Do we need to worry about credit risk correlation?

Abel Elizalde*
CEMFI and UPNA

Journal of Fixed Income, forthcoming.
December 2005.

Abstract

Yes we do. This paper shows that any firm's credit risk is, to a very large extent, driven by common risk factors affecting all firms. Using a reduced form model and sequential Kalman filtering estimation we decompose the credit risk of a sample of corporate bonds (14 US firms, 2001-2003) into different unobservable risk factors. A single common factor accounts for more than 50% of all (but two) of the firms' credit risk levels, with an average of 68% across firms. This factor represents the credit risk levels underlying the US economy and is strongly correlated with main US stock indexes.

Keywords: Credit risk, correlation, factor models.

JEL Classification: C19, G12, G13.

*I would like to thank Juan Ayuso, Max Bruche, Jose Luis Fernández, Santiago Forte, Lane P. Hughston, Manuel Moreno, Rafael Repullo, and Enrique Sentana for their help and comments. Thanks are also due to the Publications department at Bank for International Settlements (Monetary and Economic Department) for permission to reproduce some pictures from BIS Quarterly Review. All remaining errors are my own. First version: May 2005. CEMFI, Casado del Alisal 5, 28014 Madrid, Spain. Email: abel_elizalde@hotmail.com. Web: www.abelelizalde.com.

1 Introduction

Credit risk plays a central role in credit derivatives pricing, portfolio management, and bank capital regulation. As a consequence, and due to parallel efforts of academic researchers, practitioners and regulators, our understanding about credit risk has experienced a big step forward in recent years. New modelling techniques and empirical research have been produced at a rapid pace. However, as we move forward in the modelling and measurement of credit risk new issues arise. Without doubt, the most important one has to do with credit risk correlations.

How is the credit risk of different firms related? The answer to this simple question has substantial consequences for the pricing of multiname credit derivatives, the management of portfolios of defaultable instruments and credit risk regulatory capital charges. Still, I would claim that we don't even grasp the magnitude of the correlation problem: How important are credit risk correlations? Important enough to worry about them? What is the principal force driving them?

This paper presents an empirical exercise which helps to discern how much of the credit risk can be explained by common factors. Using bond prices for a sample of 14 US firms during 2001-2003, we apply a reduced form model and an estimation procedure based on Kalman filtering in order to extract the realization of unobservable common, sector and idiosyncratic risk factors affecting the firms' credit risk. We find that credit risk correlations, measured by the impact of a few common factors on the firms' credit risk, explain a large part of such risk. In particular, a single common factor affecting the credit risk of all firms is found to explain between 15 and 91% of the firms' credit risk, with an average of 68% across firms.

Such a single common factor driving credit risk correlations is strongly and negatively correlated with US stock indexes (Dow Jones and S&P 500). Furthermore, it can be shown to resemble the evolution of different variables capturing the underlying credit risk levels in the US economy.

Two main types of models have been used in the credit risk literature: structural and reduced (or intensity) form models. Structural models derive default probabilities and credit risk dependencies between firms from the evolution of structural variables representing their assets and liabilities structure (see Elizalde 2005b for a survey.) In contrast, reduced form models rely on the market prices of the firms' defaultable securities to capture the evolution of the firms' credit risk. In particular, the firms' instantaneous default probability is modelled as the hazard rate or intensity of arrival of a Poisson process, whose first jump represents the time of default.

From the different approaches the literature has proposed to model and estimate credit risk correlations in reduced form models, conditionally independent defaults (CID) models represent, in our view, the most flexible and reliable approach (see Elizalde 2005a for a survey). CID models introduce correlation in the firms' credit risk making their default intensities dependent on a set of common and firm specific factors; conditioned on the realization of the common factors, the firms' default intensities are independent.

Using a reduced form model, we identify as credit risk the product of the intensity of default and the loss given default: probability times severity of default. We use interchangeably credit risk and credit spreads to refer to this product. The credit risk of each firm is decomposed into different credit risk sources, some common to all firms and others idiosyncratic to each one. Each credit risk source is represented by an unobservable Gaussian risk factor whose realization is estimated through a filtering mechanism from the market prices of the firms' bonds. Simultaneously, the impact of the factors on each firm's credit risk is obtained, completing the credit risk correlation structure across firms. Our estimation procedure allows us to have a clear economic interpretation of each of the risk factors.

Among the common credit risk factors, we consider two factors extracted from the evolution of the term structure of interest rates, which are found to represent its long-term rate and slope. Our results show that for a given long-term interest rate, a higher slope of the term structure of interest rates implies higher credit spreads for all firms; and that for a given slope, higher long-term rates imply lower credit spreads. However, as we argue below, the importance of these two risk factors on the firms' credit risk is, for most firms, relatively small when compared with an alternative common risk factor.

Apart from the previous two common factors extracted from the term structure of interest rates, we also include another risk factor affecting the credit risk of all firms in the sample. These three-common-factors structure represents the first (of two) considered specifications to capture default correlations. The three common factors not only explain more than 96% of bond prices, they also account for most of the firms' credit risk (between 43 and 89% with an average of 72% across firms).

Additionally to the three common risk factors we include a risk factor affecting the firms' credit risk in the same sector of activity and a firm idiosyncratic risk factor independent across firms. Finally, we perform a regression analysis of the impact of several liquidity variables on our measure of the firms' credit risk, finding that bond liquidity measures have a positive, although weak, impact on the evolution of credit risk.

Even though the previous three-common-factors structure captures the impact of interest rates on credit risk, a more natural question to evaluate the importance of correlations might be: How much credit risk can we explain with the simplest possible correlation structure? For that purpose, we re-estimate the model considering just one common factor, without the two interest rates risk factors. We obtain that such a single common factor explains more than 95% of bond prices and between 15 and 91% of the firms' credit risk. Moreover, it is able to explain, except in two cases, more than 50% of the firms' credit risk. Its estimated realization resembles the evolution of several variables directly related with credit risk levels in the US economy (credit spreads over government bond yields, profit warnings, number of defaults, downgrade ratio, ...) and it is negatively correlated with major US stock indexes.

The previous results imply that credit risk correlations matter and should be accounted for in any credit risk model. The diversification of portfolios completely invested in US corporate debt is questionable, and credit risk correlation is key in pricing multiname credit derivatives whose underlying firms are all US corporates.

The period we analyze, from July 2001 to November 2003, is particularly interesting regarding the evolution of credit risk levels in the US economy because it includes a whole credit risk cycle with its peak around October/November 2002.

This paper is closely related to other CID models such as Duffee (1999) and Driessen (2005). Duffee (1999) estimates a CID model in which firms' default intensity rates only depend on a firm idiosyncratic factor and on two common risk factors extracted from the term structure of interest rates. As long as the firms' credit risk depend on common factors different from the interest rates factors, Duffee's specification is not able to capture all the correlation between firms' default probabilities. Xie, Wu and Shi (2004) estimate Duffee's model for a sample of US corporate bonds and perform a careful analysis of the model pricing errors. A principal component analysis reveals that the first factor explains most of the variation of pricing errors. Regressing bond pricing errors on several macroeconomic variables, the authors find that returns on the S&P 500 stock index explain a significant part of their variation. Therefore, Duffee's model leaves out some important aggregate factors that affect the credit risk of all firms. It seems reasonable to expand existing intensity models to include, as in our specification, economy wide common risk factors.

Driessen (2005) proposes a model in which the firms' default intensity rate is a linear function of two common factors, two factors derived from the term structure of interest rates, a firm idiosyncratic factor and a liquidity factor. Driessen focuses on liquidity and tax effects on the firms' intensities, as well as on the estimation of default event risk premiums. While the author considers that all firms with the same rating are affected in the same way by common factors, we allow for the effect of each common factor to differ across firms, which increases the flexibility of the credit risk correlation structure.

Related CID papers include Bakshi, Madan and Zhang (2006), Janosi, Jarrow and Yildirim (2002) and Zhang (2003). In contrast to the previous CID papers, which estimate default intensity rates using defaultable instruments market data, Chava and Jarrow (2004) estimate a CID model using data on US firms' bankruptcies from 1962 to 1999.

There exist a number of empirical papers using duration analysis on large databases of firms' rating histories providing evidence of the importance of common risk factors in firms' default probabilities. However, they do not successfully quantify such importance. Crowder, Davis and Giampieri (2005) argue that although common factors are a key component of portfolio credit risk, their quantification is very controversial. Couderc and Renault (2005) reckon that there is a need for additional research to quantify the effects of common factors on credit risk and to analyze their impact on risk management models. Our paper represents one of the very first attempts to answer the previous queries.

Previous studies have also pointed out the existence of a common macroeconomic risk factor driving most of the firms' credit risk. Moody's Investors Service (1997), Nickell et al. (2000), Das et al. (2005), De Servigny and Renault (2002) and Koopman, Lucas and Monteiro (2005) document the strong impact of macro-factors on default probabilities; Basel Committee on Banking Supervision (2005) and Covitz and Han (2005) find similar evidence for losses given default.

The rest of the paper is organized as follows. The second section presents the characteristics of the model specification we use to introduce credit and interest rate dependencies in the determinants of the firms' bond prices. The third section describes the dataset and the estimation procedure. Finally, the fourth section presents the results and section five concludes.

2 Model

We denote the physical and risk neutral probability measures as $\tilde{\mathbf{P}}$ and $\mathbf{P}$ respectively, the default-free short-rate as r_t, and assume an arbitrage free market. For our purposes we shall use the class of equivalent probability measures $\mathbf{P}$ where non-dividend paying asset processes discounted with the risk free interest rate r_t are $\mathbf{P}$-martingales. Unless otherwise stated, all probabilities and expectations are taken under the risk neutral measure $\mathbf{P}$.

Consider the case of a single firm first. Default will be modelled as the first arrival of a Poisson process with a given intensity rate λ_t, which represents the instantaneous default probability of a firm which has not defaulted before t. If τ denotes the (future) default time:

$$\lambda_t = \lim_{h \to 0} \frac{\mathbf{P}\left[\tau \in (t, t+h] \mid \tau > t\right]}{h}.$$

Reduced-form models rely on the dynamics of the intensity λ_t to capture the evolution of the firms' default risk.

