J. Finan. Intermediation 19 (2010) 529–563

Contents lists available at ScienceDirect

J. Finan. Intermediation

journal homepage: www.elsevier.com/locate/jfi

Predicting credit spreads ☆

C.N.V. Krishnan [a,*], Peter H. Ritchken [a], James B. Thomson [b]

[a] Weatherhead School of Management, Case Western Reserve University, Cleveland, OH 44106, USA
[b] Federal Reserve Bank of Cleveland, Cleveland, OH 44106, USA

ARTICLE INFO

Article history:
Received 24 June 2007
Available online 28 July 2009

ABSTRACT

Predictions of firm-level credit spreads based on the current spot and forward credit spreads can be significantly improved upon by using the information contained in the shape of the credit-spread curve. However, the current credit-spread curve is not a sufficient statistic for predicting future out-of-sample credit spreads; predictions can be significantly improved upon by exploiting the information contained in the shape of the riskless yield curve. In the presence of credit-spread and riskless factors, other macroeconomic, marketwide, and firm-specific risk variables do not significantly improve predictions of credit spreads. These results have important implications for credit-spreads modeling as well as for better understanding corporate capital structure and risk management policies.

1. Introduction

A vast literature exists that is concerned with predicting future riskless interest rates. Fama and Bliss (1987), Campbell and Shiller (1991), Cochrane and Piazzesi (2005), Diebold and Li (2006) and others show that the current yield curve contains significant information on future yields. More recently, studies have examined whether auxiliary variables can be used in conjunction with yield-curve information to improve forecasts. Monch (2005) and Ludvigson and Ng (2009), for example, find that macroeconomic variables have predictive power over future riskless yields, over and above what is contained in the current yield curve.

☆ The views expressed in this paper are those of the authors and not necessarily those of the Federal Reserve Bank of Cleveland or the Federal Reserve System. We gratefully acknowledge the comments and suggestions of Francis Longstaff, S. Viswanathan (the editor) and two anonymous referees. We thank the participants of the Second Annual Risk Management Institute Conference in Singapore in 2008, and those of a seminar at the University of California, Irvine, for valuable comments. The authors thank the FDIC Center for Financial Research for financial support.

* Corresponding author.
E-mail addresses: cnk2@cwru.edu (C.N.V. Krishnan), phr@cwru.edu (P.H. Ritchken), JB.Thomson@clev.frb.org (J.B. Thomson).

1042-9573/$ - see front matter
doi:10.1016/j.jfi.2009.02.004

In contrast to the many studies concerned with the predictability of riskless yields, much less work has been done on predicting credit spreads at the firm level. Rather, the focus of most studies in this area has been to establish structural models that price corporate securities as claims on the assets of the firm.[1] The effectiveness of these models is often measured by their ability to *output* credit-spread curves that closely match the actual credit-spread curves.[2] For the most part, these studies do not address predictability.

Our focus in this paper is different from the above studies on credit spreads in that we take the current credit-spread curve of a firm, together with its historical values, as *input*. Rather than attempting to explain the level or shape of the current credit-spread curve by leverage, distance to default, expected recovery rates, riskless rates, volatility, liquidity, taxes, etc., our goal is to forecast what credit spreads will be in the future taking today's credit-spread curve as data. Further, we examine what auxiliary information, if any, help in predicting credit spreads.

We address two important questions in this study. First, can forecasts based on spot or forward credit spreads be improved upon by using information contained in today's credit-spread curve? In this regard, our goal is somewhat similar to that of Cochrane and Piazzesi (2005) for riskless yields. They found that the Fama and Bliss (1987) model for predicting riskless yields could be substantially improved upon by incorporating information from multiple yields drawn from the current riskless yield curve. Second, does the credit-spread curve act as a sufficient statistic for predicting future credit spreads, or can forecasts be significantly improved by incorporating additional firm-specific, market-wide, or macroeconomic information? Such an exercise would significantly enhance current knowledge on the overall determinants of *future* credit spreads, and would have important implications for credit spread models. If additional information is found to significantly improve credit spread forecasts, then it would be of interest to identify the nature of these additional factors and the degree to which they help with predictions, over and above the credit-spread factors. The existence of additional factors would cast doubt on the applicability of affine models for credit spreads because affine models have the property that all information necessary for predicting future credit spreads is impounded in the current credit-spread curve. Further, if the current period riskless term structure contains significant information on future credit spreads, then interpretations of extant empirical studies in corporate finance that make connections between corporate debt issuance and risk management decisions and the term structure of interest rates can change in significant ways.

To conduct this study, we first construct credit-spread curves at the firm level, using the prices of multiple corporate bonds issued by a firm. We analyze firms across industries and credit ratings spanning 16 years. For the most part, empirical studies at the firm level define the credit spread as the difference in basis points between the yield of the corporate bond and the yield of an equivalent Treasury security, thereby not addressing the term structure of credit spreads. In contrast, we extract the *full term structure* of credit spreads for each firm, avoiding the implicit assumption that credit spread shocks are parallel. In the riskless market three factors are often used to model riskless yield curves. These factors permit level, slope, and curvature shocks. Our construction of credit-spread curves permits these types of shocks, which collectively allow for weaker correlations among credit spreads of different maturities. Indeed, we find that shocks to credit-spread curves are not parallel and not very strongly correlated. For example, the 2-year and the 5-year credit spreads change in the same direction only about 60% of the time, and the correlation between their changes is low enough to warrant more than one factor for credit spreads.

Once credit-spread curves have been constructed firm by firm, we investigate the performances of the current spot and forward credit spreads as predictors of future credit spreads, which we compare to richer models that incorporate information on the full credit-spread curve as well as to forecasts produced using additional blocks of variables. To account for differences in the likelihood of default, expected recovery given default, and liquidity and risk premia we run panel regressions on groups

[1] Examples include Black and Cox (1976), Kim et al. (1993), Longstaff and Schwartz (1995), Leland and Toft (1996), Collin-Dufresne and Goldstein (2001) and Acharya and Carpenter (2002).

[2] Empirical studies include Huang and Huang (2003) and Eom et al. (2004), who find that very large pricing errors are the rule. For discussions on the shape of credit-spread curves, see Helwege and Turner (1999), He et al. (2004) and Agrawal and Bohn (2008). For a more recent approach, see, for example, Cremers et al. (2008).

of firms double-sorted by industry and credit ratings, and conduct our analyses firm by firm as well. Moreover, Hotchkiss et al. (2002) argue that bond trading volume (and hence, liquidity) can vary with both credit rating and industry. Therefore, double sorting by industry and credit rating also controls for liquidity. Credit spread predictability could also be influenced by time-varying risk premia and default and recovery rates. Including financial market and macroeconomic information and considering industry and credit ratings effects in our analyses allows us to account for time-varying risk premia and default and recovery rates without explicitly modeling them.

Our results are based on extensive in-sample and out-of-sample tests. It is widely believed that significant in-sample evidence of predictability does not guarantee significant out-of-sample predictability. Indeed, the literature is replete with warnings about using in-sample inferences to show predictability. However, the inclusion of irrelevant variables, while increasing the in-sample fit, does not affect the reliability of in-sample tests of predictability.[3] This point has been emphasized by several authors including Inoue and Kilian (2004), who show that neither data mining nor parameter instability provide plausible explanations for in-sample tests to reject the no-predictability null hypothesis more often than for out-of-sample tests. Indeed, they conclude that in-sample predictability is typically more credible than results of out-of-sample tests, because out-of-sample analysis requires sample splitting, which in turn involves loss of information, and, hence, lower power in small samples. On the other hand, until recently, even studies on affine equilibrium riskless term structure models, such as that of Dai and Singleton (2000) focused on in-sample fits as opposed to out-of-sample forecasting; and those models that did focus on out-of-sample forecasts, such as Duffee (2002), reported disappointing results, with predictions not better than those from a simple random walk model. To be conservative, we conduct both in-sample and out-of-sample forecasts using panel and firm-by-firm regressions.

Our results from in-sample and out-of-sample tests are consistent and are as follows. First, forward credit spreads are not unbiased estimators of future credit spreads, and in general provide poor predictions of future credit spreads. Predictions given by the current spot model are superior. Second, predictions given by the spot credit spread can be substantially improved upon by using a model that incorporates the level, slope, and curvature factors of the credit-spread curve. That is, the shape of the credit-spread curve is informative about the future level of any particular credit spread. Third, we find that the credit-spread curve is not a sufficient statistic for forecasting future credit spreads. Forecasts can be significantly improved upon by incorporating information from the riskless term structure. Fourth, in the presence of current credit-spread and riskless factors, other information such as stock-market information, macroeconomic factors and firm-specific risk variables are less informative for forecasting future credit spreads. Our most parsimonious model, which uses information only from the current credit-spread and riskless-yield curves, significantly outperforms the spot model: for example, for forecasting 6-months-ahead 5 year out-of-sample credit spreads this model produced smaller mean-squared prediction errors (MSPEs) for over 80% of our firms compared to the spot model. With such a model we are able to predict future out-of-sample 6-months-ahead 5-year credit spreads with no unconditional bias and an average absolute prediction error of 31 basis points. Given firm credit-spread curves and riskless yield factors, no other sets of variables could be identified that was consistently informative.

Our use of panel regressions on groups of firms double-sorted by credit ratings and industry addresses the issue of possible noise in firm-level estimation. To further address this issue, we also perform our predictability analysis with aggregate B-credit-rated-index yield curves taken from Bloomberg. As additional checks, we examine out-of-sample predictions of firm level credit spreads of different maturities and different forecast horizons, and also replace our macro variables with macro forecasts from the Survey of Professional Forecasters. Our results are robust.

The one empirical study that comes closest to our study is Krishnan et al. (2006), although their study focuses on banking firms and bank regulation. They investigate whether the spot or forward credit-spread curves provide unbiased estimates of future spot credit spreads. They strongly reject

[3] By construction, the F-test of predictability is designed under the hypothesis that the regressor is irrelevant, and, as more and more irrelevant variables are included, the critical value of the F-test will increase to account for this. Thus, the inclusion of irrelevant variables has no effect on the asymptotic size of predictability tests.

the expectations hypothesis and find that the current credit spread *slope* is informative about future credit spreads. Our paper shows that *multiple* points on the term structure of credit spreads contain significant information on future credit spreads. Krishnan et al. (2006) do not examine out-of-sample predictions of credit spreads. We show that information other than that contained in the credit-spread curve significantly enhance out-of-sample predictions of credit spreads.

The remainder of the paper is structured as follows. In Section 2 we describe our model for extracting the riskless-yield curve and credit-spread curves at the firm level. In Section 3, we describe the data and provide descriptive statistics. In Sections 4 and 5, we investigate in-sample measures of predictive content for riskless yields and aggregated B-rated credit spreads, and for firm credit spreads. In Section 6, we perform out-of-sample tests of predictive content, and conduct robustness checks. Section 7 concludes with implications of our results for future research.

2. Extracting riskless-yield and credit-spread curves

2.1. Extracting riskless yield curve

Let $P^{(n)}(t)$ be the date t price of a zero-coupon riskless bond that pays \$1 in n periods time. Then

$$P^{(n)}(t) = e^{-y_f^{(n)}(t)n}, \tag{1}$$

where $y_f^{(n)}(t)$ is the riskless yield to maturity.

Our first task is to extract riskless factors from the riskless yield curves. We begin with the unsmoothed Fama–Bliss yield curves which we describe in detail in the next section. For each period, we fit the following model to the yields:

$$y_f^{(n)}(t) = \beta_{1f}(t) + \beta_{2f}(t)F_2^{(n)} + \beta_{3f}(t)F_3^{(n)}, \tag{2}$$

where

$$F_2^{(n)} = \frac{(1-e^{-\lambda_t n})}{\lambda_t n},$$
$$F_3^{(n)} = \frac{(1-e^{-\lambda_t n})}{\lambda_t n} - e^{-\lambda_t n}.$$

The Nelson and Siegel (1987) model has a discount function that begins from 1 at date 0, and approaches 0 as the horizon extends to infinity. Bliss (1997) shows that such a model performs very well in fitting the cross-section of riskless bond prices relative to a large class of alternative parameterizations. Our purpose of using this methodology is to reduce the dimensionality of the yield curve without losing much information. Specifically, at each point in time the resulting fitted curve is characterized by three variables,

$$\beta_f(t) = \left(\beta_{1f}(t), \beta_{2f}(t), \beta_{3f}(t)\right)'.$$

The loading on $\beta_{1f}(t)$ is 1, a constant that can be viewed as a permanent or long-term factor that affects all maturities equally; the loading on $\beta_{2f}(t)$, $F_2^{(n)}$, is a function that rapidly decreases to zero as n increases and hence can be viewed as a short-term factor; the loading on $\beta_{3f}(t)$, $F_3^{(n)}$, a function that begins at zero, increases, and then decreases to zero, can be viewed as a mid-term factor. Diebold and Li (2006) point out that these three factors can be viewed as controlling the level, slope, and curvature of the yield curve. Indeed, since $y_{f,t}^{(\infty)} = \beta_{1f}(t)$ and $y_{f,t}^{(\infty)} - y_{f,t}^{(0)} = -\beta_{2f}(t)$, the first two betas correspond to level and slope. Increasing $\beta_{3f}(t)$ has no effect on the short and long rates but does affect the middle rates, so it captures curvature effects. Diebold and Li show that the three time varying parameters can be interpreted as factors. Unlike factor analysis, however, in which one estimates both the unobserved factors and the factor loadings, the Nelson–Siegel Diebold–Li procedure imposes a particular functional form on the factor loadings, and this facilitates their precise estimation.

2.2. Constructing credit-spread curves

We use the Diebold–Li parameterization for a firm credit-spread curve so as to fit corporate bond prices of individual firms at given dates. Let $\pi_j^{(n)}(t)$ be the date t price of a zero-coupon bond issued by firm j that promises to pay \$1 n periods later. Then

$$\pi_j^{(n)}(t) = e^{-y_j^{(n)}(t)n}, \tag{3}$$

where $y_j^{(n)}(t)$ is the risky yield to maturity. The date t n-year credit spread is given by, $s_j^{(n)}(t)$, where

$$s_j^{(n)}(t) = y_j^{(n)}(t) - y_f^{(n)}(t). \tag{4}$$

In order to tease out a unique credit-spread curve for a particular firm at a particular date, we have to decompose risky coupon bond prices into a portfolio of risky zero coupon bond prices in a way which is independent of any specific coupon bond. Jarrow (2004) identifies sufficient conditions on recovery rates that allows such a decomposition to hold. These conditions include the more typical conditions assumed in most models of recovery, such as recovery of Treasury, recovery of par, and recovery of market value just prior to default. Under any of these recovery assumptions, and, assuming that, at any point in time, all bonds used are straight coupon bonds with no special features and of the same seniority, a credit-spread curve can be constructed for the firm. Indeed, as described in the next section, we are careful to choose only plain-vanilla senior fixed-coupon bonds for our analysis.

In this study, we take the credit-spread curve of a firm, together with its historical values, as input data in our prediction models. Therefore, it is very important to tease out a credit-spread curve that prices corporate bonds very accurately. At any date, we use the actual riskless-yield curve and tease out the credit-spread curve every firm every month so as to accurately price corporate bonds with maturities spanning the yield curve in that firm-month. To do this, we adopt the Nelson–Siegel Diebold–Li methodology. Specifically, we assume:

$$s_j^{(n)}(t) = \beta_{1j}(t) + \beta_{2j}(t)F_2^{(n)} + \beta_{3j}(t)F_3^{(n)}. \tag{5}$$

Then:

$$\pi_j^{(n)}(t) = e^{-y_f^{(n)}(t)n - \left(\beta_{1j}(t) + \beta_{2j}(t)F_2^{(n)} + \beta_{3j}(t)F_3^{(n)}\right)n}. \tag{6}$$

Let $\beta_j(t)$ represent the credit-spread vector for firm j at date t. Then:

$$\beta_j(t) = \left(\beta_{1j}(t), \beta_{2j}(t), \beta_{3j}(t)\right).$$

Given this vector and the actual observed riskless yield $y_f^{(n)}(t)$, for each maturity n, the price of all corporate zero-coupon bonds can be obtained and then the price of all corporate coupon bonds can be established. Given the prices of an array of corporate bonds issued by firm j, together with the contemporaneous riskless yield curve, we can infer the credit-spread curve's state variables, $\beta_j(t)$. To do this, assume firm j has $N_j(t)$ bonds trading in period t. We look for firms which have five or more bond issues outstanding with maturities spanning at least 7 years. We choose the credit-spread state variables to minimize the resulting sum of squared errors between theoretical and observed coupon-bond prices. That is, for each date, t, and for each firm, j, we solve:

$$\beta_j^*(t) = \arg\min_{\beta_j(t)} \sum_{i=1}^{N_j(t)} \epsilon_{i,j}^2(t), \tag{7}$$

where $\epsilon_{i,j}(t)$ is the actual price of bond i of firm j trading at date t less its estimated value.

