Università degli Studi di Bergamo

Dottorato di Ricerca in

Metodi computazionali per le previsioni e decisioni economiche e finanziarie

Fractional Models to Credit Risk Pricing

Candidato:
Arturo Leccadito

Supervisor:
Prof. Giovanni Urga

To my parents.

I'm not even supposed
to be here today!

Dante Hicks — "Clerks"

Harold Crick was a man
of infinite numbers,
endless calculations
and incredibly few words.

Karen Eiffel — "Stranger
Than Fiction"

Roger: What's wrong with
men like me?
Eve: They don't believe in
marriage.
Roger: I've been married
twice.
Eve: See what I mean?

Roger Thornhill and Eve
Kendall — "North by North-
west"

Contents

List of Figures

Introduction

In this thesis we propose a fractional version of two well-know structural models to credit risk pricing: the Merton and Black and Cox models. A brief review of the Credit Risk models known in the literature and commonly used by practitioners is presented in chapter 1. The notion of Fractional Brownian Motion along with idea of long range dependence in finance is introduced. Also the empirical literature on credit spreads is briefly investigated.

In chapter 2 we develop a new theoretical framework for credit spread dynamics. Specifically, we propose a continuous-time credit risk model that is consistent with the observed behaviour of credit spread and can be used for pricing credit sensitive instruments, including credit derivatives. We assume that the value of the firm obeys to a Geometric Fractional Brownian Motion. Prices for the equity, the bond and credit spreads are derived and a sensitivity analysis is performed. The sensitivity analysis shows that the fractional versions of the two models above predict credit spreads which are consistent to the ones observed in the market. Moreover, to provide a justification for these models, an empirical analysis is carried out, which employs Moody's Long-Term Corporate Bond Yield Averages for the period December 1992–November 2003 and Lehman Brothers Eurodollar Indices for the period June 1996–July 2006. Long-range dependence (LRD) in the series of Treasury and corporate bond yields as well as credit spreads is thus investigated. In particular semiparametric methods (mainly periodogram-based techniques) are implemented and used in both a fractional integration and fractional cointegration analysis of the series involved. Yields and spreads are shown to be generally long memory nonstationary processes. This is substantially confirmed by the recent LM test proposed by Nielsen (2005). Although a number of empirical studies on credit spread dynamics have been carried out (see for instance Pedrosa and Roll, 1998; Prigent et al., 2001; Kiesel et al., 2001; Bierens et al., 2005), no research has been performed to date to investigate the long memory properties of credit spread. Our investigation aims at critically reassessing whether the dynamics of credit spread is consistent with the assumptions behind mainstream credit risk models, which are widely used by financial institutions.

Chapter 3 is about the most popular type of credit derivatives, i.e. Credit Default Swaps (CDS) and a new type which is becoming increasingly popular, i.e. Constant Maturity Credit Default Swaps (CMCDS). CDS represent one of the main reasons why credit markets have grown so fast over the past few years. They are contracts that provide insurance against the risk of a default by particular company. In such a contract, in exchange for protection, the buyer pays a premium through the life of the CDS or until default occurs. and in case of default of the reference issuer, she receives par and delivers an eligible defaulted obligation. The premium (called the CDS spread) in a CDS spread contract is determined by matching the discounted cash flows of a fixed leg paid by the protection buyer and a loss leg which corresponds to the net payment made by the protection seller to the protection buyer in case of default. CMCDS offer investors access to floating credit spread products and represent an efficient instrument for the investors to express views on the level and slope of credit curves. Similar to a CDS contract, by entering a CMCDS, the buyer pays a premium (spread) in exchange of protection. The essential difference between a CMCDS and CDS arises in the the payment leg: while in a CDS the spread is fixed, in a CMCDS the spread is floating. In particular the spread is equal to the prevailing reference CDS spread at each reset date times a factor known as the participation rate. As a consequence, in a CMCDS contract the loss leg is paired with a floating leg, where spread payments are indexed against a reference constant maturity CDS spread at each reset date. Another

important feature of these contracts is that investor who sell protection in a CMCDS contract are subject to a lower mark-to-market volatility than protection sellers in a CDS contract. It is interesting to note that default risk can be hedged away by constructing a portfolio short CMCDS protection and long CDS protection.

The main contribution of the chapter is the development of a set of tools that allow, for each company, the calculation of a proxy premium spread for CMCDS and the analysis of trading strategies based on CDS and CMCDS. In particular, the chapter aims to identify possible imbalances that may exist in the credit markets when pairing CDS and CMCDS on the same name. The general idea is to form a swap type of trading strategy whereby a fixed premium payment is netted against a floating one, both representing protection premia against default. Note that the aforementioned strategy completely eliminates credit risk (only counterparty risk is taken). A large database (source: Markit partners) of single-name CDS premia covering the period January 2001–November 2006 is used to produce the corresponding CMCDS prices using common market models.

Finally in chapter 4 we address an interesting issue which is related to the topic discussed in chapter 2. As pointed out by Dolado et al. (2005) and by Mikosch and Stărică (2004) for the case of multiple breaks, short-memory processes affected by shifts in trends or in the mean may display similar properties as those prescribed by LRD. In particular Mikosch and Stărică (2004) claim that long-range dependence in volatility may be explained by nonstationarities in the data (to be precise by changes in the parameters over different subsamples). The fact that for a time series the estimated fractional integration parameter $\hat{d}$ falls in the long memory interval, thus, does not necessarily imply Long Range Dependence: it is well known that the presence of structural breaks may cause a $I(0)$ process to look like an $I(d)$ one. Since we are dealing with credit spreads, our sample is very likely to include extraordinary financial and credit events that could result in one or more structural breaks. Recently some tests to detect if a time series is truly long memory have been proposed in the econometric literature. These include for instance the test based on temporal aggregation of Ohanissian et al. (2007). Further, we implement the test proposed by Dolado et al. (2005) of a process being $I(d)$, under the null, against the alternative of being $I(0)$ with deterministic components subject to structural breaks. The statistic they propose has been labeled as the Structural Break-Fractional Dickey-Fuller (SB-FDF) test, since it is based on the same principles behind the Dickey-Fuller unit root test. Finally the two test of Shimotsu (2006) are illustrated and implemented. One of these is based on sample splitting and the other on $\hat{d}$th differencing the data, where $\hat{d}$ is an estimate of the long memory parameter d. The tests are applied to both the datasets of credit indeces and credit spreads used for chapter 2. The conclusions for the Lehman Brothers is that that both yields and spreads are genuine long memory processes. As far as Moody's credit indeces are concerned, the yields are genuine long memory processes according to both the tests procedures. Further, the test based on dth differencing never rejects the null of true long memory also for the credit spreads series. The SB–FDF test, on the other hand, rejects the null at the 1% confidence level for four series of credit spreads out of nine. In every case the test detects a jump, the jump occurs in the second half of the sample. The same tests are also applied to the CDS data used in chapter 3. First the long memory parameter is estimated for the most liquid maturities and for each obligor using the first differences of log-CDS quotes. Then the tests of true vs. spurious long memory are run only for the issuers whose data displays the long memory property. As far as the 5-years contract is concern, looking only at the obligors whose log-CDS quotes display LRD, the rejection frequencies are between 20 and 28%.

Credit Risk Pricing: Review of the Literature

The rise of credit risk measurement and the credit derivatives market started in the early 1990s and has grown ever since: credit risk is today one of the most intensely studied topics in quantitative finance. The phenomenal growth of the credit markets, i.e. the enormous diffusion of derivative securities with credit risk, options and forward contracts subject to counterparty default risk, credit-risky bonds and credit derivatives, has generated a powerful array of new instruments for managing credit risk. The increasing popularity of credit derivatives is due to the fact that they allow market participants to easily trade and manage pure credit risk. In the academic literature models for the pricing of risky debt can be subdivided into two classes: firm's value and reduced form models.

The philosophy underlying firm's value models is to assume there is a fundamental process usually interpreted as the total value of the assets of the firm that has issued the bonds the researcher is interested in. A stochastic process for the evolution of the firms' underlying assets is assumed and this is the driving force behind the dynamics of the prices of all securities issued by the firm. The well-known *structural* approach due to the seminal paper of Merton (1974) assumes that the company has issued only shares and a zero-coupon bond. The firm defaults if the value of its assets is lower than the promised debt payment at maturity. As a consequence all claims on the firm's value are modeled as derivative securities with the firm's value as underlying. Merton's model has been extended for instance by Black and Cox (1976), Geske (1977), Shimko et al. (1993) and Leland (1994) to allow for more realistic assumptions, such as the possibility of default before maturity, coupon payments, stochastic interest rates, etc.

In the *reduced-form* approach the default process of risky debt is directly modeled making some assumptions on the evolution of the risk-free rate and the recovery rate in the event of default. Usually, the time of default is modeled directly as the time of the first jump of a Poisson process with random intensity. In particular Jarrow and Turnbull (1995) consider the simplest case where the default is driven by a Poisson process with constant intensity with known payoff at default, whereas Duffie and Singleton (1997) developed a similar model where the payoff in default is also cash, but denoted as a fraction of the value of the defaultable security just before default. Some papers developed within this framework use a rating-based approach in which default is attained through changes in credit ratings driven by a Markovian transition matrix. Among the others it is worth citing Das and Tufano (1996), Jarrow et al. (1997) and Duffee (1998).

1.1   Firm's Value Models

The first class of models views the firms' liabilities as contingent claims issued against the firm underlying assets, with the payoffs to all the firm liabilities in bankruptcy completely specified. Denote by V be the total value of the assets of the firm. A stochastic process for the evolution of the firm underlying assets and the conditions under which a default is triggered as well as the payoff of the risky debt in the event of default are specified. For instance it is possible to assume that V follows a geometric Brownian motion:

$$\mathrm{d}V_t = \mu V_t \mathrm{d}t + \sigma V_t \mathrm{d}B_t, \tag{1.1}$$

where B_t is a standard Brownian motion. Default can be triggered in two ways: a default occurs at maturity if V is insufficient to pay back the outstanding debt (in this way during the lifetime of the contract a default can not be triggered) or more realistically one can assume that a default is triggered as soon as the value of V falls below a barrier. Let us use the following standard notation: $B(t,T)$ and $\overline{B}(t,T)$ are the prices in t of default-free and defaultable zero coupon bond respectively, both with maturity T and S_t is the price in t of the firm's equity. The main intuition in Merton (1974) is to assume the value of the firm as the driving force behind the dynamics of the prices of all securities issued by the firm. Therefore both $\overline{B}$ and S are function of V and more generally all claims on the firm's value are evaluated as derivative securities with the firm's value as underlying. This model assumes that the company has a simple capital structure that consists of equity and a zero-coupon bond. Moreover the term structure of interest rate is deterministic and flat and the firm pays no dividend over the life of the debt. Suppose, for instance, that the firm has only issued zero coupon bonds with maturity T and total face value D. Indicate (for ease of notation) with S and $\overline{B}$ the prices of all shares issued by the firm and default may happen only at maturity. Thus one gets the following payoffs at time T:

$$\overline{B}(T,T) = \min(D, V_T) \tag{1.2}$$
$$S_T = \max(V_T - D, 0). \tag{1.3}$$

Equations (1.2) and (1.3) are straightforward: in case of default ($V < D$) one gets $\overline{B} = V$ and $S = 0$, otherwise $\overline{B} = D$ and $S = V - D$. Note that in both cases

$$\overline{B}(T,T) + S_T = V_T$$

and we can rewrite (1.2) as

$$\overline{B}(T,T) = D - \max(D - V_T, 0). \tag{1.4}$$

Equation (1.3) implies that the equity is an European call option with underlying V, maturity T and strike D, $c(V_t, D, T)$:

$$S_t = c(V_t, D, T), \tag{1.5}$$

and therefore, under model (1.1), the value of the share is simply given by the Black-Scholes formula, i.e.

$$S_t = V_t N(d_1) - D\mathrm{e}^{-r(T-t)} N(d_2), \tag{1.6}$$

where r is the interest rate $N(\cdot)$ is the Normal c.d.f. and

$$d_1 = \frac{\log(V_t/D) + (r + \frac{1}{2}\sigma^2)(T - t)}{\sigma\sqrt{T - t}}$$
$$d_2 = d_1 - \sigma\sqrt{T - t}.$$

Moreover from (1.4), the corporate debt is a given by a combination of D riskless bond and a short position in a put option on the firm's assets. Therefore its value is given by:

$$\overline{B}(t,T) = DB(t,T) - p(V_t, D, T), \tag{1.7}$$

or alternatively

$$\overline{B}(t,T) = V_t - c(V_t, D, T). \tag{1.8}$$

This means that under the assumptions of the model

$$\overline{B}(t,T) = V_t - S_t = D\mathrm{e}^{-r(T-t)} - \left[D\mathrm{e}^{-r(T-t)} N(-d_2) - V_t N(-d_1) \right]$$
$$= V_t N(-d_1) + D\mathrm{e}^{-r(T-t)} N(d_2).$$

Note that the risk neutral probability that the firm defaults equals the probability Q that the equityholders will not exercise their option to buy the assets of the firm for the strike D at time T. This probability is

$$Q(V_T < D) = 1 - N(d_2) = N(-d_2).$$

The yield to maturity for a corporate bond is defined implicitly from:

$$\overline{B}(t,T) = De^{-\overline{Y}(t,T)(T-t)}$$

and therefore the credit yield spread (defined as the difference between the yield to maturity of a corporate bond and the yield to maturity of a default-free bond, $Y(t,T)$) is

$$s(t,T) = \overline{Y}(t,T) - Y(t,T),$$

so that

$$\overline{B}(t,T) = DB(t,T)e^{-s(t,T)(T-t)}. \tag{1.9}$$

In Merton model

$$s(t,T) = \overline{Y}(t,T) - r$$

$$= -\frac{\log(\overline{B}(t,T)/D)}{T-t} - r = -\frac{\log(V_t/DN(-d_1) + e^{-r(T-t)}N(d_2))}{T-t} - r$$

$$= -\frac{1}{T-t}\log(V_t/De^{r(T-t)}N(-d_1) + N(d_2)).$$

Defining leverage by

$$\ell = \frac{D}{V_te^{r(T-t)}},$$

we can rewrite the credit spread arising from the model as

$$s(t,T) = -\frac{1}{T-t}\log\left(\frac{1}{\ell}N(h_1) + N(h_2)\right), \tag{1.10}$$

with

$$h_1 = -d_1 = \frac{\log\ell}{\sigma\sqrt{T-t}} - \frac{\sigma}{2}\sqrt{T-t}$$

$$h_2 = d_2 = -\frac{\log\ell}{\sigma\sqrt{T-t}} - \frac{\sigma}{2}\sqrt{T-t}.$$

The Merton model has some important implications. First of all, when the value of the firm is much bigger that the debt D, then the put option is deep out-of-the-money. This means that the probability of default is low and corporate debt trades as if it is default-free. Conversely, when $V_t \ll D$ then in (1.8) $c(V_t, D, T)$ is small and $\overline{B}(t,T)$ is approximately equal to the value of the equity. As a consequence the volatility of the corporate debt depends on the volatility of the underlying asset when the the put option trades deep in-the-money. Moreover, if the default-free interest rate increases, then the credit spread decreases. To see this from (1.9) and (1.7) derive

$$s(t,T) = -\frac{1}{T-t}\log\left(1 - \frac{p(V_t,D,T)}{DB(t,T)}\right).$$

Thus

$$\frac{\partial s(t,T)}{\partial r} = \left(\frac{\partial p(V_t,D,T)}{\partial r}\right)\left(\frac{\partial s(t,T)}{\partial p(V_t,D,T)}\right)$$

Since

$$\frac{\partial p(V_t, D, T)}{\partial r} = -D\mathrm{e}^{-r(T-t)}(T-t)N(-d_2),$$

$$\frac{\partial s(t,T)}{\partial p(V_t, D, T)} = \frac{1}{(T-t)(DB(t,T) - p(V_t, D, T))} = \left(\frac{1}{(T-t)\overline{B}(t,T)}\right)$$

it is possible to conclude

$$\frac{\partial s(t,T)}{\partial r} = -\frac{D}{\overline{B}(t,T)}N(-d_2) \leq 0.$$

When

$$B(t,T) = \mathrm{e}^{-r(T-t)}$$

it follows that

$$s(t,T) = -\frac{1}{T-t}\log\left(\frac{\overline{B}(t,T)}{D}\right) - r$$

$$= -\frac{1}{T-t}\log\left(\frac{V_t - S_t}{D}\right) - r$$

and thus

$$\frac{\partial s(t,T)}{\partial r} = \frac{1}{(T-t)}\left(\frac{1}{\overline{B}(t,T)}\right)\left(\frac{\partial S_t}{\partial r}\right) - 1$$

$$= \frac{1}{(T-t)}\left(\frac{\mathrm{e}^{[s(t,T)+r](T-t)}}{D}\right)\left(D(T-t)\mathrm{e}^{-r(T-t)}N(d_2)\right) - 1$$

$$= \mathrm{e}^{s(t,T)(T-t)}N(d_2) - 1 = \frac{-V_t N(-d_1)}{\overline{B}(t,T)}.$$

The last equality follows from

$$\mathrm{e}^{s(t,T)(T-t)}N(d_2) = \frac{DB(t,T)N(d_2)}{\overline{B}(t,T)} = \frac{\overline{B}(t,T) - V_t N(-d_1)}{\overline{B}(t,T)}$$

and

$$\overline{B}(t,T) = V_t N(-d_1) + DB(t,T)N(d_2).$$

However there is an intuitive explanation: an increase in the default-free spot interest rate, keeping the value of the firm constant, makes the probability of default decline. This, in turn, makes the corporate bond price increase and, consequently, the spread decline. Besides, increases in volatility raise credit spreads because

$$\frac{\partial s(t,T)}{\partial \sigma} = \frac{1}{(T-t)DB(t,T)}\frac{\partial p(V_t, D, T)}{\partial \sigma} = \frac{1}{\sqrt{T-t}DB(t,T)}N'(d_1) > 0.$$

Another implication of the model is that the credit spread tends to zero as the maturity of the zero-coupon bond tends to zero, at least when $V > D$. In fact, first take (1.10) and set $t = 0$. Note that

$$\begin{cases} \lim_{T\to 0} N(h_1) = 1 - \lim_{T\to 0} N(h_2) = 0 & \text{if } D < V \\ \lim_{T\to 0} N(h_1) = 1 - \lim_{T\to 0} N(h_2) = 1 & \text{if } D > V \\ \lim_{T\to 0} N(h_1) = \lim_{T\to 0} N(h_2) = \frac{1}{2} & \text{if } D = V \end{cases}.$$

Therefore for $D \geq V$

$$\lim_{T\to 0} s(T) = +\infty,$$

Figure 1.1: Credit spreads against maturity when $\sigma = 0.25$, $D/V = 2/3$, $r = 0.05$.
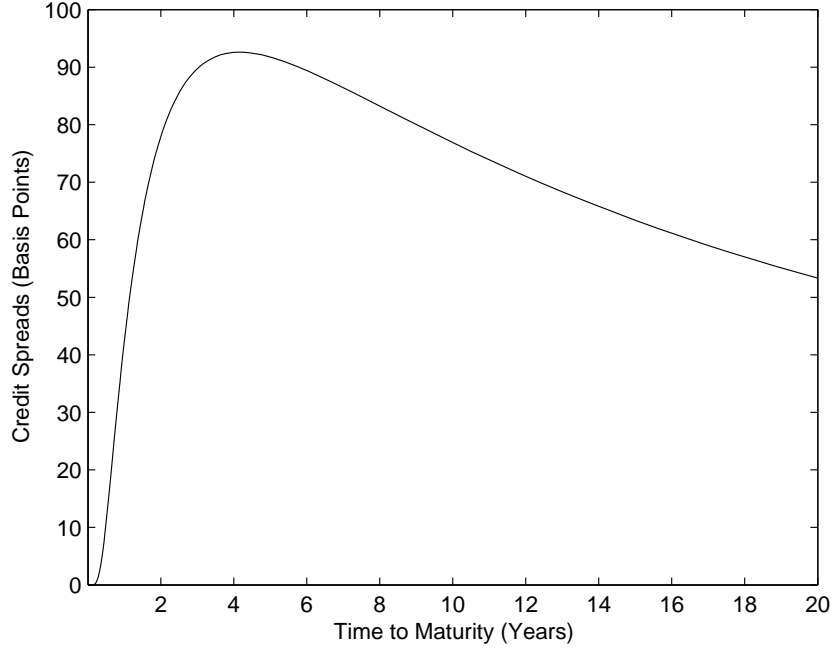
whereas for $D < V$

$$\lim_{T \to 0} s(T) \overset{Hop.}{=} -\lim_{T \to 0} \frac{\left(\frac{\partial}{\partial T} h_2\right) N'(h_2) + \frac{1}{\ell}\left(\frac{\partial}{\partial T} h_1\right) N'(h_1) - \frac{\left(\frac{\partial}{\partial T} \ell\right)}{\ell^2} N(h_1)}{N(h_2) + \frac{1}{\ell} N(h_1)}$$

$$= -\lim_{T \to 0} \frac{\left(\frac{\partial}{\partial T} h_2\right) \ell N'(h_2) + \left(\frac{\partial}{\partial T} h_1\right) N'(h_1) + r N(h_1)}{\ell N(h_2) + N(h_1)}$$

$$= -\lim_{T \to 0} \frac{\left[\frac{\partial}{\partial T}(h_1 + h_2)\right] N'(h_1) + r N(h_1)}{\ell N(h_2) + N(h_1)}$$

$$= -\lim_{T \to 0} \frac{-\frac{\sigma}{2} T^{-1/2} N'(h_1) + r N(h_1)}{\ell N(h_2) + N(h_1)} = 0.$$

Note that we used

$$\ell N'(h_2) = N'(h_1).$$

As a consequence, short maturity credit spreads are near zero in Merton model when $D < V$. From

$$\frac{\partial}{\partial T} s(T) = -\frac{1}{T}\left[s(T) + \frac{-\frac{\sigma}{2} T^{-1/2} N'(h_1) + r N(h_1)}{\ell N(h_2) + N(h_1)}\right]$$

it is clear that credit spreads rise with maturity for low values of leverage. The explanation is straightforward: the bigger the leverage, the bigger the probability of default. Clearly, an increase in the probability of default results in a decrease of the corporate bond price and therefore in an increase in the credit spread. For values of leverage below unity credit spreads, as functions of time to maturity, are hump shaped, rising at first and then falling (see Figure 1.1). For leverage levels above unity credit spreads tend to explode near zero and fall with maturity (see Figure 1.2). A big shortcoming of the model are the inconsistencies in the treatment of debt with multiple maturities: they default at different times, each at its own maturity, while in reality they

Figure 1.2: Credit spreads against maturity when $\sigma = 0.25$, $D/V = 2$, $r = 0.05$.
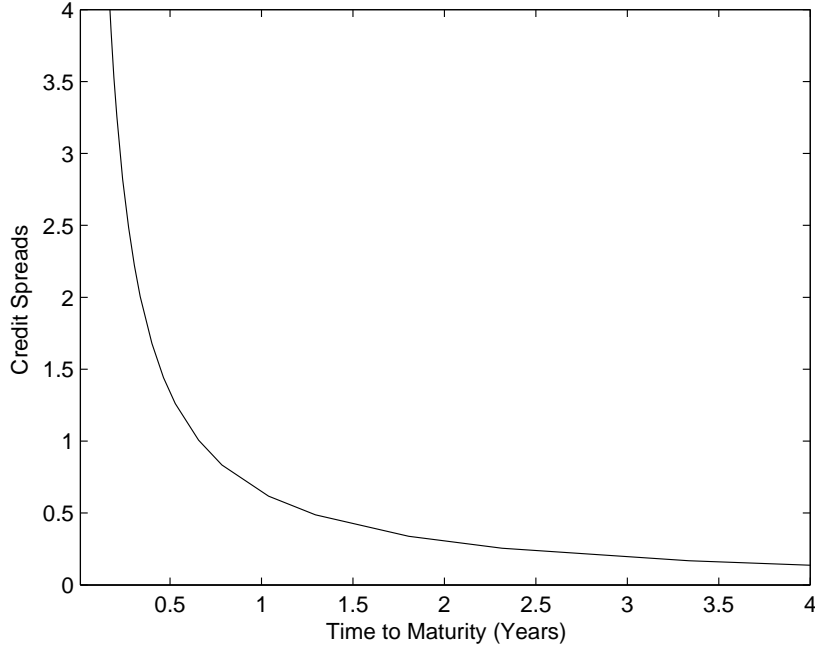
default all together. Moreover, the assumption of constant interest rates and a flat yield curve are contrary to realistic term structures of interest rates. Defaultable bonds must have an embedded exposure to interest rates that is here modeled unrealistically in terms of movements in a flat yield curve. It is also possible to verify that credit spreads rise with leverage. In fact

$$\frac{\partial}{\partial \ell} s(\ell) = -\frac{1}{T} \frac{\frac{\partial}{\partial \ell} \left[N(h_2) + \frac{1}{\ell} N(h_1) \right]}{N(h_2) + \frac{1}{\ell} N(h_1)}$$

$$= \frac{\ell N'(h_2) + N(h_1) - N'(h_1)}{T \left[\ell^2 N(h_2) + \ell N(h_1) \right]} = \frac{N(h_1)}{T \left[\ell^2 N(h_2) + \ell N(h_1) \right]}$$

$$= \frac{N(h_1) \mathrm{e}^{s(\ell)T}}{T \ell^2} > 0.$$

Another unrealistic feature of the model is that it implies risk premia[1] bigger than those one can find on the market. Consider, for instance a 2 year maturity bond with $D/V = 2/3$, volatility of the asset $\sigma = 0.2$,

---

[1] If a claim which pays 1 if the event X happens is given, then the expected cash flow under the market probability is $P(X)$. The price of the claim can be derived by the risk neutral probability Q:

$$\frac{Q(X)}{1 + r}.$$

Define the gross rate of return on the claim, $1 + R(X)$, as the ratio of the expected cash flow to the price:

$$1 + R(X) = \frac{P(X)}{Q(X)}(1 + r).$$

The excess return

$$\frac{P(X)}{Q(X)} - 1 = \frac{R(X) - r}{1 + r}$$

is the risk premium.

$r = 0.1$ and equity risk premium of 10%. The expected rate of return is $\mu = r + 0.1 = 0.2$. Recall now that the risk neutral probability of no default is $N(d_2)$. Thus we need to evaluate the ratio

$$\frac{N(d_{2,\mu})}{N(d_2)} - 1 = 0.0427 = 427 \quad \text{basis points,}$$

where $d_{2,\mu}$ is obtained from the expression for d_2 replacing r by μ. Realistic values for the risk premium would be below 100 basis points.

Finally the model reflects the interaction of market and credit risk, i.e. the fact that market and credit risk are not separable. To see this, suppose that the value of the firm's assets unexpectedly decreases, giving rise to market risk. The decrease in the asset's value increases the probability of default, giving rise to credit risk. The converse is also true.

In practice the model is calibrated by estimating the current value of the company's assets and the volatility of the assets from the market value of the company's equity and the equity's instantaneous volatility (see Jones et al., 1984). Recently Hull et al. (2004) proposed to estimate the model's parameters from the implied volatilities of options on the company's equity.

As already mentioned, a more realistic assumption can be made about the conditions under which default is triggered. For instance the company could be assumed to default the first time its assets fall below some default barrier, say $L < V_t$. This is the case of the first passage time model proposed by Black and Cox $(1976)^2$. In this framework the equity can be viewed as a down-and-out call option. If the assets hit the barrier L before time T, the option ceases to exist and the bondholders receive the assets' value or some recovered portion of it. If the barrier is not hit before T, then the equity at maturity is, as in Merton's model, the payoff on a European call option. In this way the model implies strict absolute priority. Define

$$\tau = \min\{k \geq t : V_k < L\}$$

Denoting by $\mathbb{E}_t^Q[\cdot]$ the expectation under the risk neutral probability Q conditional to the information available up to time t and by I the indicator function, the equity value, under the assumption of a geometric Brownian motion (1.1) for V, is:

$$S_t = \mathbb{E}_t^Q \left[\mathrm{e}^{-r(T-t)}(V_T - D)^+ I_{\{\tau > T\}} \right]$$
$$= \left[V_t N(x_1) - D\mathrm{e}^{-r(T-t)} N(x_1 - \sigma\sqrt{T-t}) \right]$$
$$- \left[V_t \left(\frac{L}{V_t}\right)^{2\theta} N(y_1) - D\mathrm{e}^{-r(T-t)} \left(\frac{L}{V_t}\right)^{2\theta-2} N(y_1 - \sigma\sqrt{T-t}) \right],$$

where

$$\theta = \frac{r + \frac{1}{2}\sigma^2}{\sigma^2}$$

$$x_1 = \begin{cases} \frac{\log(V_t/D)+(r+\frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}} & \text{if } D \geq L \\ \frac{\log(V_t/L)+(r+\frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}} & \text{if } D < L \end{cases},$$

$$y_1 = \begin{cases} \frac{\log(L^2/(V_tD))+(r+\frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}} & \text{if } D \geq L \\ \frac{\log(L/V_t)+(r+\frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}} & \text{if } D < L \end{cases}.$$

---

^{2}Actually in the paper the barrier takes an exponential form $Le^{-\gamma(T-t)}$, where $T-t$ is the time to maturity. For simplicity we consider the constant barrier case $\gamma = 0$. Alternatively it could be reasonable to assume that a default is triggered as soon as the firm's value is worth less than a similar but default-free investment, i.e. when

$$V_\tau < L \times B(\tau, T).$$

Denoting by

$$m_t^T = \min_{t \leq s \leq T} V_s,$$

the running minimum of firm values, and by

$$\bar{r} = r - \frac{1}{2}\sigma^2$$

the risk neutral probability of default is given by

$$1 - Q(\tau > T, V_T > D)$$
$$= 1 - Q\left[m_t^T > L, V_T > D\right]$$
$$= 1 - Q\left[\min_{t \leq s \leq T}(\bar{r}(T-s) + \sigma B_{T-s}) > \log(L/V_t), \bar{r}(T-t) + \sigma B_{T-t} > \log(D/V_t)\right].$$

Using the joint distribution of an arithmetic Brownian and its running minimum[3] , we get

$$1 - Q(\tau > T, V_T > D)$$
$$= N\left(\frac{\log(D/V_t) - \bar{r}(T-t)}{\sigma\sqrt{T-t}}\right) + \left(\frac{L}{V_t}\right)^{\frac{2\bar{r}}{\sigma^2}} N\left(\frac{\log(L^2/(DV_t)) + \bar{r}(T-t)}{\sigma\sqrt{T-t}}\right).$$

Clearly this default probability is higher than the corresponding probability in the Merton model, which is obtained as the special case $L = 0$. Note that the barrier option value is, unlike the vanilla call value, not monotone in firm volatility σ. It follows that equity investors do not always benefit from an increase in asset volatility, as in the classical Merton model. Note also that, when $L \geq D$, bond investors are fully protected: they receive at least the face value D upon default and the bond is not subject to default risk any more. The corresponding payoff to bond investors at maturity is

$$\min(V_T, D)I_{\{\tau > T\}} + V_T I_{\{\tau \leq T\}}$$
$$= D - (D - V_T)^+ + (V_T - D)^+ I_{\{\tau \leq T\}}.$$

This position is equivalent to a portfolio composed of a riskfree loan with face value D maturing at T, a short European put on the firm with strike D and maturity T and a long European down-and-in call on the firm with strike D and maturity T. Clearly it holds $V_T = S_T + \overline{B}(T,T)$. The price of the bond can be derived using the corresponding Black-Scholes formulae. Suppose now that $1 - q = y \in [0,1]$ is the fraction lost in case of default, i.e. if default is triggered the bondholders receive qV_τ due to bankruptcy costs. This implies that the price in t of the corporate bond is given by the following formula

$$\overline{B}(t,T) = \mathbb{E}_t^Q\left[e^{-r(\tau-t)}qV_\tau I_{\{\tau \leq T\}} + \min(V_T, D)e^{-r(T-t)}I_{\{\tau > T\}}\right]. \tag{1.11}$$

Deviations from priority can be taken into account with a different payoff distribution at $V = L$. For instance we could assume[4]

$$S_\tau = yV_\tau = yL,$$

---

[3] Note that using the reflection principle we have

$$Q(\tau > T) = Q(m_t^T > L)$$
$$= Q\left[\min_{t \leq s \leq T}(\bar{r}(T-s) + \sigma B_{T-s}) > \log(L/V_t)\right]$$
$$= N\left(\frac{\log(L/(V_t)) - \bar{r}(T-t)}{\sigma\sqrt{T-t}}\right) + \left(\frac{L}{V_t}\right)^{\frac{2\bar{r}}{\sigma^2}} N\left(\frac{\log(L/V_t) + \bar{r}(T-t)}{\sigma\sqrt{T-t}}\right).$$

[4]Again, if the default is triggered as soon as V_τ hits the barrier $L \times B(\tau,T)$, we could impose

$$S_\tau = y \times L \times B(\tau,T)$$
$$\overline{B}(\tau,T) = q \times L \times B(\tau,T).$$

i.e. we could assume a deviation from absolute priority by yL in favour of the shareholders. Note that equation (1.11) can be adapted to every defaultable claim paying the amount X at maturity T and X' in case of default before maturity. Taking into account the possibility of a stochastic risk-free interest rate, one gets for the price in t, U_t of this derivative

$$U_t = \mathbb{E}_t^Q \left[\exp\left(-\int_t^\tau r_u \mathrm{d}u \right) X' I_{\{\tau \leq T\}} + \exp\left(-\int_t^T r_u \mathrm{d}u \right) X I_{\{\tau > T\}} \right]. \tag{1.12}$$

If the amount X' is paid at maturity T in case of a default prior to maturity it is apparent that

$$U_t = \mathbb{E}_t^Q \left[\exp\left(-\int_t^T r_u \mathrm{d}u \right) \left(X' I_{\{\tau \leq T\}} + X I_{\{\tau > T\}} \right) \right]. \tag{1.13}$$

To solve the inconsistencies in the treatment of debt with multiple maturities of Merton model in Longstaff and Schwartz (1995) default occurs at the first time that asset values reach a threshold level and at this time all maturities still outstanding default simultaneously. The problem of no short maturity credit spreads is eliminated by taking into account the possibility of early default which usually yields higher credit spreads. In particular it is assumed that interest rates are stochastic and follow a Vasicek process with movements that are correlated with the firm's assets:

$$\mathrm{d}V_t = r_t V_t \mathrm{d}t + \sigma_V V_t \mathrm{d}B_V(t)$$
$$\mathrm{d}r_t = \alpha(\theta - r_t)\mathrm{d}t + \sigma_r \mathrm{d}B_r(t)$$
$$\mathrm{d}B_V(t)\mathrm{d}B_r(t) = \rho \mathrm{d}t.$$

If during the life of the debt V reaches the default threshold, creditors receive a fraction q of the face value D at maturity. Assuming also that the value of the firm is independent of the capital structure of the firm, which implies that changes in capital structure (such as payments of coupon and principal) have no effect on V, closed form solution for both fixed and floating rate debt are derived.

Merton's model has been developed among the others by Collin-Dufresne and Goldstein (2001) who try to overcome the problems of many structural models which preclude the firm issuing the debt from altering its capital structure. It is well known, however, that firms adjust outstanding debt levels in response to changes in firm value. The authors take into account this fact and propose a framework in which leverage ratios are mean reverting. They assume that the dynamics for the firm value under the risk neutral measure is

$$\mathrm{d}V_t = (r - \zeta)V_t \mathrm{d}t + \sigma V_t \mathrm{d}B_t, \tag{1.14}$$

where ζ is the payout rate, i.e. the constant fraction of firm value paid out to security holders. The latter quantity is assumed not to be affected by changes in leverage. Applying Ito's Lemma the dynamics for $z = \log V$ can be easily derived:

$$\mathrm{d}z_t = \left(r - \zeta - \frac{\sigma^2}{2} \right) \mathrm{d}t + \sigma \mathrm{d}B_t.$$

As in Black and Cox (1976) default is assumed to be triggered the first time firm value reaches some exogenously specified threshold. The dynamics for the log-default threshold, k_t is as follows:

$$\mathrm{d}k_t = \lambda(z_t - \nu - k_t)\mathrm{d}t. \tag{1.15}$$

This model implies that every time k_t is less than $(z_t - \nu)$ the firm wishes to increase k_t and vice-versa. In other words, the firm tends to issue debt when their leverage ratio falls below some target and is more hesitant to modify maturing debt when the leverage is above that target. Note that the solution of (1.15) is

$$k_t = [k_0 + \nu]\mathrm{e}^{-\lambda t} + \lambda \int_0^t \mathrm{e}^{-\lambda(t-s)} z_s \mathrm{d}s - \nu =: Q_t - \nu$$

and thus

$$D_t = e^{k_t} = e^{Q_t - \nu}$$

meaning that the book value of debt depends on the average of past log-firm values. The log-leverage is $\mathcal{L}_t = k_t - z_t$ and its dynamics is

$$d\mathcal{L}_t = dk_t - dz_t = \lambda \left(\overline{\mathcal{L}} - \mathcal{L}_t \right) dt - \sigma dB_t$$

where

$$\overline{\mathcal{L}} = \frac{-r + \zeta + \frac{\sigma^2}{2}}{\lambda} - \nu.$$

Since the solution of this SDE is

$$\mathcal{L}_t = e^{-\lambda t} \mathcal{L}_0 + \overline{\mathcal{L}}[1 - e^{-\lambda t}] - \sigma \int_0^t e^{-\lambda(t-u)} dB_u,$$

we get

$$\mathbb{E}^Q \left[\mathcal{L}_t | \mathcal{L}_0 \right] =: M(t) = e^{-\lambda t} \mathcal{L}_0 + \overline{\mathcal{L}}[1 - e^{-\lambda t}]$$

$$\text{Var}^Q \left[\mathcal{L}_t | \mathcal{L}_0 \right] =: S^2(t) = \frac{\sigma^2}{2\lambda}[1 - e^{-2\lambda t}].$$

Starting from $\mathcal{L}_s$ we have

$$\mathcal{L}_t = e^{-\lambda(t-s)} \mathcal{L}_s + \overline{\mathcal{L}}[1 - e^{-\lambda(t-s)}] - \sigma \int_s^t e^{-\lambda(t-u)} dB_u,$$

and thus

$$\text{Var}^Q \left[\mathcal{L}_t | \mathcal{L}_s \right] =: S^2(t-s) = \frac{\sigma^2}{2\lambda}[1 - e^{-2\lambda(t-s)}].$$

When $\mathcal{L}_s = 0$

$$\mathbb{E}^Q \left[\mathcal{L}_t | \mathcal{L}_s = 0 \right] =: L(t-s) = \overline{\mathcal{L}}[1 - e^{-\lambda(t-s)}].$$

Default is triggered at the random time τ when $\mathcal{L}_t$ reaches zero for the first time. In event of default the bondholders receive only the fraction $1 - y$ of the face value. Assuming constant interest rate r, equation (1.13) entails the following time zero price for a defaultable zero coupon bond maturing at time t:

$$\overline{B}(0,t) = e^{-rt} \mathbb{E}^Q \left[I_{\{\tau > t\}} + (1-y)I_{\{\tau < t\}} \right] = e^{-rt} \left[1 - yQ(\tau < t) \right]. \tag{1.16}$$

To get the risk neutral probability of default before maturity, note that $\mathcal{L}$ is a Markov process with transition density denoted as $\pi(\mathcal{L}_t, t | \mathcal{L}_s, s)$. Define $g\left(\mathcal{L}_s = \underline{\mathcal{L}}, s | \mathcal{L}_0, 0 \right)$ as the probability density that the first passage time through the boundary $\underline{\mathcal{L}}$ occurs at time s. It follows

$$\pi(\mathcal{L}_t, t | \mathcal{L}_0, 0) = \int_0^t \pi(\mathcal{L}_t, t; \tau = s | \mathcal{L}_0, 0) ds + \pi(\mathcal{L}_t, t; \tau > t | \mathcal{L}_0, 0)$$

$$= \int_0^t \pi(\mathcal{L}_t, t | \tau = s; \mathcal{L}_0, 0) \pi(\tau = s; | \mathcal{L}_0, 0) ds + \pi(\mathcal{L}_t, t; \tau > t | \mathcal{L}_0, 0)$$

$$= \int_0^t \pi(\mathcal{L}_t, t | \mathcal{L}_s = \underline{\mathcal{L}}, s) g\left(\mathcal{L}_s = \underline{\mathcal{L}}, s | \mathcal{L}_0, 0 \right) ds + \pi(\mathcal{L}_t, t; \tau > t | \mathcal{L}_0, 0).$$

For $\mathcal{L}_t > \underline{\mathcal{L}} > \mathcal{L}_0$ the last term vanishes and therefore

$$\pi(\mathcal{L}_t, t | \mathcal{L}_0, 0) = \int_0^t \pi(\mathcal{L}_t, t | \mathcal{L}_s = \underline{\mathcal{L}}, s) g\left(\mathcal{L}_s = \underline{\mathcal{L}}, s | \mathcal{L}_0, 0 \right) ds. \tag{1.17}$$

This is because the the a continuous process starting below the boundary ($\mathcal{L}_0 < \underline{\mathcal{L}}$) and ending up above the boundary ($\mathcal{L}_t > \underline{\mathcal{L}}$) at some intermediate time s must touch the boundary $\underline{\mathcal{L}}$ for the first time.

Integrating both sides of (1.17) by $\int_0^\infty \mathrm{d}\mathcal{L}_t$ and setting the default boundary $\underline{\mathcal{L}} = 0$ one gets

$$Q(\mathcal{L}_t > 0 | \mathcal{L}_0) = \int_0^t g\left(\mathcal{L}_s = \underline{\mathcal{L}}, s \middle| \mathcal{L}_0, 0\right) \mathrm{d}s \; Q(\mathcal{L}_t > 0 | \mathcal{L}_s = \underline{\mathcal{L}} = 0),$$

and therefore

$$N\left(\frac{M(t)}{S(t)}\right) = \int_0^t g\left(\mathcal{L}_s = \underline{\mathcal{L}}, s \middle| \mathcal{L}_0, 0\right) \mathrm{d}s \; N\left(\frac{L(t-s)}{S(t-s)}\right). \tag{1.18}$$

To get a solution for the first passage density, the authors discretize time into n equal intervals of length $\Delta t = t/n$. The integral on the right hand side of (1.18) is approximated by estimating values at the midpoint of the interval. In particular defining

$$a_i = \frac{M(i\Delta t)}{S(i\Delta t)},$$

$$b_i = \frac{L(i\Delta t)}{S(i\Delta t)}$$

the first two terms from equation (1.18) can be approximated as

$$N(a_1) = \Delta t \; g\left(\mathcal{L}_{\Delta t/2} = \underline{\mathcal{L}}, \Delta t/2 \middle| \mathcal{L}_0, 0\right) N(b_{1/2})$$

$$N(a_2) = \Delta t \; g\left(\mathcal{L}_{\Delta t/2} = \underline{\mathcal{L}}, \Delta t/2 \middle| \mathcal{L}_0, 0\right) N(b_{3/2})$$

$$+ \Delta t \; g\left(\mathcal{L}_{3\Delta t/2} = \underline{\mathcal{L}}, 3\Delta t/2 \middle| \mathcal{L}_0, 0\right) N(b_{1/2}).$$

Continuing in this manner one has n equations in the n unknowns

$$g\left(\mathcal{L}_{(i-1/2)\Delta t} = \underline{\mathcal{L}}, (i-1/2)\Delta t \middle| \mathcal{L}_0, 0\right). \quad i = 1, \ldots, n$$

Going back to equation (1.16), if one defines

$$t_j := \frac{jt}{n} = j\Delta t,$$

it is apparent that

$$Q(\tau < t) =: Q(\mathcal{L}_0, t_j) = \sum_{i=1}^{j} q_i$$

where

$$q_i := \Delta t \; g\left(\mathcal{L}_{(i-1/2)\Delta t} = \underline{\mathcal{L}}, (i-1/2)\Delta t \middle| \mathcal{L}_0, 0\right).$$

Solving the linear system above one gets

$$q_1 = \frac{N(a_1)}{N(b_{1/2})}$$

$$q_i = \frac{N(a_i) - \sum_{k=1}^{i-1} q_k N(b_{i-k+1/2})}{N(b_{1/2})} \quad i = 2, 3, \ldots, n.$$

Since a coupon bond can be thought as a portfolio of discount bonds, the evaluation is straightforward. Consider a coupon bond with promised coupon payment at dates t_j, $j = 1, \ldots, N$ and face value one to be paid in $t_N = T$. When the interest rate is the constant r the time zero value of the bond is

$$\overline{CB}(0, T) = \sum_{j=1}^{N} C e^{-rt_j} \mathbb{E}^Q \left[I_{\{\tau > t_j\}} + (1 - y_{coup}) I_{\{\tau < t_j\}} \right]$$

$$+ e^{-rT} \mathbb{E}^Q \left[I_{\{\tau > T\}} + (1 - y) I_{\{\tau < T\}} \right]$$

$$= \sum_{j=1}^{N} C e^{-rt_j} [1 - y_{coup} Q(\mathcal{L}_0, t_j)] + e^{-rT} [1 - y Q(\mathcal{L}_0, T)].$$

The analysis is extended to the case of stochastic interest rate. The process for z is

$$\mathrm{d}z_t = \left(r_t - \zeta - \frac{\sigma^2}{2}\right)\mathrm{d}t + \sigma\mathrm{d}B_1(t).$$

and a Vasicek model for the interest rate is assumed:

$$\mathrm{d}r_t = \alpha(\theta - r_t)\mathrm{d}t + \eta\mathrm{d}B_2(t).$$

Besides it is assumed that the log-default threshold follows the process

$$\mathrm{d}k_t = \lambda[z_t - \nu - \phi(r_t - \theta) - k_t]\mathrm{d}t$$

where ϕ is a non-negative parameter. Thus the drift of the log-default threshold is a decreasing function of the spot rate. This is because debt issuances are likely to drop during high interest rate period. A closed form solution for the price of the discount bond is derived as

$$\overline{B}(0,T) = \mathbb{E}^Q\left[\mathrm{e}^{-\int_0^T r_s\mathrm{d}s}(1 - yI_{\{\tau<T\}})\right] = B(0,T)\left[1 - yQ^T(r_0, \mathcal{L}_0, T)\right],$$

where $B(0,T)$ is the price[5] of a risk free bond:

$$B(0,T) = \mathrm{e}^{A_\alpha^{(T)} - r_0 B_\alpha^{(T)}},$$

with

$$A_\alpha^{(T)} = \left[\frac{\eta^2}{2\alpha^2} - \theta\right]T + \left[\frac{\theta}{\alpha} - \frac{\eta^2}{\alpha^3}\right][1 - \mathrm{e}^{-\alpha T}] + \frac{\eta^2}{4\alpha^3}[1 - \mathrm{e}^{-2\alpha T}]$$

and

$$B_\alpha^{(T)} = \frac{1}{\alpha}(1 - \mathrm{e}^{-\alpha T})$$

The expression for $Q^T(r_0, \mathcal{L}_0, T)$ is derived again by the means of the integral equation satisfied by the first hitting time density.

The same idea is used in Leland and Toft (1996). They assume the dynamics (1.14) for the firm's asset and that default is triggered when V falls to the value constant L. The time zero price of a bond which continuously pays the constant coupon flow $c(t)$, has principal $p(t)$ and pays the fraction $q(t)$ of the asset value L in case of default is

$$\overline{B}(0,t) = c(t)\int_0^t \mathrm{e}^{-rs}[1 - G(V_s = L, s)]\mathrm{d}s + \mathrm{e}^{-rs}p(t)[1 - G(V_s = L, s)]$$

$$+ q(t)L\int_0^t \mathrm{e}^{-rs}g(V_s = L, s)\mathrm{d}s \quad (1.19)$$

---

[5] If the model for the interest rate is specified as

$$\mathrm{d}r_t = (\alpha(t) - \beta(t)r_t)\mathrm{d}t + \eta(t)\mathrm{d}B_t$$

with $\alpha(t)$, $\beta(t)$ and $\eta(t)$ deterministic functions, then the price of the bond is

$$B(0,T) = \mathrm{e}^{A_K^{(T)} - r_0 B_K^{(T)}},$$

where

$$A_K^{(T)} = \frac{1}{2}\int_0^T \mathrm{e}^{2K(v)}\eta^2(v)\left(\int_v^T \mathrm{e}^{-K(y)}\mathrm{d}y\right)^2\mathrm{d}v - \int_0^T\int_0^t \mathrm{e}^{-K(t)+K(u)}\alpha(u)\mathrm{d}u\mathrm{d}t$$

$$B_K^{(T)} = \int_0^T \mathrm{e}^{-K(t)}\mathrm{d}t$$

and

$$K(t) = \int_0^t \beta(u)\mathrm{d}u.$$

where $g(V_s = L, s)$ is again the density of the first passage for V to the boundary L and $G(V_s = L, s)$ is the cumulative distribution function of the first passage time to bankruptcy. The first term in the right hand side of (1.19) is the discounted expected value of the coupon flow, paid at time s with probability $1 - G(s)$, the second term is the discounted expected value of the principal repayment and the third term represents the discounted expected value paid in case of bankruptcy. Integrating the first term in the right hand side (1.19) by parts yields

$$\overline{B}(0,t) = \frac{c(t)}{r} + \mathrm{e}^{-rt}\left[p(t) - \frac{c(t)}{r}\right][1 - G(t)] + \left[q(t)L - \frac{c(t)}{r}\right]H(t),$$

where

$$H(t) = \int_0^t \mathrm{e}^{-rs} g(V_s = L, s)\mathrm{d}s.$$

Formulae for $G(t)$ and $H(t)$ are provided:

$$G(t) = N[h_-(t)] + \mathrm{e}^{-2ab}N[h_+(t)]$$
$$H(t) = \mathrm{e}^{(z-a)b}N[q_-(t)] + \mathrm{e}^{-(z+a)b}N[q_+(t)]$$

where

$$q_\pm(t) = \frac{-b \pm z\sigma^2 t}{\sigma\sqrt{t}}; \qquad\qquad h_\pm(t) = \frac{-b \pm a\sigma^2 t}{\sigma\sqrt{t}};$$
$$a = \frac{r - \zeta}{\sigma^2} - \frac{1}{2}; \qquad\qquad b = \log\left(\frac{V_0}{L}\right);$$

and

$$z = \left(a^2 + \frac{2r}{\sigma^2}\right)^{1/2}.$$

Next, the firm is assumed to continuously issue new debt that it will redeem at par at maturity T if not in bankruptcy. New bond principal is issued at a rate $p = P/T$ per year where P is the total principal value of all outstanding bonds. As long as the firm remains solvent the total outstanding debt will be P. The coupons are paid at a rate of $c = C/T$ per year, implying that the total coupon paid by all the outstanding bonds is C per year. The fraction of firm asset value lost in bankruptcy is y and the remaining value, $(1 - y)L$ is distributed to bondholders. Equal seniority for all outstanding debt is assumed and therefore the recovery rate per year is assumed to be $q(t) = q/T$, implying $q = 1 - y$. Thus the value of all outstanding bond is

$$D(V; L, T) = \int_0^T \overline{B}(0,t)\mathrm{d}t$$
$$= \int_0^T \left[\frac{C}{Tr} + \mathrm{e}^{-rt}\left(\frac{P}{T} - \frac{C}{Tr}\right)(1 - G(t)) + \left(\frac{1-y}{T}L - \frac{C}{Tr}\right)H(t)\right]\mathrm{d}t$$
$$= \frac{C}{r} + \left(P - \frac{C}{r}\right)\left[\frac{1 - \mathrm{e}^{-rT}}{rT} - I(T)\right] + \left[(1-y)L - \frac{C}{r}\right]J(T),$$

where

$$I(T) = \frac{1}{T}\int_0^T \mathrm{e}^{-rt}G(t)\mathrm{d}t \qquad\qquad\qquad (1.20\text{a})$$

$$J(T) = \frac{1}{T}\int_0^T H(t)\mathrm{d}t. \qquad\qquad\qquad (1.20\text{b})$$

The integral (1.20a) can be solved by parts to get

$$I(T) = \frac{1}{T} \left[-\frac{1}{r}\mathrm{e}^{-rt}G(t) \Big|_0^T + \frac{1}{r} \int_0^T \mathrm{e}^{-rt}g(t)\mathrm{d}t \right]$$

$$= \frac{1}{rT} \left[-\mathrm{e}^{-rT}G(T) + \int_0^T \mathrm{e}^{-rt}g(t)\mathrm{d}t \right]$$

$$= \frac{1}{rT} \left[H(T) - \mathrm{e}^{-rT}G(T) \right].$$

The integral in (1.20b) is not so straightforward but the authors show that

$$J(T) = \frac{1}{z\sigma\sqrt{T}} \left(-q_-(T)\mathrm{e}^{(z-a)b}N[q_-(T)] + q_+(T)\mathrm{e}^{-(z+a)b}N[q_+(T)] \right).$$

The total value of the firm, v, is assumed to be the sum of the asset value and the value of tax benefit, less the value of bankruptcy costs. Thus the value of equity is

$$S(V; L, T) = v(V; L) - D(V; L, T).$$

The level L is determined endogenously using the equation

$$\frac{\partial S(V; L, T)}{\partial V} \bigg|_{V=L} = 0$$

This condition maximises with respect to L both the value of the equity and the value of the firm subject to the condition

$$S(V) \geq 0 \quad \text{for all } V \geq L$$

and provide a closed form solution for the bankruptcy-triggering level L, which turns out to be independent of time.

Firm's value models are suitable if there exists a strong relationship between the prices of the different securities issued by the firm. In this case these models perform very well when one wants to answer question that arise from corporate finance, like the analysis of the relative powers of shareholders and creditors. The weaknesses of these models are to be found in the difficulty to find a meaningful process for the firm's value. which is often not tradeable and therefore not observable. In fact it is rarely possible to know the market value of the firm's assets because the typical firm has numerous complex debt contracts outstanding traded on an infrequent basis. This implies among the other things, that the rate of return for the firm's asset and hence the volatility cannot be computed.

Furthermore in some situations for the pricing of credit risk derivatives it could be important to consider directly a stochastic process for the prices of defaultable bonds. This is the case of reduced-form models.

1.2   Reduced-form Models

In reduced-form models default is treated as an unpredictable event driven by a hazard-rate process, i.e. the time of default is assumed to be the time of the first jump of a Poisson process. A Poisson process, $\{N_t\}$, with intensity λ is defined as a non-decreasing integer-valued process with initial value $N_0 = 0$, whose increments are independent and satisfy for $s > t$

$$P(N_s - N_t = n) = \frac{1}{n!}[(s-t)\lambda]^n \mathrm{e}^{-(s-t)\lambda}. \tag{1.21}$$

To derive (1.21), assume that the probability of a jump in a small time interval Δt is

$$P(N_{t+\Delta t} - N_t = 1) = 1 - P(N_{t+\Delta t} - N_t = 0) = \lambda \Delta t,$$

so that more 1 jump cannot occur in $[t, t + \Delta t]$. Now subdivide the interval $[t, s]$ in k subintervals of length $\Delta t = (s - t)/k$. Assume that the probability of a jump in each of these subinterval is $\Delta t \lambda$ and that a jump in a subinterval is independent of a jump in another subinterval. Then

$$P(N_s - N_t = n) = \binom{k}{n} \left(\frac{s - t}{k} \lambda \right)^n \left(1 - \frac{s - t}{k} \lambda \right)^{k - n}$$

$$\overset{k \to \infty}{\longrightarrow} \frac{1}{n!} [(s - t)\lambda]^n \mathrm{e}^{-(s - t)\lambda}.$$

One can think of N_t as the process for the number of arrivals (in our case defaults) up to time t: it is a process in the integers $0, 1, 2, \ldots$ where the times of the jumps are $T_1, T_2, \ldots$. Note that the inter-arrival times of a Poisson process $(T_{n+1} - T_n)$ are exponentially distributed and the density is

$$P\left[(T_{n+1} - T_n) \in (t, t + \mathrm{d}t)\right] = \lambda \mathrm{e}^{-\lambda t}\mathrm{d}t.$$

A more realistic model for the number of defaults could be an inhomogeneous Poisson process, which has the same properties of an homogeneous Poisson process except that the intensity is a function of time. Repeating the construction used to get (1.21) yields

$$P(N_s - N_t) = \frac{1}{n!} \left(- \int_t^T \lambda_u \mathrm{d}u \right)^n \exp\left(- \int_t^T \lambda_u \mathrm{d}u \right),$$

i.e. the constant $(s - t)\lambda$ is replaced by the integral $\int_t^s \lambda_u \mathrm{d}u$.

The first model presented in the following is the one proposed by Jarrow and Turnbull (1995). They first analyze a two-period discrete time economy, with a default-free and corporate bonds (the latter issued by the same firm). All the bonds have face value equal to one and maturity $T = 1, 2$. They define the function

$$e(t) = \overline{B}(t, t),$$

which represents the time t payoff of the corporate bond. They first assume a binomial process for the spot interest rate. Moreover they select a discrete time binomial process to approximate a continuous time Poisson process. In particular, at time $t = 1$ default occurs with probability $\lambda \mu_0$ and in this case the bondholders will get the fraction q. The probability that default happens at time $t = 2$ conditional on the information available up to time $t = 1$ is $\lambda \mu_1$. If the firm defaults at time $t = 1$, it remains in default and the payoff is again q. Thus

$$\mathbb{E}_0^Q[e(1)] = q\lambda\mu_0 + (1 - \lambda\mu_0), \tag{1.22a}$$

$$\mathbb{E}_0^Q[e(2)] = q\lambda\mu_0 + (1 - \lambda\mu_0)[q\lambda\mu_1 + (1 - \lambda\mu_1)]. \tag{1.22b}$$

Equation (1.22b) is derived from

$$\mathbb{E}_1^Q[e(2)] = \begin{cases} q & \text{if bankrupt at time 1} \\ q\lambda\mu_1 + (1 - \lambda\mu_1) & \text{otherwise} \end{cases}.$$

An important assumption is the independence between the process for the spot interest rate and the bankruptcy process under the risk neutral probability. It can be shown that this implies for $t = 1, 2$

$$\overline{B}(t, T) = B(t, T)\mathbb{E}_t^Q[e(T)]. \tag{1.23}$$

We can rewrite (1.23) as

$$\begin{aligned}\overline{B}(t, T) &= B(t, T)\left[Q_t(\tau > T) + q(1 - Q_t(\tau > T))\right] \\ &= B(t, T)\left[q + (1 - q)Q_t(\tau > T)\right].\end{aligned} \tag{1.24}$$

In a continuous time setting they start with a process for the forward rate[6] for both the riskless and defaultable bond. Moreover the function $e(\cdot)$ is specified as

$$e(t) = \begin{cases} q & \text{if } t \geq \tau \\ 1 & \text{if } t < \tau \end{cases},$$

where as usual τ is the time of default. They assume that τ has an exponential distribution over $[0, \infty)$ with parameter λ and that the Poisson bankruptcy process has constant intensity μ under the risk neutral probability. This implies the independence of the bankruptcy process from the default-free interest rate process under the risk neutral probability, because in this case the time of bankruptcy has an exponential distribution with parameter $\lambda\mu$, which is independent of the spot interest rate process. As a consequence, equation (1.23) is still valid and it can be shown that

$$\overline{B}(t,T) = B(t,T)\mathbb{E}_t^Q[e(T)]$$

$$= B(t,T) \begin{cases} q & \text{if } t \geq \tau \\ e^{-\lambda\mu(T-t)} + q\left(1 - e^{-\lambda\mu(T-t)}\right) & \text{if } t < \tau \end{cases}.$$

Note that risk neutral probability of no-default at time T, conditional to the fact that no-default has been triggered at time t is

$$Q_t(\tau > T) = e^{-\lambda\mu(T-t)}.$$

To see that (1.24) is valid in a general setting, denote by $\beta(t)$ the money market account process and by r_t the default-free spot rate. Clearly

$$\beta(t) = \begin{cases} \exp\left(\sum_{i=0}^{t} r_i\right) & \text{in discrete time} \\ \exp\left(\int_0^t r_u \mathrm{d}u\right) & \text{in continuous time} \end{cases},$$

and the price of the risk-free z.c.b. satisfies

$$\frac{B(t,T)}{\beta(t)} = \mathbb{E}_t^Q\left[\frac{B(T,T)}{\beta(T)}\right],$$

which implies

$$B(t,T) = \mathbb{E}_t^Q\left[\frac{\beta(t)}{\beta(T)}\right].$$

---

[6]The default-free forward rate is defined implicitly by

$$f(t,T) \equiv \begin{cases} -\log\left(\frac{B(t,T+1)}{B(t,T)}\right) & \text{in discrete time} \\ -\frac{\partial \log B(t,T)}{\partial T} & \text{in continuous time} \end{cases}.$$

The same definition applies also to the defaultable forward rate. Note that in the continuous case

$$B(t,T) = \exp\left(-\int_t^T f(t,u)\mathrm{d}u\right).$$

Moreover the default-free spot interest rate can be obtained as

$$r(t) = f(t,t).$$

In fact, denoting by Q the risk neutral measure

$$\frac{\partial}{\partial T}B(t,T)\Big|_{T=t} = \frac{\partial}{\partial T}\mathbb{E}_t^Q\left[\exp\left(-\int_t^T r_u\mathrm{d}u\right)\right]\Big|_{T=t}$$

$$= \mathbb{E}_t^Q\left[-r_T\exp\left(-\int_t^T r_u\mathrm{d}u\right)\right]\Big|_{T=t} = -r_T$$

but also

$$\frac{\partial}{\partial T}B(t,T)\Big|_{T=t} = \frac{\partial}{\partial T}\exp\left(-\int_t^T f(t,u)\mathrm{d}u\right)\Big|_{T=t} = -B(t,T)f(t,T)|_{T=t} = -f(T,T).$$

The price of the corporate z.c.b. is given by

$$\overline{B}(t,T) = \mathbb{E}_t^Q \left[\frac{\beta(t)}{\beta(T)} \left(q I_{\{\tau \leq T\}} + I_{\{\tau > T\}} \right) \right],$$

and since the stochastic process for default-free spot rates and the bankruptcy process are statistically independent under Q,

$$\overline{B}(t,T) = \mathbb{E}_t^Q \left[\frac{\beta(t)}{\beta(T)} \right] \mathbb{E}_t^Q \left[q I_{\{\tau \leq T\}} + I_{\{\tau > T\}} \right],$$

which implies (1.24).