Consider a defaultable zero-coupon bond with maturity T and face value 1 which, in case of default at τ generates a recovery payment R_τ, and whose price is free of any type of risk different from interest or credit risks (e.g. liquidity risk). Its price at time $t \leq T$ is given by

$$Q(t,T) = E_t\left[e^{-\int_t^T r_s ds} Q(T,T)\right] = E_t\left[e^{-\int_t^T r_s ds} \mathbf{1}_{\{\tau > T\}}\right] + E_t\left[e^{-\int_t^\tau r_s ds} R_\tau\right]. \quad (1)$$

$\mathbf{1}_{\{\cdot\}}$ represents the indicator function.

To model the recovery payment R_τ, we use the so-called Recovery of Market Value RMV specification.[1] RMV assumes a recovery payment equal to an exogenous fraction of the market value of the bond just before default. Duffie and Singleton (1999) show that using RMV a defaultable zero-coupon bond (1) can be priced as if it was default-free by replacing the usual short-rate r_t with a default adjusted

[1] Lando (2004, Chapter 5) presents a detailed analysis of existing recovery assumptions.

short-rate π_t :

$$Q(t,T) = E_t\left[\exp\left(-\int_t^T \pi_s ds\right)\right], \tag{2}$$

where $\pi_t = r_t + \lambda_t L_t$. L_t is the expected loss rate in the market value if default occurs at t, conditional on the information available up to t.[2]

We refer to credit risk as the product of default probability and expected loss, $\lambda_t L_t$, because it measures not only the probability of default but also its severity. We denote $s_t = \lambda_t L_t$ and we refer to it as credit spread. If $s_t = 0$, (2) represents the price of a default-free zero-coupon bond.

Reduced form models are able to estimate the spread s_t implied in bond prices, but not to differentiate the part coming from default risk λ_t from the part coming from the loss given default L_t. A common practice is to assume an exogenously given recovery rate based on historical data. The advantage of using a RMV assumption is that (although it does not solve the identification problem) it does not require an exogenous recovery rate. We shall estimate s_t directly.

Next, consider the case with several firms. Assume J different firms and I independent unobservable Markov processes $X_{1,t}, ..., X_{I,t}$ representing credit risk factors. The default-free rate r_t and credit spreads $s_{1,t}, ..., s_{J,t}$ will be expressed as linear functions of the previous Markov processes. We use unobservable factors because, as shown by Collin-Dufresne, Goldstein and Martin (2001), observable financial and economic variables which should in theory determine credit spread changes have a small explanatory power. See also Blanco, Brennan and Marsh (2005) and Couderc and Renault (2005).

The risk factors $X_{1,t}, ..., X_{I,t}$ follow independent Vasicek processes:

$$dX_{i,t} = \kappa_i (\theta_i - X_{i,t})\, dt + \sigma_i d\tilde{W}_{i,t}, \qquad \text{for } i = 1, ..., I,$$

where $\tilde{W}_{1,t}, ..., \tilde{W}_{I,t}$ are independent $\tilde{\mathbf{P}}$-standard Brownian motions. θ_i, κ_i, and σ_i are the mean reversion level, mean reversion rate and absolute instantaneous volatility

[2] The recovery rate is just one minus the loss given default.

of $X_{i,t}$ respectively. The market price of risk for each factor $X_{i,t}$ is given by $\eta_{i,t} = \xi_i + \gamma_i X_{i,t}$.

In our first three-common-factors specification, we express the interest rate r_t and the different firms' credit spreads $s_{1,t}, ..., s_{J,t}$ as

$$r_t = X_{1,t} + X_{2,t}, \tag{3}$$

$$s_{j,t} = a_{1,j}X_{1,t} + a_{2,j}X_{2,t} + a_{3,j}X_{3,t} + a_{4,j}X_{4,t} + X_{5,j,t}, \tag{4}$$

for each firm $j = 1, ..., J$. $X_{1,t}$ and $X_{2,t}$ represent two unobservable factors affecting both the default-free interest rate and the credit spread $s_{j,t}$ of every firm and will be derived from the term structure of interest rates. $X_{3,t}$ and $X_{4,t}$ are unobservable factors which affect the credit spreads $s_{j,t}$ of all firms and all firms in the same sector (or industry) respectively. $X_{5,j,t}$ is different for each firm j and represents an idiosyncratic or firm specific credit risk factor.

Appendix A derives closed form solutions for bond prices, both default-free and defaultable, as functions of the risk factors and the parameters of the model.

The firms' credit risk correlation structure is given by the dependence of their credit spreads $s_{j,t}$ on the common factors $X_{1,t}$, $X_{2,t}$, $X_{3,t}$ and $X_{4,t}$. The coefficients $a_{1,j}$, $a_{2,j}$, $a_{3,j}$ and $a_{4,j}$, constant across time and different for each firm j, capture the impact of the different risk factors on the firms' credit spreads.

The literature on credit risk correlation has criticized the CID approach arguing that it generates low levels of correlation when compared with empirical ones. However, Yu (2005) suggests that this apparent low correlation is not a problem of the approach itself but a problem of the choice of state or latent variables, owing to the inability of a limited set of state variables to fully capture the dynamics of changes in default intensities. The problem of low correlation in these models may arise because of the insufficient specification of the common factor structure, which may not capture all the sources of common variation in the model, leaving them to the idiosyncratic component, which in turn would not be independent across firms. As shown by Driessen (2005), as one increases the number of common factors, the correlation among the firms' idiosyncratic factors goes to zero. Our specification of the firms' credit risk (4) tries to capture all credit risk correlation through the common risk factors, providing them with a clear economic interpretation.

3 Data and estimation procedure

The dataset consists of daily data covering the period from July 25, 2001 to November 20, 2003. From the Federal Reserve Board interest rates dataset we have available zero-coupon rates of US government securities for six different maturities, which we use to estimate the two first factors of the model, $X_{1,t}$ and $X_{2,t}$.[3]

The second source of data consists of Bloomberg ask, bid and mid bond prices of 33 bonds from 14 different US firms: Owens, CSC Holdings, Comcast Cable, Walt Disney, Ford Motor Credit, General Motors (GM), Hertz Corporation, Marriott, Norfolk Southern, Sears Roebuck, TCI Communications, Time Warner, USX Corporation, and Union Pacific Corporation. For each firm we have two to three bonds. The firms and bonds in our sample satisfy the requirements contained in Appendix B. All but two firms have an S&P rating of BBB at the end of the sample period; Owens is rated B and CSC is rated BB at the end of the period.

The estimation strategy is divided in two steps. First, and using only the data of the US default-free interest rates, we estimate the risk factors $X_{1,t}$ and $X_{2,t}$. The second step uses the corporate bond dataset and the already estimated risk factors $X_{1,t}$ and $X_{2,t}$ to estimate the rest of the risk factors and model parameters. In the same way, the estimation of the common, sector and idiosyncratic factors from the corporate bonds dataset is subdivided in successive steps.[4] The estimation strategy shall guarantee that the interpretation of the different risk factors coincide with the one given in the previous section.

[3] Our results also hold using USD swap rates as default-free rates.

[4] Although directly estimating all risk factors and parameters of the model seems a natural approach, it is not actually practical when many firms and factors are considered. The number of parameters involved increases rapidly and quickly becomes unmanageable. As a consequence, we break down the estimation of the model in successive steps.

We assume that the credit spread $s_{j,t}$ is similar to all bonds of each firm j, which allows us to use several bonds of the same firm to estimate the model. The model implied pricing errors will be bond specific.

3.1 Interest rates risk factors

Since we have assumed the default-free short-rate r_t to be the sum of two independent factors $X_{1,t}$ and $X_{2,t}$, we can express the price of the default-free zero-coupon bond at time t, with maturity T and face value 1, as

$$Q_{df}(t,T) = E_t\left[\exp\left(-\int_t^T (X_{1,s} + X_{2,s})\, ds\right)\right], \tag{5}$$

for which we have a closed form solution. The zero-coupon rate from time t to time T is given by

$$y_{df}(t,T) = -\frac{\ln Q_{df}(t,T)}{T-t}. \tag{6}$$

The estimation strategy of the two risk factors extracted from the term structure of interest rates builds on Bobadilla (1999) and consists of a Maximum Likelihood estimation of $X_{1,t}$ and $X_{2,t}$ using a Kalman Filter (KF) procedure.[5] Zero-coupon rates $y_{df}(t,T)$ depend linearly on the risk factors $X_{1,t}$ and $X_{2,t}$, which allows for a direct application of the KF. The state space model for the KF estimation procedure as well as the discretization of a Vasicek process can be found in Appendices C and D.

In order for $X_{1,t}$ and $X_{2,t}$ to contain information about the whole term structure of interest rates, we use data of $y_{df}(t,T)$ for six different maturities T to estimate them (3 months, 6 months, 1 year, 3 years, 5 years, and 20 years). The standard deviation of the pricing errors for the risk-free rates $y_{df}(t,T)$, denoted by σ_ε, is common for the different maturities T.

[5] A classical reference on Kalman filtering is Harvey (1989). See also Jegadeesh and Pennacchi (1996).

3.2 Credit risk factors

We use the estimated risk factors $X_{1,t}$ and $X_{2,t}$ as inputs in the estimation process to obtain the other risk factors.[6] Since, as shown in Appendix A, coupon bond prices do not depend linearly on the risk factors we use an extended KF (EKF) procedure to estimate each credit risk factor $X_{3,t}$, $X_{4,t}$ and $X_{5,j,t}$. See Appendix E for details.

First, using the bonds of all firms we estimate the common factor $X_{3,t}$, as well as the parameters $a_{1,j}$, $a_{2,j}$ and $a_{3,j}$ for each firm j. That is, in this step of the estimation we consider that for each firm j

$$s_{j,t} = a_{1,j}X_{1,t} + a_{2,j}X_{2,t} + a_{3,j}X_{3,t}, \tag{7}$$

where the factors $X_{1,t}$ and $X_{2,t}$ are already known. We assume pricing errors implied by the previous specification of $s_{j,t}$ are independent across bonds (within and across firms) and normally distributed with zero mean and standard deviation $\sigma_{\varepsilon,j}$ (different for each firm j but common for all bonds of each firm). We keep this assumption for the estimation of the rest of the risk factors throughout the paper.