We also use the Diebold–Li parameterization to extract B index credit spread factors from credit-spread curves constructed from Bloomberg's fair-market-value corporate yield curves. The time series of these factors may be helpful in predicting the future levels of an individual firms credit spread. We chose the B rated index, since firms in this index will be very sensitive to changing economic conditions. Let $\beta_I(t)$ represent the vector of the level, slope, and curvature of the credit-spread curve obtained from firms belonging to a certain credit rating. Then

$\beta_I(t) = (\beta_{1I}(t), \beta_{2I}(t), \beta_{3I}(t)).$

In summary, our goal in the data preparation phase is to construct a time series of monthly riskless state variables or factors, $\beta_f(t)$, credit-spread-index factors, $\beta_I(t)$, and a panel of firm credit-spread factors, $\beta_j(t), j = 1, \ldots, N$, where N is the number of firms from which the credit-spread curves could be constructed.

There are other ways of constructing credit-spread curves consistent with the data. For example, Krishnan et al. (2005) use a three factor model in which interest rates evolved according to a two factor double mean reverting process and the credit spread process was modeled as a one factor mean reverting process with constant volatility. Given structures for all the market prices of risks, analytical solutions for risky bond prices were generated, and, using the full panel data, the parameters were estimated using a Kalman filter. While this type of procedure provides consistent arbitrage-free pricing over consecutive time periods, the actual fit of corporate bond prices for each firm and at each point in time is not necessarily good. In contrast, here we want to make sure that our extracted firm-by-firm credit-spread curves, implied out at each point in time has the property that the fitted corporate bond prices very closely matches their actual values for each firm. Moreover, the amount of transaction price data to estimate the credit-spread curve for each individual firm each month is limited. The three parameter Nelson–Siegel Diebold Li model provides the flexibility to accommodate an array of possible credit-spread curve shapes even with relatively sparse data. Thus, our choice of the Nelson–Siegel Diebold–Li model reflects our desire for a simple parsimonious model, capable of taking on a multitude of shapes to fit observed bond prices.

2.3. Forecasting future firm-specific credit-spread curves

Diebold and Li (2006) explore the time series properties of the extracted beta factors for the riskless yield curve. They find that the sequence of parameter values, $\beta_f(t),\ t = 1, \ldots, T$, are highly autocorrelated. They establish a vector autoregressive model that forecasts future beta values, which are then used to estimate future yields. Specifically:

$$\beta_f(t+h) = \eta_{f1} + \eta_{f2}\beta_f(t) + \eta_{f3}X(t) + \epsilon(t+h), \tag{8}$$

where η_{f1} is a 3×1 vector, η_{f2} is a 3×3 matrix, $X(t)$ is a set of control variables, say of size k, η_{f3} is a matrix of size $3 \times k$, and $\epsilon(t+h)$ is the residual vector. These models could then be enhanced by adding additional independent variables in the right-hand side. For example, one could add in a panel of macro-economic variables.[4]

Let $\beta_f(t+h|t)$ be the date t forecast for the beta state values at date $t+h$, and let $y_f^{(n)}(t+h|t)$ be the date t forecast of the yield at date $t+h$, of a riskless bond that matures n years later. Then:

$$y_f^{(n)}(t+h|t) = \beta_{1f}(t+h|t) + \beta_{2f}(t+h|t)F_2^{(n)} + \beta_{3f}(t+h|t)F_3^{(n)}. \tag{9}$$

For the riskless yield curve, Diebold and Li compare these forecasts with forecasts generated by several other competing models including some of the affine term structure models, the Fama and Bliss (1987) model, and the Cochrane and Piazzesi (2005) model, and concluded that this approach provided superior one year ahead forecasts.

We adopt the Diebold–Li forecasting approach for credit-spreads at the firm level. In principle, we could run regressions of the form in Eq. (8) at the firm level. In practice, we complement the firm-by-firm regressions with panel regressions that exploit commonalities across firms of similar types. We postulate models of the form

$$\beta_j(t+h) = \eta_{j0} + \eta_1\beta_j(t) + \eta_2\beta_f(t) + \eta_3\beta_I(t) + \eta_4 M(t) + \eta_5 F_j(t) + \epsilon_j(t+h), \tag{10}$$

where $M(t)$ is a vector of macroeconomic and marketwide variables and $F_j(t)$ is a vector of firm-specific variables. Once the predicted values of future beta values are obtained, theoretical credit spreads can be computed, and these theoretical values can be compared to the future credit spreads. Firms in

[4] For further discussion of this approach, see Diebold et al. (2006).

each of these panel regression models can further be grouped by credit ratings, industry, or along other possible dimensions. Even the simplest case of the above model where $\eta_2 = \eta_3 = \eta_4 = \eta_5 = 0$ is of some interest because it allows us to examine the predictability of credit spreads based on information contained in the shape of today's credit-spread curve alone.

3. Data and descriptive statistics

3.1. Riskless-yield-curve factors

Our first data set consists of month-end price quotes for Treasury issues for the period 1970–2005 taken from the Center for Research in Security Prices (CRSP) Government Bond Files. We eliminate bonds with option features and bonds with special liquidity problems that arise because their maturities fall within one year. We use the Fama and Bliss (1987) bootstrapping procedure on the riskless-bond data collected from CRSP to compute raw yields from the filtered data. This method establishes forward rates in order to price bonds of successively longer maturities correctly, given the yields fitted to previously included issues. From these forward rates, the averages over increasing maturities are computed to obtain the zero-yield curve. This resulting curve, called the unsmoothed Fama–Bliss curve, prices the included bonds exactly.[5] Once the yield curve is established, we compute the yields to maturity for zero bonds of any maturity. We chose different maturities up to 10 years that are six months apart. For each selected month, therefore, we have 20 yields of maturities ranging from 6 months to 10 years. Using this data, we construct the vector of riskless beta factors, $\beta_f(t)$, for each date.

Fig. 1 shows that the three beta factors do indeed correspond closely to level, slope, and curvature effects. The solid lines in Fig. 1 show the time series of the estimated parameter values over the period 1970–2005, and are based on fixing $\lambda = 0.7308$, the value recommended by Diebold and Li. The dashed lines show the level, slope and curvature of the data-based riskless-yield curve. The data-based riskless-yield-curve level is taken as the 10-year rate. The slope is the difference between the 10-year and 3-month rate. The curvature is defined as twice the 2-year yield minus the sum of the 3-month and 10-year yield.

3.2. B-rated credit-spread-index factors

Our second data set consists of the B-credit-rated-index yield curves for industrial firms taken from Bloomberg. These yield curves are available daily from 1992 and are constructed using prices from new-issue calendars, trading/portfolio systems, dealers, brokers and evaluation services. Option-adjusted spread analysis is employed to construct option-free yield curves.

We choose a below-investment-grade index of industrial firms (the B-rated index) because these credit spreads could be extremely sensitive to prospects of changes in market conditions. Further, using an aggregate index eliminates noise from idiosyncratic firm-level shocks. We extract the monthly level, slope, and curvature factors, $\beta_I(t)$, for the B-rated-index credit spreads from these zero-coupon yield curves. Fig. 2 repeats the analysis of Fig. 1, and shows that the B-index factors indeed are closely related to the B-rated index levels (long credit spreads), slopes (long-credit spreads minus short-credit spreads), and credit-spread curvatures.

Thus, the two figures above show that the beta factors correspond very well with the level, slope, and curvature effects for both the riskless and the risky term structures.

3.3. Macroeconomic variables

Some models show that risky debt valuation explicitly incorporates macroeconomic conditions. Tang and Yan (2006), for example, develop a structural model that allows credit spreads to be affected by the interaction of macroeconomic conditions and firm characteristics. Their model improves upon existing structural models such as that examined by Huang and Huang (2003). Amato and Luisi (2006) develop a reduced form model where instantaneous credit spreads are assumed to be affine functions

[5] We thank Rob Bliss for providing us with the FORTRAN programs and data that allowed us to make this computation.

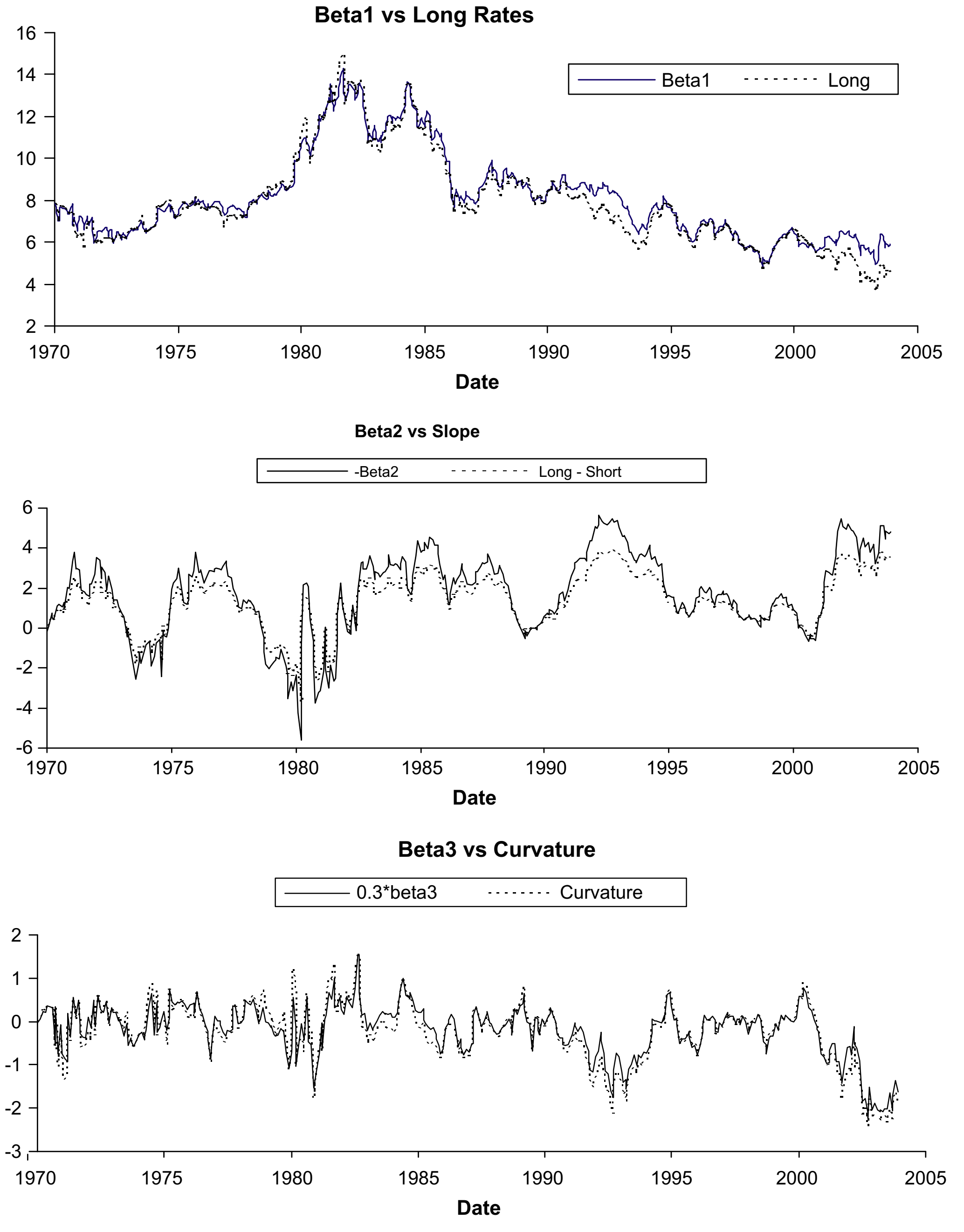

Fig. 1. Riskless term structure: beta values as level, slope, and curvature factors. The plots compare the time series of riskless-yield-curve level, slope, and curvature parameter estimates with the corresponding actual riskless long rates, slopes, and curvatures. The time series of beta values are obtained using the Nelson–Siegel model as discussed in the paper. The data consists of the unsmoothed monthly Fama–Bliss riskless yields over the period from January 1970 to December 2005.

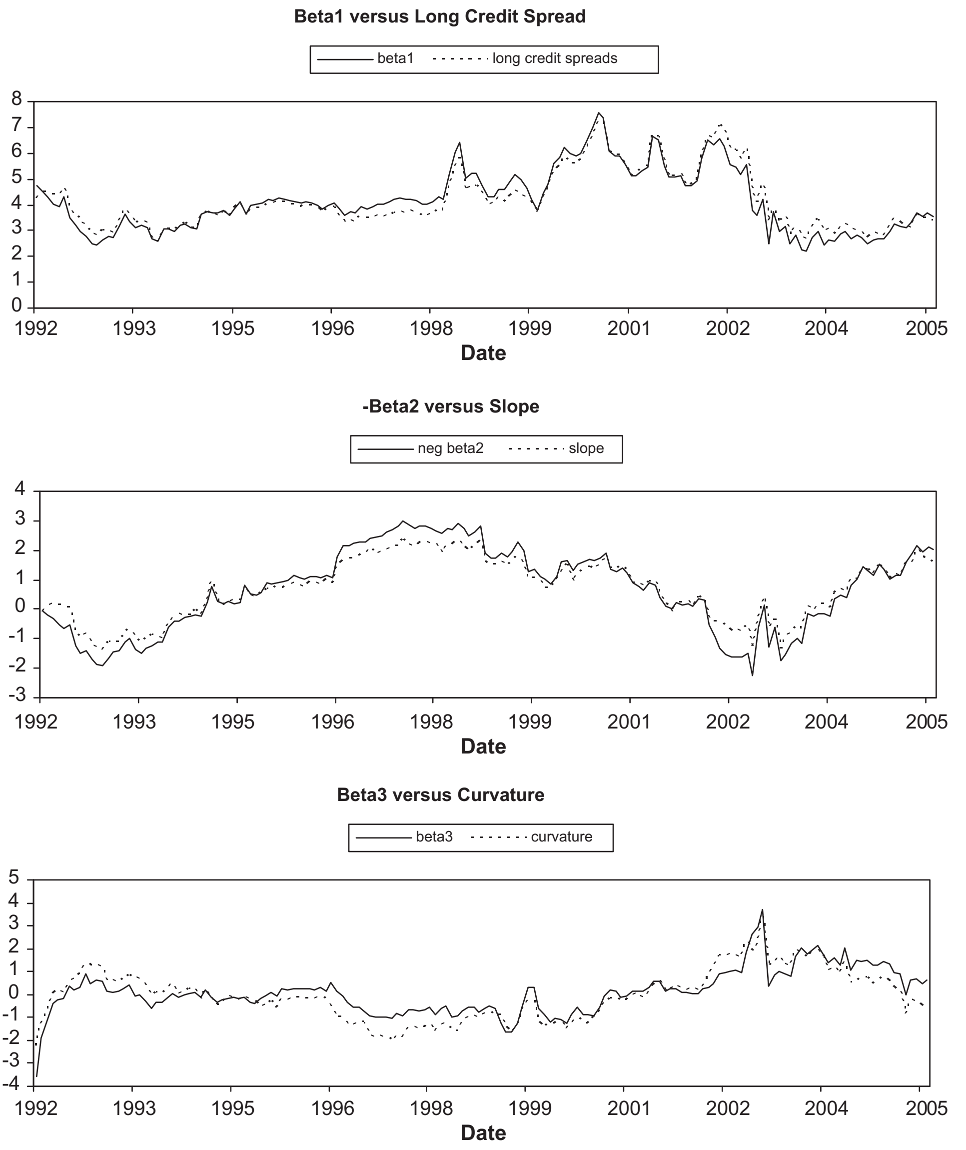

Fig. 2. B-Rated index term structure of credit spreads: beta values as level, slope, and curvature factors. The plots compare the time series of B-rated-index credit-spread level, slope, and curvature factor estimates with the corresponding actual long credit spreads, credit-spread slopes, and credit-spread curvatures. The data consists of the B-rated-index term structure curves obtained from Bloomberg for the period from January 1992 to December 2005.

of both observable macroeconomic variables and latent factors. Neither of these studies explore predictability, rather they show how economy-wide factors can affect credit spreads, which we exploit in our study.