It could be interesting to consider a slight modification to this model and consider the possibility of multiple defaults. Assume that after the first default subsequent defaults may happen and every time there is a default, a reorganisation takes place and the bondholders lose a fraction y of the face value of their claims, but the claims continue to live and the issuer continues to operate. To model this situation, it is useful to consider the inhomogeneous Poisson process $\{N_t\}$ with possibly time varying intensity λ_t. Assuming a stochastic risk-free interest rate r_t and using (1.12) with $X' = 0$ and $X = (1-y)^{N_T}$ the price of the corporate bond is:

$$\overline{B}(t,T) = \mathbb{E}_t^Q \left[\exp\left(-\int_t^T r_u \mathrm{d}u \right) (1-y)^{N_T} \right]. \tag{1.25}$$

Assuming again the independence of the bankruptcy process from the default-free interest rate process under the risk neutral probability, (1.25) becomes

$$\overline{B}(t,T) = \mathbb{E}_t^Q \left[\exp\left(-\int_t^T r_u \mathrm{d}u \right) \right] \mathbb{E}_t^Q \left[(1-y)^{N_T} \right]. \tag{1.26}$$

Let us evaluate the second term of the right hand side of (1.26) when λ_t is constant, i.e. when the Poisson process is homogeneous. We have

$$\mathbb{E}_t^Q \left[(1-y)^{N_T} \right] = \sum_{n=0}^{+\infty} (1-y)^n \frac{1}{n!} [\lambda(T-t)]^n \mathrm{e}^{-\lambda(T-t)}$$

$$= \mathrm{e}^{(1-y)\lambda(T-t)} \mathrm{e}^{-\lambda(T-t)} = \mathrm{e}^{-y\lambda(T-t)}.$$

This implies that the price of the corporate bond is

$$\overline{B}(t,T) = B(t,T) \mathrm{e}^{-y\lambda(T-t)},$$

and the yield spread

$$s(t,T) = -\frac{1}{T-t} \log\left(\frac{\overline{B}(t,T)}{B(t,T)} \right) = y\lambda.$$

For the inhomogeneous Poisson process again it is sufficient to replace $(T-t)\lambda$ with the integral $\int_t^T \lambda_u \mathrm{d}u$ and therefore (1.26) becomes

$$\overline{B}(t,T) = \mathbb{E}_t^Q \left[\exp\left(-\int_t^T r_u \mathrm{d}u \right) \right] \exp\left(-y \int_t^T \lambda_u \mathrm{d}u \right)$$

Plugging the second term in the expectation yields

$$\overline{B}(t,T) = \mathbb{E}_t^Q \left[\exp\left(-\int_t^T (r_u + y\lambda_u) \mathrm{d}u \right) \right]. \tag{1.27}$$

Equation (1.27) tells us that the price of a defaultable bond of maturity T equals the price a default-free bond of the same maturity in a world where the risk-free short rate is

$$R_t = r_t + y\lambda_t.$$

A result similar to (1.27) can be found in Duffie and Singleton (1999). They denote by h_t the hazard rate[7] for default at time t and with y_t the expected fractional loss in the market value of the defaultable claim if default were to occur at time t, conditional on the information up to time t. Both this quantity are to be considered under the risk neutral probability measure. They conclude, assuming the mean-loss rate[8] process $h_t y_t$ to be given exogenously, that under technical conditions, the t price of the claim with payoff X in T is

$$U_t = \mathbb{E}_t^Q \left[\exp\left(-\int_t^T R_u \mathrm{d}u \right) X \right], \tag{1.28}$$

where $R_t = r_t + h_t y_t$ is the default-adjusted short rate process. This means that a bond may be priced as if it were default-free, replacing r with R. Let us illustrate their model in a discrete time setting. Denote by h_t the conditional probability under Q of default between t and $t+1$ and by φ_t the recovery (in units of money) in the event of default at time t. Consider a defaultable claim promising the amount X_{t+T} at maturity $t+T$. If the firm has not defaulted by time t, then the price, U_t of the claim in t is given by the expected value of $\mathrm{e}^{-r_t}\varphi_{t+1}$ in case of default between t and $t+1$ and the expected value of $\mathrm{e}^{-r_t}U_{t+1}$ otherwise. This means

$$U_t = h_t \mathbb{E}_t^Q \left[\mathrm{e}^{-r_t}\varphi_{t+1} \right] + (1-h_t)\mathbb{E}_t^Q \left[\mathrm{e}^{-r_t}U_{t+1} \right],$$

or equivalently

$$U_t = h_t \mathrm{e}^{-r_t}\mathbb{E}_t^Q \left[\varphi_{t+1} \right] + (1-h_t)\mathrm{e}^{-r_t}\mathbb{E}_t^Q \left[U_{t+1} \right]. \tag{1.29}$$

The next step is to consider the "recovery of market value" hypothesis: the risk neutral expected recovery at t is a fraction, $1 - y_t$, of the risk neutral expected survival-contingent market value at time $t+1$, i.e.

$$\mathbb{E}_t^Q \left[\varphi_{t+1} \right] = (1 - y_t)\mathbb{E}_t^Q \left[U_{t+1} \right]. \tag{1.30}$$

In a continuous time setting this hypothesis is written as

$$\overline{B}(t,T) = L_\tau \overline{B}(\tau^-, T),$$

where $\overline{B}(\tau^-, T)$ is the value of the corporate bond just before default. Note that this is equivalent to say that the claim made by bondholders in the event of default is equal to the value of the bond immediately prior to default whereas in Jarrow and Turnbull (1995) one has $\varphi_t = (1 - y_t)B(t,T)$, where L is an exogenously specified fractional recovery process and, as usual, $B(t,T)$ is the price at time t of an otherwise identical default-free bond. This is called "recovery of treasury" hypothesis. An alternative assumption on the recoveries to creditors during bankruptcy proceedings is the "recovery of face value" hypothesis: in case of default the creditor receives the fraction $y = 1 - q$ of face value immediately upon default. Plugging eq. (1.30) in (1.29) yields

$$U_t = (1 - h_t y_t)\mathrm{e}^{-r_t}\mathbb{E}_t^Q \left[U_{t+1} \right].$$

For annualized rates with small time periods, one can use the approximation $\mathrm{e}^c \simeq 1 + c$, to get

$$U_t = \mathrm{e}^{-R_t}\mathbb{E}_t^Q \left[U_{t+1} \right], \tag{1.31}$$

where $R_t \simeq r_t + h_t y_t$. Equation (1.31) implies

$$U_t = \mathbb{E}_t^Q \left[\exp\left(-\sum_{j=0}^{T-1} R_{t+j} \right) X_{t+T} \right]$$

---

[7] The hazard rate h_t can be thought so that $h_t \Delta t$ is the probability of default between times t and $t + \Delta t$ as seen at time t assuming no default between time zero and time t.

[8] Note that this is a measure for the exposure to default.

which is the discrete time version of (1.28). In the continuous time setting the authors introduce a vector of state variables Z_t driving the term structure of defaultable bond prices. In practice these are to be inferred from the existing level of credit spreads. These state variable are assumed to follow a vector stochastic differential equation that generalizes the Cox, Ingersoll, and Ross model for the instantaneous spot rate to the vector case. Let A a diagonal matrix of speeds of mean reversion and Θ a vector of long term levels to which the state variables are reverting. Let Σ be the covariance matrix between the state variables and B_t be a standard vector Brownian motion of dimension equal to that of Z. Thus the stochastic differential equation governing the evolution is given by:

$$\mathrm{d}Z_t = A(\Theta - Z_t)\mathrm{d}t + \Sigma \operatorname{diag}(Z_t)^{1/2}\mathrm{d}B_t. \tag{1.32}$$

Both the instantaneous spot rate and the mean-loss rate are assumed to be linear in the state variables:

$$r_t = \delta_0 + \delta' Z_t,$$
$$h_t y_t = \gamma_0 + \gamma' Z_t.$$

Since the price of a defaultable bond requires that we evaluate the expectation of an exponential of the integral of a linear function of the state variables, these prices may be obtained in closed form by evaluating the joint characteristic function. For the process (1.32) it is possible to derive in closed form an exponential affine expression for the joint characteristic function of the integral of the process:

$$\mathbb{E}^Q\left[\exp\left(\mathrm{i}\xi\int_0^t Z_u \mathrm{d}u\right)\right] = \exp\left[a(t,\xi) + b(t,\xi)' Z_0\right],$$

where the coefficient function $a(t,\xi)$ and $b(t,\xi)$ may be explicitly solved for. Equation (1.27) can be obtained also using a PDE approach. As usual denote by $\{N_t\}$ a Poisson process with intensity $\lambda(t)$. Assume a jump-diffusion model for the equity under the risk neutral measure:

$$\mathrm{d}S = [r + \lambda(t)]S\mathrm{d}t + \sigma \mathrm{d}B - S\mathrm{d}N. \tag{1.33}$$

This model is consistent with the risk neutral valuation[9] as

$$\mathbb{E}[S] = rS\mathrm{d}t.$$

Moreover, the instantaneous variance of the equity return process under jump-to-default will be $\sigma^2 + \lambda$ since

$$\mathbb{E}\left[\left(\frac{\mathrm{d}S}{S}\right)^2\right] = (\sigma^2 + \lambda)\mathrm{d}t.$$

Therefore, if λ changes over time, the equity process will have stochastic variance even when diffusion volatility σ is constant. Denote by $\overline{B} = \overline{B}(S,t)$ the debt of maturity T and assume that on default the bond pays a recovery fraction $1-y$ of its value just before default, which is the "recovery of market value" condition. Let us write

$$\mathrm{d}S = \mathrm{d}S^c - S\mathrm{d}N$$

where S^c is the continuous part of S. To derive the dynamics of the bond price, we need the modified version of Ito's lemma suitable for jumps in process:

$$\mathrm{d}\overline{B} = \overline{B}_S \mathrm{d}S^c + \overline{B}_t \mathrm{d}t + \frac{1}{2}\overline{B}_{SS}(\mathrm{d}S^c)^2$$
$$+ [\overline{B}(S - \mathrm{d}N, t) - \overline{B}(S,t)]\mathrm{d}N.$$

---

[9] It is worth to summarise here the rule of thumbs:

$$(\mathrm{d}N)^2 = \mathrm{d}N \qquad\qquad \mathbb{E}(\mathrm{d}N) = \lambda\mathrm{d}t$$
$$(\mathrm{d}t)^2 = (\mathrm{d}N)(\mathrm{d}B) = 0 \qquad\qquad (\mathrm{d}B)^2 = \mathrm{d}t.$$

Note that the last term is

$$(1 - y)\overline{B} - \overline{B} = -y\overline{B}$$

if there is a jump in the interval of length $\mathrm{d}t$. Therefore it holds

$$\mathrm{d}\overline{B} = \left[\overline{B}_S(r + \lambda)S + \frac{1}{2}\overline{B}_{SS}\sigma^2 S^2 + \overline{B}_t \right]\mathrm{d}t$$
$$+ \overline{B}_S\sigma S\mathrm{d}B - y\overline{B}\mathrm{d}N.$$

Since in a risk neutral world the bond must earn the risk free rate of return it follows

$$\mathbb{E}[\mathrm{d}N] = rD\mathrm{d}t$$

and thus $\overline{B}$ must fulfil

$$(r + y\lambda)\overline{B} = \overline{B}_S(r + \lambda)S + \frac{1}{2}\overline{B}_{SS}\sigma^2 S^2 + \overline{B}_t.$$

Using the Feynman-Kac formula the solution to the PDE is exactly (1.27). It is natural to assume that default intensity λ is inversely related to the level of the stock price S. This is because as the stock price falls, leverage will increase, and λ will increase, resulting also in higher variance of the stock price. From equation (1.27), given the assumption that λ is inversely related to S, it is easy to show that debt value is an increasing and concave function of equity S. Since the firm value is the sum of debt and equity, it follows that the debt will be concave in the firm value as well. This corresponds to what is seen in structural models: recall equation (1.4) and notice that the negative position in the put (a convex function) makes debt a concave function of firm value.

Note that, since equity value and firm value specifications are reflections of each other, we could begin with firm value and derive the resultant stock value process. For instance let

$$\frac{\mathrm{d}V}{V} = \mu\mathrm{d}t + \eta\mathrm{d}B + (Y - 1)\mathrm{d}N.$$

In this formulation $Y - 1$ represents the percentage change from the jump in firm value, meaning that in case of default firm value V moves to VY. Denoting by $\ell = \frac{\overline{B}}{V}$ the leverage ratio and assuming again the RMV hypothesis with zero recovery for the shareholders, the percentage change in firm value is

$$\frac{\mathrm{d}V}{V}\mathrm{d}N = \frac{V_{\text{def}}}{V} - 1 = (1 - y)\ell - 1$$

because the firm value in case of default is

$$V_{\text{def}} = S_{\text{def}} + \overline{B}_{\text{def}} = \overline{B}_{\text{def}} = (1 - y)\overline{B}.$$

We can now apply the jump-extended Ito's lemma to get the dynamics for the equity:

$$\mathrm{d}S = \mu_S\mathrm{d}t + \eta_S\mathrm{d}B + \{S[(1 - y)\ell - 1]V - S\}\mathrm{d}N$$
$$= \mu_S\mathrm{d}t + \eta_S\mathrm{d}B - S\mathrm{d}N$$

where μ_S and η_S are derived using Ito's lemma. The term $S[(1 - y)\ell - 1]V$ is zero because recovery value is less than debt value. However starting from the dynamics for the equity (1.33), since the identity $V = S + \overline{B}$ holds, we can derive the process for the firm value as

$$\mathrm{d}V = \mathrm{d}B + \mathrm{d}\overline{B}$$
$$= \left[(1 + \overline{B}_S)(r + \lambda)S + \frac{1}{2}\overline{B}_{SS}\sigma^2 S^2 + \overline{B}_t \right]\mathrm{d}t$$
$$+ (1 + \overline{B}_S)\sigma S\mathrm{d}B - (y\overline{B} + S)\mathrm{d}N.$$

This means that the local volatility of $\mathrm{d}V/V$ is $(1 + \overline{B}_S)\frac{\sigma S}{V}$, which is not a constant, even if σ is constant. Thus, an implication of the model is that the volatility of firm value changes over time. It is possible to show that, keeping constant the face value of the debt,

$$\lim_{S\to\infty} (1 + \overline{B}_S)\frac{\sigma S}{V} = \sigma.$$

The explanation for this result is that if the firm is almost all equity, then its volatility will be that of equity. The same idea of an analysis based on the stochastic process for the equity can be found also in Das and Sundaram (2005). As in the reduced-form approach, they describe default likelihood by the means of a hazard-rate process. From the structural model approach, they borrow the boundary conditions on equity in the event of default. Hence default is identified with zero equity value. The model for the evolution of equity prices is based on a trinomial branching process with two non-default branches and one absorbing default branch. Given the stock price at t the evolution over the next period is

$$S(t + \Delta t) = S(t)\exp[\sigma_s X_s(t)\sqrt{\Delta t}],$$

where $X_s(t)$ is a discrete random variable taking on values in the set $\{+1, -1, -\infty\}$ and σ_s is the the parameter governing the volatility of the equity process on the non-default path. When $X_s(t) \to -\infty$ the firm suddenly defaults and its stock price goes to zero. Next, an univariate Heath, Jarrow, and Morton lattice for the evolution of the term structure is considered. Forward rates are assumed to follow the process

$$f(t + \Delta t) = f(t, T) + \alpha(t, T)\Delta t + \sigma(t, T)X_f\sqrt{\Delta t},$$

where the deterministic functions of time α and σ are the drift and the volatility of the process X_f is a random variable taking values in the set $\{-1, +1\}$. The two processes for the term structure and the defaultable equity price are connected together on a bivariate lattice with 6 emanating branches, i.e. 6 different states are possible at each node. The next step requires the specification of the joint process requires a probability measure over the random vector $[X_s(t), X_f(t)]$. Basically the probability measure is chosen so that normalized equity prices and bond prices are martingales. This is achieved by setting

$$\mathbb{E}\left[\frac{S(t + \Delta t)}{S(t)}\right] = \mathbb{E}[\exp(\sigma_s X_s(t))] = \exp(r(t)\Delta t),$$

$$\mathbb{E}[X_f(t)] = 0,$$

$$\mathrm{Var}[X_f(t)] = 1.$$

The next condition is on the correlation:

$$\mathrm{Corr}(X_s, X_f) = \rho.$$

Rather than adding an extra dimension to the lattice model by embedding a separate process, one-period default probability functions at each node on the bivariate lattice are looked at as function of equity prices and interest rates at each node. This is because equity prices already reflect credit risk, and default probabilities are empirically known to be connected to the term structure.

Some papers developed within the reduced-form framework use a rating-based approach in which default is attained through changes in credit ratings driven by a Markovian transition matrix. Among the others it is worth citing Das and Tufano (1996), Jarrow et al. (1997) and Duffee (1998).

In particular in Jarrow et al. (1997) the distribution for the default time is modeled via a discrete time, time-homogeneous Markov chain on a finite state space $S = \{1, \ldots, K\}$. S has to be interpreted as the space of the possible credit classes, so that the state 1 represents the highest class, $K-1$ the lowest and K represents bankruptcy. The Markov chain $\{\eta_t : 0 \le t \le T^*\}$ is specified under the market probability by the $K \times K$ transition matrix $\mathbf{M} = [m_{i,j}]_{i,j=1,\ldots,K}$ with $m_{i,j} \ge 0$ for all i, j, $i \ne j$ and $m_{i,i} \equiv 1 - \sum_{\substack{j=1 \\ j\ne i}}^{K} m_{i,j}$ for all i. The state K is assumed to be absorbing, i.e. $m_{K,i} = 0$ for all $i \ne K$ and $m_{K,K} = 1$. $m_{i,j}$ represents the probability

under the market measure of going from state i to state j in one time step. Under the equivalent martingale measure, however, the transition probability from the state i time at time t to the state j time at time $t+1$ can depend on the entire history of the process up to time t. Hence, under the martingale probabilities, the process need not be Markov. Therefore the transition matrix is $\tilde{\mathbf{M}}_{t,t+1} = [\tilde{m}_{i,j}(t,t+1)]_{i,j=1,\ldots,K}$ where for every $t \geq 0$

$$\tilde{m}_{i,j}(t,t+1) \geq 0 \quad i,j \quad i \neq j$$

$$\tilde{m}_{i,i}(t,t+1) \equiv 1 - \sum_{\substack{j=1 \\ j \neq i}}^{K} \tilde{m}_{i,j}(t,t+1)$$

$$\tilde{m}_{K,i}(t,t+1) = 0 \quad \tilde{m}_{K,K}(t,t+1) = 1$$

$$\tilde{m}_{i,j}(t,t+1) > 0 \Leftrightarrow m_{i,j} > 0 \quad \text{for } 0 \leq t \leq T^* - 1.$$

To facilitate empirical implementation the following assumption is made:

$$\tilde{m}_{i,j}(t,t+1) = \pi_i(t)m_{i,j} \quad \text{for all } i, j, i \neq j, \tag{1.34}$$

where $\pi_i(t)$ is a deterministic function of time such that $\tilde{m}_{i,j}(t,t+1) \geq 0$ for all $i, j, i \neq j$ and $\sum_{\substack{j=1 \\ j \neq i}}^{K} \tilde{m}_{i,j}(t,t+1)$ for $i = 1,\ldots,K$. The proportionality adjustments $\pi_i(t)$ are used to transform the actual probabilities to those used in valuation and can be interpreted as risk premia. The independence of j in the π's implies that moving from i to 1 receives the same risk premium as moving from i to K. Note that $\pi_i(t) \geq 0$ for all i and t and $\pi_K(t) \equiv 1$. Besides we can restate the relationship between $\tilde{\mathbf{M}}$ and $\mathbf{M}$ as follows:

$$\tilde{\mathbf{M}}_{t,t+1} - \mathbf{I} = \mathbf{\Pi}(t)[\mathbf{M} - \mathbf{I}],$$

where $\mathbf{I}$ is the $K \times K$ identity matrix and

$$\mathbf{\Pi}(t) = \text{diag}(\pi_1(t),\ldots,\pi_{K-1}(t),1).$$

Now, given this structure, it is straightforward to compute the risk neutral probability of default occurring after date T for a firm in the class i at time t:

$$Q_t^i(\tau > T) = \sum_{j=1}^{K-1} \tilde{m}_{i,j}(t,T) = 1 - \tilde{m}_{i,K}(t,T). \tag{1.35}$$

Cleary (1.35) implies

$$Q_t^K(\tau > T) = 0 \quad \text{for all } 0 \leq t \leq T \leq T^*. \tag{1.36}$$

Note that the matrix $\tilde{\mathbf{M}}_{t,T} = [\tilde{m}_{i,j}(t,T)]_{i,j=1,\ldots,K}$ satisfies:

$$\tilde{\mathbf{M}}_{t,T} = \prod_{h=0}^{T-t-1} \tilde{\mathbf{M}}_{t+h,t+h+1}.$$

To obtain the t price of a of a zero-coupon bond issued by a firm in credit class i at time t assuming that if bankrupt, the firm pays only $q < 1$, just plug (1.35) in (1.24). To get an expression for the credit risk spread, evaluate first the forward rate:

$$f^i(t,T) \equiv -\log\left(\frac{\overline{B}^i(t,T+1)}{\overline{B}^i(t,T)}\right)$$

$$= f(t,T) + I_{\{\tau > t\}} \log\left(\frac{q + (1-q)Q_t^i(\tau > T)}{q + (1-q)Q_t^i(\tau > T+1)}\right)$$

The first term in the right hand side is due to the fact that it is assumed that bondholders will receive the amount q for sure at the maturity of the contract if bankruptcy has occurred prior to time t. Note that from (1.36) one gets $f^K(t,T) = f(t,T)$. The spread is simply

$$s^i(t,T) = \begin{cases} I_{\{\tau>t\}} \log\left(\frac{q+(1-q)Q_t^i(\tau>T)}{q+(1-q)Q_t^i(\tau>T+1)}\right) & \text{for } i = 1,\ldots,K-1 \\ 0 & \text{for } i = K \end{cases}.$$

Moreover the change in the firm's forward rate over the interval $[t,t+1]$ can be derived as follows. Consider a firm whose time t credit rating $\eta_t = i$. Then

$$f^{\eta_{t+1}}(t+1,T) - f^i(t,T) = [f(t+1,T) - f(t,T)]$$
$$+ \left[s^{\eta_{t+1}}(t+1,T) - s^i(t,T)\right]$$

for $\eta_{t+1} \in \{1,\ldots,K\}$ with pseudo-probabilities $\tilde{m}_{i,\eta_{t+1}}(t,t+1)$. The first component of the change in the risky firm's forward rates is due to changes in the default-free forward rate structure and a shortening of time to maturity, whereas the second one is due to changes in the default probability arising from an unpredictable change in credit class and a (predictable) shortening of the time to maturity. This risk has at most K different outcomes and to hedge this credit class risk, in general, one needs at most K of this firm's credit risky securities.

Finally for the spot rate we note that $Q_t^i(\tau>t) = 1$ and $Q_t^i(\tau>t+1) = 1 - \tilde{m}_{i,K}(t,t+1)$ to get

$$r^i(t) = r(t) + I_{\{\tau>t\}} \log\left(\frac{1}{1-(1-q)\tilde{m}_{i,K}(t,t+1)}\right).$$

Again, $r^K(t) = r(t)$. We have seen that, given estimates of the empirical transition matrix $\mathbf{M}$, and the risk premium $\pi_i(t)$ for $0 \le t \le T^* - 1$, one can easily derive the price of the risky zero-coupon bonds. However one may be interested in deriving the risk premium given the term structure of credit-risky zeros. Suppose also the recovery rate q is known. Rewriting (1.24) as

$$\overline{B}^i(t,T) = B(t,T)\left[1 - (1-q)Q_t^i(\tau \le T)\right]$$

and simple algebra yields

$$Q_0^i(\tau \le T) = \frac{B(0,T) - \overline{B}^i(0,T)}{(1-q)B(0,T)} \tag{1.37}$$
$$\text{for } i = 1,\ldots,K \text{ and } T = 1,\ldots,T^*$$

Since

$$Q_0^i(\tau \le 1) = \tilde{m}_{i,K}(0,1) = \pi_i(0)m_{i,K},$$

it follows form (1.37)

$$\pi_i(0) = \frac{B(0,T) - \overline{B}^i(0,T)}{(1-q)m_{i,K}B(0,T)}.$$

This allows us to evaluate

$$\tilde{\mathbf{M}}_{0,1} = \mathbf{I} + \mathbf{\Pi}(0)[\mathbf{M} - \mathbf{I}].$$

In order to obtain the matrix $\mathbf{\Pi}(1)$, first note that

$$\tilde{\mathbf{M}}_{0,2} = \tilde{\mathbf{M}}_{0,1}\tilde{\mathbf{M}}_{1,2} = \tilde{\mathbf{M}}_{0,1}\left[\mathbf{I} + \mathbf{\Pi}(1)(\mathbf{M} - \mathbf{I})\right],$$

and that

$$Q_0^i(\tau \le 2) = \tilde{m}_{i,K}(0,2) = \sum_{j=1}^K \tilde{m}_{i,j}(0,1)\tilde{m}_{j,K}(1,2) = \sum_{j=1}^K \tilde{m}_{i,j}(0,1)\pi_j(1)m_{j,K}.$$

Now it is sufficient to solve the linear system

$$\sum_{j=1}^{K} \tilde{m}_{i,j}(0,1)\pi_j(1)m_{j,K} = \frac{B(0,2) - \overline{B}^i(0,2)}{(1-q)B(0,2)}$$

where the only unknowns are the variables $\pi_j(1)$. This procedure is repeated for $t = 1, \ldots, T^* - 1$. It can be shown that if $\tilde{\mathbf{M}}_{0,t}^{-1}$ exists and its entries are $\bar{m}_{i,j}(0,t)$, then

$$\pi_i(t) = \sum_{j=1}^{K} \bar{m}_{i,j}(0,t)\frac{B(0,t+1) - \overline{B}^i(0,t+1)}{(1-q)B(0,t+1)m_{i,K}}.$$

In a credit risk framework, it could be useful to extend the Heath, Jarrow, and Morton model to the case of a defaultable bond. Using the usual notation, $f^i(t,T)$ is the forward rate prevailing at time t for a defaultable bond in the rating class i and $f(t,T)$ is the forward rate for risk-free bond. It is assumed that under the market measure

$$f^i(t,T) = f^i(0,T) + \int_0^t \mu^i(u,T)\mathrm{d}u + \int_0^t \sigma^i(u,T)\mathrm{d}B_u.$$

To derive the price of the bond or other credit derivatives, it is important to understand the dynamics of the forward rate under the risk neutral measure. In other words one needs to derive some restrictions on the drift and the diffusion that make discounted prices martingales under the risk neutral measure. By recovering the prices from forward rates, and applying Ito's lemma one deduces that the martingale condition determines completely the forward rate drifts $\mu^i(u,T)$. Under the "recovery of market value" hypothesis, i.e. when recovery in default is defined as a proportion of the predefault price of the bond, this drift restriction is seen to be identical to its counterpart in default free bond pricing:

$$\mu^i(t,T) = \sigma^i(t,T)\left(\int_t^T \sigma^i(u,T)\mathrm{d}u\right)'.$$

Under the "recovery of treasury" hypothesis when recovery is a proportion of the price of a Treasury or default free bond, this restriction is not so straighforward. In particular

$$\mu^i(t,T) = \sigma^i(t,T)\left(\int_t^T \sigma^i(u,T)\mathrm{d}u\right)' + h_t(1-y_t)\frac{\overline{B}^i(t,T)}{B(t,T)}\left[f^i(t,T) - f(t,T)\right].$$

It could be very important to impose these drift restrictions in simulations or in the construction of risk neutral trees to determine claim valuations.

Another interesting model is proposed in Madan and Unal (1998). Basically they model the hazard rate of default as a decreasing function of equity value measured in units of the money market account, $S^*(t) = S(t)/\beta(t)$. This discounted price is modeled as a martingale:

$$\mathrm{d}S^*(t) = \sigma S^*(t)\mathrm{d}B(t).$$

The intensity of default is modeled by

$$h_t = \phi(S^*(t)) = \frac{c}{\left(\log\left(\frac{S^*(t)}{\delta}\right)\right)^2}$$

where c is a positive constant and δ is a parameter to be empirically estimated. If $\delta > S^*(t)$ then h is an increasing function of the equity value and thus increases in equity value are seen as signals of greater risk increasing the value of equity viewed as an option. On the other hand if δ is below the current equity values

then h is a decreasing function of S^* and the hazard rate tend to infinity as S^* approaches δ. Further, the authors show that under Q:

$$G(t,T) := Q^T(\tau > T) = \mathbb{E}_t^Q \left[\exp \left(- \int_t^T h_u \mathrm{d}u \right) \right]. \tag{1.38}$$

Since we are in a Markov setting, it follows that this survival probability is of the form

$$G(t,T) = \psi(S^*, t).$$

Define

$$g(S^*, t) = \exp \left(- \int_0^t h_u \mathrm{d}u \right) \psi(S^*, t) =: \varphi(S^*, t) \psi(S^*, t).$$

Note that since this quantity is $\mathscr{F}_t$-measurable we have

$$g(S^*, t) = \mathbb{E}_t^Q \left[\exp \left(- \int_0^T h_u \mathrm{d}u \right) \right].$$

Since for $s < t$

$$\mathbb{E}_s^Q \left[g(S^*, t) \right] = \mathbb{E}_s^Q \left[\exp \left(- \int_0^T h_u \mathrm{d}u \right) \right] = g(S^*, s),$$

it follows that $g(S^*, t)$ is a Q-martingale.
Let us apply Ito's lemma to get

$$\mathrm{d}g = \psi_{S^*} \varphi \mathrm{d}S^* + \frac{1}{2} \psi_{S^* S^*} (\mathrm{d}S^*)^2 + [\psi_t - \phi(S^*)\psi] \, \mathrm{d}t$$

$$= \psi_{S^*} \varphi \sigma S^* \mathrm{d}B + \frac{1}{2} \sigma^2 (S^*)^2 \psi_{S^* S^*} \mathrm{d}t + [\psi_t - \phi(S^*)\psi] \, \mathrm{d}t.$$

Since $g(S^*, t)$ is a Q-martingale, ψ satisfies the partial differential equation

$$\frac{1}{2} \sigma^2 (S^*)^2 \psi_{S^* S^*} + \psi_t = \phi(S^*)\psi$$

subject to the boundary condition

$$\psi(S^*, T) = 1.$$

The solution is given by

$$\psi(S^*, t) = G_a \left(\frac{2}{d^2(S^*, T)} \right)$$

$$d(S^*, T) = \frac{\log(S^*/\delta)}{\sigma \sqrt{T}} - \frac{\sigma \sqrt{T}}{2}$$

$$a = \frac{2c}{\sigma^2}$$

where $G_a(y)$ satisfies the ordinary differential equation

$$y^2 G'' + \left(\frac{3y}{2} - 1 \right) G' - aG = 0$$

subject to $G(0) = 1$, $G'(0) = -a$.
Although reduced form models have many advantages (the most important is that they generate realistic short maturity credit spreads), they lack a structural definition of the default event because the hazard rate

is specified as an exogenous process. Therefore no explanation of the presence of structural changes in firm-specific variables is given.

Madan and Unal (2000) bridge the gap between the two streams of literature. They propose a structural hazard rate model in closed form: default is assumed to be a consequence of a single jump loss event that drives the equity value to zero and requires cash outlays that cannot be externally financed. Therefore in this model, default arises from the arrival of an unforeseen event, e.g. the outcome of lawsuits, sudden default of a creditor, supplier or a customer and catastrophes in production lines. Moreover they assume that the defaultable bond with face value 1, maturity T in case of default at time τ pays *at maturity* the amount $q = 1 - y$. From (1.13) it follows:

$$\overline{B}(t,T) = \mathbb{E}_t^Q\left[\exp\left(-\int_t^T r_u \mathrm{d}u\right)(1 - yI_{\{\tau \leq T\}})\right].$$

While Jarrow and Turnbull (1995) evaluate this expectation assuming the independence under Q between the interest rate process and the default process, Madan and Unal (2000) allow for dependence between the two and use the rikless bond as a numeraire, rather than the money market account. In other words they use the T-forward measure[10], Q^T, and obtain

$$\begin{aligned}\overline{B}(t,T) &= B(t,T)\mathbb{E}_t^{Q^T}\left[1 - yI_{\{\tau \leq T\}}\right] \\ &= B(t,T)\left[1 - yQ^T(\tau \leq T)\right].\end{aligned}$$

As a consequence

$$\frac{\overline{B}(t,T)}{B(t,T)} = Q^T(\tau > T) + \left[1 - Q^T(\tau > T)\right]q.$$

Now, it is easy to show that (1.38) is true also under Q^T because of the continuity of $\mathrm{d}Q/\mathrm{d}Q^T = B(t,T)/\beta(t)$

$$\begin{aligned}\frac{U_t}{\beta(t)} := G(t,T) &= \mathbb{E}_t^Q\left[\exp\left(-\int_t^T h_u \mathrm{d}u\right)\right] \\ &= \mathbb{E}_t^{Q^T}\left[B(t,T)/\beta(t)\exp\left(-\int_t^T h_u \mathrm{d}u\right)\right].\end{aligned}$$

Since

$$\mathbb{E}_t^{Q^T}\left[B(t,T)/\beta(t)\exp\left(-\int_t^T h_u \mathrm{d}u\right)\right] = \frac{U_t}{B(t,T)}$$

it follows

$$G(t,T) = \mathbb{E}_t^{Q^T}\left[\exp\left(-\int_t^T h_u \mathrm{d}u\right)\right] = \mathbb{E}_t^Q\left[\exp\left(-\int_t^T h_u \mathrm{d}u\right)\right]. \tag{1.39}$$

Therefore, according to (1.39) the survival probability is obtained as the expected value of the promised payoff of the corporate bond discounted at the hazard rate. In contrast, the price of the riskless bond is given by the expected value of the promise payoff discounted by the risk-free rate r_t. Since

$$\overline{B}(t,T) = B(t,T) - B(t,T)y[1 - G(t,T)]$$

and thus

$$B(t,T)y[1 - G(t,T)] = \overline{B}(t,T) - B(t,T),$$

---

[10] Under this measure every claim settling at time T' with payoff X has price in $t < T'$

$$U_t = \begin{cases} B(t,T)\mathbb{E}_t^{Q^T}[B^{-1}(T',T)X] & \text{if } T' \leq T \\ B(t,T)\mathbb{E}_t^{Q^T}[B^{-1}(T,T')X] & \text{if } T' \geq T \end{cases}.$$

the quantity $1 - G(t,T) = Q^T (\tau \leq T)$ can be thought as the "price of default" for a given loss level y.

Next, the assumption that the firm faces at the random time z the payment of the amount D is made. Default is triggered if D dominates the existing continuously evolving equity, S. The equity value of the firm is given by the value of its assets less its liabilities. The firm holds cash assets with market value V, that does not depend on the current level of interest rates and other assets depending on r with current market value $g(t,r)$. Therefore

$$S = V + g(t,r) - \bar{v}(t,r),$$

where $\bar{v}(t,r)$ denotes the liabilities obtained as the present value of promised payment discounted at the risk-free rate r. When the terms $g(t,r)$ and $\bar{v}(t,r)$ have different sensitivities to r a duration gap arises. Default arises when $D > S_T$ and a payment at time T is due, i.e. when a loss of magnitude D dominates the existing equity. Now note that

$$h_t = h(V,r) = \frac{P\left[z \in (t, t+\mathrm{d}t) \quad \text{and} \quad D \geq S\right]}{\mathrm{d}t}$$

and assume that the loss arrival process is Poisson with constant rate (per unit time) λ and the loss distribution, $F_D(\cdot)$ is exponential with mean loss μ_D. It follows

$$h(V,r) = \lambda\left[1 - F_D(S)\right] = \lambda\left[1 - F_D(V + g(t,r) - \bar{v}(t,r))\right].$$

A first order approximation of this function around $\log V_0$ and r_0 yields

$$h(V_t, r_t) \simeq h(V_0, r_0) - \lambda f_D(S_0) V_0 (\Delta \log V) - \lambda f_D(S_0)(g_r^0 - \bar{v}_r^0)(\Delta r)$$

with $\Delta \log V = \log V_t - \log V_0$, $\Delta r = r_t - r_0$ and $g_r^0 - \bar{v}_r^0$ denotes the partial of $g - \bar{v}$ with respect to r evaluated at r_0. Since

$$f_D(u) = \frac{1}{\mu_D} \exp\left(-\frac{d}{\mu_D}\right),$$

and simple algebra yields

$$h_t = a - b \log V_t + c r_t \tag{1.40}$$

where

$$a = \lambda\left[1 - F_D(S_0)\right] + b \log V_0 - c r_0,$$
$$b = \frac{\lambda}{\mu_D} \exp\left(-\frac{S_0}{\mu_D}\right) V_0,$$
$$c = -\frac{b}{V_0}(g_r^0 - \bar{v}_r^0) = \frac{b}{V_0}\mathscr{D}$$

and $\mathscr{D}$ is the firm's net asset duration.

Equation (1.40) represents a two factor structural hazard rate model. The two factor are the firm's cash asset, V, and the level of default free interest rate. The parameters involved may be interpreted in terms of firm specific features such as the equity level and the duration $\mathscr{D}$.

Assuming a geometric Brownian motion for the firm's non-financial asset value

$$\mathrm{d}V = rV\mathrm{d}t + \sigma_V V \mathrm{d}B_V$$

and a Vasicek interest rate model

$$\mathrm{d}r = \alpha(\theta - r)\mathrm{d}t + \sigma_r \mathrm{d}B_r,$$

with

$$\rho\mathrm{d}t = \mathrm{d}B_V \mathrm{d}B_r,$$

simple closed forms are derived for the prices of defaultable coupon bonds.

1.3   Fractional Brownian Motion and Long Range Dependence in Finance

Fractional Brownian motion has been introduced in the seminal paper Mandelbrot and Van Ness (1968) already in late sixties and has been widely studied for instance in Sottinen and Valkeila (2001) and Sottinen (2001). Before stating its properties it is important to introduce formally the concept of long range dependence. Let $X = \{X_n\}_{n \in \mathbb{N}}$ be a stationary process with autocorrelation function (ACF), $\rho(\cdot)$. We say that X exhibit the statistical property of long range dependence if

$$\lim_{n \to \infty} \frac{\rho(n)}{c_\rho n^{-\alpha}} = 1. \tag{1.41}$$

This means that the dependence between X_t and X_{t+n} varies slowly as n tends to infinity. In particular the autocorrelation function decays at a hyperbolic rate, which is much slower than the geometric rate of processes such as finite-order stationary ARMA processes. Sometimes the long range dependence is defined via the equivalent requirement:

$$\sum_{n=-\infty}^{\infty} \rho(n) = \infty,$$

which means that the ACF decays so slowly that the autocorrelations are not summable. Fractional Brownian motion is part of a more general class of processes that exhibit long range dependence under certain conditions. These process, called self-similar processes, were introduced by Kolmogorov (1961) as models for turbulence. A centered process $X = \{X_t\}_{t \in [0,T]}$ is said to be self-similar with Hurst exponent H if X_t has the same distribution of $a^{-H} X_{at}$ for all $a > 0$ i.e.

$$X_t \overset{d}{=} a^{-H} X_{at} \quad \forall a > 0.$$

Alternatively one can require that the finite-dimensional distributions satisfy

$$(T^H X_{t_1}, \ldots, T^H X_{t_n}) \overset{d}{=} (X_{Tt_1}, \ldots, X_{Tt_n})$$

for every $T > 0$, any choice of $t_i \geq 0$, $i = 1, \ldots, n$ and $n \geq 1$. If, in addition, the process is square integrable and has stationary increments,

$$\mathrm{Cov}(X_t, X_s) = \frac{\mathrm{Var}(X_1)}{2} \left(t^{2H} + s^{2H} - |t - s|^{2H} \right). \tag{1.42}$$

In fact, assuming for simplicity $\mathbb{E}(X_t) = 0$, it follows

$$\mathrm{Cov}(X_t, X_s) = \frac{1}{2} \left\{ \mathbb{E}(X_t^2) + \mathbb{E}(X_s^2) - \mathbb{E}\left[(X_t - X_s)^2\right] \right\}$$
$$= \frac{1}{2} \left\{ \mathbb{E}(X_t^2) + \mathbb{E}(X_s^2) - \mathbb{E}\left(X_{t-s}^2\right) \right\},$$

which implies (1.42).
Because of (1.42), one usually assumes $H \in (0,1)$. The cases $H < 0$ and $H > 1$ are impossible, the former implying $\mathrm{Var}(X_0) = \infty$ because from (1.42) $\mathrm{Var}(X_t) = t^{2H}\mathrm{Var}(X_1)$ and the latter implying

$$\mathrm{Corr}(X_t, X_1) = \frac{t^{2H} - (t-1)^{2H}}{t^H} > 1$$

for t big enough. Moreover the cases $H = 0$ and $H = 1$ are not interesting because they imply respectively $X = 0$ identically and $\mathrm{Corr}(X_t, X_1) = 1$ for t big enough.
To show the connection to long range dependence let X be a centered square integrable process with stationary

increments. Consider the process $\{Y_n\}_{n\in\mathbb{N}}$ with $Y_n = X_n - X_{n-1}$. Then the autocorrelation function for the stationary increments is

$$\rho(n) = \frac{\text{Cov}(Y_{t+n}, Y_t)}{\text{Var}(Y_t)} = \frac{\text{Cov}(X_{t+n} - X_{t+n-1}, X_t - X_{t-1})}{\text{Var}(X_t - X_{t-1})}$$

$$= \frac{1}{2}\left[(n+1)^{2H} + (n-1)^{2H} - 2n^{2H}\right].$$

It follows

$$\lim_{n\to\infty} \frac{\rho(n)}{H(2H-1)n^{2H-2}} = \lim_{m\to 0} \frac{(1+m)^{2H} + (1-m)^{2H} - 2}{2H(2H-1)m^2}$$

$$\stackrel{H\,op.}{=} \lim_{m\to 0} \frac{(1+m)^{2H-1} - (1-m)^{2H-1}}{2(2H-1)m}$$

$$\stackrel{H\,op.}{=} \lim_{m\to 0} \frac{(1+m)^{2H-2} + (1-m)^{2H-2}}{2} = 1.$$

Therefore it $H \in \left(\frac{1}{2}, 1\right)$, the increment process $\{Y_n\}_{n\in\mathbb{N}}$ exhibit long range dependence with $\alpha = 2 - 2H$ and $c_\rho = H(2H-1)$.

It is worth noticing, however, that there exist centered processes with stationary and independent increments satisfying the self-similarity property for $H > \frac{1}{2}$, for instance the α-stable Lévy processes with $H = \frac{1}{\alpha}$. These processes, however, have infinite variance.

1.3.1   Properties

Now we can introduce formally the notion of fractional Brownian motion. It can be thought as a moving average of $\mathrm{d}B_t$, in which the past increments of the standard Brownian motion are weighted by the kernel $(t-s)^{H-\frac{1}{2}}$. In particular the fractional Brownian motion admits the following representation:

$$B_{H,K}(t) = K \cdot V_H^{1/2} \int_{\mathbb{R}} f_t(s)\mathrm{d}B_s, \tag{1.43}$$

with

$$f_t(s) = \frac{1}{\Gamma(H+1/2)} \left\{|t-s|^{H-1/2} I_{(-\infty,t]}(s) - |s|^{H-1/2} I_{(-\infty,0]}(s)\right\}$$

where

$$V_H = \Gamma(2H+1)\sin(\pi H)$$

is a normalization factor such that $\mathbb{E}\left[(B_{H,K}(t) - B_{H,K}(0))^2\right] = K^2$. K is a scale factor and for $K = 1$ the process is called standard and denoted by B_H. Thus the fractional Brownian motion is a Gaussian process with zero mean, stationary increments, variance

$$\mathbb{E}\left[B_{H,K}^2(t)\right] = K^2 t^{2H}$$

and covariance

$$\mathbb{E}\left[B_{H,K}(t)B_{H,K}(s)\right] = \frac{K^2}{2}\left(t^{2H} + s^{2H} - |t-s|^{2H}\right).$$

Therefore the parameter H negotiates whether the fractional Brownian motion has independent increments ($H = 1/2$), positive covariance between two increments over non-overlapping time intervals ($1/2 < H \le 1$), or negative covariance between increments ($0 < H < 1/2$). Another possible representation of the standard fractional Brownian motion, that involves integrating over the finite interval $[0, t]$, is the following

$$B_H(t) = \int_0^t z_t(s)\mathrm{d}B_s, \tag{1.44}$$

where

$$z_t(s) = c \left[\left(\frac{t}{s} \right)^{H-1/2} (t-s)^{H-1/2} - \left(H - \frac{1}{2} \right) s^{1/2-H} \int_s^t u^{H-3/2} (u-s)^{H-1/2} \mathrm{d}u \right]$$

and c is a normalization constant depending on H. Another property, useful to define the stochastic integral with respect to the fractional Brownian motion is the so-called index of p-variation. Let Π: $0 = t_0 < t_1 < \ldots < t_n = T$ be a partition of the interval $[0, T]$. Let

$$v_p(B_H, \Pi) := \lim_{|\Pi| \to 0} \sum_{k=1}^n |B_H(t_k) - B_H(t_{k-1})|^p$$

be the p-variation of the fractional Brownian motion over the the interval $[0, T]$, where $|\Pi| = \max_{1 \le k \le n}(t_k - t_{k-1})$. The index of p-variation is simply the smallest $p > 0$ for which the p-variation is finite, i.e.

$$v(B_H, [0, T]) := \inf\{p > 0 : v_p(B_H, [0, T]) < \infty\}$$

if this set is non-empty and $v(B_H, [0, T]) = \infty$ otherwise. Now, it is possible to show that for $H \in \left[\frac{1}{2}, 1 \right)$ one has

$$v(B_H) = v(B_H, [0, T]) = \frac{1}{H}.$$

As a recap, here are reported the properties of the standard fractional Brownian motion B_H with self-similarity parameter $H \in \left(\frac{1}{2}, 1 \right)$:

- $B_H(t)$ has stationary increments;

- $B_H(0) = 0$ and $\mathbb{E}[B_H(0)] = 0$;

- $\mathbb{E}[B_H^2(t)] = |t|^{2H}$ for all t;

- $\mathbb{E}[B_H(t)B_H(s)] = \frac{1}{2} \left(t^{2H} + s^{2H} - |t-s|^{2H} \right)$ for all t and s;

- $B_H(t)$ is Gaussian;

- $B_H(t)$ has continuous sample paths;

- If $H \ne \frac{1}{2}$, then $B_H(t)$ is not a semimartingale;

- If $H \ne \frac{1}{2}$, then $B_H(t)$ is non-Markovian;

- $B_H(t)$ has index of p-variation $v(B_H) = \frac{1}{H}$.

Since for a semimartingale M, $v(M) \in [0, 1] \cup 2$, it follows that B_H is not a semimartingale unless $H = 1/2$. As a consequence one cannot hope to use the Ito theory to define the stochastic integral with respect to B_H. To solve this problem two different integrals have been proposed. The first is based on pathwise, i.e. ω-by-ω, integrals very close to the Riemann-Stieltjes integrals, which leads for "reasonable" integrands $f(t, \omega)$ to

$$\int_0^T f(t, \omega) \delta B_H(t) := \lim_{|\Pi| \to 0} \sum_{k=1}^n f(t_{k-1}, \omega)(B_H(t_k) - B_H(t_{k-1})).$$

The second one is called the Wick-Ito integral and has the form

$$\int_0^T f(t, \omega) \mathrm{d}B_H(t) := \lim_{|\Pi| \to 0} \sum_{k=1}^n f(t_{k-1}, \omega) \diamond (B_H(t_k) - B_H(t_{k-1})),$$

where $\diamond$ denotes the Wick product (see Hu and Øksendal, 2003). First of all, define

$$\phi(s,t) = H(2H-1)|s-t|^{2H-2}$$

and consider the space

$$L_\phi^2(\mathbb{R}) := \left\{ f : \mathbb{R} \to \mathbb{R} \quad \text{so that} \quad |f|_\phi^2 := \iint_{\mathbb{R}^2} f(s)f(t)\phi(s,t)\mathrm{d}s\mathrm{d}t < \infty \right\}$$

equipped with the inner product

$$(f,g)_\phi := \iint_{\mathbb{R}^2} f(s)g(t)\phi(s,t)\mathrm{d}s\mathrm{d}t.$$

It is possible to show that $L_\phi^2(\mathbb{R})$ is separable Hilbert space (i.e. has countable yet dense subsets) and that

$$\Gamma_\phi[f(u)] = c_H \int_u^\infty (t-u)^{H-3/2} f(t)\mathrm{d}t$$

with $c_H = \sqrt{\frac{H(2H-1)\Gamma(\frac{3}{2}-H)}{\Gamma(H-\frac{1}{2})\Gamma(2-2H)}}$, is an isometry from $L_\phi^2(\mathbb{R})$ to $L^2(\mathbb{R})$. To define the integration of the function $f \in L_\phi^2(\mathbb{R})$, let the sequence of approximating functions $\{f_m(t)\} \stackrel{m\to\infty}{\longrightarrow} f(t)$ be $f_m(t) = \sum_i a_i^{(m)} I_{[t_i,t_{i+1}]}(t)$. For these simple integrands set

$$\int_\mathbb{R} f_m(t)\mathrm{d}B_H(t) = \sum_i a_i^{(m)}(B_H(t_{i+1}) - B_H(t_i)).$$

Define

$$\int_\mathbb{R} f(t)\mathrm{d}B_H(t) = \int_\mathbb{R} f(t)\delta B_H(t) = \lim_{m\to\infty} \int_\mathbb{R} f_m(t)\mathrm{d}B_H(t).$$

The limit exists in $L^2(\mu_\phi)$ because of the isometry

$$\mathbb{E}\left[\left(\int_\mathbb{R} f_m(t)\mathrm{d}B_H(t)\right)^2\right] = |f_m|_\phi^2,$$

where μ_ϕ is the probability law of B_H.

1.3.2 Long Range Dependence in Finance

Although introduction of the Geometric Brownian motion-based Black-Scholes formulation of vanilla options by Black, Scholes, and Merton marked the advent of mathematical finance, it is well known that their model has many limits. Remember that, the Black & Scholes pricing model assumes the following dynamics for the asset S:

$$\mathrm{d}S_t = S_t\left(\mu\mathrm{d}t + \sigma\mathrm{d}B_t\right), \tag{1.45}$$

where B_t is a standard Brownian motion. As a consequence the log-returns

$$R_{t_k} := \log\frac{S_{t_k}}{S_{t_{k-1}}} = \left(\mu - \frac{\sigma^2}{2}\right)(t_k - t_{k-1}) + \sigma(B_{t_k} - B_{t_{k-1}})$$

are independent normal random variables. However, in recent years, many empirical evidence have shown that this model cannot describe the behaviour of financial assets. There are two main reasons to explain why: the empirical distributions of the log-price variations are far from being normal and the actual returns show some form of dependence. To explain the first reason, many empirical studies indicate that log-returns

are not normal especially when the observation intervals $t_k - t_{k-1}$ are short: after the 1987 market crash, industry and researchers began to take note of the heavy-tail distribution of financial assets and several models has been developed using for instance heavy-tailed Lévy processes. The second problem leads to long-range dependence. For many decades, the general consensus has been to assume a Markovian process for the asset, implying that all information is contained within current asset price. However, many studies of financial time series indicate that the stock market prices exhibit the so-called long range dependency property. To deal with this feature, stochastic volatility models have been introduced. Even though these models can produce quasi long-range dependence, they are very difficult to implement and often do not lead to tractable pricing solutions as they lead to high-dimensional PDE's with variable coefficients. On the other hand, fractional Brownian motion offers a natural way of modelling long-range dependence giving tractable solutions to pricing financial instruments. The dependence structure of log-returns is therefore described via the Hurst parameter H. The assumption of a geometric Brownian motion (gBm) for the firms' underlying assets, eq. (1.1), proposed in the original paper by Merton could be the explanation for the poor performance of structural model in predicting credit spreads.

With this problem in mind, Della Ratta and Urga (2005) try to understand whether credit spreads are short or long memory processes, and whether, in the latter case, they are stationary or nonstationary. The theoretical model for the dynamics of the credit spread is described by the following stochastic differential equation:

$$ds_t = \mu(s_t, t)dt + \sigma(s_t, t)dB_H(t), \tag{1.46}$$

where s_t is the relevant spread, μ_t and σ_t are the (possibly time-varying) drift and diffusion term respectively, and B_H is a standard fractional Brownian motion characterized by a slowly decaying autocorrelation function depending on the Hurst exponent $H \in (0, 1]$. Their dataset consists of daily observations for the 30-year Historical US Treasury Constant Maturity Yields and Moody's Aaa, Aa, A and Baa Long-Term Corporate Bond Yield Averages over the period from December 1992 to November 2003. The latter are average yields calculated by Moody's for bonds in a given rating class. Spreads are calculated as the difference between corporate yields and Treasury yields, as well as between different corporate yields. First difference of the series involved are also taken into account. The use of rating-specific indices rather than data on individual issues is justified by the fact that the market for individual corporate bonds is often illiquid which can make credit spreads not a perfect indicator of credit risk. Moreover, though individual bonds may exhibit inconsistencies between spreads and ratings because ratings adjust slowly to changes in firm value, on average ratings tend to be good indicators of credit quality.

First they show, using a Jarque-Bera test that all the series involved are far from being normal. Moreover differences in credit spreads, are positively skewed, implying that the loss tail of the return distribution contains more probability than the normal distribution, and leptokurtic. This last feature implies a small chance of very large returns combined with a large probability of small returns, which is very common in credit risk.

Besides, using the Ljung-Box test it is shown that yields and spreads are not serially correlated. As far as first differences are concerned, yield differences are not serially correlated, whereas spread differences have a significant negative first order autocorrelation coefficient.

Two empirical counterparts for the model (1.46) are considered. The first is based on a fractional white noise for a discrete time series:

$$y_t - \mu = (1 - L)^{-d}\epsilon_t, \quad \epsilon_t \sim WN(0, \sigma^2) \tag{1.47}$$

where L is the lag operator, i.e. $Ly_t = y_{t-1}$, $\{\epsilon_t\}_t$ is a sequence of $i.i.d.(0, \sigma^2)$ random variables (White Noise) and d is the difference parameter which is allowed to take on non-integer values only in the interval $(-1/2, 1/2)$ to guarantee covariance stationarity. Clearly $y_t \sim I(d)$. When a fractionally integrated series y_t has long memory, the Hurst parameter can be related to the difference parameter in (1.47) via

$$d = H - \frac{1}{2}.$$

If $d = 0$, or equivalently $H = \frac{1}{2}$, $y_t = \epsilon_t$ and thus the process is serially uncorrelated, whereas if $d > 0$, implying $H > \frac{1}{2}$, the process has long memory. Note that assuming $\mu = 0$ the process (1.47) can be written

in another form. Using the binomial theorem for integer powers

$$(1 - L)^d = \sum_{k=0}^{\infty} (-1)^k \binom{d}{k} L^k$$

and allowing d to be non-integer

$$(1 - L)^d = \sum_{k=0}^{\infty} \frac{\Gamma(k - d)}{\Gamma(-d)\Gamma(k + 1)} L^k$$

and thus a fractional white noise process becomes

$$\sum_{k=0}^{\infty} \frac{\Gamma(k - d)}{\Gamma(-d)\Gamma(k + 1)} y_{t-k} = \epsilon_t.$$

Alternatively it can be seen as an infinite order MA process:

$$y_t = B(L)\epsilon_t$$

with

$$B(L) = \sum_{k=0}^{\infty} \frac{\Gamma(k + d)}{\Gamma(d)\Gamma(k + 1)} L^k.$$

In every case the autocovariance is

$$\gamma_k = \frac{\sigma^2 \Gamma(1 - 2d)\Gamma(k + d)}{\Gamma(d)\Gamma(1 - d)\Gamma(k + 1 - d)}.$$

Note that

$$\gamma_k \simeq ck^{2d-1}$$

for some constant c as $k \to \infty$.

The model (1.47) can be generalized to the Autoregressive Fractionally Integrated Moving Average process. An ARFIMA(p, d, q) process is defined as

$$\phi(L)(1 - L)^d(y_t - \mu) = \theta(L)\epsilon_t, \tag{1.48}$$

where $\phi(L)$ and $\theta(L)$ involve autoregressive and moving average coefficients of order p and q respectively:

$$\phi(L) = 1 - \sum_{i=1}^{p} \phi_i L^i$$

$$\theta(L) = 1 + \sum_{i=1}^{q} \theta_i L^i,$$

with roots lying outside the unit circle and d can be non-integer. Besides, an ARFIMA$(0, d, 0)$ is equivalent to model (1.47). Again for $d \neq 0$ the process exhibits long memory. In particular, if $d > 0$, the sum of the autocorrelations diverges to infinity, the dependence is positive and the long memory is called persistent. On the contrary, when $d < 0$ the sum of the autocorrelations converges to zero, the dependence is negative and the long memory is called antipersistent (or mean reverting).

To test for the presence of long memory in the series of yields and spreads, the method suggested in Geweke and Porter-Hudak (1983) is typically used. The main idea of this procedure is to assume that the the spectral density[11] in the neighbourhood of zero frequency is

$$f(\lambda) = C \left[4 \sin^2 \left(\frac{\lambda}{2} \right) \right]^{-d} = C \left| 1 - e^{-i\lambda} \right|^{-2d}, \tag{1.49}$$

---

[11] The spectral density is defined as the Fourier transform of the autocovariance function $\gamma(n)$:

$$f(\lambda) = \frac{1}{2\pi} \sum_{n=-\infty}^{+\infty} \gamma(n)e^{-in\lambda},$$

for some unknown constant C. Now consider the Fourier transform and the periodogram

$$\omega(\lambda) = \frac{1}{\sqrt{2\pi T}} \sum_{t=1}^{T} y_t e^{it\lambda} \tag{1.50}$$

$$I(\lambda) = |\omega(\lambda)|^2 . \tag{1.51}$$

where T is the number of observations and evaluate these quantities and (1.49) at the harmonic ordinates

$$\lambda_j = \frac{2\pi}{T} j. \tag{1.52}$$

Taking the logarithm of both sides of (1.49) and adding $\log I(\lambda_j)$ to both sides one gets the following regression model:

$$\log f(\lambda_j) = c + da_j + U_j, \quad j = 1, 2, \ldots, m$$

where m is a bandwidth parameter,

$$U_j = \log \left\{ \left[4 \sin^2 \left(\frac{\lambda_j}{2} \right) \right]^d I(\lambda_j)/C \right\} + \eta$$

is the disturbance term,

$$c = \log C - \eta$$

and

$$a_j = -\log \left[4 \sin^2 \left(\frac{\lambda_j}{2} \right) \right]. \tag{1.53}$$

$\eta = 0.5772$ the Euler-Mascheroni constant[12]. Hence the fractional difference parameter is estimated by least squares. Geweke and Porter-Hudak (1983) showed that using a periodogram estimate of $f(\lambda_j)$, the least squares estimate $\hat{d}$ using the above regression is normally distributed in large samples if $m = T^\alpha$ with $0 < \alpha < 1$:

$$\hat{d} \sim N \left(d, \frac{\pi^2}{6 \sum_{j=1}^{m} (a_j - \bar{a})^2} \right)$$

where

$$\bar{a} = \frac{\sum_{j=1}^{m} a_j}{m}.$$

Under the null hypothesis of no long memory ($d = 0$), the statistic

$$\hat{d} \left(\frac{\pi^2}{6 \sum_{j=1}^{m} (a_j - \bar{a})^2} \right)^{-1/2}$$

has a limiting standard normal distribution.

Denote by d' the long memory parameter of the first differences of the series and by T the number of

---

and vice versa

$$\gamma(n) = \int_{-\pi}^{\pi} f(\lambda) e^{in\lambda} d\lambda.$$

Note that since $\forall n \in \mathbb{N}$ one has $\gamma(n) = \gamma(-n)$, $\sin(-n\lambda) = -\sin(n\lambda)$, $\cos(-n\lambda) = \cos(n\lambda)$ and $e^{in\lambda} = \cos(n\lambda) + i\sin(n\lambda)$, we can rewrite the spectral density as:

$$f(\lambda) = \frac{1}{2\pi} \left[\rho(n) + 2 \sum_{n=1}^{+\infty} \gamma(n) \cos(n\lambda) \right].$$

[12] It holds

$$\eta = \lim_{n \to \infty} \left(\sum_{k=1}^{n} \frac{1}{k} - \log k \right).$$

observations. It follows that the long memory parameter of the series in levels is $d = d' + 1$. Della Ratta and Urga (2005) test the null hypothesis that $d' = 0$ and show show that the estimated value of d' for the first differences of yields and spreads is not statistically different from zero when $m = T^{0.5}$ or $m = T^{0.6}$ and the null $d' = 0$ is therefore not rejected. Thus the estimated values for d are not statistically different from one, implying that yields and spreads are long memory and nonstationary processes. However, by looking at the point estimates, they notice that the d' estimate for the first differences of yields is generally negative, whilst the d' estimate for the first differences of spreads is generally positive. Therefore first differences of spreads are more likely to be long memory than first differences of yields.

To have a confirmation of this, the procedure proposed in Robinson (1995b) is considered for the first differences of the series. In Robinson (1995b) the fractional difference parameter is estimated by least squares using the regression:

$$\log f(\lambda_j) = c + da_j + U_j, \quad j = l, 2, \ldots, m$$

The parameter l is such that $0 \leq l < m < T$ and is chosen to trim low frequencies out. The corresponding estimator is called log-periodogram regression estimator (LPE) and is given by

$$\hat{d}(l) = \sum_{j=l+1}^{m} (a_j - \bar{a}) \log I(\lambda_j) / S_l$$

with a_j given by (1.53) and

$$\bar{a} = \frac{\sum_{j=l+1}^{m} a_j}{m - l},$$

$$S_l = \sum_{j=l+1}^{m} (a_j - \bar{a})^2.$$

It can be shown that this estimator is more efficient than the GPH estimator and robust to nonnormality. Results obtained in this way are consistent with the GPH results and show that yields and spreads are long memory non stationary and their first differences are stationary.