The transition equation of the state space model is a discretized version of the Vasicek process followed by the common factor $X_{3,t}$ (see Appendix D). In this case, not all the parameters in the model are identified and, in order to identify the model, we normalize the value of the instantaneous volatility of the common risk factor σ_3 (see Appendix F). The identification problem is also present when estimating the sector risk factors $X_{4,t}$, for which we also normalize the value of σ_4, but not in the estimation of the idiosyncratic risk factor $X_{5,j,t}$ for each firm j.

Next, we estimate the realization of the sector risk factor $X_{4,t}$, the parameters governing its dynamics, and the parameter $a_{4,j}$ for each firm j. In this estimation step, we consider that for each firm j the credit spread is given by

$$s_{j,t} = a_{1,j}X_{1,t} + a_{2,j}X_{2,t} + a_{3,j}X_{3,t} + a_{4,j}X_{4,t}, \tag{8}$$

[6] The estimated realization of all risk factors corresponds to the KF unsmoothed filtered processes: the realization of the factors at time t is estimated using information up to time t.

where the factors $X_{1,t}$, $X_{2,t}$ and $X_{3,t}$, as well as the parameters $a_{1,j}$, $a_{2,j}$ and $a_{3,j}$ are taken as given from previous estimation steps. Firms within the same sector of activity will be affected by the same sector factor, but firms in different sectors of activity will be affected by different sector factors. Although each firm is only affected by (at most) one sector risk factor, we will be estimating several different sector risk factors.

The inclusion of a sector-specific credit risk factor is supported by previous empirical evidence. Chava and Jarrow (2004) demonstrate the importance of including industry effects in default rate estimation in an empirical paper which investigates the forecasting accuracy of bankruptcy hazard rate models for US economy over the time period 1962-1999. Lucas (1995) enumerates time periods in the recent US history in which firms belonging to the same industry experience credit problems simultaneously. Crowder, Davis, and Giampieri (2005) and Couderc and Renault (2005) present further evidence.

In order for $X_{4,t}$ to have the interpretation of a risk factor which only affects firms in the same sector, we estimate it using data for the firms in the same sector.[7]

Finally, we estimate the realization of the idiosyncratic risk factor $X_{5,j,t}$, as well as the parameters governing its dynamics, for each firm j, using only bonds of firm j and taking as given all previously estimated risk factors and parameters.

Our step-by-step estimation method implies that in each stage, the parameters estimated in previous stages are taken as given, i.e. as if they were the true parameters of the model, which will (artificially) reduce the standard errors of the estimated parameters.

[7] The fact that we estimate the common risk factor $X_{3,t}$ assuming that pricing errors are independent across firms and then we introduce a sector risk factor $X_{4,t}$ is, technically speaking, inconsistent. The reason is that the sector risk factors are introducing correlation in the pricing errors left from the estimation of the common risk factor $X_{3,t}$. Using as initial values the parameters estimated with the proposed sequential method for the two estimation stages for $X_{3,t}$ and $X_{4,t}$, we proceeded to estimate the previous risk factors, the parameters governing their dynamics, as well as the parameters $a_{1,j}$, $a_{2,j}$, $a_{3,j}$ and $a_{4,j}$ in a single estimation step, assuming the pricing errors are independent across firms. The results, in terms of pricing accuracy, estimation of the realization of the risk factors $X_{3,t}$ and $X_{4,t}$, and the parameters $a_{1,j}$, $a_{2,j}$, $a_{3,j}$ and $a_{4,j}$, do not significantly vary with respect to the sequential estimation approach, which is significantly less time consuming.

4 Results

Table 1 shows the estimates and standard errors of the parameters of the two first risk factors $X_{1,t}$ and $X_{2,t}$ extracted from the term structure of interest rates. The first factor $X_{1,t}$ is highly and positively correlated (0.97) with the long-term (20 years) zero-coupon interest rate, whereas the second factor $X_{2,t}$ is negatively correlated (-0.57) with the slope of the term structure of interest rates (difference between 20 years and 3 months zero-coupon rates). This result is common in the literature, and indicates that we can decompose the whole information of the term structure of interest rates in two risk factors (see Duffee 1999 and literature therein, and Zhang 2003.) Table 2 contains the (in sample) mean and mean absolute pricing errors for the zero-coupon interest rates (6) for each of the six maturities considered using the realizations of $X_{1,t}$ and $X_{2,t}$.

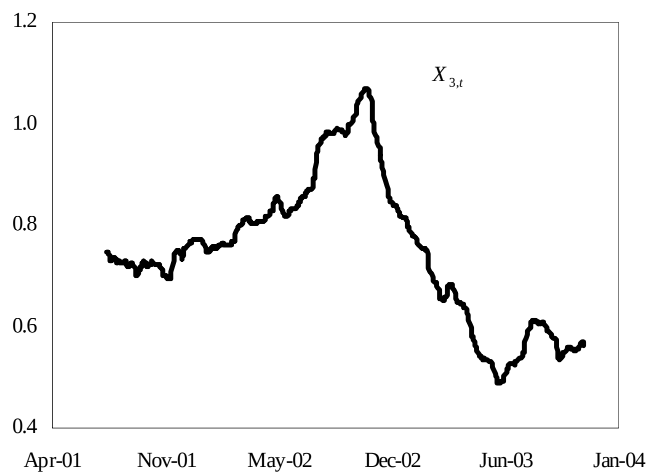

Figure 1: Estimated realization of $X_{3,t}$.

Table 3 presents the estimated values of the parameters governing the dynamics of $X_{3,t}$, and Figure (1) represents its estimated realization, showing that it reaches maximum levels around October/November 2002. Table 4 contains the values and standard deviations of the estimated parameters for $a_{1,j}$, $a_{2,j}$ and $a_{3,j}$ for each firm j.

The common risk factor $X_{3,t}$ affects the credit risk of all firms positively, although in different degrees.

Table 5 presents, among other things, the average value of $a_{1,j}X_{1,t} + a_{2,j}X_{2,t}$ for each firm during the whole sample period. The effect of the interest rate risk factors $X_{1,t}$ and $X_{2,t}$ on credit spreads $s_{j,t}$ is, on average and during the sample period analyzed, positive for high yield firms becoming negative as we increase the credit rating of the firm. This pattern is observed in all BBB rated firms except Ford and Hertz.

The signs of the parameters $a_{1,j}$ and $a_{2,j}$ imply that for a given long-term interest rate, a higher slope, i.e. a lower short-term rate, implies higher levels of the spread $s_{j,t}$; and, for a given slope, higher long-term rates imply lower levels of $s_{j,t}$. Empirical results about the impact of the term structure of interest rates on firms' credit spreads are mixed. In a comprehensive study using Standard and Poor's rating histories of over 10000 companies during the period 1981-2003, Couderc and Renault (2005) find that the level and slope of the term structure of interest rates affect default intensities in the very same way as we do. Bakshi, Madan and Zhang (2006) and Janosi, Jarrow and Yildirim (2002) find a positive impact of short rates on credit spreads. Driessen (2005) finds a negative, although small, relationship between spreads and default-free rates. Similarly, Zhang (2003) finds that (for Argentina) the default intensity rate is positively related to both the slope and the level at the long end of the term structure of interest rates.[8]

For the estimation of the sector risk factor $X_{4,t}$ one needs to group the firms in sectors of activity. The grouping serves to study the risk dependence structure between the firms included in the same sector. We divide the firms into three sectors: Railroad (Norfolk and Union Pacific), Cable TV (CSC and Comcast) and Film (Disney and Warner). For the rest of the firms we do not estimate any sector risk factor.

Table 6 shows the estimates for the parameters of the three sector risk factors and

[8] From a theoretical perspective, Repullo and Suarez (2000) analyze the effects of changes in the short term interest rate on credit risk.

Table 7 shows the estimated values of $a_{4,j}$. While both railroad firms are affected positively by the Railroad credit risk factor, the effect of the Cable TV and Film risk factors has a different sign for each firm within those sectors.

Finally, we estimate the idiosyncratic risk factor $X_{5,j,t}$ for each firm j. Table 5 shows the average impact of each risk factor on the firms' credit spread, as well as the average estimated credit spread $s_{j,t}$ and the firms' credit rating at the end of the sample period. As expected, the average level of $s_{j,t}$ is higher for the two firms with lower ratings, Owens and CSC.

The average cross-firm correlation of the idiosyncratic factors is 0.13, which indicates that common factors are able to capture the credit risk dependence structure, and that the idiosyncratic factors can be considered independent across firms.

Table 8 presents the model mean absolute pricing errors (expressed as percentage of bonds prices) in each estimation step. The three common risk factors $X_{1,t}$, $X_{2,t}$ and $X_{3,t}$ which affect all firms in the sample are enough to explain, on average, more than 96% of the firms' bond prices.[9] Including all risk factors, the model is able to explain more than 98% of the bond prices of all firms, which indicates that the contribution of the sector and idiosyncratic credit risk factors is limited (except for Union Pacific) compared with the contribution of the other three common risk factors in explaining bond prices.

Table 9 presents, by firm, the average percentage of bond prices not explained by the default free term structure of interest rates which is explained when we include the different risk factors.[10] If liquidity and tax considerations were nonexistent,[11] the values in Table 9 could be interpreted as the percentage of credit risk explained by the model when we include different risk factors in the firms' credit spreads $s_{j,t}$. The

[9] The percentage of the firms' bond prices explained by the risk factors is computed as 100 minus the average mean absolute pricing error.

[10] In order to compute the part of bond prices explained by the default free term structure of interest rates, we price the bonds assuming $r_t = X_{1,t} + X_{2,t}$ and $s_{j,t} = 0$. This corresponds to the case in which bonds have been priced as if they were default-free bonds using the estimation results for the two-factor affine model for the risk free rate given in Table 1.

[11] See Amato and Remolona (2003, 2005) and literature therein for an analysis of the importance of the different components of the credit spread.

three common risk factors ($X_{1,t}$, $X_{2,t}$ and $X_{3,t}$) account for most of the firms' credit risk, between 43 and 89% with an average and standard deviation across firms of 72% and 14% respectively.

