A non-parametric investigation of risk premia

Peroni, Chiara

school of economics, university of east anglia, rimini center for economic analysis (RCEA)

1 June 2007

A Non-parametric Investigation of Risk Premia

Chiara Peroni*

Abstract

This paper investigates features of credit risk using non-parametric techniques, studying determinants of risk premia using a non-parametric term-structure model of the corporate spread. The model, which measures the extra return of defaultable corporate bonds on their government counterparts, involves the rate of inflation, a key macroeconomic variable that is found to explain the spread non-linearly. This approach demonstrates the usefulness of non-linear approaches in contrast with standard linear approaches. The model is also useful to forecast the future course of the spread.

Credit risk, usually defined in terms of *default risk*, measures the possibility of borrowers not being able to pay neither contractual interest nor the principal on their debt obligations. When buying securities, investors try to assess the quality of the borrower in order to reduce the probability of incurring in financial losses, and the higher this risk of credit the higher the required promised payment to compensate for its bearing. Usually, returns on risky securities are higher than returns on securities regarded as "safe" (e.g., yields on corporate bonds are higher than yields on government bonds).

This extra return is known as *risk premium*, a key financial variable which conveys information on the market perception of credit risk and of economic conditions. Its determinants, however, are not very well understood.

To study risk premia, empirical research has often focused on *corporate spreads*. Because these are differences between yields on corporate debt subject to default risk and comparable government bonds free of such risk, they are easily interpreted as direct measures of the *risk premium*. Historical default probabilities, however, are too low to account for the size of observed spreads. Studies report a large non-default component in corporate spreads, which is often left unexplained. Thus, capturing large observed spreads and explaining the link between spreads and default risk are key empirical challenges in this area of finance and constitutes the main aim of this paper.

This paper also aims to check the appropriateness of a widely popular class of model in the analysis of risky assets, Reduced Form Models. These constitutes pricing frameworks, which model the term-structure of activities subject to default risk as a direct extension

*Correspondence address: School of Economics, University of East Anglia, NR4 7TJ, Norwich, UK. E-mail address: c.peroni@uea.ac.uk.

of risk-free yield curves, and are state of the art in this field. Reduced Form Models derive analytical and simple expressions for risky yields, which are linear in a set of state variables (factors), but impose a strong parametric assumption of linearity on the state variables' time series processes and cross section of yield-prices alike.

This article proposes instead a non-parametric multi-factor term-structure model of the spread to better explain the formation of large risk premia. In this model, corporate yields are functions of two sets of factors related to the risk of credit, one based on key macroeconomic variables, the other related to the risk-free term-structure. Functions describing how factors determine spreads are non-linear. It is by using non-parametric techniques, which allow flexible estimation of the model by relying only on the data available, that this study tries to overcome limitations of standard linear approaches. The analysis focuses on the cross-section relation of yields to factors.

This paper is structured as follows. Section 1 examines the theoretical determinants of risk premia in a Reduced Form framework. After discussing the data (Section 2), section 4 analyses results from the estimation and testing of a non-parametric model of the spread. Section 5 compares forecasting performances of several models, after which section 6 summarises results and gives concluding remarks.

1 Background

Reduced-Form Models (RFMs), introduced by Duffie and Singleton (1997), are frameworks for the measurement and pricing of credit risk. They are widely popular with academics and practitioners because they offer easily interpretable intuition, but they rely on strong parametric assumptions. RFMs belong to the *affine* class of *term-structure models*, which capture movements of interest rates (yields) and establish determinants of their evolution over time (Dai and Singleton, 2000). In these models, the dynamics of interest rates depends on the evolution of a set of observed, or unobserved, variables (or factors). These variables can be identified with nodes of the term structure itself, or may be macroeconomic variables. Their underlying dynamics is described by affine (linear) *Itô* diffusion processes. The main advantage of affine models is their tractability: the linearity assumption in the factors' dynamics yields an analytical representation of the term structure and bond price formulas which are easy to interpret and well suited to empirical testing.

RFMs describe the term-structure of yields on defaultable activities as a direct extension of government yield curves. This is done by replacing the risk-free instantaneous rate of interest r, used in conventional term-structure modelling, with a default-adjusted (or risk-adjusted) discount rate R. The price in t of a defaultable zero-coupon zero-recovery bond of maturity T can be written as if the promised payoff were default-free:

$$P_t = E_t^Q[\exp\{-\int_t^T R_s ds\}], \tag{1}$$

where Q indicates assessment under risk-neutral probabilities (Duffie and Singleton, 1997). The risk-adjusted rate R is given by the sum of the risk-free rate r plus a term, λ, which captures default risk and depends on the probability of default (hereafter referred to as the *default rate*):

$$R_s = r_s + \lambda_s; \tag{2}$$

Key assumption here is that default is a "surprise" event, which occurs unexpectedly, and is exogenous to the model. The dynamics of the risk-adjusted rate is modelled as an affine multi-factors diffusion process, examples of which can be found in Duffee (1999), and Duffie and Singleton (1999). This gives representation of yields as linear functions of factors. As a result, credit spreads, which are differences between yields on risky assets and yields on risk-free assets, are also linear functions of factors. This is shown as follows, under the simplifying assumption of risk-neutral independence.

Substituting equation 2 in 1 the price of the zero-coupon risky bond can be decomposed into the product of a risk-free price component (Δ) and a price component that depends on the default rate (D):

$$P_t^{(\tau)} = D^{(\tau)}(\lambda)\Delta^{(\tau)}(r), \tag{3}$$

where τ is the time to maturity ($\tau \equiv T - t$).

These zero-coupon bond prices are as follows:

$$D^{(\tau)}(\lambda) = A_D(\tau)\exp\{-B_D(\tau)\lambda\}, \tag{4}$$

$$\Delta^{(\tau)}(r) = A_\Delta(\tau)\exp\{-B_\Delta(\tau)r\}; \tag{5}$$

(The coefficients A_Δ, B_Δ, A_D, B_D are deterministic functions of underlying diffusion parameters and time to maturity.) The credit spread implied by RFMs is computed as the difference between the yield on the risky bond and the yield on the risk-free bond, using equations 4 and 5. The resulting (observable) spread is linear in the default rate:

$$s(\tau) = -\frac{\log A_D(\tau)}{\tau} + \frac{B_D(\tau)\lambda}{\tau}. \tag{6}$$

Following Duffee (1999), I now allow the default rate λ to depend (linearly) on the risk-free rate r. This yields a simple and tractable representation of the spread, which is now *linear* in the interest rate AND default rate:

$$s(\tau) = -\frac{\log A(\tau)}{\tau} + \frac{B_{D,\lambda}(\tau)}{\tau}\lambda + \frac{(B_{D,r}(\tau) - B_\Delta(\tau))r}{\tau}. \tag{7}$$

(Notice that the equation above can be inverted to write the default premium as a linear function of observable variables.)

One of the best features of RFMs is this parallel between default-free and defaultable bond price formulas, which gives simple linear expression for yields and spreads. This linearity assumption, however, has been questioned for government yield curves. Substantial evidence, based primarily on the estimation of underlying diffusion processes, suggests that it is too restrictive (Aït-Sahalia, 1996; Ahn and Gao, 1999; Arapis and Gao, 2006). The possible presence of non-linearities in data has important implications, suggesting that affine models are misspecified and bond prices formulas are not correct.

There is another well-known shortcoming to these models. While affine models represent yields as linear functions of factors, they fail to provide a clear economic interpretation of those factors. To address this problem, models of risk-free yields have recently been extended to include macroeconomic variables as factors (Ang and Piazzesi, 2003), but this idea has not been applied in the area of credit risk. Some empirical evidence suggests, however, that risk premia are affected by general economic conditions. The observed negative correlations between Treasury rates and credit spreads (documented in Duffee, 1998) is often interpreted as evidence that risk premia are correlated with the business cycle. (During economic downturns — associated to low interest rate levels — the risk of bankruptcy increases, driving up yields on risky securities.) Despite this evidence, and to the best of my knowledge, the only (empirical) model of spreads which explicitly considers the effect of macroeconomic indicators is Wadhwani (1986). Wadhwani (1986) explains why the inflation rate does affect default risk and spreads, arguing that inflation has an adverse effect on firms' interest payments, creating cash-flow problems and as a consequence increasing the number of bankruptcies.

This paper addresses these issues by analysing corporate yield-spread indices using non-parametric techniques. In particular, it uses *non-parametric kernel regression* methods, which permit to specify and estimate non-linear models. This approach has the advantage of being data-driven: regression functions are determined by the data, instead of having a pre-specified functional form (as in parametric regression). This delivers reliable estimators, and makes non-parametric regression a flexible tool to discover non-linearities in the data. These techniques are also useful for data description, as well as for building and checking parametric models.