To estimate the parameter d in the ARFIMA(p, d, q) framework the AML (approximate maximum likelihood) procedure of Haslett and Raftery (1989) is used. Note that this procedure as well as the one proposed by Sowell (1992) (exact maximum likelihood) can be used for a stationary ARFIMA model $(-1/2 < d < 1/2)$. However, for many economic and financial time series, the data usually seem to lie on the borderline separating stationarity from non-stationarity. As a consequence, one usually needs to decide whether or not to difference the original time series before estimating a stationary ARFIMA model. To solve this problem Beran (1995) extended the estimation of ARFIMA models for any $d > -1/2$ by considering the following model:

$$\phi(L)(1 - L)^{\delta}((1 - L)^n y_t - \mu) = \theta(L)\epsilon_t,$$

where $-1/2 < \delta < 1/2$ $\phi(L)$ and $\theta(L)$ are defined as above. The integer n is the number of times that y_t must be differenced to achieve stationarity. Thus the difference parameter is given by $d = \delta + n$. The choice $n = 0$ or $n = 1$ is usually satisfactory for modelling economic and financial time series.

For each yield and spread series and their first differences the long memory parameter d is estimated for the nine combinations of ARIMA(p, d, q) models where p and q are between 0 and 2. Then the model which minimises the Bayesian Information Criterion is chosen. From the empirical results, the best model for the series of yields is an ARFIMA$(0, 1, 0)$, which confirms the long memory nonstationary property of these series. As a consequence the best model for the first differences of yields is an ARFIMA$(0, 0, 0)$, implying that these series are short memory stationary. As far as the series of spread are concerned, the best model is an ARFIMA$(2, d, 0)$, with d always statistically greater than 1. The series of spreads over Treasury are therefore long memory nonstationary processes. As a consequence the differenced series are long memory stationary processes and this is confirmed by the fact that the best model for the first differences of the spreads over Treasury is an ARFIMA$(2, d, 0)$, where d is statistically greater than 0 and less than $1/2$. However, for the

series of spreads between different corporate yields the authors seem to get contradictory results. In fact only one series over six is a long memory stationary process while the other series are short memory stationary processes. Thus these results seem to contradict the one obtained from the GPH and Robinson analysis but the authors justify by noticing that for each of the spread series the sum of the AR coefficients is close to unity, which indicates that the nonstationarity feature highlighted in the GPH and Robinson results has in fact been captured by the AR coefficients rather than by the d parameter in the ARFIMA framework. Finally the first differences of spreads between corporate yields are all short memory stationary processes, with some differences in the estimated long memory parameter (in three cases d is not statistically different from zero and in the other three cases $-\frac{1}{2} < d < 0$). Moreover the authors point out that, since 9 out of 10 spread series are driven by an ARMA$(1, 1)$, and in light of estimation for d, it seems reasonable to assume that spreads are driven by a mixture of short memory and long memory processes.

Credit spread indices are investigated also in Prigent et al. (2001). The data used in this paper consist of 3561 daily observations from January 1986 to the end of March 2000. Moody's indices for seasoned corporate bond yields with a Aaa or Baa rating and the 10 year (constant maturity) Treasury yield constructed by the Federal Reserve have been collected. As usual spreads are calculated as the difference between corporate and Treasury yields. As in Della Ratta and Urga (2005), the authors prefer to model directly the spreads using equation (1.46). The justification of this choice is that when one constructs first a model of risk-free rates and a model of yields on defaultable securities, and then derives spreads as a difference between the two, error in both models add up and this may lead to an inaccurate description of the process. Moreover modelling directly the spreads allows to capture both the liquidity and the credit risk component of credit spreads. To determine the behaviour of the drift and diffusion terms nonparametric techniques are applied. First of all the densities of the spread processes are estimated using a Gaussian kernel estimator:

$$\hat{f}_i(x) = \frac{1}{n h_i} \sum_{t=1}^{n} \phi \left(\frac{x - s_t^i}{h_i} \right),$$

where s_t^i is the time t spread for a bond in the class i, $\phi(\cdot)$ is the standard normal density function, h_i is the window width and n the number of observations. The window width is chose as

$$h_i = c \hat{\sigma}_i n^{-1/5}$$

where c is an arbitrary constant (depending on the level of smoothness one is willing to achieve for the density) and $\hat{\sigma}_i$ is the empirical standard deviation of spreads of class i. Next, the estimators corresponding to first-order and second-order approximation proposed in Stanton (1997) for the drift term and the diffusion are evaluated. From the analysis of these result it is clear that the drift is not constant in the level of spreads but tend to decrease with spread levels. Since the value of the drift is close to zero for values of the spread around their mean, there is clear evidence of mean reversion. Moreover they find that the mean reversion is faster in Aaa bonds than in Baa bonds. The justification is that Aaa spreads are explained in a greater proportion by liquidity and therefore they revert more quickly to their average than lower rate bonds. This fact has been acknowledged by many practitioners who suggest that, when one expects the end of a crisis (where spreads are far above their long-term mean) it is reasonable to invest first in Aaa bonds (which recover faster) and then move progressively to more speculative securities.

Also a parametric estimation of corporate credit spread is proposed following the model of Chan et al. (1992):

$$s_{t+1}^i - s_t^i = \alpha_D + \beta_D s_t^i + \sigma_D \left| s_t^i \right|^{\gamma_D} \epsilon_{t+1},$$

where ϵ_{t+1} are assumed to be i.i.d. normal variables. When $\alpha_D > 0$ and $\beta_D < 0$ this is a mean reverting process with $-\beta_D < 0$ as speed of mean reversion and $-\alpha/\beta_D$ as long term mean. The assumption of normality makes the estimation of the parameters quite straightforward.

To conclude, the paper proposes a new model for credit spread indices based on these empirical findings and which guarantees the positivity of the spreads and at the same time captures jumps and mean reversion:

$$\mathrm{d}Y_t = \alpha(\theta - Y_t)\mathrm{d}t + \sigma \mathrm{d}B_t + \mathrm{d}N_t,$$

where $Y_t = \log s_t$, B_t is the standard Brownian motion and N_t is a compound Poisson process.

Using more or less the same econometric methodology of Della Ratta and Urga (2005), Cheng (2004) investigates long memory dynamics in the daily and weekly exchange rates of six Asia Pacific countries: Australia, Japan, New Zealand, Singapore, South Korea and Taiwan. Three semiparametric frequency domain estimators, GPH, LPE and the Gaussian semiparametric estimator (GSE) proposed by Robinson (1995a), and an exact maximum likelihood estimator are used assuming that the processes involved are described by a fractional white noise. Af far as the GSE technique is concerned, as in (1.49) it is assumed that

$$\lim_{\lambda \to 0^+} \frac{f(\lambda)}{G\lambda^{1-2H}} = 1$$

for $0 < G < \infty$ (recall that $d = H - 1/2$). Clearly, this is a consequence of (1.41) and implies that for a long memory process the spectral density tends to infinity at zero frequency. Now consider the objective function

$$Q(G,H) = \frac{1}{m} \sum_{j=1}^{m} \left\{ \log G\lambda_j^{1-2H} + \frac{\lambda_j^{2H-1}}{G} I(\lambda_j) \right\},$$

where the bandwidth parameter m is an integer less than $n/2$. Parameters are estimated as

$$(\hat{G}, \hat{H}) = \arg \min_{\substack{0 < G < \infty \\ 0 < H < 1}} Q(G,H).$$

This implies that an approximate form of frequency domain Gaussian likelihood is maximized.

Although there exist considerable disparities between the results obtained by the four estimators in Cheng (2004), the author finds strong evidence of long memory in all the exchange rates considered in case of weekly data. On the other hand, he generally fails to detect long memory when daily exchange rate series are considered.

Many studies indicate positive long range dependence, i.e. $H > \frac{1}{2}$: for instance for the daily exchange rate between USD and JPY between January 1972 and December 1990 the estimated[13] Hurst index is $\hat{H} = 0.642$; Peters (1994) confirms positive long range dependence for the Dow Jones and the S&P500.

Crato and Rothman (1994) investigate the dynamics of Sterling pound quarterly exchange rates against the currencies of nine countries (US dollar, Japanese yen, deutschmark, French franc, Italian lira, Canadian dollar, Dutch guilder, Swiss franc and Swedish kroner) over the period January 1973- March 1990 and find long memory and mean reversion in five of the nine series studied.

Cheung and Lai (2001) examine for long memory dynamics in the yen-based real exchange rates of 8 industrialised countries and, using the frequency domain maximum likelihood estimator of Fox and Taqqu (1986) on monthly exchange rates from 1973 to 1997, find long memory and mean reversion behaviour in all 8 exchange rate series.

However Jacobsen (1996) does not find evidence of LRD investigating return series of stock indexes of some European country, U.S. and Japan. To reach these results the author uses first a test based on the rescaled range statistic proposed in Hurst (1951). This statistic is simply the range of partial sums of deviations of a time series from its mean, rescaled by its standard deviation. Let $\{y_t\}$ be a time series, for instance the series of the returns from a stock and $\bar{y}_n = \frac{1}{n} \sum_{t=1}^{n} y_t$ the average return after n periods. Consider the statistic R_n, defined by the difference between the maximum and the minimum accumulated deviation from the mean after n periods, i.e.

$$R_n = \max_{1 \le k \le n} \sum_{t=1}^{k} (y_t - \bar{y}_n) - \min_{1 \le k \le n} \sum_{t=1}^{k} (y_t - \bar{y}_n).$$

This statistic is made dimensionless by dividing by the usual standard deviation:

$$S_n = \sqrt{\frac{1}{n} \sum_{t=1}^{n} (y_t - \bar{y}_n)^2}.$$

---

[13] Comprehensive surveys for the techniques developed to estimate H can be found in Taqqu et al. (1995) and Taqqu and Teverovsky (1996).

Listing 1.1: Rescaled Range Statistic: Matlab Code

```
1  function V=rescaled_range(y)
2  n=length(y);
3  R=max(cumsum(y-mean(y)))-min(cumsum(y-mean(y)));
4  V=R/((n-1)^.5*std(y));
```

and R_n/S_n is called the rescaled range or simply the R/S statistic.
The asymptotic distribution of the normalized rescaled range

$$V_n = \frac{1}{\sqrt{n}} \frac{R_n}{S_n} = \frac{R_n}{\sqrt{\sum_{t=1}^{n} (y_t - \bar{y}_n)^2}}$$

under the assumption that the y_t are i.i.d. is

$$F_{V_n}(v) \xrightarrow{n \to \infty} F_V(v) = 1 + 2 \sum_{j=1}^{\infty} (1 - 4j^2 v^2) e^{-2(jv)^2}, \tag{1.54}$$

where the random variable V is the range of a Brownian bridge on the unit interval. Listing 1.1 shows the Matlab code to compute the Rescaled Range statistic.
It is possible to show that

$$\mathbb{E}(V) = \sqrt{\frac{\pi}{2}} \quad \text{and} \quad \mathbb{E}(V^2) = \frac{\pi^2}{6}.$$

Using the asymptotic p-values that can be found for instance in Lo (1991, Table II p. 1288), one can test the null hypothesis of no long term dependence. Although this procedure gives reliable results even if the time series exhibits skewness and kurtosis, it has a major drawback. The rescaled range is very sensitive to short range dependence. To see this, consider the simple form of short range dependence and let X_t be a stationary $AR(1)$ process:

$$X_t = \rho X_{t-1} + \eta_t, \quad \eta_t \sim WN\left(0, \sigma_\eta^2\right), \quad |\rho| \in (0,1).$$

It can be shown that under this $AR(1)$ process the limiting distribution of $V_n/\sqrt{n}$ is ξV, where $\xi = \sqrt{(1+\rho)/(1-\rho)}$. Thus under short term dependence, the statistic is biased implying that any incompatibility between the data and the predicted behaviour of the R/S statistic under the null hypothesis need not be attributed to the long memory property but may be a symptom of short term dependence. In particular Lo (1991) proves that in case of short time dependence, the limiting distribution of the rescaled range changes by a multiplicative constant depending on the short term dependence structure. To solve this problem the modified rescaled range is introduced. Let $\{y_t\}$ be a strong-mixing process[14]

$$y_t = \mu + \epsilon_t$$

where, as usual ϵ_t is a zero mean random variable. Under the null that $\{y_t\}$ is short term dependent, the asymptotic distribution of the statistic

$$V_{n,q} = \frac{R_n}{\sqrt{n}\sigma_{n,q}}$$

is again (1.54) and in particular

$$\lim_{n \to \infty} P\{V_{n,q} \in [.809, 1.862]\} = 0.95.$$

---

[14] See Lo (1991) for the details. However here it is sufficient to know that the strong-mixing property implies a form of asymptotic independence.

<div align="center">Listing 1.2: Modified Rescaled Range Statistic: Matlab Code</div>

```matlab
1   function V_q=mod_rescaled_range(y,q)
2   n=length(y);
3   deviations=y-mean(y);
4
5   R=max(cumsum(deviations))-min(cumsum(deviations));
6   for t=1:q
7       g(t)=deviations(t+1:n)'*deviations(t:n-1)/n;
8   end
9   g=g';
10
11  sigma_q=((n-1)/n*var(y)+2*(1-(1:q)/(q+1))*g)^.5;
12  V_q=R/(n^.5*sigma_q);
```

$V_{n,q}$ is called (normalized) modified rescaled range statistic (MRS) and the term in the denominator which replaces the standard deviation, is

$$\sigma_{n,q}^2 = \frac{1}{n}\sum_{t=1}^{n}(y_t - \bar{y}_n)^2 + \frac{2}{n}\sum_{t=1}^{q}w_t(q)\left[\sum_{s=t+1}^{n}(y_s - \bar{y}_n)(y_{s-t}-\bar{y}_n)\right]$$

$$= S_n^2 + 2\sum_{t=1}^{q}w_t(q)\hat{\gamma}_t,$$

where S_n^2 is the usual estimator of the sample variance, $\hat{\rho}$ is the estimator of the sample autocovariance of y

$$\hat{\gamma}_t = \frac{1}{n}\sum_{s=t+1}^{n}(y_s - \bar{y}_n)(y_{s-t}-\bar{y}_n)\quad t=1,\ldots,q$$

and

$$w_t(q) = 1 - \frac{t}{q+1}.$$

This means that the statistic $\sigma_{n,q}^2$ involves not only the sum of squared deviations of y_t but also its weighted autocovariances up to lag q. The weights are such that $\sigma_{n,q}^2$ is positive. Note that $\sigma_{n,q}^2$ is an estimator of 2π times the unnormalized spectral density function of y_t. The intuition behind this statistic is the following. If the series is subject to short term dependence, then the variance of the partial sum is not simply the sum of the variance, but also includes autocovariances. Listing 1.2 shows the Matlab code to compute the Modified Rescaled Range statistic.

The interval $[.809, 1.862]$ can be used as as the 95% (asymptotic) acceptance region for testing the null hypothesis

$$H_0 = \{\text{no long-range dependence, i.e., } H = 0.5\}$$

against the composite alternative

$$H_1 = \{\text{there is long-range dependence, i.e. } 1/2 < H < 1\}.$$

However this procedure presents the drawback that one has to choose the parameter q. If one chooses small values for q, not all the short term dependence is captured and hence the modified rescaled range statistic may continue to give biased result. On the contrary, choosing too large values for q may make the statistic V_q insensitive to long memory. This could influence both the actual size of the test, $P(\text{reject } H_0|H_0)$, and its power, $P(\text{reject } H_0|H_1)$. In particular Willinger et al. (1999) show that for large values of q,

$$V_{n,q} \simeq q^{1/2-H}.$$

This means that, when $H > 1/2$, the test statistic $V_{n,q}$ decreases as q increases and, for q large enough, it will be well within the confidence interval for the null hypothesis, i.e., $V_{n,q} \in [.809, 1.862]$.

To solve this problem one may compute the classical rescaled range statistics under particular models of short term dependence, for instance an $AR(1)$ or a $MA(1)$ model. To be more precise, first one imposes an $AR(1)$ or a $MA(1)$ model for the short term dependence and then applies the statistic $V_{n,0}$ to the time series of the corresponding residuals. Although in this case one takes a strong position on the form of short term dependence which need not be correct, this solution may be helpful when when the modified rescaled range statistic shows mixed results on the existence of long term dependence, i.e. when rejection of the null hypothesis of no long term dependence depends on the choice of q. Once the short term dependence model has been chosen, then it is possible to analyse the residuals for long term dependence using again the rescaled range statistic.

Going back to Jacobsen (1996), particular models of short term dependence are imposed to estimate the Hurst exponent using the graphical technique proposed by Mandelbrot and Wallis (1969), the so called R/S analysis. Hurst (1951) empirically found that, as $n \to \infty$ the relationship

$$\mathbb{E}\left[\frac{R_n}{S_n}\right] = a_1 n^H \tag{1.55}$$

holds for some positive finite constant a_1 that does not depend on n. For a short range dependent process it has been shown that

$$\mathbb{E}\left[\frac{R_n}{S_n}\right] = a_2 n^{1/2} \tag{1.56}$$

when $n \to \infty$. The discrepancy between (1.55) and (1.56) is generally referred to as the Hurst effect or the Hurst phenomenon. Therefore, it is possible to hypothesize a relationship of the form

$$\log\left[\frac{R_n}{S_n}\right] = \log a + H \log n. \tag{1.57}$$

The analysis based on (1.57) was proposed by Mandelbrot and Wallis (1969) and can be described as follows. First of all, the original sample of T observations is subdivided into K blocks, each of size $n = \lfloor T/K \rfloor$. For instance, if one has a series of 480 monthly returns, then it could be divided into $K = 80$ blocks considering six months lags ($n = 6$). The next step is to evaluate the rescaled range statistic in each interval $\{y_1, \ldots, y_n\}$, $\{y_{n+1}, \ldots, y_{2n}\}$, ..., $\{y_{jn+1}, \ldots, y_{(j+1)n}\}$, ..., $\{y_{(K-1)n+1}, \ldots, y_{nK}\}$ and average the values obtained. Let $Q(n)$ be this average, i.e.

$$Q(n) = \frac{1}{K} \sum_{j=0}^{K-1} \frac{R_{j,n}}{S_{j,n}},$$

where $R_{j,n}$ and $S_{j,n}$ denote the range and the standard deviation for the jth interval of length n:

$$R_{j,n} = \max_{1 \leq k \leq n} \sum_{t=1}^{k} (y_{jn+t} - \bar{y}_j) - \min_{1 \leq k \leq n} \sum_{t=1}^{k} (y_{jn+i} - \bar{y}_j)$$

$$S_{j,n} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_{jn+i} - \bar{y}_j)^2}.$$

and

$$\bar{y}_j = \frac{1}{n} \sum_{i=1}^{n} y_{jn+i}.$$

This procedure is repeated for different lags n. The Hurst exponent can hence be obtained as

$$H = \lim_{n \to \infty} \frac{\log[Q(n)]}{\log n}.$$

Listing 1.3: The graphical technique of Mandelbrot and Wallis (1969): Matlab Code

```matlab
function H=Mandelbrot_Wallis(y)
T=length(y);

for n=2:T
    K=floor(T/n);
    R_S=zeros(K,1);
    for j=1:K
    R_S(j)=rescaled_range(y((j-1)*n+1:j*n));
    end
    Q(n-1)=mean(R_S);
end
Q=Q';

B=regress(log(Q),[ones(T-1,1),log([2:1:T]')]);
H=B(2);
```

and its estimate can be obtained from the slope of the following regression

$$\log[Q(n)] = \alpha + \beta \log n \quad n = 2, \ldots, T$$

using ordinary least squares. See Listing 1.3 for the Matlab code.

It is well known that the range R_n for a sequence of independent Normal random variables approaches $\sqrt{n}$ for large n. As a consequence the Hurst exponent is asymptotically close to $1/2$ for an independent process. Again, in case of positive or negative long term dependence H will converge to values larger or smaller than $1/2$. In case of short term dependence H would converge to $1/2$. The R/S analysis has some advantages over other methods such as the analysis of autocorrelations (for instance it can be used if the time series exhibits large skewness and kurtosis), and, if compared to Lo's method that only indicates whether long-range dependence is present or not, it provides an estimate albeit rough of the Hurst parameter.

The dataset used in Jacobsen (1996) consists of monthly (continuously compounded) return series of stock market indexes of the Netherlands, Germany, the U.K., Italy, France, the U.S. and Japan. over the period December 1952 through December 1990 (456 observations) . Using the methodology described before for these data the author does not find evidence of long range dependence. Only in the cases of Italy and Germany the rescaled range statistic suggests long range dependence but the author explains that this is probably due to the presence of short term dependence. Besides he points out that the R/S analysis is biased in the case of short term dependence as well. In fact after adjusting for short term dependence the estimates for the Hurst exponent are lower than the ones for the original series.

The same result has been reached by Hiemstra and Jones (1997). Analysing $N = 1,952$ common stocks and using the modified rescaled range approach, they find little evidence of long memory. Their dataset consists of daily return of stocks listed on the New York and American Stock Exchanges over the period July 2, 1962 to December 31, 1991. Only those stocks for which there are 750 or more contiguous stock-return observations are included in the sample. Returns from firm i are adjusted for both dividends and stock splits and are given by

$$r_{i,t} = \frac{r_{i,t} \cdot f_{i,t} + d_{i,t}}{p_{i,t}} - 1, \quad \text{for } i = 1, \ldots, N \text{ and } t \in T_i = \{\tau_1, \ldots, \tau_{n_i}\}$$

where $n_i \geq 750$ denotes the length of the series corresponding to the firm, T_i the set of days for which returns are observed, $p_{i,t}$ the last sale price at time t, $d_{i,t}$ a cash adjustment and $f_{i,t}$ a price adjustment factor at time t. The MRS methodology is applied to the series using a variety of q-length, $q_i = \left\lfloor n_i^{1/4} \right\rfloor, \left\lfloor n_i^{1/3} \right\rfloor, \left\lfloor n_i^{1/2} \right\rfloor$ and

$$q_i^* = \left\lfloor \left(\frac{3n_i}{2}\right)^{1/3} \left(\frac{2\hat{\rho}}{1 - \hat{\rho}^2}\right)^{2/3} \right\rfloor$$

where $\lfloor x \rfloor$ denotes the largest integer less than or equal to x and $\hat{\rho}$ is, as usual, the first order sample autocorrelation estimated from the series. The last length is based on a data-dependent formula from Andrews (1991) and is proved to perform well when the data is generated by an $AR(1)$ process.

In Mandelbrot (1967) empirical studies based on the classical R/S analysis lead the author to suggest that representative values of the Hurst parameter for asset returns might be around $H = 0.55$. Mandelbrot's empirical findings were essentially confirmed by a later large-scale study by Greene and Fielitz (1977) who found significant long range dependence in many series of securities. In particular they used the classical R/S-method on 200 daily stock return series of securities listed on the New York Stock Exchange and reported that many of the series are characterized by long-range dependence. Lo (1991) himself, using his modified R/S-statistic and the CRSP daily stock returns from 1962 to 1987 (6400 observations), re-examined the question of long-run memory in asset returns raised in Greene and Fielitz (1977). Analyzing the entire series, as well as fractions of it (1/2 and 1/4 of the original series), and using $q = 90, 180, 270$ and 360 trading periods, Lo's main finding is that the daily stock returns do not exhibit long-range dependence. Moreover, Lo observed that the values of the test-statistic $V_{n,q}$ do not change much for the different values of q, which he takes as strong supporting evidence that the test statistic can be trusted. Lo attributes the findings of Greene and Fielitz (1977) to the fact that the classical R/S analysis is sensitive to the presence of short-range dependence and concludes that traditional short-range dependent models are adequate to describe actual stock returns. However the analysis of Lo is reconsidered in Willinger et al. (1999). They point out that, in his analysis of the CRSP data, Lo relies exclusively on the modified R/S-statistic, and, as pointed out earlier, $V_{n,q}$ may be not reliable and has a strong preference for accepting the null hypothesis of no long-run memory, if used without other methodologies. The authors using the same dataset as in Lo (1991), i.e. the daily CRSP data for the equal-weighted returns indices over the period 1962-1987, find an estimate of the Hurst parameter H of about 0.62. But as H-values are very low, the evidence is not absolutely conclusive. However they argue that Lo's conclusion is the result of the excessively conservative nature of his proposed test-statistic in rejecting the null hypothesis of no long-range dependence. More empirical studies can be found in Shiryaev (1999).

Another test to verify the hyphotesis of long range dependence has been studied by Lobato and Savin (1998). They employ an approximation to the Lagrange multiplier test and use the statistic

$$LM_T(m) = m \frac{\sum_{j=1}^{m} v_j |\omega(\lambda_j)|}{\sum_{j=1}^{m} |\omega(\lambda_j)|}$$

where m is a bandwidth parameter, $\omega(\lambda_j)$ and λ_j are given respectively by (1.50) and (1.52) and

$$v_j = \log j - \frac{\sum_{j=1}^{m} \log j}{m},$$

to test the null hypothesis $H_0 : H = 0.5$. The statistic has nice asymptotic properties: its asymptotic distribution is a χ_1^2.

It is also interesting to consider the estimator for H proposed by Robinson (1994). As usual it is assumed that for some $H \in \left(\frac{1}{2}, 1\right)$ the spectrum is

$$f(\lambda) \sim L\left(\frac{1}{\lambda}\right) \lambda^{1-2H} \quad \text{as } \lambda \to 0^+,$$

where the symbol $\sim$ indicates that the ratio of left and right hand sides tends to 1 and $L(\lambda)$ is a slowly varying function at infinity, i.e. a positive function satisfying

$$\lim_{\lambda \to \infty} \frac{L(q\lambda)}{L(\lambda)} = 1 \quad \text{for all } q > 0.$$

Moreover it is shown that as $T \to \infty$ and for the bandwidth m such that

$$\frac{1}{m} + \frac{m}{n} \to 0 \quad \text{as } T \to \infty$$

it holds

$$\frac{\hat{F}(\lambda_m)}{F(\lambda_m)} \xrightarrow{p} 1 \quad \text{as } T \to \infty$$

where $\xrightarrow{p}$ denotes the convergence in probability,

$$F(\lambda) = \int_0^\infty f(\theta) \mathrm{d}\theta,$$

and

$$\hat{F}(\lambda) = \frac{2\pi}{T} \sum_{j=1}^{[T\lambda/2\pi]} I(\lambda_j)$$

is the discretely averaged periodogram. Since Robinson shows that

$$F(\lambda) \sim L\left(\frac{1}{\lambda}\right) \frac{\lambda^{2(1-H)}}{2(1-H)} \quad \text{as } \lambda \to 0^+,$$

then for any $q > 0$

$$\frac{F(q\lambda)}{F(\lambda)} \sim q^{2(1-H)} \frac{L(1/q\lambda)}{L(1/\lambda)} \sim q^{2(1-H)}$$

as $\lambda \to 0^+$, it is reasonable to consider the following estimator for H:

$$\hat{H}_{mq} = 1 - \frac{1}{2\log q} \log\left[\frac{\hat{F}(q\lambda_m)}{\hat{F}(\lambda_m)}\right].$$

The variety of contradictory results offered by the empirical studied examined here can be explained by the means of an extension of fBm referred to as multifractional Brownian motion (mBm). This process has been introduced by Peltier and Lévy Véhel (1995) and studied among the others by Bianchi (2005). The main idea is to substitute the constant over time Hurst exponent with a suitable time dependent function $H(t)$. For the mBm the moving average representation (1.43) becomes

$$M_{H(t),K(t)}(t) = K(t) \cdot V_{H(t)}^{1/2} \int_{\mathbb{R}} f_t(s) \mathrm{d}B_s,$$

with

$$f_t(s) = \frac{1}{\Gamma(H(t) + 1/2)} \left\{ |t-s|^{H(t)-1/2} I_{(-\infty,t]}(s) - |s|^{H(t)-1/2} I_{(-\infty,0]}(s) \right\}.$$

$H : [0,\infty) \to (0,1]$ is a Hölder function[15]. This process is no longer stationary nor self-similar but, as $H(t)$ is the punctual Hölder exponent of the mBm at point t, the process is locally asymptotically self-similar with index $H(t)$, meaning that

$$\lim_{h \to 0^+} \frac{M_{H(t+hu),K(t+hu)}(t+hu) - M_{H(t),K(t)}(t)}{h^{H(t)}} \stackrel{d}{=} B_{H(t),K(t)}(u), \quad u \in \mathbb{R}.$$

As a consequence at any point t there exists a fBm with parameter $H(t)$ tangent to the mBm. Bianchi (2005) states that dependence and independence could be both present in financial time series, depending on the local behaviour of $H(t)$ in the time span ones looks at. Therefore he extends and adapts a class of estimators of the parameter H of the fBm in order to estimate the time-dependent long memory parameter of a multifractional process.

---

[15] Given two metric spaces (X, d_X) and (Y, d_Y), a function $f : X \to Y$ is a Hölder function with exponent $\alpha > 0$ if for each $x, y \in X$ such that $\mathrm{d}_X(x,y) < 1$ there exists a constant k such that

$$\mathrm{d}_Y(f(x), f(y)) \le k\mathrm{d}_X(x,y)^\alpha.$$

1.4   Merton Model and fBM

In this section we try to reconsider the Merton model assuming that the process for the firm's assets is driven by a fractional Brownian motion. Suppose again that the company has a simple capital structure that consists of equity and a zero-coupon bond with maturity T, that the term structure of interest rate is deterministic and flat and the firm pays no dividend over the life of the debt. Now we make an additional, more general, assumption on the dynamics of the firm's value. In particular we assume that V follows a geometric fractional Brownian motion with $H \in \left(\frac{1}{2}, 1\right)$, i.e.

$$\mathrm{d}V_t = \mu V_t \mathrm{d}t + \sigma V_t \mathrm{d}B_H(t).$$

The solution of this stochastic differential equation (see Hu and Øksendal, 2003, for the proof), based on the Wick products is

$$V_t = V_0 \exp\left(\mu t + \sigma B_H(t) - \frac{1}{2}\sigma^2 t^{2H}\right). \tag{1.58}$$

In order to get a solution for the price of the equity issued by the firm (and hence the price for the corporate bond), we notice that formula (1.5) and formula (1.8) are still valid but we need to evaluate the call under the fBM.

First we introduce some "prediction" formula for the fBM. Let $\mathscr{F}_t^H$ the σ-algebra generated by the fBM. The conditional expectation of a fBM process $\mathbb{E}\left[B_H(T)|\,\mathscr{F}_t^H\right]$, is difficult to compute due to correlation with the past. However Gripenberg and Norros (1996) derived the following formula:

$$\mathbb{E}\left[B_H(T)|\,\mathscr{F}_t^H\right] = B_H(t) + \int_0^t \Psi_T(t,s)\mathrm{d}B_H(s),$$

where

$$\Psi_T(t,s) = \frac{\sin\left(\pi(H-\frac{1}{2})\right)}{\pi} s^{\frac{1}{2}-H} \int_t^T \frac{(u^2-ut)^{H-\frac{1}{2}}}{u-s}\mathrm{d}u.$$

Alternatively it is possible to show that from (1.44) one gets

$$\mathbb{E}\left[B_H(T)|\,\mathscr{F}_t^H\right] = \int_0^t z_T(s)\mathrm{d}B_s.$$

To avoid this difficult computations, a system of quasi-conditional expectations is developed for the fBM. In particular the quasi-conditional expectation, $\tilde{\mathbb{E}}$ is such that

$$\tilde{\mathbb{E}}\left[B_H(T)|\,\mathscr{F}_t^H\right] = B_H(t)$$

and we say that $B_H(t)$ is a quasi-martingale. Moreover, for $f \in L_\phi^2(\mathbb{R})$

$$\tilde{\mathbb{E}}\left[\int_0^T f(s,\omega)\mathrm{d}B_H(s)\,\middle|\,\mathscr{F}_t^H\right] = \int_0^t f(s,\omega)\mathrm{d}B_H(s),$$

i.e. the stochastic integral is a quasi-martingale as well.

If the constant interest rate is denoted by r, then from an application of the Girsanov theorem for fractional Brownian motion (see Hu and Øksendal, 2003, Theorem 3.18) it follows that

$$\hat{B}_H(t) := B_H(t) + \frac{\mu - r}{\sigma}t \tag{1.59}$$

is a fractional Brownian motion with respect to the measure $\hat{\mu}_\phi$ defined on $\mathscr{F}_t^H$ by

$$\mathrm{d}\hat{\mu}_\phi(\omega) = \exp\left(-\int_0^T K(s)\mathrm{d}B_H(s) - \frac{1}{2}|K|_\phi^2\right)\mathrm{d}\mu_\phi(\omega),$$

where $K(s) = K(T, s)$ satisfies

$$\int_0^T K(T, s)\phi(t, s)\mathrm{d}s = \frac{\mu - r}{\sigma} \quad \text{for } 0 \leq t \leq T.$$

More importantly, under $\hat{\mu}_\phi$ for every bounded $\mathscr{F}_t^H$-measurable claim one has

$$F(t) = \mathrm{e}^{-r(T-t)}\tilde{\mathbb{E}}_{\hat{\mu}_\phi}\left[F(T)|\,\mathscr{F}_t^H\right].$$

It is also true that for $\theta \in \mathbb{R}$

$$B_H^*(t) = B_H(t) + \theta t^{2H} \tag{1.60}$$

is a fractional Brownian motion under a measure μ_ϕ^*. Now we need the following

Lemma 1.4.1. *If* $\mathbb{E}[f(B_H(T))] < \infty$, *then* $\forall t \leq T$

$$\tilde{\mathbb{E}}_{\mu_\phi}\left[f(B_H(T))|\,\mathscr{F}_t^H\right] = \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi(T^{2H} - t^{2H})}} \exp\left(-\frac{(x - B_H(t))^2}{2(T^{2H} - t^{2H})}\right) f(x)\mathrm{d}x. \tag{1.61}$$

Moreover denoting by

$$Z(t) = \exp\left(-\theta B_H(t) - \frac{\theta^2}{2}t^{2H}\right) = \exp\left(-\theta B_H^*(t) + \frac{\theta^2}{2}t^{2H}\right) \tag{1.62}$$

it holds

$$\tilde{\mathbb{E}}_{\mu_\phi^*}\left[f(B_H(T))|\,\mathscr{F}_t^H\right] = \frac{1}{Z(t)}\tilde{\mathbb{E}}_{\mu_\phi}\left[f(B_H(T))Z(T)|\,\mathscr{F}_t^H\right]. \tag{1.63}$$

Proof. First for $\lambda \in \mathbb{R}$ consider the process

$$\mathrm{d}X(t) = \lambda X(t)\mathrm{d}B_H(t); \quad X(0) = 1.$$

From (1.58) and recalling that $X(t)$ is a martingale

$$\tilde{\mathbb{E}}_{\mu_\phi}\left[\mathrm{e}^{\lambda B_H(T) - \frac{1}{2}\lambda^2 T^{2H}}\,\bigg|\,\mathscr{F}_t^H\right] = X(t) = \mathrm{e}^{\lambda B_H(t) - \frac{1}{2}\lambda^2 t^{2H}}$$

and thus

$$\tilde{\mathbb{E}}_{\mu_\phi}\left[\mathrm{e}^{\lambda B_H(T)}\,\bigg|\,\mathscr{F}_t^H\right] = \mathrm{e}^{\lambda B_H(t) + \frac{\lambda^2}{2}(T^{2H} - t^{2H})}$$

Now let

$$f(B_H(T)) = \frac{1}{2\pi}\int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}B_H(T)\xi}\hat{f}(\xi)\mathrm{d}\xi.$$

This means that we write f as the inverse Fourier transform of $\hat{f}$.
Now

$$\tilde{\mathbb{E}}_{\mu_\phi}\left[f(B_H(T))|\,\mathscr{F}_t^H\right] = \frac{1}{2\pi}\int_{\mathbb{R}} \tilde{\mathbb{E}}_{\mu_\phi}\left[\mathrm{e}^{\mathrm{i}B_H(T)\xi}\,\bigg|\,\mathscr{F}_t^H\right]\hat{f}(\xi)\mathrm{d}\xi$$

$$= \frac{1}{2\pi}\int_{\mathbb{R}} \exp\left(\mathrm{i}\xi B_H(t) - \frac{\xi^2}{2}\kappa(t, T, H)\right)\hat{f}(\xi)\mathrm{d}\xi,$$

where $\kappa(t, T, H) = (T^{2H} - t^{2H})$.
Since $\mathrm{e}^{-\frac{\xi^2}{2}\kappa(t,T,H)}$ is an inverse Fourier transform and can be obtained as the Fourier transform of

$$z(y) = \frac{1}{\sqrt{2\pi\kappa(t, T, H)}}\mathrm{e}^{\frac{-x^2}{2\kappa(t,T,H)}},$$

and since the Fourier transform of a convolution is the product of the Fourier transform of the two functions we have

$$\tilde{\mathbb{E}}_{\mu_\phi}\left[f(B_H(T))\middle|\,\mathscr{F}_t^H\right] = \int_\mathbb{R} z(B_H(t) - x)f(x)\mathrm{d}x$$

which is (1.61). The proof for (1.63) is similar:

$$\tilde{\mathbb{E}}_{\mu_\phi}\left[f(B_H(T))Z(T)\middle|\,\mathscr{F}_t^H\right] = \frac{1}{2\pi}\mathrm{e}^{-\frac{\theta^2}{2}T^{2H}}\int_\mathbb{R}\tilde{\mathbb{E}}_{\mu_\phi}\left[\mathrm{e}^{(\mathrm{i}\xi-\theta)B_H(T)}\middle|\,\mathscr{F}_t^H\right]\hat{f}(\xi)\mathrm{d}\xi$$

$$= \frac{1}{2\pi}\mathrm{e}^{-\frac{\theta^2}{2}T^{2H}}\int_\mathbb{R}\mathrm{e}^{(\mathrm{i}\xi-\theta)B_H(t)}\exp\left(\frac{(\mathrm{i}\xi-\theta)^2}{2}\kappa(t,T,H)\right)\hat{f}(\xi)\mathrm{d}\xi$$

$$= Z(t)\frac{1}{2\pi}\int_\mathbb{R}\exp\left(\mathrm{i}\xi B_H(t) - \left(\frac{\xi^2}{2} + \mathrm{i}\xi\theta\right)\kappa(t,T,H)\right)\hat{f}(\xi)\mathrm{d}\xi,$$

whereas

$$\tilde{\mathbb{E}}_{\mu_\phi^*}\left[f(B_H(T))\middle|\,\mathscr{F}_t^H\right] = \frac{1}{2\pi}\int_\mathbb{R}\tilde{\mathbb{E}}_{\mu_\phi^*}\left[\mathrm{e}^{\mathrm{i}\xi\left(B_H^*(T)-\theta T^{2H}\right)}\middle|\,\mathscr{F}_t^H\right]\hat{f}(\xi)\mathrm{d}\xi.$$

Since for $\lambda \in \mathbb{R}$ it holds

$$\tilde{\mathbb{E}}_{\mu_\phi^*}\left[\mathrm{e}^{\lambda B_H^*(T)}\middle|\,\mathscr{F}_t^H\right] = \mathrm{e}^{\lambda B_H^*(t)+\frac{\lambda^2}{2}(T^{2H}-t^{2H})},$$

it follows

$$\tilde{\mathbb{E}}_{\mu_\phi^*}\left[f(B_H(T))\middle|\,\mathscr{F}_t^H\right]$$

$$= \frac{1}{2\pi}\int_\mathbb{R}\exp\left(\mathrm{i}\xi B_H^*(t) - \frac{\xi^2}{2}\kappa(t,T,H)\right)\mathrm{e}^{-\mathrm{i}\xi\theta T^{2H}}\hat{f}(\xi)\mathrm{d}\xi$$

$$= \frac{1}{2\pi}\int_\mathbb{R}\exp\left(\mathrm{i}\xi B_H(t) - \left(\frac{\xi^2}{2} + \mathrm{i}\xi\theta\right)\kappa(t,T,H)\right)\hat{f}(\xi)\mathrm{d}\xi.$$

and hence (1.63). $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Finally, following Necula (2002)[16] we can derive closed form solution for the price of a European call option and hence the value of the equity.

Theorem 1.4.2. *Assuming a geometric fractional Brownian motion with $H \in \left(\frac{1}{2}, 1\right)$ for the firm asset, and that the firm has only issued zero coupon bonds with maturity T and total face value D, the value of the equity at time t is*

$$S_t^H = V_t N(d_1^H) - D\mathrm{e}^{-r(T-t)}N(d_2^H) \tag{1.64}$$

where

$$d_1^H = \frac{\log(V_t/D) + r(T-t) + \frac{1}{2}\sigma^2\kappa(t,T,H)}{\sigma\sqrt{\kappa(t,T,H)}} \tag{1.65a}$$

$$d_2^H = \frac{\log(V_t/D) + r(T-t) - \frac{1}{2}\sigma^2\kappa(t,T,H)}{\sigma\sqrt{\kappa(t,T,H)}}. \tag{1.65b}$$

where $\kappa(t,T,H) = T^{2H} - t^{2H}$.

Proof. Under the measure $\hat{\mu}_\phi$ recalling (1.58) and (1.59) we have

$$V_T = V_t \exp\left(r(T-t) + \sigma(\hat{B}_H(T) - \hat{B}_H(t)) - \frac{1}{2}\sigma^2\kappa(t,T,H)\right).$$

---

[16] A formula for the price of a European call option at $t = 0$ is derived in Hu and Øksendal (2003). The formula is extended for every $t \in [0,T]$ in Necula (2002).

It is clear that

$$S_t = \mathrm{e}^{-r(T-t)} \tilde{\mathbb{E}}_{\hat{\mu}_\phi} \left[(V_T - D)^+ \middle| \mathscr{F}_t^H \right]$$
$$= \mathrm{e}^{-r(T-t)} \tilde{\mathbb{E}}_{\hat{\mu}_\phi} \left[V_T I_{\{V_T > D\}} \middle| \mathscr{F}_t^H \right] - D\mathrm{e}^{-r(T-t)} \tilde{\mathbb{E}}_{\hat{\mu}_\phi} \left[I_{\{V_T > D\}} \middle| \mathscr{F}_t^H \right].$$

Now, after simple algebraic manipulations

$$\tilde{\mathbb{E}}_{\hat{\mu}_\phi} \left[I_{\{V_T > D\}} \middle| \mathscr{F}_t^H \right] = \tilde{\mathbb{E}}_{\hat{\mu}_\phi} \left[I_{\left\{ \hat{B}_H(T) > -d_2^H \sqrt{\kappa(t,T,H)} + \hat{B}_H(t) \right\}} \middle| \mathscr{F}_t^H \right].$$

Thus equation (1.61) guarantees that

$$\tilde{\mathbb{E}}_{\hat{\mu}_\phi} \left[I_{\{V_T > D\}} \middle| \mathscr{F}_t^H \right] = 1 - N \left(\frac{-d_2^H \sqrt{\kappa(t,T,H)} + \hat{B}_H(t) - \hat{B}_H(t)}{\sqrt{\kappa(t,T,H)}} \right)$$
$$= 1 - N \left(-d_2^H \right) = N \left(d_2^H \right).$$

Let $Z(t)$ be as in (1.62) with $\theta = -\sigma$. Then

$$\tilde{\mathbb{E}}_{\hat{\mu}_\phi} \left[V_T I_{\{V_T > D\}} \middle| \mathscr{F}_t^H \right] = \mathrm{e}^{rT} V_0 \tilde{\mathbb{E}}_{\hat{\mu}_\phi} \left[Z_T I_{\{V_T > D\}} \middle| \mathscr{F}_t^H \right].$$

From (1.63) we have

$$\tilde{\mathbb{E}}_{\hat{\mu}_\phi} \left[V_T I_{\{V_T > D\}} \middle| \mathscr{F}_t^H \right] = \mathrm{e}^{rT} V_0 Z(t) \tilde{\mathbb{E}}_{\mu_\phi^*} \left[I_{\{V_T > D\}} \middle| \mathscr{F}_t^H \right]$$
$$= V_t \tilde{\mathbb{E}}_{\mu_\phi^*} \left[I_{\{V_T > D\}} \middle| \mathscr{F}_t^H \right].$$

Now

$$\tilde{\mathbb{E}}_{\mu_\phi^*} \left[I_{\{V_T > D\}} \middle| \mathscr{F}_t^H \right] = N \left(d_1^H \right)$$

follows from (1.61) and

$$V_T = V_t \exp \left(r(T - t) + \sigma(\hat{B}_H^*(T) - \hat{B}_H^*(t)) + \frac{1}{2}\sigma^2 \kappa(t,T,H) \right).$$

$\square$

This result deserves some comments. Formula (1.64) is very similar to the classical Black-Scholes formula: time to maturity $T - t$ is replaced by $\kappa(t,T,H) = (T^{2H} - t^{2H})$. Note, however, that the price of an European call in a fBM framework converges to the classical Black-Scholes price for $H \downarrow \frac{1}{2}$ and thus

$$\lim_{H \downarrow \frac{1}{2}} S_t^H = S_t,$$

where S_t is given by (1.6).
However there is an important difference: option prices evaluated at different times t but same increment $T - t$ will give different results with fractional Black-Scholes formula, whereas the classical formula will give identical prices, i.e. if $T_2 - t_2 = T_1 - t_1$, then

$$S^H(t_1, T_1) \neq S^H(t_2, T_2)$$

unless $H = \frac{1}{2}$. This is due to the non-Markovian of fBM which forces S_t^H to be dependent on the underlying process evolution during the increment between different t's.
The price of the corporate bond and the spread are straightforward:

$$\overline{B}^H(t, T) = V_t N(-d_1^H) + D\mathrm{e}^{-r(T-t)} N(d_2^H);$$
$$s^H(t, T) = -\frac{1}{T - t} \log(V_t / D\mathrm{e}^{-r(T-t)} N(-d_1^H) + N(d_2^H)).$$

Again, both $\overline{B}^H(t,T)$ and $s^H(t,T)$ converge to the corrisponding value under the standard BM for $H \downarrow \frac{1}{2}$. Under the fBM is still valid the property for which credit spreads are decreasing function of the default-free interest rate, because we still have

$$\frac{\partial s(t,T)}{\partial r} = -\frac{D}{\overline{B}^H(t,T)} N(-d_2^H) \leq 0.$$

It could be interesting consider the first passage time model under the fractional Brownian motion. In this case the equity is a down-and-out call option with strike D and barrier level L.

Unfortunately in the literature a closed form solution for this contingent claim has been found only in the case $r = 0$ and when both the current value of the firm assets and the total face value are above the barrier. Following Necula (2003) we have the following result:

Theorem 1.4.3. *If $r = 0$, $V_t > L$ and $D > L$ the price of the equity is*

$$S_t^H = V_t N(d_1^H) - D e^{-r(T-t)} N(d_2^H) - \left[LN(y_1^H) - \frac{D}{L} V_t N(y_2^H) \right] \tag{1.66}$$

where d_1^H and d_2^H are obtained from (1.65a) and (1.65b) setting $r = 0$ and

$$y_1^H = \frac{\log(\frac{L^2}{V_t D}) + \frac{1}{2}\sigma^2 \kappa(t,T,H)}{\sigma \sqrt{\kappa(t,T,H)}}$$

$$y_2^H = \frac{\log(\frac{L^2}{V_t D}) - \frac{1}{2}\sigma^2 \kappa(t,T,H)}{\sigma \sqrt{\kappa(t,T,H)}}.$$

The main intuition behind (1.66) is that a down-and-out call with strike price D, barrier L and maturity T, whose price in t (when $r = 0$) is given by

$$\tilde{\mathbb{E}} \left[(V_T - D)^+ I_{\{\tau > T\}} \big| \mathscr{F}_t^H \right]$$
$$= \tilde{\mathbb{E}} \left[V_T I_{\{V_T > D, \tau > T\}} \big| \mathscr{F}_t^H \right] - D \tilde{\mathbb{E}} \left[I_{\{V_T > D, \tau > T\}} \big| \mathscr{F}_t^H \right],$$
$$\tau = \inf\{t : V_t = L\},$$

has the same payoff as a portfolio that consists in a long position of one call with strike D and maturity T and a short position of $\frac{D}{L}$ puts with strike price $\frac{L^2}{D}$ and maturity T. To see this, first consider the case of a standard Brownian motion B_t, $0 \leq t \leq T$ on some probability space $(\Omega, \mathscr{F}, P)$. When $r = 0$, under the risk neutral measure Q equivalent to P, the process for $X_t = \log V_t$ is

$$dX_t = -\frac{\sigma^2}{2} dt + \sigma dB_t.$$

Now consider the measure $\tilde{Q}$ equivalent to Q such that

$$Z = \frac{d\tilde{Q}}{dQ} = \exp\left[\frac{\sigma}{2} B_T - \frac{\sigma^2}{8} T \right]$$

is the Radon-Nikodym derivative of $\tilde{Q}$ with respect to Q, and

$$\tilde{B}_t = B_t - \frac{\sigma}{2} t$$

is a standard Brownian motion under $\tilde{Q}$. Clearly under $\tilde{Q}$, $X_t = \log V_t$ is a Brownian motion as

$$dX_t = \sigma d\tilde{B}_t.$$

Therefore we have

$$X_T - X_0 = \sigma \tilde{B}_T = \sigma B_T - \frac{\sigma^2}{2} T$$

and thus

$$Z = \exp\left[\frac{\sigma}{2}\left(\frac{X_T - X_0}{\sigma} + \frac{\sigma T}{2}\right) - \frac{\sigma^2}{8} T\right]$$

$$= \exp\left[\frac{X_T - X_0}{2} + \frac{\sigma^2}{8} T\right].$$

Let $L = \mathrm{e}^l$ and define

$$T_l = \inf\{t > 0 : V_t = L\} = \inf\{t > 0 : X_t = l\}.$$

For $l < X_0$ it holds

$$I_{\{X_t > l\}} = I_{\{T_l > t\}} + I_{\{X_t > l\}} I_{\{T_l \le t\}}$$

and therefore for every bounded function $f(\cdot)$, noticing that

$$Y_t = \begin{cases} X_t & \text{if } T_l > t \\ 2l - X_t & \text{if } T_l \le t \end{cases}$$

is a Brownian motion (reflection principle), one has

$$f(X_t) I_{\{T_l > t\}} = f(X_t)\left(I_{\{X_t > l\}} - I_{\{X_t > l\}} I_{\{T_l \le t\}}\right)$$

$$= f(X_t) I_{\{X_t > l\}} - f(2l - Y_t) I_{\{2l - Y_t > l\}} I_{\{T_l \le t\}}$$

$$= f(X_t) I_{\{X_t > l\}} - f(2l - Y_t) I_{\{Y_t < l\}} I_{\{T_l \le t\}}.$$

Denoting

$$T_l^Y = \inf\{t > 0 : Y_t = l\}$$

it follows

$$T_l^Y = T_l,$$

and thus

$$f(X_t) I_{\{T_l > t\}} = f(X_t) I_{\{X_t > l\}} - f(2l - Y_t) I_{\{Y_t < l\}} I_{\{T_l^Y \le t\}}$$

$$= f(X_t) I_{\{X_t > l\}} - f(2l - Y_t) I_{\{Y_t < l\}}.$$

Finally we can conclude, since X_t is a $\tilde{Q}$-Brownian motion

$$\mathbb{E}^{\tilde{Q}}\left[f(X_t) I_{\{T_l > t\}}\right] = \mathbb{E}^{\tilde{Q}}\left[f(X_t) I_{\{X_t > l\}}\right] - \mathbb{E}\left[f(2l - X_t) I_{\{X_t < l\}}\right]. \tag{1.67}$$

Now we can apply (1.67) to get the price in zero of a down-and-out call with maturity T:

$$\mathbb{E}^Q\left[(V_T - D)^+ I_{\{T_l > T\}}\right] = \mathbb{E}^{\tilde{Q}}\left[\mathrm{e}^{-\frac{X_T - X_0}{2} - \frac{\sigma^2}{8} T}(\mathrm{e}^{X_T} - D)^+ I_{\{T_l > T\}}\right]$$

$$= \mathbb{E}^{\tilde{Q}}\left[\mathrm{e}^{-\frac{X_T - X_0}{2} - \frac{\sigma^2}{8} T}(\mathrm{e}^{X_T} - D)^+ I_{\{V_T > L\}}\right]$$

$$- \mathbb{E}^{\tilde{Q}}\left[\mathrm{e}^{-\frac{2l - X_T - X_0}{2} - \frac{\sigma^2}{8} T}(\mathrm{e}^{2l - X_T} - D)^+ I_{\{V_T < L\}}\right]$$

$$= \mathbb{E}^{\tilde{Q}}\left[\frac{1}{Z}(\mathrm{e}^{X_T} - D)^+ I_{\{V_T > L\}}\right] - \mathbb{E}^{\tilde{Q}}\left[\mathrm{e}^{X_T - l}\frac{1}{Z}(\mathrm{e}^{2l - X_T} - D)^+ I_{\{V_T < L\}}\right]$$

$$= \mathbb{E}^Q\left[(\mathrm{e}^{X_T} - D)^+ I_{\{V_T > L\}}\right] - \mathbb{E}^Q\left[\mathrm{e}^{X_T - l}(\mathrm{e}^{2l - X_T} - D)^+ I_{\{V_T < L\}}\right]$$

$$= \mathbb{E}^Q\left[(V_T - D)^+ I_{\{V_T > L\}}\right] - \mathbb{E}^Q\left[\frac{V_T}{L}\left(\frac{L^2}{V_T} - D\right)^+ I_{\{V_T < L\}}\right].$$

Simple algebraic manipulation yields

$$\mathbb{E}^Q\left[(V_T - D)^+ I_{\{T_l > T\}}\right] =$$

$$= \mathbb{E}^Q\left[(V_T - D)I_{\{V_T > D\}}I_{\{V_T > L\}}\right] - \mathbb{E}^Q\left[\frac{D}{L}\left(\frac{L^2}{D} - V_T\right)I_{\{V_T < L^2/D\}}I_{\{V_T < L\}}\right].$$

Now, for $D > L$ the price of the barrier option is

$$\mathbb{E}^Q\left[(V_T - D)I_{\{V_T > D\}}\right] - \mathbb{E}^Q\left[\frac{D}{L}\left(\frac{L^2}{D} - V_T\right)I_{\{V_T < L^2/D\}}\right],$$

meaning that its payoff is equal to the one of a portfolio with a long position in a call struck at D and a short position in D/L puts struck at L^2/D. This result can be extended to the case of a fractional Brownian motion and taking the quasi-conditional expectations instead of the normal expectations one gets equation (1.66).

2

Fractional Models to Credit Risk Pricing

2.1   Introduction

In the academic literature models for the pricing of risky debt can be subdivided into two classes: firm's value and reduced form models.

The philosophy underlying firm's value models is to assume there is a fundamental process usually interpreted as the total value of the assets of the firm that has issued the bonds we are interested in. The value of the firm is assumed to move around stochastically and hence a stochastic process for the evolution of the firms' underlying assets is assumed. This is the driving force behind the dynamics of the prices of all securities issued by the firm. The well-known structural approach due to the seminal paper of Merton (1974) assumes that the company has issued only shares and a zero-coupon bond. The firm defaults if the value of its assets is lower than the promised debt payment at maturity. As a consequence all claims on the firm's value are modeled as derivative securities with the firm's value as underlying. Merton's model has been extended for instance by Black and Cox (1976), Geske (1977), Shimko et al. (1993) and Leland (1994) to allow for more realistic assumptions, such as the possibility of default before maturity, coupon payments, stochastic interest rates. The most common alternative to structural models is given by the reduced-form approach, which directly models the default process of risky debt. In combination with assumptions on the evolution of the risk-free rate and the recovery rate in the event of default, this is used to value risky debt. See, for example, Jarrow and Turnbull (1995), Duffie and Singleton (1997) and Madan and Unal (1998).

It is commonly agreed that structural models of credit risk have a poor performance in predicting corporate bonds prices. One of the main critiques to the classical Merton model and structural models in general, is that they predict credit spreads that are lower than the ones observed in the market. This is due to the fact that the assumptions behind the model are far for realistic. In particular the assumption of a geometric Brownian motion (gBm) for the firms' underlying assets proposed in the original paper by Merton could be the explanation for this poor performance. Many empirical evidence have shown that, when a gBm is used to describe log-returns, such a specification cannot describe the behaviour of financial assets mainly because the actual returns look to show some form of dependence.

The contribution of this chapter to the literature is twofold. First of all, we investigate the empirical properties of credit spreads, with specific reference to their long memory features. Although a number of empirical studies on credit spread dynamics have been carried out (see Pedrosa and Roll, 1998; Prigent et al., 2001; Kiesel et al., 2001), no research has been performed to date to investigate the long memory properties of credit spread, the only exception being Della Ratta and Urga (2005). Applied analysis on credit spreads so far has been carried out within the classical $I(0)$ vs $I(1)$ framework, i.e. by testing for stationarity vs nonstationarity of spreads only. It is well known that the distinction between $I(0)$ and $I(1)$ processes can be too restrictive. In contrast to $I(0)$ time series in which shocks die out at an exponential rate, or an $I(1)$ series with an infinite persistence (no mean reversion), Adenstedt (1974), Granger (1980) and Granger and Joyeux (1980) proposed an $I(d)$ time series with $0 < d < 1$ in which shocks dissipate at a slow hyperbolic rate. We show that credit spreads are likely to be long memory nonstationary processes, i.e. $I(d)$ processes with $d > 0.5$.

Secondly, we propose the fractional version of two structural credit risk models. The presence of long memory in credit spreads time series would provide a justification for the theoretical models proposed, that, in turn, would be able to explain realized credit spreads better than traditional credit risk structural models. A sensitivity analysis on bond prices and credit spreads predicted by such models shows that indeed the predicted credit spreads are closer to the real ones than those predicted by the original Merton model.

2.2 Merton and Black and Cox Fractional Models

In the structural approach to credit risk the firm liabilities are thought as contingent claims issued against the firm underlying assets. A stochastic process for the evolution of the firm underlying assets, V, and the conditions under which a default is triggered as well as the payoff of the risky debt in the event of default are specified. Merton (1974) assumes that the firm has only issued zero coupon bonds with maturity T and total face value D and that default may happen only at maturity. Denote by $\overline{B}(t,T)$ and S_t the prices in t of a defaultable zero coupon bond and the equity respectively. Both $\overline{B}$ and S are function of V and more generally all claims on the firm's value are evaluated as derivative securities with the firm's value as underlying. The term structure of interest rate is assumed to be deterministic and flat and the firm pays no dividend over the life of the debt. In case of default, bondholders are assumed to have absolute priority, i.e. bond value at time T is $\overline{B}(T,T) = \min(D, V_T)$ and the equity is simply a call option, $S_T = \max(V_T - D, 0)$. Whereas the original model assumes a Geometric Brownian motion for the firm value, here we consider the following dynamics for V:

$$\mathrm{d}V_t = \mu V_t \mathrm{d}t + \sigma V_t \mathrm{d}B_H(t) \tag{2.1}$$

where $B_H(t)$ denotes a fractional Brownian motion and $H \in \left(\frac{1}{2}, 1\right)$ is the Hurst parameter. The fractional Brownian motion is a Gaussian process with zero mean, stationary increments, variance

$$\mathbb{E}\left[B_H^2(t)\right] = t^{2H}$$

and covariance

$$\mathbb{E}\left[B_H(t)B_H(s)\right] = \frac{1}{2}\left(t^{2H} + s^{2H} - |t-s|^{2H}\right).$$

Usually it is assumed that $B_H(0) = 0$. For any $H \in (0,1)$ the process $B_H(t)$ is self-similar in the sense that

$$B_H(t) \stackrel{d}{=} a^{-H} B_H(at).$$

The parameter H negotiates whether the fractional Brownian motion has independent increments ($H = 1/2$), positive covariance between two increments over non-overlapping time intervals ($1/2 < H < 1$), or negative covariance between increments ($0 < H < 1/2$). If $1/2 < H < 1$ we say that $B_H(t)$ has a long range dependence, since

$$\sum_{n=-\infty}^{\infty} \gamma(n) = \infty, \tag{2.2}$$

where

$$\gamma(n) = \mathrm{Cov}\left[B_H(1), B_H(n+1) - B_H(n)\right] = \frac{1}{2}\left[|n+1|^{2H} + |n-1|^{2H} - 2|n|^{2H}\right].$$

When $H = 1/2$, $\gamma(n) = 0$ for all $n \neq 0$ whereas for $1/2 < H \leq 1$

$$\gamma(n) \sim H(2H-1)|n|^{2H-2}, \quad \text{as} \quad |n| \to \infty, \tag{2.3}$$

where "$\sim$" means that the ratio of the left and right hand sides tends to one. Therefore, when $H = 1/2$, (2.1) implies that log-returns are independent normal random variables. However, in recent years, many empirical evidence have shown that this assumption cannot be used to describe the behaviour of financial assets because: 1) the empirical distributions of the log-price variations are far from being normal; 2) the actual returns look to show some form of dependence. In this work we do not deal with the first issue. To

address the second point, stochastic volatility models have been proposed. Even though these models can produce quasi long-range dependence, they are very difficult to implement and often do not lead to tractable pricing solutions. On the other hand, fractional Brownian motion offers a natural way of modelling long-range dependence giving tractable solutions to pricing financial instruments.

In what follows we denote by $N(\cdot)$ the cumulative probability distribution function of a standard normal random variable:

$$N(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} \mathrm{e}^{-\frac{1}{2}u^2} \mathrm{d}u,$$

and by $n(\cdot) = N'(\cdot)$ the density function. The next result follows directly from the fractional Black-Scholes formula derived in Hu and Øksendal (2003) using the Wick–Itô–Skorohod calculus:

Theorem 2.2.1. *Assuming absolute priority for the bondholders, a geometric fractional Brownian motion (2.1) with $H \in \left(\frac{1}{2}, 1\right)$ for the firm asset, and that the firm has only issued zero coupon bonds with maturity T and total face value D, when the risk-free rate is constant and equal to r, the value of the equity at time 0 is*

$$S_0 = V_0 N(d_1) - D\mathrm{e}^{-rT} N(d_2) \tag{2.4}$$

where

$$d_1 = \sigma^{-1} T^{-H} \left(\log(V_0/D) + rT\right) + \frac{1}{2}\sigma T^H$$
$$d_2 = d_1 - \sigma T^H$$

and $N(\cdot)$ denotes the standard Normal cumulative distribution function. The price of the bond is

$$\overline{B}(0,T) = V_0 - S_0. \tag{2.5}$$

The spread is

$$s(0,T) = -\frac{1}{T} \log\left(\frac{\overline{B}(0,T)}{D}\mathrm{e}^{rT}\right) = -\frac{1}{T} \log\left(\frac{\overline{B}(0,T)}{\ell V_0}\right), \tag{2.6}$$

where $\ell = \frac{D\mathrm{e}^{-rT}}{V_0}$ is the firm leverage.

Note that as $H \downarrow \frac{1}{2}$ the price of the equity (2.4) converges to the Black-Scholes price for a call option and the model reduces to the classical Merton model.