In our analysis so far, we have considered the joint impact of the three common risk factors $X_{1,t}$, $X_{2,t}$ and $X_{3,t}$ on the explanation of bonds' prices and firms' credit risk. The reason is that we have estimated their impact on the firms' bond prices in a single estimation step in which the realization of the common risk factor $X_{3,t}$ is also derived. In order to discern the relative importance of $X_{3,t}$ versus $X_{1,t}$ and $X_{2,t}$ we estimate the alternative specification in which, instead of three common factors as in (7), we consider a single common factor:

$$s_{j,t} = b_j Z_t, \tag{9}$$

where Z_t is a Vasicek unobservable risk factor. This alternative one-common-factor specification of the credit spread $s_{j,t}$ only incorporates one common credit risk factor and does not consider the impact on the firms' credit spreads of the two risk factors $X_{1,t}$ and $X_{2,t}$ extracted from the term structure of interest rates. A single common risk factor for credit spreads represents the simplest possible credit risk correlation structure.[12] Tables 12 and 13 present, respectively, the model mean absolute pricing errors and the percentage of bond prices not explained by the risk free interest rate which is explained when we consider (9). These results can be compared to the ones in Tables 8 and 9 when $s_{j,t}$ is specified as in our benchmark case (7).

The percentage mean absolute pricing errors obtained with a single common factor (9) are of the same order of magnitude than the ones when $s_{j,t}$ includes three factors (7); higher for some firms and lower for others, but below 5% in all cases. In terms of the percentage of firms' credit risk explained, the results in both cases are close on average (72 vs. 68% respectively). However, there are significative differences for some firms. In particular, for Comcast, Marriott, Norfolk, TCI and Warner the

[12] See Tables 10 and 11 for the estimation of the parameters governing the dynamics of Z_t and of the parameters b_j for each firm j respectively.

percentage of the credit risk explained by (7) is significantly higher than the one explained by (9), indicating the importance of the two factors $X_{1,t}$ and $X_{2,t}$ extracted from the term structure of interest rates. The reverse is true for Disney.

In any case, with the alternative specification (9) which considers just one common credit risk factor the model is able to explain between 15 and 91% of the firms' credit risk, with an average across firms of 68% and a standard deviation of 21%. A single credit risk factor is able to explain, except in two cases, more than 50% of the firms' credit risk.

Figure (2) compares the estimated realization of the risk factors $X_{3,t}$ and Z_t obtained with the two previous specifications of the credit spread respectively. It is clear that the estimated realization of the credit risk factor Z_t resembles the evolution of $X_{3,t}$; their correlation coefficient is 0.53.

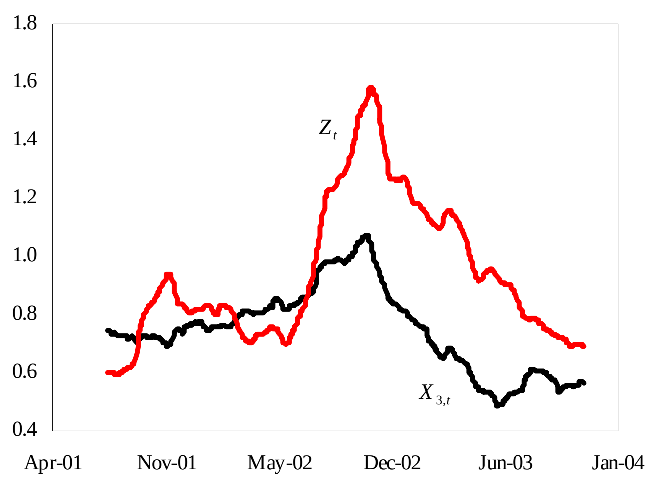

Figure 2: Estimated realization of the common risk factors $X_{3,t}$ and Z_t when the specification of the firms' credit spreads is $s_{j,t} = a_{1,j}X_{1,t} + a_{2,j}X_{2,t} + a_{3,j}X_{3,t}$ and $s_{j,t} = b_j Z_t$ respectively.

Figure (3) presents the evolution of the common risk factor Z_t together with that of the two major stock indexes in the US: S&P 500 and Dow Jones. The dynamics of Z_t are inversely related to those of the S&P 500 and Dow Jones, with correlation coefficients of -0.76 and -0.77 respectively.

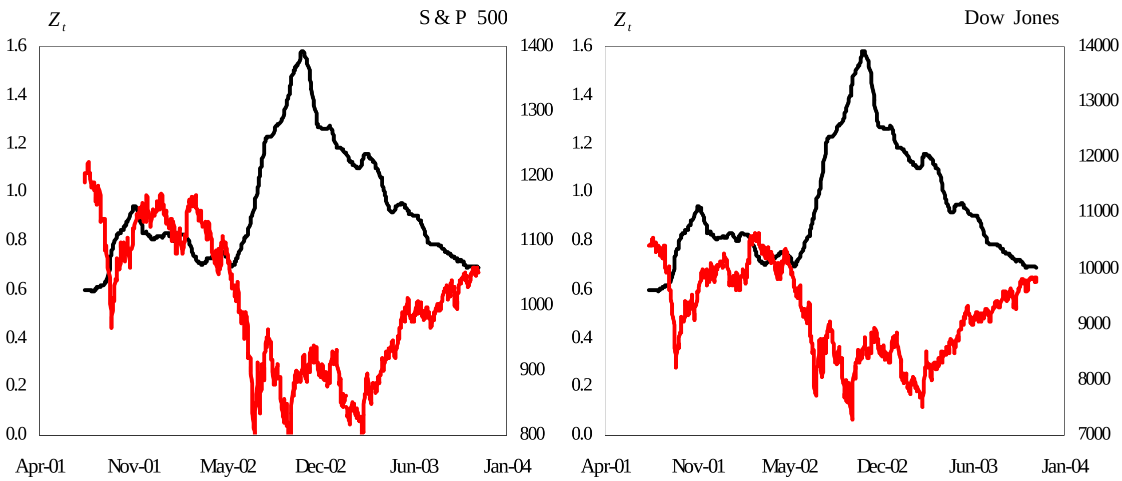

Figure 3: Evolution of the common risk factor Z_t, S&P 500 and Dow Jones. Source: Yahoo finance (close prices adjusted for dividends and splits).

The evolution of credit risk factor Z_t (and hence the one of $X_{3,t}$) also resembles that of several variables directly related with credit risk levels in the US economy (profit warnings, number of defaults, downgrades as a percentage of all rating actions, and KMV default probabilities for the US economy; see Figures 4 and 5.) Therefore, the common risk factor explaining most of the credit risk of the firms in our sample seems to be common to all firms in the US economy, which suggests that the results in this paper can probably be extrapolated to other US corporates.

Pricing errors and liquidity analysis

Within the components $s_{j,t}$ might include, apart from credit risk, one can think of liquidity risk. Our dataset allows us using bid-ask spreads as a measure for the bonds liquidity to assess the impact of liquidity on both the spreads $s_{j,t}$ and pricing errors.

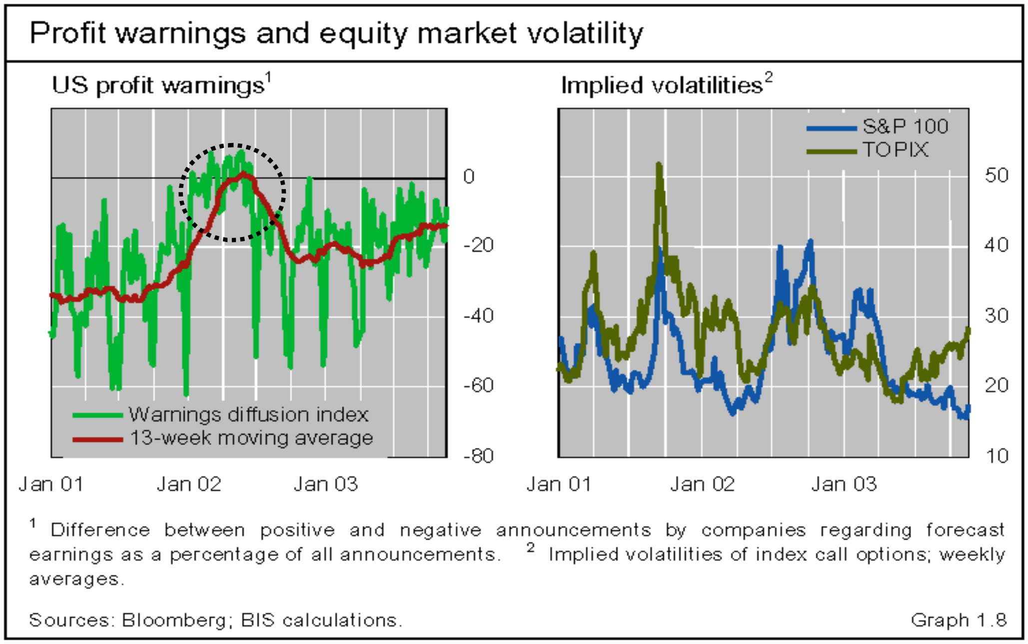

Figure 4: BIS Quarterly Review, 2003, December, Graph 1.8. Dotted line added.

Several other studies have tried to discern to what extent liquidity considerations are able to explain bond credit spreads. Longstaff, Mithal and Neis (2005) estimate the default intensity rate of a firm using data of bonds and credit default swaps (CDS). Substracting from the total bond corporate spread the estimated default component which matches CDS premiums they obtain (what they term) the non-default component. They find that the majority of corporate spread is due to default risk and that the nondefault component is time varying and positively related to several liquidity measures. Driessen (2005) uses a liquidity common factor based on bond age and finds a positive dependence of credit spreads on the liquidity factor.

For each firm, we construct daily liquidity spreads as the average ratio between its bonds' price bid-ask spreads and the bond mid price (Table 14). Except for Hertz, higher liquidity spreads are found in low rated firms: Owens and CSC.