The model of equation (7) suggests a simple linearity test of RFMs based on the cross-section relation of yield-spreads to factors. Here, this model is tested against a non-linear non-parametric alternative, with the aim of assessing the validity of affine specifications. The estimation of model (7) is complicated by the inclusion the crucial but unobservable variable λ (the default rate). In theoretical models, λ depends on the probabilities of default, but historical default rates are too low to account for observed spreads, even for very short-term securities (this fact is often referred to as the *credit spread puzzle*). It has been argued that this puzzle emerges because credit spreads are determined by more factors, in addition to the risk of default. Empirical studies have found evidence of non-default components in risk premia, often linked to liquidity and taxation, and reported a range of estimates on its size (Elton et al., 2001). Recently, however, Longstaff et al. (2004) estimates that default risk accounts for a large part of the corporate spread (this uses credit-default swaps). This seems to suggest that the problem lies in the *measurement* of default risk, in particular in the measurement of the market's *perception* of risk, rather than in its relative explanatory ability. This paper attempts to measure the link between default risk and spreads by choosing appropriate factors which are observable. The idea is to use a macroeconomic variable as the observable. Thus, the spread model estimated here includes a key macro variable related to default risk: *the inflation rate*. This extends RFMs to include macroeconomic variables, allowing us to study the relationship between macroeconomic conditions and the risk premium.

2 Data

In this analysis I use a data set constructed from government yield curve data compiled by the Federal Reserve and corporate yields indices from Moody's database, which were obtained from the FRED database.[1] The data set is made up of monthly time series ranging from April 1953 to June 2006. Spreads are computed as differences between yields on investment-grade corporate bond (rated by Moody's as Aaa and Baa) and constant-maturity-Treasury yields of comparable maturity. Table 1 reports summary statistics for the spread series.[2]

Figure 1 presents time series of Aaa and Baa spreads (right panel) and of the *relative spread* (right panel), which measures the difference between Baa and Aaa yields. The spread series look very correlated. They never intersect and the Baa spread is always higher, and more variable, than the Aaa spread. Aaa spreads have been well above zero over most of the sample period, and have been ever increasing since the early 80s. Baa spreads have been high and volatile during the decades 1970-2000 and featured an increasing trends in recent years. Most high observations for Aaa spreads were recorded in the years following the Asian crisis, whereas high values of Baa spreads occur during the 70s and 80s. The relative spread has peaked in the first half of the 80s, after the 1979 turning point in U.S. monetary policy. Following a period of decrease, it has increased again in recent years, following the outbreak of the Asian crisis and Russia's default (1998). One can also see the large movement which occurred in the aftermath of the 9/11/2001 terrorist attacks.

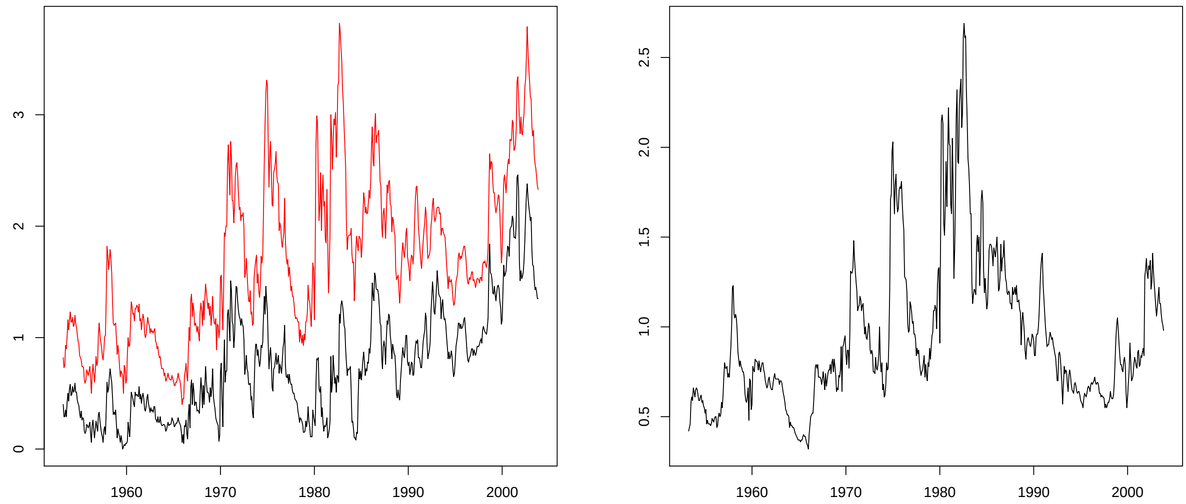
Statistics	S^{Aaa}	S^{Baa}
Mean	0.75	1.69
Standard dev.	0.50	0.72
Minimum	0.00	0.40
Maximum	2.46	3.82
Kurtosis	3.33	2.57
Skewness	0.82	0.44
JB stat	71.92	24.23
ADF stat	-4.48	-3.86

Table 1: *Summary Statistics for Spread Data*

Legend: S^{Aaa}, S^{Baa} denote spread on Aaa-rated bonds, and on Baa-rated bonds; JB is the Jarque-Bera test for normality, and ADF is the Augmented Dickey-Fuller test for unit root.

[1] http://research.stlouisfed.org/fred2/

[2] Both Aaa and Baa spreads exhibit moderate skewness and kurtosis. The Jarque-Bera test rejects normality in both series. The unit-root hypothesis is also rejected at 5% confidence level by a Dickey-Fuller (DF) test. (These results, however, should be interpreted carefully. In general, DF regressions are based on linear AR models, so their results could be misleading if the true dynamics is non-linear.)



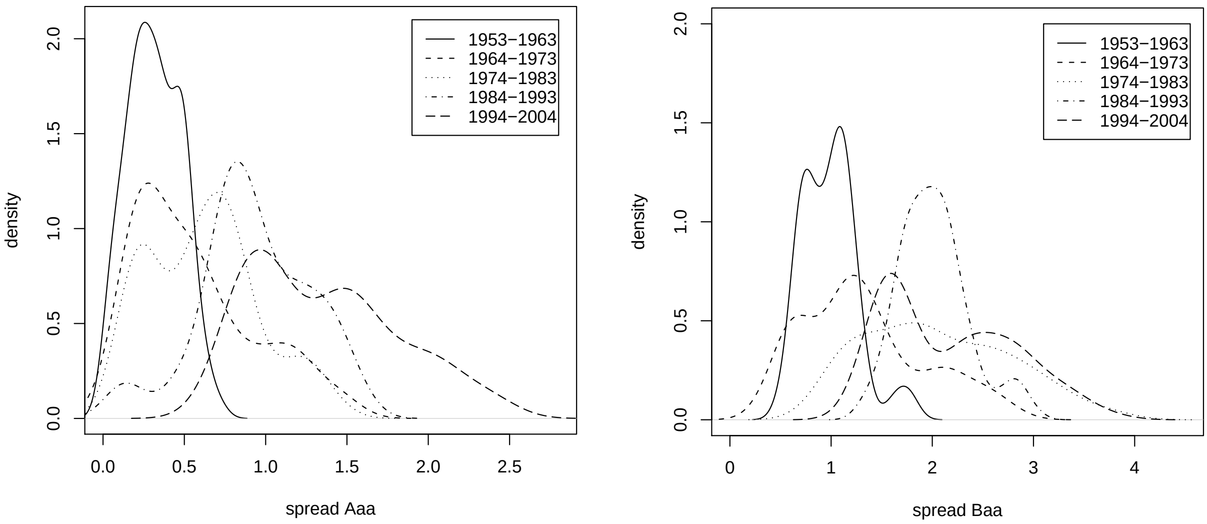
(a) (b)

Figure 1: Time series of **Aaa** (in black) and **Baa** (in red) spreads (a) and relative **Baa-Aaa** spread (b).

Kernel density estimates of the spreads, in figure 3, evidence non-normal features in the data, such as heavy tails and skewness. These features are not unusual in financial data and imply high probabilities of observing extreme values. Although kernel density plots are useful summaries of the data, they ignore the time series nature of the data. Indeed, figure 2 shows that when sub-samples of the original data are considered instability of the marginal distributions is apparent. This is relevant as many techniques used in the analysis of the term-structure assume the existence of a time invariant stationary distribution for interest rates. This suggests looking at the conditional densities, which give a better description of the data generating process than the marginal density when data are time series and provide insight into the dependence structure in the data. They also incorporate relevant information on features such as non-normality, and non-constancy of mean and variance. Figure 4 shows that densities of spreads conditional on their past values shares features with estimated marginal densities, such as skewness and heavy-tails, but do vary over the conditioning variable. (Here I only report conditional densities for a lag of 6-months.) There is also a suggestion of bimodality in the upper boundary of the data. (One should be aware, however, that densities in this region is noisy, due to the sparsity of observations).[3]

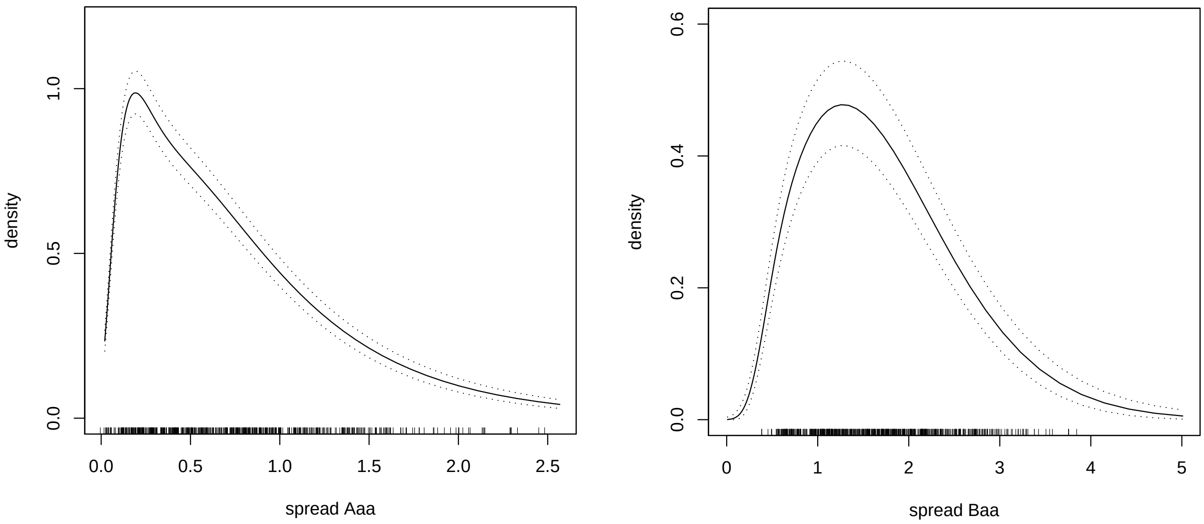
So, spread densities have non-normal features, and do not have constant feature both over time.