Remark 2.2.2. *Since the price of the equity at time t is given by*

$$S_t = V_t N\left(\frac{\log(V_t/D) + r(T-t) + \frac{1}{2}\sigma^2\left(T^{2H} - t^{2H}\right)}{\sigma\sqrt{T^{2H} - t^{2H}}}\right)$$
$$- D\mathrm{e}^{-r(T-t)} N\left(\frac{\log(V_t/D) + r(T-t) - \frac{1}{2}\sigma^2\left(T^{2H} - t^{2H}\right)}{\sigma\sqrt{T^{2H} - t^{2H}}}\right),$$

equity prices evaluated at different times t but same time to maturity $T - t$ will give different results with fractional Black-Scholes formula, whereas the classical formula will give identical prices, i.e. if $T_2 - t_2 = T_1 - t_1$, then, denoting by $S(t,T)$ price at time t with maturity T and Hurst exponent H,

$$S^H(t_1, T_1) \neq S^H(t_2, T_2)$$

unless $H = \frac{1}{2}$. This is due to the fact that fBM is non-Markovian, which forces S_t^H to be dependent on the underlying process evolution during the increment between different t's.

Table 2.1: Sensitivity Analysis for the fractional Merton model.

Equity	Spread
$\frac{\partial}{\partial r}S_0 = DT\mathrm{e}^{-rT}N(d_2)$	$\frac{\partial}{\partial r}s = N(d_2)\mathrm{e}^{sT} - 1 = -\frac{V_0 N(-d_1)}{B}$
$\frac{\partial}{\partial \sigma}S_0 = V_0 T^H n(d_1)$	$\frac{\partial}{\partial \sigma}s = T^{H-1}n(d_2)\mathrm{e}^{sT}$
$\frac{\partial}{\partial \ell}S_0 = -V_0 N(d_2)$	$\frac{\partial}{\partial \ell}s = \frac{1 - N(d_2)\mathrm{e}^{sT}}{T\ell} = \frac{N(-d_1)\mathrm{e}^{sT}}{T\ell^2}$
$\frac{\partial}{\partial T}S_0 = V_0 \sigma H T^{H-1}n(d_1) + rD\mathrm{e}^{-rT}N(d_2)$	$\frac{\partial}{\partial T}s = \sigma H T^{H-2}n(d_2)\mathrm{e}^{sT} + \frac{-s+r\left[\mathrm{e}^{sT}N(d_2)-1\right]}{T}$
$\frac{\partial}{\partial H}S_0 = V_0 \log T \sigma T^H n(d_1)$	$\frac{\partial}{\partial H}s = \log T \sigma T^{H-1}n(d_2)\mathrm{e}^{sT}$

Table 2.1 reports the results of the sensitivity analysis for the fractional Merton model.

Some observations are in order. First of all the fractional Merton model shares some features of the original model. Note that the corporate bond is equivalent to a long position in a default free bond with nominal D and a short position in a put option written on V and with strike D or, alternatively, to a long position in the asset value and a short position in a call written on V and with strike D. Therefore, when the value of the firm is much bigger that the debt D, then the put option is deep out-of-the-money, the probability of default, $N(-d_2)$ is low and corporate debt trades as if it is default-free. Conversely, when $V_t \ll D$ then the call component is small and the bond is approximately equal to the value of the equity. As a consequence the volatility of the corporate debt depends on the volatility of the underlying asset when the the put option trades deep in-the-money. It is clear from Table 2.1 that the credit spread is a decreasing function of the risk-free rate. The intuition is as follows: an increase in the default-free spot interest rate, keeping the value of the firm constant, makes the probability of default decline. This, in turn, makes the corporate bond price increase and, consequently, the spread decline. Moreover credit spreads are an increasing function of the firm's leverage. The explanation is straightforward: the bigger the leverage, the bigger the probability of default. Clearly, an increase in the probability of default results in a decrease of the corporate bond price and therefore in an increase in the credit spread. Of course spreads are increasing functions of the firm volatility because equity investors do always benefit from an increase in asset volatility which makes the equity price increase. On the other hand this results in a decrease in the corporate bond price and, thus, in an increase in the spread. Another implication of the model is that the credit spreads tend to zero as the maturity of the zero-coupon bond tends to zero when $V > D$ and explode when $V \leq D$.

In Figure 2.1–Figure 2.3 we plotted the value of the spreads as a function of time to maturity for three values of firm leverage ($\ell \in \{0.8, 1, 1.2\}$) and three values of the parameter H ($H \in \{0.55, 0.8, 0.95\}$). The risk-free is $r = 0$, so that $\ell = D/V$. In every graph, part (a) plots the credit spreads against time for $T \in [0, 30]$ and part (b) for $T \in [0, 1]$, to show the different behaviour for time to maturity bigger and smaller than one.

It is clear that when H is close to $\frac{1}{2}$, i.e. under the classical Merton model, credit spreads are hump shaped with respect to maturity (rising at first and then falling), for values of leverage below unity and are decreasing for values of leverage bigger than one. For values of the memory parameter H bigger than $\frac{1}{2}$, spreads are increasing ($\ell = 80\%$) and first decreasing and then increasing ($\ell = 100\%$ and $\ell = 120\%$).

However one of the main critiques to the classical Merton model is that it predicts credit spreads which are lower than the ones observed in the market. In the fractional Brownian motion framework, on the other hand, when $T > 1$ spreads are increasing function of the long memory parameter H. This implies that the theoretical spreads in the fractional framework are bigger that the spreads predicted by the classical Merton model ($H = \frac{1}{2}$). The fractional model entails much more realistic spreads because it takes into account the dependency structure of financial returns. For $T = 1$ the spreads are a constant function of H and for $T < 1$ they are a decreasing function of H, meaning that for short maturities credit spreads are underestimated. This result can be explained by the fact that the fractional Merton model is affected, for short maturities, by the same issue of the classical model: since the firm value is modelled as a diffusion process, a sudden drop in the firm value is impossible, implying that the the firm's probability of default on very short term

Figure 2.1: Credit spreads resulting in the fractional Merton model against maturity when $\sigma = 0.2$, $D/V = 0.8$, $r = 0$.

(a) $0 < T < 30$

(b) $0 < T < 1$.

Figure 2.2: Credit spreads resulting in the fractional Merton model against maturity when $\sigma = 0.2$, $D/V = 1$, $r = 0$.

(a) $0 < T < 30$

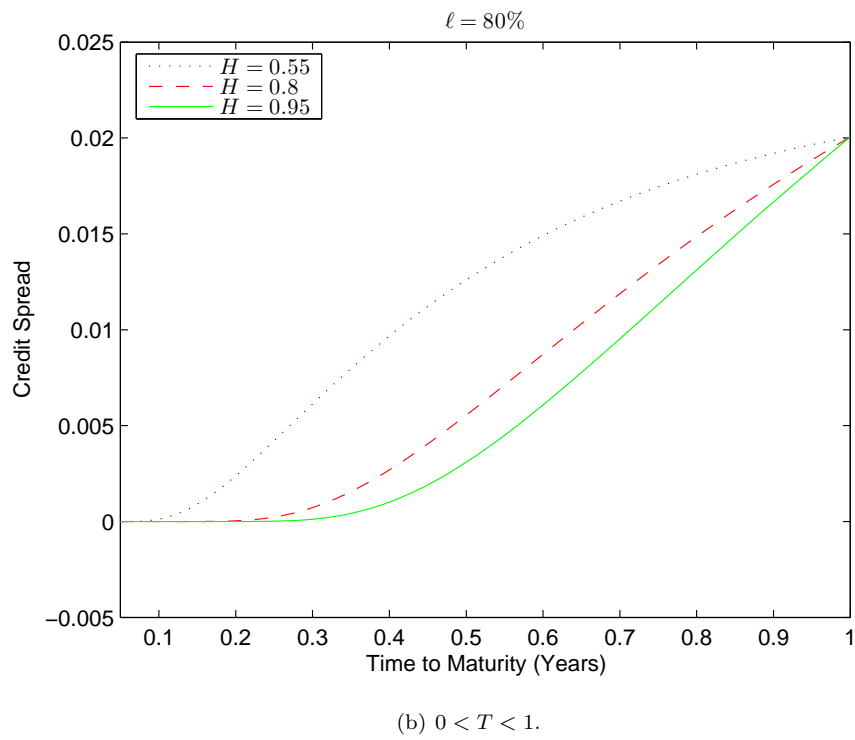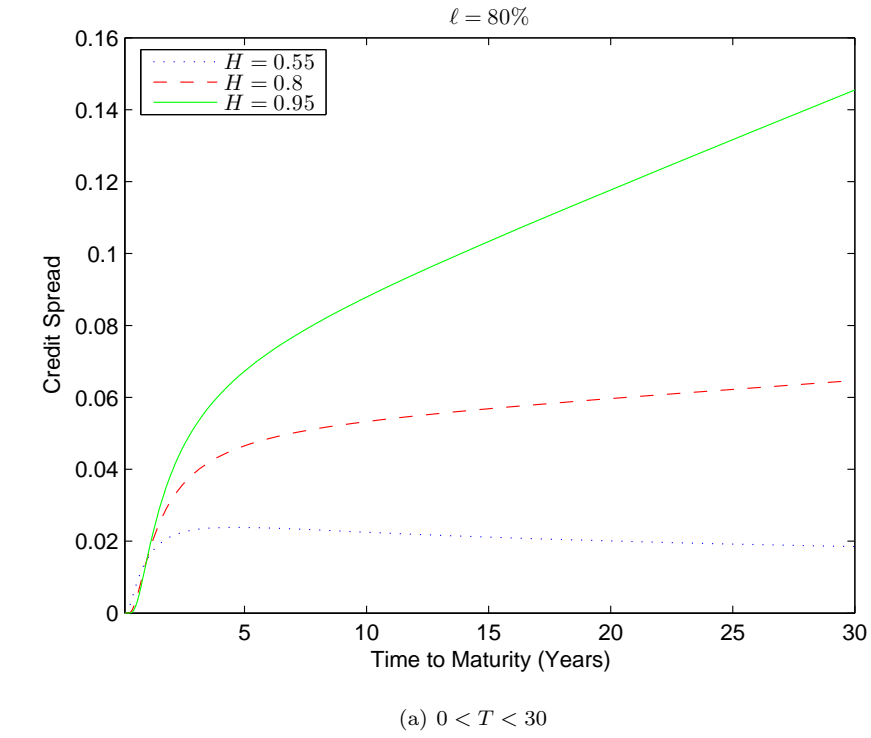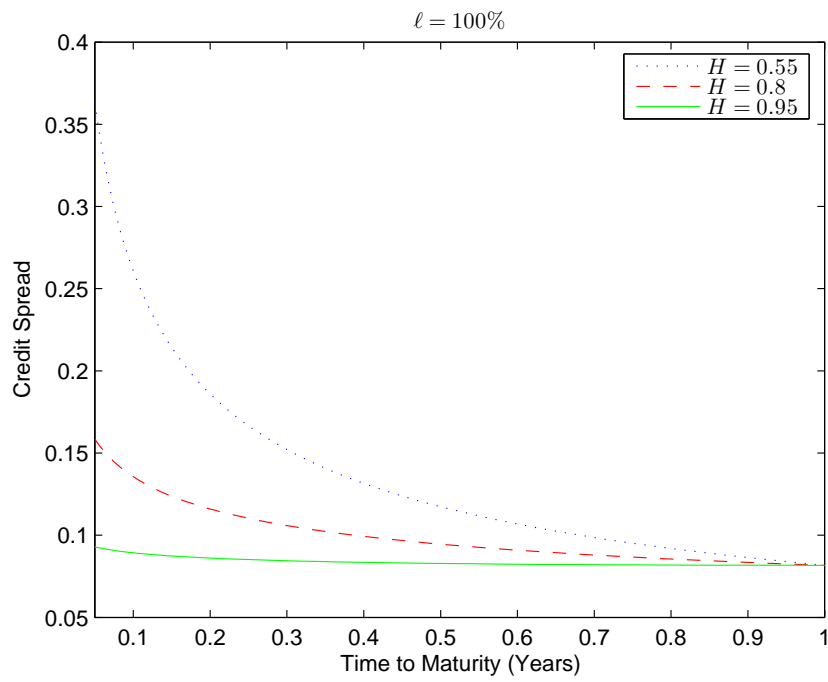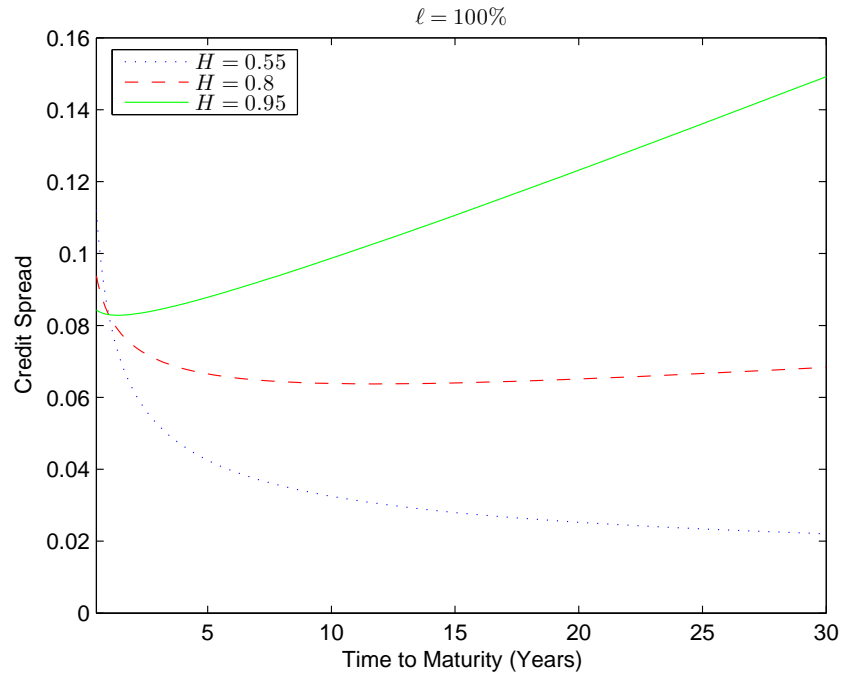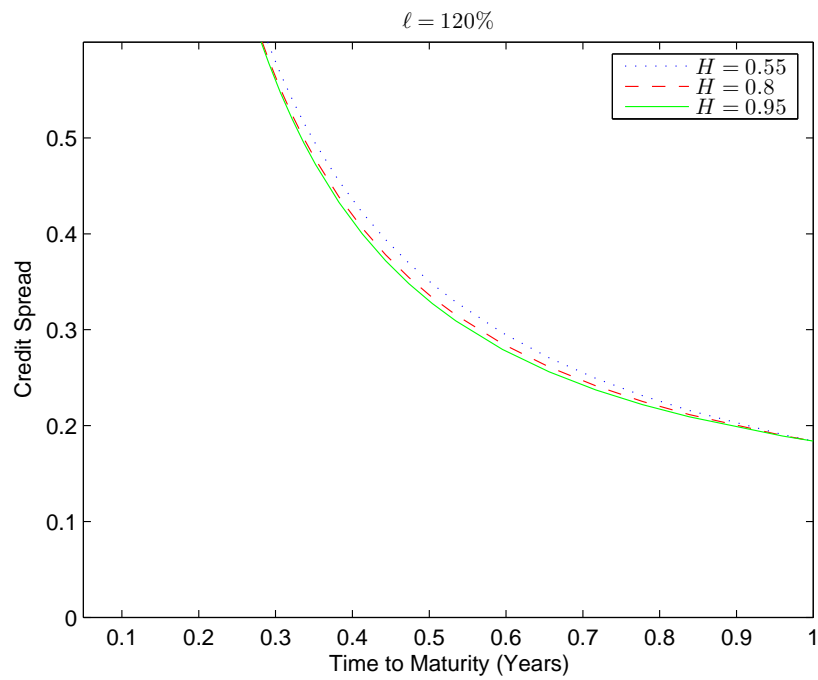(b) $0 < T < 1$.

Figure 2.3: Credit spreads resulting in the fractional Merton model against maturity when $\sigma = 0.2$, $D/V = 1.2$, $r = 0$.
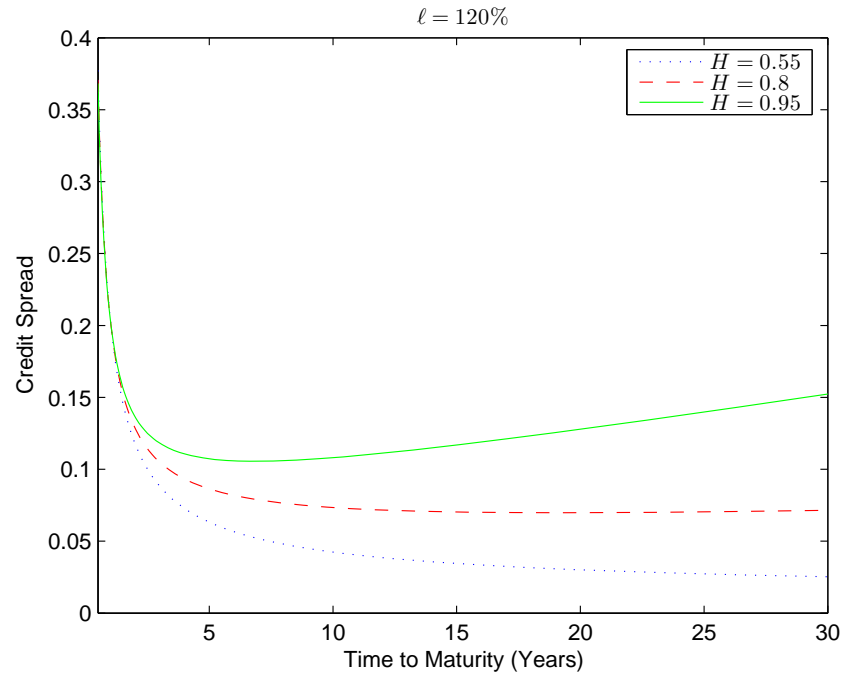

(a) $0 < T < 30$

(b) $0 < T < 1$.

debt should be zero. Therefore, as we approach time to maturity, theoretical credit spreads approach zero, systematically underestimating market spreads.

A more realistic assumption can be made about the conditions under which default is triggered. The company issuing debt, for instance, can be assumed to default not only at maturity T but the first time its assets fall below some default barrier, say L. To be more precise, if the assets hit the barrier L before time T, the option ceases to exist and the bondholders receive the assets value or some recovered portion of it. If the barrier is not hit before T, then the equity at maturity is, as in Merton model, the payoff on a European call option. This is the well known first passage time model proposed by Black and Cox (1976). A straightforward implication of the model is that the equity can be viewed as a down-and-out call option. To be more precise the time of default is given by the stopping time $\tau = \min\{k \geq t : V_k < L\}$. Denoting by $\mathbb{E}_t^Q[\cdot]$ the expectation under the risk neutral probability Q conditional to the information available up to time t and by I the indicator function, the equity value at time t is given by

$$S_t = \mathbb{E}_t^Q \left[\mathrm{e}^{-\int_t^T r(u)\mathrm{d}u} (V_T - D)^+ I_{\{\tau > T\}} \right].$$

It is well known that this expression leads to a nice closed-form solution when the interest rate is constant and V follows a geometric Brownian motion:

$$S_t = \left[V_t N(x_1) - D\mathrm{e}^{-r(T-t)} N(x_1 - \sigma\sqrt{T-t}) \right]$$

$$- \left[V_t \left(\frac{L}{V_t} \right)^{2\theta} N(y_1) - D\mathrm{e}^{-r(T-t)} \left(\frac{L}{V_t} \right)^{2\theta-2} N(y_1 - \sigma\sqrt{T-t}) \right],$$

where

$$\theta = \frac{r + \frac{1}{2}\sigma^2}{\sigma^2}$$

$$x_1 = \begin{cases} \frac{\log(V_t/D) + (r+\frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}} & \text{if } D \geq L \\ \frac{\log(V_t/L) + (r+\frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}} & \text{if } D < L \end{cases},$$

$$y_1 = \begin{cases} \frac{\log(L^2/(V_t D)) + (r+\frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}} & \text{if } D \geq L \\ \frac{\log(L/V_t) + (r+\frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}} & \text{if } D < L \end{cases}.$$

Denoting by $\bar{r} = r - \frac{1}{2}\sigma^2$, a closed-form solution is available for the risk neutral probability of default:

$$1 - Q(\tau > T, V_T > D)$$

$$= N\left(\frac{\log(D/V_t) - \bar{r}(T-t)}{\sigma\sqrt{T-t}} \right) + \left(\frac{L}{V_t} \right)^{\frac{2\bar{r}}{\sigma^2}} N\left(\frac{\log(L^2/(DV_t)) + \bar{r}(T-t)}{\sigma\sqrt{T-t}} \right).$$

Intuitively, this default probability is higher than the corresponding probability in the classical approach, which is obtained as the special case where $L = 0$. The corresponding payoff to bond investors at maturity is

$$\min(V_T, D)I_{\{\tau > T\}} + V_T I_{\{\tau \leq T\}}$$

$$= D - (D - V_T)^+ + (V_T - D)^+ I_{\{\tau \leq T\}}.$$

This position is equivalent to a portfolio composed of a risk-free loan with face value D maturing at T, a short European put on the firm with strike D and maturity T and a long European down-and-in call on the firm with strike D and maturity T. The price of the bond can be derived using the corresponding Black-Scholes formulae.

In order to derive a fractional version of this model, we need some assumptions. First of all we assume that $r = 0$ and that both the amount of the debt and the value of the firm are bigger than the threshold L. While

the former assumption may be restrictive, the latter is very realistic. Now we can extend to the case of a fBM the result in Carr et al. (1998): under the previous assumptions and in an arbitrage-free economy a position in a down-and-out call option is equivalent to being long a call struck at D and short D/L puts struck at L^2/D:

Theorem 2.2.3. *Assuming a geometric fractional Brownian motion with $H \in \left(\frac{1}{2}, 1\right)$ for the firm asset, that the firm has only issued zero coupon bonds with maturity T and total face value D and that it could default not only at maturity but also at the first time its assets fall below the barrier $L < V_0$, when the risk-free rate is $r = 0$, the value of the equity at time 0 for $D > L$ is*

$$S_0 = V_0 N(d_1) - De^{-rT} N(d_2) - \left[LN(y_1) - \frac{D}{L} V_0 N(y_2) \right] \tag{2.7}$$

where

$$y_1 = \sigma^{-1} T^{-H} \log\left(\frac{L^2}{V_0 D} \right) + \frac{1}{2} \sigma T^H$$

$$y_2 = y_1 - \sigma T^H.$$

The price of the bond is

$$\overline{B}(0, T) = V_0 - S_0. \tag{2.8}$$

The spread is

$$s(0, T) = -\frac{1}{T} \log\left(\frac{\overline{B}(0, T)}{D} \right) = -\frac{1}{T} \log\left(\frac{\overline{B}(0, T)}{\ell V_0} \right), \tag{2.9}$$

where $\ell = \frac{D}{V_0}$ is the firm leverage.

Table 2.2 reports the results of the sensitivity analysis for the fractional Black and Cox model.

Table 2.2: Sensitivity Analysis for the fractional Black and Cox model.

Equity	Spread
$\frac{\partial}{\partial \sigma} S_0 = T^H \left[V_0 n(d_1) - Ln(y_1) \right]$	$\frac{\partial}{\partial \sigma} s = T^{H-1} e^{sT} \left[n(d_2) - \frac{V_0}{L} n(y_2) \right]$
$\frac{\partial}{\partial \ell} S_0 = V_0 \left[N(y_2) - N(d_2) \right] + \frac{Ln(y_1)}{\sigma \ell T^H} \left[1 - \frac{L}{V_0} \right]$	$\frac{\partial}{\partial \ell} s = \frac{1}{T} \left[\frac{1}{\overline{B}(0,T)} \left(\frac{\partial S_0}{\partial \ell} \right) + \frac{1}{\ell} \right]$
$\frac{\partial}{\partial T} S_0 = \sigma H T^{H-1} \left[V_0 n(d_1) - Ln(y_1) \right]$	$\frac{\partial}{\partial T} s = \sigma H T^{H-2} e^{sT} \left[n(d_2) - \frac{V_0}{L} n(y_2) \right] - \frac{s}{T}$
$\frac{\partial}{\partial H} S_0 = \log T \sigma T^H \left[V_0 n(d_1) - Ln(y_1) \right]$	$\frac{\partial}{\partial H} s = \log T \sigma T^{H-1} e^{sT} \left[n(d_2) - \frac{V_0}{L} n(y_2) \right]$

Note that, denoting by s_{M} and s_{BC} the spreads resulting from the fractional Merton and Black and Cox model respectively, we have

$$\frac{\partial}{\partial x} s_{\mathrm{BC}} = \left(\frac{\partial}{\partial x} s_{\mathrm{M}} \right) \left[n(d_2) - \frac{V_0}{L} n(y_2) \right], \quad x \in \{\sigma, H\}.$$

Simple algebra yields

$$n(d_2) - \frac{V_0}{L} n(y_2) = n(d_2) \left\{ 1 - \exp\left[-\frac{2}{\sigma^2 T^{2H}} \log(L/V_0) \log(L/D) \right] \right\} > 0$$

since $L < D$ and $L < V_0$. This means that, as in the fractional Merton model, in the fractional Black and Cox model spreads are increasing functions of both the volatility and the memory parameter.

In Figure 2.4–Figure 2.6 we plotted the value of the spreads as a function of time to maturity for three values of firm leverage ($\ell \in \{0.8, 1, 1.2\}$) and three values of the parameter H ($H \in \{0.55, 0.8, 0.95\}$). In

particular we chose $V = 100$, $L = 70$ and $D \in \{80, 100, 120\}$. In every graph, part (a) plots the credit spreads against time for $T \in [0, 30]$ and part (b) for $T \in [0, 1]$, to show the different behaviour for time to maturity bigger and smaller than one. The spreads predicted by the Black and Cox model are below those predicted by Merton fractional model because $S_{BC} \leq S_M$, which implies $\overline{B}_{BC} \geq \overline{B}_M$ and therefore $s_{BC} \leq s_M$.

Clearly, for $\ell = 80\%$ spreads keep the usual hump shape, whereas for $\ell \geq 1$ spreads behave like a decreasing function of T, even for values of H close to unity.

2.3   Econometric Methodology

In what follows we will investigate whether credit spreads can be described by long range dependent series, in order to validate the analysis of section 2.2. The econometric methodology regarding both fractional integration and fractional cointegration is presented.

2.3.1   Fractional Integration

Consider a stationary sequence X_t. Assume that the spectral density of the sequence, i.e. $f(\lambda)$, $-\pi < \lambda \leq \pi$ such that $\gamma(s) = \int_{-\pi}^{\pi} f(\lambda) \cos(s\lambda) \mathrm{d}s$, $s = 0, \pm 1, \ldots$, satisfies for $0 < G < \infty$

$$f(\lambda) \sim G \lambda^{-2d} \quad \text{as} \quad \lambda \to 0^+, \tag{2.10}$$

where $d \in \left(-\frac{1}{2}, \frac{1}{2}\right)$ is the fractional differencing parameter which can be related to H via

$$d = H - \frac{1}{2}.$$

The main idea behind a fractional differentiation is that for many time series the distinction between $I(0)$ and $I(1)$ processes can be too restrictive. Even though many series show autocorrelation up to very long lags as in (2.2), taking the first differences may be excessive and lead to overdifferencing. Thus, the main idea of some authors (see Granger, 1980; Granger and Joyeux, 1980, among the others) was to introduce an $I(d)$ time series with $0 < d < 1$ in which shocks dissipate at a slow hyperbolic rate, in contrast to $I(0)$ time series in which shocks die out at an exponential rate, or an $I(1)$ series with an infinite persistence (no mean reversion). If the series X_t is $I(d)$, it becomes stationary after d-th differencing, i.e.

$$\Delta^d X_t = e_t,$$

where e_t is white noise, $\Delta = 1 - L$ and L denotes the lag operator, $L^k X_t = X_{t-k}$, and $(1 - L)^d$ has the binomial expansion

$$(1 - L)^d = \sum_{k=0}^{\infty} (-1)^k \frac{\Gamma(d+1)}{\Gamma(d-k+1)k!} L^k,$$

where $\Gamma(\cdot)$ is the Gamma function.

Note that the non-summability condition (2.2) for $1/2 < H \leq 1$ implies that the spectral density is unbounded at zero frequency. The spectral density (2.10) has, in fact, a pole when $0 < d < 1/2$ and a zero when $-1/2 < d < 0$ at $\lambda = 0$.

One of the most common procedure to estimate the parameter d in the frequency domain with stationary data, $-1/2 < d < 1/2$, is the log-periodogram regression proposed by Geweke and Porter-Hudak (1983), GPH henceforth. This is a semiparametric method in that no assumption on the behaviour of the spectral density apart from the origin is made (in other words, parameterization of the short run component is avoided). Suppose that equation (2.10) is approximately valid for the spectral density of a stationary long memory process when the first m Fourier frequencies, λ_j, $j = 1, \ldots, m$, are considered. The reason why only the first frequencies are considered is that long range dependency can be captured by studying the behaviour of big

Figure 2.4: Credit spreads resulting in the fractional Black and Cox model against maturity when $\sigma = 0.2$, $V = 100$, $D = 80$ and $L = 70$.

(a) $0 < T < 30$

(b) $0 < T < 1$.

Figure 2.5: Credit spreads resulting in the fractional Black and Cox model against maturity when $\sigma = 0.2$, $V = 100$, $D = 100$ and $L = 70$.

(a) $0 < T < 30$

(b) $0 < T < 1$.

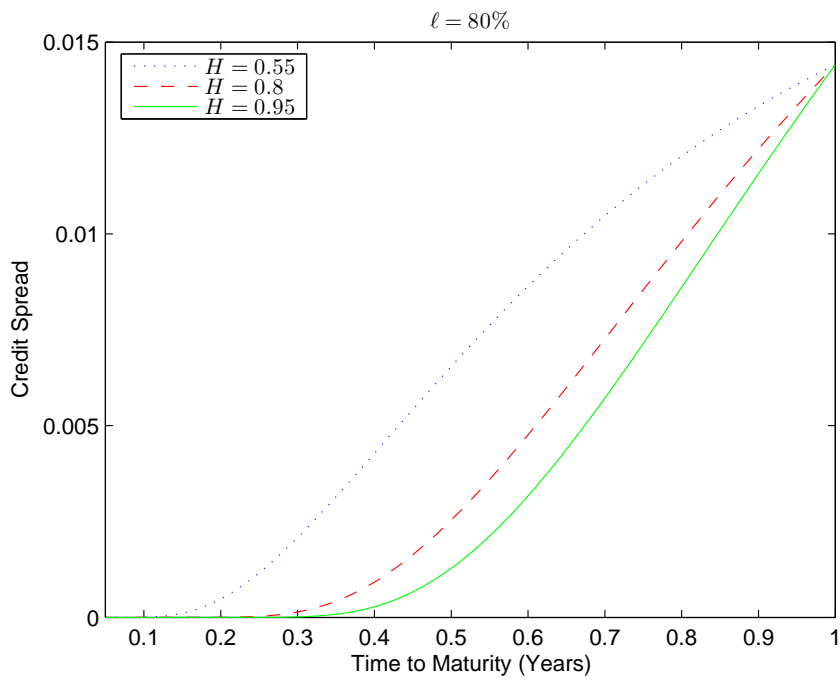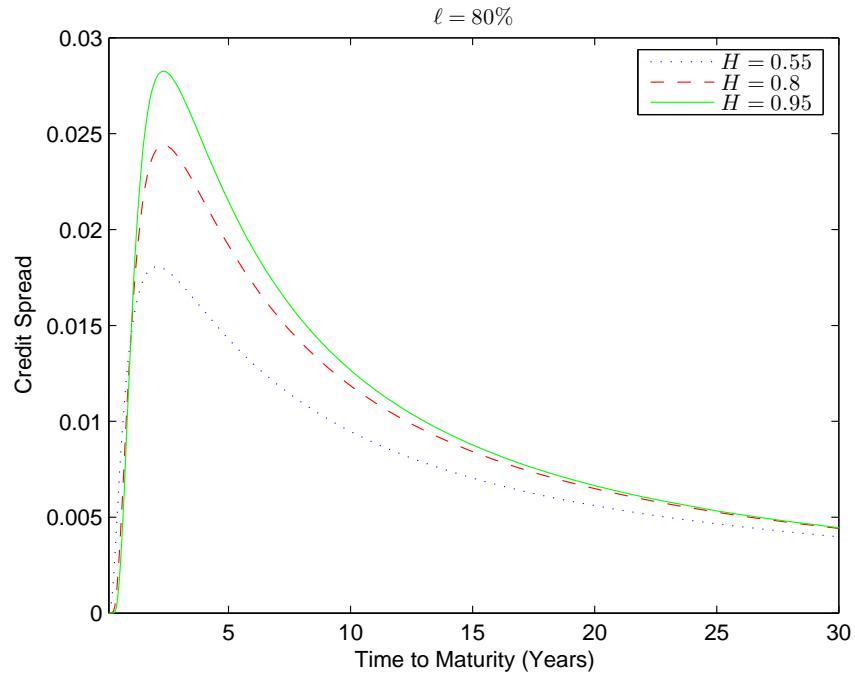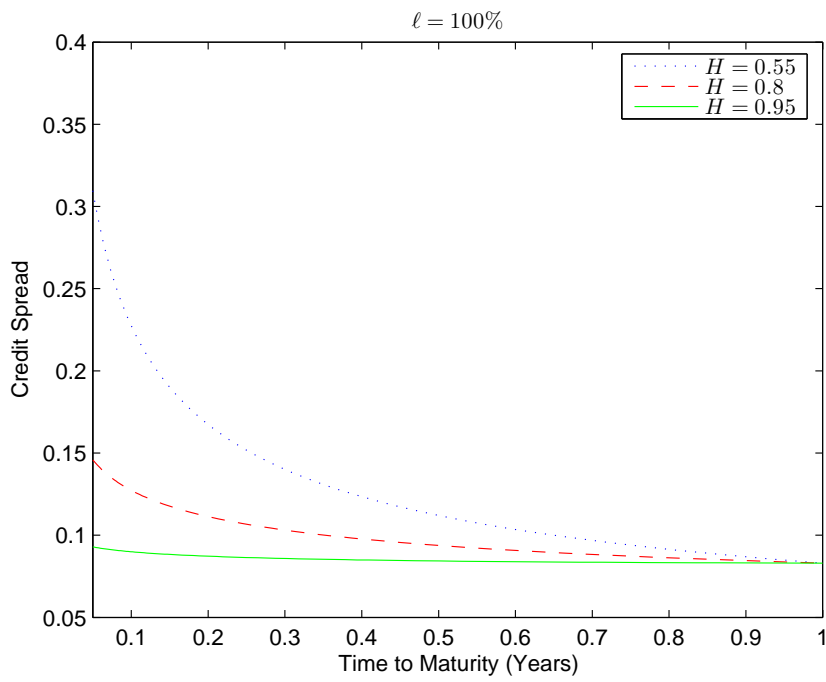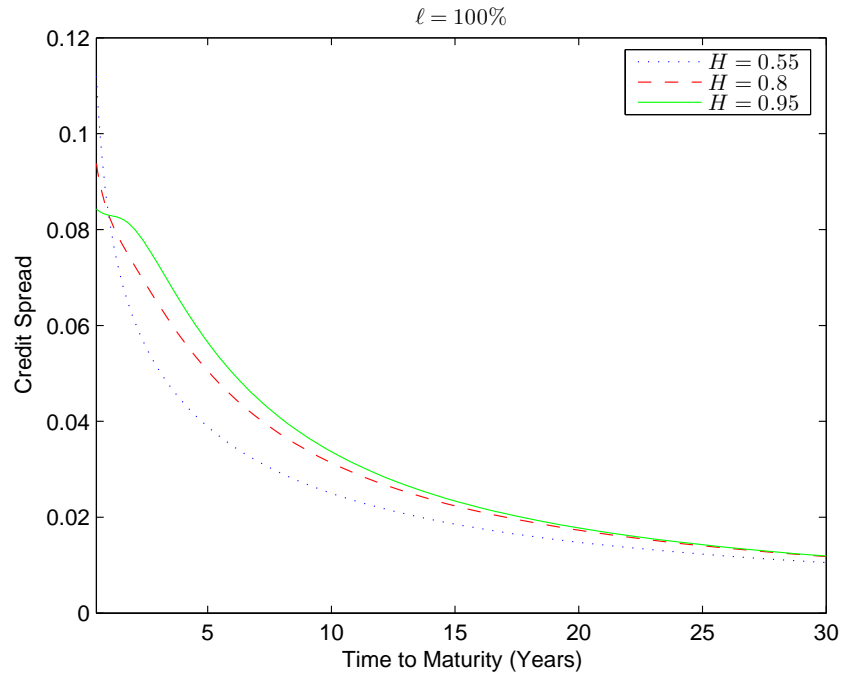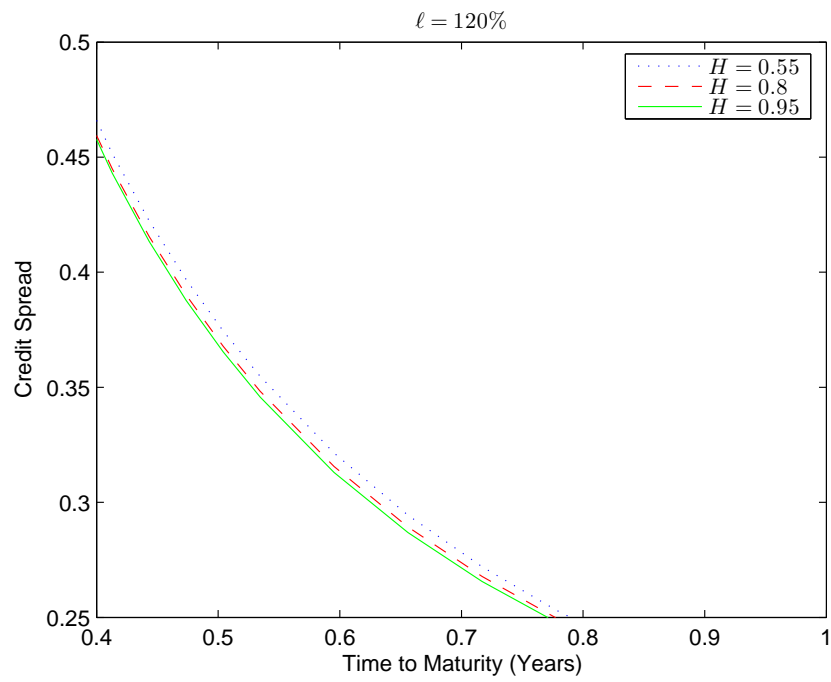**Figure 2.6:** Credit spreads resulting in the fractional Black and Cox model against maturity when $\sigma = 0.2$, $V = 100$, $D = 120$ and $L = 70$.
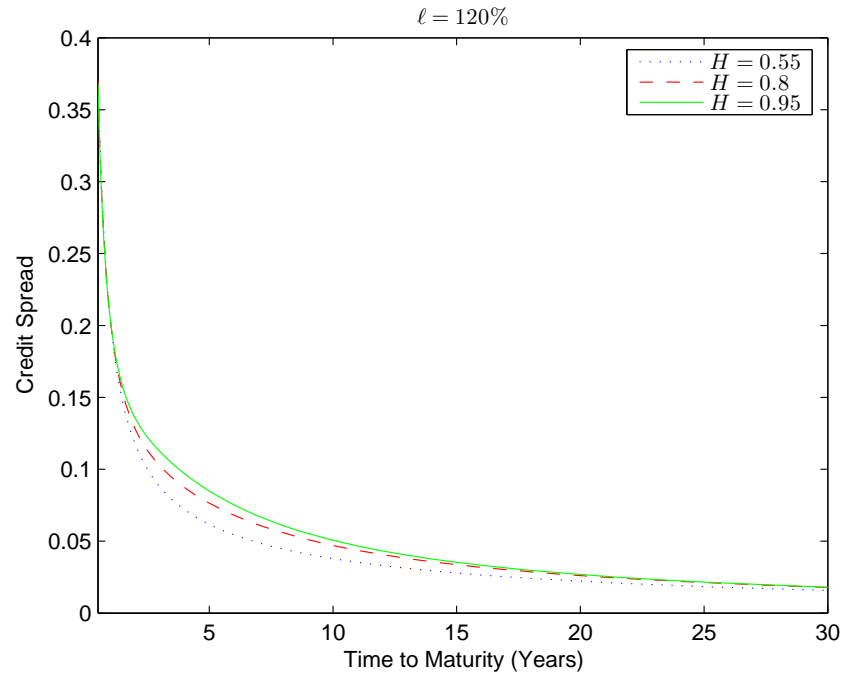

(a) $0 < T < 30$

(b) $0 < T < 1$.

lags in the autocorrelation function or, equivalently low frequencies in the spectral density.

Define the discrete Fourier transform of X_t for n observations and for $\lambda_j = 2\pi j/n$:

$$w(\lambda_j) = \frac{1}{\sqrt{2\pi n}} \sum_{t=1}^{n} X_t \mathrm{e}^{\mathrm{i}\lambda_j t}$$

and the periodogram

$$I(\lambda_j) = |w(\lambda_j)|^2 .$$

Taking the logarithm of both sides of (2.10) yields

$$\log f(\lambda_j) \approx \log G - 2d \log \lambda_j \quad j = 1, \ldots, m$$

or equivalently

$$\log I(\lambda_j) \approx c + da_j + \epsilon_j \quad j = 1, \ldots, m. \tag{2.11}$$

with $\epsilon_j = \log \frac{I(\lambda_j)}{f(\lambda_j)}$, $c = \log G$ and $a_j = -2 \log \lambda_j$.

Since the random variables $I(\lambda_j)/f(\lambda_j)$ are, at least for $d = 0$, asymptotically i.i.d. distributed, eq. (2.11) represents a linear regression model, with regressand $\log I(\lambda_j)$, regressors $-2 \log \lambda_j$ and slope d, which can be estimated by ordinary least squares (OLS). This is exactly the GPH estimator, except that here the quantity $2\sin(\lambda_j/2)$ is replaced by its first order Taylor expansion λ_j.

Robinson (1995b) proposed to trim the very low l frequencies and to consider the logs of a pooled periodogram of J periodogram values as the dependent variable in the log-periodogram regression. To be more precise, let J and $(m-l)/J$ be integers and

$$Y_h^{(J)} = \log \left(\sum_{j=1}^{J} I(\lambda_{h+j-J}) \right), \quad h = l+J, l+2J, \ldots, m.$$

The parameter d is now estimated by OLS from the regression

$$Y_h^{(J)} = c^{(J)} + da_h + \epsilon_h^{(J)}, \quad h = l+J, l+2J, \ldots, m$$

and therefore we have

$$\widehat{d} = \left(\sum_{h=l(J)}^{m} W_h^2 \right)^{-1} \sum_{h=l(J)}^{m} W_h Y_h^{(J)}, \tag{2.12}$$

where

$$W_h = a_h - \bar{a}, \quad \bar{a} = \frac{J}{m-l} \sum_{h=l(J)}^{m} a_h.$$

When both l and m tend to infinity with the sample size n but more slowly and under the assumption of Gaussianity, Robinson (1995b) has derived the limiting distribution for the GPH estimator $\widehat{d}$:

$$\sqrt{m} \left(\widehat{d} - d \right) \rightarrow_d N \left(0, \frac{\pi^2}{24} \right). \tag{2.13}$$

If in the log-periodogram regression we use $A_j = -\log \left[4\sin^2(\lambda_j/2) \right]$ as regressors,

$$\left(\widehat{d} - d \right) \rightarrow_d N \left(0, \sigma_{\mathrm{GPH}}^2 \right),$$

with

$$\sigma_{\mathrm{GPH}} = \sqrt{\frac{\pi^2}{6 \sum_{j=l+1}^{m} (A_j - \bar{A})^2}}$$

where $\bar{A} = \frac{1}{m-l} \sum_{j=l+1}^{m} A_j$. This suggests how to construct a test statistic for

$$H_0 : \widehat{d} = d.$$

The test statistic is

$$t_d = \frac{\widehat{d}(X) - d}{\sigma_{\text{GPH}}} \tag{2.14}$$

or

$$\tau_d = \frac{\widehat{d-1}(\Delta X) + 1 - d}{\sigma_{\text{GPH}}} \tag{2.15}$$

and it has to be compared with standard normal percentiles. The test statistic (2.15) has to be used when data is not stationary, in which case the long memory parameter for the series X is given by the sum of the the long memory parameter for the first differences of X and one.

This model easily extends to the multivariate case. Suppose that now X_t is a K-dimensional vector, whose k-th component is X_{kt}. We assume that for $k = 1, \ldots, K$

$$f_{kk}(\lambda) \sim G_k \lambda^{-2d_k} \quad \text{as} \quad \lambda \to 0^+,$$

where $-1/2 < d_k < 1/2$ and $f_{kk}(\cdot)$ denotes the spectral density of the k-th series. Now we have K regressions

$$Y_{kh}^{(J)} = c_k^{(J)} + d_k a_h + \epsilon_{kh}^{(J)},$$
$$h = l + J, l + 2J, \ldots, m \quad k = 1, \ldots, K,$$

where

$$Y_{kh}^{(J)} = \log\left(\sum_{j=1}^{J} I_{kk}(\lambda_{h+j-J})\right), \quad h = l + J, l + 2J, \ldots, m \quad k = 1, \ldots, K,$$

and

$$I_{kk}(\lambda) = \frac{1}{2\pi n} \left| \sum_{t=1}^{n} X_{kt} e^{i\lambda t} \right|^2 \quad k = 1, \ldots, K. \tag{2.16}$$

Denoting by $\widehat{\mathbf{d}}$ and $\widehat{\mathbf{c}}$ the estimates for $\mathbf{d} = (d_1, \ldots, d_K)'$ and $\mathbf{c} = \left(c_1^{(J)}, \ldots, c_K^{(J)}\right)'$ respectively, it follows

$$\begin{pmatrix} \widehat{\mathbf{c}} \\ \widehat{\mathbf{d}} \end{pmatrix} = \text{vec}\left(Y^{(J)'} S (S'S)^{-1}\right) \tag{2.17}$$

where $S = (S_{l+J}, S_{l+2J}, \ldots, S_m)'$ and $Y^{(J)} = (Y_1^{(J)}, \ldots, Y_K^{(J)})$, with $S_h = (1, a_h)'$ and $Y_k^{(J)} = \left(Y_{k,l+J}^{(J)}, Y_{k,l+2J}^{(J)}, \ldots, Y_{k,m}^{(J)}\right)'$, $k = 1, \ldots, K$.

Once again, for Gaussian stationary and invertible time series the estimator $\widehat{\mathbf{d}}$ is consistent and asymptotically multivariate normal. Thus it becomes straightforward to test the homogeneous restriction

$$H_0 : P\mathbf{d} = \mathbf{0}$$

where P is a $A \times K$ matrix of rank $A < K$. The test statistic is

$$\widehat{\mathbf{d}}' P' \left[[0, P]\{(S'S) \otimes \widehat{\Omega}^{-1}\}(0, P)' \right]^{-1} P\widehat{\mathbf{d}},$$

where $\widehat{\Omega} = \frac{J}{m-l} \sum_{h=l(J)}^{m} \widehat{\epsilon_h^{(J)}} \left(\widehat{\epsilon_h^{(J)}}\right)'$, $\widehat{\epsilon_h^{(J)}} = \left(\widehat{\epsilon_{1h}^{(J)}}, \ldots, \widehat{\epsilon_{Kh}^{(J)}}\right)'$ are the residuals of the log-periodogram regression and $\otimes$ denotes the Kronecker product. The test statistics has asymptotic χ_A^2 under H_0. For instance

it could be useful to test if the difference parameter is common to every series. In this case P is the $(K-1) \times (K+1)$ matrix

$$
P = \begin{pmatrix}
1 & -1 & 0 & \cdots & \cdots & \cdots \\
0 & 1 & -1 & 0 & \cdots & \cdots \\
\vdots & \ddots & \ddots & \ddots & \ddots & \ddots \\
0 & \cdots & \cdots & 0 & 1 & -1
\end{pmatrix}.
$$

Basically the same hypothesis is tested in Nielsen (2005). He proposed a Lagrange multiplier (LM) test for the K-dimensional series $\{X_t, t = 1, \ldots, n\}$ generated by

$$
\Delta^{d+\theta} X_t = e_t \mathbb{I}\,(t \geq 1) \equiv e_t^{\#} \quad t = 0, \pm 1, \pm 2, \ldots, \tag{2.18}
$$

where $\mathbb{I}(\cdot)$ denotes the indicator function. This definition (type II $I(d+\theta)$ process) entails that the results that follow are valid for all d and θ, whereas for a type I process they would be valid only for $d+\theta \in (-1/2, 1/2)$. The difference between the two types of processes is discussed in Robinson (2005).
The hypothesis to be tested is

$$
\mathrm{H}_0 : \theta = 0. \tag{2.19}
$$

From (2.18), the Gaussian log-likelihood function is

$$
L(\theta, \Sigma) = -\frac{n}{2} \log\left(2\pi|\Sigma|\right) - \frac{1}{2} \sum_{t=1}^{n} \Delta^{d+\theta} X_t' \Sigma^{-1} \Delta^{d+\theta} X_t.
$$

Denoting by $\eta = (\mathrm{vec}(\Sigma)', \theta')'$, the multivariate LM test statistic for testing H_0 is

$$
LM = \left.\frac{\partial L(\eta)}{\partial \eta'}\right|_{\theta=0, \Sigma=\widehat{\Sigma}} \left[-\left.\frac{\partial^2 L(\eta)}{\partial \eta \partial \eta'}\right|_{\theta=0, \Sigma=\widehat{\Sigma}} \right]^{-1} \left.\frac{\partial L(\eta)}{\partial \eta}\right|_{\theta=0, \Sigma=\widehat{\Sigma}}.
$$

Since it can be shown that

$$
\left.\frac{\partial L(\theta, \Sigma)}{\partial \theta}\right|_{\theta=0, \Sigma=\widehat{\Sigma}} = \mathrm{tr}\left(\widehat{\Sigma}^{-1} S_{10}\right)
$$

and that the relevant block in the Hessian matrix is

$$
-\left.\frac{\partial^2 L(\theta, \Sigma)}{\partial \theta^2}\right|_{\theta=0, \Sigma=\widehat{\Sigma}} = \mathrm{tr}\left(\widehat{\Sigma}^{-1} M_{11}\right)
$$

it follows that

$$
LM = \frac{\mathrm{tr}\left(\widehat{\Sigma}^{-1} S_{10}\right)^2}{\mathrm{tr}\left(\widehat{\Sigma}^{-1} M_{11}\right)}. \tag{2.20}
$$

Consider the processes $Z_t = \Delta^d X_t$, $Z_{t-1}^* = \sum_{j=1}^{t-1} j^{-1} Z_{t-j}$ and $Z_{t-2}^{**} = \sum_{j=1}^{t-2} j^{-1} Z_{t-j-1}^*$ and define $Z = (Z_1, \ldots, Z_n)'$, $Z^* = (Z_1^*, \ldots, Z_{n-1}^*)'$, $Z^{**} = (Z_1^{**}, \ldots, Z_{n-2}^{**})'$, $\underline{Z} = (Z_2, \ldots, Z_n)'$ and $\underline{\underline{Z}} = (Z_3, \ldots, Z_n)'$. The matrices used in (2.20) are defined as follows: $\widehat{\Sigma} = \frac{1}{n} Z'Z$ is a consistent estimate of the covariance matrix Σ, $S_{10} = \sum_{t=2}^{n} Z_{t-1}^* Z_t' = Z^{*'}\underline{Z}$, $M_{11} = S_{11} + \frac{1}{2}(S_{20} + S_{20}')$, with $S_{11} = Z^{*'}Z^*$ and $S_{20} = \sum_{t=3}^{n} Z_{t-2}^{**} Z_t' = Z^{**'}\underline{\underline{Z}}$. If the assumption of Gaussianity is made, under the null the test statistic (2.20) is asymptotically χ_1^2.
The test can be readily extended to the case of different θ for each variable. Suppose now that

$$
\Delta^{d_k+\theta_k} X_{kt} = e_{kt}^{\#}, \quad k = 1, \ldots, K, \quad t = 0, \pm 1, \pm 2, \ldots. \tag{2.21}
$$

Now $\theta = (\theta_1, \ldots, \theta_K)'$ is a K-vector. Denote by $L_K(\theta, \Sigma)$ the log-likelihood for Gaussian data generated by (2.21). It can be shown that

$$
\left.\frac{\partial L_K(\theta, \Sigma)}{\partial \theta}\right|_{\theta=0, \Sigma=\widehat{\Sigma}} = J_K' \,\mathrm{vec}\left(\widehat{\Sigma}^{-1} S_{10}'\right)
$$

and

$$
-\left.\frac{\partial^2 L_K(\theta, \Sigma)}{\partial\theta\partial\theta'}\right|_{\theta=0,\Sigma=\widehat{\Sigma}} = S_{11} \odot \widehat{\Sigma}^{-1} + \left(\widehat{\Sigma}^{-1} S_{20}'\right) \odot I_K
$$

where $\odot$ denotes the Hadamard product, I_K is the K-dimensional identity matrix and J_K is the $K^2 \times K$ $J_K = (\operatorname{vec} E_{11}, \ldots, \operatorname{vec} E_{KK})$, $E_{ii} = e_i e_i'$, where e_i is the i-th unit K-vector. Nielsen (2005) shows that, if the assumption of Gaussianity is made, under the null (2.19), the test statistic

$$
LM_K = \operatorname{vec}\left(\widehat{\Sigma}^{-1} S_{10}'\right)' J_K \left(S_{11} \odot \widehat{\Sigma}^{-1} + \left(\widehat{\Sigma}^{-1} S_{20}'\right) \odot I_K\right)^{-1} J_K' \operatorname{vec}\left(\widehat{\Sigma}^{-1} S_{10}'\right)
$$

is asymptotically χ_K^2. This due to the fact that now the number of restriction tested is K. The test can accommodate also deterministic trends. Suppose that we observe instead of X_t the series

$$
X_t^0 = \beta z_t + X_t
$$

where z_t is a vector of deterministic components, for instance $z_t = 1$ or $z_t = (1, t)'$. Assuming that the matrix $\sum_{t=1}^n \tilde{z}_t \tilde{z}_t'$ is positive definite for n sufficiently large, where $\tilde{z}_t = \Delta^d z_t$, we can estimate β by OLS regressing Z_t on $\tilde{z}_t$. The test statistic is then based on the residuals $\tilde{X}_t = X_t^0 - \hat{\beta} z_t$.

If one wants to allow for short run dynamics, the following vector autoregressive (VAR) could be considered:

$$
\Phi(L) e_t = \varepsilon_t \quad t = 0, \pm 1, \pm 2, \ldots
$$

where ε_t is $I(0)$, $\Phi(z)$ is a matrix polynomial of order p such that $\Phi(1)$ has full rank and e_t, the process in (2.18), is stationary. Now the test is based on the residuals from the regression

$$
Z_t = \hat{\Phi}_1 Z_{t-1} + \ldots + \hat{\Phi}_p Z_{t-p} + \hat{\varepsilon}_t \quad t = p+1, \ldots, n.
$$

Let $\hat{\varepsilon}_{t-1}^* = \sum_{j=1}^{t-1} j^{-1} \hat{\varepsilon}_{t-j}$, $\hat{\varepsilon}_{t-2}^{**} = \sum_{j=1}^{t-2} j^{-1} \hat{\varepsilon}_{t-j-1}^*$, $\hat{\varepsilon} = (\hat{\varepsilon}_{p+1}, \ldots, \hat{\varepsilon}_n)'$, $\hat{\varepsilon}^* = \left(\hat{\varepsilon}_{p+1}^*, \ldots, \hat{\varepsilon}_{n-1}^*\right)'$ and $\hat{\varepsilon}^{**} = \left(\hat{\varepsilon}_{p+1}^{**}, \ldots, \hat{\varepsilon}_{n-2}^{**}\right)'$. Define also $\underline{\hat{\varepsilon}} = (\hat{\varepsilon}_{p+2}, \ldots, \hat{\varepsilon}_n)'$ and $\underline{\underline{\hat{\varepsilon}}} = (\hat{\varepsilon}_{p+3}, \ldots, \hat{\varepsilon}_n)'$ and consider for $t = p+1, \ldots, n$ the pK-vector $\tilde{Z}_{t-1} = \left(Z_{t-1}', \ldots, Z_{t-p}'\right)'$ and the $pK \times (n-p)$ matrix $\tilde{Z} = \left(\tilde{Z}_{p+1}, \ldots, \tilde{Z}_n\right)'$. Now the test statistics are

$$
LM(p) = \frac{\operatorname{tr}\left(\widehat{\Sigma}^{-1} \hat{S}_{10}\right)^2}{\operatorname{tr}\left(\widehat{\Sigma}^{-1} M_{11} - \hat{S}_{Z1}' S_{ZZ}^{-1} \hat{S}_{Z1}\right)} \tag{2.22}
$$

$$
LM_K(p) = \operatorname{vec}\left(\widehat{\Sigma}^{-1} \hat{S}_{10}'\right)' J_K \left(\hat{S}_{11} \odot \widehat{\Sigma}^{-1} + \left(\widehat{\Sigma}^{-1} \hat{S}_{20}'\right) \odot I_K - \left(\hat{S}_{Z1}' S_{ZZ}^{-1} \hat{S}_{Z1}\right) \odot \widehat{\Sigma}^{-1}\right)^{-1}
$$

$$
\times J_K' \operatorname{vec}\left(\widehat{\Sigma}^{-1} \hat{S}_{10}'\right) \tag{2.23}
$$

with $\widehat{\Sigma} = \frac{1}{n-p} \hat{\varepsilon}' \hat{\varepsilon}$, $\hat{S}_{10} = \hat{\varepsilon}^{*\prime} \underline{\hat{\varepsilon}}$, $\hat{S}_{11} = \hat{\varepsilon}^* \hat{\varepsilon}^{*\prime}$, $\hat{S}_{20} = \hat{\varepsilon}^{**\prime} \underline{\underline{\hat{\varepsilon}}}$, $\hat{M}_{11} = \hat{S}_{11} + \frac{1}{2}(\hat{S}_{20} + \hat{S}_{20}')$, $S_{ZZ} = \tilde{Z}\tilde{Z}'$ and $\hat{S}_{Z1} = \sum_{t=p+2}^{n-1} \tilde{Z}_{t-1} \hat{\varepsilon}_{t-1}^* = \underline{\tilde{Z}} \hat{\varepsilon}^{*\prime}$ where in the last equality we defined $\underline{\tilde{Z}} = \left(\tilde{Z}_{p+1}, \ldots, \tilde{Z}_{n-1}\right)'$.

As discussed in Velasco (2000), for $0 < d < 1/2$ the log-periodogram regression estimator is still consistent even when the data is not Gaussian. Instead of Gaussianity a fourth order stationary linear process condition is required. To achieve asymptotic normality also tapering is needed. Define the tapered discrete Fourier transform which uses the full cosine bell as

$$
w_k^g(\lambda) = \left(2\pi \sum_{t=1}^n g_t^2\right)^{-1/2} \sum_{t=1}^n g_t X_{kt} e^{i\lambda t} \quad k = 1, \ldots, K,
$$

where

$$
g_t = \frac{1}{2}\left\{1 - \cos\left(\frac{2\pi t}{n}\right)\right\}.
$$

Note that $\sum_{t=1}^{n} g_t = \frac{1}{2}$ and $\sum_{t=1}^{n} g_t^2 = \frac{3}{8} n$. Next, the vector of difference parameter is estimated as in (2.17), except that the periodogram in (2.16) is replaced by

$$I_{kk}^g(\lambda) = \frac{1}{2\pi n} |w_k^g(\lambda)|^2 \quad k = 1, \ldots, K. \tag{2.24}$$

The use of (2.24) reduces the bias of the periodogram on the tails, because tapering downweights the observations at both extremes of the sample and does not change the central part.

One of the negative features of the GPH estimator is its finite-sample bias. In Andrews and Guggenberger (2003) a bias-reduced estimator is proposed which implies the use of the frequencies to the power $2k$ for $k = 1, \ldots, r$ and for some positive integer r, as additional regressors in the log-periodogram regression. The bias of this estimator goes to zero faster than the GPH estimator, but its variance is increased by a multiplicative constant. Let $\log I$, and a denote the column m-vectors whose j-th elements are $\log I(\lambda_j)$ and a_j and let Q be the $m \times r$ matrix with (j,k) element given by $[Q]_{j,k} = \lambda_j^{2k}$. The bias-reduced estimator $\widehat{d}_r$ can be obtained as the OLS estimator from the regression of $\log I(\lambda_j)$ on 1, a_j, $\lambda_j^2, \ldots, \lambda_j^{2r}$, for $j = 1, \ldots, m$ and therefore, using the partitioned regression formula, is given by

$$\widehat{d}_r = (a^{*\prime} M_{Q^*} a^*)^{-1} a^{*\prime} M_{Q^*} \log I,$$

with

$$\begin{aligned}
M_{Q^*} &= I_m - Q^*(Q^{*\prime} Q^*)^{-1} Q^{*\prime} \mathbb{I}\,(r \geq 1) \\
a^* &= a - m^{-1} \iota a' \iota \\
Q^* &= Q - m^{-1} \iota Q' \iota
\end{aligned} \tag{2.25}$$

In (2.25) I_m denotes the m-identity matrix and ι denotes a column m-vector of ones. Note that when $r = 0$ $M_{Q^*} = I_m$ and $\widehat{d}_r$ reduces to the GPH estimator. If m goes to infinity at a slower rate than the MSE-optimal rate (see Andrews and Guggenberger, 2003, for details),

$$\sqrt{m}\left(\widehat{d}_r - d\right) \rightarrow_d N\left(0, \sigma_r^2\right), \tag{2.26}$$

with

$$\sigma_r^2 = \frac{\pi^2}{24} c_r \tag{2.27}$$

$$c_r = \left(1 - \mu_r' \Gamma_r^{-1} \mu_r\right)^{-1} \tag{2.28}$$

where $\mu_r = (\mu_{r,1}, \ldots, \mu_{r,r})'$,

$$\mu_{r,k} = \frac{2k}{(2k+1)^2}, \quad \text{for } k = 1, \ldots, r$$

and Γ_r is an $r \times r$ matrix with (j,k) element given by

$$[\Gamma_r]_{j,k} = \frac{4jk}{(2j+2k+1)(2j+1)(2k+1)}, \quad \text{for } j, k = 1, \ldots, r.$$

The test statistic for

$$H_0 : \widehat{d} = d$$

is

$$t_d = \sqrt{m} \frac{\widehat{d}_r(X) - d}{\sigma_r} \tag{2.29}$$

or

$$\tau_d = \sqrt{m} \frac{\widehat{d_r - 1}(\Delta X) + 1 - d}{\sigma_r} \tag{2.30}$$

and it has to be compared with standard normal percentiles.