The analysis that follows is carried out for the three-common-factor specification of spreads (4). Table 15 presents the results of two different specifications regressing,

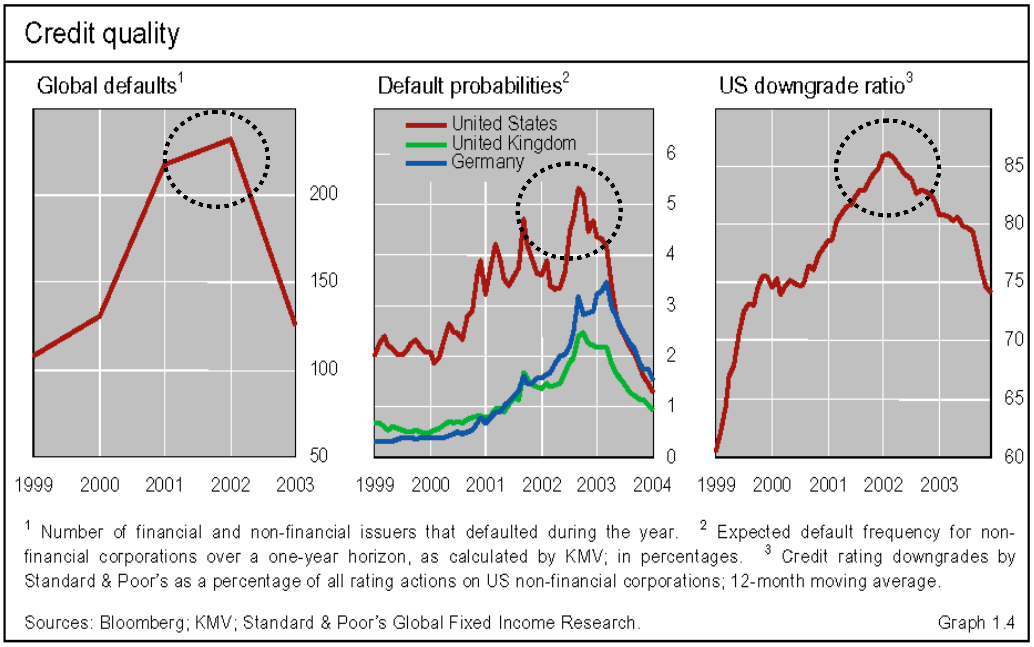

Figure 5: BIS Quarterly Review, 2004, March, Graph 1.4. Dotted lines added.

by bond, credit spreads $s_{j,t}$ with respect to ratings (at the end of the sample), liquidity spreads and other possible measures of liquidity (time to maturity, age and notional amount issued).

When we include all variables (column 2) we obtain an R^2 of 16%, that liquidity spreads positively affect credit spreads $s_{j,t}$, and that, as expected, a higher rating implies a higher credit spread. The coefficients for the rest of the liquidity measures are not significant. If we take out the rating and leave just liquidity variables in the regression (column 4), the R^2 is reduced to 7%, indicating the low impact of liquidity measures on the credit spreads $s_{j,t}$ implied by the model. Therefore, although weak, we find a positive relationship between credit and liquidity spreads.

In order to analyze the model pricing errors we regress, by bond, absolute pricing errors (as a percentage of bond prices) on time to maturity, age, notional amount issued, rating, liquidity spread, credit spread $s_{j,t}$, common risk factor $X_{3,t}$ and idiosyncratic risk factor $X_{5,j,t}$. As before, see Table 16, the R^2 of the regression is low, 11%, and we find that higher liquidity spreads and higher credit spreads $s_{j,t}$ imply higher pricing errors.

5 Conclusion

This paper studies the importance of credit risk correlations in bond market prices. By decomposing the firms' credit spreads on different credit risk factors the model is able to compute the importance of each credit risk factor on the evolution of the firms' credit risk, identifying their credit risk correlation structure.

The main result is the high portion of credit risk explained by a very limited number of common risk factors, which shows that credit risk is, to a very large extent, driven by common factors affecting all firms. Therefore, credit risk correlations matter.

The firms' credit risk correlation structure mainly stems from a single risk factor which we found positively related to other credit risk measures in the US economy and negatively related with major US stock indexes. Other factors which potentially can affect the firms' credit risk, such as the evolution of the term structure of interest rates or factors associated with the credit risk of the firms' sector of activity, are found to play a lower role in credit risk correlations. As a consequence, our results support evidence in Couderc and Renault (2005) suggesting that while interest rates seem to be unhelpful explanatory covariates of default across rating classes, stock market information contains significant explanatory power.

In order to test the robustness of our results and derive implications about the credit risk correlation structure of firms in different credit rating categories, the model should be estimated using firms representing all the spectrum of credit risk ratings. In the same way, the sample period analyzed is one in which credit risk levels in the US reach their highest values in, at least, the last 10 years. Testing the model in periods with less extreme credit levels is also a natural extension for further research.

Appendix

A Closed form solutions for bond prices

Consider a state variable X_t which follows a Vasicek process under the physical probability measure $\tilde{\mathbf{P}}$:

$$dX_t = \kappa (\theta - X_t) dt + \sigma d\tilde{W}_t,$$

where $\tilde{W}_t$ is a $\tilde{\mathbf{P}}$-Brownian motion. If η_t is its market price of risk and $dW_t = d\tilde{W}_t + \eta_t dt$, then W_t is a $\mathbf{P}$-Brownian motion. Assuming $\eta_t = \xi + \gamma X_t$, the dynamics of X_t under the risk neutral probability measure $\mathbf{P}$ are given by

$$dX_t = (\kappa (\theta - X_t) - (\xi + \gamma X_t) \sigma) dt + \sigma dW_t.$$

Denote by $Q(t, T; X_t)$ the price (2) at time $t \leq T$ of a zero-coupon bond with maturity T and face value 1. Consider the particular case where $\pi_t = X_t$:

$$Q(t, T; X_t) = E_t \left[\exp \left(- \int_t^T X_s ds \right) \right].$$

Using the Feynman-Kac theorem, it can be shown that the bond price Q follows the partial differential equation (PDE)

$$\frac{\partial Q(t, T; x)}{\partial t} + \frac{1}{2}\sigma^2 x \frac{\partial^2 Q(t, T; x)}{\partial x^2} + (\kappa (\theta - x) - (\xi + \gamma x) \sigma) \frac{\partial Q(t, T; x)}{\partial x} - xQ(t, T; x) = 0,$$

with boundary condition

$$Q(T, T; x) = 1, \qquad \forall x.$$

The solution of the previous PDE is given (Moreno 2003) by:

$$Q(t, T; X_t) = G(t, T) \exp(-F(t, T) X_t),$$

where

$$\begin{aligned}
G(t, T) &= \exp \left(-\frac{\sigma^2}{4q} F(t, T)^2 + \chi (F(t, T) - (T - t)) \right), \\
F(t, T) &= \frac{1}{q} \left[1 - e^{-q(T-t)} \right], \\
q &= \kappa + \gamma \sigma, \\
\chi &= \omega - \frac{\sigma^2}{2q^2}, \\
\omega &= \frac{(\kappa \theta - \xi \sigma)}{q}.
\end{aligned}$$

If π_t is a linear combination of I independent state variables, $\pi_t = e_1 X_{1,t} + ... + e_I X_{I,t}$, the price of a zero-coupon bond (2) can be expressed as

$$\begin{aligned} Q(t,T) &= E_t\left[\exp\left(-\int_t^T (e_1 X_{1,s} + ... + e_I X_{I,s})\, ds\right)\right] \\ &= E_t\left[\exp\left(-\int_t^T e_1 X_{1,s} ds\right)\right] ... E_t\left[\exp\left(-\int_t^T e_I X_{I,s} ds\right)\right] \\ &= Q(t,T;Y_{1,t}) ... Q(t,T;Y_{I,t}), \end{aligned} \tag{A1}$$

where $Y_{i,t} = e_i X_{i,t}$, for $i = 1,...,I$. As in the case in which the interest rate was equal to one state variable, one can use the Feynman-Kac theorem to obtain closed form solutions for $Q(t,T;Y_{1,t})$, ..., $Q(t,T;Y_{I,t})$ because the processes governing the dynamics of $Y_{1,t},...,Y_{I,t}$ are independent Vasicek diffusions.

If we have a coupon bond with maturity T, face value M and semiannual coupon c, its price at t can be computed as $Q(c;t,T)$:

$$Q(c;t,T) = \sum_{n \in N} \frac{c}{2} Q(t,t_n) + MQ(t,T), \tag{A2}$$

where N is the set of all coupon dates. Using the formula (A1) for the closed form solution of zero-coupon bonds we can express the price of a coupon bond as a closed function of the model's parameters and the Vasicek risk factors $X_{1,t},...,X_{I,t}$ which drive the dynamics of the risk adjusted rate π_t. Given the closed form solution (A1) of zero-coupon bonds, it is clear that the price of a coupon bond (A2) is a nonlinear function of the risk factors.

B Bond data filtering process

We identify a firm by the six first digits of its CUSIP (Committee on Uniform Securities Identification Procedures) code, which identifies debt issuers in the US and Canada. The firms and bonds in our sample satisfy the following requirements:

1. Bonds: (i) Maturity between 1 and 35 years. (ii) Face value higher than USD 300 million. (iii) Semiannual fixed coupon rate. (iv) No embedded options (call provisions, put provisions, extendible bonds, convertible bonds, ...). (v) Nonsubordinated seniority: notes, senior notes, senior unsubordinated and unsubordinated bonds. (vi) Sufficiently liquid bond prices, considered as BGN by Bloomberg. BGN prices are a weighted average of quotes submitted by at least five brokers or dealers. The weighting method is proprietary. (vii) Bonds with available ask, bid and mid prices. (viii) Of the firm's bonds with a remaining maturity in the range [n years, $(n+1)$ years), n an integer, all but one bond are dropped. In case we have more than one bonds with the same maturity year, we keep the one with the highest amount outstanding. And in case we have more than one bonds with the same maturity year and amount outstanding, we keep the one with most recent maturity.

2. Firms: (i) Firms with Standard and Poor's rating available at the end of the sample. (ii) Firms with at least two bonds satisfying the previous characteristics.