Therefore, our third data set consists of macroeconomic information. It includes a Real Activity Index, $RA(t)$, an Inflation Index, $I(t)$, and two aggregate stock market variables: stock market momentum, $R_M(t)$, and stock market volatility, $\sigma_M(t)$.

The Real Activity and the Inflation indices are the first principal components of several observable time-series of macroeconomic variables, following Ang and Piazzesi (2003). Before conducting the principal component analysis, we purge these variables of the riskless factors. They therefore represent variables that are orthogonal to the riskless term-structure information. Increased real activity could spell investor confidence or, alternatively, signal inflationary pressures, both of which affect credit spreads, but differently. As inflation increases, so do the riskless yields, and the credit spread can increase as well. However, the effects of rising inflation on the long and short credit spreads can be different.

Stock market momentum is the 12-month cumulative holding-period return of the CRSP Value-weighted Index return. Collin-Dufresne et al. (2001) find a negative relationship between equity market returns and the level and slope of credit spreads. On the other hand, with stock market momentum, there could be flight of funds to the stock market, as a result of which credit spreads could rise. Stock market volatility is the monthly volatility of the CRSP value-weighted portfolio using the daily returns of the index within each month, following French et al. (1987). Asset volatility, generally approximated by equity volatility, includes both an idiosyncratic and a market component. Campbell et al. (2001) demonstrate that these two components could have different impacts on credit spreads. As stock market volatility increases, idiosyncratic risk could increase and credit spreads should rise. On the other hand, as stock market volatility increases, there could be a flight to the perceived relative safety of the bond markets, causing credit spreads to fall.

Let

$$M(t) = (RA(t), I(t), R_M(t), \sigma_M(t))$$

represent our 4-vector of macro-economic variables, each of which is described in detail in Appendix A.

3.4. Firm-specific risk variables

Our fourth data set consists of firm-specific information. It includes leverage, $L_j(t)$, book-to-market ratio, $BM_j(t)$, stock return momentum, $R_j(t)$, and stock return volatility, $\sigma_j(t)$, the data for all of which are taken from the quarterly Compustat and CRSP databases.

As leverage increases, bond risk increases and the credit spreads should increase. As Pastor and Veronesi (2003) show, the book-to-market ratio decreases with expected profitability. Also Fama and French (1992) show that the book-to-market ratio is a risk factor; as the book-to-market ratio increases, credit spreads should increase. According to structural models, an increase in stock return (stock return momentum) raises the equity holders' option value and reduces the default probability, which, in turn, should decrease credit spreads. Avramov et al. (2007) found stock market momentum to be a primary driver of credit spreads. As stock return volatility increases, firm value volatility increases, and default risk increases, which, in turn, should increase credit spreads. Each of these variables is described in detail in Appendix A.

Let

$$F_j(t) = (L_j(t), BM_j(t), R_j(t), \sigma_j(t))$$

represent the vector of firm variables for firm j.

We use additional firm-specific information for grouping purposes. These are credit ratings and industry. Credit spreads depend on the likelihood of default and on the recovery rate given default. Credit ratings determine the probability of default. Chava and Jarrow (2004) show that recovery rates vary with industry. We assign firms to either of two credit rating groups, namely, investment-grade and below-investment-grade firms.[6] We assign firms to one of two industry groups, namely, manufacturing and service.

[6] Investment grade firms are those that are rated *BBB* or above according to the S&P long-term firm rating found in the quarterly Compustat, augmented with credit-ratings data from Bloomberg, while the below-investment-grade firms are those that are rated below *BBB*.

3.5. Corporate bond data

Our final data set consists of the prices of corporate bonds from firms that are or were part of the S&P500 Index for any of the years in our sample period from January 1990 to December 2005. The S&P 500 Index is maintained by a team of Standard & Poor's economists and index analysts, who ensure that its composition remains a leading indicator of US stock market, reflecting the risk and return characteristics of the broader large-cap universe. On average, almost two changes are made to the S&P 500 Index each month, so the number of firms in the index over this period is quite large. Whereas some deletions were caused by mergers and acquisitions or spin-offs, others were deleted because of low share-price or market-capitalization. Indeed, our sample includes a few firms that subsequently defaulted. Our primary source of trade-price data on bonds is Bloomberg. We collect bond prices for all fixed-rate US dollar-denominated senior plain-vanilla bonds that are non-callable, non-putable, non-convertible, not part of an unit (e.g., sold with warrants) and have no sinking fund. We also excluded bonds with asset-backed and credit-enhancement features. This ensures that our credit spreads relate more directly to the creditworthiness of the issuer rather than the collateral. We use only transaction prices. Further, we eliminate bonds with principal repayment dates that are inconsistent with the maturity date that can be calculated using data on future coupon payment dates, and bonds with prices that are above otherwise equivalent Treasury securities. Appendix B details the key data filters used in the pre- and post-credit-spread generation stages to enable reproducibility. Although the Bloomberg database contains a large number of quoted bond prices, it contains a large number of traded bond prices as well. Indeed, for just the S&P 500 Index firms over the period 1990–2005, we collected 14,049 firm-months of traded price data from Bloomberg, for firms for which there were a minimum of five prices of bonds of different maturities that span at least 7 years in any given month. Almost all our data came from Bloomberg. We use DataStream as a second source to cross-check some of the traded prices obtained from Bloomberg. In almost all cases, when we had prices from both sources the prices were identical or almost identical. We resorted to DataStream, wherever possible, to fill in our requirement of five prices of bonds of different maturities that span at least 7 years in any given firm-month, where we fell short in terms of getting data from Bloomberg. As with Bloomberg, we obtain trade prices from DataStream. However, the proportion of trade price data obtained from DataStream to augment trade price data from Bloomberg is only around 1%.

Panel A of Table 1 shows the number of firms and firm-months in our sample, step-by-step through our screening process. The first column shows the data particulars of the initial set of firms. In the first screen, we require a minimum of five prices of bonds of different maturities that span at least 7 years, for each month, to estimate the credit-spread level, slope and curvature parameters. In the second screen, we drop all firm-months that do not have at least six consecutive months of reasonable credit-spread level, slope, and curvature parameter estimates. This is a requirement for our out-of-sample prediction analysis. In the third screen, we need equity returns data for the past 12 months from the CRSP database, and data from the quarterly Compustat to obtain our firm-specific risk measures. Our final sample comprises 241 firms and 11,894 firm-months of data.

Panel B shows that manufacturing and service-sector firms are well-represented in our final sample. Most of the firms are investment grade, as would be expected from S&P 500 firms. However, we have a sizeable sample of below-investment-grade firms as well as three firms that defaulted.

Historically, the corporate bond market has been the main source of credit-risk data. In recent years high-quality data on credit spreads have become available from the Credit Default Swap (CDS) market and from the secondary loan market. We choose bond market price data for this study for a few reasons. First, the bond market has a longer history of data available, which allows us to incorporate macroeconomic variables and establishing behavior over a longer period. Indeed, the quality of credit default swap data before 2000 is questionable, and our corporate bond data extends back a decade beyond 2000. Second, the corporate bond market still has very wide coverage of names, which gives us access to a larger universe. Third, although some authors claim that default swap spreads are less confounded by illiquidity, tax, and various market microstructure effects, trading in the CDS market is infrequent, with each issuer having about one trade or quote per trading

Table 1
Data sample.

Initial sample		Sample after 1st screen to obtain firm credit-spread factors		Sample after 2nd screen after obtaining credit-spread factors		Sample after 3rd screen of obtaining Compustat/CRSP data	
Firms	Firm months	Firms	Firm months	Firms	Firm months	Firms	Firm months
Panel A: The screening process							
387	14,234	340	14,049	256	13,589	241	11,894

Industry	Firms	Firm months	Credit rating	Firms	Firm months
Panel B: the final sample					
Manufacturing-sector	92	4034	Investment grade	194	9709
Service-sector	149	7860	Below investment grade	47	2185

Notes: Panel A shows the number of firms and firm × months in our sample, step-by-step through our screening process. Our sample comprises industrial, banking, and services sector firms in the S&P 500 Index at any time during the period 1990–2005. The first column shows data particulars of the initial set of firms resulting from our data collection screens that are detailed in Appendix B. In the first screen, we require a minimum of five prices of bonds of different maturities that span at least 7 years for each month to estimate the credit spread level, slope and curvature factors. In the second screen, we drop all firm-months that do not have at least six consecutive months of reasonable credit spread level, slope, and curvature factor estimates. In the third screen, we need data from the Center for Research in Security Prices (CRSP) and quarterly Compustat to obtain our firm-specific risk measures. This results in our final sample. Panel B shows the proportion of firms and firm months in our final dataset falling under two industry and rating cohorts. We obtain bond ratings from quarterly Compustat (augmented with data from Bloomberg), and assign numerical scores for the ratings starting with a score of 1 for AAA rating, 2 for AA rating and so on. We then segregate all firms into two overall groups: those whose bonds are rated BBB and above based on the average score of all bonds of that firm (the investment grade firms), and those that are rated below BBB (the below investment grade firms).

day; CDS spreads are often larger than corporate bond spreads, which would be unlikely if CDS spreads contain no liquidity premium.[7]

For each of our eligible firms, and for each month we perform the optimization routine given by Eq. (7). Bond prices derived from our optimization routine fit the data extremely well, as shown by the histogram of errors in Fig. 3.

Over 90% of the bonds could be fitted to within 1.0 dollar of their price, and over 75% (50%) of all bonds were priced within an error of 65 (33) cents. The average absolute error was 45 cents. In sum, our firm credit-spread curves implied out from data on corporate bonds have the property of being able to accurately replicate the market prices of bonds drawn from across the maturity spectrum.

Fig. 4 illustrates the time series of yield curves for a representative firm in our sample. The graph clearly shows that a firm's risky bond yield curve can indeed be upward-sloping, downward-sloping or hump-shaped, and change shape over time.

Furthermore, credit spreads along the maturity spectrum do not always move in the same direction. For example, for our data set, the 3-year and 5-year credit spreads moved in the same direction 77.5% of the time, 3-year and 7-year credit spreads moved in the same direction 67% of the time, and 3-year and 10-year credit spreads moved in the same direction 63% of the time. The results confirm that shocks to the credit-spread curve need not be parallel; they also suggest that the primary drivers of short-term credit spreads may be quite different from the drivers of longer-dated credit spreads. This is the reason we use the credit spread level, slope, and curvature (beta) factors, which represent the entire term structure of credit spreads, in the analysis in this paper.

Panel A of Table 2 compares firm specific characteristics by our industry and credit ratings groups. Leverage and book-to-market ratio are, on average, significantly higher for the below investment grade firms than for the investment grade firms, and for the service-sector firms than for the manufacturing firms. Stock return momentum is, on average, significantly higher for the investment-grade firms than for the below-investment-grade firms, and for the service-sector firms than for the manufacturing firms.

[7] For example, 19 of the 33 reference entities studied by Blanco et al. (2005) have average CDS spreads larger than their corresponding corporate-bond yield spreads. For more discussions on this issue see Berndt et al. (2005), Ericsson et al. (2005), Longstaff et al. (2005), Pan and Singleton (2008) and Tang and Yan (2007).

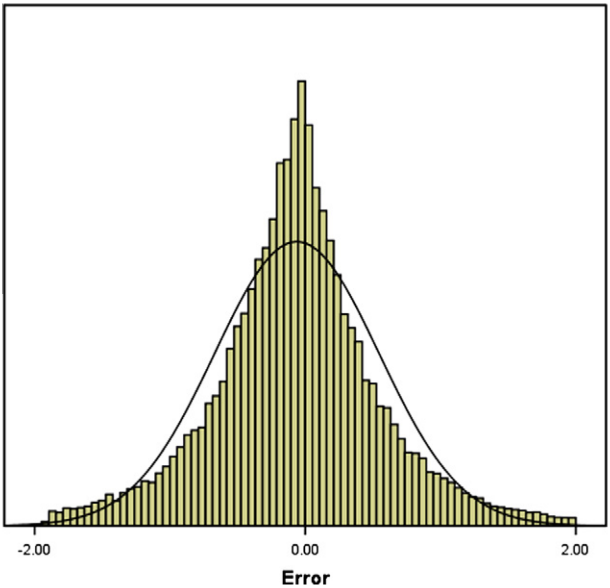

Fig. 3. Histogram of pricing errors. This figure shows the histogram of bond pricing errors obtained from our optimization routines for each firm-month combination in our sample. The bond pricing errors (calculated as the bond trade price from Bloomberg minus the fitted bond price obtained from our model) are in dollars.

Panel B of Table 2 shows the average credit-spread level (β_1), slope ($-\beta_2$), and curvature (β_3) factors across firm-months for firms grouped according to their industry or credit ratings. The level of credit spreads for low grade firms is on average 18 basis points higher than for investment grade bonds, and the difference is significant. The slopes and curvatures of service-sector firms are significantly different from those of manufacturing-sector firms.

Overall this table shows that firm characteristics as well as the credit-spread factors are different both across our two credit ratings groups, and across our two industry groups. Thus, it would be

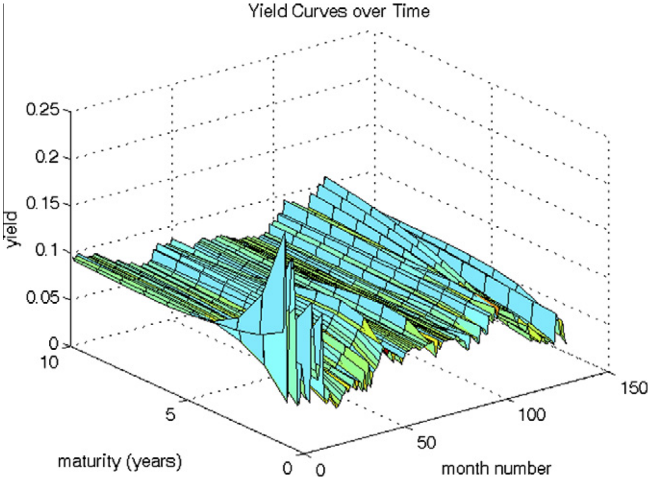

Fig. 4. Corporate bond yield curves for an illustrative firm. This figure shows the time series of risky yield curves for a representative firm, Altria, constructed using the estimated credit-spread level, slope, and curvature parameters.

Table 2
Firm characteristics and credit-spread factors.