The asymptotic theory developed by Robinson (1995b) for stationary $I(d)$ processes has been extended to non-stationary ($d \geq 1/2$) and non-invertible ($d \leq -1/2$) series by Velasco (1999a,b). Consider $X_t \sim I(d)$ with $1/2 + \gamma \leq d < 3/2 + \gamma$, with γ integer. Then we will apply the log-periodogram regression to the series $W_t = \Delta^\gamma X_t$. Denote by $\widehat{d}(W)$ the corresponding estimate. The estimate for the original series will be

$$\widehat{d}(X) = \widehat{d}(W) + \gamma.$$

In other words, we first differenciate the original series γ times, apply the log-periodogram regression to the differencied series and finally adjust the resulting estimate with the number of differences taken. Basically the same idea applies to the case of non-invertible series, for which the data is first integrated and then the number of integrations taken is subtracted by the resulting estimate.

To capture also short range dependencies in the process X_t, Autoregressive Fractionally Integrated Moving Average (ARFIMA) processes have been proposed defined as

$$\phi(L)\Delta^d X_t = \theta(L)e_t,$$

where $\phi(L)$ and $\theta(L)$ involve autoregressive and moving average coefficients of order p and q respectively:

$$\phi(L) = 1 - \sum_{i=1}^{p} \phi_i L^i$$

$$\theta(L) = 1 + \sum_{i=1}^{q} \theta_i L^i,$$

with roots lying outside the unit circle. The requirement on $\phi(\cdot)$ implies stationarity and the one on $\theta(\cdot)$ implies invertibility. Using a, The spectral density is

$$f(\lambda; \zeta) = \frac{\sigma^2}{2\pi} \left|1 - e^{i\lambda}\right|^{-2d} \left|\frac{\theta(e^{i\lambda})}{\phi(e^{i\lambda})}\right|^2, \quad -\pi < \lambda \leq \pi, \tag{2.31}$$

where $\zeta = (d, \phi_1, \ldots, \phi_p, \theta_1, \ldots, \theta_q, \sigma^2)'$. The vector of parameter ζ can be estimated by maximizing an approximate form of frequency-domain Gaussian likelihood. This amounts to minimize the Whittle objective function

$$L(\zeta) = \sum_{j=1}^{n-1} \left\{ \log f(\lambda_j; \zeta) + \frac{I(\lambda_j)}{f(\lambda_j; \zeta)} \right\}.$$

A problem with this approach is to choose the autoregressive and moving average orders p and q. Under- or over-specifying p and q could, indeed, lead to bias in the estimates of d. To overcome these difficulties we could use the local Whittle estimator, $\tilde{d}$ proposed by Künsch (1987), which is based on (2.10) and maximizes the frequency-domain Gaussian likelihood only for frequencies in the neighborhood of zero. The Whittle log-likelihood is $-m/2$ times

$$Q(G, d) = \frac{1}{m} \sum_{j=1}^{m} \left\{ \log\left[G\lambda_j^{-2d}\right] + \frac{\lambda_j^{2d}}{G} I(\lambda_j) \right\}.$$

Concentrating $Q(G, d)$ with respect to G entails the minimization of the function $Q(\hat{G}, d)$, with

$$\hat{G} = \arg \min_{0 < G < +\infty} Q(G, d) = \frac{1}{m} \sum_{j=1}^{m} \lambda_j^{2d} I(\lambda_j).$$

The corresponding estimator is:

$$\tilde{d} = \arg \min_{-\frac{1}{2} < d < \frac{1}{2}} \left\{ \log\left[\frac{1}{m} \sum_{j=1}^{m} \lambda_j^{2d} I(\lambda_j)\right] - 2\frac{d}{m} \sum_{j=1}^{m} \log \lambda_j \right\}.$$

Robinson (1995a) shows that, if m goes to infinity but slower that n,

$$\sqrt{m}\left(\tilde{d} - d\right) \rightarrow_d N\left(0, \frac{1}{4}\right).$$

Thus, the test statistic for for

$$H_0 : \tilde{d} = d$$

is

$$t_d = 2\sqrt{m}\left(\tilde{d}(X) - d\right) \tag{2.32}$$

or, for first differenced data,

$$\tau_d = 2\sqrt{m}\left(\widetilde{d-1}(\Delta X) + 1 - d\right) \tag{2.33}$$

and it has to be compared with standard normal percentiles.

Looking at (2.13), it is clear that the local Whittle estimator is asymptotically more efficient than the GPH estimator for the same choice of m.

One of the disadvantages of the local Whittle estimator is its finite sample bias, which can be large. This can be avoided by using the local polynomial Whittle (LPW) estimator of Andrews and Sun (2004). The main idea is to approximate the logarithm of the short-run component of the spectrum in a shrinking neighborhood of frequency zero by a polynomial, instead of a constant. This means that instead of (2.10), the spectral density is assumed to be

$$f(\lambda) \sim \lambda^{-2d}\varphi(\lambda) \quad \text{as} \quad \lambda \to 0^+.$$

The logarithm of $\varphi(\lambda)$ near zero is approximated for some integer r by a constant plus an even polynomial of degree $2r$:

$$\varphi(\lambda) \approx \log G - \sum_{k=1}^{r} \theta_k \lambda^{2k}.$$

An even polynomial is chosen in order to reflect the symmetry of the spectrum about zero. Denoting by $\theta = (\theta_1, \ldots, \theta_r)'$, the local polynomial Whittle log-likelihood is $-m/2$ times

$$Q_r(G, d, \theta) = \frac{1}{m}\sum_{j=1}^{m}\left\{\log\left[G\lambda_j^{-2d}e^{-p_{r,j}}\right] + \frac{\lambda_j^{2d}}{G}I(\lambda_j)e^{p_{r,j}}\right\},$$

where

$$p_{r,j} = p_{r,j}(\theta) = \sum_{k=1}^{r}\theta_k\lambda_j^{2k}.$$

Concentrating $Q_r(G, d, \theta)$ with respect to G implies that the LPW estimator is given by

$$\tilde{d}_r = \arg\min_{-\frac{1}{2}<d<\frac{1}{2}}\left\{\log\left[\frac{1}{m}\sum_{j=1}^{m}\lambda_j^{2d}I(\lambda_j)e^{p_{r,j}}\right] - \frac{1}{m}\sum_{j=1}^{m}p_{r,j} - 2\frac{d}{m}\sum_{j=1}^{m}\log\lambda_j\right\}.$$

The asymptotic distribution of the LPW estimator is

$$\sqrt{m}\left(\tilde{d}_r - d\right) \rightarrow_d N\left(0, \frac{c_r}{4}\right), \tag{2.34}$$

where c_r is given by (2.28). The test statistic for

$$H_0 : \widehat{d} = d$$

is

$$t_d = 2\sqrt{m}\frac{\tilde{d}_r(X) - d}{\sqrt{c_r}} \tag{2.35}$$

or

$$\tau_d = 2\sqrt{m}\frac{\widetilde{d_r - 1}(\Delta X) + 1 - d}{\sqrt{c_r}} \tag{2.36}$$

2.3.2   Fractional Cointegration

Consider again the time series X_t and suppose that it is $I(d)$. Suppose that there exists α such that $\alpha' X_t$ is $I(\gamma)$ with $d \geq \gamma$ positive real numbers. In this case we say that X_t is fractionally cointegrated and denoted by $CI(d, d - \gamma)$. Suppose that $X_t = (y_t, x_t')'$ is a $(K+1)$-vector with $x_t = (x_{1t}, \ldots, x_{Kt})'$.

It is common in the literature to compute first the residuals from an equilibrium relation between non-stationary series integrated of order $d \in (1/2, 3/2)$, and then to apply the methodologies proposed in §2.3.1, to the residuals (or to the differenced residuals). Let us start with the case $K = 1$. Suppose $x_t \equiv x_{1t} \sim I(d)$ and $y_t \sim I(d)$ are observable with $d \in (1/2, 3/2)$, and that $u_t \sim I(\delta)$, where

$$y_t = \beta x_t + u_t, \quad t = 1, \ldots, n$$

and $0 \leq \delta < d$. Consider a consistent estimator for β and denote by $\hat{u}_t$ the observed residuals. If the equilibrium errors are stationary, $\delta \leq 1/2$, we can estimate δ as follows:

$$\widehat{\delta}(\hat{u}) = \left(\sum_{h=l+1}^{m} W_h^2 \right)^{-1} \sum_{h=l+1}^{m} W_h \log I_{\hat{u}\hat{u}}(\lambda_h).$$

This is (2.12) with $J = 1$. For non-stationary residuls we have

$$\widehat{\delta}(\Delta\hat{u}) = 1 + \left(\sum_{h=l+1}^{m} W_h^2 \right)^{-1} \sum_{h=l+1}^{m} W_h \log I_{\Delta\hat{u}\Delta\hat{u}}(\lambda_h).$$

Hassler et al. (2006) derive the condition under which these estimator are $\log n$-consistent and asymptotically normal, provided that $\hat{\beta}$ converges fast enough. In particular they assume that

$$\begin{cases} \hat{\beta} - \beta = O_P(n^{\delta-d}) & \text{if } \delta + d \geq 1 \\ \hat{\beta} - \beta = O_P(n^{1-2d}) & \text{if } \delta + d < 1 \end{cases}, \tag{2.37}$$

and

$$m \sim An^a, \quad l \sim Bn^b, \quad 0 < b < a < 1, \quad 0 < A, B < \infty. \tag{2.38}$$

Assumption (2.37) requires different rate of convergence for the β-estimator, depending on the overall memory of regressors and errors, $\delta + d$. In particular, when $\delta + d < 1$, a slower rate of convercence is required. Assumption (2.38) restricts the bandwidth numbers l and m to a power of n. Usually only small values are chosen for l ($l = 0$ or $l = 1$.)

They show[1] that for Gaussian u_t and x_t, under (2.37) and (2.38) as $n \to \infty$

$$\begin{cases} \log n \left(\widehat{\delta}(\hat{u}) - \delta \right) \to_p 0 & \text{if } 0 \leq \delta < 1/2 \text{ and } \delta < d - 1/2 < 1 \\ \log n \left(\widehat{\delta}(\Delta\hat{u}) - \delta \right) \to_p 0 & \text{if } 1/2 < \delta < d - 1/2 < 1 \end{cases}.$$

Moreover asymptotic Normality as in (2.13) for both $\widehat{\delta}(\hat{u})$ and $\widehat{\delta}(\Delta\hat{u})$ is showed and, most important, $\log n$-consistency even for non-Gaussian data when the pooled-tapered estimator is used:

$$\begin{cases} \log n \left(\widehat{\delta}^{(J)}(\hat{u}) - \delta \right) \to_p 0 & \text{if } d + \delta < 1 \text{ and } 0 \leq \delta < d < 3/2 \\ \log n \left(\widehat{\delta}^{(J)}(\Delta\hat{u}) - \delta \right) \to_p 0 & \text{if } 1/2 < \delta < d < 3/2 \end{cases}, \quad \text{as } n \to \infty.$$

Note that $\widehat{\delta}(\hat{u})$ is consistent only if $0 \leq \delta < 1/2$ with $\delta < d - 1/2 < 1$ as for differenced residuals, when l and m are chosen appropriately. Note that assumption (2.37) is fulfilled when $\hat{\beta}$ is the OLS estimates for

---

[1]Actually some more technical conditions than the ones presented here are required. (See Hassler et al., 2006, for the whole set of assumptions.)

$\delta \in [0, 3/2) - \{1/2\}$. Eq. (2.37) is satisfied also the narrow band frequency domain LS (FDLS) estimator proposed by Robinson and Marinucci (2001) when a bandwidth is chosen appropriately. This estimator improves the asymptotic and finite sample properties of OLS estimates. However it is important to notice that the authors use a different definition of $I(d)$ process and say that a process a_t, is $I(d)$ if there exist a zero mean scalar $I(0)$ process, ξ_t, $t \in \mathbb{Z}$, and a scalar μ such that

$$a_t = \mu + \Delta^{-d} \xi_t^{\#} \quad t \in \mathbb{Z}, \quad d > 0. \tag{2.39}$$

Let us present their estimator in our setting assuming now that $K \geq 1$. Suppose that

$$y_t = \beta' x_t + u_t = \sum_{i=1}^{K} \beta_i x_{it} + u_t, \quad t = 1, \ldots, n \tag{2.40}$$

for $x_{it} \sim I(d_i)$, $d_i \in (1/2, 3/2)$, $y_t \sim I(d_{\max})$ and $u_t \sim I(\delta)$, $0 \leq \delta < d_{\min}$ where $d_{\min} = \min_i d_i$ and $d_{\max} = \max_i d_i$. Note that the formulation (2.39) does not need the requirements on d_i because of the truncation. Thus following Robinson and Marinucci (2001) we can estimate β by the frequency domain least square statistic

$$\hat{\beta}_m = \hat{F}_{xx}(m) \hat{F}_{xy}(m), \quad 1 \leq m \leq n/2, \tag{2.41}$$

where for the column vector or scalar sequence b_t, $t = 1, \ldots, n$, possibly identical to a_t, $\hat{F}_{ab}(m)$ denotes the averaged cross-periodogram:

$$\hat{F}_{ab}(m) = 2\Re \left\{ \frac{2\pi}{n} \sum_{j=1}^{m} I_{ab}(\lambda_j) \right\} - \frac{2\pi}{n} I_{ab}(\pi) \mathbb{I} \left(m = \frac{n}{2} \right),$$

and

$$I_{ab}(\lambda) = w_a(\lambda) w_b'(-\lambda)$$

is the cross-periodogram. Note that, denoting by $[\cdot]$ the integer part and by $\bar{a} = n^{-1} \sum_{t=1}^{n} a_t$ we have

$$\hat{F}_{ab} \left(\left[\frac{n}{2} \right] \right) = \frac{1}{n} \sum_{t=1}^{n} (a_t - \bar{a})(b_t - \bar{b})', \tag{2.42}$$

and thus

$$\hat{\beta}_{[n/2]} = \left(\sum_{t=1}^{n} (x_t - \bar{x})(x_t - \bar{x})' \right)^{-1} \sum_{t=1}^{n} (x_t - \bar{x})(y_t - \bar{y})'$$

is the OLS estimate with intercept. Moreover $\hat{F}_{ab}(m)$ can be looked at as the contribution from the first m frequencies to the mean-corrected sample covariance (2.42). Denoting by $\hat{\beta}_{im}$ the i-th element of $\hat{\beta}_m$, the authors show that one has for $d_i > 1/2$, $\delta \geq 0$, $d_i + \delta < 1$

$$\begin{cases} \hat{\beta}_{im} - \beta_i = \mathrm{O}_{\mathrm{P}}(n^{\delta - d_i} m^{1 - d_{\min} - \delta}) \\ \hat{\beta}_{i[n/2]} - \beta_i = \mathrm{O}_{\mathrm{P}}(n^{1 - d_{\min} - d_i}) \end{cases}, \quad i = 1, \ldots, K.$$

if the condition

$$\frac{1}{m} + \frac{m}{n} \to 0 \quad \text{as } n \to \infty \tag{2.43}$$

is satisfied. This indicates that the FDLS estimator converges faster than the OLS estimator. Moreover, under (2.43) and when $d_1 = \cdots = d_K > 1/2$, $\delta > 0$ and $d_i + \delta > 1$

$$\begin{cases} \hat{\beta}_{im} - \beta_i = \mathrm{O}_{\mathrm{P}}(n^{\delta - d_i}) \\ \hat{\beta}_{i[n/2]} - \beta_i = \mathrm{O}_{\mathrm{P}}(n^{\delta - d_i}) \end{cases}, \quad i = 1, \ldots, K$$

meaning that in this case, as long as at least an arbitrarily slowly increasing number m of frequencies is included, omission of higher frequencies does not change the speed of convergence.

When $-1/2 < \delta < 1/2$, to estimate the vector $(\delta, d_1 - 1, \ldots, d_K - 1)'$, derive the residuals from (2.40) and simply use (2.17) and thus solve $K + 1$ regressions with dependent variables

$$\left(\log I_{\hat{u}\hat{u}}(\lambda_j), \log I_{\Delta x_1 \Delta x_1}(\lambda_j), \ldots, \log I_{\Delta x_K \Delta x_K}(\lambda_j)\right).$$

When the residuals are likely to be non-stationary, estimate the vector $(\delta - 1, d_1 - 1, \ldots, d_K - 1)'$ by (2.17) and thus solve $K + 1$ regressions with dependent variables

$$\left(\log I_{\Delta\hat{u}\Delta\hat{u}}(\lambda_j), \log I_{\Delta x_1 \Delta x_1}(\lambda_j), \ldots, \log I_{\Delta x_K \Delta x_K}(\lambda_j)\right).$$

Define by $\widehat{\mathbf{d}}(\hat{u})$ and $\widehat{\mathbf{d}}(\Delta\hat{u})$ the estimates of the vector $(\delta, d_1, \ldots, d_K)'$ based on original and differenced residuals respectively. Assume that

$$\begin{cases} \hat{\beta}_i - \beta_i = \mathrm{O}_{\mathrm{P}}(n^{\delta - d_i}) & \text{if } \delta + d_i \geq 1 \\ \hat{\beta}_i - \beta_i = \mathrm{O}_{\mathrm{P}}(n^{1 - d_{\min} - d_i}) & \text{if } \delta + d_i < 1 \end{cases}, \quad i = 1, \ldots, K. \tag{2.44}$$

Clearly under assumptions (2.38)–(2.44), $\widehat{\delta}$ is still $\log n$-consistent and under the additional assumption of Normality

$$\begin{cases} 2\sqrt{m}\left(\widehat{\mathbf{d}}(\hat{u}) - \mathbf{d}\right) \to_d N(0, \Omega) & \text{if } 0 \leq \delta < 1/2 \\ 2\sqrt{m}\left(\widehat{\mathbf{d}}(\Delta\hat{u}) - \mathbf{d}\right) \to_d N(0, \Omega) & \text{if } 1/2 < \delta < 1 \end{cases}.$$

To test the null of cointegration in fractionally cointegrated model we could use the Lagrange multiplier test of Nielsen (2004). Suppose that $X_t = (y_t, x_t')'$, with $y_t \sim I(d)$, is generated by the fractionally cointegrated system

$$y_t = \beta' x_t + z_t \quad t = 1, 2, \ldots \tag{2.45}$$

$$\Delta^{\delta + \theta} z_t = u_{1t}^{\#} \quad t = 1, 2, \ldots \tag{2.46}$$

$$\Delta^d x_t = u_{2t}^{\#} \quad t = 1, 2, \ldots \tag{2.47}$$

where $\delta = d - b$ with $d \geq b \geq 3/4 + \epsilon$, for some $\epsilon > 0$ and $u_t = (u_{1t}, u_{2t}')'$ is an error component. Under the null

$$\mathrm{H}_0 : \theta = 0$$

X_t is $CI(d, b)$. Moreover we assume that

$$\phi(L) u_{1t} = e_{1t} \quad t = 1, 2, \ldots$$
$$\Phi(L) u_{2t} = e_{2t} \quad t = 1, 2, \ldots$$

where $\phi(z)$ and $\Phi(z)$ are polynomials of order p with coefficient gathered in $\gamma = (\gamma_1', \gamma_2')'$, $\Phi(1)$ has full rank, meaning that there is no cointegration among the components of x_t and $e_t = (e_{1t}, e_{2t}')' \sim \mathrm{iid}(0, \Sigma)$ with

$$\Sigma = \begin{bmatrix} \sigma_{11}^2 & \sigma_{21}' \\ \sigma_{21} & \Sigma_{22} \end{bmatrix}.$$

Assuming Gaussianity of the errors the log-likelihood function is

$$L(\theta, \beta, \Sigma, \gamma) = -\frac{n}{2} \log |\Sigma| - \frac{1}{2} \sum_{t=1}^{n} \begin{pmatrix} \phi(L)\Delta^{\gamma + \theta} z_t \\ \Phi(L)\Delta^d x_t \end{pmatrix}' \Sigma^{-1} \begin{pmatrix} \phi(L)\Delta^{\gamma + \theta} z_t \\ \Phi(L)\Delta^d x_t \end{pmatrix}.$$

Define

$$e_{1.2t} = e_{1t} - \sigma_{21}' \Sigma_{22}^{-1} e_{2t}$$

which is e_{1t} centered about its mean conditional on e_{2t}, and the corresponding variance

$$\sigma_{1.2}^2 = \sigma_{11}^2 - \sigma_{21}' \Sigma_{22}^{-1} \sigma_{21}.$$

It can be shown that the MLE of β under the null can be obtained as the non-linear least squares estimator in the augmented regression

$$\Delta^\delta y_t = \beta' \Delta^\delta x_t + (1 - \phi(L)) \Delta^\delta (y_t - \beta' x_t) + c' \Phi(L) \Delta^d x_t + e_{1.2t}. \tag{2.48}$$

Thus the estimate we get is not the one we would have gotten estimating (2.45) by OLS. The two estimator coincide only if $\sigma_{21} = 0$ and $\phi(L) = 1$. When $\sigma_{21} \neq 0$ or $\phi(L) \neq 1$ the OLS estimator is biased because of endogeneity and serial correlation. When there is no autoregressive term in the equilibrium errors, i.e. when $\phi(L) = 1$, (2.48) reduces to

$$\Delta^\delta y_t = \beta' \Delta^\delta x_t + \sum_{k=0}^{p} c_k' \Delta^d x_{t-k} + e_{1.2t}. \tag{2.49}$$

The normalized score statistic under the null

$$S_n = \frac{1}{\sqrt{n}} \frac{\partial L(\theta, \beta, \Sigma, \gamma)}{\partial \theta} \bigg|_{\theta=0, \beta=\hat{\beta}, \Sigma=\hat{\Sigma}, \gamma=\hat{\gamma}}$$

can be shown to be

$$S_n = \frac{1}{\sqrt{n} \hat{\sigma}_{1.2}^2} \sum_{t=1}^{n} \sum_{j=1}^{t-1} j^{-1} \hat{e}_{1,t-j} \hat{e}_{1.2t}$$

with

$$\hat{e}_{1t} = \hat{\phi}(L) \Delta^\delta (y_t - \beta' x_t)$$

$$\hat{e}_{1.2t} = \hat{\phi}(L) \Delta^\delta (y_t - \beta' x_t) - c' \hat{\Phi}(L) \Delta^d x_t$$

$$\hat{\sigma}_{1.2} = \sqrt{\frac{1}{n} \sum_{t=1}^{n} \hat{e}_{1.2t}' \hat{e}_{1.2t}}.$$

A numerical approximation to the one sided test H_0 against $H_1 : \theta > 0$ is

$$\widehat{LM} = \frac{\sqrt{n} \sum_{t=1}^{n} \sum_{j=1}^{t-1} j^{-1} \hat{e}_{1,t-j} \hat{e}_{1.2t}}{\sqrt{\sum_{t=1}^{n} \left(\sum_{j=1}^{t-1} j^{-1} \hat{e}_{1,t-j} \right)^2 \sum_{t=1}^{n} \hat{e}_{1.2t}^2}} \tag{2.50}$$

which is to be compared with the quantiles of the standard normal distribution.

A fractionally cointegrated system can be estimated using the results of Dittmann (2004), which proposes an alternative to Granger (1986) error correction model for fractionally integrated systems. We can use a three steps procedure. Basically the first two step coincide with the procedure presented before: the residuals from (2.40) are calculated and the difference parameter $\hat{\delta}(\hat{u})$ estimated via the GPH method, for instance. In the third step one computes

$$\varpi_t \equiv \Delta^{\hat{\delta}(\hat{u})} \hat{u}_t \quad t = 1, \ldots, n \tag{2.51}$$

and verifies that ϖ_t is a stationary process, using for instance the KPSS test.

2.4   Empirical Application

We consider two different datasets. The first one is given by the Moody's Long-Term Corporate Bond Yield Averages. The second one consists of Lehman Brothers Eurodollar Indices.

Figure 2.7: Treasury, Aaa, Aa, A and Baa yields.
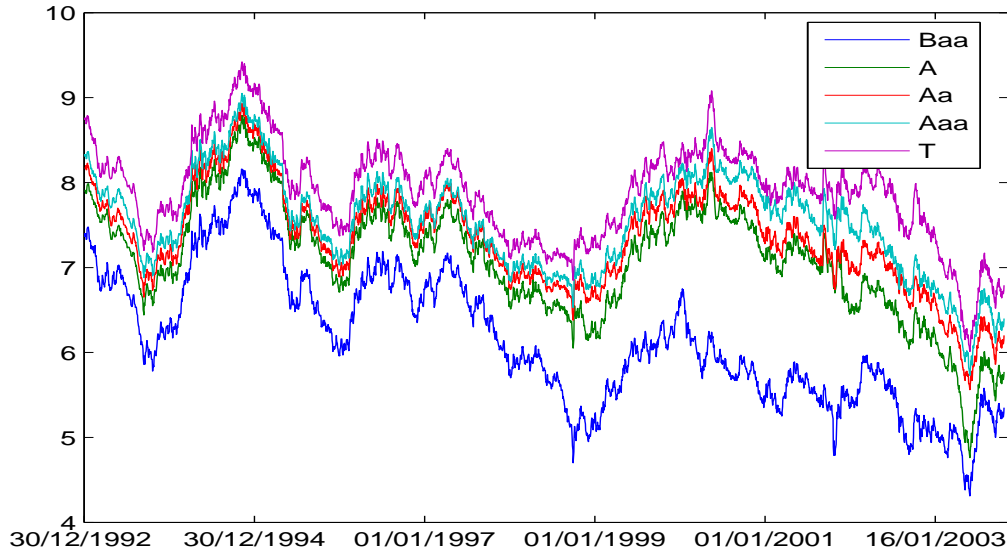
2.4.1   Dataset 1

First we employ the following data: 30-year Historical US Treasury Constant Maturity Yields and Moody's Aaa, Aa, A and Baa Long-Term Corporate Bond Yield Averages. The data covers the period from December 1992 to November 2003, for $n = 2703$ observations. Spreads are calculated as the difference between corporate yields and Treasury yields, as well as between different corporate yields. Thus we have 15 series: Treasury yields (denoted by T), corporate yields (Aaa, Aa, A, and Baa), spreads over Tresasury (sTAaa, sTAa, sTA, and sTBaa), spreads between corporate yields (sAaaAa, sAaaA, sAaaBaa, sAaA, sAaBaa, and sABaa). These series are plotted in Figure 2.7–Figure 2.9.

Table 2.3 and Table 2.4 report summary statistics and normality tests for the series involved and for their first differences respectively.

Tests for normality show that distributions of yields and spreads, and their first differences, are highly non-normal. Differences in credit spreads, with the exceptions of dsTreasAaa, dsAaBaa and dsABaa are positively skewed. This implies that the probability of a loss from the return is bigger than the probability of a loss from return from a normal distribution. Moreover, differences are also leptokurtic.

Table 2.5 and Table 2.6 report unit root and stationarity tests for yields and spreads and their first differences.

Yields and spreads seem to be non-stationary. We have ambiguous results only for sAaaAa and sABaa. In both cases the KPSS test rejects the null of stationarity, but for the former the null of a unit root is rejected by the PP test at a 1% level of significance, whereas for the latter the null of a unit root is rejected by the ADF with lag length 4 at a 5% level of significance and by the PP test at a 1% level of significance. However first differences of the series are clearly stationary, as shown by Table 2.6. As a consequence, we use the differenced series to estimate the difference parameter.

First we estimate the parameter d for yields and spreads using different values of m and J. In particular, when $J = 1$ (no pooling), $m = [n^\alpha]$ with $\alpha \in \{0.4, 0.5, 0.6, 0.7, 0.8\}$. As pointed out in Diebold and Inoue (2001) the choice of m is very important because even though a large value for m, would result in reducing standard error, this would induce bias in the estimator. This is because eq. (2.10), on which the log-periodogram estimate is based, is valid only for frequencies close to zero. Even though a popular choice is $m = \sqrt{n}$, many authors suggest $m = n^{4/5}$ (see for instance Hassler et al., 2006, pag. 189), which is the mean-square optimal choice. Table 2.7 and Table 2.8 report the d estimates for the yields and the spreads when $l = 0$.
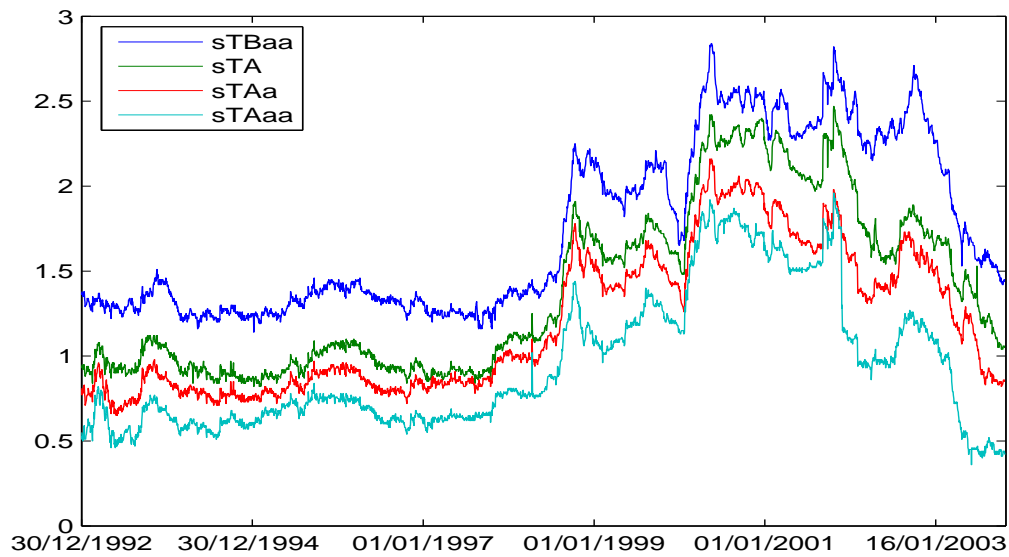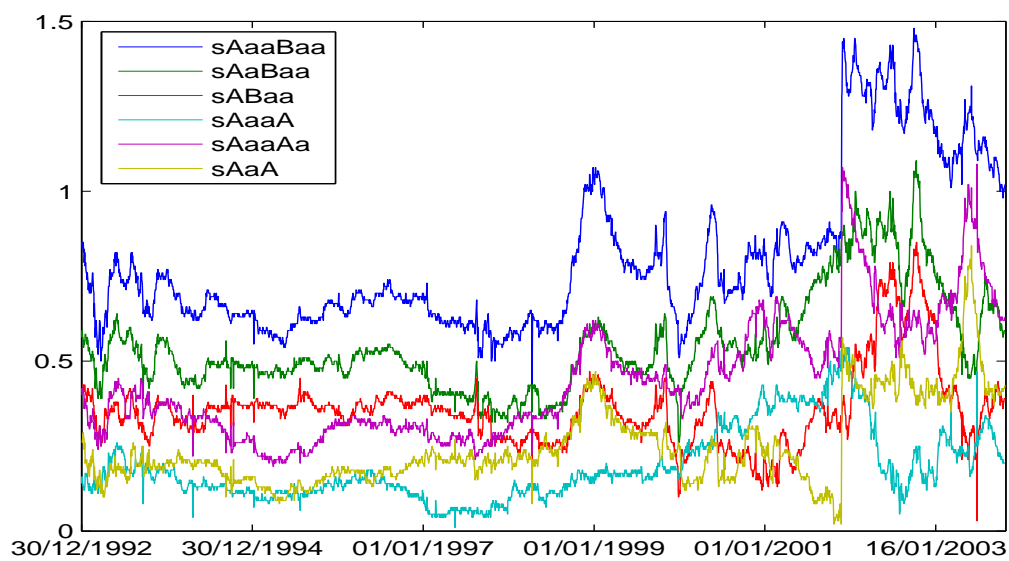
Figure 2.8: Aaa, Aa, A and Baa spreads over Treasury yields.

Figure 2.9: Spreads between corporate yields.

Table 2.3: Summary Statistics and Normality Tests for yields and spreads. Both the Jarque–Bera test and the Normality test proposed by Doornik and Hansen (1994) are computed. In both cases the null hypothesis is that the series is normally distributed and the test statistic is χ^2_2. The p-value is in square bracket. One asterisk denotes significance at 5% level and two asterisks denote significance at 1% level.

	mean	std	skewness	ex. Kurtosis	JB test	DH test
T	6.1536	0.78947	0.30498	-0.57126	78.655 [0.0000]**	140.41 [0.0000]**
Aaa	7.0975	0.68853	-0.40296	0.37445	88.943 [0.0000]**	69.261 [0.0000]**
Aa	7.3463	0.60072	-0.12787	0.079963	8.0862 [0.0175]*	7.5408 [0.0230]*
A	7.5368	0.58879	-0.1924	-0.11221	18.095 [0.0001]**	21.806 [0.0000]**
Baa	7.8947	0.57677	-0.23461	0.29	34.269 [0.0000]**	26.234 [0.0000]**
sTAaa	0.94396	0.40039	0.85386	-0.47124	353.46 [0.0000]**	1291.5 [0.0000]**
sTAa	1.1927	0.41573	0.54433	-1.1281	276.81 [0.0000]**	934.56 [0.0000]**
sTA	1.3833	0.49232	0.61294	-1.0284	288.36 [0.0000]**	1033.2 [0.0000]**
sTBaa	1.7412	0.49927	0.51375	-1.3408	321.38 [0.0000]**	1197.4 [0.0000]**
sAaaAa	0.24873	0.12367	1.3572	2.0211	1289.9 [0.0000]**	1343 [0.0000]**
sAaaA	0.43932	0.17651	1.0234	0.68153	524.14 [0.0000]**	901.43 [0.0000]**
sAaaBaa	0.79721	0.2279	1.2453	0.5105	727.98 [0.0000]**	2312.4 [0.0000]**
sAaA	0.19058	0.10543	1.0911	0.49992	564.49 [0.0000]**	1330.5 [0.0000]**
sAaBaa	0.54848	0.14479	1.1672	1.0723	743.29 [0.0000]**	1182.2 [0.0000]**
sABaa	0.35789	0.11427	1.5693	3.7498	2693.1 [0.0000]**	1260.6 [0.0000]**

Table 2.4: Summary Statistics and Normality Tests for first differences of yields and spreads. Both the Jarque–Bera test and the Normality test proposed by Doornik and Hansen (1994) are computed. In both cases the null hypothesis is that the series is normally distributed and the test statistic is χ_2^2. The p-value is in square bracket. One asterisk denotes significance at 5% level and two asterisks denote significance at 1% level.

	mean	std	skewness	ex. Kurtosis	JB test	DH test
ΔT	-0.00074	0.051925	0.30619	2.0323	507.20 [0.0000]**	254.79 [0.0000]**
ΔAaa	-0.00079	0.048771	0.31655	2.6307	824.26 [0.0000]**	382.40 [0.0000]**
ΔAa	-0.00073	0.046392	0.43888	1.9246	503.78 [0.0000]**	210.14 [0.0000]**
ΔA	-0.00071	0.04719	0.43378	2.0412	553.84 [0.0000]**	230.94 [0.0000]**
ΔBaa	-0.00072	0.047793	0.42962	1.9843	526.41 [0.0000]**	221.83 [0.0000]**
ΔsTAaa	-4.81E-05	0.024443	-2.0015	55.643	350377.5 [0.0000]**	12448 [0.0000]**
ΔsTAa	1.48E-05	0.022006	0.45617	8.3799	7999.7 [0.0000]**	1970.7 [0.0000]**
ΔsTA	3.70E-05	0.023	0.57711	14.692	24451.1 [0.0000]**	3923.6 [0.0000]**
ΔsTBaa	2.22E-05	0.022466	0.5992	9.0243	9330.2 [0.0000]**	2012.1 [0.0000]**
ΔsAaaAa	6.29E-05	0.01623	5.8072	144.87	2.3780E+06 [0.0000]**	6833 [0.0000]**
ΔsAaaA	8.51E-05	0.017467	4.9598	130.89	1.9397E+06 [0.0000]**	10617 [0.0000]**
ΔsAaaBaa	7.03E-05	0.018208	4.3241	112.28	1.4277E+06 [0.0000]**	11489 [0.0000]**
ΔsAaA	2.22E-05	0.013233	0.66513	52.071	3.0545E+05 [0.0000]**	16289 [0.0000]**
ΔsAaBaa	7.40E-06	0.014959	-0.19375	10.042	11371 [0.0000]**	2719.2 [0.0000]**
ΔsABaa	-1.48E-05	0.01553	-0.10372	28.283	90065 [0.0000]**	8920.3 [0.0000]**

Table 2.5: Unit root and stationarity tests for yields and spreads. The Dickey-Fuller (DF) or augmented Dickey-Fuller with the constant (ADF), Phillips-Perron with the constant (PP), a two KPSS tests without trend are carried out. n is the lag length in the ADF and it is chosen by the AIC. In the first KPSS test the Bartlett kernel with bandwidth parameter $\left[4\left(\frac{n}{100}\right)^{1/4}\right]$ is chosen for the estimation of the long run variance. In the second test the automatic bandwidth selection procedure of Hobijn et al. (1998) is considered. One asterisk denotes significance at 5% level and two asterisks denote significance at 1% level.

	DF-ADF	n	KPSS	KPSS HFO	PP
T	-1.8998	0	18.272**	4.8192**	-1.8811
Aaa	-1.6007	1	9.8415**	2.6134**	-1.5274
Aa	-1.8819	1	8.5263**	2.2743**	-1.771
A	-1.9308	1	6.0752**	1.623**	-1.8066
Baa	-2.0792	1	5.3553**	1.4406**	-2.0162
sTAaa	-1.4212	39	12.07**	3.16**	-1.2784
sTAa	-1.415	96	17.553**	4.5677**	-1.253
sTA	-1.4234	73	17.974**	4.6607**	-1.1279
sTBaa	-1.219	4	19.013**	4.9377**	-1.1825
sAaaAa	-2.37	41	11.285**	3.0905**	-2.8852**
sAaaA	-1.9799	1	18.582**	4.9404**	-2.0816
sAaaBaa	-1.6043	57	15.306**	4.0569**	-1.983
sAaA	-1.9006	10	12.082**	3.1866**	-2.2683
sAaBaa	-2.292	1	11.63**	3.121**	-2.4266
sABaa	-3.0041*	4	2.3387**	0.63969**	-3.1083**

Table 2.6: Unit root and stationarity tests for the first differences of yields and spreads. The Dickey-Fuller (DF) or augmented Dickey-Fuller with the constant (ADF), Phillips-Perron with the constant (PP), a two KPSS tests without trend are carried out. n is the lag length in the ADF and it is chosen by the AIC. In the first KPSS test the Bartlett kernel with bandwidth parameter $\left[4\left(\frac{n}{100}\right)^{1/4}\right]$ is chosen for the estimation of the long run variance. In the second test the automatic bandwidth selection procedure of Hobijn et al. (1998) is considered. One asterisk denotes significance at 5% level and two asterisks denote significance at 1% level.

	DF-ADF	n	KPSS	KPSS HFO	PP
ΔT	-50.796**	0	0.038103	0.037796	-50.788**
ΔAaa	-50.254**	0	0.065246	0.06474	-50.225**
ΔAa	-49.933**	0	0.05892	0.05892	-49.894**
ΔA	-50.361**	0	0.066975	0.065249	-50.345**
ΔBaa	-50.116**	0	0.061583	0.059524	-50.087**
ΔsTAaa	-7.5675**	38	0.30846	0.3141	-63.294**
ΔsTAa	-5.8175**	95	0.2924	0.2924	-62.675**
ΔsTA	-5.3436**	72	0.34606	0.35011	-64.169**
ΔsTBaa	-23.651**	3	0.26461	0.24729	-59.851**
ΔsAaaAa	-9.4063**	40	0.035604	0.035017	-62.101**
ΔsAaaA	-63.273**	0	0.038956	0.036094	-63.162**
ΔsAaaBaa	-8.3088**	56	0.059622	0.059054	-58.503**
ΔsAaA	-18.297**	9	0.047834	0.051928	-73.938**
ΔsAaBaa	-60.905**	0	0.043784	0.044609	-60.648**
ΔsABaa	-64.597**	0	0.035307	0.032132	-64.219**

Table 2.7: d estimates for the yields with $l = 0$ and non-tapered data. For every series X, this table report the estimates $\hat{d}(\Delta X) \equiv \widehat{d-1}(\Delta X) + 1$ along with the test statistic (2.15). One asterisk denotes significance at 5% level and two asterisks denote significance at 1% level.

			T	Aaa	Aa	A	Baa
	$m = 23$	$\hat{d}(\Delta X)$	0.80087	0.97492	0.88708	0.97457	0.91151
		$\tau_{d=1}$	-1.2036	-0.15158	-0.68253	-0.15374	-0.53491
	$m = 51$	$\hat{d}(\Delta X)$	0.84594	0.94272	0.87103	0.91001	0.93016
		$\tau_{d=1}$	-1.5125	-0.56232	-1.2661	-0.88346	-0.68566
$J = 1$	$m = 114$	$\hat{d}(\Delta X)$	0.96895	1.0195	0.92161	0.97034	0.97393
		$\tau_{d=1}$	-0.48028	0.30185	-1.2127	-0.45883	-0.40322
	$m = 252$	$\hat{d}(\Delta X)$	0.9419	0.98311	0.92178	0.94766	0.99618
		$\tau_{d=1}$	-1.3749	-0.39957	-1.851	-1.2385	-0.090393
	$m = 556$	$\hat{d}(\Delta X)$	0.97927	0.99219	0.97855	0.98767	1.0323
		$\tau_{d=1}$	-0.7327	-0.27612	-0.75813	-0.43585	1.141
	$m = 22$	$\hat{d}(\Delta X)$	1.0896	1.1732	1.2483	1.2048	1.1473
		$\tau_{d=1}$	0.3291	0.63644	0.91224	0.75256	0.54123
	$m = 50$	$\hat{d}(\Delta X)$	0.91791	1.0427	1.0301	1.031	1.019
		$\tau_{d=1}$	-0.52307	0.27201	0.19174	0.19757	0.12082
$J = 2$	$m = 114$	$\hat{d}(\Delta X)$	1.0018	1.0379	1.0097	1.0277	1.0257
		$\tau_{d=1}$	0.01843	0.39641	0.10166	0.28978	0.2693
	$m = 252$	$\hat{d}(\Delta X)$	0.97895	1.0072	0.98284	0.99102	1.0219
		$\tau_{d=1}$	-0.343	0.11798	-0.27969	-0.14631	0.35682
	$m = 556$	$\hat{d}(\Delta X)$	0.99833	1.0078	1.0004	1.0049	1.0229
		$\tau_{d=1}$	-0.041077	0.19234	0.010702	0.12011	0.56376
	$m = 21$	$\hat{d}(\Delta X)$	0.87718	1.1111	1.0974	1.124	1.0247
		$\tau_{d=1}$	-0.32127	0.29064	0.25485	0.32441	0.064606
	$m = 51$	$\hat{d}(\Delta X)$	0.83253	0.99082	0.96514	0.97111	0.95098
		$\tau_{d=1}$	-0.83107	-0.04555	-0.17298	-0.14337	-0.24328
$J = 3$	$m = 114$	$\hat{d}(\Delta X)$	0.98173	1.0239	0.99281	1.0057	1.015
		$\tau_{d=1}$	-0.15055	0.19682	-0.059257	0.046875	0.12323
	$m = 252$	$\hat{d}(\Delta X)$	0.96576	0.99235	0.97424	0.97644	1.0134
		$\tau_{d=1}$	-0.44573	-0.099633	-0.33529	-0.30662	0.1748
	$m = 555$	$\hat{d}(\Delta X)$	0.97733	0.99088	0.98522	0.9814	1.0062
		$\tau_{d=1}$	-0.44888	-0.18057	-0.29255	-0.36831	0.12348

Table 2.8: d estimates for the spreads with $l = 0$ and non-tapered data. For every series X, this table report the estimates $\hat{d}(\Delta X) \equiv \widehat{d-1}(\Delta X) + 1$ along with the test statistic (2.15). One asterisk denotes significance at 5% level and two asterisks denote significance at 1% level.

			sTAaa	sTAa	sTA	sTBaa	sAaaAa	sAaaA	sAaaBaa	sAaA	sAaBaa	sABaa
	$m = 23$	$\hat{d}(\Delta X)$	0.92097	1.05	1.027	0.99384	0.67217	0.65317	0.90996	0.98424	0.87955	1.0693
		$\tau_{d=1}$	-0.47769	0.30223	0.16305	-0.037205	-1.9816*	-2.0965*	-0.54425	-0.095236	-0.7281	0.4191
	$m = 51$	$\hat{d}(\Delta X)$	1.0332	1.0306	1.072	1.1047	0.80634	0.82961	0.95578	0.994	0.81231	0.91661
		$\tau_{d=1}$	0.32557	0.30043	0.70673	1.0283	-1.9013	-1.6728	-0.43409	-0.058928	-1.8426	-0.81866
$J = 1$	$m = 114$	$\hat{d}(\Delta X)$	1.0458	0.99239	1.0342	1.0863	0.87859	0.91356	0.97636	0.99051	0.97782	0.88735
		$\tau_{d=1}$	0.70796	-0.11771	0.52927	1.3347	-1.8782	-1.3371	-0.36577	-0.14687	-0.34314	-1.7426
	$m = 252$	$\hat{d}(\Delta X)$	1.0549	1.0704	1.0868	1.1103	0.98959	1.0024	1.0507	0.96664	1.0448	0.97755
		$\tau_{d=1}$	1.2982	1.6663	2.0547*	2.6091**	-0.24635	0.057549	1.1992	-0.78952	1.0607	-0.53117
	$m = 556$	$\hat{d}(\Delta X)$	1.0348	0.99189	0.9903	1.0914	0.97802	0.96553	1.0733	0.90041	0.98604	0.98274
		$\tau_{d=1}$	1.2285	-0.28653	-0.34269	3.2291**	-0.77674	-1.2181	2.59**	-3.5191**	-0.49324	-0.61002
	$m = 22$	$\hat{d}(\Delta X)$	0.84579	0.83857	0.98118	0.92743	0.69372	0.60359	0.79276	0.98821	0.71874	0.85606
		$\tau_{d=1}$	-0.56658	-0.59311	-0.069149	-0.26662	-1.1253	-1.4565	-0.76143	-0.043302	-1.0334	-0.52884
	$m = 50$	$\hat{d}(\Delta X)$	1.0075	1.0331	1.1084	1.1012	0.80097	0.84875	0.95023	0.98799	0.77459	0.93089
		$\tau_{d=1}$	0.047495	0.21117	0.69051	0.6447	-1.2683	-0.96381	-0.31713	-0.07656	-1.4364	-0.44038
$J = 2$	$m = 114$	$\hat{d}(\Delta X)$	1.0607	1.0256	1.0642	1.0635	0.89574	0.9342	0.93122	0.9979	0.9928	0.9255
		$\tau_{d=1}$	0.63486	0.26751	0.67237	0.66457	-1.0912	-0.68868	-0.71985	-0.022027	-0.075375	-0.77971
	$m = 252$	$\hat{d}(\Delta X)$	1.0906	1.0847	1.1043	1.1113	0.99528	1.0063	1.0284	0.95773	1.0317	1.0113
		$\tau_{d=1}$	1.4769	1.3808	1.6996	1.8135	-0.076846	0.10267	0.46259	-0.6889	0.51733	0.18443
	$m = 556$	$\hat{d}(\Delta X)$	1.0334	0.99621	1.0075	1.0822	0.98171	0.96581	1.0503	0.88844	0.98291	0.98403
		$\tau_{d=1}$	0.82167	-0.093305	0.18488	2.0223*	-0.44978	-0.84092	1.2378	-2.7438**	-0.42024	-0.39273
	$m = 21$	$\hat{d}(\Delta X)$	1.0932	1.0585	1.2068	1.0247	0.68511	0.53466	0.83354	0.94063	0.72614	0.93701
		$\tau_{d=1}$	0.24371	0.15295	0.54093	0.064676	-0.82366	-1.2172	-0.43542	-0.1553	-0.71632	-0.16477
	$m = 51$	$\hat{d}(\Delta X)$	1.0793	1.0999	1.1471	1.1204	0.80996	0.82075	0.96278	0.98229	0.75197	0.94084
		$\tau_{d=1}$	0.39336	0.49595	0.72994	0.59768	-0.94307	-0.88952	-0.1847	-0.087908	-1.2309	-0.29361
$J = 3$	$m = 114$	$\hat{d}(\Delta X)$	1.0777	1.0638	1.0863	1.085	0.87889	0.92741	0.91604	0.97628	0.98823	0.92951
		$\tau_{d=1}$	0.64038	0.52583	0.71137	0.7004	-0.99806	-0.59817	-0.69193	-0.1955	-0.096991	-0.5809
	$m = 252$	$\hat{d}(\Delta X)$	1.0853	1.0891	1.0949	1.1083	0.99974	1.0001	1.0329	0.94609	1.0055	1.0203
		$\tau_{d=1}$	1.1101	1.1597	1.2348	1.4097	-0.0034405	0.00074993	0.42821	-0.70174	0.072236	0.26444
	$m = 555$	$\hat{d}(\Delta X)$	1.0304	1.0037	0.99824	1.0812	0.98445	0.95898	1.0474	0.87589	0.9791	0.98987
		$\tau_{d=1}$	0.60192	0.072986	-0.03487	1.6083	-0.30782	-0.81216	0.93849	-2.457*	-0.4138	-0.20047

Table 2.9 and Table 2.10 report the d estimates for the yields and the spreads when $l = 1$.

Table 2.9: d estimates for the yields with $l = 1$ and non-tapered data. For every series X, this table report the estimates $\hat{d}(\Delta X) \equiv \widehat{d-1}(\Delta X) + 1$ along with the test statistic (2.15). One asterisk denotes significance at 5% level and two asterisks denote significance at 1% level.

			T	Aaa	Aa	A	Baa
	$m = 23$	$\hat{d}(\Delta X)$	1.0353	1.0736	0.89097	0.99382	0.93387
		$\tau_{d=1}$	0.17203	0.35853	-0.53124	-0.030098	-0.32219
	$m = 51$	$\hat{d}(\Delta X)$	0.94903	0.97784	0.8668	0.90459	0.94177
		$\tau_{d=1}$	-0.43916	-0.19088	-1.1476	-0.82201	-0.50171
$J = 1$	$m = 114$	$\hat{d}(\Delta X)$	1.0294	1.0439	0.92612	0.97535	0.98441
		$\tau_{d=1}$	0.42055	0.62844	-1.0566	-0.35257	-0.22294
	$m = 252$	$\hat{d}(\Delta X)$	0.96888	0.99244	0.92397	0.94852	1.0026
		$\tau_{d=1}$	-0.70267	-0.17059	-1.7165	-1.1622	0.058915
	$m = 556$	$\hat{d}(\Delta X)$	0.99411	0.99716	0.98197	0.98975	1.037
		$\tau_{d=1}$	-0.20233	-0.097479	-0.61933	-0.35224	1.2698
	$m = 23$	$\hat{d}(\Delta X)$	1.0026	1.1408	1.1049	1.0806	1.025
		$\tau_{d=1}$	0.008386	0.44602	0.33236	0.25526	0.079348
	$m = 51$	$\hat{d}(\Delta X)$	0.92601	1.0564	1.0201	0.98881	0.99223
		$\tau_{d=1}$	-0.42809	0.32632	0.11613	-0.064734	-0.044986
$J = 2$	$m = 115$	$\hat{d}(\Delta X)$	1.0316	1.0563	1.0043	1.0119	1.0314
		$\tau_{d=1}$	0.3109	0.55491	0.041876	0.11683	0.30894
	$m = 253$	$\hat{d}(\Delta X)$	0.98455	1.0027	0.98385	0.97935	1.0147
		$\tau_{d=1}$	-0.24245	0.041635	-0.25345	-0.3241	0.23073
	$m = 557$	$\hat{d}(\Delta X)$	0.98428	0.98638	0.98776	0.97939	1.0106
		$\tau_{d=1}$	-0.37764	-0.32737	-0.29403	-0.49522	0.25581
	$m = 22$	$\hat{d}(\Delta X)$	1.3385	1.2626	1.3282	1.2539	1.2117
		$\tau_{d=1}$	0.7788	0.60426	0.75521	0.5842	0.48704
	$m = 52$	$\hat{d}(\Delta X)$	0.9806	1.0903	1.0449	1.0139	1.012
		$\tau_{d=1}$	-0.08864	0.41274	0.20518	0.063583	0.054933
$J = 3$	$m = 115$	$\hat{d}(\Delta X)$	1.0333	1.0352	1.0093	1.0102	1.0356
		$\tau_{d=1}$	0.26047	0.27487	0.072759	0.079342	0.27822
	$m = 253$	$\hat{d}(\Delta X)$	0.99747	1.0067	0.99208	0.99224	1.0216
		$\tau_{d=1}$	-0.031876	0.084883	-0.099657	-0.097712	0.272
	$m = 556$	$\hat{d}(\Delta X)$	1.0008	1.0079	1.0011	0.99927	1.0173
		$\tau_{d=1}$	0.01551	0.15227	0.022004	-0.014192	0.3351

Table 2.11 and Table 2.12 report the d estimates for the yields and the spreads when $l = 0$ and the data is tapered.

Table 2.13 and Table 2.14 report the d estimates for the yields and the spreads when $l = 1$ and the data is tapered.

From the analysis of Table 2.7 through Table 2.14 it is clear that yields and spreads are likely to be long memory and nonstationary processes. In each table we report also the test statistic for the null hypothesis of $d = 1$. First of all, the null is never rejected for the yields series, no matter if we consider tapered or non-tapered data, $l = 0$ or $l = 1$, $J = 1$, $J = 2$ or $J = 3$. As far as spreads are concerned, the number of rejections of the null increases when moving from non-tapered to tapered data. On the other hand, this

Table 2.10: d estimates for the spreads with $l = 1$ and non-tapered data. For every series X, this table report the estimates $\hat{d}(\Delta X) \equiv \widehat{d-1}(\Delta X) + 1$ along with the test statistic (2.15). One asterisk denotes significance at 5% level and two asterisks denote significance at 1% level.

			sTAaa	sTAa	sTA	sTBaa	sAaaAa	sAaaA	sAaaBaa	sAaA	sAaBaa	sABaa
	$m=23$	$\hat{d}(\Delta X)$	0.74987	0.90468	0.85302	0.87147	0.80844	0.52991	0.91703	0.99587	0.83407	1.3064
		$\tau_{d=1}$	-1.2187	-0.46443	-0.71613	-0.62622	-0.93333	-2.2904*	-0.40423	-0.020122	-0.80843	1.4926
	$m=51$	$\hat{d}(\Delta X)$	0.98108	0.96415	1.0047	1.0691	0.88827	0.80681	0.96923	1.0014	0.78319	0.98951
		$\tau_{d=1}$	-0.16299	-0.30889	0.040815	0.59539	-0.96263	-1.6645	-0.2651	0.011927	-1.868	-0.090371
$J=1$	$m=114$	$\hat{d}(\Delta X)$	1.024	0.95789	0.99952	1.0671	0.92361	0.91376	0.98493	0.99356	0.9812	0.91611
		$\tau_{d=1}$	0.34372	-0.60219	-0.006876	0.96001	-1.0925	-1.2334	-0.21554	-0.092087	-0.26894	-1.1997
	$m=252$	$\hat{d}(\Delta X)$	1.0452	1.059	1.0735	1.1025	1.0187	1.0086	1.0599	0.96614	1.0507	0.99742
		$\tau_{d=1}$	1.0194	1.3323	1.66	2.315*	0.42313	0.19378	1.3516	-0.76431	1.1452	-0.058244
	$m=556$	$\hat{d}(\Delta X)$	1.0292	0.98286	0.97963	1.0867	0.99183	0.967	1.0788	0.89724	0.98653	0.99265
		$\tau_{d=1}$	1.003	-0.58871	-0.69992	2.9775**	-0.28077	-1.1336	2.708**	-3.5303**	-0.46291	-0.25249
	$m=23$	$\hat{d}(\Delta X)$	0.93695	0.93291	1.0259	0.88609	0.70816	0.55502	0.84467	1.0097	0.71308	1.0857
		$\tau_{d=1}$	-0.19975	-0.21254	0.081996	-0.36087	-0.92458	-1.4097	-0.4921	0.030827	-0.90899	0.27151
	$m=51$	$\hat{d}(\Delta X)$	1.0691	1.0495	1.1185	1.096	0.86277	0.83437	0.97586	1.0125	0.79056	0.99666
		$\tau_{d=1}$	0.3999	0.28669	0.68558	0.55523	-0.79403	-0.95836	-0.13968	0.072256	-1.2118	-0.019327
$J=2$	$m=115$	$\hat{d}(\Delta X)$	1.0495	1.0361	1.0838	1.0714	0.93674	0.95939	0.95783	0.96523	0.97625	0.94174
		$\tau_{d=1}$	0.48755	0.35526	0.82554	0.70339	-0.62323	-0.40004	-0.41539	-0.34249	-0.23392	-0.57398
	$m=253$	$\hat{d}(\Delta X)$	1.0602	1.0759	1.0841	1.0892	1.0275	1.0134	1.0508	0.94644	1.0174	1.0334
		$\tau_{d=1}$	0.94483	1.1915	1.3194	1.399	0.43186	0.21027	0.79669	-0.84045	0.2729	0.52394
	$m=557$	$\hat{d}(\Delta X)$	1.0254	0.9916	0.98942	1.0748	0.98586	0.95837	1.0537	0.87611	0.9766	0.99081
		$\tau_{d=1}$	0.61105	-0.20193	-0.2543	1.7968	-0.33985	-1.0004	1.2893	-2.9768**	-0.56217	-0.22084
	$m=22$	$\hat{d}(\Delta X)$	0.82371	0.73205	0.89577	0.76287	0.62827	0.57538	0.80618	1.0203	0.61229	0.97652
		$\tau_{d=1}$	-0.40559	-0.6165	-0.2398	-0.54557	-0.85525	-0.97695	-0.44592	0.046625	-0.89202	-0.054012
	$m=52$	$\hat{d}(\Delta X)$	1.0593	1.0357	1.1252	1.0662	0.82906	0.86221	0.98179	0.99795	0.75392	0.96371
		$\tau_{d=1}$	0.27088	0.16321	0.57235	0.30249	-0.78121	-0.62973	-0.083205	-0.0093506	-1.1247	-0.16586
$J=3$	$m=115$	$\hat{d}(\Delta X)$	1.0742	1.0027	1.0552	1.0367	0.93158	0.95869	0.96859	0.98545	0.97857	0.95627
		$\tau_{d=1}$	0.58004	0.021021	0.4311	0.2871	-0.53473	-0.32284	-0.24551	-0.11371	-0.1675	-0.34179
	$m=253$	$\hat{d}(\Delta X)$	1.0805	1.0819	1.0811	1.0863	1.0135	1.0144	1.0398	0.9437	1.0301	1.0493
		$\tau_{d=1}$	1.0135	1.0316	1.0212	1.0866	0.16971	0.18157	0.50066	-0.70873	0.3787	0.62016
	$m=555$	$\hat{d}(\Delta X)$	1.0246	0.99001	0.99484	1.071	0.97738	0.9567	1.0557	0.86794	0.99075	0.99221
		$\tau_{d=1}$	0.47749	-0.19368	-0.10012	1.3757	-0.43851	-0.83947	1.0799	-2.5603*	-0.17931	-0.15106

Table 2.11: d estimates for the yields with $l = 0$ and tapered data. For every series X, this table report the estimates $\hat{d}(\Delta X) \equiv \widehat{d-1}(\Delta X) + 1$ along with the test statistic (2.15). One asterisk denotes significance at 5% level and two asterisks denote significance at 1% level.

