dropped," *Risk,* 10(5), 70-75.

Jones, E. P., S. P. Mason, and E. Rosenfeld, 1984, "Contingent Claims Analysis of Corporate Capital Structures: An Empirical Investigation," *Journal of Finance*, 39, 611-625.

Keenan, S. C., 2000, "Historical Default Rates of Corporate Bond Issuers: 1920-1999," Moody's Special Comment.

Lando, D., 1998, "On Cox Processes and Credit Risky Securities," *Review of Derivatives Research,* 2, 99-120.

Li, H., 1998, "Pricing of Swaps with Default Risk," *Review of Derivatives Research*, 2, 231-250.

Li, H., and Y. Xu, 2001, "Can Survival Bias Explain the Equity Premium Puzzle?" Working Paper, Cornell University; forthcoming in *Journal of Finance*.

Liu, J., F. Longstaff, and R. Mandell, 2000, "The Market Price of Credit Risk: An Empirical Analysis of Interest Rate Swap Spreads," Working Paper, UCLA.

Merton, R. C., 1974, "On the Pricing of Corporate Debt: The Risk Structure of Interest Rates," *Journal of Finance*, 29, 449-470.

Schultz, P., 2001, "Corporate Bond Trading Costs: A Peek Behind the Curtain," *Journal of Finance*, 56, 677-698.

Shumway, T., 1997, "The Delisting Bias in CRSP Data," *Journal of Finance*, 52, 327-340.

Shumway, T., 2001, "Forecasting Bankruptcy More Accurately: A Simple Hazard Model," *Journal of Business,* 74, 101-124.

Shumway, T., and V. Warther, 1999, "The Delisting Bias in CRSP's Nasdaq Data and Its Implications for the Size Effect," *Journal of Finance,* 54, 2361-2379.

Vassalou, M., and Y. Xing, 2002, "Default Risk in Equity Returns," Working Paper, Columbia University.

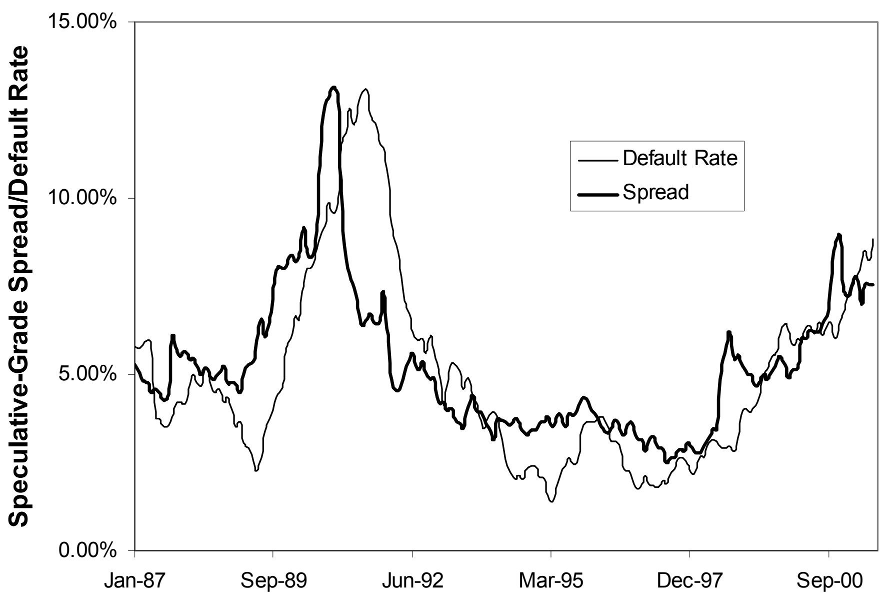

Figure 1: Speculative-grade trailing 12-month default rate and average speculative-grade spread against Treasuries. The default rates are obtained from Moody's. The spreads are obtained from Datastream. The sample period is January 1987 to August 2001.