[3]Conditional density plots have been produced using the package `hdrcde`, by R. Hyndman.



(a) (b)

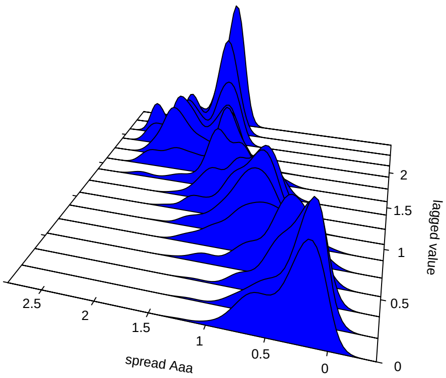
Figure 2: Density comparison for Aaa spread (a) and Baa spread (b).



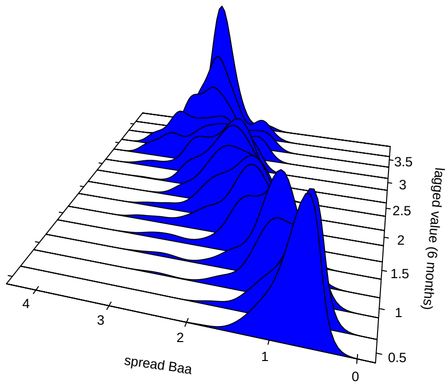
(a) (b)

Figure 3: Kernel density estimates of Aaa (a) and Baa spreads (b).

(The histogram at the basis of each plot represents the frequency of the observations; dotted lines are variability bands.)



(a)



(b)

Figure 4: Stacked conditional densities of Aaa (a) and Baa spreads (b) on past values (6 months).

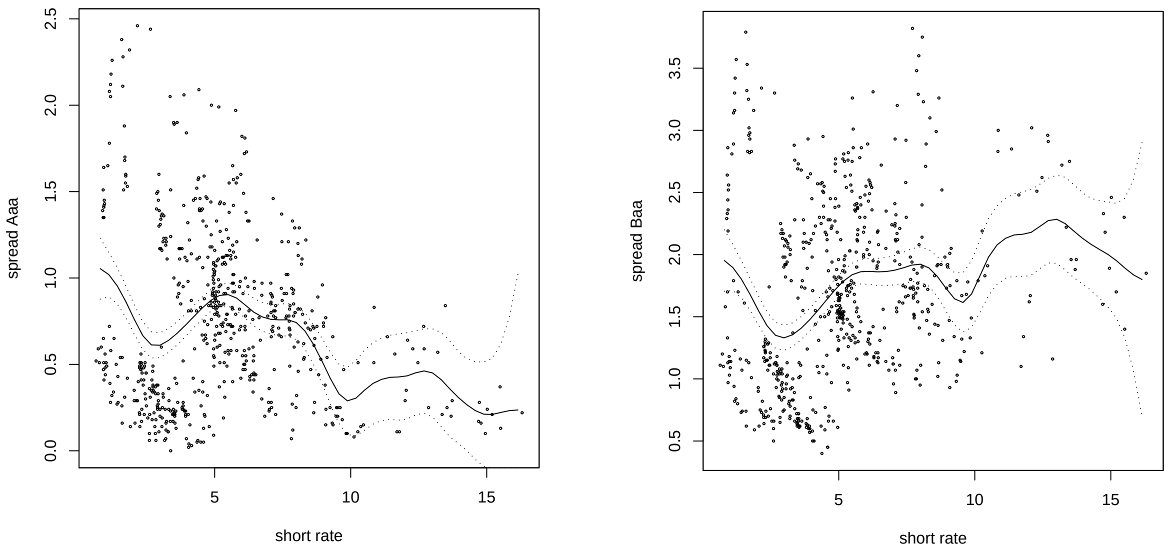
3 Spread and factors: a first look

This section presents a graphical analysis of the relationships of spreads and two important variables: risk-free short-term interest rate and inflation rate. (These variables enter the multi-factor term-structure model of the interest rate analysed in following sections.) The choice of the interest rate is motivated by existing theoretical and empirical literature on default risk, which suggests that risk-free rates are among determinants of yields on risky bonds (e.g., see Duffee, 1999). Inflation is a macroeconomic indicator related with default risk. (Here, the short-rate is measured by the 3-months Treasury-bill rate, whereas inflation is measured by the 12-months percent change in the Consumer Price Index released by the Bureau of Labor Statistics.)

The following estimates a simple regression model of spreads to the risk-free short-term rate, which allows departures from linearity. The model is estimated using a locally-linear regression technique (Fan and Gijbels, 1996), where the bandwidth is selected by cross-validation (Hardle and Marron, 1985). (Further details can be found in appendix A.)

The scatter-plots in figure 5, which shows spreads against the short-rate, evidence several clusters. These correspond to two groups of observations which associate low spread values with low interest rates, and mid-range spread values with mid-range interest rates. In the top-left area of the graphic, there is a group of observations characterised by the association of high spreads with low interest rates; this group consists of data for the period post 9/11.

The following non-linear model summarises the relationship of spreads to the risk-free



(a) (b)

Figure 5: Aaa (a) and Baa spread (b) vs. short-rate: scatter-plot and non-parametric regression curve (with indication of its variability).

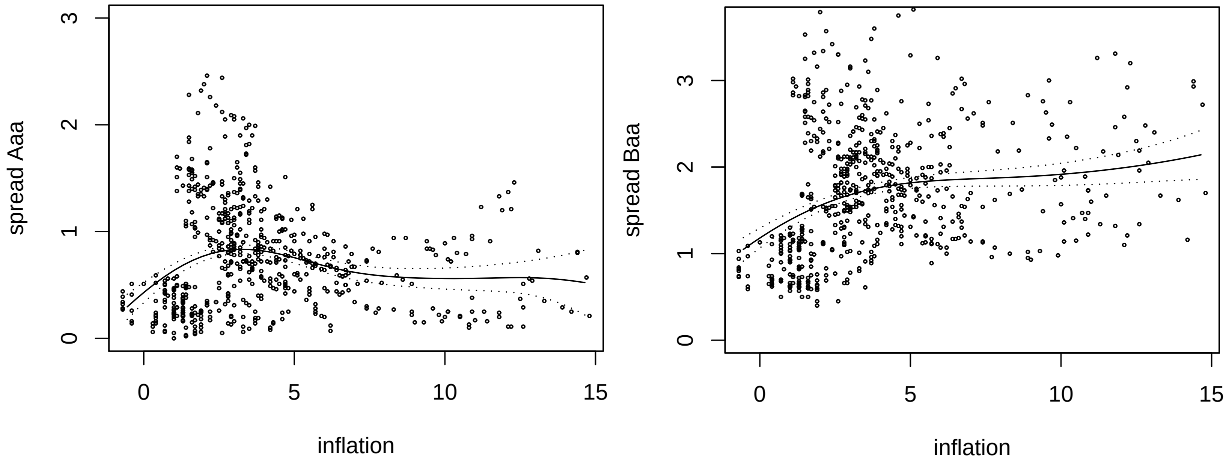
rate:

$$s_t^i = m(r_t) + \epsilon_t; \tag{8}$$

here, s is the spread, r the risk-less rate, $i = Aaa, Baa$ (in the following, the index i is omitted); m is a function whose shape is smooth but unrestricted, hence allowing departures from linearity, and ϵ is an *iid* error term. The continuous curves in figure 5 represent the non-parametric estimates of the regression function m, with variability bands. These estimates show a non-linear relation between short-term yields and spreads. The relation between Aaa spreads and the short-rate highly non-linear (left panel). The relation between Baa spreads and the short-rate, negative for low level of the short rate, becomes positive for higher values of the short-rate.[4]

Figure 6 present scatter-plots and kernel regression estimates of spreads against inflation rates. The relation between Aaa spreads and inflation is increasingly concave for low to mid-range inflation rates, turning negative for higher inflation rates ($> 5\%$). Smooth regressions indicates a positive concave relationship between Baa spreads and inflation rates, implying that yield spreads increases with inflation at decreasing rates. This seems in line with the view that increasing inflation correspond to increasing spreads.

[4]The estimation of the same model, leaving out the post 9/11 observations, provided a clearer non-linear pattern. There is evidence of two regimes in the Aaa spread to short-rate relationship: this is increasing for low interest rates, and decreasing for high interest rates (that is, greater than 5%). The relationship between Baa spreads and short-rate is increasing and concave.