			T	Aaa	Aa	A	Baa
	$m = 23$	$\hat{d}(\Delta X)$	1.0794	1.1038	1.0122	1.0886	1.0329
		$\tau_{d=1}$	0.47993	0.62732	0.073489	0.53561	0.19873
	$m = 51$	$\hat{d}(\Delta X)$	0.9769	0.99086	0.94898	0.97784	0.97821
		$\tau_{d=1}$	-0.22673	-0.089711	-0.50083	-0.21758	-0.21394
$J = 1$	$m = 114$	$\hat{d}(\Delta X)$	1.0054	0.9199	0.95393	0.97763	0.97777
		$\tau_{d=1}$	0.084006	-1.2391	-0.71274	-0.34599	-0.34382
	$m = 252$	$\hat{d}(\Delta X)$	0.98594	0.94146	0.97549	0.96459	0.99369
		$\tau_{d=1}$	-0.3326	-1.3853	-0.58009	-0.83783	-0.14941
	$m = 556$	$\hat{d}(\Delta X)$	0.9836	0.95627	0.99043	0.9805	1.0066
		$\tau_{d=1}$	-0.57937	-1.5454	-0.33825	-0.689	0.23245
	$m = 22$	$\hat{d}(\Delta X)$	1.0903	1.1548	1.0586	1.1207	1.0611
		$\tau_{d=1}$	0.33182	0.56892	0.21532	0.44332	0.2246
	$m = 50$	$\hat{d}(\Delta X)$	0.99765	1.0371	0.98705	0.99847	0.99164
		$\tau_{d=1}$	-0.015005	0.23667	-0.082552	-0.009731	-0.053293
$J = 2$	$m = 114$	$\hat{d}(\Delta X)$	1.0041	0.93739	0.95375	0.95836	0.97683
		$\tau_{d=1}$	0.042722	-0.65527	-0.4841	-0.43581	-0.24247
	$m = 252$	$\hat{d}(\Delta X)$	0.98426	0.95082	0.9613	0.95778	0.96824
		$\tau_{d=1}$	-0.25658	-0.80147	-0.63075	-0.68805	-0.5176
	$m = 556$	$\hat{d}(\Delta X)$	0.98484	0.96438	0.9765	0.97148	0.99067
		$\tau_{d=1}$	-0.37275	-0.87604	-0.57799	-0.70142	-0.22945
	$m = 21$	$\hat{d}(\Delta X)$	1.1436	1.2969	1.1541	1.1949	1.0546
		$\tau_{d=1}$	0.3755	0.77655	0.40303	0.50977	0.14291
	$m = 51$	$\hat{d}(\Delta X)$	0.95943	0.9916	0.97575	0.97473	0.96302
		$\tau_{d=1}$	-0.20131	-0.041709	-0.12033	-0.12542	-0.18352
$J = 3$	$m = 114$	$\hat{d}(\Delta X)$	1.0032	0.92731	0.96023	0.96149	0.98751
		$\tau_{d=1}$	0.026266	-0.599	-0.32773	-0.31739	-0.10294
	$m = 252$	$\hat{d}(\Delta X)$	0.99589	0.95255	0.96919	0.96247	0.97723
		$\tau_{d=1}$	-0.053537	-0.61759	-0.401	-0.48846	-0.29638
	$m = 555$	$\hat{d}(\Delta X)$	0.97881	0.96188	0.97832	0.96462	0.98357
		$\tau_{d=1}$	-0.41944	-0.75476	-0.42926	-0.70044	-0.32534

Table 2.12: d estimates for the spreads with $l = 0$ and tapered data. For every series X, this table report the estimates $\hat{d}(\Delta X) \equiv \widehat{d-1}(\Delta X) + 1$ along with the test statistic (2.15). One asterisk denotes significance at 5% level and two asterisks denote significance at 1% level.

			sTAaa	sTAa	sTA	sTBaa	sAaaAa	sAaaA	sAaaBaa	sAaA	sAaBaa	sABaa
	$m=23$	$\hat{d}(\Delta X)$	0.97058	0.96698	1.0416	0.8911	0.69883	0.8081	0.60979	1.1058	0.77158	0.59618
		$\tau_{d=1}$	-0.17786	-0.1996	0.2512	-0.65824	-1.8204	-1.16	-2.3586*	0.63965	-1.3807	-2.4409*
	$m=51$	$\hat{d}(\Delta X)$	1.0374	1.065	1.1789	1.1409	0.89207	1.0805	1.0218	0.88402	0.88381	0.88352
		$\tau_{d=1}$	0.36733	0.63831	1.7565	1.3829	-1.0596	0.79042	0.21421	-1.1386	-1.1407	-1.1435
$J=1$	$m=114$	$\hat{d}(\Delta X)$	1.1106	1.071	1.1172	1.128	0.93064	0.97942	1.0287	0.97902	0.96461	0.85941
		$\tau_{d=1}$	1.7107	1.0986	1.8124	1.9799*	-1.073	-0.31839	0.44416	-0.32454	-0.5474	-2.1749*
	$m=252$	$\hat{d}(\Delta X)$	1.132	1.1836	1.186	1.1225	1.0037	0.99484	1.0374	0.93534	0.95569	0.92352
		$\tau_{d=1}$	3.1225**	4.3453**	4.4004**	2.8991**	0.087883	-0.12205	0.88423	-1.5301	-1.0485	-1.8097
	$m=556$	$\hat{d}(\Delta X)$	1.02	1.0401	1.063	1.0987	0.91555	0.94069	0.97827	0.83367	0.96585	0.95882
		$\tau_{d=1}$	0.70726	1.4167	2.2276*	3.4893**	-2.9841**	-2.0958*	-0.76778	-5.8775**	-1.2067	-1.4552
	$m=22$	$\hat{d}(\Delta X)$	0.96895	0.98127	1.0732	0.86534	0.77152	0.7885	0.56508	1.2826	0.73123	0.55281
		$\tau_{d=1}$	-0.11409	-0.06881	0.26884	-0.49475	-0.83946	-0.77708	-1.598	1.0383	-0.9875	-1.6431
	$m=50$	$\hat{d}(\Delta X)$	1.0326	1.0731	1.1363	1.1284	0.94519	1.1126	1.0066	0.97601	0.87644	0.8499
		$\tau_{d=1}$	0.20771	0.46576	0.86871	0.81818	-0.34925	0.71751	0.041853	-0.15285	-0.78738	-0.95646
$J=2$	$m=114$	$\hat{d}(\Delta X)$	1.0805	1.0439	1.1015	1.0938	0.96059	1.0102	0.99244	1.0037	0.92699	0.84934
		$\tau_{d=1}$	0.84255	0.45924	1.0624	0.9813	-0.41247	0.10692	-0.079132	0.038292	-0.76416	-1.5769
	$m=252$	$\hat{d}(\Delta X)$	1.1098	1.1472	1.1692	1.1222	0.99875	1.011	1.0186	0.93698	0.942	0.90653
		$\tau_{d=1}$	1.7894	2.3985*	2.7579**	1.9915*	-0.020376	0.17942	0.30339	-1.0269	-0.94523	-1.5233
	$m=556$	$\hat{d}(\Delta X)$	1.0183	1.0255	1.0658	1.0917	0.90921	0.94521	0.99247	0.84796	0.9632	0.94889
		$\tau_{d=1}$	0.45093	0.62662	1.6177	2.2556*	-2.233*	-1.3476	-0.1852	-3.7395**	-0.90508	-1.257
	$m=21$	$\hat{d}(\Delta X)$	0.95851	0.96511	1.0426	0.84411	0.7998	0.78809	0.65575	1.2381	0.67676	0.61542
		$\tau_{d=1}$	-0.10852	-0.091265	0.1114	-0.40777	-0.52367	-0.5543	-0.90044	0.6229	-0.8455	-1.006
	$m=51$	$\hat{d}(\Delta X)$	1.0536	1.1105	1.17	1.1285	0.92703	1.0581	1.0939	0.93573	0.86702	0.90717
		$\tau_{d=1}$	0.26585	0.54859	0.84367	0.63778	-0.36215	0.28831	0.46586	-0.31894	-0.65993	-0.46067
$J=3$	$m=114$	$\hat{d}(\Delta X)$	1.0889	1.0992	1.144	1.1274	0.92305	0.98988	0.99682	1.0123	0.92951	0.85291
		$\tau_{d=1}$	0.7324	0.8172	1.1863	1.0497	-0.63411	-0.083381	-0.02624	0.10151	-0.58086	-1.2122
	$m=252$	$\hat{d}(\Delta X)$	1.1289	1.1832	1.1992	1.1519	0.99718	1.004	1.018	0.93562	0.92674	0.91203
		$\tau_{d=1}$	1.678	2.3846*	2.5927*	1.9766*	-0.036721	0.052639	0.23472	-0.83792	-0.95354	-1.145
	$m=555$	$\hat{d}(\Delta X)$	1.027	1.0473	1.076	1.1006	0.90895	0.93513	0.99848	0.83203	0.96065	0.95801
		$\tau_{d=1}$	0.53412	0.93722	1.5036	1.9924*	-1.8026	-1.2843	-0.030047	-3.3255**	-0.77899	-0.83126

Table 2.13: d estimates for the yields with $l=1$ and tapered data. For every series X, this table report the estimates $\hat{d}(\Delta X) \equiv \widehat{d-1}(\Delta X)+1$ along with the test statistic (2.15). One asterisk denotes significance at 5% level and two asterisks denote significance at 1% level.

			T	Aaa	Aa	A	Baa
	$m=23$	$\hat{d}(\Delta X)$	1.1884	1.2771	1.155	1.2036	1.1253
		$\tau_{d=1}$	0.9178	1.3499	0.75544	0.99195	0.61038
	$m=51$	$\hat{d}(\Delta X)$	1.0009	1.0365	0.99467	1.0026	1.0038
		$\tau_{d=1}$	0.0079118	0.31429	-0.045894	0.022692	0.032615
$J=1$	$m=114$	$\hat{d}(\Delta X)$	1.0197	0.9328	0.97565	0.98891	0.98967
		$\tau_{d=1}$	0.28106	-0.96113	-0.34828	-0.15855	-0.14768
	$m=252$	$\hat{d}(\Delta X)$	0.9912	0.94918	0.98725	0.969	1.0003
		$\tau_{d=1}$	-0.19874	-1.1472	-0.28793	-0.69992	0.0075866
	$m=556$	$\hat{d}(\Delta X)$	0.98626	0.96081	0.99699	0.98349	1.0105
		$\tau_{d=1}$	-0.47201	-1.3464	-0.10337	-0.56734	0.36221
	$m=23$	$\hat{d}(\Delta X)$	1.1827	1.2709	1.1755	1.2025	1.1098
		$\tau_{d=1}$	0.57871	0.85821	0.5561	0.6416	0.3478
	$m=51$	$\hat{d}(\Delta X)$	0.98636	1.0224	0.99647	0.98728	0.97068
		$\tau_{d=1}$	-0.078932	0.12977	-0.020433	-0.073616	-0.16963
$J=2$	$m=115$	$\hat{d}(\Delta X)$	1.0152	0.92745	0.95922	0.96206	0.99032
		$\tau_{d=1}$	0.15	-0.71469	-0.40176	-0.37374	-0.095404
	$m=253$	$\hat{d}(\Delta X)$	1.0219	0.9704	0.99306	0.98199	0.99199
		$\tau_{d=1}$	0.34416	-0.4645	-0.10899	-0.28258	-0.12572
	$m=557$	$\hat{d}(\Delta X)$	0.9963	0.96762	0.98787	0.97465	0.99454
		$\tau_{d=1}$	-0.088829	-0.77797	-0.29139	-0.60916	-0.13124
	$m=22$	$\hat{d}(\Delta X)$	1.3521	1.2921	1.196	1.2095	1.1056
		$\tau_{d=1}$	0.81	0.67203	0.45085	0.48195	0.24288
	$m=52$	$\hat{d}(\Delta X)$	1.0366	1.0708	1.0233	1.0017	0.98548
		$\tau_{d=1}$	0.16712	0.3234	0.10638	0.0078825	-0.066381
$J=3$	$m=115$	$\hat{d}(\Delta X)$	1.0083	0.92826	0.94957	0.93996	0.97668
		$\tau_{d=1}$	0.064737	-0.56064	-0.39417	-0.46927	-0.18223
	$m=253$	$\hat{d}(\Delta X)$	1.0106	0.96208	0.98062	0.97014	0.98304
		$\tau_{d=1}$	0.13345	-0.47735	-0.24392	-0.37584	-0.21347
	$m=556$	$\hat{d}(\Delta X)$	0.98997	0.9647	0.98136	0.97179	0.98751
		$\tau_{d=1}$	-0.19448	-0.6843	-0.36142	-0.54701	-0.24205

Table 2.14: d estimates for the spreads with $l = 1$ and tapered data. For every series X, this table report the estimates $\hat{d}(\Delta X) \equiv \widehat{d-1}(\Delta X) + 1$ along with the test statistic (2.15). One asterisk denotes significance at 5% level and two asterisks denote significance at 1% level.

			sTAaa	sTAa	sTA	sTBaa	sAaaAa	sAaaA	sAaaBaa	sAaA	sAaBaa	sABaa
	$m = 23$	$\hat{d}(\Delta X)$	0.86505	0.8581	0.94091	0.79418	0.83822	0.75347	0.62092	0.99287	0.69988	0.56592
		$\tau_{d=1}$	-0.65751	-0.69136	-0.28792	-1.0028	-0.78825	-1.2012	-1.847	-0.034751	-1.4622	-2.1149*
	$m = 51$	$\hat{d}(\Delta X)$	1.0027	1.033	1.1588	1.1389	0.98358	1.103	1.0998	0.7973	0.87553	0.92263
		$\tau_{d=1}$	0.023083	0.28432	1.3682	1.197	-0.14149	0.88733	0.86003	-1.7464	-1.0724	-0.66664
$J = 1$	$m = 114$	$\hat{d}(\Delta X)$	1.1025	1.0576	1.1019	1.1246	0.97604	0.9788	1.0628	0.95275	0.96755	0.87341
		$\tau_{d=1}$	1.4664	0.82384	1.457	1.7813	-0.34274	-0.30318	0.89768	-0.67571	-0.46409	-1.8105
	$m = 252$	$\hat{d}(\Delta X)$	1.1291	1.1848	1.1835	1.1201	1.0308	0.99568	1.0537	0.91961	0.95622	0.93445
		$\tau_{d=1}$	2.9156**	4.1726**	4.143**	2.7123**	0.69581	-0.097571	1.2127	-1.8148	-0.98834	-1.4798
	$m = 556$	$\hat{d}(\Delta X)$	1.0139	1.0345	1.0567	1.0966	0.92509	0.93878	0.98391	0.82166	0.96636	0.96557
		$\tau_{d=1}$	0.47584	1.1835	1.9465	3.3193**	-2.5735*	-2.103*	-0.55262	-6.1267**	-1.1558	-1.1829
	$m = 23$	$\hat{d}(\Delta X)$	0.89271	0.86956	0.94986	0.77017	0.81107	0.80397	0.7709	1.2047	0.70425	0.67184
		$\tau_{d=1}$	-0.3399	-0.41325	-0.15884	-0.72813	-0.59854	-0.62103	-0.72581	0.64859	-0.93696	-1.0396
	$m = 51$	$\hat{d}(\Delta X)$	1.042	1.0632	1.1423	1.1319	0.94575	1.0717	1.1563	0.85854	0.90909	0.97809
		$\tau_{d=1}$	0.24298	0.36558	0.8236	0.76345	-0.31391	0.41514	0.90425	-0.81851	-0.52602	-0.12678
$J = 2$	$m = 115$	$\hat{d}(\Delta X)$	1.0948	1.0903	1.1329	1.1433	0.94662	1.0034	1.052	0.98472	0.95561	0.88498
		$\tau_{d=1}$	0.93353	0.88919	1.3089	1.4114	-0.52588	0.033477	0.51238	-0.15049	-0.43726	-1.1331
	$m = 253$	$\hat{d}(\Delta X)$	1.1244	1.1939	1.2087	1.1395	1.0068	1.0072	1.0359	0.94463	0.94727	0.92958
		$\tau_{d=1}$	1.9528	3.0428**	3.2749**	2.1893*	0.10712	0.11225	0.56383	-0.86897	-0.82741	-1.1051
	$m = 557$	$\hat{d}(\Delta X)$	1.0082	1.0404	1.0672	1.1002	0.90628	0.93846	0.99459	0.82947	0.9682	0.96534
		$\tau_{d=1}$	0.19659	0.97044	1.6136	2.4077*	-2.2519*	-1.4787	-0.13008	-4.0975**	-0.7641	-0.83286
	$m = 22$	$\hat{d}(\Delta X)$	0.9163	0.86516	0.95769	0.75386	0.81551	0.8946	0.80897	1.307	0.69999	0.65095
		$\tau_{d=1}$	-0.19256	-0.31024	-0.097341	-0.5663	-0.42447	-0.2425	-0.43951	0.7063	-0.69025	-0.80307
	$m = 52$	$\hat{d}(\Delta X)$	1.0687	1.0744	1.1692	1.1628	0.97588	1.1609	1.1934	0.84064	0.94133	1.004
		$\tau_{d=1}$	0.31377	0.34001	0.77312	0.74403	-0.11023	0.73544	0.88385	-0.72831	-0.26812	0.018248
$J = 3$	$m = 115$	$\hat{d}(\Delta X)$	1.0883	1.0526	1.1059	1.122	0.97594	1.0441	1.0511	0.95795	0.95823	0.89021
		$\tau_{d=1}$	0.68975	0.41134	0.82757	0.95377	-0.18802	0.34434	0.39943	-0.32866	-0.32645	-0.85802
	$m = 253$	$\hat{d}(\Delta X)$	1.0954	1.1608	1.1712	1.121	1.0127	1.0202	1.0389	0.92055	0.94951	0.93231
		$\tau_{d=1}$	1.201	2.0244*	2.155*	1.5237	0.15936	0.25447	0.48964	-1.0002	-0.63562	-0.85205
	$m = 555$	$\hat{d}(\Delta X)$	0.99814	1.0306	1.0598	1.0936	0.90888	0.94024	1.0034	0.8293	0.97406	0.96729
		$\tau_{d=1}$	-0.035999	0.59346	1.1587	1.8141	-1.7666	-1.1586	0.066679	-3.3095**	-0.50301	-0.63409

number decreases when moving from $l = 0$ to $l = 1$ and as J increases. Moreover the rejections seem to be concentrated to the highest values of m, in particular to $m = \left[n^{0.7}\right]$. However, even when the null is rejected, it is quite clear that the fractional difference parameter belong to the interval $\left(\frac{1}{2}, 1\right)$, which, again, entails long memory and non-stationarity.

Table 2.15 and Table 2.16 report the bias-reduced log-periodogram regression estimates for the long-memory parameter for yields and spreads respectively. Table 2.17 and Table 2.18 report the results for tapered data.

Table 2.15: Bias-reduced log-periodogram regression estimates for the long-memory parameter for the yields. For every series X, this table report the estimates $\widehat{d_r}(\Delta X) \equiv \widehat{d_r - 1}(\Delta X) + 1$ along with the test statistic (2.30). One asterisk denotes significance at 5% level and two asterisks denote significance at 1% level.

			T	Aaa	Aa	A	Baa
	$m = 23$	$\widehat{d_r}(\Delta X)$	0.43996	0.98478	0.95773	1.1716	1.013
		$\tau_{d=1}$	-2.7922**	-0.07587	-0.21073	0.85568	0.065032
	$m = 51$	$\widehat{d_r}(\Delta X)$	0.83033	0.95312	0.97245	1.0198	0.93057
		$\tau_{d=1}$	-1.2596	-0.34801	-0.20454	0.1472	-0.51548
$r = 1$	$m = 114$	$\widehat{d_r}(\Delta X)$	0.84397	0.95255	0.84036	0.883	0.9117
		$\tau_{d=1}$	-1.7319	-0.52663	-1.772	-1.2987	-0.98013
	$m = 252$	$\widehat{d_r}(\Delta X)$	0.94686	0.97995	0.89534	0.9393	0.96861
		$\tau_{d=1}$	-0.87701	-0.33091	-1.7272	-1.0017	-0.51807
	$m = 556$	$\widehat{d_r}(\Delta X)$	0.94624	0.9794	0.92868	0.95405	0.99853
		$\tau_{d=1}$	-1.3178	-0.50506	-1.7482	-1.1263	-0.03605
	$m = 23$	$\widehat{d_r}(\Delta X)$	-0.21458	0.46407	0.49389	0.74231	0.71132
		$\tau_{d=1}$	-4.8444**	-2.1376*	-2.0187*	-1.0278	-1.1514
	$m = 51$	$\widehat{d_r}(\Delta X)$	0.7266	1.0213	0.93382	1.0812	0.96691
		$\tau_{d=1}$	-1.6238	0.12625	-0.39308	0.48229	-0.19654
$r = 2$	$m = 114$	$\widehat{d_r}(\Delta X)$	0.82385	0.91065	0.86506	0.91174	0.92051
		$\tau_{d=1}$	-1.5642	-0.79341	-1.1983	-0.78372	-0.7059
	$m = 252$	$\widehat{d_r}(\Delta X)$	1.0167	1.0495	0.95689	1.0069	1.0118
		$\tau_{d=1}$	0.22057	0.65341	-0.56919	0.091542	0.15547
	$m = 556$	$\widehat{d_r}(\Delta X)$	0.93907	0.99071	0.91994	0.94768	0.98016
		$\tau_{d=1}$	-1.1949	-0.18226	-1.5701	-1.026	-0.38913
	$m = 23$	$\widehat{d_r}(\Delta X)$	-0.15001	0.58748	0.77727	0.93175	0.86655
		$\tau_{d=1}$	-3.9316**	-1.4103	-0.76147	-0.23334	-0.45623
	$m = 51$	$\widehat{d_r}(\Delta X)$	0.55066	1.0115	0.8993	1.138	1.0543
		$\tau_{d=1}$	-2.2875*	0.058579	-0.51266	0.70236	0.27668
$r = 3$	$m = 114$	$\widehat{d_r}(\Delta X)$	0.81547	0.99903	0.98418	1.0371	0.94946
		$\tau_{d=1}$	-1.4045	-0.00741	-0.12043	0.28217	-0.38469
	$m = 252$	$\widehat{d_r}(\Delta X)$	0.82901	0.9523	0.83025	0.88369	0.88551
		$\tau_{d=1}$	-1.935	-0.53975	-1.9209	-1.3162	-1.2956
	$m = 556$	$\widehat{d_r}(\Delta X)$	0.9197	0.97881	0.87348	0.91639	0.97081
		$\tau_{d=1}$	-1.3497	-0.35613	-2.1267*	-1.4055	-0.49065

Table 2.19 and Table 2.20 report the local Whittle estimates for yields and spreads respectively.

Table 2.21 and Table 2.22 report the local polynomial Whittle estimates for yields and spreads.

Basically, the local Whittle estimations and the related tests to verify whether the fractional difference parameter is different from 1, confirm the results we got from the log-periodogram techniques.

Table 2.23 reports the result of Nielsen (2005) LM test for yields when setting $d = 1$ in eq. (2.18) or $\mathbf{d} = \iota$ in eq. (2.21) for the multivariate case. Panel A reports univariate tests whereas Panel B reports multivariate tests. Table 2.24 reports the result of Nielsen (2005) LM test for spreads when setting $d = 1$ in eq. (2.18) or $\mathbf{d} = \iota$ in eq. (2.21) for the multivariate case. Panel A reports univariate tests whereas Panel B reports

Table 2.16: Bias-reduced log-periodogram regression estimates for the long-memory parameter for the spreads. For every series X, this table report the estimates $\widehat{d_r}(\Delta X) \equiv \widehat{d_r - 1}(\Delta X) + 1$ along with the test statistic (2.30). One asterisk denotes significance at 5% level and two asterisks denote significance at 1% level.

			sTAaa	sTAa	sTA	sTBaa	sAaaAa	sAaaA	sAaaBaa	sAaA	sAaBaa	sABaa
	$m=23$	$\widehat{d_r}(\Delta X)$	1.059	1.2045	1.1696	1.1346	0.58068	0.4825	0.95328	0.83931	0.99171	0.70896
		$\tau_{d=1}$	0.29406	1.0197	0.84552	0.67103	-2.0906	-2.5801**	-0.23293	-0.80115	-0.04135	-1.451
	$m=51$	$\widehat{d_r}(\Delta X)$	1.0181	1.1011	1.1258	1.1767	0.65606	0.7657	0.86367	0.9308	0.81977	1.0235
		$\tau_{d=1}$	0.1341	0.75079	0.93369	1.3118	-2.5535*	-1.7395	-1.0122	-0.51376	-1.338	0.17415
$r=1$	$m=114$	$\widehat{d_r}(\Delta X)$	1.0001	1.0159	1.0415	1.1504	0.82088	0.7753	0.91838	0.98174	0.94027	0.98043
		$\tau_{d=1}$	0.001257	0.17631	0.46028	1.6691	-1.9882*	-2.4941*	-0.90597	-0.20267	-0.66295	-0.21718
	$m=252$	$\widehat{d_r}(\Delta X)$	1.0154	1.0297	1.0648	1.1113	0.89699	0.92836	0.9838	1.029	1.0138	0.92827
		$\tau_{d=1}$	0.25436	0.49016	1.0699	1.8372	-1.7001	-1.1823	-0.26743	0.47928	0.22732	-1.1837
	$m=556$	$\widehat{d_r}(\Delta X)$	1.0584	1.0942	1.1083	1.1508	1.0004	0.99194	1.0359	0.99125	1.0257	1.0122
		$\tau_{d=1}$	1.4325	2.309*	2.656**	3.6977**	0.009627	-0.19766	0.8793	-0.21438	0.63031	0.29836
	$m=23$	$\widehat{d_r}(\Delta X)$	0.87372	0.88378	1.0416	1.1535	0.40838	0.73727	1.2329	0.90874	1.589	0.42791
		$\tau_{d=1}$	-0.50366	-0.46355	0.16578	0.61243	-2.3597*	-1.0479	0.92909	-0.364	2.3494*	-2.2818*
	$m=51$	$\widehat{d_r}(\Delta X)$	0.83194	1.2083	1.1303	0.99718	0.57858	0.61812	0.88608	1.0193	0.92044	0.8081
		$\tau_{d=1}$	-0.99815	1.2369	0.77409	-0.01676	-2.5029*	-2.2681*	-0.67664	0.11448	-0.47253	-1.1398
$r=2$	$m=114$	$\widehat{d_r}(\Delta X)$	1.0159	1.15	1.2048	1.185	0.69743	0.73443	0.98665	0.9786	0.84315	0.93522
		$\tau_{d=1}$	0.1416	1.3324	1.8187	1.6428	-2.6868**	-2.3583*	-0.11852	-0.19004	-1.3928	-0.57528
	$m=252$	$\widehat{d_r}(\Delta X)$	1.0709	1.0085	1.0451	1.1236	0.82695	0.83462	0.95248	0.99267	0.98219	0.89694
		$\tau_{d=1}$	0.93612	0.11246	0.59499	1.6318	-2.2847*	-2.1834*	-0.62734	-0.09679	-0.23519	-1.3606
	$m=556$	$\widehat{d_r}(\Delta X)$	0.98481	1.0475	1.0505	1.0738	0.93566	0.98306	1.0153	1.0256	1.0303	0.90585
		$\tau_{d=1}$	-0.29798	0.93148	0.98967	1.4482	-1.2618	-0.33218	0.3	0.50133	0.59342	-1.8464
	$m=23$	$\widehat{d_r}(\Delta X)$	1.5487	1.3743	1.5428	1.5189	0.23828	0.99229	0.84172	0.78267	1.143	0.005531
		$\tau_{d=1}$	1.8758	1.2798	1.8556	1.7739	-2.6041**	-0.02635	-0.54113	-0.74301	0.48902	-3.3999**
	$m=51$	$\widehat{d_r}(\Delta X)$	0.95502	1.1039	0.96958	0.98069	0.54429	0.28705	0.9367	0.852	1.1204	0.93087
		$\tau_{d=1}$	-0.22897	0.52883	-0.15485	-0.09833	-2.32*	-3.6295**	-0.32224	-0.75342	0.61293	-0.35194
$r=3$	$m=114$	$\widehat{d_r}(\Delta X)$	1.1002	1.0591	1.0812	1.1034	0.72119	0.88721	0.87987	0.95507	0.70713	0.98158
		$\tau_{d=1}$	0.76294	0.45	0.6181	0.78734	-2.1221*	-0.8585	-0.91432	-0.34194	-2.2291*	-0.14024
	$m=252$	$\widehat{d_r}(\Delta X)$	1.0443	0.99721	1.0992	1.1288	0.84288	0.78257	0.92227	0.93698	0.91356	0.90651
		$\tau_{d=1}$	0.50145	-0.03153	1.1221	1.457	-1.7781	-2.4606*	-0.87962	-0.71311	-0.97813	-1.0579
	$m=556$	$\widehat{d_r}(\Delta X)$	1.0576	1.0672	1.1092	1.1191	0.91318	0.94757	1.0003	1.0163	1.075	0.93473
		$\tau_{d=1}$	0.9682	1.1297	1.8361	2.0023*	-1.4593	-0.88135	0.005632	0.27326	1.2605	-1.0972

Table 2.17: Bias-reduced log-periodogram regression estimates for the long-memory parameter for the tapered yields. For every series X, this table report the estimates $\widehat{d_r}(\Delta X) \equiv \widehat{d_r - 1}(\Delta X) + 1$ along with the test statistic (2.30). One asterisk denotes significance at 5% level and two asterisks denote significance at 1% level.

			T	Aaa	Aa	A	Baa
	$m = 23$	$\widehat{d_r}(\Delta X)$	1.0024	1.1394	1.0789	1.1753	1.115
		$\tau_{d=1}$	0.011889	0.69514	0.39359	0.87415	0.5734
	$m = 51$	$\widehat{d_r}(\Delta X)$	1.142	1.2702	1.075	1.1663	1.1118
		$\tau_{d=1}$	1.0541	2.0063*	0.55669	1.2346	0.83001
$r = 1$	$m = 114$	$\widehat{d_r}(\Delta X)$	0.95423	0.97232	0.91717	0.96366	0.95208
		$\tau_{d=1}$	-0.50803	-0.3072	-0.91939	-0.40339	-0.53191
	$m = 252$	$\widehat{d_r}(\Delta X)$	1.0047	0.90765	0.95592	0.9703	0.9811
		$\tau_{d=1}$	0.077832	-1.5241	-0.72747	-0.4901	-0.31192
	$m = 556$	$\widehat{d_r}(\Delta X)$	0.95056	0.91582	0.93804	0.93192	0.95831
		$\tau_{d=1}$	-1.212	-2.0635*	-1.5188	-1.6689	-1.0219
	$m = 23$	$\widehat{d_r}(\Delta X)$	0.57062	0.56423	0.75148	0.80936	0.97394
		$\tau_{d=1}$	-1.7126	-1.7381	-0.99125	-0.76037	-0.10396
	$m = 51$	$\widehat{d_r}(\Delta X)$	1.1155	1.1977	1.0708	1.1782	1.1187
		$\tau_{d=1}$	0.68593	1.1745	0.42056	1.0584	0.70505
$r = 2$	$m = 114$	$\widehat{d_r}(\Delta X)$	1.1015	1.1257	1.0095	1.0886	1.0543
		$\tau_{d=1}$	0.90096	1.1166	0.084281	0.78687	0.48217
	$m = 252$	$\widehat{d_r}(\Delta X)$	1.067	0.97652	0.99023	1.0336	1.0116
		$\tau_{d=1}$	0.88498	-0.31003	-0.12903	0.44317	0.15374
	$m = 556$	$\widehat{d_r}(\Delta X)$	0.97466	0.91166	0.9651	0.95789	0.97538
		$\tau_{d=1}$	-0.49701	-1.7325	-0.68451	-0.82585	-0.48275
	$m = 23$	$\widehat{d_r}(\Delta X)$	0.69049	0.69371	0.76844	0.96686	0.98444
		$\tau_{d=1}$	-1.0582	-1.0471	-0.79165	-0.11331	-0.0532
	$m = 51$	$\widehat{d_r}(\Delta X)$	0.97134	1.0078	1.0225	1.1186	1.0561
		$\tau_{d=1}$	-0.14589	0.0395	0.11453	0.60389	0.28557
$r = 3$	$m = 114$	$\widehat{d_r}(\Delta X)$	1.0581	1.2127	1.0331	1.1137	1.0558
		$\tau_{d=1}$	0.44239	1.6192	0.25205	0.86526	0.42461
	$m = 252$	$\widehat{d_r}(\Delta X)$	0.93191	0.94418	0.88974	0.94565	0.95005
		$\tau_{d=1}$	-0.77059	-0.63164	-1.2478	-0.61508	-0.56529
	$m = 556$	$\widehat{d_r}(\Delta X)$	1.0347	0.93926	0.97937	0.98571	1.0169
		$\tau_{d=1}$	0.58312	-1.0209	-0.34674	-0.24024	0.2839

Table 2.18: Bias-reduced log-periodogram regression estimates for the long-memory parameter for the tapered spreads. For every series X, this table report the estimates $\widehat{d_r}(\Delta X) \equiv \widehat{d_r - 1}(\Delta X) + 1$ along with the test statistic (2.30). One asterisk denotes significance at 5% level and two asterisks denote significance at 1% level.

			sTAaa	sTAa	sTA	sTBaa	sAaaAa	sAaaA	sAaaBaa	sAaA	sAaBaa	sABaa
	$m=23$	$\widehat{d_r}(\Delta X)$	1.0499	1.0849	1.1271	1.0558	0.71379	0.70085	0.40559	1.1063	0.92604	0.6375
		$\tau_{d=1}$	0.24902	0.42349	0.63361	0.27796	-1.427	-1.4915	-2.9636**	0.53021	-0.36874	-1.8073
	$m=51$	$\widehat{d_r}(\Delta X)$	1.0891	1.116	1.1309	1.1679	0.80449	0.98151	0.69377	1.0852	0.7438	0.66089
		$\tau_{d=1}$	0.66164	0.86119	0.97167	1.2465	-1.4515	-0.13726	-2.2735*	0.63243	-1.9021	-2.5176*
$r=1$	$m=114$	$\widehat{d_r}(\Delta X)$	1.0619	1.0069	1.0845	1.1648	0.89106	1.0071	1.1244	0.78772	1.0251	0.93859
		$\tau_{d=1}$	0.68701	0.076283	0.93757	1.8293	-1.2092	0.079201	1.3813	-2.3563	0.27913	-0.68165
	$m=252$	$\widehat{d_r}(\Delta X)$	1.1492	1.0831	1.0954	1.1709	0.88639	0.97115	1.108	0.97535	0.997	0.89844
		$\tau_{d=1}$	2.4614*	1.3715	1.5743	2.8211**	-1.8749	-0.47618	1.7822	-0.40673	-0.04949	-1.676
	$m=556$	$\widehat{d_r}(\Delta X)$	1.1498	1.2127	1.172	1.1267	1.0019	0.99946	1.0228	0.92428	1.0191	0.97122
		$\tau_{d=1}$	3.6721**	5.2135**	4.2164**	3.1066**	0.046261	-0.01335	0.55856	-1.8562	0.46713	-0.70548
	$m=23$	$\widehat{d_r}(\Delta X)$	0.99939	1.1089	1.1159	1.1086	0.50581	0.68613	0.40375	1.3656	1.0456	0.65495
		$\tau_{d=1}$	-0.0024164	0.43452	0.46215	0.43323	-1.9711	-1.2519	-2.3782*	1.458	0.18194	-1.3763
	$m=56$	$\widehat{d_r}(\Delta X)$	0.98155	1.061	1.1676	0.88518	0.54465	0.68823	0.40612	1.1152	0.81036	0.47599
		$\tau_{d=1}$	-0.1096	0.36245	0.99516	-0.68195	-2.7045	-1.8517	-3.5273**	0.68417	-1.1263	-3.1123**
$r=2$	$m=114$	$\widehat{d_r}(\Delta X)$	1.0597	1.1767	1.2581	1.2714	0.89595	1.1089	0.90328	0.83456	0.85663	0.78908
		$\tau_{d=1}$	0.53054	1.569	2.2919*	2.4096*	-0.92391	0.96682	-0.8589	-1.4691	-1.2731	-1.873
	$m=252$	$\widehat{d_r}(\Delta X)$	1.1245	1.0689	1.1108	1.1502	0.9282	0.99472	1.071	0.91405	0.98613	0.88569
		$\tau_{d=1}$	1.644	0.90998	1.4635	1.9826*	-0.94791	-0.06969	0.93697	-1.1348	-0.18313	-1.5092
	$m=556$	$\widehat{d_r}(\Delta X)$	1.1282	1.1542	1.1574	1.0999	0.99174	0.98391	1.056	0.98117	0.92893	0.85883
		$\tau_{d=1}$	2.5138*	3.0248**	3.0877**	1.9594	-0.16193	-0.31546	1.0974	-0.36923	-1.3938	-2.7684**
	$m=23$	$\widehat{d_r}(\Delta X)$	1.8031	1.6837	1.6949	1.4422	0.31379	0.57444	0.31805	1.8157	0.93309	0.59212
		$\tau_{d=1}$	2.7455**	2.3375*	2.3756*	1.5118	-2.346*	-1.4549	-2.3314*	2.7888**	-0.22875	-1.3945
	$m=51$	$\widehat{d_r}(\Delta X)$	0.83486	0.81456	0.8459	0.59809	0.54911	0.40078	0.25638	1.2843	0.93727	0.52875
		$\tau_{d=1}$	-0.84071	-0.94406	-0.78451	-2.0461*	-2.2954*	-3.0506**	-3.7857**	1.4473	-0.31934	-2.3991*
$r=3$	$m=114$	$\widehat{d_r}(\Delta X)$	1.1328	1.1289	1.2537	1.1108	0.79883	1.0708	0.72726	1.1609	0.69125	0.67726
		$\tau_{d=1}$	1.011	0.9808	1.9314	0.84322	-1.5312	0.53859	-2.0759*	1.2248	-2.35*	-2.4565*
	$m=252$	$\widehat{d_r}(\Delta X)$	1.0616	1.0343	1.1587	1.1524	0.93044	1.0115	1.0045	0.8616	0.99915	0.85044
		$\tau_{d=1}$	0.69742	0.38783	1.7955	1.7251	-0.78716	0.13007	0.05048	-1.5661	-0.00959	-1.6925
	$m=556$	$\widehat{d_r}(\Delta X)$	1.1816	1.1603	1.1932	1.1933	0.95299	1.015	1.1181	0.98059	0.97508	0.88751
		$\tau_{d=1}$	3.0532**	2.6945**	3.2476**	3.2499**	-0.79013	0.25156	1.9846*	-0.32619	-0.41887	-1.8909

Table 2.19: Local Whittle estimates for the yields. For every series X, this table report the estimates $\tilde{d}(X)$ and $\tilde{d}(\Delta X) \equiv \widetilde{d-1}(\Delta X) + 1$ along with the test statistics (2.32) and (2.33) respectively. One asterisk denotes significance at 5% level and two asterisks denote significance at 1% level.

		T	Aaa	Aa	A	Baa
$m = 23$	$\tilde{d}(X)$	1.1577	1.167	1.1097	1.1454	1.1219
	$t_{d=1}$	1.5123	1.6017	1.0521	1.3948	1.1689
$m = 51$	$\tilde{d}(X)$	0.91182	0.94809	0.92025	0.90521	0.88131
	$t_{d=1}$	-1.2594	-0.74138	-1.1391	-1.3539	-1.6952
$m = 114$	$\tilde{d}(X)$	0.99674	0.99071	0.96684	0.96289	0.96243
	$t_{d=1}$	-0.0697	-0.19833	-0.708	-0.7925	-0.80225
$m = 252$	$\tilde{d}(X)$	0.9941	0.99666	0.98005	0.97504	0.99569
	$t_{d=1}$	-0.18737	-0.10596	-0.63354	-0.79245	-0.13684
$m = 556$	$\tilde{d}(X)$	0.98095	0.97878	0.97273	0.96478	0.97887
	$t_{d=1}$	-0.89842	-1.0006	-1.2861	-1.6611	-0.99643
$m = 23$	$\tilde{d}(\Delta X)$	0.96063	1.0084	1.0473	1.026	0.99513
	$\tau_{d=1}$	-0.37759	0.08021	0.45354	0.24948	-0.04672
$m = 51$	$\tilde{d}(\Delta X)$	0.93782	1.0165	1.0079	0.9944	0.97732
	$\tau_{d=1}$	-0.88811	0.2356	0.11253	-0.08003	-0.32394
$m = 114$	$\tilde{d}(\Delta X)$	1.0372	1.0134	1.0083	1.0101	1.0269
	$\tau_{d=1}$	0.79448	0.28525	0.17759	0.21594	0.57464
$m = 252$	$\tilde{d}(\Delta X)$	0.98669	0.98386	0.972	0.97201	1.0024
	$\tau_{d=1}$	-0.42253	-0.51249	-0.889	-0.88861	0.077778
$m = 556$	$\tilde{d}(\Delta X)$	0.98806	0.99025	0.97874	0.97453	0.99326
	$\tau_{d=1}$	-0.56288	-0.45975	-1.0024	-1.2014	-0.3177

multivariate tests.

Table 2.20: Local Whittle estimates for the spreads. For every series X, this table report the estimates $\tilde{d}(X)$ and $\tilde{d}(\Delta X) \equiv \widetilde{d-1}(\Delta X) + 1$ along with the test statistics (2.32) and (2.33) respectively. One asterisk denotes significance at 5% level and two asterisks denote significance at 1% level.

		sTAaa	sTAa	sTA	sTBaa	sAaaAa	sAaaA	sAaaBaa	sAaA	sAaBaa	sABaa
$m = 23$	$\tilde{d}(X)$	0.93646	0.98644	1.0337	0.96975	0.79523	0.84721	0.87801	0.9524	0.78504	0.89519
	$t_{d=1}$	-0.60941	-0.1301	0.3232	-0.29017	-1.9641	-1.4655	-1.1701	-0.45658	-2.0619*	-1.0053
$m = 51$	$\tilde{d}(X)$	1.0659	1.0772	1.1409	1.104	0.88337	0.96783	0.98093	0.9698	0.79327	0.88544
	$t_{d=1}$	0.94154	1.1032	2.0125*	1.4849	-1.6658	-0.45951	-0.27232	-0.43128	-2.9527**	-1.6362
$m = 114$	$\tilde{d}(X)$	1.0828	1.0495	1.0834	1.0701	0.91027	0.99811	0.94407	0.9712	0.98375	0.90814
	$t_{d=1}$	1.7686	1.0575	1.7815	1.4974	-1.9161	-0.04037	-1.1943	-0.61498	-0.3471	-1.9617*
$m = 252$	$\tilde{d}(X)$	1.0842	1.0728	1.0877	1.0965	1.0034	1.0272	1.0454	0.9167	1.0119	1.0119
	$t_{d=1}$	2.6731**	2.312*	2.783**	3.0635**	0.10737	0.86367	1.4401	-2.6448**	0.37758	0.37762
$m = 556$	$\tilde{d}(X)$	1.0083	0.99229	1.0046	1.0537	0.94755	0.9577	1.0333	0.85168	0.9668	0.96479
	$t_{d=1}$	0.38937	-0.36375	0.219	2.5311*	-2.4737*	-1.995	1.5712	-6.9946**	-1.5658	-1.6603
$m = 23$	$\tilde{d}(\Delta X)$	0.95017	0.97652	1.0213	0.96277	0.62929	0.7002	0.82383	0.99918	0.78182	0.89659
	$\tau_{d=1}$	-0.47793	-0.22524	0.20405	-0.35709	-3.5558**	-2.8756**	-1.6898	-0.00782	-2.0927*	-0.99187
$m = 51$	$\tilde{d}(\Delta X)$	1.0623	1.0774	1.1407	1.1004	0.80978	0.88246	0.91923	0.95135	0.7861	0.89481
	$\tau_{d=1}$	0.88943	1.1058	2.01*	1.4336	-2.7169**	-1.6788	-1.1536	-0.69483	-3.055**	-1.5023
$m = 114$	$\tilde{d}(\Delta X)$	1.0776	1.05	1.0822	1.0677	0.86685	0.94205	0.90191	0.96217	0.97921	0.91494
	$\tau_{d=1}$	1.6573	1.0669	1.7545	1.4466	-2.8433**	-1.2375	-2.0946*	-0.80773	-0.44394	-1.8163
$m = 252$	$\tilde{d}(\Delta X)$	1.0873	1.0787	1.0903	1.1	0.99018	1.0055	1.0362	0.913	1.0149	1.0148
	$\tau_{d=1}$	2.7718**	2.4972*	2.8678**	3.1736**	-0.31183	0.17487	1.1506	-2.7621**	0.47373	0.46939
$m = 556$	$\tilde{d}(\Delta X)$	1.0273	1.008	1.0162	1.0703	0.96682	0.96454	1.054	0.85775	0.98494	0.97835
	$\tau_{d=1}$	1.2862	0.37586	0.76442	3.3159**	-1.5646	-1.6724	2.545*	-6.7085**	-0.71027	-1.0208

Table 2.21: Local polynomial Whittle estimates for the yields. For every series X, this table report the estimates $\tilde{d}_r(\Delta X) \equiv \widetilde{d_r - 1}(\Delta X) + 1$ along with the test statistic (2.36). One asterisk denotes significance at 5% level and two asterisks denote significance at 1% level.

			T	Aaa	Aa	A	Baa
	$m = 23$	$\tilde{d}_r(\Delta X)$	0.59787	1.0289	1.0835	1.1342	1.0622
		$\tau_{d=1}$	-2.5714*	0.18507	0.53378	0.85836	0.3978
	$m = 51$	$\tilde{d}_r(\Delta X)$	0.95966	0.98565	1.0727	1.049	0.99381
		$\tau_{d=1}$	-0.38413	-0.13667	0.69195	0.46704	-0.05896
$r = 1$	$m = 114$	$\tilde{d}_r(\Delta X)$	0.98557	1.0478	1.012	1.0122	0.98501
		$\tau_{d=1}$	-0.20547	0.68098	0.17144	0.17397	-0.21333
	$m = 252$	$\tilde{d}_r(\Delta X)$	1.0263	1.0096	1.0043	1.0075	1.0237
		$\tau_{d=1}$	0.55581	0.20244	0.0911	0.15857	0.50094
	$m = 556$	$\tilde{d}_r(\Delta X)$	0.97396	0.96149	0.95921	0.95598	0.98334
		$\tau_{d=1}$	-0.81882	-1.2109	-1.2825	-1.3839	-0.52367
	$m = 23$	$\tilde{d}_r(\Delta X)$	0.59787	1.029	1.0835	1.1342	1.0622
		$\tau_{d=1}$	-2.0571*	0.1481	0.42705	0.68671	0.31827
	$m = 51$	$\tilde{d}_r(\Delta X)$	0.95967	0.98566	1.0727	1.0491	0.99382
		$\tau_{d=1}$	-0.30722	-0.10925	0.55359	0.37397	-0.04709
$r = 2$	$m = 114$	$\tilde{d}_r(\Delta X)$	0.92865	1.0034	1.0279	0.99909	0.97961
		$\tau_{d=1}$	-0.81264	0.038178	0.31805	-0.01041	-0.23218
	$m = 252$	$\tilde{d}_r(\Delta X)$	1.1153	1.0893	1.0781	1.0825	1.0799
		$\tau_{d=1}$	1.9526	1.5118	1.3227	1.3978	1.3524
	$m = 556$	$\tilde{d}_r(\Delta X)$	1.003	1.0026	0.98883	0.99634	1.0225
		$\tau_{d=1}$	0.075935	0.066491	-0.28093	-0.09195	0.56505
	$m = 23$	$\tilde{d}_r(\Delta X)$	0.59787	1.0289	1.0835	1.1342	1.0622
		$\tau_{d=1}$	-1.7632	0.12694	0.36609	0.58861	0.27278
	$m = 51$	$\tilde{d}_r(\Delta X)$	0.95967	0.98566	1.0727	1.049	0.99382
		$\tau_{d=1}$	-0.2633	-0.09363	0.47451	0.31982	-0.04035
$r = 3$	$m = 114$	$\tilde{d}_r(\Delta X)$	0.92885	1.0035	1.0279	0.99903	0.97963
		$\tau_{d=1}$	-0.69458	0.033868	0.27191	-0.00942	-0.19887
	$m = 252$	$\tilde{d}_r(\Delta X)$	0.94258	0.9848	0.96609	0.96355	0.95871
		$\tau_{d=1}$	-0.83336	-0.2206	-0.49212	-0.529	-0.59923
	$m = 556$	$\tilde{d}_r(\Delta X)$	1.0166	1.014	1.0013	1.0037	1.0352
		$\tau_{d=1}$	0.35723	0.30117	0.027176	0.079197	0.75857

Table 2.22: Local polynomial Whittle estimates for the spreads. For every series X, this table report the estimates $\tilde{d}_r(\Delta X) \equiv \widetilde{d_r - 1}(\Delta X) + 1$ along with the test statistic (2.36). One asterisk denotes significance at 5% level and two asterisks denote significance at 1% level.

			sTAaa	sTAa	sTA	sTBaa	sAaaAa	sAaaA	sAaaBaa	sAaA	sAaBaa	sABaa
$r=1$	$m=23$	$\tilde{d}_r(\Delta X)$	0.98958	0.97615	1.0086	0.99845	0.54751	0.47622	0.82339	0.82933	0.91142	0.73433
		$\tau_{d=1}$	-0.06661	-0.1525	0.055122	-0.00988	-2.8934**	-3.3493**	-1.1293	-1.0913	-0.56639	-1.6988
	$m=51$	$\tilde{d}_r(\Delta X)$	1.0171	1.0154	1.0741	1.0876	0.56199	0.70758	0.76569	0.87679	0.73	0.87479
		$\tau_{d=1}$	0.16271	0.14695	0.70548	0.83395	-4.1707**	-2.7844**	-2.2311*	-1.1732	-2.5709*	-1.1923
	$m=114$	$\tilde{d}_r(\Delta X)$	1.033	1.0679	1.0977	1.1247	0.89921	0.86401	0.93223	0.94564	0.81838	0.93568
		$\tau_{d=1}$	0.46956	0.96614	1.3904	1.7752	-1.4349	-1.936	-0.96484	-0.77381	-2.5855	-0.91569
	$m=252$	$\tilde{d}_r(\Delta X)$	1.0724	1.0714	1.0981	1.1015	0.87836	0.96398	0.92438	0.99362	1.0167	0.94442
		$\tau_{d=1}$	1.5321	1.5118	2.0761*	2.148*	-2.5746**	-0.76236	-1.6006	-0.13499	0.35271	-1.1764
	$m=556$	$\tilde{d}_r(\Delta X)$	1.0993	1.114	1.1232	1.1413	0.99786	0.99568	1.0267	0.93327	1.0414	1.0535
		$\tau_{d=1}$	3.1224**	3.5855**	3.8729**	4.4421**	-0.06738	-0.1358	0.84045	-2.0978*	1.3027	1.6806
$r=2$	$m=23$	$\tilde{d}_r(\Delta X)$	0.98958	0.97615	1.0086	0.99845	0.54751	0.47622	0.82339	0.82933	0.91143	0.73433
		$\tau_{d=1}$	-0.05329	-0.12201	0.044046	-0.00791	-2.3147*	-2.6794**	-0.90344	-0.87307	-0.45311	-1.359
	$m=51$	$\tilde{d}_r(\Delta X)$	1.0171	1.0154	1.0741	0.93765	0.56202	0.7076	0.76571	1.0611	0.73001	0.8748
		$\tau_{d=1}$	0.13027	0.11762	0.56448	-0.47496	-3.3363**	-2.2273*	-1.7847	0.46549	-2.0567*	-0.95375
	$m=114$	$\tilde{d}_r(\Delta X)$	1.0442	1.1329	1.2093	1.1724	0.68801	0.77586	0.89559	0.92398	0.76173	0.89184
		$\tau_{d=1}$	0.50347	1.5142	2.3839*	1.9636*	-3.5532**	-2.5527*	-1.1891	-0.86574	-2.7136**	-1.2318
	$m=252$	$\tilde{d}_r(\Delta X)$	1.0812	1.0714	1.0949	1.0872	0.8368	0.88461	0.88516	0.98193	0.94894	0.89768
		$\tau_{d=1}$	1.3758	1.2087	1.607	1.477	-2.7634**	-1.9539	-1.9445	-0.30606	-0.86465	-1.7326
	$m=556$	$\tilde{d}_r(\Delta X)$	1.0477	1.0674	1.0754	1.0699	0.94827	0.9943	0.96456	0.96952	1.0056	0.95539
		$\tau_{d=1}$	1.1989	1.6961	1.8976	1.7576	-1.301	-0.14334	-0.89135	-0.76673	0.13968	-1.1219
$r=3$	$m=23$	$\tilde{d}_r(\Delta X)$	0.98958	0.97615	1.0086	0.99845	0.54751	0.47622	0.82339	0.82933	0.91143	0.7338
		$\tau_{d=1}$	-0.04568	-0.10458	0.037746	-0.00678	-1.984*	-2.2967*	-0.77438	-0.74835	-0.38838	-1.1672
	$m=51$	$\tilde{d}_r(\Delta X)$	1.0171	1.0154	1.0741	0.93768	0.56202	0.70761	0.76571	1.0611	0.73001	0.8748
		$\tau_{d=1}$	0.11166	0.10083	0.48383	-0.40693	-2.8597**	-1.9091	-1.5297	0.39899	-1.7629	-0.81749
	$m=114$	$\tilde{d}_r(\Delta X)$	1.0442	1.1331	1.2093	1.1727	0.68853	0.77597	0.8957	0.92382	0.76191	0.89197
		$\tau_{d=1}$	0.43131	1.2994	2.0429*	1.6855	-3.0405**	-2.1869*	-1.0182	-0.74369	-2.3242*	-1.0545
	$m=252$	$\tilde{d}_r(\Delta X)$	1.0808	1.0636	1.1323	1.1033	0.88548	0.86043	0.86742	0.90521	0.8094	0.88199
		$\tau_{d=1}$	1.1723	0.92236	1.9203	1.4987	-1.6622	-2.0256*	-1.9243	-1.3758	-2.7664**	-1.7128
	$m=556$	$\tilde{d}_r(\Delta X)$	1.0973	1.0815	1.1109	1.0933	0.90859	0.98106	0.96213	0.97777	1.034	0.97422
		$\tau_{d=1}$	2.0975*	1.756	2.39*	2.0112*	-1.9708*	-0.40835	-0.81647	-0.47927	0.73193	-0.55574

Table 2.23: Nielsen (2005) LM test for yields. For the univariate case, we set $d = 1$ in eq. (2.18). For the multivariate case, we set $\mathbf{d} = \iota$ in eq. (2.21). Panel A reports univariate tests whereas Panel B reports multivariate tests.

Panel A

		T	Aaa	Aa	A	Baa
$p = 0$	LM	0.24403	0.98816	0.74011	0.25188	1.4597
	pval	0.62131	0.32019	0.38963	0.61575	0.22698
$p = 1$	LM	0.2163	0.16096	0.68366	0.7104	0.076076
	pval	0.64187	0.68828	0.40833	0.39931	0.78269
$p = 2$	LM	0.16922	0.37611	0.5547	0.66573	0.12288
	pval	0.68081	0.53969	0.4564	0.41454	0.72593
$p = 3$	LM	0.10598	0.23807	0.30477	0.38305	0.04553
	pval	0.74476	0.62561	0.58091	0.53598	0.83103
$p = 4$	LM	0.003461	0.020401	0.057072	0.054574	0.013225
	pval	0.95308	0.88642	0.81119	0.81529	0.90845

Panel B

$p = 0$	LM	167.6335	LMK	169.8412
	pval	0	pval	0
$p = 1$	LM	0.050609	LMK	3.7343
	pval	0.822	pval	0.5883
$p = 2$	LM	0.005161	LMK	3.9713
	pval	0.9427	pval	0.5536
$p = 3$	LM	0.003228	LMK	2.2841
	pval	0.9547	pval	0.8086
$p = 4$	LM	0.099125	LMK	1.9812
	pval	0.7529	pval	0.8517

The tables show that every time we do not allow for short run dynamics (i.e. $p = 0$), and we consider the univariate procedure, first differencing is sufficient to achieve stationarity for yields, but not for spreads. The null of stationarity after first differencing is rejected for the series sBaa when $p = 1$ and $p = 2$ and for the series sAaA when $p = 1$. Also in the multivariate case for yields and for spreads over treasury when $p = 0$ the null hypothesis of stationarity after first differencing is rejected. This is not the case when we allow for for short run dynamics ($p > 0$).

Table 2.25 reports the results of some unit root and stationarity tests for the residuals from the cointegrating regression for all possible bivariate systems of yields.

Table 2.24: Nielsen (2005) LM test for spreads. For the univariate case, we set $d = 1$ in eq. (2.18). For the multivariate case, we set $\mathbf{d} = \iota$ in eq. (2.21). Panel A reports univariate tests whereas Panel B reports multivariate tests for spreads over Treasury only.

Panel A

	sAaa	sAa	sA	sBaa	sAaaAa	sAaaA	sAaaBaa	sAaA	sAaBaa	sABaa
$p = 0$	43.789	44.47	53.892	10.476	55.159	68.764	8.1254	254.31	39.43	79.58
	3.66E-11	2.58E-11	2.12E-13	0.00121	1.11E-13	1.11E-16	0.004365	0	3.40E-10	0.00E+00
$p = 1$	3.0048	1.177	1.6468	8.4342	0.14955	0.015179	7.6248	26.407	0.092168	0.00071
	0.083017	0.27797	0.19939	0.003682	0.69897	0.90195	0.005757	2.77E-07	0.76144	0.97875
$p = 2$	3.2169	2.8481	4.0746	6.6772	0.36829	0.074718	0.087621	5.5076	0.096976	0.23447
	0.072881	0.09148	0.043533	0.009765	0.54394	0.78459	0.76722	0.018934	0.75549	0.62823
$p = 3$	1.047	1.5122	2.4191	1.6061	0.05593	0.039755	0.053543	1.8383	0.092241	0.15987
	0.30619	0.2188	0.11986	0.20505	0.81305	0.84196	0.81701	0.17516	0.76135	0.68927
$p = 4$	0.51029	0.74678	0.92112	0.375	0.13419	0.020081	0.2416	0.79341	0.000393	0.095139
	0.47501	0.3875	0.33718	0.54029	0.71412	0.88731	0.62305	0.37307	0.98419	0.75774

Panel B

$p = 0$	LM	232.1004	LMK	282.6042
	pval	0	pval	0
$p = 1$	LM	0.004572	LMK	10.178
	pval	0.94609	pval	0.037538
$p = 2$	LM	0.067499	LMK	4.669
	pval	0.79501	pval	0.32297
$p = 3$	LM	0.038253	LMK	1.4772
	pval	0.84494	pval	0.83067
$p = 4$	LM	0.03184	LMK	1.0435
	pval	0.85838	pval	0.90313

Table 2.25 shows that there is no evidence of cointegration between Treasury and corporate yields. The only exception is the system Aaa–Aa for which the DF and the KPSS tests lead to different conclusions. This surprising result could be explained by the fact that the usual concept of cointegration may be too restrictive. Treasury and corporate yields may in fact be fractionally cointegrated.

Table 2.26 through Table 2.29 report the results of the three step procedure of Dittmann (2004) when residuals from the narrow band FDLS estimation are used. Table 2.26 and Table 2.27 report the results for the residuals, whereas Table 2.28 and Table 2.29 report the results for the differenced residuals.

Table 2.30 through Table 2.33 report the results of the three step procedure of Dittmann (2004) when tapered residuals from the narrow band FDLS estimation are used. Table 2.30 and Table 2.31 report the results for the residuals, whereas Table 2.32 and Table 2.33 report the results for the differenced residuals.

From the analysis of Table 2.26 through Table 2.33, it is clear that the fractional differencing parameter is always statistically different (and less) than 1 only for the systems Aaa-Aa and Aa-A. From Table 2.26 (non-tapered data, $l = 0$ and $J = 1$), it seems that the some fractional cointegration is present also in the system Aaa-A. When tapered data is used, some evidence of fractional cointegration in the system T-Aaa, can be found as well.

These results are different from the ones in Della Ratta and Urga (2005) in which the authors find fractional cointegration in all the bivariate systems except T-Baa and Aaa-Baa. This discrepancy is due to the fact that they use $m = [n^{0.9}]$.

On the other hand, looking at the Dittmann (2004) estimation procedure, the residuals from the cointegrating relation seem to be non-stationary even after the fractional differention, according to the KPSS test, only for the system T-Aaa. For each other bivariate system the null hypothesis of stationarity cannot be reject by using the KPSS test.

2.4.2 Dataset 2

Our second dataset consists of Lehman Brothers Eurodollar Aaa, Aa, A and Baa Indices and U.S. Global Treasury Index. The indices include primarily corporates bonds (even though they can include government-related and securitized bonds). The data covers the period from June 1996 to July 2006, for $n = 2613$ observations. Spreads are calculated as the difference between corporate yields and Treasury yields, as well as between different corporate yields. Thus we have 15 series: Lehman Brothers U.S. Global Treasury Index (denoted by T), corporate yields (Aaa, Aa, A, and Baa), spreads over Tresasury (sTAaa, sTAa, sTA, and sTBaa), spreads between corporate yields (sAaaAa, sAaaA, sAaaBaa, sAaA, sAaBaa, and sABaa). These series are plotted in Figure 2.10–Figure 2.12.

Table 2.34 and Table 2.35 report summary statistics and normality tests for the series involved and for their first differences respectively. All the series are highly non-normal.

Table 2.36 and Table 2.37 report unit root and stationarity tests for yields and spreads and their first differences.

The series are clearly non-stationary. As far as first differences are concerned, only for the series ΔT the KPSS tests reject the null of stationarity (at a 5% level of significance). Thus we can conclude that first differences of the series are stationary.

Table 2.38 and Table 2.39 report the d estimates for the yields and the spreads when $l = 0$.

Table 2.40 and Table 2.41 report the d estimates for the yields and the spreads when $l = 1$.

Table 2.25: Cointegration analysis: unit root and stationarity tests for bivariate systems. Cointegration analysis is performed for all possible bivariate systems $X - Y$. For each pair of variables X and Y, the Dickey-Fuller without the constant (DF), augmented Dickey-Fuller without the constant (ADF), Phillips-Perron with the constant (PP), a two KPSS tests without trend are carried out on the estimated OLS residuals of the regression of Y on X and a constant. In the first KPSS test the Bartlett kernel with bandwidth parameter $\left[4 \left(\frac{n}{100} \right)^{1/4} \right]$ is chosen for the estimation of the long run variance. In the second test the automatic bandwidth selection procedure of Hobijn et al. (1998) is considered.

	T - Aaa	T - Aa	T- A	T - Baa	Aaa - Aa	Aaa - A	Aaa - Baa	Aa - A	Aa - Baa	A - Baa
DF	-1.1126	-1.2185	-1.2286	-1.3477	-5.1035**	-3.3781*	-2.7935	-3.3225*	-2.8467	-3.642**
ADF1	-0.76146	-0.98468	-1.0283	-1.2659	-4.3831**	-2.8898*	-2.6082	-2.4953	-2.4872	-2.9817*
ADF2	-0.81788	-0.98766	-1.0388	-1.3245	-4.5047**	-2.875*	-2.7763	-2.2745	-2.4988	-2.9505*
ADF3	-0.90426	-0.99256	-1.0319	-1.374	-4.3949**	-2.8312	-2.8307	-2.1635	-2.5038	-2.9765*
ADF4	-0.91223	-0.98571	-1.0051	-1.3784	-4.4816**	-2.7972	-2.8929	-2.0932	-2.5522	-3.1397*
ADF5	-0.90747	-0.97528	-0.93669	-1.3485	-4.5053**	-2.7663	-2.8436	-2.0015	-2.5385	-3.1477*
PP	-0.9226	-1.036	-1.042	-1.407	-4.681**	-2.926*	-2.869*	-2.304	-2.656	-3.249*
KPSS	4.009**	4.8108**	4.9315**	5.0898**	2.6821**	10.335**	7.4391**	9.3078**	7.949**	1.5756**
KPSS HFO	1.0565**	1.2671**	1.2948**	1.3483**	0.80581**	2.8043**	2.0253**	2.459**	2.1548**	0.43417*

Table 2.26: Fractional cointegration analysis: Dittmann (2004) estimation procedure for bivariate systems using the narrow band FDLS residuals. For all possible bivariate systems $X - Y$ first the narrow band FDLS estimator (2.41) of the regression of Y on X is computed for different m. The last one is the OLS estimator (with intercept). Next the GPH estimator and the test statistic are computed for the residuals. The bandwidth parameter in the GPH estimation is set to $[n^{4/5}]$ and we choose $l = 0$ and $J = 1$. Finally the KPSS test without trend and with the Bartlett kernel and bandwidth parameter $\left[4\left(\frac{n}{100}\right)^{1/4}\right]$ for the estimation of the long run variance, is performed on the fractionally differenced residuals (2.51).