Firm risk variables	Industry			Credit rating		
	Service	Manufacturing	Difference	Below investment grade	Investment grade	Difference
Panel A						
Leverage	0.74	0.34	0.40*** (38.14)	0.70	0.58	0.12*** (7.00)
Book-to-market ratio	0.72	0.47	0.25*** (29.01)	0.76	0.61	0.15*** (11.32)
Stock momentum	0.17	0.15	0.02** (2.01)	0.13	0.17	−0.04*** (−6.99)
Stock volatility (annualized)	0.21	0.20	0.01 (1.70)	0.21	0.21	0.00 (0.85)
Credit-spread factor						
Panel B						
Level	239.69	244.54	−4.85 (−1.59)	255.79	238.08	17.71*** (4.60)
Slope	28.72	21.79	6.93*** (6.89)	26.50	26.34	0.16 (0.13)
Curvature	−15.20	−28.54	13.34*** (8.03)	−20.18	−17.70	−2.48 (−1.24)

***, **, * and denote significantly different from zero at the 1%, 5%, and 10% levels, respectively.
Notes: Panel A shows the average firm characteristics – the average leverage, book-to-market ratio, monthly stock momentum, and stock volatility (reported on an annualized basis), for firms grouped by credit ratings (investment grade and below investment grade), and industry (manufacturing and service). Leverage is the ratio of debt outstanding on the balance sheet of the firm (Compustat Quarterly data item 51) and the market value of its common stock, computed monthly as the product of the number of shares outstanding and the closing share price each month (Compustat Quarterly data items 61 and 14). The book value of equity is defined as stockholders' equity plus balance sheet deferred taxes and investment tax credit, minus book value of preferred stock, which are, respectively, Compustat Quarterly data items 60, 52, and 55. The market value of equity is defined as the number of shares outstanding multiplied by the end of respective month closing stock price, which are, respectively, Compustat Quarterly data items 61 and 14. Stock momentum is the cumulative 12 monthly holding period returns from end of month t-13 through month t-2 from CRSP. Stock volatility is the sum of the daily squared holding period returns divided by number of observations in a month, from CRSP. The monthly volatility is presented as annualized numbers. Panel B shows the average credit spread level, slope and curvature factors across firm-months for firms segregated by their industry or credit ratings. The level (beta 1), slope (negative of beta 2) and curvature (beta 3) factors are measured in basis points. Difference of means, along with their *t*-statistics in parenthesis, between manufacturing firms and service-sector firms, and between the below-investment-grade firms and investment-grade firms are shown.

meaningful to run analyses separately groups of firms that are double-sorted based on their credit ratings and industry groupings.

4. In-sample prediction of riskless yields and index credit spreads

4.1. Predicting riskless yields

Panel A of Table 3 shows the results of the regressions of 6-month-ahead riskless yields against the current yield and with two additional yields. Specifically:

$$y_f^{(n)}(t+h) = \gamma_0^{(n)} + \gamma_1^{(n)} y_f^{(3)}(t) + \gamma_2^{(n)} y_f^{(5)}(t) + \gamma_3^{(n)} y_f^{(10)}(t) + \epsilon^{(n)}(t+h). \tag{11}$$

For example, the future 5-year yield is regressed against the current 5-year yield, the 3-year yield and the 10-year yield. The results show that the future level of each of the three yields depends not only on their current level, but also on the shape of the yield curve.

The shape of the yield curve is captured by the riskless-yield-curve level, slope and curvature factors. So, in the first three columns of panel B, we use the 3-vector of current month riskless factors, and then add successively our blocks of 1-month lagged factors, the current-month macro variables, and the 1-month lagged macro variables. Specifically:

$$y_f^{(n)}(t+h) = \eta_{0f}^{(n)} + \eta_{1f}^{(n)} \beta_f(t) + \eta_{2f}^{(n)} \beta_f(t-1) + \eta_{3f}^{(n)} M(t) + \eta_{4f}^{(n)} M(t-1) + \epsilon^{(n)}(t+h). \tag{12}$$

The impact of incorporating different blocks of variables is assessed for the 3-, 5-, and 10-year yields. Notice that the current 3-vector of riskless parameters can explain more than 72% of the variability of the future riskless yields across the maturity spectrum. The panel also shows that collectively, the lagged riskless parameters do not increase explanatory power. The macro variables, however, do add to the explanatory power, as do the lagged macro variables. Although statistically significant, the contribution of macro variables to the R^2 value is small, decreasing from about 4% for the three year yield to about 1% for the 10-year yield.

Since the source of predictability of riskless yields stems from the predictability of riskless beta factors, the last three columns of panel B report the same statistics as panel B except that the dependent variables are now the 6-month ahead riskless level, slope and curvature factors. More than 77% of the variability of the future riskless level and slope factors can be explained by the current level, slope and curvature factors, while 45% of the variability of the future riskless curvature can be explained by the current level, slope and curvature parameters. The vector of current-month macro-variables increases the explanatory power.

Panel C reports the R^2 values as well as the significance of each block of explanatory variables based on partial R^2 values when future 6-month ahead riskless yields are regressed on current month level, slope and curvature factors successively. Panel C shows that future yields depend not only on their current level, but also on the shape of the yield curve. That is, the slope and the curvature of the current period yield curve also contain statistically significant information (over and above the level) that help predict future yields, albeit with varying degrees depending on the maturity of the future yields, thereby more explicitly corroborating the evidence presented in panel A.

The implications of the above results are consistent with the findings of Cochrane and Piazzesi (2005), Ludvigson and Ng (2009), and Diebold et al. (2006), namely that the shape of the yield curve is useful for predicting future levels of yields and that macroeconomic factors can be used to improve the forecasts. We now proceed to investigate whether similar results hold for the predictability of credit spreads.

4.2. Predicting B-rated index credit spreads

The 6-months-ahead 3-, 5-, and 10-year credit spreads for the B-rated index are regressed against the current credit-spread-index factors, their lagged values, the riskless factors, their lagged values, the macro variables, and their lagged values in a hierarchical regression:

$$s_I^{(n)}(t+h) = \eta_{0I}^{(n)} + \eta_{1I}^{(n)}\beta_I(t) + \eta_{2I}^{(n)}\beta_I(t-1) + \eta_{3I}^{(n)}\beta_f(t) + \eta_{4I}^{(n)}\beta_f(t-1) + \eta_{5I}^{(n)}M(t) + \eta_{6I}^{(n)}M(t-1) + \epsilon_I^{(n)}(t+h). \quad (13)$$

The R^2 as well as the significance of each block of explanatory variables based on partial R^2 values are reported in the first three columns of panel A of Table 4.

We find that the current beta factors explain more than 44% of the variability of future index credit spreads across the maturity spectrum. The lagged factors contribute a small but significant amount to explanatory power. More important is the role of the riskless beta factors, which explain about 20% of the remaining variability, with the exact amount depending on the credit maturity. Lagged riskless factors play an insignificant role, but current macro variables explain over 30% of the remaining variability, with their lags explaining a small additional amount.

The last three columns of panel A report the results of hierarchical regressions where the future credit spreads in the regression equation (13) are replaced with the credit-spread factors. Results are similar to those reported when we examined future credit spreads. Most of the variability of future credit spread factors can be explained by the current period credit-spread beta factors, current riskless-curve beta factors, and the current macro variables. These three blocks of explanatory variables explain about 75% of the variability of future credit-spread levels, about 80% of the variability of future credit-spread slopes, and about 65% of the variability of future credit-spread curvature factors.

Since the current beta factors are significant in predicting future credit spreads, we dig deeper to examine the importance of each current period credit-spread factor. Panel B reports the R^2 values as well as the significance of each block of explanatory variables based on partial R^2 values when future 6-month ahead credit spreads are regressed on current month level, slope and curvature factors successively. The results show that future credit spreads depend not only on their current level, but

Table 3
Predicting 6-months-ahead riskless yields and factors.

	6-Months-ahead 3-year yield	6-Months-ahead 5-year yield	6-Months-ahead 10-year yield
Panel A			
Current month 3-year yield	4.556*** (5.61)	3.937*** (5.93)	3.634*** (7.11)
Current month 5-year yield	−6.795*** (−4.60)	−6.323*** (−5.23)	−6.531*** (−7.00)
Current month 10-year yield	3.203*** (4.66)	3.358*** (5.95)	3.871*** (8.79)
Intercept	0.002 (0.96)	0.002 (1.04)	0.003 (1.42)

	6-Months-ahead 3-year yield	6-Months-ahead 5-year yield	6-Months-ahead 10-year yield	6-Months-ahead level	6-Months-ahead slope	6-Months-ahead curvature
Panel B						
Current month 3-vector of riskless factors	0.727	0.726	0.752	0.781	0.773	0.456
1-Month lagged 3-vector of riskless factors	0.729	0.728	0.754	0.790*	0.803***	0.465
Current month 4-vector of macro variables	0.771***	0.755***	0.765**	0.799**	0.865***	0.574***
1-Month lagged 4-vector of macro variables	0.781**	0.767**	0.776**	0.807*	0.871*	0.603**

	6-Months-ahead 3-year yield	6-Months-ahead 5-year yield	6-Months-ahead 10-year yield
Panel C			
Current month level	0.308	0.445	0.636
Current month slope	0.671***	0.703***	0.745***
Current month curvature	0.727***	0.726**	0.752*

***, **, and * denote significantly different from zero at the 1%, 5%, and 10% levels, respectively.
Notes: Panel A shows the regression coefficients and, in parentheses, the associated *t*-statistics that are based on robust standard errors, when the future 6-month ahead 3, 5 and 10-year riskless yields are regressed on the current month 3-, 5-, and 10-year riskless yields. The sample period is 1970–2005. The first three columns of panel B report the R^2 values as well as the significance of each block of explanatory variables based on partial R^2 values when future 6-month ahead riskless yields are regressed on four blocks of variables: the 3-vector of current month level, slope and curvature factors, the 1-month lagged 3-vector of level, slope, and curvature factors, the 4-vector of current month macro-variables, and finally, the 1-month lagged 4-vector of macro variables. The last three columns report the same statistics as the first three columns except that the dependant variables are now the 6-month-ahead riskless level, slope, and curvature factors. Panel C reports the R^2 values as well as the significance of each block of explanatory variables based on partial R^2 values when future 6-month ahead riskless yields are regressed on current month level, slope and curvature factors successively. The significance of partial R^2 values is indicated based on the *p*-values of the partial *F*-statistic of the block of variables. The sample period in panels B and C is 1990–2005.

also on the slope and the curvature of the current period credit-spread curve. Thus, the "parallel shift" assumption does not hold: the slope and the curvature of the current period yield curve also contain statistically significant information (over and above the level) that help predict future credit spreads.

In general, the results of this table are analogous to those of the previous table on riskless yield predictions. The shape of the current credit-spread curve contains information for predicting future credit-spread curves, and forecasts can be further enhanced using contemporaneous information on the riskless-yield curve and macroeconomic variables.

5. In-sample prediction of firm credit spreads

The above analyses for riskless yield curves and aggregated credit-spread curves illustrate that the shape of the yield and credit-spread curves are informative for prediction purposes, and that there

Table 4
Predicting 6-months-ahead B-rated-index credit spreads and factors.

	6-Months-ahead 3-year credit spread	6-Months-ahead 5-year credit spread	6-Months-ahead 10-year credit spread	6-Months-ahead level	6-Months-ahead slope	6-Months-ahead curvature
Panel A						
Current month index factors	0.481	0.449	0.452	0.508	0.737	0.564
1-Month lagged index factors	0.505*	0.474*	0.478*	0.531*	0.760***	0.577
Current month riskless factors	0.596***	0.582***	0.603***	0.666***	0.821***	0.659***
1-Month lagged riskless factors	0.597	0.582	0.603	0.666	0.825	0.663
Current month macro variables	0.734***	0.735***	0.759***	0.799***	0.839***	0.682**
1-Month lagged macro variables	0.758**	0.757**	0.776**	0.809*	0.846*	0.688

	6-Months-ahead 3-year credit spread	6-Months-ahead 5-year credit spread	6-Months-ahead 10-year credit spread
Panel B			
Current month index level	0.158	0.196	0.228
Current month index slope	0.181**	0.221**	0.255**
Current month index curvature	0.481***	0.449***	0.452***

***, **, and * denote that partial R^2 is significantly different from zero at the 1%, 5%, and 10% levels, respectively.
Notes: The first three columns of panel A report the R^2 values as well as the significance of each block of explanatory variables based on partial R^2 values, when future 6-months-ahead 3, 5 and 10-year B-rated-index credit spreads are regressed on six blocks of variables: the 3-vector of current month level, slope, and curvature index factors, the 1-month lagged 3-vector of level, slope, and curvature index factors, the 3-vector of current month riskless factors, their 1-month lagged values, the 4-vector of current month macro variables, and finally their 1-month lagged values. The last three columns report the same statistics except that the dependant variables are now the 6-months-ahead B-rated credit spread level, slope and curvature factors. Panel B reports the R^2 values as well as the significance of each block of explanatory variables based on partial R^2 values when future 6-month-ahead credit spreads are regressed on current month level, slope and curvature factors successively. The significance of partial R^2 values is indicated based on the p-values of the partial F-statistic of the block of variables. The sample period analyzed is 1992–2005.

may be a significant additional role played by macroeconomic variables. We now turn to our main focus which is on predictability of credit spreads at the firm level.

5.1. The importance of current and lagged credit-spread curve factors

We forecast the h-month ahead credit spreads, based on their current values, and on other points on the credit-spread curve. Specifically, for $n = 3, 5, 10$ years we consider,

$$s_j^{(n)}(t+h) = \gamma_{0,j}^{(n)} + \gamma_1^{(n)} s_j^{(n)}(t) + \gamma_2^{(n)} s_j^{(\text{short})}(t) + \gamma_3^{(n)} s_j^{(\infty)}(t) + \epsilon_j^{(n)}(t+h), \tag{14}$$

where $h = 6$ months, the long credit spread $s_j^{(\infty)}(t) = \beta_{j1}(t)$, and the short credit spread $s_j^{(\text{short})}(t) = \beta_{j1}(t) + \beta_{j2}(t)$. We use panel regression methodology, with firm fixed effects, run over groups of firms that are double-sorted based on their credit rating and industry. In such panel regressions, the explanatory variables and residuals can be correlated, in which case White standard errors are biased, while standard errors corrected for clustering by firms that are also robust to heteroskedasticity are well-specified [see Petersen (2009)]. So we compute standard errors that are robust to heteroskedasticity and firm-clustering. Since the results for the 3- and 10-year spreads are qualitatively similar to those of the 5-year spreads, we tabulate only the results for the 5-year spreads. The R^2 as well as the significance of each block of explanatory variables based on partial R^2 values are reported in panel A of Table 5.

Does the first factor (the parallel factor) alone help predict future credit spreads? Not surprisingly, we find that the future 5-year credit spread depends on the current 5-year credit spread. However, future credit spreads also depend on the shape of the credit-spread curve. Specifically, future credit spreads depend on the current long and short credit spreads as well. A partial F-test reveals that these two additional points add significantly to the explanatory power of future 5 year credit spreads. In other words, although the current period level is very important in predicting future credit spreads, there is significant additional information left on the table: the slope and curvature of the current credit-spread curve are informative as well.

Table 5
Predicting 6-months-ahead 5-year credit spreads.

Block of variables	6-Months-ahead 5-year credit spread		
Panel A: panel regressions			
Firm fixed effects	0.351	0.351	0.351
Current month same-maturity credit spread	0.579***		
Current month long credit spread	0.601***		
Current month short credit spread	0.626***		
Current month Factors		0.626***	0.626***
1-Month lagged factors		0.634**	0.634**
2-Month lagged factors		0.636	
3-Month lagged factors		0.638	
Riskless factors			0.685***
B-Rated index factors			0.691**
Macro variables			0.697**
Firm variables			0.704**

Block of variables	Average adjusted R^2	Proportion of firms that are significant at the		Average adjusted R^2	Percentage of firms that are significant at		Average absolute error	Average RMSE
		10% level	5% level		10% level	5% level		
Panel B: firm-by-firm regressions								
Current month factors	0.435	81.7	77.9	0.435	81.7	77.9	28.82	43.19
1-Month lagged factors	0.460	35.8	28.5	0.460	35.8	28.5	28.48	42.49
2-Month lagged factors	0.477	21.9	14.0					
3-Month lagged factors	0.462	16.3	10.6					
Riskless factors				0.595	74.5	69.6	25.77	39.30
B-Rated index factors				0.675	55.6	47.7	25.31	38.96
Macro variables				0.702	60.5	52.6	25.02	38.63
Firm variables				0.731	56.9	52.7	25.02	38.46

***, **, and * denote that partial R^2 is significantly different from zero at the 1%, 5%, and 10% levels, respectively.
Notes: Panel A reports the R^2 values as well as the significance of each block of explanatory variables based on partial R^2 values for three different specifications when the 6-months-ahead 5-year credit spreads are regressed on blocks of explanatory variables. Panel regressions (after controlling for firm fixed effects) are run over groups of firms double-sorted by credit ratings and industry, and the results consolidated. The significance of the partial R^2 values is based on the p-values of the partial F-statistic of the block of variables. Panel B shows the average R^2 and the average adjusted R^2 values when firm-by-firm regressions are used. Also shown are the proportion of firms for which each block of variables was significant at the 10% and 5% levels, the average absolute error, and the average root mean square error (RMSE).