(a) (b)

Figure 6: Aaa spread and Baa spread vs. inflation rate: scatter-plot and non-parametric regression curve (with variability bands).

Scatter-plots and regression lines give interesting insights into the relationships of interest: correlations between interest rate and spread, and inflation and spread, vary among rating classes, and are non-linear. This contrast with the existing literature, which indicates a negative sign for the correlation between spreads and short-rate (see, for, example, Duffee, 1998). This analysis, however, remains purely descriptive. Yields are known as better described by multi-factors models, as opposed to single-factor models. This is addressed in the following section.

4 Testing linearity in Reduced Form Models

This section explores the role of risk-free yields and inflation rate in determining spreads in the context of a multi-factors model of risky yields. It also tests a fundamental implication of affine models of credit risk: the linearity of yield-spreads in the underlying variables, or factors.

As shown in Section 1, RFMs imply a spread which is linear in factors. The model of equation 7 suggests a simple linear regression model, in which spreads depend on the risk-free rate and default risk. In order to estimate the model, it is essential to include variables able to capture default risk which are observable. Thus, this section assumes that corporate spreads are explained by factors capturing the risk-free term-structure (short rate and term spread) and macroeconomic conditions (the inflation rate). This yields the following regression model:

$$s_t = \alpha + \beta_1 r_t + \beta_2 \pi_t + \varepsilon_t; \tag{9}$$

Here, s denotes the corporate spread, r the risk-free short rate (the 3-months Treasury bill), y the term spread (i.e. the difference between the 10-years Treasury note yield and the 3-month Treasury Bill yield) and π the inflation rate; ε is a random *iid* error term. The *term spread*, also known as the slope of the yield curve, is included along the short rate to better capture the the risk-free term-structure.[5]

Our first objective is to identify a suitable alternative model to test the validity of the affine specification above. Consider the following model, in which the spread is described as a non-linear function of factors:

$$s_t = m(r_t, y_t, \pi_t) + \epsilon_t; \tag{10}$$

Here, variables are the same as above; m is a smooth function of factors, and ϵ an *iid* error term with zero mean and standard deviation σ. This model is non-parametric, in the sense that it does not place assumptions on the functional form of the relationship between spread and factors. Hence, it provides a general non-linear alternative to model (9), generality which would be greatly restricted by specifying a non-linear but parametric functional form.

In principle, model (10) can be estimated using standard non-parametric kernel regression techniques. (An extension of those illustrated in previous section for analysing bivariate relations.) When modelling relations among several variables, however, non-parametric analysis is more difficult due to increased computational costs and, more fundamentally, to a problem known as the *curse of dimensionality.* This indicates the worsening performance of local smoothers as the number of variables increases and motivates the use of dimension-reduction models, such as *Generalised Additive Models* (GAMs) (Hastie and Tibshirani, 1986) which only involve one dimensional functions.[6] GAMs retain the flexibility of non-parametric regression, while reducing the impact of the so-called curse of dimensionality. For this reason, the alternative is specified using a GAM framework as follows:

$$s_t = m_\pi(\pi_t) + m_r(r_t) + m_y(y_t) + \epsilon'_t; \tag{11}$$

This model is more tractable than the one of equation 10, because the response variable is the sum of one-dimensional non-parametric functions (also called smooth terms). These non-parametric terms are estimated using an iterative procedure, called *back-fitting* algorithm (Buja et al., 1989), whose basic building block is the local-linear smoother. (Further details are provided in Appendix B.) The estimation output is essentially graphical, as

[5] Single-factor models of the term-structure have been criticised for their inability to explain the observed variability of the yield curve through time and across maturities. Authors have argued that the term-structure dynamics is too complex to be summarised by a single-source of uncertainty. To address this, multi-factor representations, where the yield are explained by several state variables, have been introduced. Short and long-term rates are being used to explain intermediate maturities in the non-defaultable bond market (e.g. Schaefer and Schwartz, 1984; Knight et al., 2006).

[6] The convergence rate of a non-parametric estimator depends on the number of regressors; thus, the "distance" between the non-parametric estimator and the true value collapses at a much slower speed as the number of independent variables increases (Härdle, 1989, Section 4.1).

plots of individual smooth terms allow us to to examine the possible non-linear effect of each explanatory variable.

Tables 2 and 3 present estimation results for, respectively, the linear model of equation (9) and the non-parametric model of equation (16). Both models have been estimated separately for Aaa and Baa spreads.[7] Table 3 reports values of approximate significance tests for each smooth term. All non-parametric terms are statistically significants. The F test is analogous to the F statistic used to compare linear models, as it is based on (approximate) degrees of freedom and residual sum of squares. (For the computation of degrees of freedom in a non-parametric setting, see Hastie and Tibshirani, 1990, Chapter 2.) Although in the non-parametric setting a null distribution of the test is not available, F tables can be used as guidance (see Hastie and Tibshirani, 1990, Sections 5.4.4 and 5.4.5).

Figure 7 present estimates of the smooth terms of equation (16). (Plots on the left column are for the model estimated with Aaa spreads data; those on the right column refer to Baa spreads data.)

The inflation effect is highly non-linear; the shape of the smooth functions looks very similar in both models. The slope of the function is positive at very low inflation rates, turns negative at medium-range rates (between 3% and 7%), and again positive at high levels of inflation (above 7%). In the region characterised by high inflation rates, where only few observations are available, confidence bands are large, but not dramatically large, indicating that the estimation is reasonably stable.

The term-spread component is S-shaped in both models. (One should note, however, that variability bands are large at the boundaries, and are consistent with a U-shaped effect, at least for the Aaa spread.) Negative values of the variable (corresponding to yield curve inversion) are associated to a flat response in both models. In general, this effect is weaker for the Aaa spread.

The short rate effect is non-linear. At low and medium-range levels, the non-linear pattern is similar in both models, first decreasing and then increasing. For higher rates (i.e. greater than 7.5%), the smooth function has a decreasing trend in the Aaa spread model, and an increasing trend in the Baa data. (Here, confidence bands are large for values of the interest rate greater than 10%.)

A negative slope of the short rate effect is consistent with the view that economic downturns lead both to lowering interest rates and a deterioration in credit quality, thus to widening spreads. The positive slope in the medium-range interest rates region may be linked to a recovery in the economy and to the fear of inflationary pressures, leading to increasing spreads.

The overall impression is that there is a decreasing trend in the relationship of the Aaa spread to the short rate. On the other hand, a widening of the Baa spread is associated to increasing short rates. So, rising interest rates seem to lead investors away from defaultable bonds, lowering bonds prices, and increasing yields. If this is associated to rising term-spread, which suggests that inflation is expected to rise, investors also expect a deterioration in credit quality, and require higher yields to hold lower-rated corporate bonds. (Recall

[7]The GAM models have been estimated using the package `gam` in `R`.

that Baa spreads are the lowest rated of investment-graded bonds.)

Model	term	estimate (se)	t value	$Pr\ (>\lvert t\rvert)$	Adj. R^2	F-stat
Aaa spread	int.	0.619 (0.049)	12.541	0.000	0.169	44.32
	π	0.026 (0.009)	2.911	0.003		
	r	-0.034 (0.009)	-3.539	0.000		
	y	0.157 (0.015)	9.886	0.000		
Baa spread	int.	0.783 (0.06)	12.496	0.000	0.342	111.4
	π	0.069 (0.01)	5.919	0.000		
	r	0.037 (0.01)	3.032	0.002		
	y	0.334 (0.02)	16.507	0.000		

Table 2: **Determinants of Credit Spread: linear regression.**

Model	term	Npar df	Npar F	$Pr\ F$	$RSS\ (df)$	AIC
Aaa spread	$m(\pi)$	2.6	18.25	0	94.28 (558)	620.40
	$m(r)$	3.4	24.38	0		
	$m(y)$	3.2	6.82	0		
Baa spread	$m(\pi)$	2.6	9.49	0	159.75 (558)	922.05
	$m(r)$	3.4	20.77	0		
	$m(y)$	3.1	5.10	0.001		

Table 3: **Determinants of Credit Spread: GAM model: approximate significance of smooth terms.**

Legend: Npar df denotes degrees of freedom, and Npar-F the approximate F value for each smooth term; pr F are corresponding p-values. RSS is the sum of squared residuals, with degrees of freedom (dof) in parentheses, and AIC is the Akaike Information Criterion for the overall model.

A positive relationships between the spread and the inflation rate is consistent with the view that rising inflation increases the number of bankruptcies. Thus, bond-holders request compensation for the higher default risk, which leads to an increase in the corporate spread. The mechanism through which inflation affects the probability of default has been effectively illustrated in Wadhwani (1986): an increase in inflation causes cash-flow problems for firms, due to an unexpected rise in (nominal) interest rate payments. Interpretation here is complicated by the fact that the response of spreads to inflation is not constant but appears to vary with different levels of inflation. Indeed, it can be observed that the positive relationship spread-inflation corresponds to very high and very low levels of inflation. This can be interpreted as follows: (a) very low levels of inflation, usually associated to a stagnating economy (deflation), increase the market's perception of risk and boost corporate yields; (b) investors' perceive very high rate of inflation as a sign of worsening economic conditions, and this, in addition to the cash-flow effect described above, determines an upward pressure on spreads. Indeed, historically, the US economy was characterised by high levels of inflation in the 70s, the period known as "stagflation", a mixture of high inflation and stagnation. In contrast, mid-range inflation rates are perceived by investors as symptoms of a growing economy, thus, corporate spreads narrow. Indeed, these levels of inflation have been observed during good times for the U.S. economy, such as mid 80s and mid 90s.