C Kalman Filter State Space Model: interest rates risk factors estimation

Following Appendix A, the default-free zero-coupon rate (6) can be expressed as a linear function of the two unobservable risk factors $X_{1,t}$ and $X_{2,t}$:

$$y_{df}\left(t,T\right)=-\frac{\ln Q_{df}\left(t,T\right)}{T-t}=-\frac{\ln\left(G_1\left(t,T\right)G_2\left(t,T\right)\right)}{T-t}+\frac{F_1\left(t,T\right)}{T-t}X_{1,t}+\frac{F_2\left(t,T\right)}{T-t}X_{2,t},$$

which implies the following state space model for the KF estimation procedure:

- The measurement equation, which represents the observable variable as a linear function of the unobserved factors, is given by:

$$Y_t^{df}=B+HX_t+\varepsilon_t,$$

  where

  - Y_t^{df} is a vector of zero-coupon interest rates $y_{df}\left(t,T\right)$ for $N=6$ different maturities T,
  - X_t is a vector containing $X_{1,t}$ and $X_{2,t}$,
  - B is an $Nx1$ vector such that each component B_n, which represents a given maturity T_n-t for $n=1,...,N$, is given by

$$B_n=-\frac{1}{T_n-t}\left[\ln G_1\left(T_n-t\right)+\ln G_2\left(T_n-t\right)\right],$$

    where

$$\begin{aligned}
\ln G_i\left(T_n-t\right) &= -\frac{\left(\sigma_i\right)^2}{4q_i}F_i\left(T_n-t\right)^2+\chi_i\left(F_i\left(T_n-t\right)-\left(T_n-t\right)\right)\\
F_i\left(T_n-t\right) &= \frac{1}{q_i}\left[1-e^{-q_i\left(T_n-t\right)}\right],\\
q_i &= \kappa_i+\gamma_i\sigma_i,\\
\chi_i &= \omega_i-\frac{\sigma_i^2}{2q_i^2},\\
\omega_i &= \frac{\left(\kappa_i\theta_i-\xi_i\sigma_i\right)}{q_i},
\end{aligned}$$

    for $i=1,2$.

– H is an $Nx2$ matrix with elements

$$H_{n,i} = \frac{1}{T_n - t} F_i \left(T_n - t\right),$$

for $i = 1, 2$ and $n = 1, ...N$.

– ε_t is an N-dimensional zero-mean normal random variable capturing bond pricing errors. Its covariance matrix is a time invariant NxN matrix $\sigma_\varepsilon^2 I$, where σ_ε is constant and I is the identity matrix.

- The transition equation, which represents the dynamics of the unobservable factors, uses the dynamics of the state variables under the physical probability measure $\tilde{\mathbf{P}}$.

  Using the discretizations of $X_{1,t}$ and $X_{2,t}$ derived in Appendix D, the transition equation can be expressed as

$$X_{t+h} = a + F X_t + \phi_t, \tag{C1}$$

where

$$\begin{aligned} X_t &= \begin{pmatrix} X_{1,t} \\ X_{2,t} \end{pmatrix}, \\ a &= \begin{pmatrix} \theta_1 \left(1 - e^{-\kappa_1 h}\right) \\ \theta_2 \left(1 - e^{-\kappa_2 h}\right) \end{pmatrix}, \\ F &= \begin{pmatrix} e^{-\kappa_1 h} & 0 \\ 0 & e^{-\kappa_2 h} \end{pmatrix}. \end{aligned}$$

h is the time difference, in years, between two consecutive observations, and ϕ_t follows a two dimensional zero-mean normal distribution under $\tilde{\mathbf{P}}$ with covariance matrix

$$\begin{pmatrix} \frac{(\sigma_1)^2}{2\kappa_1} \left(1 - e^{-2\kappa_1 h}\right) & 0 \\ 0 & \frac{(\sigma_2)^2}{2\kappa_2} \left(1 - e^{-2\kappa_2 h}\right) \end{pmatrix}.$$

From Bobadilla (1999), Duffee (1999), Duffie, Pan and Singleton (2000) and Driessen (2005) follows that all the parameters in the specification of the KF for the interest rates can be identified.

D Discretization of a Vasicek process

In the case of a Vasicek process, the dynamics of X_t are explicitly solvable. The solution is an Ornstein-Uhlenbeck process, and can be derived as follows. Consider the process $e^{\kappa t}$, and using the integration by parts formula we apply Itô's lemma to obtain the dynamics of the process $e^{\kappa t}X_t$:

$$\begin{aligned} d\left(e^{\kappa t}X_t\right) &= e^{\kappa t}dX_t + X_t d\left(e^{\kappa t}\right) + dX_t d\left(e^{\kappa t}\right) \\ &= e^{\kappa t}\left(\kappa\left(\theta - X_t\right)dt + \sigma d\tilde{W}_t\right) + X_t\left(\kappa e^{\kappa t}dt\right) \\ &= \theta\kappa e^{\kappa t}dt + e^{\kappa t}\sigma d\tilde{W}_t, \end{aligned}$$

and integrating from $t-h$ to t

$$\begin{aligned} e^{\kappa t}X_t &= e^{\kappa(t-h)}X_{t-h} + \theta\int_{t-h}^{t}\kappa e^{\kappa s}ds + \sigma\int_{t-h}^{t}e^{\kappa s}d\tilde{W}_s \\ &= e^{\kappa(t-h)}X_{t-h} + \theta e^{\kappa t} - \theta e^{\kappa(t-h)} + \sigma\int_{t-h}^{t}e^{\kappa s}d\tilde{W}_s. \end{aligned}$$

Therefore

$$X_t = \theta + e^{-\kappa h}\left(X_{t-h} - \theta\right) + \sigma e^{-\kappa t}\int_{t-h}^{t}e^{\kappa s}d\tilde{W}_s,$$

where, using Itô's isometry $\int_{t-h}^{t}e^{\kappa s}d\tilde{W}_s$ is a square integrable martingale with expectation 0. This implies that, under the physical probability measure $\tilde{\mathbf{P}}$

$$\begin{aligned} \tilde{E}_{t-h}\left[X_t\right] &= \theta + e^{-\kappa h}\left(X_{t-h} - \theta\right), \\ \tilde{V}_{t-h}\left[X_t\right] &= \frac{\sigma^2}{2\kappa}\left[1 - e^{-2\kappa h}\right]. \end{aligned}$$

Thus, we can express X_t as

$$X_t = \theta + e^{-\kappa h}\left(X_{t-h} - \theta\right) + \left(\frac{\sigma^2}{2\kappa}\left[1 - e^{-2\kappa h}\right]\right)^{\frac{1}{2}}\phi_t,$$

where $\phi_t \sim N\left(0,1\right)$ under $\tilde{\mathbf{P}}$.

E Credit risk factors: Extended Kalman Filter

The price (A2) of a coupon bond can be expressed as a linear combination of a set of zero-coupon bonds. Since (A2) does not depend linearly on the risk factors, the measurement equation in the KF state space model is a linearization of that formula around the unobservable risk factor we want to estimate. This estimation technique receives the name of extended KF, EKF. For more detailed descriptions of the EKF see, among others, Harvey (1989) and Bishop and Welch (2003). De Jong (2000) shows the reliability of EKF procedures in affine models through a Monte Carlo experiment.

The transition equation of the state space model does not vary, and, as in the case of the interest rates risk factors KF state space model, is given by (C1).

F Identification

When estimating the realization of the factors $X_{3,t}$ and $X_{4,t}$ one cannot identify all parameters involved. We follow Driessen (2005) to analyze the identification of the parameters when estimating common factors. The contribution of a common factor $X_{i,t}$ to the instantaneous credit spread $s_{j,t}$ of a given firm j is $Y_{i,j,t} = a_{i,j}X_{i,t}$. The processes of $Y_{i,j,t} = a_{i,j}X_{i,t}$ under the physical and risk neutral measures are, respectively:

$$\begin{aligned} d\left(a_{i,j}X_{i,t}\right) &= \left(a_{i,j}\kappa_i\theta_i - a_{i,j}\kappa_i X_{i,t}\right)dt + a_{i,j}\sigma_i d\tilde{W}_{i,t}, \\ d\left(a_{i,j}X_{i,t}\right) &= \left(\left(a_{i,j}\kappa_i\theta_i - \xi_i a_{i,j}\sigma_i\right) - X_{i,t}\left(a_{i,j}\kappa_i + \gamma_i a_{i,j}\sigma_i\right)\right)dt + a_{i,j}\sigma_i dW_{i,t}, \end{aligned}$$

which implies

$$\begin{aligned} dY_{i,j,t} &= \left(a_{i,j}\kappa_i\theta_i - \kappa_i Y_{i,j,t}\right)dt + a_{i,j}\sigma_i d\tilde{W}_{i,t}, \\ dY_{i,j,t} &= \left(\left(a_{i,j}\kappa_i\theta_i - \xi_i a_{i,j}\sigma_i\right) - Y_{i,j,t}\left(\kappa_i + \gamma_i\sigma_i\right)\right)dt + a_{i,j}\sigma_i dW_{i,t}. \end{aligned}$$

The identifiable parameters are the parameters in the process under $\tilde{\mathbf{P}}$ and $\mathbf{P}$, given by the five reduced-form parameters $\left(a_{i,j}\kappa_i\theta_i,\kappa_i,a_{i,j}\left(\kappa_i\theta_i - \xi_i\sigma_i\right),\kappa_i + \gamma_i\sigma_i,a_{i,j}\sigma_i\right)$. These five reduced-form parameters are a function of six structural parameters $a_{i,j}$, κ_i, θ_i, σ_i, ξ_i and γ_i. For another firm m the identifiable parameters are $\left(a_{i,m}\kappa_i\theta_i,\kappa_i,a_{i,m}\left(\kappa_i\theta_i - \xi_i\sigma_i\right),\kappa_i + \gamma_i\sigma_i,a_{i,m}\sigma_i\right)$. Apart from κ_i and without doing any normalization it is not possible to recover all structural parameters from the reduced-form parameter estimates. In order to solve the problem we normalize σ_i to 0.08, a value close to the ones obtained in Duffee (1999) and Driessen (2005) for common factors.

The same applies to the common factor Z_t.

References

[1] Amato, J., and Remolona, E., 2003, "The Credit Spread Puzzle," BIS Quarterly Review, December, 51-63.

[2] Amato, J., and Remolona, E., 2005, "The Pricing of Unexpected Credit Losses," Working Paper, BIS.