With just the current 5-year credit spread as an independent variable, the coefficient on the spread is significantly different from zero but significantly lower than one, implying that the changes in credit spreads do not follow a random walk. About 35% of the variability of the 5-year credit spread is attributable to differences in firms. That is, there is significant within-firm variability (the remaining 65%) that still needs to be explained. We see that once the effects of different firms have been removed, we can account for 42% of the variance in future 5-year credit spreads with three points on the current credit-spread curve. The importance of the shape of the credit-spread curve parallels the findings of Cochrane and Piazzesi (2005) for the riskless term structure, and are consistent with our findings for riskless yield curves and B-index credit-spread curves.

Rather than use the current level of specific points on the credit-spread curve as independent variables, we could use the three credit spread state variables, namely the level, β_1, the slope, β_2, and the curvature, β_3, as the independent variables. This leads to a panel regression model that gives the same R^2 number as reported in the first column, namely, 0.626. All three factors are significant for our three maturities of 3, 5 and 10 years.

We next incorporate information on lagged credit spreads. Specifically, we consider panel regression models of the form:

$$s_j^{(n)}(t+h) = \gamma_{0,j}^{(n)} + \gamma_1^{(n)}(L)\beta_{1j}(t) + \gamma_2^{(n)}(L)\beta_{2j}(t) + \gamma_3^{(n)}(L)\beta_{3j}(t) + \epsilon_j^{(n)}(t+h). \quad (15)$$

In this autoregressive distributed lag model $\gamma_i^{(n)}(L)\beta_{kj}$ represents a lag polynomial, so that

$$\gamma_i^{(n)}(L)\beta_{kj}(t) = \gamma_{i1}^{(n)}\beta_{kj}(t) + \gamma_{i2}^{(n)}\beta_{kj}(t-1) + \cdots + \gamma_{ip}^{(n)}\beta_{kj}(t-p),$$

where p is the number of lagged variables.

The middle column of Table 5 reports on the significance of each block of variables. With firm fixed effects and the current credit spread factors included in the model, the R^2 is 0.626, as reported in the first regression. Adding 1-month lagged credit-spread factors increases the R^2 to 0.634. A partial F-test at the 1% level rejects the null hypothesis that the lags are not significant. Incremental additions of 2 and 3 month lags are, however, not statistically significant.

We repeat the analysis, but this time run the above regression model firm by firm. The left columns of panel B of Table 5 reports the results. The first column shows the average adjusted R^2 values over all firms. The inclusion of all three lags, increases the average adjusted R^2 value from 0.435 to 0.462. The next two columns reports the proportion of firms for which the incremental contribution of the lagged variables was significant at the 10% and 5% levels of significance. For 35.8% (28.5%) of our firms incorporating the first lag of credit-spread factors is significant at the 10% (5%) level. The proportions drop as the lag increases.

Overall, the shape of the current credit-spread curve contains significant information for forecasting future credit spreads; incorporating information contained in the lagged-period credit-spread curve may modestly enhance the predictive power of future credit spreads.

5.2. The importance of auxiliary information

We now evaluate whether the credit-spread curve reflects all known information relevant for forecasting future credit spreads, or whether there are other variables that can be used to improve predictability. Beyond our set of current and lagged credit spread factors, other blocks of explanatory variables include the three vector of riskless factors, the three vector of B-index credit spread factors, the four vector of macro variables, and the four vector of firm-specific risk variables.

We examine the sequential importance of our blocks of variables by running a hierarchical panel regression model of the form:

$$s_{jk}^{(n)}(t+h) = \eta_{0,jk} + \eta_{1k}\beta_{jk}(t) + \eta_{2k}\beta_{jk}(t-1) + \eta_{3k}\beta_f(t) + \eta_{4k}\beta_I(t) + \eta_{5k}M(t) + \eta_{6k}F_{jk}(t) + \epsilon_{jk}^{(n)}(t+h), \quad (16)$$

where j references the firm and $k = 1,2,3,4$ represents the industry-ratings group. Our goal is to investigate the hypothesis that $\eta_{3k} = \eta_{4k} = \eta_{5k} = \eta_{6k} = 0$.

Table 6
Relative importance of explanatory variables for predicting credit spreads.

	Short credit spread	0.5 Year	1 Year	3 Year	5 Year	7 Year	10 Year	20 Year	Long credit spread
Panel A									
Firm fixed effects	0.198	0.203	0.209	0.261	0.351	0.327	0.307*	0.237***	0.194***
3-Vector of current credit spread factors	0.321***	0.336***	0.357***	0.513***	0.626***	0.594***	0.520***	0.375***	0.302***
3-Vector of lagged credit spread factors	0.345***	0.361***	0.383***	0.536***	0.634***	0.607***	0.537***	0.392***	0.318***
3-Vector of riskless factors	0.357***	0.378***	0.405***	0.595***	0.685***	0.627***	0.545***	0.401***	0.333***
3-Vector of index factors	0.357	0.378	0.406*	0.599**	0.691**	0.632**	0.555***	0.425***	0.358***
4-Vector of macro variables	0.362**	0.384**	0.412**	0.605**	0.697**	0.639***	0.565***	0.434***	0.365***
4-Vector of firm variables	0.364	0.386	0.414	0.609*	0.704**	0.647**	0.572**	0.440**	0.370**
Panel B									
Firm fixed effects	0.198	0.203	0.209	0.261	0.351	0.327	0.307*	0.237***	0.194***
4-Vector of firm variables	0.200	0.206	0.214	0.294***	0.388***	0.400***	0.367***	0.271***	0.214***
4-Vector of macro variables	0.249***	0.264***	0.281***	0.398***	0.481***	0.491***	0.470***	0.358***	0.276***
3-Vector of index factors	0.268***	0.289***	0.316***	0.488***	0.555***	0.516***	0.478***	0.385***	0.318***
3-Vector of riskless factors	0.281***	0.304***	0.335***	0.527***	0.586***	0.532***	0.489***	0.395***	0.327***
3-Vector of lagged credit spread factors	0.337***	0.359***	0.388***	0.588***	0.680***	0.625***	0.554***	0.424***	0.353***
3-Vector of current credit spread factors	0.364***	0.386***	0.414***	0.609***	0.704***	0.647***	0.572***	0.440***	0.370***

***, **, and * denote that partial R^2 is significantly different from zero at the 1%, 5%, and 10% levels, respectively.
Notes: This table reports the R^2 values as well as the significance of each block of explanatory variables based on partial R^2 values, when future 6-months-ahead credit spreads of different maturities are regressed sequentially on the successive blocks of explanatory variables. The sequence of the blocks of explanatory variables is reversed in panel B as compared to panel A. Panel regression controlling for firm fixed effects are run by groups that are based on ratings and industry, and the results consolidated. The significance of each block of explanatory variable is based on the *p*-values of the partial *F*-statistics.

The rightmost column of panel A of Table 6 report the R^2 values as well as the significance of each block of explanatory variables based on partial R^2 values, as successive blocks are added to the set of explanatory variables. Once the fixed firm effects as well as the current and lagged credit-spread factors are removed, the riskless credit spread factors account for almost 14% of the remaining unexplained variability. And once this has been accounted for, the B-rated index factors, the macro variables, and the firm variables each account for about 2%, which is statistically significant at the 5% level. We also examine the explanatory power of our sets of variables for the 3- and 10-year credit spreads for all firms as well as firms segregated by industry group and ratings. Collectively, we can explain anywhere from 60% to 80% of the variability of future credit spreads by our set of independent variables, depending on the group.

Panel B shows the results for firm-by-firm regressions. The average adjusted R^2 values increase as additional blocks are added. After the firm's current and lagged credit-spread factors are accounted for, the impact of the block of riskless-yield factors is significant at the 5% level for 70% of all firms. The additional blocks – B-index factors, macro variables, and firm-specific variables – are all incrementally significant in about 50% of the firms. We study those firms for which macro and firm-specific variables are significant at the 10% level, given all other blocks of explanatory variables, and classify them according to industry and ratings. The firms for which macro and firm-effects are significant do not concentrate in any of these groups. Moreover, for the firms for which macro variables were significant, the median incremental contribution to adjusted R^2 values is of the order of 10% for each of the future credit-spread factors. The same conclusion can be drawn from the marginal impact of firm-specific risk variables. The firms for which firm-risk variables contribute significantly in the presence of other factors, do not concentrate in any industry or credit-ratings categories, and their contribution to the explanatory power is of a similar magnitude as the macro variables.

The last two columns report the average absolute errors and the average root mean square errors of each regression model, averaged over all firm-months. As the table shows, the average absolute prediction error drops to about 25 basis points when the current and lagged 3-vector of beta factors, and the current three vector of riskless factors are used to predict 6-month-ahead 5-year credit spreads. In comparison, the average absolute prediction error using the spot (forward) credit spread is 31.47 (44.34) basis points. The table also shows that the root-mean-squared prediction errors drop to about 39 basis points when the credit-spread and riskless factors are used. Beyond this, as the number of independent variables increases, the errors decrease, but at a very slow rate.

The results show that the shape of the current credit-spread curve contains significant information for forecasting future credit spreads, and additional information is contained in the riskless-yield-curve factors. The results also indicate that much smaller incremental contributions could come from additional factors.

The results of our hierarchical regressions have to be interpreted carefully. The contribution of macro and firm-specific-risk variables to the explanatory power of credit spreads, in the presence of all other factors is small. By themselves, however, macro and firm variables can explain about 48% of the variability of future 5-year credit spreads. However, much of this explanatory power is subsumed by information contained in the credit-spread and riskless factors. To emphasize this point, in Table 6, we perform panel regressions of the form in Eq. (16) over an array of maturities. In the top panel, we report the sequential R^2 values as well as the significance of each block of explanatory variables based on partial R^2 values, as blocks of variables are added. The bottom panel repeats the analysis, but reverses the order of the blocks so that firm variables are the first block.

Consider the results for the 3-year credit spread. With credit spread and riskless factors, 40% of the variability is explained. Macroeconomic and firm-specific variable blocks, being the last two, collectively explain an additional amount of less than 1%. In contrast, panel B shows that firm and macro variables explain 28% of the variability. The riskless, B-index, and credit-spread factors, however, as the last blocks, collectively explain an incremental amount of over 13%. Thus, credit spread and riskless factors are extremely informative, and contain most of the information that is necessary for forecasting credit spreads. This result holds consistently across the maturity spectrum.

6. Out-of-sample prediction of firm credit spreads

We now consider rolling out-of-sample predictions. We begin by using information over an initial training period to estimate parameters for our regression models. Then, using all historical information known to the market up to date *t*, we predict future credit spreads. We repeat this procedure over consecutive months, using all our models and using panel regressions as well as firm-by-firm regressions.

We consider several models for predicting future credit spreads. The first model, the spot model, uses the current credit spread to predict future credit spread. The second model, the forward model, uses the appropriate forward credit spread to predict future credit spread. These two models are the benchmark random-walk models against which the performance of additional prediction models is evaluated. In our first model, M_1, predictions are based on the current credit-spread level, slope, and curvature factors. Successive models include information on additional blocks of variables in a hierarchical fashion. Model M_2 uses information on both current period credit-spread factors and lagged credit-spread factors; Model M_3, includes information on the riskless factors; Model M_4, includes information on the B-index credit spread factors; Model M_5, includes information on the macro variables; and Model M_6 includes all these variables and adds firm-specific-risk variables on top of all the above-mentioned variables.

The top panel of Table 7 reports the average bias and standard deviation as well as the average absolute error and standard deviation, when panel regression methodology (controlling for firm fixed effects run on groups of firms double sorted on the basis of credit ratings and industry) is used to estimate the 5-year credit spreads six months ahead. Model M_3 has the lowest average bias, the lowest average absolute error, and the lowest standard deviation.

We dig deeper to examine the importance of each current period credit-spread factor in Model M_3. We find that the average bias, average absolute error, and standard deviation all increase if we drop either the current period credit-spread slope or the credit-spread curvature factor from Model M_3. This corroborates the earlier in-sample prediction results show that future credit spreads depend not only on their current level, but also on the slope and the curvature of the current period credit-spread curve.

Table 7
Out-of-sample 6-months-ahead 5-year-credit-spread prediction errors using panel regressions.

	Errors		Absolute errors		SIC
	Average	Standard deviation	Average	Standard deviation	
Panel A					
Spot	7.96	30.84	37.34	30.61	515.7
Forward	−6.26	37.57	46.97	34.48	537.0
M_1	6.34	31.63	40.06	29.85	512.1
M_2	6.75	31.84	39.52	29.51	516.2
M_3	5.40	29.17	35.86	28.30	509.1
M_4	6.39	29.16	36.01	28.74	522.7
M_5	6.32	29.98	36.82	29.65	534.0
M_6	10.42	29.84	37.27	30.43	553.6

	25th Percentile	50th Percentile	75th Percentile	Proportion of firms, for which model has lower MSPE than Spot Model
Panel B				
Forward	1.05	1.51	2.39	0.22
M_1	0.88	1.12	1.51	0.40
M_2	0.91	1.10	1.45	0.36
M_3	0.70	0.92	1.24	0.58
M_4	0.72	0.94	1.22	0.56
M_5	0.79	1.01	1.41	0.48
M_6	0.82	1.04	1.44	0.42

Notes: This table shows the statistics of the 6-months-ahead out-of-sample prediction errors of 5-year credit spreads (where prediction error is defined as the actual 5-year credit spread minus predicted 5-year credit spreads in basis points). Panel regression methodology with firm fixed effects is used where the panel regressions are run by groups that are based on ratings and industry, and the results consolidated. Panel A shows the average out-of-sample 6-months-ahead prediction errors (the average bias) and the average of the absolute errors, along with standard deviations for different prediction models. The last column reports the Schwarz Information Criterion (SIC) values for the eight models that we examine. Panel B reports statistics of the distribution of the ratio of mean-squared prediction errors (MSPEs) for each model relative to that of the Spot Model by firms, and the proportion of firms for which each model has a lower MSPE relative of that of the Spot Model. Our credit spread prediction models, successively, include more blocks of variables. Model M_1 uses only the current credit-spread factors, M_2 adds the lagged credit-spread factors, M_3 adds the riskless factors, M_4 adds the B-rated-index factors, M_5 adds the macro variables, and finally, M_6 adds the firm variables on top of all the aforementioned blocks of variables.

For each model, we compute the MSPEs, and then compute the ratio of this value relative to the MSPE for the spot model. The bottom panel shows the quartiles of the ratios, followed by the proportion of firms for which a model produced MSPEs that were smaller than those of the spot model. The forward model underperforms the spot model, as do models M_1 and M_2. Model M_3, again, is the best. Indeed, 58% of the firms had smaller MSPEs using M_3 than using the spot model, which is significantly different from 0.5 at the 1% level of significance. Higher-order models performed worse than Model M_3 in the out-of-sample analysis.

To check whether our best model for predicting credit spreads, Model M_3, is not "overparameterized" in a statistical sense, we examine the Schwarz Information Criterion (SIC, 1978). The SIC trades off the residual sums of squares with a penalty for the loss of degrees of freedom from adding extra parameters or variables. This criterion embodies a stiffer penalty than other information criteria, such as the Akaike's Information Criterion or the Hannan–Quinn Criterion, to discourage overfitting. The preferred model is the one with the lowest SIC value. The last column of panel A reports the SIC values for the eight models that we examine. Model M_3 remains the best model for predicting out-of-sample credit spreads.