		T - Aaa	T - Aa	T- A	T - Baa	Aaa - Aa	Aaa - A	Aaa - Baa	Aa - A	Aa - Baa	A - Baa
$m = 23$	$\hat{\beta}_m$	0.74243	0.63936	0.56868	0.54938	0.86181	0.82778	0.78703	0.96271	0.92503	0.9561
	$\hat{\delta}(\hat{u})$	0.98741	0.98462	0.98743	1.0212	0.92114	0.92985	1.0344	0.8899	0.97552	0.96388
	$t_{d=1}$	-0.44504	-0.54344	-0.44409	0.74761	-2.7866**	-2.4789*	1.2171	-3.8907**	-0.86504	-1.2765
	KPSS	0.445	0.39127	0.37659	0.29511	0.31628	0.33678	0.20062	0.27833	0.23695	0.18659
$m = 51$	$\hat{\beta}_m$	0.74829	0.6464	0.57828	0.56105	0.86315	0.83101	0.7934	0.96457	0.9304	0.96007
	$\hat{\delta}(\hat{u})$	0.98909	0.98569	0.98748	1.0202	0.92156	0.92646	1.0352	0.88891	0.97519	0.96386
	$t_{d=1}$	-0.38541	-0.5055	-0.44253	0.71274	-2.7719**	-2.5987**	1.2445	-3.9257**	-0.8766	-1.2769
	KPSS	0.44318	0.39026	0.37776	0.29763	0.31383	0.33882	0.19815	0.27005	0.23158	0.17879
$m = 114$	$\hat{\beta}_m$	0.75012	0.64879	0.58112	0.56418	0.86389	0.83195	0.79485	0.96476	0.93129	0.96067
	$\hat{\delta}(\hat{u})$	0.98961	0.98609	0.98749	1.02	0.92185	0.92528	1.0354	0.88881	0.97508	0.96389
	$t_{d=1}$	-0.36731	-0.49162	-0.44197	0.70594	-2.7616**	-2.6405**	1.2524	-3.9291**	-0.88061	-1.276
	KPSS	0.44265	0.38987	0.3781	0.29819	0.31238	0.33967	0.19752	0.26918	0.23071	0.17757
$m = 252$	$\hat{\beta}_m$	0.75106	0.64999	0.58257	0.56568	0.86425	0.83242	0.7955	0.96485	0.93156	0.96081
	$\hat{\delta}(\hat{u})$	0.98987	0.98629	0.9875	1.0199	0.92201	0.92465	1.0355	0.88877	0.97504	0.9639
	$t_{d=1}$	-0.35801	-0.4844	-0.44166	0.70353	-2.7561**	-2.6627**	1.2561	-3.9307**	-0.8819	-1.2758
	KPSS	0.44238	0.38966	0.37828	0.29842	0.31166	0.34014	0.19724	0.26876	0.23045	0.17726
$m = 556$	$\hat{\beta}_m$	0.75151	0.65056	0.58331	0.56649	0.8644	0.83267	0.79589	0.96494	0.93179	0.96093
	$\hat{\delta}(\hat{u})$	0.99	0.98639	0.98751	1.0199	0.92208	0.9243	1.0356	0.88872	0.97501	0.9639
	$t_{d=1}$	-0.35355	-0.48089	-0.44151	0.70252	-2.7536**	-2.675**	1.2584	-3.9323**	-0.88304	-1.2756
	KPSS	0.44225	0.38956	0.37837	0.29854	0.31135	0.3404	0.19706	0.26834	0.23022	0.17701
$m = 1351$	$\hat{\beta}_m$	0.75171	0.65084	0.58367	0.5669	0.86447	0.83277	0.79608	0.96494	0.93189	0.96097
	$\hat{\delta}(\hat{u})$	0.99005	0.98644	0.98751	1.0199	0.92211	0.92416	1.0356	0.88872	0.975	0.9639
	$t_{d=1}$	-0.35159	-0.47914	-0.44143	0.7021	-2.7525**	-2.6799**	1.2595	-3.9322**	-0.8835	-1.2755
	KPSS	0.44219	0.38951	0.37841	0.29859	0.31121	0.34051	0.19698	0.26835	0.23013	0.17694

Table 2.27: Fractional cointegration analysis: Dittmann (2004) estimation procedure for bivariate systems using the narrow band FDLS residuals. For all possible bivariate systems $X - Y$ first the narrow band FDLS estimator (2.41) of the regression of Y on X is computed for different m. The last one is the OLS estimator (with intercept). Next the GPH estimator and the test statistic are computed for the residuals. The bandwidth parameter in the GPH estimation is set to $[n^{4/5}]$ and we choose $J = l = 1$. Finally the KPSS test without trend and with the Bartlett kernel and bandwidth parameter $\left[4\left(\frac{n}{100}\right)^{1/4}\right]$ for the estimation of the long run variance, is performed on the fractionally differenced residuals (2.51).

		T - Aaa	T - Aa	T- A	T - Baa	Aaa - Aa	Aaa - A	Aaa - Baa	Aa - A	Aa - Baa	A - Baa
$m = 23$	$\hat{\beta}_m$	0.74243	0.63936	0.56868	0.54938	0.86181	0.82778	0.78703	0.96271	0.92503	0.9561
	$\hat{\delta}(\hat{u})$	0.98774	0.98584	0.98878	1.0257	0.93441	0.92952	1.0409	0.88566	0.97728	0.97376
	$t_{d=1}$	-0.42108	-0.48645	-0.38547	0.88216	-2.2531*	-2.4212	1.4056	-3.9282**	-0.78036	-0.90133
	KPSS	0.44426	0.38887	0.37405	0.28797	0.29696	0.33725	0.19358	0.2817	0.23506	0.17716
$m = 51$	$\hat{\beta}_m$	0.74829	0.6464	0.57828	0.56105	0.86315	0.83101	0.7934	0.96457	0.9304	0.96007
	$\hat{\delta}(\hat{u})$	0.98928	0.9866	0.98837	1.0241	0.93471	0.92582	1.0415	0.88455	0.97678	0.97373
	$t_{d=1}$	-0.36827	-0.46051	-0.39939	0.82715	-2.2431*	-2.5484	1.4254	-3.9661**	-0.79767	-0.90255
	KPSS	0.44277	0.3885	0.37606	0.29144	0.29485	0.33974	0.19142	0.27332	0.22993	0.16981
$m = 114$	$\hat{\beta}_m$	0.75012	0.64879	0.58112	0.56418	0.86389	0.83195	0.79485	0.96476	0.93129	0.96067
	$\hat{\delta}(\hat{u})$	0.98974	0.98688	0.98826	1.0237	0.93493	0.92453	1.0417	0.88444	0.97663	0.97375
	$t_{d=1}$	-0.35233	-0.45066	-0.40338	0.81505	-2.2353*	-2.5926	1.4316	-3.9698**	-0.80268	-0.90179
	KPSS	0.44235	0.38832	0.37666	0.29225	0.29359	0.34073	0.19087	0.27243	0.2291	0.16866
$m = 252$	$\hat{\beta}_m$	0.75106	0.64999	0.58257	0.56568	0.86425	0.83242	0.7955	0.96485	0.93156	0.96081
	$\hat{\delta}(\hat{u})$	0.98998	0.98703	0.9882	1.0236	0.93506	0.92385	1.0418	0.88439	0.97659	0.97376
	$t_{d=1}$	-0.34415	-0.44546	-0.40541	0.81013	-2.231*	-2.616	1.4346	-3.9715**	-0.80427	-0.90158
	KPSS	0.44213	0.38821	0.37696	0.2926	0.29295	0.34128	0.19061	0.27201	0.22885	0.16837
$m = 556$	$\hat{\beta}_m$	0.75151	0.65056	0.58331	0.56649	0.8644	0.83267	0.79589	0.96494	0.93179	0.96093
	$\hat{\delta}(\hat{u})$	0.9901	0.98711	0.98817	1.0235	0.93511	0.92348	1.0418	0.88434	0.97655	0.97376
	$t_{d=1}$	-0.34023	-0.44291	-0.40643	0.80778	-2.2291*	-2.6289	1.4364	-3.9733**	-0.80567	-0.90141
	KPSS	0.44203	0.38816	0.37712	0.29278	0.29268	0.34158	0.19046	0.27159	0.22864	0.16814
$m = 1351$	$\hat{\beta}_m$	0.75171	0.65084	0.58367	0.5669	0.86447	0.83277	0.79608	0.96494	0.93189	0.96097
	$\hat{\delta}(\hat{u})$	0.99015	0.98714	0.98815	1.0235	0.93514	0.92332	1.0418	0.88434	0.97653	0.97376
	$t_{d=1}$	-0.33851	-0.44163	-0.40693	0.80668	-2.2282*	-2.6341	1.4374	-3.9732**	-0.80624	-0.90136
	KPSS	0.44199	0.38813	0.37719	0.29286	0.29256	0.3417	0.19038	0.2716	0.22855	0.16807

Table 2.28: Fractional cointegration analysis: Dittmann (2004) estimation procedure for bivariate systems using the first differenced narrow band FDLS residuals. For all possible bivariate systems $X - Y$ first the narrow band FDLS estimator (2.41) of the regression of Y on X is computed for different m. The last one is the OLS estimator (with intercept). Next the GPH estimator and the test statistic are computed for the first differenced residuals. The bandwidth parameter in the GPH estimation is set to $[n^{4/5}]$ and we choose $l = 0$ and $J = 1$. Finally the KPSS test without trend and with the Bartlett kernel and bandwidth parameter $\left[4\left(\frac{n}{100}\right)^{1/4}\right]$ for the estimation of the long run variance, is performed on the fractionally differenced residuals (2.51).

		T - Aaa	T - Aa	T- A	T - Baa	Aaa - Aa	Aaa - A	Aaa - Baa	Aa - A	Aa - Baa	A - Baa
$m = 23$	$\hat{\beta}_m$	0.74243	0.63936	0.56868	0.54938	0.86181	0.82778	0.78703	0.96271	0.92503	0.9561
	$\hat{\delta}(\Delta\hat{u})$	1.0147	0.98232	0.96133	1.0723	0.96185	0.962	1.0657	0.90746	0.99318	0.99013
	$\tau_{d=1}$	0.51848	-0.62488	-1.3665	2.5564*	-1.3483	-1.3429	2.3217*	-3.2699**	-0.24109	-0.34869
	KPSS	0.38731	0.39583	0.42804	0.22093	0.25929	0.29206	0.16819	0.26428	0.21831	0.16228
$m = 51$	$\hat{\beta}_m$	0.74829	0.6464	0.57828	0.56105	0.86315	0.83101	0.7934	0.96457	0.9304	0.96007
	$\hat{\delta}(\Delta\hat{u})$	1.015	0.97985	0.9618	1.0743	0.96162	0.9629	1.0662	0.90745	0.99437	0.98578
	$\tau_{d=1}$	0.53124	-0.71208	-1.3498	2.6247**	-1.3563	-1.3108	2.3384*	-3.2703**	-0.1988	-0.50235
	KPSS	0.38848	0.40183	0.42817	0.21943	0.25816	0.28866	0.16648	0.25599	0.21205	0.15926
$m = 114$	$\hat{\beta}_m$	0.75012	0.64879	0.58112	0.56418	0.86389	0.83195	0.79485	0.96476	0.93129	0.96067
	$\hat{\delta}(\Delta\hat{u})$	1.0152	0.9782	0.96191	1.0748	0.96148	0.96316	1.0663	0.90745	0.9946	0.98516
	$\tau_{d=1}$	0.53552	-0.77031	-1.346	2.6426**	-1.3612	-1.3018	2.3424*	-3.2705**	-0.19084	-0.52444
	KPSS	0.38884	0.4055	0.42828	0.21905	0.25755	0.28767	0.16608	0.25512	0.21096	0.15874
$m = 252$	$\hat{\beta}_m$	0.75106	0.64999	0.58257	0.56568	0.86425	0.83242	0.7955	0.96485	0.93156	0.96081
	$\hat{\delta}(\Delta\hat{u})$	1.0152	0.97738	0.96196	1.075	0.96141	0.96329	1.0663	0.90744	0.99467	0.98501
	$\tau_{d=1}$	0.53778	-0.79919	-1.3442	2.6516**	-1.3638	-1.2972	2.3442*	-3.2706**	-0.18848	-0.52988
	KPSS	0.38902	0.40733	0.42835	0.21885	0.25725	0.28718	0.1659	0.25471	0.21063	0.1586
$m = 556$	$\hat{\beta}_m$	0.75151	0.65056	0.58331	0.56649	0.8644	0.83267	0.79589	0.96494	0.93179	0.96093
	$\hat{\delta}(\Delta\hat{u})$	1.0152	0.97742	0.96198	1.0752	0.96138	0.96336	1.0664	0.90744	0.99472	0.98488
	$\tau_{d=1}$	0.53887	-0.79803	-1.3434	2.6566**	-1.3649	-1.2948	2.3453*	-3.2707**	-0.18644	-0.53426
	KPSS	0.38911	0.40735	0.42839	0.21874	0.25713	0.28692	0.16579	0.2543	0.21035	0.15849
$m = 1351$	$\hat{\beta}_m$	0.75171	0.65084	0.58367	0.5669	0.86447	0.83277	0.79608	0.96494	0.93189	0.96097
	$\hat{\delta}(\Delta\hat{u})$	1.0153	0.97754	0.962	1.0753	0.96136	0.96338	1.0664	0.90744	0.99475	0.98484
	$\tau_{d=1}$	0.53936	-0.79364	-1.3429	2.6591**	-1.3654	-1.2939	2.3458*	-3.2707**	-0.18564	-0.53554
	KPSS	0.38915	0.40715	0.4284	0.21869	0.25707	0.28681	0.16573	0.25431	0.21023	0.15846

Table 2.29: Fractional cointegration analysis: Dittmann (2004) estimation procedure for bivariate systems using the first differenced narrow band FDLS residuals. For all possible bivariate systems $X - Y$ first the narrow band FDLS estimator (2.41) of the regression of Y on X is computed for different m. The last one is the OLS estimator (with intercept). Next the GPH estimator and the test statistic are computed for the first differenced residuals. The bandwidth parameter in the GPH estimation is set to $[n^{4/5}]$ and we choose $J = l = 1$. Finally the KPSS test without trend and with the Bartlett kernel and bandwidth parameter $\left[4\left(\frac{n}{100}\right)^{1/4}\right]$ for the estimation of the long run variance, is performed on the fractionally differenced residuals (2.51).

		T - Aaa	T - Aa	T- A	T - Baa	Aaa - Aa	Aaa - A	Aaa - Baa	Aa - A	Aa - Baa	A - Baa
$m = 23$	$\hat{\beta}_m$	0.74243	0.63936	0.56868	0.54938	0.86181	0.82778	0.78703	0.96271	0.92503	0.9561
	$\hat{\delta}(\Delta\hat{u})$	1.0089	0.97473	0.95241	1.0704	0.9693	0.96063	1.0694	0.90359	0.99315	0.99937
	$\tau_{d=1}$	0.30463	-0.86817	-1.635	2.4187*	-1.0545	-1.3524	2.3852*	-3.3121**	-0.23527	-0.02149
	KPSS	0.39919	0.41107	0.44648	0.22347	0.24959	0.2939	0.16459	0.26739	0.21834	0.15427
$m = 51$	$\hat{\beta}_m$	0.74829	0.6464	0.57828	0.56105	0.86315	0.83101	0.7934	0.96457	0.9304	0.96007
	$\hat{\delta}(\Delta\hat{u})$	1.0092	0.97197	0.95273	1.0722	0.96912	0.96161	1.0699	0.90362	0.99448	0.9949
	$\tau_{d=1}$	0.31493	-0.96309	-1.6238	2.481*	-1.0609	-1.3187	2.4029*	-3.311**	-0.1898	-0.1753
	KPSS	0.40048	0.41776	0.44683	0.22211	0.24847	0.29037	0.16288	0.25891	0.21195	0.15159
$m = 114$	$\hat{\beta}_m$	0.75012	0.64879	0.58112	0.56418	0.86389	0.83195	0.79485	0.96476	0.93129	0.96067
	$\hat{\delta}(\Delta\hat{u})$	1.0093	0.97016	0.95279	1.0727	0.969	0.96189	1.0701	0.90362	0.99472	0.99425
	$\tau_{d=1}$	0.31844	-1.025	-1.6218	2.4972*	-1.065	-1.3091	2.4071*	-3.3111**	-0.1813	-0.1974
	KPSS	0.40087	0.4218	0.44701	0.22177	0.24786	0.28935	0.16248	0.25803	0.21084	0.15112
$m = 252$	$\hat{\beta}_m$	0.75106	0.64999	0.58257	0.56568	0.86425	0.83242	0.7955	0.96485	0.93156	0.96081
	$\hat{\delta}(\Delta\hat{u})$	1.0093	0.96927	0.95282	1.0729	0.96894	0.96203	1.0701	0.90362	0.9948	0.9941
	$\tau_{d=1}$	0.32029	-1.0557	-1.6208	2.5054*	-1.0671	-1.3044	2.409*	-3.3111**	-0.17878	-0.20284
	KPSS	0.40107	0.42382	0.44711	0.22159	0.24757	0.28884	0.1623	0.25761	0.2105	0.151
$m = 556$	$\hat{\beta}_m$	0.75151	0.65056	0.58331	0.56649	0.8644	0.83267	0.79589	0.96494	0.93179	0.96093
	$\hat{\delta}(\Delta\hat{u})$	1.0093	0.96929	0.95283	1.0731	0.96891	0.9621	1.0702	0.90362	0.99486	0.99397
	$\tau_{d=1}$	0.3212	-1.0549	-1.6204	2.5099*	-1.068	-1.3018	2.4101*	-3.3111**	-0.1766	-0.20723
	KPSS	0.40116	0.42385	0.44717	0.22149	0.24745	0.28857	0.16219	0.25719	0.21021	0.1509
$m = 1351$	$\hat{\beta}_m$	0.75171	0.65084	0.58367	0.5669	0.86447	0.83277	0.79608	0.96494	0.93189	0.96097
	$\hat{\delta}(\Delta\hat{u})$	1.0094	0.96942	0.95284	1.0731	0.9689	0.96213	1.0702	0.90362	0.99488	0.99393
	$\tau_{d=1}$	0.3216	-1.0505	-1.6202	2.5123*	-1.0684	-1.3008	2.4107*	-3.3111**	-0.17574	-0.20851
	KPSS	0.4012	0.42363	0.44719	0.22144	0.24739	0.28846	0.16214	0.2572	0.2101	0.15087

Table 2.30: Fractional cointegration analysis: Dittmann (2004) estimation procedure for bivariate systems using the narrow band FDLS tapered residuals. For all possible bivariate systems $X - Y$ first the narrow band FDLS estimator (2.41) of the regression of Y on X is computed for different m. The last one is the OLS estimator (with intercept). Next the GPH estimator and the test statistic are computed for the tapered residuals. The bandwidth parameter in the GPH estimation is set to $[n^{4/5}]$ and we choose $l = 0$ and $J = 1$. Finally the KPSS test without trend and with the Bartlett kernel and bandwidth parameter $\left[4\left(\frac{n}{100}\right)^{1/4}\right]$ for the estimation of the long run variance, is performed on the fractionally differenced residuals (2.51).

		T - Aaa	T - Aa	T- A	T - Baa	Aaa - Aa	Aaa - A	Aaa - Baa	Aa - A	Aa - Baa	A - Baa
$m = 23$	$\hat{\beta}_m$	0.74243	0.63936	0.56868	0.54938	0.86181	0.82778	0.78703	0.96271	0.92503	0.9561
	$\hat{\delta}(\hat{u})$	0.96096	1.0136	0.99213	1.0513	0.93205	0.98667	1.0286	0.83335	0.9934	0.98539
	$t_{d=1}$	-1.3797	0.47901	-0.2781	1.8138	-2.401*	-0.47102	1.0104	-5.8889**	-0.23329	-0.51639
	KPSS	0.50544*	0.33671	0.36775	0.24961	0.30034	0.25993	0.20713	0.32178	0.21808	0.1665
$m = 51$	$\hat{\beta}_m$	0.74829	0.6464	0.57828	0.56105	0.86315	0.83101	0.7934	0.96457	0.9304	0.96007
	$\hat{\delta}(\hat{u})$	0.96307	1.0152	0.99303	1.0553	0.93111	0.98661	1.0296	0.83248	0.99252	0.9847
	$t_{d=1}$	-1.305	0.53794	-0.24613	1.9527	-2.4343*	-0.47305	1.0445	-5.9197**	-0.26445	-0.54065
	KPSS	0.50223*	0.3349	0.36732	0.24507	0.29997	0.25818	0.20435	0.311	0.2139	0.16019
$m = 114$	$\hat{\beta}_m$	0.75012	0.64879	0.58112	0.56418	0.86389	0.83195	0.79485	0.96476	0.93129	0.96067
	$\hat{\delta}(\hat{u})$	0.96373	1.0158	0.99325	1.0564	0.93059	0.9866	1.0295	0.83239	0.99237	0.9846
	$t_{d=1}$	-1.2815	0.55745	-0.23868	1.9935*	-2.4527*	-0.47344	1.0412	-5.9228**	-0.26963	-0.54433
	KPSS	0.50123*	0.33431	0.36732	0.24373	0.29975	0.25765	0.20405	0.30987	0.21317	0.15922
$m = 252$	$\hat{\beta}_m$	0.75106	0.64999	0.58257	0.56568	0.86425	0.83242	0.7955	0.96485	0.93156	0.96081
	$\hat{\delta}(\hat{u})$	0.96408	1.016	0.99334	1.057	0.93034	0.9866	1.0294	0.83235	0.99233	0.98457
	$t_{d=1}$	-1.2694	0.56693	-0.23527	2.0137*	-2.4616*	-0.4736	1.04	-5.9243**	-0.2712	-0.54525
	KPSS	0.50072*	0.33404	0.36733	0.24307	0.29964	0.25738	0.20391	0.30934	0.21295	0.15897
$m = 556$	$\hat{\beta}_m$	0.75151	0.65056	0.58331	0.56649	0.8644	0.83267	0.79589	0.96494	0.93179	0.96093
	$\hat{\delta}(\hat{u})$	0.96424	1.0162	0.99339	1.0573	0.93023	0.9866	1.0296	0.83231	0.99229	0.98455
	$t_{d=1}$	-1.2637	0.57136	-0.23364	2.0247*	-2.4654*	-0.47368	1.0455	-5.9258**	-0.27258	-0.546
	KPSS	0.50047*	0.33392	0.36735	0.24271	0.29959	0.25724	0.20363	0.3088	0.21275	0.15878
$m = 1351$	$\hat{\beta}_m$	0.75171	0.65084	0.58367	0.5669	0.86447	0.83277	0.79608	0.96494	0.93189	0.96097
	$\hat{\delta}(\hat{u})$	0.96431	1.0162	0.99341	1.0575	0.93019	0.98659	1.0297	0.83231	0.99227	0.98454
	$t_{d=1}$	-1.2611	0.57354	-0.23289	2.0303*	-2.4671*	-0.47371	1.0497	-5.9257**	-0.27312	-0.54622
	KPSS	0.50037*	0.33385	0.36736	0.24252	0.29957	0.25718	0.20344	0.30881	0.21268	0.15872

Table 2.31: Fractional cointegration analysis: Dittmann (2004) estimation procedure for bivariate systems using the narrow band FDLS tapered residuals. For all possible bivariate systems $X - Y$ first the narrow band FDLS estimator (2.41) of the regression of Y on X is computed for different m. The last one is the OLS estimator (with intercept). Next the GPH estimator and the test statistic are computed for the tapered residuals. The bandwidth parameter in the GPH estimation is set to $[n^{4/5}]$ and we choose $J = l = 1$. Finally the KPSS test without trend and with the Bartlett kernel and bandwidth parameter $\left[4\left(\frac{n}{100}\right)^{1/4}\right]$ for the estimation of the long run variance, is performed on the fractionally differenced residuals (2.51).

		T - Aaa	T - Aa	T- A	T - Baa	Aaa - Aa	Aaa - A	Aaa - Baa	Aa - A	Aa - Baa	A - Baa
$m = 23$	$\hat{\beta}_m$	0.74243	0.63936	0.56868	0.54938	0.86181	0.82778	0.78703	0.96271	0.92503	0.9561
	$\hat{\delta}(\hat{u})$	0.93611	0.98865	0.96550	1.02754	0.9067	0.96202	1.00633	0.81032	0.97528	0.97083
	$t_{d=1}$	-2.1948*	-0.38976	-1.18531	0.94626	-3.20515**	-1.30488	0.21750	-6.5163**	-0.84922	-1.00213
	KPSS	0.5659*	0.38336	0.41957	0.28505	0.33809	0.29203	0.23319	0.33849	0.23721	0.17992
$m = 51$	$\hat{\beta}_m$	0.74829	0.6464	0.57828	0.56105	0.86315	0.83101	0.7934	0.96457	0.9304	0.96007
	$\hat{\delta}(\hat{u})$	0.93841	0.99043	0.96644	1.03168	0.90578	0.96204	1.00749	0.80973	0.97474	0.97060
	$t_{d=1}$	-2.1157*	-0.32876	-1.15295	1.088201	-3.2369**	-1.3039	0.257192	-6.5365**	-0.86762	-1.00998
	KPSS	0.5617*	0.381031	0.418799	0.279637	0.337499	0.289797	0.229751	0.326608	0.232049	0.172625
$m = 114$	$\hat{\beta}_m$	0.75012	0.64879	0.58112	0.56418	0.86389	0.83195	0.79485	0.96476	0.93129	0.96067
	$\hat{\delta}(\hat{u})$	0.93914	0.99102	0.96666	1.03289	0.90527	0.96206	1.00741	0.80967	0.97466	0.97057
	$t_{d=1}$	-2.091*	-0.30854	-1.14545	1.129896	-3.2544**	-1.30339	0.254549	-6.5385**	-0.87064	-1.01113
	KPSS	0.5605*	0.38028	0.418702	0.27805	0.337153	0.289128	0.229367	0.325368	0.231145	0.171502
$m = 252$	$\hat{\beta}_m$	0.75106	0.64999	0.58257	0.56568	0.86425	0.83242	0.7955	0.96485	0.93156	0.96081
	$\hat{\delta}(\hat{u})$	0.93951	0.99130	0.96676	1.03349	0.90502	0.96207	1.00738	0.80964	0.97463	0.97056
	$t_{d=1}$	-2.078*	-0.29871	-1.14201	1.150492	-3.2629**	-1.30309	0.253617	-6.5395**	-0.87157	-1.01142
	KPSS	0.5598*	0.379923	0.418676	0.277266	0.336981	0.28879	0.229186	0.324783	0.230873	0.171221
$m = 556$	$\hat{\beta}_m$	0.75151	0.65056	0.58331	0.56649	0.8644	0.83267	0.79589	0.96494	0.93179	0.96093
	$\hat{\delta}(\hat{u})$	0.93969	0.99144	0.96681	1.03382	0.90492	0.96207	1.00756	0.80962	0.97461	0.97055
	$t_{d=1}$	-2.0719*	-0.29411	-1.14037	1.161758	-3.2665**	-1.30292	0.259702	-6.5404**	-0.87238	-1.01166
	KPSS	0.5595*	0.379759	0.41867	0.276837	0.336907	0.288608	0.228846	0.324188	0.230637	0.170993
$m = 1351$	$\hat{\beta}_m$	0.75171	0.65084	0.58367	0.5669	0.86447	0.83277	0.79608	0.96494	0.93189	0.96097
	$\hat{\delta}(\hat{u})$	0.93977	0.99150	0.96683	1.03399	0.90487	0.96208	1.00769	0.80962	0.97460	0.97055
	$t_{d=1}$	-2.069*	-0.29186	-1.1396	1.167526	-3.2681**	-1.30285	0.264278	-6.5404**	-0.8727	-1.01173
	KPSS	0.5593*	0.379679	0.418668	0.276618	0.336874	0.288535	0.228623	0.324204	0.230543	0.170926

Table 2.32: Fractional cointegration analysis: Dittmann (2004) estimation procedure for bivariate systems using the first differenced narrow band FDLS tapered residuals. For all possible bivariate systems $X - Y$ first the narrow band FDLS estimator (2.41) of the regression of Y on X is computed for different m. The last one is the OLS estimator (with intercept). Next the GPH estimator and the test statistic are computed for the first differenced tapered residuals. The bandwidth parameter in the GPH estimation is set to $[n^{4/5}]$ and we choose $l = 0$ and $J = 1$. Finally the KPSS test without trend and with the Bartlett kernel and bandwidth parameter $\left[4\left(\frac{n}{100}\right)^{1/4}\right]$ for the estimation of the long run variance, is performed on the fractionally differenced residuals (2.51).

		T - Aaa	T - Aa	T- A	T - Baa	Aaa - Aa	Aaa - A	Aaa - Baa	Aa - A	Aa - Baa	A - Baa
$m = 23$	$\hat{\beta}_m$	0.74243	0.63936	0.56868	0.54938	0.86181	0.82778	0.78703	0.96271	0.92503	0.9561
	$\widehat{\delta}(\Delta\hat{u})$	0.96222	1.0117	0.98838	1.0455	0.91288	0.97626	1.0175	0.83796	0.98258	0.95788
	$\tau_{d=1}$	-1.3349	0.41268	-0.41056	1.6089	-3.0786**	-0.83895	0.61834	-5.7259**	-0.61562	-1.4883
	KPSS	0.50244*	0.34009	0.3748	0.25796	0.32866	0.27324	0.21987	0.31837	0.2294	0.19246
$m = 51$	$\hat{\beta}_m$	0.74829	0.6464	0.57828	0.56105	0.86315	0.83101	0.7934	0.96457	0.9304	0.96007
	$\widehat{\delta}(\Delta\hat{u})$	0.96305	1.0127	0.989	1.0472	0.91221	0.97614	1.0195	0.83736	0.98165	0.95679
	$\tau_{d=1}$	-1.3056	0.4473	-0.3886	1.6682	-3.1023**	-0.8433	0.68919	-5.7471**	-0.64832	-1.527
	KPSS	0.50227*	0.3395	0.37487	0.25652	0.32774	0.27142	0.21568	0.30759	0.2249	0.18542
$m = 114$	$\hat{\beta}_m$	0.75012	0.64879	0.58112	0.56418	0.86389	0.83195	0.79485	0.96476	0.93129	0.96067
	$\widehat{\delta}(\Delta\hat{u})$	0.9633	1.013	0.98915	1.0478	0.91186	0.9761	1.02	0.8373	0.98148	0.95698
	$\tau_{d=1}$	-1.2967	0.45826	-0.38329	1.689	-3.1147**	-0.8447	0.70655	-5.7492**	-0.65436	-1.5203
	KPSS	0.50224*	0.33934	0.37497	0.25595	0.32719	0.27089	0.21468	0.30646	0.22413	0.184
$m = 252$	$\hat{\beta}_m$	0.75106	0.64999	0.58257	0.56568	0.86425	0.83242	0.7955	0.96485	0.93156	0.96081
	$\widehat{\delta}(\Delta\hat{u})$	0.96343	1.0131	0.98922	1.0481	0.91169	0.97608	1.0202	0.83728	0.98143	0.95703
	$\tau_{d=1}$	-1.2922	0.46343	-0.3808	1.6995	-3.1206**	-0.84543	0.71392	-5.7502**	-0.65619	-1.5184
	KPSS	0.50223*	0.33928	0.37503	0.25565	0.32691	0.27062	0.21425	0.30592	0.22389	0.18364
$m = 556$	$\hat{\beta}_m$	0.75151	0.65056	0.58331	0.56649	0.8644	0.83267	0.79589	0.96494	0.93179	0.96093
	$\widehat{\delta}(\Delta\hat{u})$	0.96349	1.0132	0.98926	1.0483	0.91162	0.97606	1.0203	0.83725	0.98138	0.95708
	$\tau_{d=1}$	-1.2901	0.46581	-0.37959	1.7053	-3.123**	-0.84582	0.71832	-5.7512**	-0.6578	-1.5167
	KPSS	0.50222*	0.33926	0.37507	0.25549	0.32679	0.27048	0.21399	0.30538	0.22369	0.18334
$m = 1351$	$\hat{\beta}_m$	0.75171	0.65084	0.58367	0.5669	0.86447	0.83277	0.79608	0.96494	0.93189	0.96097
	$\widehat{\delta}(\Delta\hat{u})$	0.96352	1.0132	0.98927	1.0483	0.91159	0.97606	1.0204	0.83725	0.98137	0.95709
	$\tau_{d=1}$	-1.2891	0.46696	-0.37902	1.7083	-3.1241**	-0.84598	0.72048	-5.7512**	-0.65844	-1.5162
	KPSS	0.50222*	0.33925	0.37508	0.2554	0.32673	0.27042	0.21387	0.3054	0.22361	0.18325

Table 2.33: Fractional cointegration analysis: Dittmann (2004) estimation procedure for bivariate systems using the first differenced narrow band FDLS tapered residuals. For all possible bivariate systems $X - Y$ first the narrow band FDLS estimator (2.41) of the regression of Y on X is computed for different m. The last one is the OLS estimator (with intercept). Next the GPH estimator and the test statistic are computed for the first differenced tapered residuals. The bandwidth parameter in the GPH estimation is set to $[n^{4/5}]$ and we choose $J = l = 1$. Finally the KPSS test without trend and with the Bartlett kernel and bandwidth parameter $\left[4\left(\frac{n}{100}\right)^{1/4}\right]$ for the estimation of the long run variance, is performed on the fractionally differenced residuals (2.51).

|		T - Aaa	T - Aa	T- A	T - Baa	Aaa - Aa	Aaa - A	Aaa - Baa	Aa - A	Aa - Baa	A - Baa
$m = 23$	$\hat{\beta}_m$	0.74243	0.63936	0.56868	0.54938	0.86181	0.82778	0.78703	0.96271	0.92503	0.9561
	$\widehat{\delta}(\Delta\hat{u})$	0.95592	1.00877	0.98279	1.04424	0.92332	0.97375	1.01981	0.82550	0.98235	0.96334
	$\tau_{d=1}$	-1.51423	0.301343	-0.59129	1.519698	-2.6341**	-0.90169	0.680466	-5.9945**	-0.60634	-1.25953
	KPSS	0.5174*	0.345378	0.385468	0.259843	0.313055	0.276501	0.217174	0.327542	0.229647	0.187109
$m = 51$	$\hat{\beta}_m$	0.74829	0.6464	0.57828	0.56105	0.86315	0.83101	0.7934	0.96457	0.9304	0.96007
	$\widehat{\delta}(\Delta\hat{u})$	0.95667	1.00963	0.98324	1.04580	0.92270	0.97362	1.02202	0.82489	0.98147	0.96231
	$\tau_{d=1}$	-1.48852	0.330942	-0.57585	1.573294	-2.6556**	-0.90637	0.756305	-6.0156**	-0.63649	-1.29483
	KPSS	0.5173*	0.344974	0.38583	0.258566	0.312159	0.274662	0.21281	0.316263	0.225089	0.180234
$m = 114$	$\hat{\beta}_m$	0.75012	0.64879	0.58112	0.56418	0.86389	0.83195	0.79485	0.96476	0.93129	0.96067
	$\widehat{\delta}(\Delta\hat{u})$	0.95690	1.00990	0.98333	1.04636	0.92237	0.97357	1.02256	0.82483	0.98131	0.96253
	$\tau_{d=1}$	-1.48077	0.340172	-0.57253	1.59264	-2.6668**	-0.90787	0.774878	-6.0177**	-0.64211	-1.28735
	KPSS	0.5174*	0.344886	0.386015	0.258022	0.31162	0.274122	0.211773	0.315084	0.224305	0.178825
$m = 252$	$\hat{\beta}_m$	0.75106	0.64999	0.58257	0.56568	0.86425	0.83242	0.7955	0.96485	0.93156	0.96081
	$\widehat{\delta}(\Delta\hat{u})$	0.95701	1.01003	0.98338	1.04665	0.92222	0.97355	1.02279	0.82480	0.98126	0.96259
	$\tau_{d=1}$	-1.47683	0.344475	-0.57107	1.602464	-2.672**	-0.90864	0.782764	-6.0187**	-0.64382	-1.28519
	KPSS	0.5174*	0.344862	0.386123	0.257739	0.311349	0.273851	0.211325	0.314528	0.224069	0.178465
$m = 556$	$\hat{\beta}_m$	0.75151	0.65056	0.58331	0.56649	0.8644	0.83267	0.79589	0.96494	0.93179	0.96093
	$\widehat{\delta}(\Delta\hat{u})$	0.95707	1.01008	0.98340	1.04680	0.92215	0.97354	1.02292	0.82477	0.98121	0.96264
	$\tau_{d=1}$	-1.47495	0.346428	-0.57038	1.607897	-2.6742**	-0.90906	0.787474	-6.0197**	-0.64533	-1.28336
	KPSS	0.5174*	0.344856	0.386182	0.257582	0.311233	0.273705	0.211055	0.313961	0.223864	0.178171
$m = 1351$	$\hat{\beta}_m$	0.75171	0.65084	0.58367	0.5669	0.86447	0.83277	0.79608	0.96494	0.93189	0.96097
	$\widehat{\delta}(\Delta\hat{u})$	0.95709	1.01011	0.98341	1.04689	0.92213	0.97353	1.02299	0.82477	0.98120	0.96266
	$\tau_{d=1}$	-1.47413	0.347369	-0.57006	1.610692	-2.6752**	-0.90923	0.789783	-6.0196**	-0.64593	-1.28281
	KPSS	0.5174*	0.344854	0.386212	0.257501	0.311181	0.273647	0.210922	0.313976	0.223783	0.178084

Figure 2.10: Treasury, Aaa, Aa, A and Baa yields.
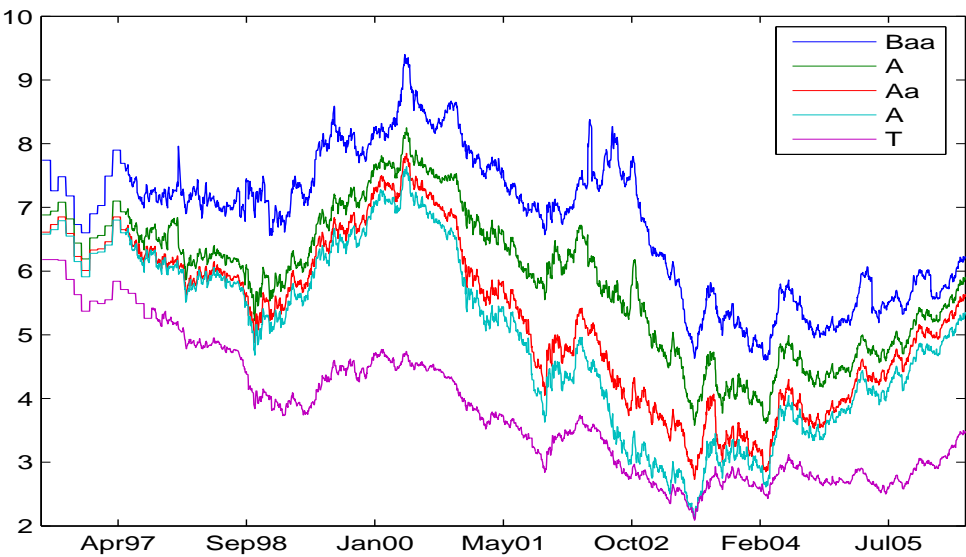
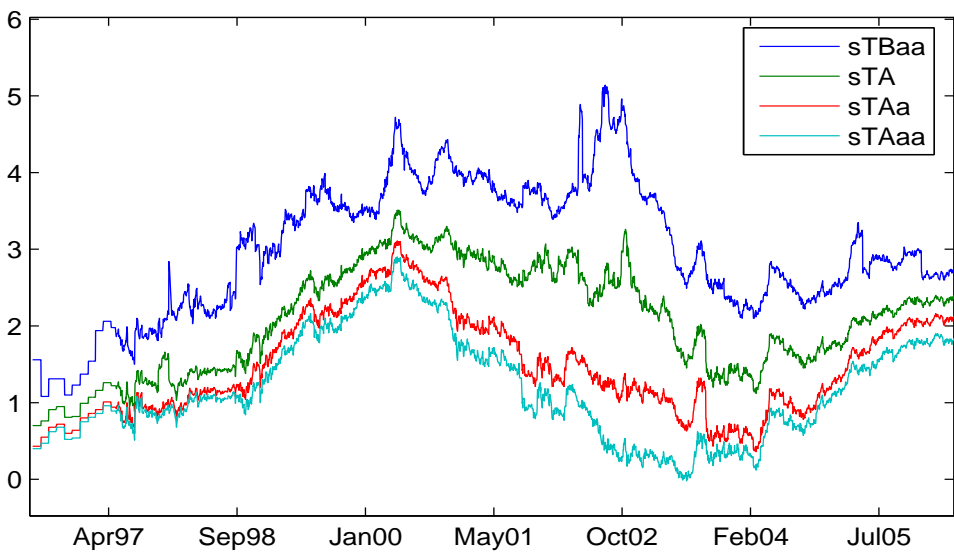**Figure 2.11:** Aaa, Aa, A and Baa spreads over Treasury yields.

Table 2.34: Summary Statistics and Normality Tests for yields and spreads. Both the Jarque–Bera test and the Normality test proposed by Doornik and Hansen (1994) are computed. In both cases the null hypothesis is that the series is normally distributed and the test statistic is χ^2_2. The p-value is in square bracket. One asterisk denotes significance at 5% level and two asterisks denote significance at 1% level.

	mean	std	skewness	ex. Kurtosis	JB test	DH test
T	3.7971	1.0329	0.49184	-0.85435	184.82	491.32
					[0.0000]**	[0.0000]**
Aaa	5.0023	1.3531	-0.1839	-1.1172	150.63	280.97
					[0.0000]**	[0.0000]**
Aa	5.2922	1.2345	-0.1256	-1.0434	125.41	209.97
					[0.0000]**	[0.0000]**
A	5.884	1.0786	-0.17544	-0.89106	99.848	164.72
					[0.0000]**	[0.0000]**
Baa	6.8015	1.0963	-0.14954	-0.84478	87.437	136.51
					[0.0000]**	[0.0000]**
sTAaa	1.2052	0.67704	0.35848	-0.79221	124.3	261.52
					[0.0000]**	[0.0000]**
sTAa	1.4951	0.6473	0.40889	-0.84553	150.65	350.45
					[0.0000]**	[0.0000]**
sTA	2.0869	0.68766	-0.07889	-1.1279	141.21	238.96
					[0.0000]**	[0.0000]**
sTBaa	3.0044	0.86577	-0.04827	-0.55195	34.183	41.948
					[0.0000]**	[0.0000]**
sAaaAa	0.28991	0.21031	1.3317	1.1194	908.8	1990
					[0.0000]**	[0.0000]**
sAaaA	0.88171	0.55901	1.1248	0.39386	567.9	1578.5
					[0.0000]**	[0.0000]**
sAaaBaa	1.7992	0.85093	1.1255	1.1131	686.55	954.28
					[0.0000]**	[0.0000]**
sAaA	0.5918	0.37765	1.0959	0.39871	540.33	1413.8
					[0.0000]**	[0.0000]**
sAaBaa	1.5093	0.67882	1.0193	1.2905	633.75	570.69
					[0.0000]**	[0.0000]**
sABaa	0.9175	0.39532	1.5203	5.4625	4255.4	633.83
					[0.0000]**	[0.0000]**

Table 2.35: Summary Statistics and Normality Tests for first differences of yields and spreads. Both the Jarque–Bera test and the Normality test proposed by Doornik and Hansen (1994) are computed. In both cases the null hypothesis is that the series is normally distributed and the test statistic is χ^2_2. The p-value is in square bracket. One asterisk denotes significance at 5% level and two asterisks denote significance at 1% level.

|	mean	std	skewness	ex. Kurtosis	JB test	DH test
ΔT	-0.00102	0.030141	-0.50965	13.697	20530 [0.0000]**	3560.2 [0.0000]**
ΔAaa	-0.00044	0.056111	0.11038	4.5905	2298.7 [0.0000]**	941.74 [0.0000]**
ΔAa	-0.00036	0.054982	0.12057	3.952	1706.1 [0.0000]**	757.68 [0.0000]**
ΔA	-0.00035	0.056562	-0.22236	6.0169	3961.7 [0.0000]**	1335 [0.0000]**
ΔBaa	-0.00054	0.095544	-5.9017	425.03	1.97E+07 [0.0000]**	85499 [0.0000]**
ΔsTAaa	5.74E-04	0.039787	0.06011	4.0579	1793.7 [0.0000]**	795.57 [0.0000]**
ΔsTAa	6.62E-04	0.039436	-0.05592	5.0338	2759.1 [0.0000]**	1080.1 [0.0000]**
ΔsTA	6.70E-04	0.042533	-0.30854	8.9381	8736 [0.0000]**	2206.3 [0.0000]**
ΔsTBaa	4.79E-04	0.09024	-6.9097	534.99	3.12E+07 [0.0000]**	1.01E+05 [0.0000]**
ΔsAaaAa	8.81E-05	0.016003	-6.9689	211.46	4.89E+06 [0.0000]**	9446.6 [0.0000]**
ΔsAaaA	9.57E-05	0.025676	-1.8983	56.592	3.50E+05 [0.0000]**	12937 [0.0000]**
ΔsAaaBaa	-9.57E-05	0.087108	-9.3362	638.72	4.44E+07 [0.0000]**	79671 [0.0000]**
ΔsAaA	7.66E-06	0.023299	-2.7988	67.146	4.94E+05 [0.0000]**	10512 [0.0000]**
ΔsAaBaa	-1.84E-04	0.086607	-9.6012	665.31	4.82E+07 [0.0000]**	81236 [0.0000]**
ΔsABaa	-1.91E-04	0.083233	-5.8308	578.27	3.64E+07 [0.0000]**	1.37E+05 [0.0000]**

Figure 2.12: Spreads between corporate yields.
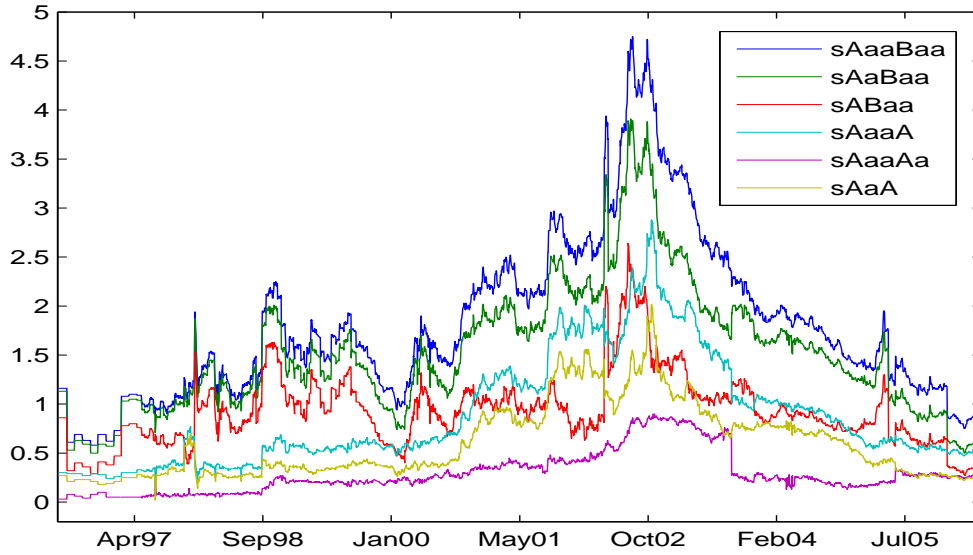
Table 2.36: Unit root and stationarity tests for yields and spreads. The Dickey-Fuller (DF) or augmented Dickey-Fuller with the constant (ADF), Phillips-Perron with the constant (PP), a two KPSS tests without trend are carried out. n is the lag length in the ADF and it is chosen by the AIC. In the first KPSS test the Bartlett kernel with bandwidth parameter $\left[4\left(\frac{n}{100}\right)^{1/4}\right]$ is chosen for the estimation of the long run variance. In the second test the automatic bandwidth selection procedure of Hobijn et al. (1998) is considered. One asterisk denotes significance at 5% level and two asterisks denote significance at 1% level.

	DF-ADF	n	KPSS	KPSS HFO	PP
T	-1.653	100	21.939**	5.8588**	-2.4475
Aaa	-1.3355	5	15.53**	4.1328**	-1.3556
Aa	-1.4069	1	15.243**	4.0624**	-1.3791
A	-1.4417	4	14.781**	3.948**	-1.4918
Baa	-1.3117	100	14.614**	3.935**	-2.0929
sTAaa	-1.2298	20	3.0473**	0.81407**	-1.3345
sTAa	-1.5499	14	3.1227**	0.8378**	-1.6637
sTA	-2.0273	9	5.7085**	1.5349**	-2.0796
sTBaa	-2.395	100	5.7847**	1.5748**	-2.6558
sAaaAa	-1.5907	40	8.2945**	2.2227**	-1.8038
sAaaA	-1.5777	12	8.2115**	2.1942**	-1.4043
sAaaBaa	-1.637	100	6.0188**	1.6326**	-2.3115
sAaA	-1.9143	13	7.1743**	1.9317**	-1.6063
sAaBaa	-1.6902	100	5.1076**	1.4035**	-2.8231
sABaa	-2.5695	100	2.5392**	0.75018**	-5.0142

Table 2.37: Unit root and stationarity tests for the first differences of yields and spreads. The Dickey-Fuller (DF) or augmented Dickey-Fuller with the constant (ADF), Phillips-Perron with the constant (PP), a two KPSS tests without trend are carried out. n is the lag length in the ADF and it is chosen by the AIC. In the first KPSS test the Bartlett kernel with bandwidth parameter $\left[4 \left(\frac{n}{100} \right)^{1/4} \right]$ is chosen for the estimation of the long run variance. In the second test the automatic bandwidth selection procedure of Hobijn et al. (1998) is considered. One asterisk denotes significance at 5% level and two asterisks denote significance at 1% level.

	DF-ADF	n	KPSS	KPSS HFO	PP
ΔT	-5.0327**	100	0.49783*	0.57193*	-46.559**
ΔAaa	-31.028**	2	0.28466	0.28348	-48.534**
ΔAa	-48.939**	0	0.23192	0.23514	-48.9**
ΔA	-30.982**	2	0.13926	0.13961	-49.026**
ΔBaa	-5.4058**	100	0.043328	0.067559	-51.144**
$\Delta sTAaa$	-12.988**	15	0.2278	0.23711	-55.581**
$\Delta sTAa$	-16.111**	9	0.20595	0.20595	-55.308**
ΔsTA	-17.149**	8	0.2161	0.20603	-53.492**
$\Delta sTBaa$	-5.4791**	100	0.087047	0.17655	-52.481**
$\Delta sAaaAa$	-9.8791**	27	0.13711	0.14612	-58.483**
$\Delta sAaaA$	-47.969**	0	0.2518	0.23199	-48.007**
$\Delta sAaaBaa$	-5.5451**	100	0.085492	0.15935	-51.546**
$\Delta sAaA$	-12.846**	11	0.16635	0.15742	-49.314**
$\Delta sAaBaa$	-5.9663**	100	0.0595	0.12429	-51.976**
$\Delta sABaa$	-6.2875**	100	0.018934	0.049348	-52.222**

Table 2.38: d estimates for the yields with $l = 0$ and non-tapered data. For every series X, this table report the estimates $\hat{d}(\Delta X) \equiv \widehat{d-1}(\Delta X) + 1$ along with the test statistic (2.15). One asterisk denotes significance at 5% level and two asterisks denote significance at 1% level.

			T	Aaa	Aa	A	Baa
	$m = 23$	$\hat{d}(\Delta X)$	1.2581	1.2515	1.1604	1.083	1.0871
		$\tau_{d=1}$	1.5604	1.5201	0.9697	0.50184	0.52656
	$m = 51$	$\hat{d}(\Delta X)$	1.1294	1.1236	1.0096	0.96692	1.0217
		$\tau_{d=1}$	1.2699	1.213	0.093876	-0.32475	0.21305
$J = 1$	$m = 112$	$\hat{d}(\Delta X)$	1.0682	1.0422	0.99158	1.0198	0.85643
		$\tau_{d=1}$	1.0449	0.64641	-0.12902	0.30295	-2.1994*
	$m = 246$	$\hat{d}(\Delta X)$	1.0208	1.0409	1.0359	1.0575	0.74505
		$\tau_{d=1}$	0.4864	0.95613	0.83888	1.3435	-5.956**
	$m = 556$	$\hat{d}(\Delta X)$	1.0148	0.99107	1.0113	1	0.97023
		$\tau_{d=1}$	0.5147	-0.31107	0.39283	0.0015018	-1.0371
	$m = 22$	$\hat{d}(\Delta X)$	1.1826	1.2629	1.2284	1.078	0.96728
		$\tau_{d=1}$	0.67096	0.96584	0.83919	0.28667	-0.12023
	$m = 50$	$\hat{d}(\Delta X)$	1.1465	1.1281	1.0176	0.92094	0.98424
		$\tau_{d=1}$	0.93375	0.81619	0.11237	-0.50375	-0.10041
$J = 2$	$m = 112$	$\hat{d}(\Delta X)$	1.0602	1.0223	0.99915	0.99495	0.85764
		$\tau_{d=1}$	0.62374	0.23103	-0.0087573	-0.052284	-1.4748
	$m = 246$	$\hat{d}(\Delta X)$	1.033	1.022	1.0354	1.034	0.78248
		$\tau_{d=1}$	0.53074	0.35311	0.56936	0.54746	-3.4981**
	$m = 540$	$\hat{d}(\Delta X)$	1.0267	0.9821	1.0001	0.98612	0.97088
		$\tau_{d=1}$	0.64605	-0.43349	0.0034196	-0.33598	-0.70518
	$m = 21$	$\hat{d}(\Delta X)$	1.1145	1.2036	1.2009	1.1472	1.0003
		$\tau_{d=1}$	0.29945	0.53258	0.52556	0.38513	0.00079007
	$m = 51$	$\hat{d}(\Delta X)$	1.1662	1.0994	1.0029	0.90823	0.88635
		$\tau_{d=1}$	0.82486	0.49336	0.014482	-0.45542	-0.56399
$J = 3$	$m = 111$	$\hat{d}(\Delta X)$	1.0624	1.0168	0.98903	0.98403	0.84686
		$\tau_{d=1}$	0.50589	0.13617	-0.088924	-0.12953	-1.2418
	$m = 246$	$\hat{d}(\Delta X)$	1.0217	1.0077	1.0134	1.0205	0.78128
		$\tau_{d=1}$	0.27872	0.098729	0.17264	0.26338	-2.8083**
	$m = 540$	$\hat{d}(\Delta X)$	1.0076	0.97012	0.97491	0.96528	0.96627
		$\tau_{d=1}$	0.14766	-0.58275	-0.48925	-0.67707	-0.65778

Table 2.39: d estimates for the spreads with $l = 0$ and non-tapered data. For every series X, this table report the estimates $\hat{d}(\Delta X) \equiv \widehat{d-1}(\Delta X) + 1$ along with the test statistic (2.15). One asterisk denotes significance at 5% level and two asterisks denote significance at 1% level.

			sTAaa	sTAa	sTA	sTBaa	sAaaAa	sAaaA	sAaaBaa	sAaA	sAaBaa	sABaa
	$m = 23$	$\hat{d}(\Delta X)$	1.2842	1.2771	1.0939	0.90145	1.2824	1.001	1.0814	0.84378	0.97525	0.91161
		$\tau_{d=1}$	1.7178	1.6751	0.56784	-0.59567	1.707	0.0061963	0.49206	-0.94429	-0.14959	-0.53428
	$m = 51$	$\hat{d}(\Delta X)$	1.0892	1.0207	0.86605	1.003	1.084	0.99138	1.0394	0.90515	0.97152	0.9241
		$\tau_{d=1}$	0.87569	0.20275	-1.315	0.029142	0.82471	-0.084613	0.38672	-0.93115	-0.27958	-0.74513
$J = 1$	$m = 112$	$\hat{d}(\Delta X)$	0.98681	0.93182	1.014	0.7928	1.0093	1.0548	0.8507	1.0269	0.81457	0.73488
		$\tau_{d=1}$	-0.2021	-1.0445	0.21447	-3.1741**	0.14258	0.83966	-2.2871	0.41208	-2.8405**	-4.0613**
	$m = 246$	$\hat{d}(\Delta X)$	0.97249	0.97061	0.97969	0.67172	0.97788	1.0765	0.68616	1.0103	0.66432	0.60893
		$\tau_{d=1}$	-0.64277	-0.68661	-0.47458	-7.6691**	-0.51672	1.7879	-7.3318	0.24046	-7.8419**	-9.136**
	$m = 541$	$\hat{d}(\Delta X)$	0.91012	0.91467	0.91893	0.92445	0.92679	1.0143	0.94304	0.95264	0.94656	0.96151
		$\tau_{d=1}$	-3.1308**	-2.9722**	-2.8237**	-2.6316**	-2.55*	0.49733	-1.9842*	-1.6495	-1.8615	-1.3408
	$m = 22$	$\hat{d}(\Delta X)$	1.2269	1.2556	1.0385	0.928	1.2656	1.0239	1.0166	0.81106	0.90499	0.84669
		$\tau_{d=1}$	0.83364	0.93919	0.14144	-0.26454	0.9759	0.087669	0.060898	-0.6942	-0.34907	-0.56329
	$m = 50$	$\hat{d}(\Delta X)$	1.0579	0.95946	0.88081	0.95231	1.0661	0.94956	0.9755	0.83398	0.91914	0.88095
		$\tau_{d=1}$	0.36926	-0.25833	-0.7595	-0.30388	0.42115	-0.32143	-0.1561	-1.0579	-0.51526	-0.75863
$J = 2$	$m = 112$	$\hat{d}(\Delta X)$	0.94896	0.91678	0.98779	0.78335	1.0151	1.0305	0.8053	0.99712	0.76965	0.68177
		$\tau_{d=1}$	-0.5288	-0.86217	-0.12648	-2.2444*	0.15628	0.31604	-2.0171	-0.029806	-2.3863*	-3.2967**
	$m = 246$	$\hat{d}(\Delta X)$	0.938	0.95109	0.96599	0.70733	0.97893	1.0677	0.68778	1.0183	0.66653	0.60789
		$\tau_{d=1}$	-0.99713	-0.78661	-0.54696	-4.7066**	-0.33879	1.0887	-5.0211	0.29487	-5.3627**	-6.3058**
	$m = 540$	$\hat{d}(\Delta X)$	0.89623	0.90968	0.91502	0.95424	0.91996	1.0055	0.95067	0.95856	0.9486	0.95037
		$\tau_{d=1}$	-2.5127*	-2.187*	-2.0578*	-1.1081	-1.9382	0.13396	-1.1946	-1.0033	-1.2447	-1.2018
	$m = 21$	$\hat{d}(\Delta X)$	1.2875	1.3558	1.1515	0.97172	1.3603	1.0681	1.027	0.8034	0.89875	0.80958
		$\tau_{d=1}$	0.75206	0.93062	0.39632	-0.073973	0.94241	0.17825	0.070726	-0.51424	-0.26485	-0.49808
	$m = 51$	$\hat{d}(\Delta X)$	1.102	1.0298	0.86285	0.87513	1.0528	0.94813	0.9567	0.84687	0.88577	0.82421
		$\tau_{d=1}$	0.50608	0.14808	-0.68059	-0.61967	0.26222	-0.25739	-0.2149	-0.75993	-0.56686	-0.87238
$J = 3$	$m = 111$	$\hat{d}(\Delta X)$	0.97016	0.96446	0.96795	0.797	0.99019	1.0275	0.8078	1.0098	0.77858	0.69052
		$\tau_{d=1}$	-0.24195	-0.28824	-0.25993	-1.6463	-0.079541	0.22333	-1.5587	0.079266	-1.7956	-2.5097*
	$m = 246$	$\hat{d}(\Delta X)$	0.93912	0.97025	0.97433	0.71867	0.97007	1.075	0.68909	1.0317	0.67452	0.6193
		$\tau_{d=1}$	-0.78171	-0.38197	-0.32961	-3.6123**	-0.38434	0.96264	-3.9921**	0.40688	-4.1791**	-4.8882**
	$m = 540$	$\hat{d}(\Delta X)$	0.90421	0.91539	0.91342	0.9544	0.91137	1.002	0.95337	0.96212	0.95584	0.96029
		$\tau_{d=1}$	-1.8682	-1.6501	-1.6886	-0.88943	-1.7285	0.039412	-0.90953	-0.73876	-0.86122	-0.77449

Table 2.40: d estimates for the yields with $l = 1$ and non-tapered data. For every series X, this table report the estimates $\hat{d}(\Delta X) \equiv \widehat{d-1}(\Delta X) + 1$ along with the test statistic (2.15). One asterisk denotes significance at 5% level and two asterisks denote significance at 1% level.