The graphs of figure 8 compare historical Aaa and Baa spread values with spread values predicted from the non-parametric model of this section. The left panel compares fitted and observed Aaa spread. As it can be observed, the model is good at capturing movements in the actual series. It is capable of predicting the increasing trend in the spread which occurred since the mid 90s (although the actual spread is clearly underestimated), and it also follows the actual series closely in the first two decades of the sample. The model's performance, however, weakens in the middle years of the estimation period; in particular, it fails to predict the sudden increases and subsequent contractions that characterised the spread series during the 80s. The right panel, which show observed and fitted values of the Baa spread, presents similar results. The non-parametric model is good at capturing features of the actual series, but its performance worsens in the middle years of the sample (e.g. it cannot detect the drop in Baa spread that follows its historical maximum in 1982).

Overall, the results show that the non-parametric models are capable of capturing the behaviour of the spread. In particular, non-linear functions of inflation and risk-free rates are capable of explaining the behaviour of the spread.

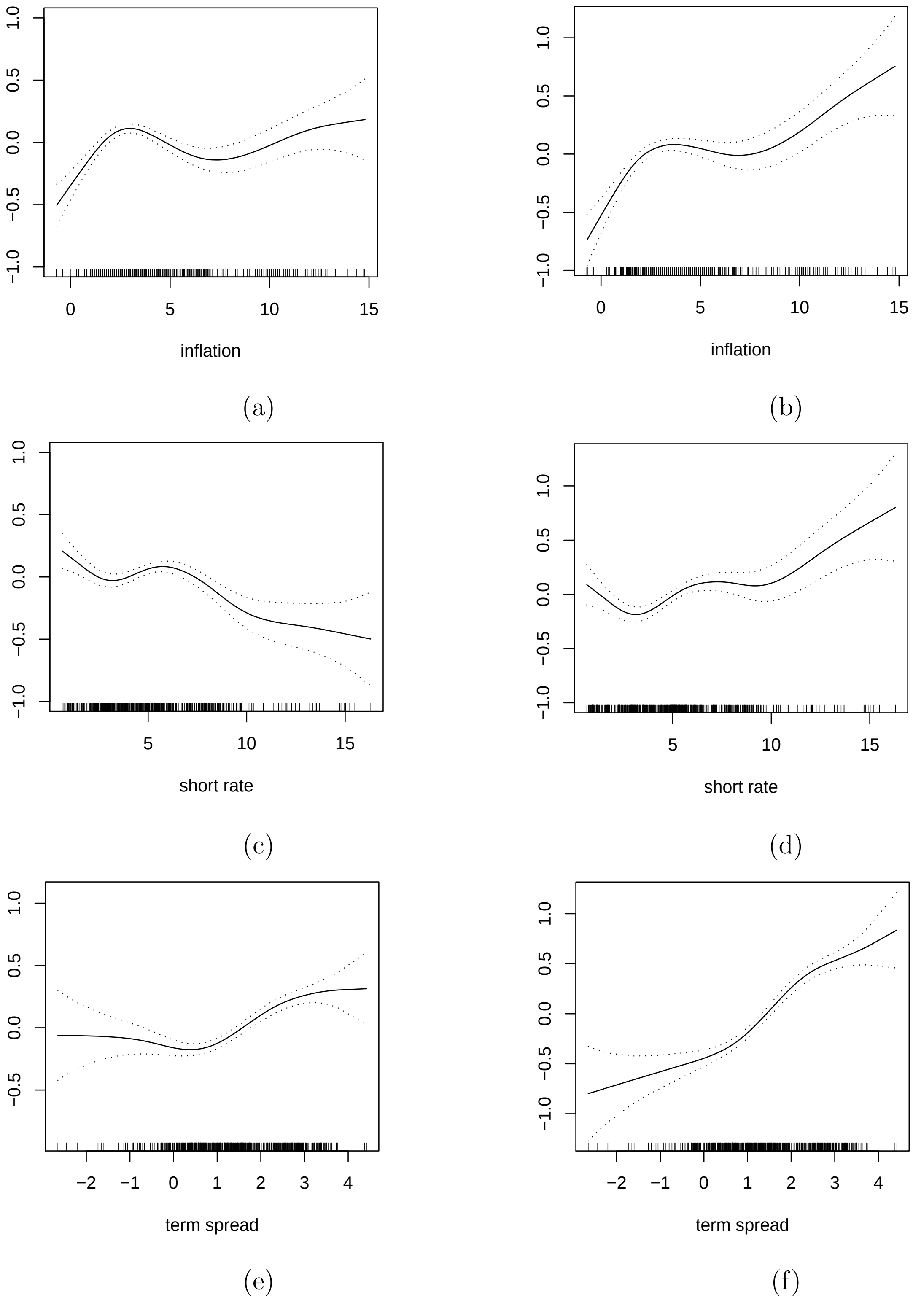


Figure 7: GAM model, estimates of smooth terms: inflation (a,b), risk-free rate (c,d), and term-spread component(e,f) for Aaa (left column) and Baa spread (right column).
(The "histogram" at the basis of each plot represents the frequency of the observations, dashed lines are variability bands.)

4.1 The GLR test

This section tests for the linearity of the non-parametric spread model. It compares the linear specification of the spread, obtained by solving RFMs for the price of bonds, against a model in which spreads are non-linear functions of factors.

The following outlines the testing procedure for the linearity of RFMs, which uses the Generalised Likelihood Ratio (GLR) approach, introduced by Fan et al. (2001). This is a general framework for testing hypotheses in a non-parametric setting, based on a generalised version of the Likelihood Ratio principle. Suppose we have a model that represents yields to maturity as linear functions of p factors:

$$y(\tau) = a(\tau) + \sum_{i=1}^{p} b_i(\tau) x_i; \tag{12}$$

Here, xs are factors, a and b parameters, and τ denotes the time to maturity. Formally, the linearity testing problem for such model is written as follows:

$$H_0 : y = \alpha + \beta_1 x_1 + ... + \beta_p x_p \quad vs. \quad H_1 : y \neq \alpha + \beta_1 x_1 + ... + \beta_p x_p, \tag{13}$$

where H_0 represents the null linear model, and H_1 the non-linear alternative. The testing procedure should comprise the following steps:

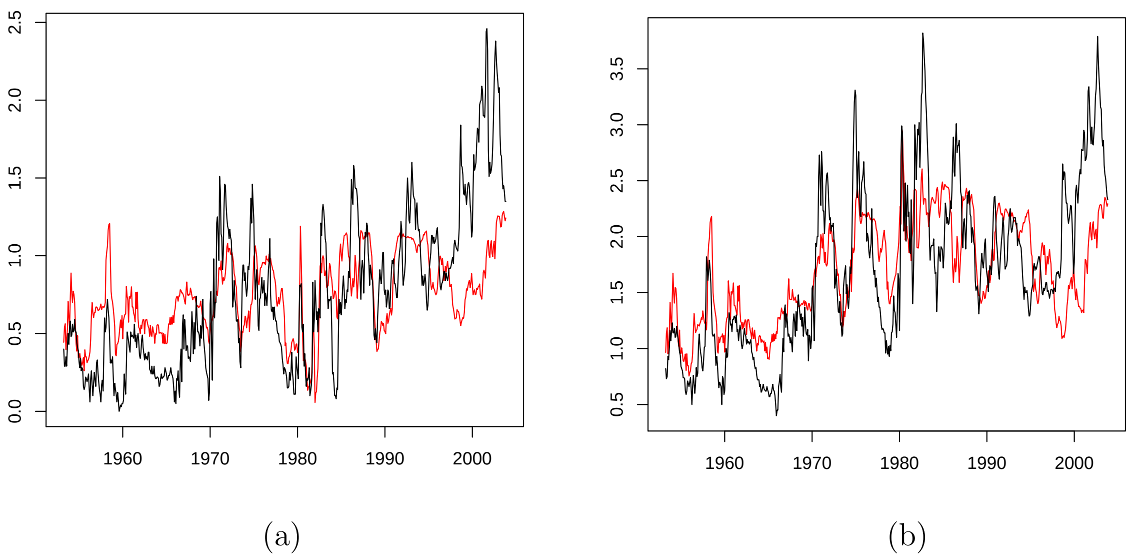


Figure 8: GAM models with inflation: comparing observed (in black) and fitted values (in red) of Aaa spread (a) and Baa spread (b).

1. write the alternative non-parametric model using a GAM model structure. This specifies H_1 as follows:

$$y = m_1(x_1) + m_2(x_2) + ... + m_p(x_p) + \varepsilon, \tag{14}$$

where the terms ms are univariate non-parametric functions of each factor;

2. estimate the alternative model, using the back-fitting algorithm, and compute the GLR statistics as:

$$\lambda_n(h) = \frac{n}{2} \log \frac{RSS_0}{RSS_1}, \tag{15}$$

where n is the number of observations, RSS_0 and RSS_1 are residuals sum of squares from the null and alternative models;

3. use bootstrap methods to simulate the null distribution of the GLR statistics, and compute p-values.