Table 8 shows the results of firm-by-firm regressions. As in the panel regression results, M_3 has the smallest bias (6 basis points) and the smallest average absolute error (31 basis points) of all the models examined, significantly lower than that of the spot model (that yields an average absolute error of 37 basis points). Information contained in variables in addition to the riskless factors do not improve the predictions of 6-month-ahead credit spreads. Moreover, the average bias, average absolute error, and standard deviation all increase if we drop either the current period credit-spread slope or the credit-spread curvature factor from Model M_3. Thus future credit spreads depend not

Table 8
Out-of-sample 6-months-ahead 5-year-credit-spread prediction errors using firm-by-firm regressions.

	Errors		Absolute errors		SIC
	Average	Standard deviation		Standard deviation	
Panel A					
Spot	7.96	30.84	515.7	30.61	515.7
Forward	−6.26	37.57	537.0	34.48	537.0
M_1	10.65	27.87	512.1	27.70	510.2
M_2	11.50	28.19	516.2	26.86	513.7
M_3	5.98	25.39	509.1	24.74	505.7
M_4	7.05	25.74	522.7	25.61	518.6
M_5	6.93	26.90	534.0	26.85	532.2
M_6	7.14	26.47	553.6	26.44	536.6
	25th Percentile	50thPercentile	75th Percentile	Proportion of firms, for which model has lower MSPE than Spot Model	
Panel B					
Forward	1.05	1.51	2.39	0.20	
M_1	0.74	0.92	1.08	0.63	
M_2	0.70	0.92	1.08	0.60	
M_3	0.63	0.79	0.95	0.84	
M_4	0.59	0.77	0.96	0.79	
M_5	0.66	0.89	1.04	0.73	
M_6	0.67	0.88	1.07	0.69	

Notes: This table shows the statistics of the 6-months-ahead out-of-sample prediction errors of 5-year credit spreads (where prediction error is defined as the actual 5-year credit spread minus predicted 5-year credit spreads in basis points). Firm-by-firm regressions are used. Panel A shows the average out-of-sample 6-months-ahead prediction errors (the average bias) and the average of the absolute errors, along with standard deviations for different prediction models. The last column reports the Schwarz Information Criterion (SIC) values for the eight models that we examine. Panel B reports statistics of the distribution of the ratio of MSPE for each model relative to that of the Spot Model by firms, and the proportion of firms for which each model has a lower MSPE relative of that of the Spot Model. Our credit spread prediction models, successively, include more blocks of variables. Model M_1 uses only the current credit-spread factors, M_2 adds the lagged credit-spread factors, M_3 adds the riskless factors, M_4 adds the B-rated-index factors, M_5 adds the macro variables, and finally, M_6 adds the firm variables on top of all the aforementioned blocks of variables.

only on their current level, but also on the slope and the curvature of the current period credit-spread curve.

The bottom panel shows that M_1, the simple credit-spread model, outperforms the spot model, in terms of MSPE, for 63% of the firms, adding lags does not improve performance, but adding riskless factors results in a model (M_3) that outperforms the spot model for 84% of the firms, and the 75th quartile of the ratio was 0.95, less than 1. Like the panel regression results, adding firm-specific and macro variables does not improve predictions of future credit spreads.

The last column of panel A reports the SIC values for the eight models that we examine. In terms of SIC also, Model M_3 is the best model for predicting out-of-sample credit spreads.

The relative performance of all the models in the firm-by-firm regression models is summarized in Fig. 5, which shows the MSPE for each model relative to the MSPE for the spot model in the form of box plots. The leftmost box and whisker plot is for the forward model, and its performance relative to the spot model is poor, with the 25th percentile exceeding 1. Using just the credit-spread factors produces a dramatic improvement, as the second box plot of the figure shows. Incorporating lagged credit spreads does not significantly improve predictions; adding riskless factors results in the best model (whose entire box plot fall below 1). Adding other blocks of variables does not improve the forecasts of future credit spreads.

The fact that the forward credit spread performs so poorly, relative to all the other models, is a result that has been obtained elsewhere, in riskless and foreign exchange markets. The fact that the forecasts can be improved using information, not only from the credit-spread curve, but also from the riskless term structure, is a new result. For each firm, and for each year we next compute the MSPEs for all our models based on firm-by-firm regressions. In panel A of Fig. 6, we compare each model's MSPE normalized by the MSPE of the spot model.

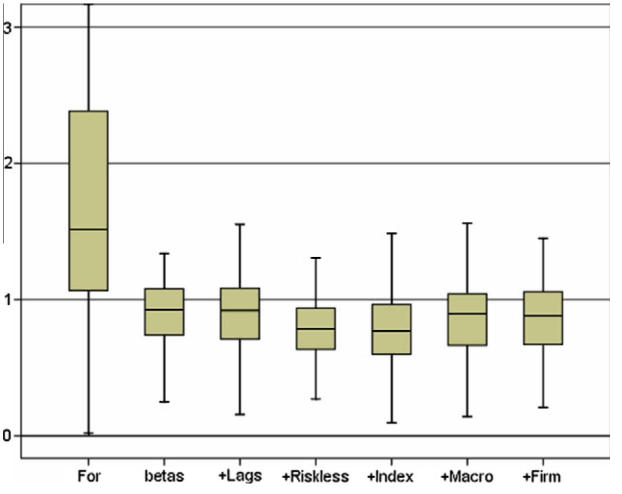

Fig. 5. Out-of-sample prediction errors of 5-year credit spreads: comparison of models. This figure shows the box plots of the ratios of mean-square prediction errors (MSPEs) for our different credit spread prediction models relative to the MSPE of the Spot Model, firm by firm, over all time periods. Our credit spread prediction models include successively more blocks of variables. The models shown are the Forward Model, Model M_1 that uses just the current credit-spread factors, Model M_2 that adds the lagged credit-spread factors, Model M_3 that adds the riskless factors, Model M_4 that adds the B-rated-index factors, Model M_5 that adds the macro variables, and finally, Model M_6 that adds the firm variables on top of all the aforementioned blocks of variables.

Panel A shows the time series of box plots of MSPE ratios for models M_1, M_2, M_3, and M_4, in that order, relative to the spot model. There is intertemporal variation, but Model M_3 outperforms the spot model in 8 of the 10 years. Panel B shows the time series of MSPE ratios for models M_3, M_5, and M_6 relative to the Spot Model. It is clear that there is no advantage in incorporating macro and firm variables. In particular, in every year except 2002, the model without macro and firm factors had the smallest MSPE ratios. To confirm the result that adding macro and firm variables does not improve out-of-sample prediction performance, we computed, for each firm and for each year, the ratio of the MSPEs for the three models that incorporated information beyond the riskless factors and normalized these values by the MSPEs of the model that incorporated information up to the riskless factors. Panel C shows these time series of MSPE ratios for models M_4, M_5 and M_6 relative to M_3. In all years except 2002, the median MSPE ratio exceeded 1 for the models that incorporated either macro or the macro and firm variables information (M_5 and M_6). The model that also uses information from the B-rated credit spreads (M_4) outperformed Model M_3 in 52% of occasions, although this is not significantly different from a tie, at the 1% level of significance. Indeed, models M_3 and M_4 had significantly lower average absolute errors in each year than either the spot or the forward models.

Overall, the results indicate that a parsimonious model that uses the information on the credit-spread and riskless factors yields predictions of future 6-months-ahead credit spreads that are significantly superior to those of the random-walk models. More information does not significantly improve the predictions; in fact, using macro and firm-specific variables only adds noise. Of course macro and firm-specific variables are important determinants of credit spreads. But the credit-spread and riskless curves essentially impound all marketwide and firm-specific information necessary for predicting future credit spreads.

6.1. Out-of-sample predictions of credit spreads of different maturities and forecast horizons

To check the robustness of our results, we examine the out-of-sample prediction errors of our various models for maturities other than 5 years and for forecast horizons other than 6 months. In this section we examine forecasts of 3-year and 10-year credit spreads; we also alter the forecast horizon from 6 months, first to 3 months and then to 12 months. Table 9 reports the average absolute errors by

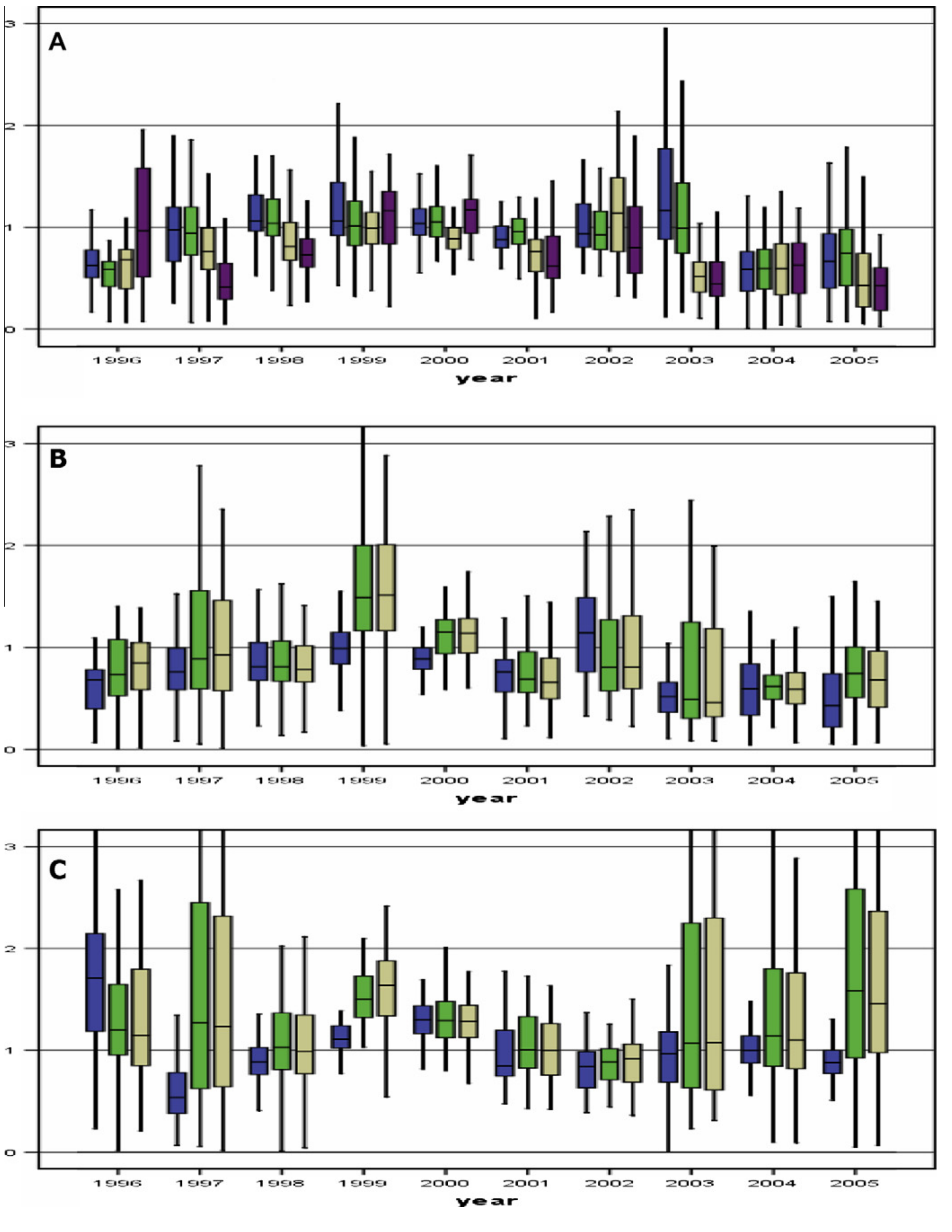

Fig. 6. Out-of-sample prediction errors of 5-year credit spreads: time series comparison of models. (A) The time series of box plots of MSPE ratios for Models M_1, M_2 M_3, and M_4, in that order, relative to the Spot Model. (B) The time series of MSPE ratios for Models M_3, M_5, and M_6 relative to the Spot Model. (C) The time series of MSPE ratios for Models M_4, M_5, and M_6 models relative to that of Model M_3. Our credit spread prediction models include successively more blocks of variables. Model M_1 uses just the current credit spread betas, M_2 adds the lagged credit-spread factors, M_3 adds the riskless factors, M_4 adds the B-rated-index factors, M_5 adds the macro variables, and finally, M_6 adds the firm variables on top of all the aforementioned blocks of variables.

Table 9
Robustness checks of out-of-sample credit-spread predictions using firm-by-firm regressions.

	6-Months-ahead 3-year credit spread		6-Months-ahead 10-year credit spread	
	Average absolute prediction error across firms	Proportion of firms, for which model has lower average MSPE than Model M_3	Average absolute prediction error across firms	Proportion of firms, for which model has lower average MSPE than Model M_3
Panel A				
Spot	60.05	0.15	44.16	0.25
Forward	61.05	0.17	51.67	0.15
M_1	54.67	0.11	40.93	0.46
M_2	54.08	0.10	40.08	0.51
M_3	48.27	N/A	40.98	N/A
M_4	48.30	0.60	41.58	0.45
M_5	52.40	0.28	42.89	0.26
M_6	51.05	0.28	44.45	0.31
	3-Months-ahead 5-year credit spread		12-Months-ahead 5-year credit spread	
	Average absolute prediction error across firms	Proportion of firms, for which model has lower average MSPE than Model M_3	Average absolute prediction error across firms	Proportion of firms, for which model has lower average MSPE than Model M_3
Panel B				
Spot	26.15	0.18	53.16	0.19
Forward	32.40	0.05	73.60	0.13
M_1	25.19	0.17	52.04	0.22
M_2	24.72	0.26	52.14	0.22
M_3	23.94	N/A	47.31	N/A
M_4	24.86	0.18	42.75	0.54
M_5	26.07	0.05	47.95	0.36
M_6	25.75	0.17	49.33	0.35

Notes: This table reports the average absolute errors by firms, and the proportion of firms for which each model has a lower MSPE relative of that of Model M_3. Our credit spread prediction models successively include more blocks of variables. Model M_1 uses only the current credit-spread factors, M_2 adds the lagged credit-spread factors, M_3 adds the riskless factors, M_4 adds the B-rated-index factors, M_5 adds the macro variables, and finally, M_6 adds the firm variables on top of all the aforementioned blocks of variables. Firm-by-firm regressions are used. Panel A shows these statistics for the 6-months-ahead 3- and 10-year credit spreads, while panel B shows these statistics for 3- and 12-months-ahead 5-year credit spreads.

firms, and the proportion of firms for which each model's MSPE is lower than that of Model M_3 based on firm-by-firm regressions.

The top panel shows that M_3 and M_4 are the best models for predicting 6 month ahead 3-year credit spreads, in terms of average absolute prediction errors as well as MSPE ratios. The best models for predicting 6-months-ahead 10-year credit spreads are M_2 and M_3. The bottom panel shows that M_3 is the best model for predicting 3 month ahead 5-year credit spreads, in terms of average absolute prediction errors as well MSPE ratios, whereas M_4 is the best model for predicting future 12-months ahead 5-year credit spreads. In each case, the proportion of firms for which Model M_3 yields a lower MSPE than either the spot or forward model is, at the 5% level, significantly higher than 50%.

These results are shown pictorially in Fig. 7. The four panels, respectively, show box plots of MSPE ratios of different models for predicting 6-months-ahead 3-year credit spread, the 6-month-ahead 10-year credit spread, the 3-month-ahead 5-year credit spread, and the 12-months-ahead 5-year credit spread models. The MSPE-ratio-box-plots in each panel are those of spot, forward, and models M_1, M_2, M_4, M_5, and M_6, all relative to the MSPE of Model M_3. The four panels show that none of the other models are significantly better than Model M_3 in terms of MSPEs because none of the box-plots of MSPE ratios fall below the 1 axis.