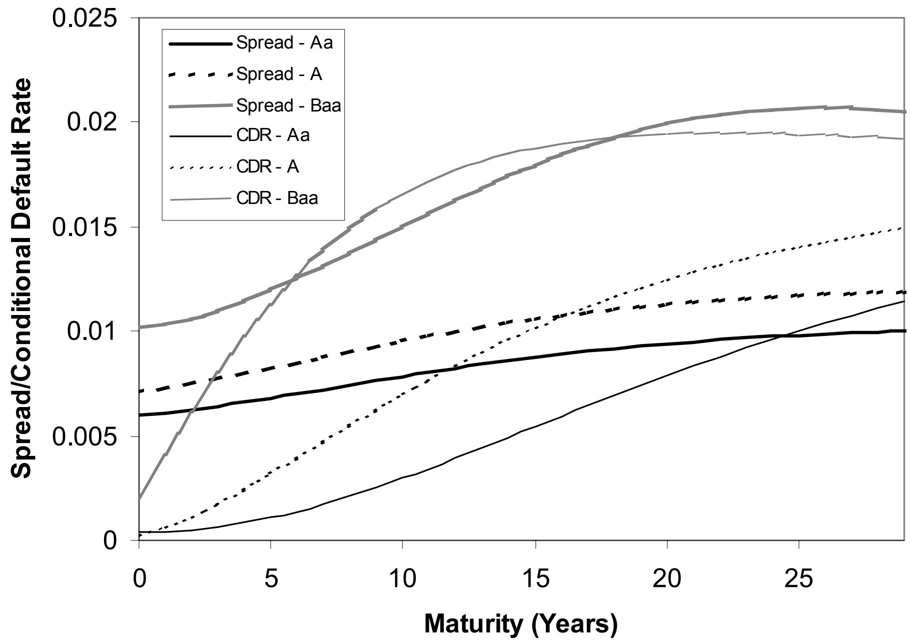

Figure 2: Term structures of spreads and conditional default rates. The spreads are computed based on estimated parameters in Duffee (1999). The conditional default rates are computed from Moody's average one-year transition matrix from 1980 to 1999.

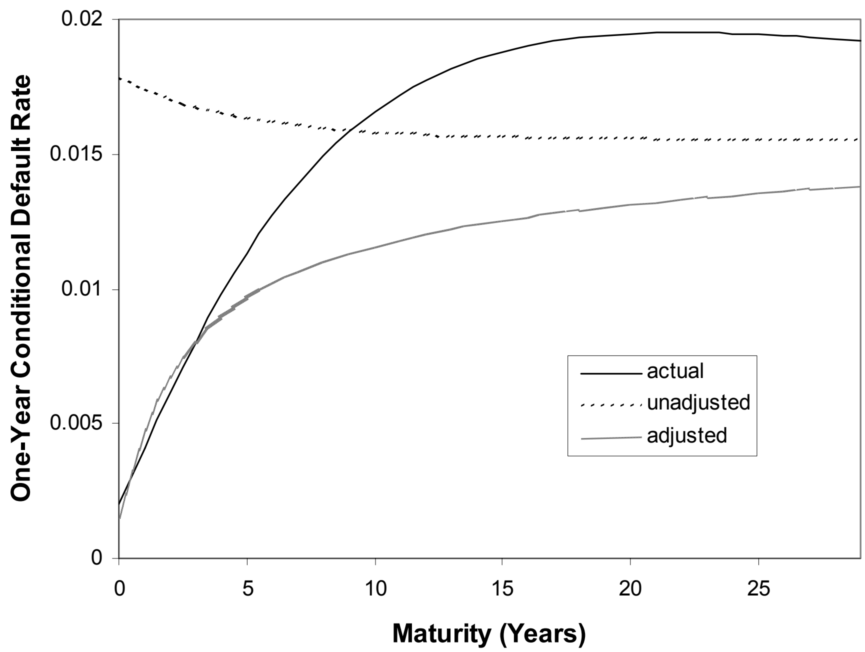

Figure 3: Term structures of conditional default rates. The "actual" series is the Baa default rate computed from Moody's average one-year transition matrix between 1980 and 1999. The "unadjusted" series is inferred from Duffee (1999)'s estimated default intensity for Baa-rated issuers. The "unadjusted" series is adjusted for liquidity premia using equation (21), yielding the "adjusted" series.