The test produces high values of the test statistics. This can be seen in table 4. The second column reports observed values of the GLR statistics. After computing the value of the GLR test-statistics, its null distribution is computed using a version of the bootstrap method detailed in Fan and Jiang (2005). (The idea is that, since the asymptotic null distribution is independent on nuisance parameters/functions, for finite samples the null distribution can be approximated by a bootstrap method.) The bootstrap procedure comprises the following steps:

1. fix the value of the bandwidth at its estimate $\hat{h}$. For the original data, compute the observed value of the test statistics $\lambda_n(\hat{h})$ according to eq. 15;

2. sample randomly and with replacement from the residuals obtained at step 1. Define the bootstrap responses $s^b = \hat{\alpha} + \hat{\beta}_1 r + \hat{\beta}_2 s + \hat{\epsilon}^b$. This forms a bootstrap sample $\{S^b; R, Y\}$;

3. use the bootstrap sample to obtain the GLR statistics $\lambda_n^b(\hat{h})$;

4. repeat steps 2 and 3 n times to obtain a sample of GLR statistics. Here the number of bootstrap replications is set at $n = 1000$.

The test rejects the null hypothesis of linearity. This can be seen, again, in table 4. Bootstrap p-values are listed in the third column. (P-values are computed as the proportion of times that the bootstrap statistics $\lambda_n^b(\hat{h})$ exceeds the observed value $\lambda_n(\hat{h})$.)

This procedure presents two difficulties: first, it is well known that p-values depend heavily on the sample size; second, the available data are time-series. Although the model-based bootstrap adopted here can be applied to the analysis of time series data (Davison and Hinkley, 1997, Chapter 8), it is useful to analyse sub-samples of data to assess the robustness of our results. Hence, we take random sub-samples of 200 observations for analysis. We repeat the procedure 100 times, as it is computationally very expensive, and obtain p-values for each sub-sample. The fourth and five columns 4 report average p-values for the sub-samples. This provides evidence that the non-parametric model is appropriate for the sub-samples at the nearly zero significance level.

We conclude that the proposed test provides formal proof against the linear specification of the yield-spread model.

5 Forecasting the spread

In previous sections, we have identified a non-linear pattern in historical corporate spread data. This section conducts a simple *out-of-sample* forecasting experiment, to help determine whether the non-parametric model estimated in the previous section is useful to forecast the future course of the corporate spread.

There, the spread (s) was modelled as a non-parametric function of the term-spread (y), the short rate of interest (r), and the inflation rate (π):

$$s_t = m_\pi(\pi_t) + m_r(r_t) + m_y(y_t) + \epsilon_t; \tag{16}$$

Based on the model above, forecasts for the out-of-sample period are formed conditionally on the values of the explanatory variables, which are assumed known. Thus, following estimation over $t = 1, ..., T$, 1-step-ahead forecasts are computed as follows:

$$\hat{s}_{T+h} = \hat{m}_\pi(\pi_{T+h}) + \hat{m}_r(r_{T+h}) + \hat{m}_y(y_{T+h}), \qquad h = 1, .., H; \tag{17}$$

where h indices out-of-sample observations. This additive framework provides a simple and parsimonious way of generalising (in the non-parametric, non-linear sense) a predictive linear model. The advantage of this GAM representation is amenability of computation. There are procedures capable of evaluating the smooth fitted functions at new (out-of-sample) values of the covariates, at least if these are in the domains of the original data (Chambers and Hastie, 1992).

In order to perform this experiment, the sample is divided into estimation and forecasting period; the 30 most recent observations are reserved for forecasting purposes (these go from 2004:1 to 2006:6).

The model of equation 16 is estimated using in-sample data; then, the fit is used to produce predicted values of the spread over the out-of-sample period. Graphical analysis enables forecasts to be compared to observed spreads. The left panel in figure 9 presents

Model	GLR test	p-value	sub-sample p-value
Aaa spread	73.04	0	0.0
Baa spread	50.60	0	0

Table 4: Testing linearity of yield-spread models

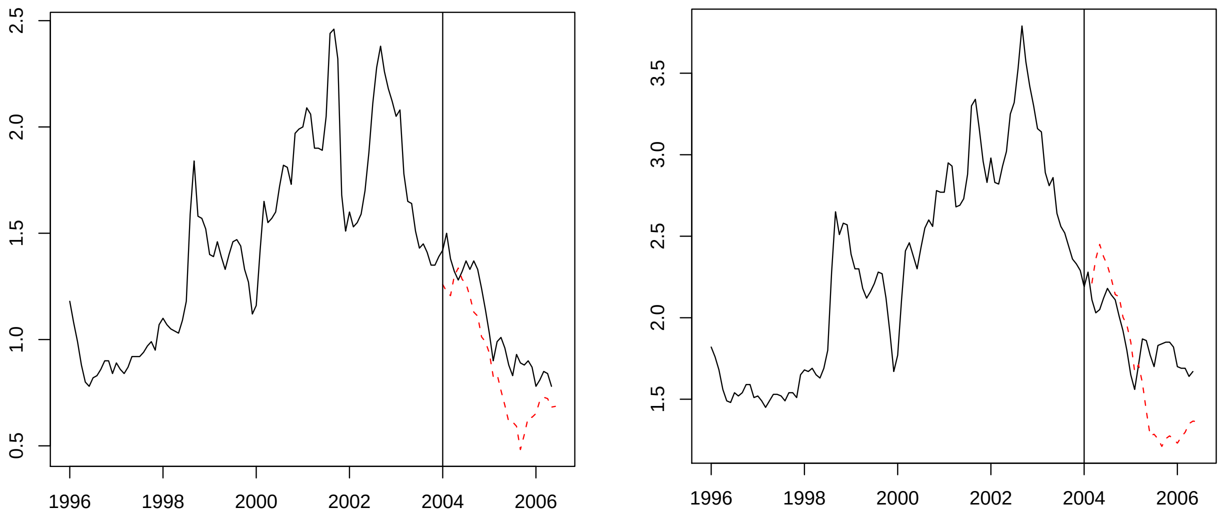
actual (s_t) and fitted values ($\hat{s}_t$) of the Aaa spread (the latter are reported in red, and refers to the forecasting period only); the right panel does the same for the Baa spread. These results are encouraging: the fitted models capture the declining trend in the spread over recent years. However, an interesting feature of the graphs is the suggested upward trend in the spread forecasts for the last few observations. A closer look at observed spreads shows that the decreasing trend that followed the 2001 peak has recently slowed down. In the last few months of the sample period, spreads have been increasing again, as well as corporate yields. Others prominent features of the data for the forecasting period are the increase in inflation, and the large increase in the short rate of interest, which led to a flattening of the yield curve. Forecasts of the spread model seem capable of capturing these tendencies in the data.

Figure 9 shows that, at the end of the forecasting period, the model underestimates realised spreads. In general, forecasts are almost never precisely accurate. Thus, to evaluate the forecasting performance of a model one should compare its predictions to those of other competing models, on the basis of some appropriate criteria. Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE) are widely used measures of forecast accuracy (Gooijer and Hyndman, 2006), given by:

$$RMSE = \sqrt{\frac{1}{H}\sum_{t=1}^{H}[y_t - f_t]^2}; \tag{18}$$

$$MAPE = 100/H \sum_{t=1}^{H} \frac{|y_t - f_t|}{y_t}; \tag{19}$$

Here f is the forecast, y the observed value of the series, and H the length of the forecasting period. These measures are based on the forecast error in t, defined as $e_t = y_t - f_t$, and on a loss function $L = l(e)$: the RMSE, for example, is based on a quadratic loss function. The MAPE is based on the average distance between forecasts and actual values, without regard to whether individual forecasts are overestimates or underestimates. The RMSE also shows the size of the error without regard to sign, but it gives greater weight to larger errors.



(a) (b)

Figure 9: **Aaa spread** (a) and **Baa spread** (b): realised and predicted values (in red).

The following compares the forecasting performance of the non-parametric spread model of equation 16 (the base model) to several alternative specifications: 1) a linear version of the base model; 2) an additive model which includes past values of the spread; 3) a benchmark linear first-order auto-regression, denoted by $AR(1)$. The linear model is as follows:

$$s_t = \alpha + \beta_\pi \pi_t + \beta_r r_t + \beta_y y_t + \varepsilon_t; \tag{20}$$

The additive models with lagged spread is specified as:

$$s_t = m_\pi(\pi_t) + m_r(r_t) + m_y(y_t) + m_l(s_{t-4}) + \epsilon'_t; \tag{21}$$

(Here, the selected past value of the spread is the spread observed at the beginning of the previous quarter.) Models estimates are used to construct forecasts and forecast errors.[8]

Tables 5 and 6 report, along with measures of (in-sample) goodness of fit, the two criteria of forecasting performance, RMSE and MAPE. For both Aaa and Baa spreads, the RMSE and MAPE have their minima for the non-parametric model with past values of the spread. The base model, however, does better than the linear model and the $AR(1)$ model when estimated using Aaa spread data. Its performance, however, worsens for Baa spread data.

[8]The $AR(1)$ model has been chosen accordingly to the AIC criterion. The model estimates and forecasts are produced using the package `stats` and `forecast` (the latter by R. Hyndman). The additive model with lags includes the spread observed at the beginning of the previous quarter (i.e. at time $t-4$). We chose to include only one lagged value because of the dimensionality problem that comes with additional explanatory terms in non-parametric estimation.