Overall, Tables 8 and 9 and Figs. 5–7, make two important points. First, the model that uses information contained in the current and lagged term structures of credit spreads as well as in the current riskless yield curve is a parsimonious model that generally leads to the best predictions for future n-month-ahead credit spreads of different maturities. Such a model leads to significantly better

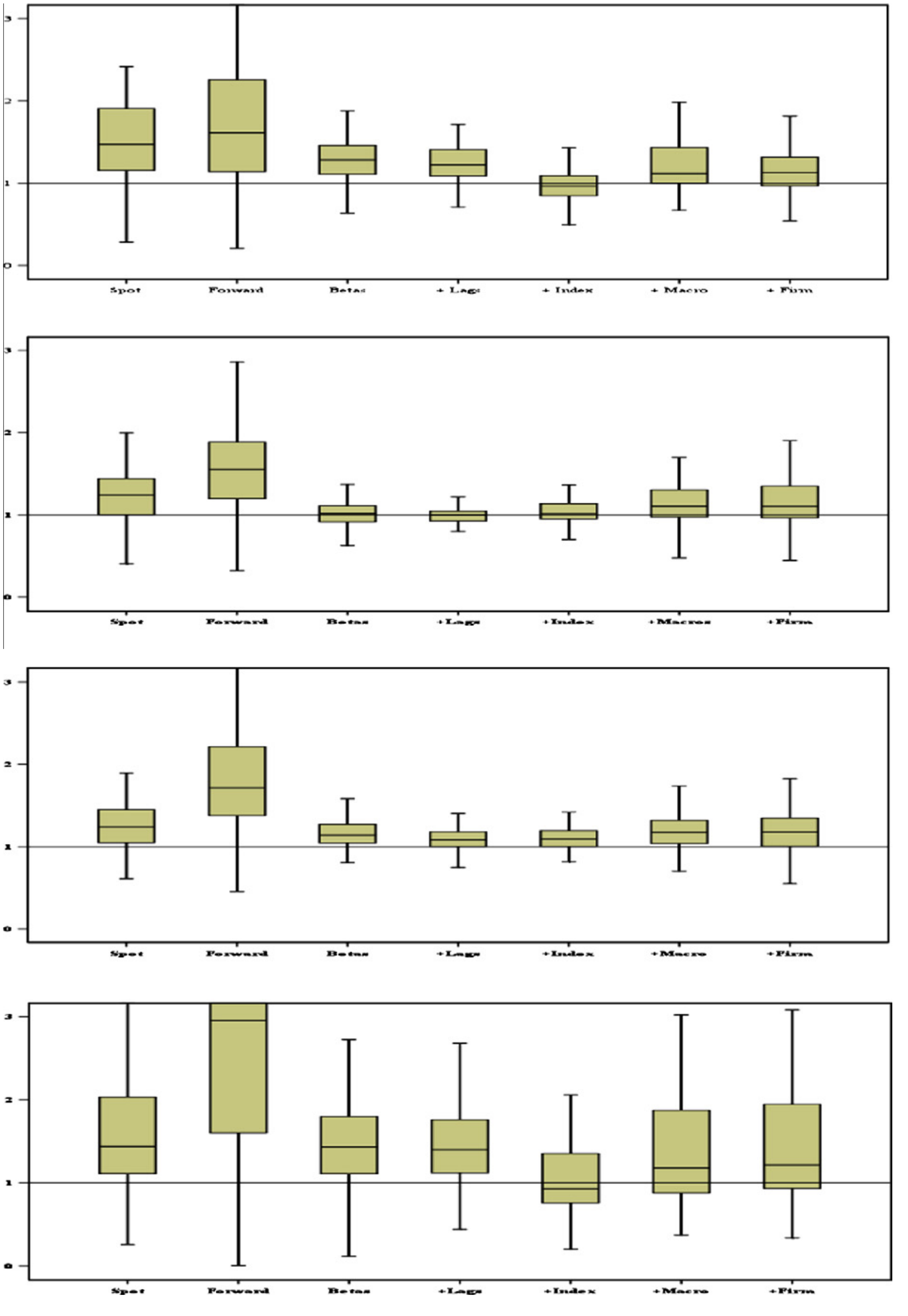

Fig. 7. Out-of-sample prediction errors: comparison of models for different credit spread predictions. The panels from top to bottom show box plots of MSPE ratios for 6-months-ahead 3-year-credit-spread forecasting models, 6-months-ahead 10-year-credit-spread forecasting models, 3-months-ahead 5-year credit spread forecasting models, and 12-months-ahead 5-year credit spread forecasting models. The MSPE ratios plots in sequence, from left to right, in each panel are those of Spot, Forward, and of Models M_1, M_2, M_4, M_5, and M_6, all relative to the MSPE of a model with just the credit spread factors, lagged factors, and riskless factors (Model M_3), firm by firm.

predictions than do the random walk models. Second, macro and firm-specific information, over and above information on credit spread and riskless factors do not improve predictions.

6.2. Why macro and firm-specific information do not improve credit-spread predictions?

Our consistent and strong finding is that macro and firm-specific information, over and above information on credit spread and riskless factors, do not improve credit-spread predictions. The explanation for this result could be that the riskless and B-rated-index factors impound contemporaneous macroeconomic information. The credit-spread factors could impound contemporaneous macroeconomic and firm-specific-risk information. We check this next.

Panel A of Table 10 examines the contemporaneous relationship between riskless level, slope and curvature factors and our macro variables. Specifically:

$$\beta_f(t) = \eta_{0f} + \eta_{1f}M(t) + \epsilon_f(t). \tag{17}$$

The table also examines the relationship between B-index credit spreads with riskless factors and macroeconomic variables:

$$\beta_I(t) = \eta_{0I} + \eta_{1I}\beta_f(t) + \eta_{2I}M(t) + \epsilon_I(t). \tag{18}$$

This table reports the adjusted R^2 values for these regressions. Consistent with Harvey (1993), for example, the riskless factors are significantly associated with contemporaneous macroeconomic information. The B-rated index credit-spread factors are significantly associated with macroeconomic

Table 10
Contemporaneous determinants of riskless factors, index factors, and credit spreads.

Dependent variables	Riskless factors	+ Macro variables
Panel A		
Riskless level		0.213***
Riskless slope		0.239***
Riskless curvature		0.276***
Index level	0.155***	0.777***
Index slope	0.692***	0.786***
Index curvature	0.428***	0.629***

Credit rating	Industry	3-Year credit spread	5-Year credit spread	10-Year credit spread
		Explanatory variables		
		Current macro and firm variables	Current macro and firm variables	Current macro and firm variables
Panel B				
High	Manufacturing	0.417***	0.595***	0.583***
	Service	0.401***	0.570***	0.498***
Low	Manufacturing	0.582***	0.702***	0.686***
	Service	0.397***	0.430***	0.439***

***, **, and * denote significantly different from zero at the 1%, 5%, and 10% levels, respectively.
Notes: The top half of panel A reports the adjusted R^2 when each of the riskless factors is regressed on the 4-vector of current month macro-variables. The 4-vector of macro-variables comprises the 1st Principal Component of the Real Activity Variables after each of these variables are purged of the riskless factors, the 1st Principal Component of the Inflation Variables after each of these variables are purged of the factors parameters, stock market momentum, and stock market volatility. The Real Activity variables are (a) the index of Help Wanted Advertising in Newspapers (HELP), (b) the unemployment rate (UE), (c) the growth rate of employment (EMPLOY), and (d) the growth rate of Industrial Production (GIP). The Inflation Variables are the growth rates of (a) the Consumer Price Index (CPI), (b) the Producer price Index (PPI) and (c) of the Commodity Price Index (COMM). The bottom half of panel A reports the adjusted R^2 when each of the B-rated-index factors are regressed on the 3-vector of current month riskless factors, and on the 4-vector of current month macro variables in a hierarchical specification. The B-rated-index credit-spread factors are constructed from the B-credit-rated-index term structure curves obtained from Bloomberg. Panel B shows the reports the adjusted R^2 when the 3-year, 5-year and 10-year credit spreads are regressed on the 8-vector of current-month macro variables and firm variables, for four groups of firms double-sorted based on their credit ratings and industry. Panel regressions with firm fixed effects are run over each group of firms. The significance of each block of explanatory variables is based on the p-value of the partial F-statistic.

information, even after controlling for the impact of riskless factors. In other words, the riskless factors impound macroeconomic information, and the B-rated index factors impound both the riskless and the macroeconomic information.

Panel B shows panel regression results, controlling for firm fixed effects, of the contemporaneous determinants of credit spreads for firms double-sorted by credit ratings and industry. Specifically, for firm j in one of the four credit ratings-industry groups (high or low rating and service or manufacturing industry), k:

$$s_{jk}^{(n)}(t) = \eta_{0,jk}^{(n)} + \eta_{1,k}^{(n)} M(t) + \eta_{2,k}^{(n)} F(t) + \epsilon_{jk}^{(n)}(t). \tag{19}$$

Controlling for firm fixed effects, the blocks of firm specific and macro variables accounted for around 40% of the variability of 3-year credit spreads, 55% of the variability of 5-year credit spreads and 50% of the variability of 10-year credit spreads, on average, across all groups. The explanatory power of the macro and firm variables, on average, is higher for manufacturing firms than for service firms, and higher for the lower-rated manufacturing firms than for higher-rated manufacturing firms. The explanatory power of macro and firm variables over credit spreads that we find is in line with the results of Avramov et al. (2007), who explain 54% (67%) of the variability of credit spreads for medium (low) grade firms using macro and firm-specific variables, and Krishnan et al. (2005), who explain about 40% of the variability of credit spreads for banks using macro and firm-specific variables.

6.3. *Using macroeconomic forecasts*

It is perhaps surprising that macro and firm variables are so efficiently embedded into the current term structures of credit-spread and riskless yields, that, at the margin, they do not improve out-of-sample forecasts of credit spreads. However, we need not be constrained to using only past information. The market knows the forecasts of future macro variables at the time of the prediction. Perhaps it might help, in terms of the accuracy of predictions of future credit spreads, if we substituted macroeconomic variables with forecasts of future macroeconomic variables.

We collect quarterly information from the Survey of Professional Forecasters (SPFs). The survey's participants forecast several macroeconomic variables, and report their forecasts at the middle of each quarter. Typically, about 40 forecasters participate. The survey data are obtained from the web site of the Federal Reserve Bank of Philadelphia. We use the median 6-months-ahead forecasts of key macro variables and match them with the month in which they are made.

The macro variables we use are the median 6-months-ahead GDP forecast, the median 6-months-ahead CPI Inflation forecast, the median 6-months-ahead Industrial Production forecast, and the median 6-months-ahead forecast of Moody's AAA corporate bond yield. We replace our current 4-vector of macro-variables with this four vector of macro-forecasts, and re-estimate the 6-month-ahead forecasts of out-of-sample credit spreads using our firm-by-firm regression models. This estimation uses data from a quarterly database rather than a monthly database because the macro-forecasts are available only once a quarter (in the middle month of the quarter).

We evaluate whether using macro forecasts rather than current macro variables improves forecasts of future credit spreads. The models we compare are the: (a) the spot model, (b) the model with credit-spread factors and riskless factors (M_3), (c) the model with credit-spread factors, riskless factors and index factors (M_4), (d) the model with credit-spread factors, lagged factors, and macro-variable forecasts, (e) the model with credit-spread factors, lagged factors, riskless factors, and macro-variable forecasts, and (f) the model with credit-spread factors, lagged factors, riskless factors, index factors, and macro-variable forecasts.

Table 11 reports average absolute errors by firms and the proportion of firms for which each out-of-sample prediction model has a lower MSPE relative of relative to MSPE of Model M_3, firm by firm.

The average absolute prediction errors are significantly lower for models M_3 and M_4 than for those of any of the other model that uses macro forecasts. The proportion of firms that have lower MSPEs is also significantly lower than 0.5 for the other models (except M_4).

Table 11
Out-of-sample 6-months-ahead 5-year-credit-spread prediction using macro forecasts.

	6-Months-ahead 5-year credit spread	
	Average absolute prediction error across firms	Proportion of firms for which model has lower average MSPE than Model M_3
Spot	36.23	0.26
M_3	31.94	N/A
M_4	32.08	0.50
M_2 + macro-forecasts	40.75	0.23
M_3 + macro-forecasts	37.46	0.20
M_4 + macro-forecasts	45.97	0.14
M_5 + macro-forecasts	53.19	0.10

Notes: This table reports the average absolute prediction errors by firms, and the proportion of firms for which each out-of-sample prediction model has a lower MSPE relative of relative to MSPE of a model with just the credit spread factors, lagged factors, and the riskless factors (Model M_3), firm by firm. The models compared are the: (a) Spot Model, (b) the model with the credit spread factors, riskless factors and Index factors, M_4, (c) the model with credit spread factors, lagged factors, and predicted macro-variables, (d) the model with credit spread factors, lagged factors, riskless factors, and predicted macro-variables, and (e) the model with credit spread factors, lagged factors, riskless factors, index factors, and predicted macro-variables. The 4-vector of macro variable forecasts used are the median 6-month ahead GDP forecast, the median 6-month ahead CPI Inflation forecast, the median 6-month ahead Industrial Production forecast, and the median 6-month ahead Moody's AAA corporate bond yield forecast from the Survey of Professional Forecasters (SPF). We obtain these quarterly forecasts from the web site of the Federal Reserve Bank of Philadelphia. Firm-by-firm regressions are run.

Fig. 8 shows the box plots of the ratios of MSPEs for our different credit-spread prediction models relative to the MSPE of the model that uses only current period credit-spread and riskless-yield factors (M_3), firm by firm and quarter by quarter.

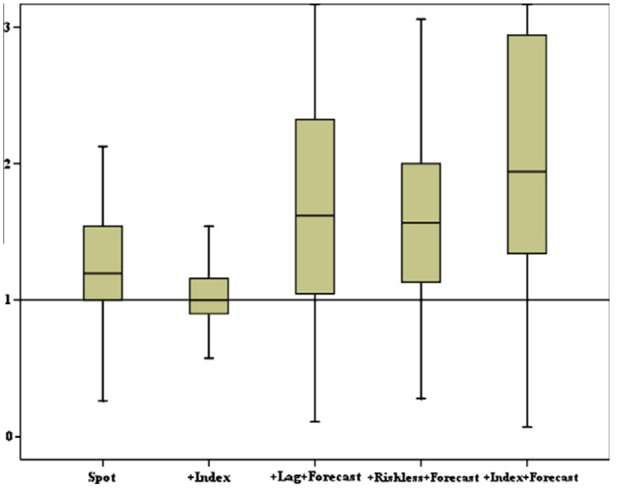

Fig. 8. Out-of-sample prediction errors of 5-year credit spreads: comparison of models using macro-forecasts. This figure shows the box plots of the ratios of MSPE for different credit spread prediction models relative to the MSPE of a model with just the credit-spread factors, lagged credit-spread factors, and riskless factors (Model M_3), firm by firm. The models compared are the: (a) Spot Model, (b) the model with the credit-spread factors, riskless factors, and B-rated-index factors, Model M_4, (c) the model with credit-spread factors, lagged credit-spread factors, and predicted macro variables, (d) the model with credit-spread factors, lagged credit-spread factors, riskless factors, and predicted macro variables, and (e) the model with credit-spread factors, lagged credit-spread factors, riskless factors, B-rated-index factors, and predicted macro variables. The 4-vector of predicted macro variables used are the median 6-months-ahead GDP forecast, the median 6-month ahead CPI Inflation forecast, the median 6-months-ahead Industrial Production forecast, and the median 6-months-ahead Moody's AAA corporate bond yield forecast from the Survey of Professional Forecasters (SPF) from the web site of the Federal Reserve Bank of Philadelphia.

The five box plots show that models M_3 or M_4 are better than any of other four models that use macro forecasts.

7. Conclusion

This study examines the predictability of credit spreads at the firm level. We construct monthly credit-spread curves for a large representative sample of 241 high- and low-credit-rated firms from both manufacturing and service sectors, over a period of 16 years from 1990 to 2005. Using a 3-factor Diebold–Li model, we fit a term structure of credit spreads for each firm every month. The fit of the bond prices derived from our optimization routine to the data are extremely good: over 90% of the bonds fit to within 1 dollar of their price. We document that credit-spread curves can be upward or downward sloping, and can take hump-shaped patterns. Credit spreads of different maturities for the same firm can move in different directions over the same time period. For example, in our data set, the 3-year and 10-year credit spreads move in the same direction only 63% of the time.