[3] Bakshi, G., Madan, D., and Zhang, F., 2006, "Investigating the Role of Systematic and Firm-Specific Factors in Default Risk: Lessons from Empirically Evaluating Credit Risk Models," *Journal of Business* 79, forthcoming.

[4] Basel Committee on Banking Supervision, 2005, "Guidance on Paragraph 468 of the Framework Document."

[5] BIS Quarterly Review, December, 2003. The full publication is available for free at the BIS website `www.bis.org/publ/quarterly.htm`.

[6] BIS Quarterly Review, March 2004. The full publication is available for free at the BIS website `www.bis.org/publ/quarterly.htm.`

[7] Bishop, G., and Welch, G., 2003, "An Introduction to the Kalman Filter," Working Paper, Department of Computer Science, University of North Carolina at Chapel Hill.

[8] Blanco, R., Brennan, S., and Marsh, I., 2005, "An Empirical Analysis of the Dynamic Relationship between Investment-grade Bonds and Credit Default Swaps," *Journal of Finance* 60, 2255-2281.

[9] Bobadilla, G., 1999, "Choosing the Right Error in Term Structure Models," Tesina CEMFI No. 9904.

[10] Chava, S., and Jarrow, R.A., 2004, "Bankruptcy Prediction with Industry Effects," *Review of Finance* 8, 537-596.

[11] Collin-Dufresne, P., Goldstein, R., and Martin, J., 2001, "The Determinants of Credit Spread Changes," *Journal of Finance* 56, 2177-2207.

[12] Couderc, F., and Renault, O., 2005, "Times-to-Default: Life Cycle, Global and Industry Cycle Impacts," FAME Research Paper No. 142.

[13] Covitz, D., and Han, S., 2005, "An Empirical Analysis of Bond Recovery Rates: Exploring a Structural View of Default," Finance and Economics Discussion Series No. 2005-10, Federal Reserve Board.

[14] Crowder, M., Davis, M., and Giampieri, G., 2005, "A Hidden Markov Model of Default Interaction," *Quantitative Finance* 5, 27-34.

[15] Das, S., Freed, L., Geng, G., and Kapadia, N., 2005, "Correlated Default Risk," Working Paper, Santa Clara University.

[16] De Jong, F., 2000, "Time-series and Cross-section Information in Affine Term Structure Models," *Journal of Economics and Business Statistics* 18, 300-318.

[17] De Servigny, A., and Renault, O., 2002, "Default Correlation: Empirical Evidence," Working Paper, Standard and Poor's.

[18] Driessen, J., 2005, "Is Default Event Risk Priced in Corporate Bonds?," *Review of Financial Studies* 18, 165-195.

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

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

[21] Duffie, D. Pan, J., and Singleton, K.J., 2000, "Transform Analysis and Asset Pricing for Affine Jump Diffusions," *Econometrica* 68, 1343-1376.

[22] Elizalde, A. "Credit Risk Models I: Default Correlation in Intensity Models," 2005a. Available at `www.abelelizalde.com`.

[23] Elizalde, A. "Credit Risk Models II: Structural Models," 2005b. Available at `www.abelelizalde.com`.

[24] Harvey, A., 1989, "*Forecasting, Structural Time Series and the Kalman Filter*," Cambridge University Press.

[25] Janosi, T., Jarrow, R., and Yildirim, Y., 2002, "Estimating Expected Losses from Liquidity Discount Bonds Implicit in Debt Prices," *Journal of Risk* 5, 1-39.

[26] Jegadeesh, N., and Pennacchi, G., 1996, "The Behavior of Interest Rates Implied by the Term Structure of Eurodollar Futures," *Journal of Money, Credit and Banking* 28, 426-446.

[27] Koopman, S., Lucas, A., and Monteiro, A., 2005, "The Multi-State Latent Factor Intensity Model for Credit Rating Transitions," Tinbergen Institute Discussion Paper No. TI 2005-071/4.

[28] Lando, D., 2004, "*Credit Risk Modelling: Theory and Applications*," Princeton University Press.

[29] Longstaff, F., Mithal, S., and Neis, E., 2005, "Corporate Yield Spreads: Default Risk or Liquidity? New Evidence from the Credit Default Swap Market," *Journal of Finance* 60, 2213-2253.

[30] Lucas, D., 1995, "Default Correlation and Credit Analysis," *Journal of Fixed Income* 4, 76-87.

[31] Moody's Investors Service, 1997, "Moody's Rating Migration and Credit Quality Correlation: 1920-1996," Special Comment.

[32] Moreno, M., 2003, "A Two-Mean Reverting-Factor Model of the Term Structure of Interest Rates," *Journal of Futures Markets* 23, 1075-1105.

[33] Nickell, P., Perraudin, W., and Varotto, S., 2000, "Stability of Rating Transitions," *Journal of Banking and Finance* 24, 203-227.

[34] Repullo, R., and Suarez, J., 2000, "Entrepreneurial Moral Hazard and Bank Monitoring: A Model for the Credit Channel," *European Economic Review* 44, 1931-1950.

[35] Xie, Y., Wu, C., and Shi, J., 2004, "Do Macroeconomic Variables Matter for the Pricing of Default Risk? Evidence from the Residual Analysis of the Reduced-Form Model Pricing Errors," Working Paper, Whitman School of Management, Syracuse University.

[36] Yu, F, 2005, "Default Correlation in Reduced-Form Models," *Journal of Investment Management* 3, 33-42.

[37] Zhang, F, 2003, "What did the Credit Market Expect of Argentina Default? Evidence from Default Swap Data," Finance and Economics Discussion Series No 2003-25, Federal Reserve Board.

Tables

Parameter	Estimate	St. Dev.
κ_1	0.1208	0.0009
θ_1	0.0041	0.0013
σ_1	0.0629	0.0005
ξ_1	−0.0018	0.0025
γ_1	0.1572	0.0196
κ_2	0.9297	0.0123
θ_2	0.1649	0.0012
σ_2	0.0199	0.0009
ξ_2	0.0450	0.0521
γ_2	0.0351	0.2756
σ_ε	0.0007	0.0001

Table 1: Estimates of the parameters in the state space model for the term structure of interest rates two factor model $X_{1,t}$ and $X_{2,t}$.

Maturity	MPE	MAPE
3 months	2.22	4.26
6 months	−1.79	2.53
1 year	−1.18	3.79
3 years	1.08	5.91
5 years	−0.32	3.67
20 years	−0.12	8.02

Table 2: (In sample) mean and mean absolute pricing errors, MPE and MAPE respectively, in basis points, implied by the model for the zero-coupon interest rates used in the estimation of $X_{1,t}$ and $X_{2,t}$.

Parameter	Estimate	Std. Dev.
κ_3	0.0046	0.0037
θ_3	0.5500	0.2268
σ_3	0.0800	
ξ_3	−0.5948	0.0200
γ_3	−0.4717	0.1687

Table 3: Estimates of parameters of the common risk factor process $X_{3,t}$.

Firm	$a_{1,j}$		$a_{2,j}$		$a_{3,j}$	
Owens	−0.3617	(0.0116)	−0.2853	(0.0177)	0.0442	(0.0016)
CSC	−1.2442	(0.0016)	−1.0422	(0.0004)	0.0742	(0.0003)
Comcast	−0.8965	(0.0053)	−0.9965	(0.0011)	0.0825	(0.0010)
Disney	−0.7459	(0.0040)	−0.8288	(0.0020)	0.0245	(0.0007)
Ford	−1.1089	(0.0016)	−0.9770	(0.0005)	0.0409	(0.0003)
GM	−1.0038	(0.0075)	−0.9295	(0.0028)	0.0292	(0.0014)
Hertz	−1.1003	(0.0683)	−0.9789	(0.0233)	0.0475	(0.0149)
Marriott	−0.8277	(0.0102)	−0.8893	(0.0039)	0.0356	(0.0016)
Norfolk	−0.2062	(0.0201)	−0.3624	(0.0252)	0.0164	(0.0034)
Sears	−0.9683	(0.0975)	−0.9487	(0.0355)	0.0392	(0.0242)
TCI	−0.9516	(0.0189)	−1.0056	(0.0014)	0.0812	(0.0018)
Warner	−0.8799	(0.0071)	−0.9945	(0.0024)	0.0848	(0.0014)
USX	−0.7512	(0.0019)	−0.8519	(0.0008)	0.0303	(0.0003)
Union Pacific	−0.7243	(0.0051)	−0.8187	(0.0025)	0.0206	(0.0008)

Table 4: Estimates of $a_{1,j}$, $a_{2,j}$ and $a_{3,j}$. Std. Dev. in brackets.

Firm	$a_{1,j}X_{1,t} + a_{2,j}X_{2,t}$	$a_{3,j}X_{3,t}$	$a_{4,j}X_{4,t}$	$X_{5,j,t}$	$s_{j,t}$	Rating
Owens	0.0086	0.0330		0.0138	0.0554	B
CSC	0.0186	0.0553	−0.0301	−0.0002	0.0436	BB
Comcast	−0.0313	0.0615	0.0022	−0.0093	0.0231	BBB
Disney	−0.0260	0.0183	0.0174	0.0001	0.0098	BBB
Ford	0.0078	0.0305		−0.0044	0.0339	BBB
GM	−0.0012	0.0218		0.0043	0.0250	BBB
Hertz	0.0060	0.0355		−0.0068	0.0347	BBB
Marriott	−0.0233	0.0266		0.0105	0.0137	BBB
Norfolk	−0.0314	0.0123	0.0101	0.0377	0.0287	BBB
Sears	−0.0106	0.0292		0.0057	0.0244	BBB
TCI	−0.0238	0.0606		−0.0012	0.0356	BBB
Warner	−0.0337	0.0633	−0.0062	0.0014	0.0248	BBB
USX	−0.0293	0.0226		0.0146	0.0079	BBB
Union	−0.0278	0.0154	0.0175	0.0001	0.0052	BBB

Table 5: Average values of the impact on each firm's credit spread $s_{j,t}$ of the different risk factors: $a_{1,j}X_{1,t} + a_{2,j}X_2$, $a_{3,j}X_{3,t}$, $a_{4,j}X_{4,t}$ and $X_{5,j,t}$. The last two columns show the average estimated credit spread and each firm's rating at the end of the sample period.