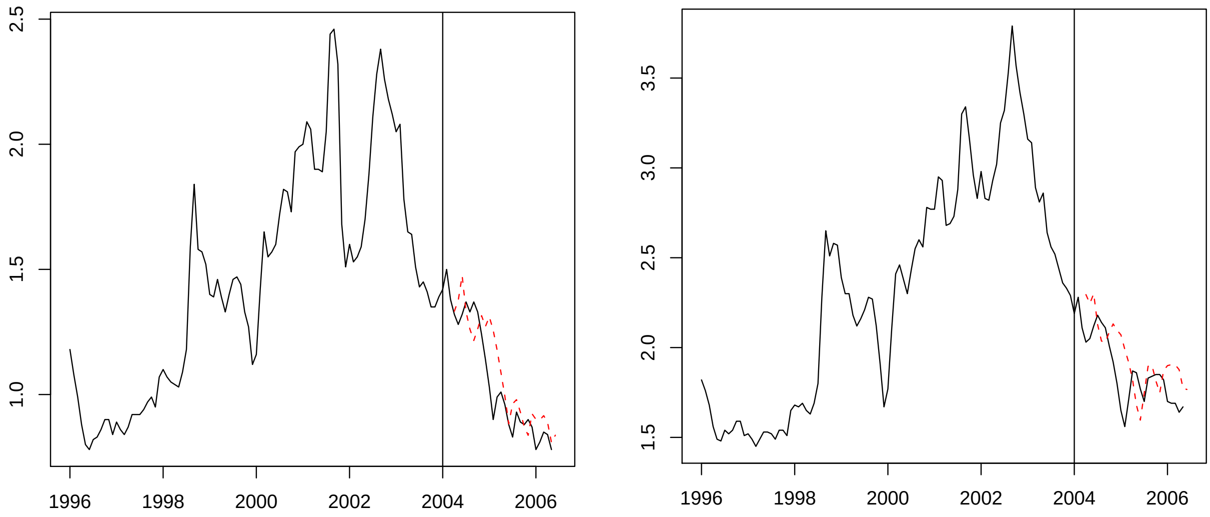
Figures 10 – 12 present comparisons of forecasts and observed values for the additive model with lags, the linear model, and the $AR(1)$ dynamics. Looking at the various graphs, one can see that the non-parametric model with lag, and the AR(1) model, tend to overestimate the observed spread. The linear model, instead, tends to underestimate it. These tendencies are not detected by accuracy criteria such as RMSE and MAPE, which do not distinguish between positive and negative forecast error. Furthermore, forecasts based on AR(1) do not capture the increase in spreads at the end of the sample. Thus, we believe that the non-parametric model of equation 16 still provides useful information for forecasting, and is preferable to models based on linear specifications.

Model	RSS	AIC	RMSE	MAPE
Model 16	98.69	648.10	0.182	15.64%
Model 20	126.71	781.93	0.248	21.60%
Model 21	25.76	-196.30	0.112	8.65%
$AR(1)$	9.60	-788.17	0.184	17.13%

Table 5: Forecasting models of the Aaa spread
Legend: RSS denotes the residual sum of squares, AIC the Akaike information criterion; RMSE is the root mean square forecast error, and MAPE the mean absolute percentage error.

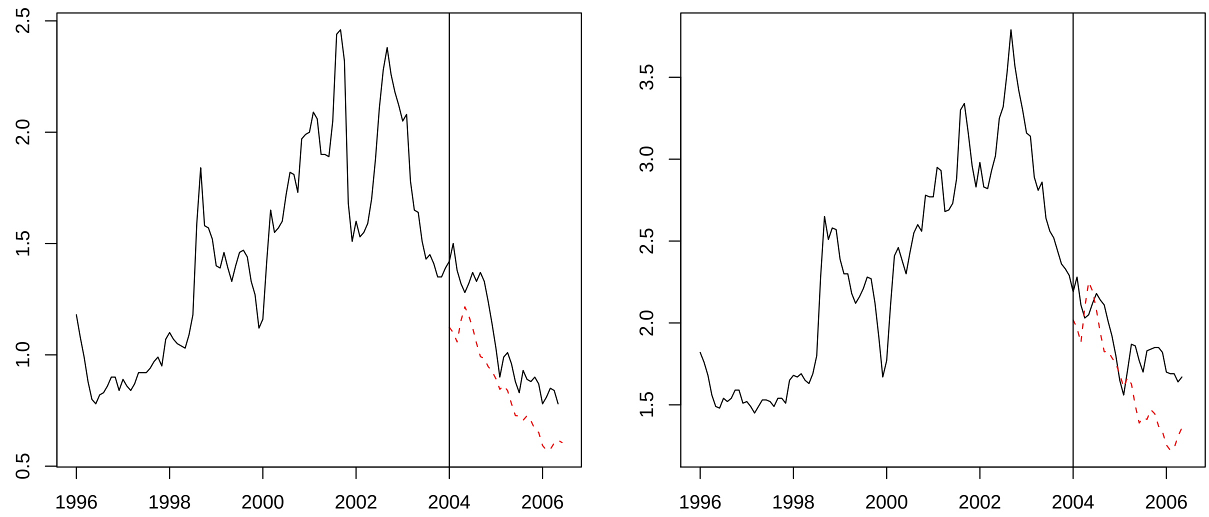
Model	RSS	AIC	RMSE	MAPE
Model 16	171.41	984.85	0.367	16.44 %
Model 20	203.84	1070	0.323	15.39%
Model 21	48.18	220.12	0.177	8.35%
$AR(1)$	17.63	-421.83	0.204	9.80%

Table 6: Forecasting models of the Baa spread
(Legend: as above.)



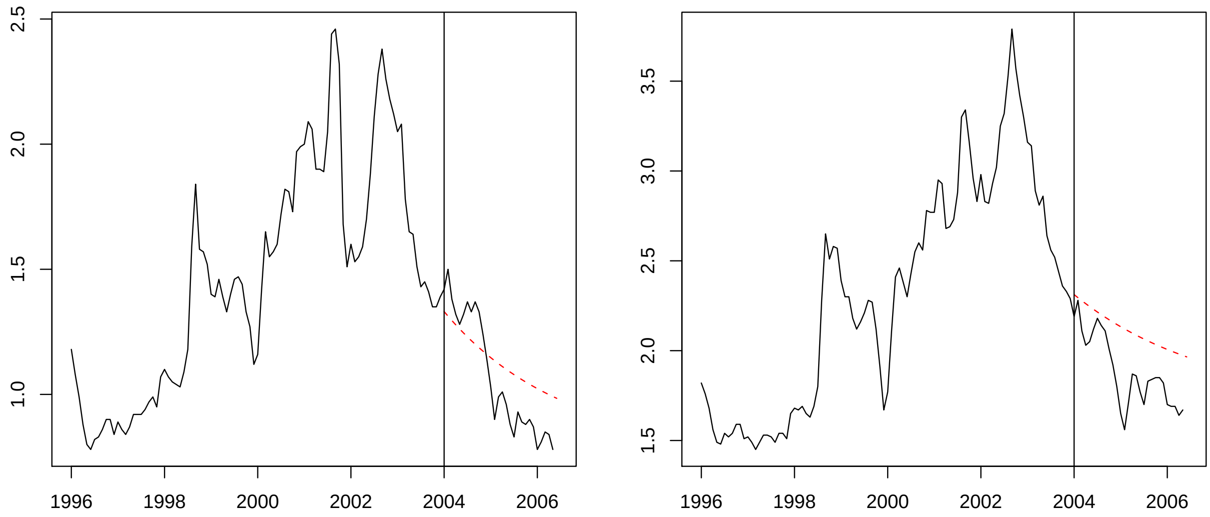
(a) (b)

Figure 10: Additive model with lags (eq. 21): observed and forecasted values (in red) for **Aaa spread** (a) and **Baa spread** (b).



(a) (b)

Figure 11: Linear model (eq. **??**): observed and forecasted values (in red) for **Aaa spread** (a) and **Baa spread** (b).



(a) (b)

Figure 12: AR(1) model: observed and forecasted values (in red) for **Aaa spread** (a) and **Baa spread** (b).

6 Conclusions

This paper tested affine RFMs of credit risk against their non-parametric counterparts with the aim of better explaining the determinants of risk premia. The analysis, based on corporate spread indices, showed that, despite their common use, there is no empirical evidence to support the restrictions imposed by affine models. On the contrary, parametric choices in affine RFMs are too restrictive.

This paper showed that functions that describe how the models' factors contribute to the determination of spreads are non-linear. This was confirmed by a formal test for the linearity of a regression function. The goodness-of-fit of the non-linear models over the estimation period, and the adequacy of the non-parametric approach, was further evaluated by comparing observed spreads and predicted values from non-parametric models. The non-parametric models have proved to be able to fit the data well.

The paper demonstrated the role of the inflation rate in explaining corporate spreads. Increasing inflation is associated to widening spreads, and this relationship is stronger when inflation is low (or high), which is consistent with the view that higher inflation increases the number of bankruptcies, leading to an increase in the spread.

This paper demonstrated also the potential advantages of forecasting the future course of the spread using non-parametric methods. Forecasts based on non-parametric models out-perform forecasts based on linear models. Linear time series models are widely used in forecasting. In general, this is motivated by the assumption that data are normally distributed. Yet, as shown in this paper, many financial variables do not have a normal distributions. Non-parametric modelling avoids the problem of structural instability in the parameters, which is well known to cause the break-down in the forecasting performance of linear predictive models, because it naturally accomodates changes in parameters. Regime changes, varying parameters, and more complex non-linearities are captured without the need of pre-specifying models' functional forms.