Once the firm-by-firm credit-spread curves are constructed, we investigate which blocks of variables are informative for predicting future credit spreads. We contribute to the literature with two new results. (a) We find that the benchmark forecasts based on the current spot and forward credit spreads can be substantially improved upon using a model that incorporates the level, slope, and curvature factors of the credit-spread curve. That is, today's credit-spread curve contains significant information on future credit spreads. The slope and the curvature of today's credit-spread curve contribute significantly to predicting future credit spreads over and above the level (the "parallel-shift" factor). (b) But the information contained in today's and past credit-spread curves is not sufficient for forecasting future credit spreads. Forecasts can be further significantly improved upon by incorporating information contained in the riskless yield curve. In the presence of these two blocks of factors, aggregate stock-market and other macroeconomic variables as well as firm-specific-risk variables do not contain significant incremental information on future credit spreads.

Our results suggest that models of risky debt should incorporate information, not only on the full term structure of credit spreads, but also on the full term structure of riskless interest rates. Indeed, our results rule out the popular affine term structure models for credit spreads. In these models, credit spreads are linear functions of state variables. Original sets of state variables that include macroeconomic variables and latent factors can be mapped onto credit spreads of different maturities that serve as equivalent state variables. Therefore, these affine models will have the property that the state variables are essentially credit-spread state variables. The implication is that the credit-spread curve essentially impounds all information necessary for forecasting credit spreads. Ludvigson and Ng (2009) discuss this issue in the context of riskless rates and go on to discuss possible ways in which theory can reconcile findings such as ours. These include developing new models that incorporate nonlinear pricing kernels, or have features of unspanned stochastic volatility, or by allowing different state variables to govern the dynamic behavior of yields and risk premia. In the context of term structures for credit spreads, our empirical results indicate the need to broaden the scope of extant credit spread models.

Our results have implications for credit-risk hedging and Value-at-Risk assessment. Specifically, the information contents of the full credit-spread curve and riskless yield curve provide information on future credit-spread levels, and the full information should be taken into account when determining economic capital. Further, our results have implications for the interpretation of empirical results in corporate finance that link corporate capital structure policies such as bond issuance, maturity selection and market timing to the term structure of interest rates, and corporate risk management policies including hedging interest rate exposure also to the term structure of interest rates. Guedes and Opler (1996), Barclay and Smith (1995) and Stohs and Mauer (1996), for example, find that maturity of debt issued is linked to riskless term structure. Baker et al. (2003) go further and conclude that managers are successful in timing new issues of debt, which confirm the survey findings of Graham and Harvey (2001).[8] The decision to select floating rate versus fixed rate debt can be influenced by

[8] For further discussions on market timing decisions of debt issuances and the link to interest rates, see Barry et al. (2005) and the references cited therein.

the correlation between operating cash flows before interest expense and interest rates. Several studies also indicate that firms issue more floating rate debt when the riskless term structure is steep. This is taken as evidence in support of market timing, rather than hedging. Faulkender (2005), for example, finds that the degree of interest rate exposure is largely driven by the slope of the riskless yield curve at the time debt is issued, and that interest rate risk management practices are primarily driven by speculation rather than by hedging considerations. Chernenko and Faulkender (2007) argue that agency issues influence the documented sensitivity of interest rate swap usage to the term structure (also see Faulkender et al. (2007)). Thus, these studies link corporate capital structure and risk management policies to the term structure of interest rates. But if the term spread is informative about future credit spreads, as our results suggest, then interpretations of such empirical results might change in significant ways. That is, corporate capital structure and risk management decisions might also be influenced by the information contained in the current term structure of interest rates on future firm-level credit spreads. Future research must reconsider interpretation of such empirical results in this light.

Appendix A. Descriptions of blocks of macro and firm variables

Macro-variables	Description
Real activity index	Real activity index, $RA(t)$, in month t is the first principal component of four underlying time series of macro-variables after purging each of them of the riskless level, slope and curvature factors. The four underlying monthly series of macro-variables are the Index of Help Wanted Advertising in Newspapers (HELP), the Unemployment Rate (UE), the growth rate of Employment (EMPLOY) and the growth rate of Industrial Production (GIP). All growth rates are measured as the 12-month difference in the logs of the index
Inflation index	Inflation index, $I(t)$, for month t is the first principal component of three underlying time series of macro-variables after purging each of them of the riskless level, slope, and curvature factors. The three underlying monthly series of macro-variables are the Consumer price Index (CPI), the Producer price Index of Finished Goods (PPI), and the Market Commodity Price Index (PCOM). All these inflation measures are measured as changes in the logs of their indices over a 12 month period
Stock market momentum	Stock market momentum, $R_{\mathrm{M}}(t)$, for month t is the 12-month cumulative holding period return from month t-13 through month t-2 of the Center for Research in Security Prices (CRSP) value-weighted index return following the methodology described in Kenneth French's web-site
Stock market volatility	Stock market volatility, $\sigma_{\mathrm{M}}(t)$, is the monthly volatility of the CRSP value-weighted index return using the daily returns of the index within each month, following the methodology of French et al. (1987)

Firm-specific variables	Description
Leverage	Leverage, $L(t)$, for month t is the ratio of debt outstanding on the balance sheet of the firm (Compustat Quarterly data item 51) and the market value of its common stock, computed monthly as the product of the number of shares outstanding and the closing share price each month (Compustat Quarterly data items 61 and 14). The book value of debt is the same number for all three months of a quarter
Book-to-market ratio	The book-to-market ratio, $BM(t)$ for month t is the ratio of the book value of equity to the market value of equity. The book value of equity is defined as stockholders' equity plus balance sheet deferred taxes and investment tax credit, minus book value of preferred stock, which are, respectively, Compustat Quarterly data items 60, 52, and 55. The market value of equity is defined as the number of shares outstanding multiplied by the end of respective month closing stock price, which are, respectively, Compustat Quarterly data items 61 and 14

Appendix A (*continued*)

Stock momentum	Stock momentum, $R(t)$, for month t is the 12-month cumulative holding period stock return from month t-13 through month t-2 taken from CRSP
Stock volatility	Stock volatility, $\sigma(t)$, is the monthly volatility using the daily stock returns in each month from CRSP

Appendix B. Data screens used in sample selection

Pre credit-spread generation stage

(1) We use only trade-price data from Bloomberg, augmented in a minor way with trade-price data from DataStream.
(2) We collect bond prices of firms that are or were part of the S&P500 Index for any of the years in our sample period from January 1990 to December 2005.
(3) So that the dataset does not suffer from survival bias, as we have not excluded bond price data on firms that subsequently defaulted.
(4) We use only fixed-rate US dollar-denominated bonds.
(5) We use only senior debt.
(6) We use only bonds that are non-callable, non-putable, non-convertible, not part of an unit (for example, sold with warrants).
(7) We use only bonds that do not have sinking fund feature. We also excluded bonds with asset-backed and credit-enhancement features.
(8) We eliminate all bonds with principal repayment dates that are inconsistent with the maturity date that can be calculated using data on future coupon payment dates.
(9) We exclude all bonds with prices that are above otherwise equivalent Treasury securities.
(10) We require a minimum of five prices of bonds of different maturities that span at least 7 years for each month to estimate the credit spread level, slope and curvature factors.

The above data screens result in 14,234 firm-months of data for 387 firms.

Post credit-spread generation stage

(11) We drop all firm-months that do not have at least six consecutive months of reasonable credit spread level, slope, and curvature factor estimates.
(12) We remove all credit spreads for which the credit spread is negative.
(13) We remove all credit spreads that exceed 5 standard deviations away from the mean for a firm, following Avramov et al. (2007) because extreme spreads can be attributed either to data errors or to credit blow-ups.
(14) We exclude firm-months that do not have past 12 months of equity returns data (to calculate the momentum risk factor) in the Center for Research in Security Prices (CRSP) database.
(15) We exclude firm-months for which the firm data from the quarterly Compustat to obtain our firm-specific risk measures are not available.

The above screens results in 11,894 firm-months of data for 241 firms.

References

Acharya, V., Carpenter, J., 2002. Corporate bond valuation and hedging with stochastic interest rates and endogenous bankruptcy. Rev. Finan. Stud. 15, 1355–1383.

Agrawal, D., Bohn, J., 2008. Humpbacks in credit spreads. J. Investment Management 6, 73–98.

Amato, J., Luisi, M., 2006. Macro factors in the term structure of credit spreads, BIS working paper 203.
Ang, A., Piazzesi, M., 2003. A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables. J. Mon. Econ. 50, 745–787.
Avramov, D., Gergana, J., Philipov, A., 2007. Understanding corporate credit risk changes. Fin. Analysts J. 63, 90–105.
Baker, M., Greenwood, R., Wurgler, J., 2003. The maturity of debt issues and predictable variation in bond returns. J. Fin. Econ. 70, 261–291.
Barclay, M., Smith, C., 1995. The maturity structure of corporate debt. J. Finance 50, 609–631.
Barry, C., Mann, C., Mihov, V., Rodriguez, M., 2005. Interest rates and timing of corporate debt issues. working paper, Texas Christian University.
Berndt, A., Douglas,R., Duffie,D., Ferguson,M., Schranz, D., 2005. Measuring default risk premia from default swap rates and EDFs, Working paper, Cornell University.
Black, F., Cox, J., 1976. Valuing corporate securities: some effects on bond indenture provisions. J. Finance 31, 351–367.
Blanco, R., Brennan, S., Marsh, I., 2005. An empirical analysis of the dynamic relationship between investment grade bonds and credit default swaps. J. Finance 60, 2255–2281.
Bliss, R., 1997. Testing term structure estimation methods. Adv. Futures Options Res. 9, 197–232.
Campbell, J., Shiller, R., 1991. Yield spreads and interest rates: a bird's eye view. Rev. Econ. Stud. 58, 494–515.
Campbell, J., Lettau, M., Malkiel, B., Xu, Y., 2001. Have individual stocks become more volatile? An empirical exploration of idiosyncratic risk. J. Finance 56, 1–43.
Chava, S., Jarrow, R., 2004. Bankruptcy prediction with industry effects. Rev. Fin. 8, 537–569.
Chernenko, S. Faulkender, M., 2007. Interest rate swap speculation and managerial compensation?, Working paper, Washington University at St. Louis.
Cochrane, J., Piazzesi, M., 2005. Bond risk premia. Am. Econ. Rev. 95, 138–160.
Collin-Dufresne, P., Goldstein, R., 2001. Do credit spreads reflect stationary leverage ratios? J. Finance 56, 1929–1957.
Collin-Dufresne, P., Goldstein, R., Martin, S., 2001. The determinants of credit spread changes. J. Finance 56, 2177–2207.
Cremers, M., Driessen, J., Maenhout, P., Weinbaum, D., 2008. Explaining the level of credit spreads: option-implied jump risk premia in a firm value model. Rev. Finan. Stud. 21, 2209–2242.
Dai, Q., Singleton, K., 2000. Specification analysis of affine term structure models. J. Finance 55, 1943–1978.
Diebold, F., Li, C., 2006. Forecasting the term structure of government bond yields. J. Econometrics 130, 337–364.
Diebold, F., Rudebusch, G., Aruoba, B., 2006. The macroeconomy and the yield curve: a dynamic latent factor approach. J. Econometrics 131, 309–338.
Duffee, G., 2002. Term premia and interest rate forecasts in affine models. J. Finance 57, 405–443.
Eom, Y., Helwege, J., Huang, J., 2004. Structural models of corporate bond pricing: an empirical analysis. Rev. Finan. Stud. 17, 499–544.
Ericsson, J., Reneby, J., Wang, H., 2005. Can structural models price default risk? Evidence from bond and credit derivative markets, Working paper, McGill University.
Fama, E., Bliss, R., 1987. The information in long-maturity forward rates. Am. Econ. Rev. 77, 680–692.
Fama, E., French, K., 1992. The cross-section of expected stock returns. J. Finance 47, 427–465.
Faulkender, M., 2005. Hedging or market timing? Selecting the interest rate exposure of corporate debt. J. Finance 60, 931–962.
Faulkender, M., Jenkins, N., Seethamraju, C. 2007. The market reaction to the strategic use of interest rate swaps, Working paper, Washington University at St. Louis.
French, K., Schwert, G., Stambaugh, R., 1987. Expected stock returns and volatility. J. Finan. Econ. 19, 3–29.
Graham, J., Harvey, C., 2001. The theory and practice of corporate finance: Evidence from the field. J. Finan. Econ. 60, 187–243.
Guedes, J., Opler, T., 1996. The determinants of the maturity of corporate debt issues. J. Finance 51, 1809–1833.
Harvey, C., 1993. The term structure forecasts economic growth. Fin. Anal. J. 49, 6–8.
He, J., Hu, W., Lang, L., 2004. Credit spread curves and credit ratings, Working paper, Chinese University of Hong Kong.
Helwege, J., Turner, C., 1999. The slope of the credit yield curve for speculative-grade issuers. J. Finance 54, 1864–1889.
Hotchkiss, E., Warga, A., Jostova, G. 2002. Determinants of corporate bond trading: a comprehensive analysis, Working paper, George Washington University.
Huang, J., Huang, M. 2003. How much of the corporate-treasury yield spread is due to credit risk? A new calibration approach, Working paper, Stanford University.
Inoue, A., Kilian, L., 2004. In-sample or out-of-sample tests of predictability: which one should we use? Econ. Rev. 23, 4 371–402.
Jarrow, R., 2004. Risky coupon bonds as a portfolio of zero-coupon bonds. Finan. Res. Lett. 1, 101–105.
Kim, I., Ramaswamy, K., Sundaresan, S., 1993. Does default risk in coupons affect the valuation of corporate bonds? Fin. Man. 22, 117–131.
Krishnan, C.N.V., Ritchken, P., Thomson, J., 2005. Monitoring and controlling bank risk: does risky debt help? J. Finance 60, 343–377.
Krishnan, C.N.V., Ritchken, P., Thomson, J., 2006. On credit spread slopes and predicting bank risk. J. Money Credit Banking 38, 1545–1598.
Leland, H., Toft, K., 1996. Optimal capital structure, endogenous bankruptcy, and the term structure of credit spreads. J. Finance 51, 987–1019.
Longstaff, F., Schwartz, E., 1995. A simple approach to valuing risky fixed and floating rate debt. J. Finance 50, 789–821.
Longstaff, F., Mithal, S., Neis, E., 2005. Corporate yield spreads: default risk or liquidity? New evidence from the credit default swap market. J. Finance 60, 2213–2253.
Ludvigson, S., Ng, S., 2009. Macro factors in bond risk premia. Rev. Finan. Stud., forthcoming.
Monch, E. 2005. Forecasting the yield curve in a data-rich environment: a no arbitrage factor-augmented VAR approach, Working paper 544, European Central Bank.
Nelson, C., Siegel, A., 1987. Parsimonious modeling of the yield curve. J. Bus. 60, 473–489.
Pan, J., Singleton, K., 2008. Default and recovery implicit in the term structure of sovereign CDS spreads. J. Finance 63, 2345–2384.

Pastor, L., Veronesi, P., 2003. Stock valuation and learning about profitability. J. Finance 58, 1749–1790.
Petersen, M., 2009. Estimating standard errors in finance panel data sets: comparing approaches. Rev. Finan. Stud. 22, 435–480.
Schwarz, G., 1978. Estimating the dimension of a model. Ann. Statist. 6, 461–464.
Stohs, M., Mauer, D., 1996. The determinants of corporate debt maturity structure. J. Bus. 69, 279–312.
Tang, D., Yan, H., 2006. Macroeconomic conditions, firm characteristics, and credit spreads. J. Finan. Services Res. 29, 177–210.
Tang, D., Yan, H., 2007. Liquidity and credit default swap spreads, Working paper, University of Texas, Austin.