$X_{4,t}$	Railroad		Film		Cable TV	
κ_4	0.4786	(0.0146)	0.0530	(0.0023)	0.4864	(0.0635)
θ_4	0.2834	(0.0095)	−0.4651	(0.0022)	0.0014	(0.0027)
σ_4	0.0800		0.0800		0.0800	
ξ_4	3.8670	(0.0615)	−1.4882	(0.0160)	−0.5441	(0.0142)
γ_4	−49.2774	(4.5363)	0.5621	(0.1784)	−1.9223	(0.7378)

Table 6: Estimates of the parameters of the sector risk factor process $X_{4,t}$. Std. Dev. in brackets.

Sector	Firm	$a_{4,j}$	
Railroad	Norfolk	0.0297	(0.0013)
	Union Pacific	0.0513	(0.0022)
Film	Disney	−0.1031	(0.0010)
	Warner	0.0369	(0.0026)
Cable TV	CSC	0.2425	(0.0066)
	Comcast	−0.0180	(0.0031)

Table 7: Estimates of $a_{4,j}$. Std. Dev. in brackets.

% MAPE	$s_{j,t} = 0$	$X_{1,t}$, $X_{2,t}$ and $X_{3,t}$	$X_{4,t}$	$X_{5,j,t}$
Owens	21.80	3.95		1.16
CSC	23.49	2.26	0.83	0.73
Comcast	9.86	1.34	1.29	0.87
Disney	3.88	1.29	0.21	0.15
Ford Credit	11.64	1.70		0.70
GM	8.79	1.12		1.05
Hertz	18.07	2.42		0.88
Marriott	9.04	2.60		0.50
Norfolk	4.49	1.49	0.99	0.51
Sears	9.81	2.53		0.65
TCI	19.58	2.31		1.59
Warner	8.33	1.24	1.14	1.13
USX	5.89	1.59		0.42
Union	4.57	2.13	0.48	0.45
Mean	11.37	2.00	0.82	0.77
Std. Dev.	6.64	0.76	0.40	0.37

Table 8: Average, by firm, mean absolute pricing error (MAPE) as a percentage of the bond prices in each estimation stage, i.e. for each specification of the spread $s_{j,t}$. The specifications considered for $s_{j,t}$ are (in columns 2, 3, 4 and 5 respectively):

$$\begin{aligned} s_{j,t} &= 0, \\ s_{j,t} &= a_{1,j}X_{1,t} + a_{2,j}X_{2,t} + a_{3,j}X_{3,t}, \\ s_{j,t} &= a_{1,j}X_{1,t} + a_{2,j}X_{2,t} + a_{3,j}X_{3,t} + a_{4,j}X_{4,t}, \\ s_{j,t} &= a_{1,j}X_{1,t} + a_{2,j}X_{2,t} + a_{3,j}X_{3,t} + a_{4,j}X_{4,t} + X_{5,j,t}. \end{aligned}$$

$s_{j,t} = 0$ corresponds to the case in which bonds have been priced as if they were default-free bonds using the estimation results for the two-factor model for the risk free rate $r_t = X_{1,t} + X_{2,t}$ given in Table 1.

	$X_{1,t}$, $X_{2,t}$ and $X_{3,t}$	$X_{4,t}$	$X_{5,j,t}$
Owens	82.89		93.52
CSC	89.82	96.38	96.76
Comcast	82.13	81.59	88.12
Disney	61.81	92.80	94.19
Ford Credit	81.65		92.14
GM	85.92		86.76
Hertz	85.61		95.05
Marriott	43.39		73.85
Norfolk	60.20	63.86	75.62
Sears	68.50		92.63
TCI	72.27		74.13
Warner	80.18	82.66	93.20
USX	69.78		91.46
Union	47.37	80.53	81.08
Mean	72.25	82.97	84.89
Std. Dev.	14.57	11.39	13.80

Table 9: Fraction of the mean absolute pricing error not explained by the risk free rate (column 2 of Table 8) explained by the different specifications of the firms credit risk spread $s_{j,t}$ when we include the different credit risk factors. The specifications considered for $s_{j,t}$ are (in columns 2, 3 and 4 respectively):

$$
\begin{aligned}
s_{j,t} &= a_{1,j}X_{1,t} + a_{2,j}X_{2,t} + a_{3,j}X_{3,t}, \\
s_{j,t} &= a_{1,j}X_{1,t} + a_{2,j}X_{2,t} + a_{3,j}X_{3,t} + a_{4,j}X_{4,t}, \\
s_{j,t} &= a_{1,j}X_{1,t} + a_{2,j}X_{2,t} + a_{3,j}X_{3,t} + a_{4,j}X_{4,t} + X_{5,j,t}.
\end{aligned}
$$

For each bond and specification of $s_{j,t}$ we compute the daily average of one minus the ratio of the absolute pricing error implied by that specification and the absolute pricing error implied when the bond is priced as if they were risk free, i.e. $s_{j,t} = 0$. Then, we compute the average of all bonds of the same firm.

Parameter	Estimate	Std. Dev.
κ_Z	0.0411	0.0095
θ_Z	0.7764	0.0119
σ_Z	0.0800	
ξ_Z	0.5790	0.1083
γ_Z	-3.2721	0.9275

Table 10: Estimates of the parameters of the common risk factor process Z_t.

Firm	b_j	
Owens	0.0668	(0.0011)
CSC	0.0471	(0.0004)
Comcast	0.0251	(0.0001)
Disney	0.0138	(0.0001)
Ford	0.0346	(0.0003)
GM	0.0267	(0.0002)
Hertz	0.0381	(0.0002)
Marriott	0.0237	(0.0001)
Norfolk	0.0144	(0.0001)
Sears	0.0228	(0.0002)
TCI	0.0325	(0.0004)
Warner	0.0267	(0.0001)
USX	0.0162	(0.0001)
Union Pacific	0.0134	(0.0001)

Table 11: Estimates of b_j. Std. Dev. in brackets.

% MAPE	$s_{j,t} = 0$	$s_{j,t} = b_j Z_t$
Owens	21.80	4.82
CSC	23.49	3.69
Comcast	9.86	2.73
Disney	3.88	0.68
Ford Credit	11.64	1.26
GM	8.79	0.75
Hertz	18.07	1.82
Marriott	9.04	2.10
Norfolk	4.49	1.64
Sears	9.81	2.15
TCI	19.58	4.11
Warner	8.33	2.43
USX	5.89	1.18
Union	4.57	1.55
Mean	11.37	2.21
Std. Dev.	6.64	1.24

Table 12: Average, by firm, mean absolute pricing error (MAPE) as a percentage of the bond prices when we consider the alternative specification of the firms' credit spread $s_{j,t} = b_j Z_t$. The second column, $s_{j,t} = 0$, corresponds to the case in which bonds have been priced as if they were default-free bonds using the estimation results for the two-factor model for the risk free rate $r_t = X_{1,t} + X_{2,t}$ given in Table 1.

	$s_{j,t} = b_j Z_t$
Owens	80.59
CSC	82.70
Comcast	64.16
Disney	78.87
Ford Credit	87.13
GM	91.26
Hertz	88.73
Marriott	15.25
Norfolk	37.78
Sears	72.12
TCI	57.93
Warner	61.57
USX	77.11
Union	56.45
Mean	68.02
Std. Dev.	21.37

Table 13: Fraction of the mean absolute pricing error not explained by the risk free rate (column 2 of Table 12) explained by the different specifications of the firms credit risk spread $s_{j,t}$ when we consider the alternative specification $s_{j,t} = b_j Z_t$. For each bond we compute the daily average of one minus the ratio of the absolute pricing error implied by the previous specification and the absolute pricing error implied when the bond is priced as if they were risk free, i.e. $s_{j,t} = 0$. Then, we compute the average of all bonds of the same firm.

Firm	Bid-ask spread	Rating
Owens	0.0058	B
CSC	0.0040	BB
Comcast	0.0030	BBB
Disney	0.0027	BBB
Ford	0.0030	BBB
GM	0.0037	BBB
Hertz	0.0044	BBB
Marriott	0.0023	BBB
Norfolk	0.0026	BBB
Sears	0.0041	BBB
TCI	0.0031	BBB
Warner	0.0032	BBB
USX	0.0024	BBB
Union Pacific	0.0027	BBB

Table 14: Average bid-ask spread $\left(\frac{ask-bid}{mid}\text{ prices}\right)$ and credit rating for each firm.

Dep. variable: $s_{j,t}$	Coeff.	Std. Err.	Coeff.	Std. Err.
Intercept	0.0649	0.0011	0.0169	0.0006
Time to maturity	0.0006	0.0000	0.0003	0.0000
Age	0.0011	0.0000	0.0009	0.0000
Notional amount	0.0000	0.0000	0.0000	0.0000
Rating	−0.0175	0.0003		
Liquidity spread	0.9486	0.0681	2.0447	0.0678
	$R^2 = 16\%$		$R^2 = 7\%$	

Table 15: Regression of credit spreads $s_{j,t}$ with respect to bonds time to maturity (in years), age (in years), amount issued, rating at the end of the sample period and liquidity spread. The rating variable takes value one for B, two for BB and three for BBB. The liquidity spread is constructed as the price ratio $\frac{ask-bid}{mid}$.

Dep. variable: APE	Coefficient	Std. Err.
Intercept	-0.0098	0.0008
Time to maturity (years)	-0.0001	0.0000
Age (years)	0.0005	0.0000
Notional amount issued	-0.0000	0.0000
Rating	0.0018	0.0002
Liquidity spread	0.5978	0.0409
Credit spread $s_{j,t}$	0.1455	0.0050
Common risk factor $X_{3,t}$	0.0081	0.0005
Idiosyncratic risk factor $X_{5,j,t}$	-0.0078	0.0055
	$R^2 = 11\%$	

Table 16: Regression of Absolute Pricing Error APE (relative to bond prices) with respect to bonds, time to maturity, age, notional amount issued, rating at the end of the sample period, liquidity spread, credit spread $s_{j,t}$, common risk factor $X_{3,t}$ and idiosyncratic risk factor $X_{5,j,t}$. The rating variable takes value one for B, two for BB and three for BBB. The liquidity spread is constructed as the price ratio $\frac{ask-bid}{mid}$.