Computations. The analysis performed in this thesis have been carried out using the open-source statistical software `R` (R Development Core Team, 2007).

Acknowledgements. Thanks are due to Peter Spencer, my PhD supervisor, and to my advisors Karim Abadir and Gabriel Talmain. I also thank seminars participants at the University of York, Manchester, and at the University of East Anglia.

References

D. Y. Ahn and B. Gao. A parametric nonlinear model of term structure dynamics. *The Review of Financial Studies*, 12:761–762, 1999.

Y. Aït-Sahalia. Non-parametric pricing of interest rate derivative securities. *Econometrica*, 64:527–560, 1996.

A. Ang and M. Piazzesi. A no-arbitrage vector autoregression of term structure dynamics with macroeconomics and latent variables. *Journal of Monetary Economics*, 50:745–787, 2003.

M. Arapis and J. Gao. Empirical comparisons in short-term interest rate models using nonparametric methods. *Journal of Financial Econometrics*, 4:310–45, 2006.

A. Buja, T. Hastie, and R. Tibshirani. Linear smoothers and additive models. *Annals of Statistics*, 17:453–555, 1989.

J. M. Chambers and T.J. Hastie. *Statistical models in S.* Chapman and Hall, New York, 1992.

Q. Dai and K. J. Singleton. Specification analysis of affine term structure models. *The Journal of Finance*, LV:1943–1978, 2000.

A. C. Davison and D. V. Hinkley. *Bootstrap methods and their application.* Cambridge University Press, 1997.

G. R. Duffee. The relation between treasury yields and corporate bond yield spread. *The Journal of Finance*, 53:2225–2241, 1998.

G. R. Duffee. Estimating the price of default risk. *The Review of Financial Studies*, 12: 197–226, 1999.

D. Duffie and K. J. Singleton. An econometric model of the term structure of interest-rate swap yields. *The Journal of Finance*, LII:1287–1321, 1997.

D. Duffie and K. J. Singleton. Modeling term structures of defaultable bonds. *The Review of Financial Studies*, 12:687–720, 1999.

E. J. Elton, M. J. Gruber, D. Agrawal, and C. Mann. Explaining the rate spread on corporate bonds. *The Journal of Finance*, 56:247–277, 2001.

J. Fan. Design-adaptive nonparametric regression. *Journal of the American Statistical Association*, 87:998–1004, 1992.

J. Fan and I. Gijbels. *Local polynomial modelling and its applications.* Chapman and Hall, 1996.

J. Fan and J. Jiang. Nonparametric inference for additive models. *Journal of the American Statistical Association*, 100:890–907, 2005.

J. Fan, C. Zhang, and J.Zhang. Generalized likelihood ratio statistics and wilks phenomenon. *The Annals of Statistics*, 23:153–93, 2001.

J. G. De Gooijer and R. J. Hyndman. 25 years of time series forecasting. *International Journal of Forecasting*, 22:443–473, 2006.

W. Härdle. *Applied nonparametric regression.* Number 19 in Econometric Society Monographs. Cambridge University Press, 1989.

W. Hardle and J. S. Marron. Optimal bandwidth selection in nonparametric regression function estimation. *The Annals of Statistics*, 13:1465–81, 1985.

T. J. Hastie and R. Tibshirani. Generalised additive models. *Statistical Science*, 1:297–318, 1986.

T. J. Hastie and R. J. Tibshirani. *Generalized Additive Models.* Chapman and Hall, 1990.

J. Knight, F. Li, and M.Yuan. A semiparametric two-factor term structure model. 4:204 – 237, 2006. Journal of Financial Econometrics.

F.A. Longstaff, S. Mithal, and E. Neis. Corporate yield spreads: default risk or liquidity? UCLA Working Papers, 2004.

E. Mammen, O. Linton, and J. P. Nielsen. The existence and asymptotic properties of a backfitting projection algorithm under weak conditions. *Annals of Statistics*, 27:1443–1490, 1999.

J. Opsomer and D. Ruppert. Fitting a bivariate additive model by local polynomial regression. *Annals of Statistics*, 25:186–211, 1997.

R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2007. URL `http://www.R-project.org`. ISBN 3-900051-07-0.

S. M. Schaefer and E. S. Schwartz. A two-factor model of the term structure: an approximate analytical solution. *Journal of Financial and Quantitative Analysis*, 19:413–424, 1984.

S. J. Sheather and M. C. Jones. A reliable data-based bandwidth selection method for kernel density estimation. *Journal of the Royal Statistical Society, Ser. B*, 53:683–90, 1991.

S. B. Wadhwani. Inflation, bankruptcy, default premia and the stock market. *The Economic Journal*, 96:120–138, 1986.

Appendix

A The local-linear approach

Consider the simple regression model of y on x:

$$y_i = m(x_i) + \epsilon_i, \quad i = 1, .., n; \tag{22}$$

Here, m is a smooth function of x, and ϵ is an *iid* error term with zero mean and standard deviation σ.

The *local-linear* regression approach (Fan, 1992; Fan and Gijbels, 1996) provides a flexible method to estimate m, based on local averaging of data points. Theoretically, it is motivated by a Taylor expansion of the regression function m around x (Fan, 1992). In practice, it proceeds by dividing the sample into intervals of equal length (along the X direction), and by running local linear regressions on each interval.

Formally, the local-linear estimator $\hat{m}_h(x)$ is given by the a^* which minimises the weighted least squares problem:

$$\min_{a,b} \sum_i \{y_i - a - b(x_i - x)\}^2 K(x_i - x, h); \tag{23}$$

The "ingredients" of the problem are as follows:

- the polynomial of order 1, $a - b(x_i - x)$ approximates the regression function around x;
- the *kernel* function $K(x - .)$ weights the observations falling into the interval around x. It is generally agreed that the choice of the kernel is not very important for the estimation's outcome. In most applications, K is the Gaussian density;

- the smoothing parameter, h (the *bandwidth*) controls the width of the kernel, hence, the degree of smoothing applied to the data. The choice of h has a large effect on estimation results, and can be done either subjectively by the researcher, or objectively by data, using methods such as Cross Validation (Hardle and Marron, 1985) or Plug-In (Sheather and Jones, 1991).

B Generalised Additive Models

Generalised Additive Models (GAMs) constitute a non-parametric version of the multiple linear regression model. The dependent variable is modelled as a sum of smooth functions of the covariates:

$$y = \alpha + m_1(x_1) + ... + m_d(x_d) + \varepsilon; \tag{24}$$

Here, the m_js are smooth functions of the explanatory variables $x_1, \ldots, x_d$, and ε is an *iid* error term. In general, the conditional mean of the response is related to the additive function via a link function η:

$$E[y|\mathbf{x}] = \eta(\alpha + \sum m_j(x_j)); \tag{25}$$

Additive models are a special case of GAMs, in which the link function is the identity, that is, $\eta(\cdot) = \alpha + \sum m_j(x_j)$.

Clearly additivity is quite a strong assumption. However, the structure of equation 24 provides a very convenient way to represent and estimate a non-parametric multiple regression model, for several reasons: (1) it is flexible enough to allow departures from linearity; (2) the model is far easier to interpret than a d-dimensional surface, and provides a graphical output; (3) the additive feature provides a basis for inference in the non-parametric context. Residual sum of squares (RSS) and approximate degrees-of-freedom (df) can be calculated, and these quantities can be used to assess the significance of each smooth term (explanatory variable) in modelling the response. Examples of additive modelling, its properties, inferences, and generalisations are presented in Hastie and Tibshirani (1990), which remains the most comprehensive source for GAMs to date.

An iterative *back-fitting algorithm* is used for fitting additive models, and consists of the following steps:

- Step 1. Initialisation ($i = 0$):
  $\hat{\alpha} = \sum y/n$; $m_j^0 = 0, \quad j = 1, .., d$;

- Step 2. $i = i + 1$;
  for each $j = 1, .., d$ define partial residuals $\epsilon_j = y - \hat{\alpha} - \sum_{k \neq j} \hat{m}_k^{(i)}(x_k)$;
  compute $\hat{m_j}^{(i+1)} = S_j(\epsilon_j)$, where S denotes a univariate smoother (usually, this is a local-linear smoother or a spline smoother);

- Step 3. Keep cycling step 2 until convergence is reached.

The idea is to produce a fit, compute partial residuals, re-fit, until some convergence criterion is satisfied. One can see that the basic building block of the algorithm is the linear smoother S, which estimates the individual functions. (In this paper, univariate functions are fitted by the local-linear smoother.) Intuitively, a justification to this approach is provided by the following fact: if the additive model is correct, then

$$E[Y - \alpha - \sum_{k \neq j} m_k(X_k)|X_j] = m_j(X_j); \tag{26}$$

Buja et al. (1989) provide a detailed analysis, and a rigourous justification, of the procedure. The implementation of the back-fitting algorithm in the programming language S is presented in Chambers and Hastie (1992, Chapter 7).

Opsomer and Ruppert (1997) provided asymptotic bias and variance properties of an additive model with two explanatory variables, and smoother matrices, under quite strong conditions. Mammen et al. (1999) proved consistency and calculated asymptotic properties under weaker conditions. In the general case, because of its iterative nature, theoretical results for back-fitting are difficult.