			T	Aaa	Aa	A	Baa
	$m = 23$	$\hat{d}(\Delta X)$	1.3759	1.2371	1.0897	1.0369	1.1007
		$\tau_{d=1}$	1.8313	1.1554	0.43725	0.17976	0.49064
	$m = 51$	$\hat{d}(\Delta X)$	1.1562	1.1007	0.95973	0.93141	1.014
		$\tau_{d=1}$	1.346	0.86782	-0.34698	-0.59092	0.12019
$J = 1$	$m = 112$	$\hat{d}(\Delta X)$	1.0733	1.0219	0.96663	1.0101	0.83367
		$\tau_{d=1}$	1.0377	0.30963	-0.47219	0.1432	-2.3536*
	$m = 246$	$\hat{d}(\Delta X)$	1.0202	1.031	1.027	1.0554	0.72618
		$\tau_{d=1}$	0.45048	0.69009	0.60237	1.233	-6.0982**
	$m = 541$	$\hat{d}(\Delta X)$	1.0143	0.98406	1.0058	0.99653	0.97097
		$\tau_{d=1}$	0.48299	-0.53958	0.19782	-0.11746	-0.98258
	$m = 23$	$\hat{d}(\Delta X)$	1.3195	1.1994	1.0626	0.96704	0.95092
		$\tau_{d=1}$	1.0122	0.63175	0.19826	-0.10442	-0.15548
	$m = 51$	$\hat{d}(\Delta X)$	1.2089	1.0938	0.97749	0.88543	0.91273
		$\tau_{d=1}$	1.2087	0.54283	-0.13025	-0.66289	-0.50492
$J = 2$	$m = 111$	$\hat{d}(\Delta X)$	1.0668	1.007	0.96148	0.96577	0.84231
		$\tau_{d=1}$	0.64348	0.067874	-0.37124	-0.32985	-1.5196
	$m = 245$	$\hat{d}(\Delta X)$	1.0198	0.99279	1.0003	1.0351	0.76307
		$\tau_{d=1}$	0.3049	-0.11113	0.0039196	0.54125	-3.6502**
	$m = 541$	$\hat{d}(\Delta X)$	0.99835	0.96241	0.96757	0.96695	0.96452
		$\tau_{d=1}$	-0.038898	-0.88873	-0.76679	-0.78145	-0.83882
	$m = 22$	$\hat{d}(\Delta X)$	1.2125	1.1807	1.1392	1.0175	0.94072
		$\tau_{d=1}$	0.48897	0.41577	0.3202	0.040309	-0.13638
	$m = 49$	$\hat{d}(\Delta X)$	1.1523	1.0693	0.98725	0.89863	0.92421
		$\tau_{d=1}$	0.66643	0.30319	-0.055805	-0.44367	-0.3317
$J = 3$	$m = 112$	$\hat{d}(\Delta X)$	1.0865	1.0385	1.0228	1.0132	0.84473
		$\tau_{d=1}$	0.66476	0.29592	0.17539	0.10168	-1.1932
	$m = 244$	$\hat{d}(\Delta X)$	1.0335	1.0266	1.0473	1.0419	0.78075
		$\tau_{d=1}$	0.41309	0.32814	0.58284	0.51635	-2.7021**
	$m = 541$	$\hat{d}(\Delta X)$	1.0209	0.98464	0.99679	0.97531	0.9764
		$\tau_{d=1}$	0.39919	-0.29327	-0.061296	-0.47139	-0.4506

Table 2.42 and Table 2.43 report the d estimates for the yields and the spreads when $l = 0$ and the data is tapered.

Table 2.44 and Table 2.45 report the d estimates for the yields and the spreads when $l = 1$ and the data is tapered.

Table 2.46 and Table 2.47 report the local Whittle estimates for yields and spreads respectively.

Once again, the main conclusion is that credit spreads are long memory nonstationary processes. The same conclusion cannot always be drawn for yields, for which the null $H_0 : d = 1$ is not very often rejected.

Table 2.48 reports the result of Nielsen (2005) LM test for yields when setting $d = 1$ in eq. (2.18) or $\mathbf{d} = \iota$ in eq. (2.21) for the multivariate case. Panel A reports univariate tests whereas Panel B reports multivariate tests. Table 2.49 reports the result of Nielsen (2005) LM test for spreads when setting $d = 1$ in eq. (2.18)

Table 2.41: d estimates for the spreads with $l = 1$ and non-tapered data. For every series X, this table report the estimates $\hat{d}(\Delta X) \equiv \widehat{d-1}(\Delta X) + 1$ along with the test statistic (2.15). One asterisk denotes significance at 5% level and two asterisks denote significance at 1% level.

			sTAaa	sTAa	sTA	sTBaa	sAaaAa	sAaaA	sAaaBaa	sAaA	sAaBaa	sABaa
	$m = 23$	$\hat{d}(\Delta X)$	1.1701	1.1398	0.9199	0.7483	1.3607	0.90304	1.0415	0.68071	0.91293	0.95979
		$\tau_{d=1}$	0.82871	0.68096	-0.39028	-1.2263	1.7573	-0.47238	0.20203	-1.5557	-0.4242	-0.19593
	$m = 51$	$\hat{d}(\Delta X)$	1.0141	0.92667	0.7551	0.95332	1.0829	0.94366	1.0112	0.84406	0.94138	0.944
		$\tau_{d=1}$	0.12169	-0.63181	-2.1099*	-0.40215	0.71385	-0.4854	0.096296	-1.3435	-0.50508	-0.48249
$J = 1$	$m = 112$	$\hat{d}(\Delta X)$	0.94058	0.87891	0.98142	0.74608	1.0006	1.0396	0.81616	1.0124	0.78285	0.72336
		$\tau_{d=1}$	-0.84075	-1.7135	-0.26292	-3.5929**	0.0087585	0.56067	-2.6013**	0.1754	-3.0727**	-3.9145**
	$m = 246$	$\hat{d}(\Delta X)$	0.94903	0.94795	0.96128	0.64073	0.9713	1.0699	0.65802	1.0011	0.63845	0.5946
		$\tau_{d=1}$	-1.1351	-1.1591	-0.86233	-8.0013**	-0.63918	1.5568	-7.616**	0.024421	-8.052**	-9.0286**
	$m = 541$	$\hat{d}(\Delta X)$	0.89596	0.90114	0.90726	0.92044	0.92141	1.0082	0.94063	0.94549	0.94637	0.97006
		$\tau_{d=1}$	-3.5212**	-3.3459**	-3.1386**	-2.6926**	-2.6597**	0.27795	-2.0093*	-1.8447	-1.815	-1.0133
	$m = 23$	$\hat{d}(\Delta X)$	1.1189	1.1007	0.81821	0.79465	1.2277	0.91241	0.99823	0.69297	0.86705	0.83902
		$\tau_{d=1}$	0.3766	0.31905	-0.57593	-0.65057	0.72131	-0.27749	-0.0056051	-0.97269	-0.42119	-0.51
	$m = 51$	$\hat{d}(\Delta X)$	1.0578	1.0043	0.77206	0.88315	1.0976	0.89983	0.9462	0.79136	0.86697	0.88421
		$\tau_{d=1}$	0.33429	0.024886	-1.3188	-0.67608	0.5647	-0.5796	-0.31129	-1.2072	-0.76971	-0.66997
$J = 2$	$m = 111$	$\hat{d}(\Delta X)$	0.94313	0.9224	0.91163	0.77666	0.99698	1.0258	0.79501	0.99304	0.75674	0.70038
		$\tau_{d=1}$	-0.54809	-0.74788	-0.85159	-2.1524*	-0.029142	0.24891	-1.9755	-0.067079	-2.3443*	-2.8874**
	$m = 245$	$\hat{d}(\Delta X)$	0.9395	0.96127	0.96641	0.68919	0.96173	1.0567	0.6659	1.0237	0.64683	0.61228
		$\tau_{d=1}$	-0.93212	-0.59672	-0.51751	-4.7884**	-0.5896	0.87381	-5.147**	0.36503	-5.4409**	-5.9731**
	$m = 541$	$\hat{d}(\Delta X)$	0.90036	0.91179	0.91001	0.94825	0.91302	0.99138	0.94922	0.9513	0.95134	0.96816
		$\tau_{d=1}$	-2.356*	-2.0858*	-2.1277*	-1.2237	-2.0566*	-0.20391	-1.2007	-1.1515	-1.1506	-0.75294
	$m = 22$	$\hat{d}(\Delta X)$	1.0739	0.96073	0.79603	0.81196	1.327	0.98387	0.97927	0.66956	0.82034	0.822
		$\tau_{d=1}$	0.16997	-0.090348	-0.46929	-0.43262	0.75237	-0.037114	-0.047689	-0.76026	-0.41335	-0.40954
	$m = 49$	$\hat{d}(\Delta X)$	1.0315	0.94349	0.71925	0.84775	1.0649	0.91049	0.89909	0.78573	0.84542	0.87466
		$\tau_{d=1}$	0.13793	-0.24731	-1.2288	-0.66636	0.28392	-0.39175	-0.44164	-0.93782	-0.67657	-0.54858
$J = 3$	$m = 112$	$\hat{d}(\Delta X)$	0.95327	0.91758	0.93779	0.7614	1.0046	1.0455	0.78145	0.98456	0.75053	0.68538
		$\tau_{d=1}$	-0.35913	-0.63334	-0.47803	-1.8335	0.034985	0.34953	-1.6795	-0.11867	-1.917	-2.4177*
	$m = 244$	$\hat{d}(\Delta X)$	0.92885	0.9427	0.9518	0.70227	0.96135	1.0662	0.6788	1.0337	0.65873	0.61628
		$\tau_{d=1}$	-0.87684	-0.70623	-0.59401	-3.6694**	-0.47632	0.81533	-3.9586**	0.41525	-4.206**	-4.7291**
	$m = 541$	$\hat{d}(\Delta X)$	0.90126	0.90693	0.89682	0.9601	0.90361	0.99342	0.95766	0.96526	0.9581	0.96811
		$\tau_{d=1}$	-1.8852	-1.777	-1.97*	-0.7618	-1.8403	-0.12563	-0.80836	-0.66331	-0.79987	-0.60893

Table 2.42: d estimates for the yields with $l = 0$ and tapered data. For every series X, this table report the estimates $\hat{d}(\Delta X) \equiv \widehat{d-1}(\Delta X) + 1$ along with the test statistic (2.15). One asterisk denotes significance at 5% level and two asterisks denote significance at 1% level.

			T	Aaa	Aa	A	Baa
	$m = 23$	$\hat{d}(\Delta X)$	1.2899	1.478	1.3751	1.1379	1.0869
		$\tau_{d=1}$	1.7523	2.8896	2.2672	0.83366	0.52532
	$m = 51$	$\hat{d}(\Delta X)$	1.1664	1.187	1.1156	0.94489	0.96329
		$\tau_{d=1}$	1.6333	1.8361	1.1345	-0.54101	-0.36041
$J = 1$	$m = 112$	$\hat{d}(\Delta X)$	1.1419	1.0675	1.0548	1.0088	0.99412
		$\tau_{d=1}$	2.1731*	1.0343	0.8397	0.13463	-0.090087
	$m = 246$	$\hat{d}(\Delta X)$	1.0404	1.029	1.0417	1.0339	1.0181
		$\tau_{d=1}$	0.94409	0.67689	0.97454	0.79283	0.4221
	$m = 541$	$\hat{d}(\Delta X)$	1.0165	0.98709	0.98741	0.94704	1.0251
		$\tau_{d=1}$	0.57379	-0.44964	-0.43852	-1.8447	0.87521
	$m = 22$	$\hat{d}(\Delta X)$	1.2432	1.3822	1.3452	1.2135	1.1297
		$\tau_{d=1}$	0.89372	1.4041	1.2683	0.7846	0.47637
	$m = 50$	$\hat{d}(\Delta X)$	1.1651	1.16	1.0681	0.92461	0.98681
		$\tau_{d=1}$	1.0521	1.0197	0.43396	-0.4804	-0.08407
$J = 2$	$m = 112$	$\hat{d}(\Delta X)$	1.1196	1.06	1.0417	1.0011	0.98983
		$\tau_{d=1}$	1.2388	0.6212	0.43199	0.011856	-0.10536
	$m = 246$	$\hat{d}(\Delta X)$	1.0399	1.0347	1.0441	1.0432	1.0183
		$\tau_{d=1}$	0.64198	0.55747	0.70873	0.69545	0.29416
	$m = 540$	$\hat{d}(\Delta X)$	1.0241	0.98924	0.99199	0.96007	1.0372
		$\tau_{d=1}$	0.58437	-0.26057	-0.19386	-0.96678	0.90108
	21	$\hat{d}(\Delta X)$	1.1945	1.3976	1.3983	1.2326	1.1202
		$\tau_{d=1}$	0.50878	1.0401	1.0417	0.60846	0.31449
	$m = 51$	$\hat{d}(\Delta X)$	1.1237	1.1377	1.0709	0.91235	0.92345
		$\tau_{d=1}$	0.61391	0.68335	0.35206	-0.43496	-0.37986
$J = 3$	$m = 111$	$\hat{d}(\Delta X)$	1.0867	1.0324	1.0131	0.9748	0.98964
		$\tau_{d=1}$	0.7033	0.26269	0.10607	-0.20434	-0.084021
	$m = 246$	$\hat{d}(\Delta X)$	1.0327	1.021	1.0283	1.0346	1.0203
		$\tau_{d=1}$	0.41999	0.27008	0.36347	0.44408	0.26123
	$m = 540$	$\hat{d}(\Delta X)$	1.023	0.97857	0.97851	0.95116	1.0339
		$\tau_{d=1}$	0.44773	-0.41788	-0.41917	-0.95264	0.66133

Table 2.43: d estimates for the spreads with $l = 0$ and tapered data. For every series X, this table report the estimates $\hat{d}(\Delta X) \equiv \widehat{d-1}(\Delta X) + 1$ along with the test statistic (2.15). One asterisk denotes significance at 5% level and two asterisks denote significance at 1% level.

			sTAaa	sTAa	sTA	sTBaa	sAaaAa	sAaaA	sAaaBaa	sAaA	sAaBaa	sABaa
	$m = 23$	$\hat{d}(\Delta X)$	1.3271	1.2381	0.99842	0.89355	1.3041	0.96375	1.0554	0.80029	0.94746	0.70537
		$\tau_{d=1}$	1.9769	1.4394	-0.0095709	-0.64347	1.8381	-0.21911	0.33513	-1.2072	-0.31758	-1.7809
	$m = 51$	$\hat{d}(\Delta X)$	1.1777	1.0788	0.88128	0.95244	1.0411	0.96008	1.002	0.85356	0.93172	0.88863
		$\tau_{d=1}$	1.7446	0.77363	-1.1655	-0.46688	0.40361	-0.39187	0.019709	-1.4376	-0.67031	-1.0934
$J = 1$	$m = 112$	$\hat{d}(\Delta X)$	1.0392	1.0195	0.99114	0.96512	0.97944	1.022	0.96871	0.95967	0.93177	0.87603
		$\tau_{d=1}$	0.60025	0.29886	-0.13579	-0.53432	-0.31491	0.33651	-0.47934	-0.61787	-1.0452	-1.8991
	$m = 246$	$\hat{d}(\Delta X)$	1.0136	1.0084	0.99483	0.98757	0.99336	1.0938	1.0029	1.0487	0.97516	0.93288
		$\tau_{d=1}$	0.31741	0.19508	-0.12067	-0.29032	-0.15505	2.1925*	0.067027	1.1367	-0.58031	-1.5681
	$m = 541$	$\hat{d}(\Delta X)$	0.91863	0.90536	0.90688	1.0478	0.94185	1.054	1.0762	1.0442	1.0736	1.0164
		$\tau_{d=1}$	-2.8344**	-3.2966**	-3.2434**	1.6645	-2.0254*	1.8825	2.6549**	1.5408	2.5629*	0.57166
	$m = 22$	$\hat{d}(\Delta X)$	1.2826	1.2477	1.0005	0.9065	1.3904	0.98571	1.0535	0.76995	0.91197	0.72771
		$\tau_{d=1}$	1.0385	0.91019	0.0018814	-0.34352	1.4345	-0.052492	0.1964	-0.84526	-0.32345	-1.0004
	$m = 50$	$\hat{d}(\Delta X)$	1.1168	1.017	0.86575	0.94063	1.0378	0.9601	1.0092	0.84993	0.95273	0.91688
		$\tau_{d=1}$	0.74401	0.10859	-0.85545	-0.37829	0.24091	-0.25422	0.058911	-0.9563	-0.30124	-0.52968
$J = 2$	$m = 112$	$\hat{d}(\Delta X)$	1.0106	0.99182	0.98783	0.95452	0.9718	1.0009	0.95521	0.95122	0.93203	0.852
		$\tau_{d=1}$	0.11014	-0.084713	-0.12603	-0.47115	-0.29217	0.0088505	-0.46396	-0.50532	-0.70419	-1.5332
	$m = 246$	$\hat{d}(\Delta X)$	1.0048	1.0096	0.99889	0.97406	0.99031	1.0843	0.9677	1.044	0.94718	0.90705
		$\tau_{d=1}$	0.07735	0.15401	-0.017772	-0.41715	-0.15576	1.3556	-0.51949	0.70838	-0.8495	-1.4947
	$m = 540$	$\hat{d}(\Delta X)$	0.9171	0.91342	0.91182	1.0423	0.93293	1.0575	1.0674	1.0441	1.0668	1.0077
		$\tau_{d=1}$	-2.0073*	-2.0964*	-2.1351*	1.0241	-1.6239	1.3935	1.6321	1.0682	1.6175	0.1855
	$m = 21$	$\hat{d}(\Delta X)$	1.351	1.3885	1.0363	0.9003	1.4829	1.0083	1.0846	0.73547	0.92579	0.7267
		$\tau_{d=1}$	0.91821	1.0161	0.094966	-0.26078	1.263	0.02161	0.22132	-0.69194	-0.1941	-0.71486
	$m = 51$	$\hat{d}(\Delta X)$	1.1372	1.0313	0.85755	0.91093	1.0163	0.94668	1.0043	0.8557	0.93772	0.89994
		$\tau_{d=1}$	0.68098	0.1551	-0.70693	-0.44203	0.08068	-0.26459	0.021183	-0.71608	-0.30905	-0.49655
$J = 3$	$m = 111$	$\hat{d}(\Delta X)$	0.99048	0.97958	0.98208	0.98401	0.94119	0.99573	0.96683	0.9722	0.94592	0.87009
		$\tau_{d=1}$	-0.077231	-0.16561	-0.1453	-0.12969	-0.47695	-0.034667	-0.26901	-0.22548	-0.43859	-1.0535
	$m = 246$	$\hat{d}(\Delta X)$	0.99393	1.007	0.9989	0.98782	0.9925	1.0923	0.967	1.0544	0.95059	0.9106
		$\tau_{d=1}$	-0.077906	0.090217	-0.014102	-0.15634	-0.096248	1.1846	-0.42368	0.69834	-0.63448	-1.1479
	$m = 540$	$\hat{d}(\Delta X)$	0.91317	0.91343	0.90697	1.0405	0.93741	1.0555	1.0629	1.0413	1.0647	1.0066
		$\tau_{d=1}$	-1.6935	-1.6884	-1.8144	0.79034	-1.2207	1.082	1.2266	0.80494	1.2625	0.12917

Table 2.44: d estimates for the yields with $l = 1$ and tapered data. For every series X, this table report the estimates $\hat{d}(\Delta X) \equiv \widehat{d-1}(\Delta X) + 1$ along with the test statistic (2.15). One asterisk denotes significance at 5% level and two asterisks denote significance at 1% level.

			T	Aaa	Aa	A	Baa
	$m = 23$	$\hat{d}(\Delta X)$	1.3479	1.5424	1.4008	1.0428	1.0549
		$\tau_{d=1}$	1.6949	2.6425	1.9528	0.2083	0.26751
	$m = 51$	$\hat{d}(\Delta X)$	1.1696	1.1666	1.0842	0.87385	0.9261
		$\tau_{d=1}$	1.4612	1.4357	0.72581	-1.0869	-0.63671
$J = 1$	$m = 112$	$\hat{d}(\Delta X)$	1.1392	1.0438	1.033	0.98458	0.97932
		$\tau_{d=1}$	1.9701	0.61991	0.46705	-0.21816	-0.29257
	$m = 246$	$\hat{d}(\Delta X)$	1.0321	1.0147	1.0301	1.0238	1.0129
		$\tau_{d=1}$	0.71463	0.32735	0.67021	0.53023	0.28768
	$m = 541$	$\hat{d}(\Delta X)$	1.0114	0.9783	0.97938	0.9383	1.0228
		$\tau_{d=1}$	0.38479	-0.7344	-0.69781	-2.0881*	0.77246
	$m = 23$	$\hat{d}(\Delta X)$	1.3497	1.5899	1.4199	1.0796	0.97171
		$\tau_{d=1}$	1.1079	1.8688	1.3302	0.25232	-0.089613
	$m = 51$	$\hat{d}(\Delta X)$	1.157	1.1578	1.0713	0.89413	0.92349
		$\tau_{d=1}$	0.90837	0.91311	0.41277	-0.61257	-0.4427
$J = 2$	$m = 111$	$\hat{d}(\Delta X)$	1.1032	1.0203	1.0108	0.94338	1.0037
		$\tau_{d=1}$	0.99439	0.19572	0.10443	-0.5457	0.0359
	$m = 245$	$\hat{d}(\Delta X)$	1.0205	0.98666	1.0077	1.0182	1.0123
		$\tau_{d=1}$	0.31548	-0.20551	0.11819	0.2798	0.19018
	$m = 541$	$\hat{d}(\Delta X)$	1.0127	0.95111	0.95982	0.93412	1.0132
		$\tau_{d=1}$	0.30058	-1.156	-0.95	-1.5577	0.31315
	$m = 22$	$\hat{d}(\Delta X)$	1.2746	1.4167	1.3819	1.1234	1.1013
		$\tau_{d=1}$	0.6318	0.9588	0.87866	0.28399	0.23303
	$m = 49$	$\hat{d}(\Delta X)$	1.114	1.1109	1.0156	0.84867	0.93783
		$\tau_{d=1}$	0.49898	0.48517	0.068277	-0.66234	-0.2721
$J = 3$	$m = 112$	$\hat{d}(\Delta X)$	1.1237	1.0571	1.0342	0.97818	0.99098
		$\tau_{d=1}$	0.95034	0.43902	0.26256	-0.16771	-0.069297
	$m = 244$	$\hat{d}(\Delta X)$	1.0343	1.0225	1.0374	1.0326	1.0211
		$\tau_{d=1}$	0.42223	0.27695	0.46133	0.40197	0.25979
	$m = 541$	$\hat{d}(\Delta X)$	1.0251	0.98493	0.99033	0.95028	1.0297
		$\tau_{d=1}$	0.47982	-0.28764	-0.18455	-0.94936	0.56727

Table 2.45: d estimates for the spreads with $l = 1$ and tapered data. For every series X, this table report the estimates $\hat{d}(\Delta X) \equiv \widehat{d-1}(\Delta X) + 1$ along with the test statistic (2.15). One asterisk denotes significance at 5% level and two asterisks denote significance at 1% level.

			sTAaa	sTAa	sTA	sTBaa	sAaaAa	sAaaA	sAaaBaa	sAaA	sAaBaa	sABaa
	$m = 23$	$\hat{d}(\Delta X)$	1.3007	1.1786	0.83758	0.822	1.356	0.87772	1.0732	0.66035	0.95341	0.82651
		$\tau_{d=1}$	1.4651	0.87002	-0.79135	-0.86724	1.7345	-0.59579	0.35673	-1.6548	-0.227	-0.84528
	$m = 51$	$\hat{d}(\Delta X)$	1.1464	1.0314	0.79479	0.92742	1.0143	0.91852	0.99064	0.80057	0.9234	0.96588
		t	1.2613	0.27034	-1.7681	-0.62533	0.12354	-0.70198	-0.080601	-1.7182	-0.65996	-0.294
$J = 1$	$m = 112$	$\hat{d}(\Delta X)$	1.0089	0.99073	0.96517	0.95466	0.96	1.0089	0.95893	0.94687	0.92718	0.90929
		$\tau_{d=1}$	0.12555	-0.13119	-0.49282	-0.64151	-0.56607	0.12565	-0.58109	-0.75179	-1.0304	-1.2835
	$m = 246$	$\hat{d}(\Delta X)$	0.99685	0.99345	0.98227	0.98463	0.98508	1.0915	1.001	1.0476	0.97641	0.95388
		$\tau_{d=1}$	-0.070165	-0.14595	-0.39497	-0.3423	-0.3322	2.0383*	0.022721	1.0609	-0.5253	-1.0271
	$m = 541$	$\hat{d}(\Delta X)$	0.90638	0.89358	0.89686	1.0488	0.93568	1.0511	1.0783	1.0433	1.0783	1.0303
		$\tau_{d=1}$	-3.1686**	-3.6018**	-3.4908**	1.6503	-2.1769*	1.7283	2.6515**	1.466	2.6506**	1.0258
	$m = 23$	$\hat{d}(\Delta X)$	1.3008	1.1765	0.77997	0.78599	1.302	0.89014	1.0532	0.66062	0.91643	0.81789
		$\tau_{d=1}$	0.95306	0.55917	-0.69708	-0.67801	0.95684	-0.34804	0.16865	-1.0752	-0.26475	-0.57695
	$m = 51$	$\hat{d}(\Delta X)$	1.1321	1.0273	0.81764	0.9077	1.001	0.92185	0.98761	0.82383	0.92797	0.96754
		$\tau_{d=1}$	0.76435	0.1579	-1.0551	-0.53407	0.0056384	-0.45218	-0.071701	-1.0193	-0.41677	-0.1878
$J = 2$	$m = 111$	$\hat{d}(\Delta X)$	1.0024	0.9762	0.94833	0.97485	0.94399	1.0073	0.96285	0.94568	0.9403	0.90483
		$\tau_{d=1}$	0.023354	-0.22932	-0.49795	-0.24234	-0.53978	0.069986	-0.35799	-0.52348	-0.57532	-0.91718
	$m = 245$	$\hat{d}(\Delta X)$	0.99183	0.99854	0.98705	0.97881	0.98026	1.0904	0.95546	1.0501	0.94464	0.91985
		$\tau_{d=1}$	-0.12589	-0.022423	-0.19945	-0.32653	-0.30409	1.3921	-0.68612	0.77173	-0.8528	-1.2348
	$m = 541$	$\hat{d}(\Delta X)$	0.89748	0.89466	0.89256	1.0374	0.93872	1.0463	1.062	1.0406	1.0719	1.0224
		$\tau_{d=1}$	-2.4241*	-2.4907*	-2.5404*	0.88403	-1.4489	1.0958	1.467	0.95942	1.7003	0.53052
	$m = 22$	$\hat{d}(\Delta X)$	1.2709	1.1372	0.81315	0.84481	1.4099	0.92604	1.0592	0.61262	0.89641	0.772
		$\tau_{d=1}$	0.62334	0.31562	-0.4299	-0.35706	0.94305	-0.17016	0.13614	-0.89125	-0.23834	-0.52456
	$m = 49$	$\hat{d}(\Delta X)$	1.1107	0.99025	0.75007	0.9255	0.99361	0.94636	1.01	0.80781	0.95436	0.99143
		$\tau_{d=1}$	0.48443	-0.042662	-1.0939	-0.32607	-0.027978	-0.23478	0.043688	-0.84117	-0.19977	-0.0375
$J = 3$	$m = 112$	$\hat{d}(\Delta X)$	1.0276	0.98392	0.96025	0.96525	0.94343	1.0237	0.95631	0.94652	0.93411	0.87899
		$\tau_{d=1}$	0.21243	-0.12353	-0.30546	-0.26703	-0.43473	0.18178	-0.33577	-0.41095	-0.50634	-0.92993
	$m = 244$	$\hat{d}(\Delta X)$	1.0033	1.0084	0.99464	0.98245	0.96651	1.0981	0.96678	1.0598	0.95237	0.92317
		$\tau_{d=1}$	0.040185	0.10374	-0.06602	-0.21631	-0.41279	1.2094	-0.40947	0.73702	-0.58703	-0.94692
	$m = 541$	$\hat{d}(\Delta X)$	0.90561	0.90769	0.89949	1.0443	0.91663	1.0469	1.0719	1.0481	1.0725	1.0215
		$\tau_{d=1}$	-1.8021	-1.7624	-1.919	0.84512	-1.5917	0.89498	1.3727	0.91814	1.3843	0.41125

Table 2.46: Local Whittle estimates for the yields. For every series X, this table report the estimates $\tilde{d}(X)$ and $\tilde{d}(\Delta X) \equiv \widetilde{d-1}(\Delta X) + 1$ along with the test statistics (2.32) and (2.33) respectively. One asterisk denotes significance at 5% level and two asterisks denote significance at 1% level.

		T	Aaa	Aa	A	Baa
$m = 23$	$\tilde{d}(X)$	1.0246	1.1363	1.1138	1.0436	0.94433
	$t_{d=1}$	0.23583	1.307	1.0918	0.41784	-0.53401
$m = 51$	$\tilde{d}(X)$	1.0562	1.0969	1.0531	1.0012	0.95645
	$t_{d=1}$	0.80258	1.3841	0.759	0.017322	-0.62198
$m = 112$	$\tilde{d}(X)$	1.035	1.0439	1.0405	1.0134	0.87084
	$t_{d=1}$	0.73995	0.92852	0.85753	0.28369	-2.7338**
$m = 246$	$\tilde{d}(X)$	0.9958	1.0184	1.0338	1.0263	0.81509
	$t_{d=1}$	-0.13178	0.57583	1.0604	0.82409	-5.8004**
$m = 541$	$\tilde{d}(X)$	0.98032	0.96551	0.97165	0.952	0.93213
	$t_{d=1}$	-0.9156	-1.6044	-1.319	-2.2331*	-3.1573**
$m = 23$	$\tilde{d}(\Delta X)$	1.1145	1.1786	1.1422	1.0511	0.9164
	$\tau_{d=1}$	1.0979	1.7132	1.3638	0.48989	-0.80189
$m = 51$	$\tilde{d}(\Delta X)$	1.1035	1.086	1.0387	0.9757	0.91543
	$\tau_{d=1}$	1.4776	1.2285	0.55322	-0.34705	-1.2079
$m = 112$	$\tilde{d}(\Delta X)$	1.0497	1.041	1.0353	1.0084	0.86556
	$\tau_{d=1}$	1.0519	0.86829	0.74792	0.17851	-2.8455**
$m = 246$	$\tilde{d}(\Delta X)$	1.0202	1.0224	1.0378	1.0319	0.81136
	$\tau_{d=1}$	0.63507	0.70141	1.1865	1	-5.9175**
$m = 541$	$\tilde{d}(\Delta X)$	1.0145	0.98156	0.98845	0.96748	0.95415
	$\tau_{d=1}$	0.67399	-0.85789	-0.5375	-1.5128	-2.1328*

or $\mathbf{d} = \iota$ in eq. (2.21) for the multivariate case. Panel A reports univariate tests whereas Panel B reports multivariate tests.

From the tables is clear that for $p = 0$ all the first differenced credit spreads series are not stationary except sBaa, sAaabaa, sAaA and sAaBaa. Further, for the series sAaaAa the null of stationarity after first differencing is rejected for the series also when $p = 1$. Among the yields, the null is rejected also for the time series T, result confirmed also by the fact that the fractional differencing parameter appears to be bigger than one from the GPH estimation.

Table 2.50 reports the results of some unit root and stationarity tests for the residuals from the cointegrating regression for all possible bivariate systems of yields.

Table 2.50 shows that there is no evidence of cointegration between Treasury and corporate yields. Ambiguous results are found for the systems Aa–Baa and A–Baa and, to some extend, for the systems T–Baa and T–Baa. The explanation could be the presence of fractional cointegration.

Table 2.47: Local Whittle estimates for the spreads. For every series X, this table report the estimates $\tilde{d}(X)$ and $\tilde{d}(\Delta X) \equiv \widetilde{d-1}(\Delta X) + 1$ along with the test statistics (2.32) and (2.33) respectively. One asterisk denotes significance at 5% level and two asterisks denote significance at 1% level.

		sTAaa	sTAa	sTA	sTBaa	sAaaAa	sAaaA	sAaaBaa	sAaA	sAaBaa	sABaa
$m = 23$	$\tilde{d}(X)$	1.1717	1.1117	0.99789	0.90988	1.0813	1.0066	0.96996	0.91087	0.89467	0.72878
	$t_{d=1}$	1.6467	1.0709	-0.02021	-0.86441	0.77936	0.062928	-0.28811	-0.85494	-1.0103	-2.6015**
$m = 51$	$\tilde{d}(X)$	1.0131	0.97652	0.9152	0.9057	1.0125	0.96721	0.93819	0.90455	0.88818	0.78432
	$t_{d=1}$	0.18725	-0.33534	-1.2113	-1.3468	0.17916	-0.46828	-0.88287	-1.3634	-1.5971	-3.0805**
$m = 112$	$\tilde{d}(X)$	1.0023	0.99495	0.98792	0.86258	1.0004	1.0323	0.843	1.0045	0.79136	0.69449
	$t_{d=1}$	0.049605	-0.10684	-0.25558	-2.9085**	0.008162	0.68448	-3.3231**	0.09521	-4.4161**	-6.4665**
$m = 246$	$\tilde{d}(X)$	0.97093	0.98067	0.97815	0.79325	0.97008	1.0266	0.75687	1.0132	0.71827	0.64924
	$t_{d=1}$	-0.91182	-0.60632	-0.68533	-6.4856**	-0.9384	0.83476	-7.6267**	0.415	-8.8376**	-11.003**
$m = 541$	$\tilde{d}(X)$	0.93805	0.94879	0.93947	0.9488	0.90683	0.98603	0.93777	0.96052	0.92321	0.93546
	$t_{d=1}$	-2.8821**	-2.3823*	-2.816**	-2.3817*	-4.3341**	-0.64987	-2.8949**	-1.8364	-3.572**	-3.0023**
$m = 23$	$\tilde{d}(\Delta X)$	1.2312	1.1598	1.0153	0.90972	1.0769	1.0036	0.97134	0.91059	0.90072	0.74891
	$\tau_{d=1}$	2.2176*	1.5332	0.1463	-0.86596	0.7375	0.034742	-0.27488	-0.8576	-0.95224	-2.4084*
$m = 51$	$\tilde{d}(\Delta X)$	1.0768	1.0244	0.94109	0.9101	1.0062	0.96765	0.93845	0.90476	0.89026	0.79554
	$\tau_{d=1}$	1.0966	0.34857	-0.84145	-1.2841	0.087909	-0.46202	-0.87916	-1.3602	-1.5674	-2.9203**
$m = 112$	$\tilde{d}(\Delta X)$	1.0147	1.0112	0.98043	0.82468	1.0069	1.0307	0.84994	1.0049	0.80597	0.71468
	$\tau_{d=1}$	0.31031	0.2364	-0.41416	-3.7108**	0.14694	0.64987	-3.1762**	0.10289	-4.1068**	-6.0391
$m = 246$	$\tilde{d}(\Delta X)$	0.98308	1.0077	0.99019	0.76281	0.97788	1.0307	0.76446	1.0173	0.73106	0.66569
	$\tau_{d=1}$	-0.53073	0.24222	-0.3077	-7.4404**	-0.69377	0.96332	-7.3886**	0.54124	-8.4363**	-10.487**
$m = 541$	$\tilde{d}(\Delta X)$	0.92833	0.94244	0.9258	0.94646	0.91755	1.0022	0.9564	0.97652	0.94539	0.96025
	$\tau_{d=1}$	-3.3342**	-2.6775**	-3.4519**	-2.4908*	-3.8353**	0.10244	-2.0284*	-1.0923	-2.5402*	-1.8489

Table 2.48: Nielsen (2005) LM test for yields. For the univariate case, we set $d = 1$ in eq. (2.18). For the multivariate case, we set $\mathbf{d} = \iota$ in eq. (2.21). Panel A reports univariate tests whereas Panel B reports multivariate tests.

Panel A

		T	Aaa	Aa	A	Baa
$p=0$	LM	15.039	1.5956	1.6096	0.58635	0.084594
	pval	0.000105	0.20653	0.20454	0.44383	0.77117
$p=1$	LM	0.18465	0.57407	0.18801	0.95166	0.60887
	pval	0.66741	0.44864	0.66457	0.3293	0.43521
$p=2$	LM	0.024477	0.1242	0.010532	0.18311	0.024272
	pval	0.87568	0.72453	0.91826	0.66871	0.8762
$p=3$	LM	0.34213	0.20205	0.16713	0.019909	0.002732
	pval	0.5586	0.65308	0.68268	0.88779	0.95831
$p=4$	LM	0.31699	0.23361	0.24655	0.087709	0.040775
	pval	0.57342	0.62886	0.61952	0.76711	0.83997

Panel B

$p=0$	LM	10.155	LMK	48.357
	pval	0.001439	pval	3.00E-09
$p=1$	LM	8.9621	LMK	12.561
	pval	0.002756	pval	0.027855
$p=2$	LM	1.0671	LMK	3.6541
	pval	0.30159	pval	0.6002
$p=3$	LM	0.029004	LMK	0.99808
	pval	0.86477	pval	0.96272
$p=4$	LM	0.14536	LMK	1.5699
	pval	0.70301	pval	0.90487

Table 2.49: Nielsen (2005) LM test for spreads. For the univariate case, we set $d = 1$ in eq. (2.18). For the multivariate case, we set $\mathbf{d} = \iota$ in eq. (2.21). Panel A reports univariate tests whereas Panel B reports multivariate tests for spreads over Treasury only.

Panel A

	sAaa	sAa	sA	sBaa	sAaaAa	sAaaA	sAaaBaa	sAaA	sAaBaa	sABaa
$p = 0$	24.31339	20.0056	11.77286	2.852134	56.06828	10.53994	0.000524	1.78096	0.631179	8.051394
	8.19E-07	7.72E-06	0.000601	0.091253	6.99E-14	0.001168	0.981734	0.182031	0.426923	0.004547
$p = 1$	6.097408	3.493377	4.736794	0.958987	11.85417	0.744657	0.069806	0.003883	0.620024	3.682575
	0.013538	0.061615	0.029524	0.327442	0.000575	0.388173	0.791619	0.950312	0.431038	0.054984
$p = 2$	1.965735	0.430146	0.75116	0.022034	1.748351	0.000364	0.394406	0.00677	0.272301	0.055102
	0.160901	0.511917	0.386109	0.881996	0.186084	0.984788	0.529993	0.934424	0.601792	0.814412
$p = 3$	0.008452	0.013447	0.033195	0.000772	0.362768	0.241951	0.03548	0.229294	0.0001	0.213746
	0.926748	0.907685	0.85543	0.977834	0.546973	0.6228	0.850593	0.632048	0.99201	0.643847
$p = 4$	0.001935	0.08217	0.028961	0.207774	0.03761	0.331287	0.065499	0.033083	0.260934	0.788442
	0.964913	0.774378	0.864869	0.648517	0.846228	0.564902	0.798006	0.855671	0.609479	0.374572

Panel B

$p = 0$	LM	25.33604	LMK	28.98468
	pval	4.82E-07	pval	7.87E-06
$p = 1$	LM	12.97074	LMK	13.03932
	pval	0.000316	pval	0.011085
$p = 2$	LM	1.23618	LMK	2.509452
	pval	0.266209	pval	0.642944
$p = 3$	LM	0.227847	LMK	0.292957
	pval	0.633125	pval	0.990264
$p = 4$	LM	0.335473	LMK	1.22171
	pval	0.562454	pval	0.874511

Table 2.50: Cointegration analysis: unit root and stationarity tests for bivariate systems. Cointegration analysis is performed for all possible bivariate systems $X - Y$. For each pair of variables X and Y, the Dickey-Fuller with the constant (DF), augmented Dickey-Fuller with the constant (ADF), Phillips-Perron with the constant (PP), a two KPSS tests without trend are carried out on the estimated OLS residuals of the regression of Y on X and a constant. In the first KPSS test the Bartlett kernel with bandwidth parameter $\left[4 \left(\frac{n}{100} \right)^{1/4} \right]$ is chosen for the estimation of the long run variance. In the second test the automatic bandwidth selection procedure of Hobijn et al. (1998) is considered.

	T - Aaa	T - Aa	T- A	T - Baa	Aaa - Aa	Aaa - A	Aaa - Baa	Aa - A	Aa - Baa	A - Baa
DF	-1.7982	-1.8528	-1.9715	-2.9225*	-2.479	-1.5925	-3.1137*	-1.8512	-3.546**	-5.3516**
ADF1	-1.6988	-1.7673	-1.9387	-2.8895*	-2.2334	-1.6715	-3.104*	-1.9196	-3.5216**	-5.2946**
ADF2	-1.642	-1.7077	-1.8504	-3.1018*	-2.0687	-1.6382	-3.4195*	-1.8036	-3.8791**	-5.9872**
ADF3	-1.5731	-1.666	-1.8045	-3.095*	-2.035	-1.6134	-3.4449**	-1.752	-3.9115**	-6.1183**
ADF4	-1.5627	-1.6539	-1.7772	-3.1314*	-1.987	-1.581	-3.4707**	-1.7518	-3.9604**	-6.2317**
ADF5	-1.5363	-1.6476	-1.7814	-2.3416	-1.9766	-1.6322	-2.4037	-1.769	-2.7033	-4.4134**
PP	-1.5876	-1.6955	-1.8444	-2.6246	-2.1118	-1.669	-2.7963	-1.841	-3.1658*	-5.0413**
KPSS	3.1**	3.2061**	4.2174**	5.0121**	3.3863**	4.9844**	5.4244**	4.8348**	4.9929**	2.8503**
KPSS HFO	0.82996**	0.86036**	1.1332**	1.3654**	0.91777**	1.3418**	1.4873**	1.3098**	1.3829**	0.84248**

3

Credit Default Swaps and Constant Maturity Credit Default Swaps

3.1   Introduction

Credit derivatives are derivative securities whose payoff depends on the credit quality (measured by the credit rating of the issuer or by the yield spread of his bonds over the yield of a comparable default-free bond) of a certain issuer. The payoff is conditioned on the occurrence of a credit event. Credit events typically include bankruptcy, failure to pay, default or restructuring. Credit derivatives widely used in practice include total return swaps, spread options, and credit default swaps[1]. The credit derivatives market has seen a phenomenal growth over the last few years: the British Bankers' Association (BBA) estimates in \$20 trillion the total gross notional value of outstanding credit derivatives at the end of 2006 and predicts that at the end of 2008 the global credit derivatives market will have expanded to \$33 trillion (BBA(2006)).

Credit default swaps (CDS) are the most common type of credit derivative: they account for around a third of the credit market as at the first quarter of 2006 (BBA(2006)). They provide insurance against the risk of a default by particular company. The first party to the contract, the protection buyer, wishes to insure against the possibility of default on a bond issued by a particular company whereas the second party to the contract, the protection seller, is willing to bear the risk associated with default by the reference entity issuing the bond. Thus the buyer of the insurance obtains the right to sell a particular bond issued by the company for its par value when a credit event occurs. The buyer of the CDS makes periodic payments to the seller until the end of the life of the CDS or until a credit event occurs. The swap is then settled by either physical delivery or in cash. If the contract requires cash settlement, the protection buyer receives the difference between the par value of the reference obligation and its market value after default. If physical delivery is required by the contract, the swap buyer delivers the bonds to the seller in exchange for their par value. In the case of either physical or cash settlement, the buyer would be required to make a final payment to the seller equal to the amount accrued since the most recent premium payment. The premium (called the CDS spread) in a CDS spread contract is determined by matching the discounted cash flows of a fixed leg paid by the protection buyer and a loss leg which corresponds to the net payment made by the protection seller to the protection buyer in case of default. As discussed in more detail in Duffie (1999) and Hull and White (2000), the credit default swap spread should be very close to the credit spread of a par yield bond issued by the reference entity over the par yield risk-free rate. This can be shown using a no arbitrage argument. Buying a par yield bond and a CDS on the reference entity an investor eliminates almost all the credit risk associated with default on the bond. This means that, denoting with y the yield on a T-year par yield bond issued by a reference entity, with r the yield on a T-year par yield riskless bond, and with S the T-year CDS spread (i.e. S is the periodical premium paid by the the protection buyer), the following relationship should hold:

$$S = y - r.$$

In fact, if S is less than $y - r$, buying a corporate bond and the credit default swap and short selling a riskless

---

[1]For an overview of credit derivatives see Tavakoli (1998); Schönbucher (2003)

bond will result in an arbitrage. If S is greater than $y - r$, then an arbitrageur will find it profitable to short a corporate bond, sell the credit default swap, and buy a riskless bond[2].

The validity of the theoretical equivalence of CDS prices and credit spreads is tested in Blanco et al. (2005). They use a dataset which consists of 33 U.S. and European investment-grade firms and find that this parity relation holds on average over time for most companies, suggesting that the bond and CDS markets price credit risk equally. Deviation from parity are found only for three European firms, for which the authors find that CDS prices are substantially higher than credit spreads for long periods of time. These cases are attributed to a combination of both imperfections in the contract specification of CDSs and measurement errors in computing the credit spread. For all the other companies they find only short-lived deviations from parity in the sample. This is because CDS prices lead credit spreads in the price discovery process, meaning that the CDS market leads the bond market in determining the price of credit risk. The relationship between credit default swaps and corporate spreads is investigated also in Longstaff et al. (2005). After developing closed-form expressions for corporate bond prices and credit default swap premium within the familiar Duffie and Singleton (1997, 1999) framework, they use the information in credit default swaps to obtain direct measures of the size of the default and nondefault components in corporate spreads. In other words they investigate what proportion of corporate yield spreads is directly attributable to default risk and how much of the spread depend on other factors such as liquidity and taxes. To answer these questions they use the information in credit default swap premia to provide direct measures of the size of the default and nondefault components in corporate yield spreads. Using CDS premia for 5-year contracts and the corresponding corporate bond prices for 68 firms traded during the period March 2001–October 2002, they find that the default component represents the majority of corporate spreads, accounting for more than 50% of the total corporate spread, even for the highest-rated investment-grade firms. Also the nondefault component is found to have a significant impact on corporate spreads. In particular, the nondefault component is time varying and mean reverts rapidly and can be explained by measures of bond-specific illiquidity such as the bidask spread and the outstanding principal amount. Finally, taxes don't seem to play an important role in explaining the nondefault component.

In Norden and Weber (2004) traditional event study methodology is applied to examine whether and to what extent stock and CDS markets responded to rating announcements during the years 2000-2002. Rating announcement events are collected from the three major rating agencies (Standard & Poor's, Moody's and Fitch) to find that CDS markets anticipate rating downgrades and that anticipation starts approximately 60 to 90 days before the announcement day. This result is consistent with Hull et al. (2004), in which credit default swap changes conditional on a ratings announcement are examined. Reviews for downgrade are found to contain significant information. However this is not the case when downgrades and negative outlooks are considered. The main conclusion is that significantly positive CDS spread changes happen before negative rating events, but positive rating events are much less significant. The latter conclusion, however, even though consistent with the results of studies on the relationship between rating events and bond yields, may be influenced by the limited number of positive rating events in the sample studied.

An investigation of the US corporate credit default swap market is provided in Schneider et al. (2007). The paper, in order to explain both the cross-section and the time series of CDS premia and to accomodate simultaneous as well as individual jumps in the risk-free and credit-risky state variables, proposes a three-factor observable jump-diffusion model for the riskless short rate and a two-factor jump-diffusion model for the default intensity of an obligor. The stochastic long-run mean of the intensity, as well as the default intensity itself, are found to be extremely persistent. The number of estimated jump events in the credit-risky components per year ranges between two and fifteen. Furthermore, the number of jumps in credit spreads is bigger for lower rated obligors. The number of jumps in the risk-free term structure is estimated in approximately four per year. As far as simultaneous jumps of the risk-free and the credit-risky components are concerned, their number is found to be approximately 0.3 per year. Pan and Singleton (2008) use a full term structure of sovereign CDS spreads to derive the market-implied default intensity and also the implicit loss rate. They argue that a lognormal process for the default intensity (as opposed to a square-root process used in Longstaff et al., 2005) is capable of capturing most of the variation in the term structure of spreads.

---

[2] Some assumptions and approximations are to be made in this arbitrage argument (see Hull et al., 2004).

The data used in the paper involves three countries (Mexico, Turkey, and Korea) with different credit ratings and covers the period March 2001–August 2006. Using MLE, the risk-neutral default intensity are found to be highly persistent and significant differences are found between the parameters in the process for the default intensity under the risk-neutral and the historical measures. This implies substantial market risk premia due to unpredictable changes in the default intensity. Furthermore, it is found for the 5-year maturity that these risk premia can be explained by measures of global risk, financial market volatility and macroeconomic policy, such as the CBOE VIX volatility index, the spread between the US Industrial 10-year BBB Yield and the 6-month Treasury bill yield and the own-country implied currency option volatility. Berndt et al. (2005) assume a lognormal model for the default intensity using a dataset of corporate CDS spreads from three sectors (broadcasting and entertainment, healthcare, and oil and gas) for the period 2000–2004. The main conclusions of the paper are that 5-year Moody's KMV EDFs explain over 74% of the variation in 5-year CDS rates across issuers and time, and that risk premia changed dramatically over time, from peaks in the third quarter of 2002, to a significant decrease at the end of 2003. In order to derive the credit risk premia, the vector of parameters governing the evolution of the default intensity process under the actual probability measure is estimated from the EDF data, whereas the vector of parameter governing the risk-neutral intensity process is estimated from 1-year and 5-year CDS rates and from 1-year EDFs.

A new class of credit derivatives is represented by constant maturity credit default swaps (CMCDS). They are very similar to CDS, in that, in exchange for protection, the buyer pays a spread. The essential difference between a CMCDS and CDS arises in the payment leg: while in a CDS the spread is fixed, in a CMCDS the spread is floating and calculated according to an indexing mechanism. In particular the spread is set equal to the prevailing reference CDS spread at each reset date times a factor known as the participation rate (PR). As a consequence, in a CMCDS contract the loss leg is paired with a floating leg, where spread payments are indexed against a reference constant maturity CDS spread at each reset date. Floating cash flows are linked to a constant-maturity term that goes under the name of *constant maturity tenor*. The reference constant maturity CDS spread does not have to have the same nominal maturity as the maturity of the contract itself. Hence one could trade a 5-year CMCDS referenced by the 3-year or 7-year CDS spread. The main attraction of CMCDS is that they offer investors access to floating credit spreads. Similar to constant maturity swaps (CMS) in fixed income market, they allow investors to take curve views and when they are combined with a CDS position, they give the investor the capability to express views on credit spreads with no default risk exposure. As a matter of fact, a short CMCDS long CDS position allows investors to isolate spread risk (i.e. the risk of changes in the premium not related to an actual credit event) and to hedge default risk. Also, CMCDS are useful for protection sellers to hedge against spread widening risk. Recall that in a CDS contract the protection seller is bound to receive a constant premium until maturity or until the occurrence of a credit event. However, if before maturity CDS spreads increase, implying that protection has become more expensive, he or she still receives the spread agreed upon at the beginning of the contract, even though the market conditions are changed. If the protection seller is worried by such a possibility, a reasonable solution could be a short position in a CMCDS contract. The investor could also benefit from the fact that the CMCDS contract has a lower mark to market than a similar CDS contract when spreads widen.

The scope of this chapter is to identify possible imbalances that may exist in the credit markets when pairing CDS and CMCDS on the same name. The general idea is to form a swap type of trading strategy whereby a fixed premium payment is netted against a floating one, both representing protection premia against default. This strategy has the advantage that default risk is eliminated and only counterparty risk is taken. A large database of single-name CDS premia is then used to produce the corresponding CMCDS prices using common market models. In doing this we use the full term structure of CDS spreads to infer default information as in Pan and Singleton (2008). In other words, single-name CDS spread data is used to replicate as much as possible the *would have been* CMCDS spreads and to investigate the possible paired trading strategy of going long CDS and shorting CMCDS and analyse the profit-loss profile across the database. The applied literature on CDS so far has focused either on issues like the validity of the theoretical equivalence of CDS prices and credit spreads or the determinants of credit default swap changes. In this paper, instead, using a dataset of CDS spreads that is large both in terms of the cross section of obligors included and in terms of the period covered, we try to identify, by the means of a statistical arbitrage analysis, trading strategies which employ CDS and CMCDS.

3.2   Pricing Credit Default Swaps

For the evaluation of a CDS contract when the payoff is contingent on default by a single reference entity in what follows we present the methodology of Hull and White (2000).

Let θ_t be the risk neutral default probability density[3] at time t, so that the probability of default in $[0, T]$ is $\int_0^T \theta_t \mathrm{d}t$. Clearly, the probability, π_t that no credit event occurs up to time t is

$$\pi_t = 1 - \int_0^t \theta_u \mathrm{d}u.$$

Consider a CDS contract with maturity T and periodic premium S to be paid at $s_1 < s_2 < \ldots < s_n = T$. Then the expected present value of the payments made by the protection buyer for a unit notional principal is

$$S \left[\sum_{i=1}^n (s_i - s_{i-1}) \pi_{s_i} \mathrm{e}^{-rs_i} + \int_0^T a_u \mathrm{e}^{-ru} \theta_u \mathrm{d}u \right], \tag{3.1}$$

where a_u is the accrual payment at time u. The first term is the discounted present value of the expected payments made at time s_i, provided the reference entity survives until s_i. The second term represents the present value of the accrual payments. Denoting by R the recovery rate, the expected present value of the payoff received by the protection buyer is equal to:

$$(1 - R) \int_0^T \mathrm{e}^{-ru} \theta_u \mathrm{d}u. \tag{3.2}$$

To find the periodic premium to be paid by the buyer of the CDS, simply equal (3.1) and (3.2) to get

$$S = \frac{(1 - R) \int_0^T \mathrm{e}^{-ru} \theta_u \mathrm{d}u}{\sum_{i=1}^n (s_i - s_{i-1}) \pi_{s_i} \mathrm{e}^{-rs_i} + \int_0^T a_u \mathrm{e}^{-ru} \theta_u \mathrm{d}u}. \tag{3.3}$$

It is also interesting to present the pricing framework considered in Longstaff et al. (2005). Let r_t denote the riskless rate, λ_t the intensity of the Poisson process governing default, and γ_t a convenience yield or liquidity process that will be used to capture the extra return investors may require, above and beyond compensation for credit risk, from holding corporate rather than riskless securities. Again, to simplify the model, each of the processes r_t, λ_t, and γ_t is stochastic but they evolve independently of each other. Moreover, bondholders recover a fraction $1 - q$ of the par value of the bond in the event of default. Consider a corporate bond with coupon rate c which pays coupon continuously. In this model its price in t is given by the sum of the present value of the coupons promised by the bond, the present value of the promised principal payment, and the present value of recovery payments in the event of a default:

$$\overline{CB}(t, T) = \mathbb{E}_t^Q \left[c \int_t^T \exp \left(- \int_t^v (r_u + \lambda_u + \gamma_u) \mathrm{d}u \right) \mathrm{d}v \right]$$

$$+ \mathbb{E}_t^Q \left[\exp \left(- \int_t^T (r_u + \lambda_u + \gamma_u) \mathrm{d}u \right) \right]$$

$$+ \mathbb{E}_t^Q \left[(1 - q) \int_t^T \lambda_v \exp \left(- \int_t^v (r_u + \lambda_u + \gamma_u) \mathrm{d}u \right) \mathrm{d}v \right].$$

---

[3]θ_t is not the same as the hazard rate λ_t. In fact $\theta_t \Delta t$ is the probability of default between t and $t + \Delta t$ as seen at time zero. Thus θ_t and λ_t are related by

$$\theta_t = \lambda_t \mathrm{e}^{- \int_0^t \lambda_u \mathrm{d}u}.$$

Observe that in each term of the sum, corporate cash flows are discounted at the adjusted discount rate $r_t + \lambda_t + \gamma_t$. Now, since the contractual[4] nature of credit default swaps makes them far less sensitive to liquidity or convenience yield effects, it is reasonable to assume that the convenience yield or illiquidity process γ_t is applicable to the cash flows from corporate bonds, but not to cash flows from credit default swap contracts. With this in mind and assuming that the premium is paid continuously, since $e^{-\int_t^v \lambda_u du}$ is the probability of no default up to time v conditional to no default up to time t, the present value of the payments made by the protection buyer in a credit default swap contract equals

$$\mathbb{E}_t^Q \left[S \int_t^T \exp\left(-\int_t^v (r_u + \lambda_u) du\right) dv \right].$$

As $\lambda_v e^{-\int_t^v \lambda_u du}$ is the probability of default between v and $v + dv$ as seen at time t, the present value of the expected payoff from the CDS is

$$\mathbb{E}_t^Q \left[\int_t^T q\lambda_v \exp\left(-\int_t^v (r_u + \lambda_u) du\right) dv \right].$$

Therefore the premium is

$$S = \frac{\mathbb{E}_t^Q \left[q \int_t^T \lambda_v \exp\left(-\int_t^v (r_u + \lambda_u) du\right) dv \right]}{\mathbb{E}_t^Q \left[\int_t^T \exp\left(-\int_t^v (r_u + \lambda_u) du\right) dv \right]} \tag{3.4}$$

and can be interpreted as a present-value-weighted average of $q\lambda_v$ where the weights are given by

$$\exp\left(-\int_t^v (r_u + \lambda_u) du\right).$$

A more formal framework for valuation of single-name credit derivatives (including credit default swaps and swaptions) is given in Jamshidian (2004) in which the general subfiltration approach of Jeanblanc and Rutkowski (2000) to modelling default risk, which includes the Cox-process setting of Lando (1998), is integrated with a numéraire invariant approach[5].

The model presented so far is quite elegant but in order to be applied one has to approximate somehow the integrals in (3.3). In what follows we describe how many banks price CDS in practice[6].

Consider again a CDS contract starting at time 0 with maturity T and periodic premium $S(0, T)$ to be paid at $s_1 < s_2 < \ldots < s_n = T$. Suppose we have a Libor discount curve $\{DF(t)\}_{t \geq 0}$ and the sequence of survival probabilities $\{SP(s_i)\}_{i=0,1,\ldots}$.

The loss leg from the CDS is

$$(1 - R) \int_0^T DF(t) SP(t) \lambda(t) dt$$

where R is the recovery rate and λ is the default intensity.

Setting $q = 1 - R$, the integral can be approximated by

$$q \sum_{j=1} DF(\tau_j) P(\text{default between } \tau_{j-1} \text{ and } \tau_j)$$

$$= q \sum_{j=1} DF(\tau_j) [SP(\tau_{j-1}) - SP(\tau_j)]$$

---

[4]Credit default swap are contracts, not securities.

[5]The increasing expanding literature on credit default swaps pricing include also Ben Ameur et al. (2006); Bielecki et al. (2005); Chu and Kwok (2003); O'Kane and Turnbull (2003)

[6]The widely employed discretized valuation formula can be found in standard textbooks like Lando (2004, chap. 8) or Schönbucher (2003, chap. 3), Arvanitis and Gregory (2001) and Cherubini et al. (2004)

where $\{\tau_1, \tau_2, \ldots\}$ is a partition (fine enough) of the interval $(0, T)$. Typically using a monthly grid results in a good approximation.

The payment leg is given by

$$S(0,T)\left(\sum_{i=1}^{n}\Delta(s_i, s_{i-1})DF(s_i)SP(s_i) + \sum_{i=1}^{n}\int_{s_{i-1}}^{s_i}\Delta(s, s_{i-1})DF(s)SP(s)\lambda(s)\mathrm{d}s\right)$$

where the second term is the accrual payment between s_{i-1} and s_i and Δ denotes the time accrual exposure with an Actual/360 convention. If we assume that the default will arrive on average in the middle of the interval, the second term can be approximated with

$$S(0,T)\sum_{i=1}^{n}\Delta(s_i, s_{i-1})DF(s_i)\frac{1}{2}[SP(s_{i-1}) - SP(s_i)].$$

Thus the fixed payment leg can be calculated as

$$S(0,T)\sum_{i=1}^{n}\Delta(s_i, s_{i-1})DF(s_i)\frac{1}{2}[SP(s_{i-1}) + SP(s_i)].$$

The credit default swap expressed in basis point at time 0 is easily calculated by equating loss and payment legs:

$$S(0,T) = \frac{q\sum_{j=1}DF(\tau_j)[SP(\tau_{j-1}) - SP(\tau_j)]}{\sum_{i=1}^{n}\Delta(s_i, s_{i-1})DF(s_i)\frac{1}{2}[SP(s_{i-1}) + SP(s_i)]}.$$

3.3   Pricing Constant Maturity Credit Default Swaps

In this section we discuss how to price CMCDS in the framework presented in the last part of section 3.2. A more formal pricing framework can be found in Brigo (2005), in which an approximated no-arbitrage market valuation formula for CMCDS is derived. The formula for CMCDS derived in Brigo (2005) is the analogous of the formula for constant maturity swaps in the default free swap market under the Libor market model. Closed-form solutions for Constant Maturity Credit Default Swaps, as well as Credit Default Swaps and Credit Default Swaptions, are derived also in Krekel and Wenzel (2006), where a Libor market model with default risk is used. Further details on CMCDS pricing can be found in Rajan et al. (2007) and in Brigo and Mercurio (2006).

To derive the participation rate, we simply exploit the fact that since the loss leg from a CMCDS is identical to the loss leg from a CDS on the same obligor and same maturity, the fixed payment legs must be identical too. Hence, when the reference CDS has maturity m,

$$\mathrm{PR}\sum_{i=1}^{n}\mathbb{E}_0[S(s_{i-1}, s_{i-1} + m)]\Delta(s_i, s_{i-1})DF(s_i)\frac{1}{2}[SP(s_{i-1}) + SP(s_i)]$$

$$= S(0,T)\sum_{i=1}^{n}\Delta(s_i, s_{i-1})DF(s_i)\frac{1}{2}[SP(s_{i-1}) + SP(s_i)].$$

Therefore, denoting by $d(s_i) = \Delta(s_i, s_{i-1})DF(s_i)$ we find

$$\mathrm{PR} = \frac{S(0,T)\sum_{i=1}^{n}d(s_i)[SP(s_{i-1}) + SP(s_i)]}{\sum_{i=1}^{n}\mathbb{E}_0[S(s_{i-1}, s_{i-1} + m)]d(s_i)[SP(s_{i-1}) + SP(s_i)]}. \tag{3.5}$$

The next step is to evaluate the expected value of future spreads in the denominator. It is clear that, when spreads evolve in a completely deterministic setting, future realised spreads will be completely determined from today's spread curve and thus the expected value equals the corresponding forward spread. However for high volatility names or long maturities a convexity adjustment is required.

3.3.1   The Forward CDS Spread

A long position in a forward default swap gives a credit protection that is active for a period of time in the future at a premium agreed upon today, but paid only during the active period of the contract.

The price for a forward contract for default protection during the time period $(t, t+m)$ is derived in Berd (2003):

$$FS(t, t+m) = \frac{S(0, t+m) - \delta(t, t+m)S(0, t)}{1 - \delta(t, t+m)} \tag{3.6}$$

where

$$\delta(t, t+m) = \frac{\int_0^t \exp\left(-\int_0^v (r_u + \lambda_u)\mathrm{d}u\right)\mathrm{d}v}{\int_0^{t+m} \exp\left(-\int_0^v (r_u + \lambda_u)\mathrm{d}u\right)\mathrm{d}v} \equiv \frac{\text{RiskyPV01}(0, t)}{\text{RiskyPV01}(0, t+m)}$$

and r_t and λ_t are the risk-free rate and the hazard rate respectively.

The derivation of (3.6) is given in section A.1.

3.3.2   The Convexity Adjustment

The usual discrepancy between the realised future rate and the current forward rate is attributed to a convexity effect. This adjustment is important mainly for long maturity contracts.

It is commonly assumed that the default intensity is described by a Ornstein-Uhlenbeck (OU) process:

$$\mathrm{d}\lambda_t = (k - \alpha\lambda_t)\mathrm{d}t + \sigma\mathrm{d}B_t. \tag{3.7}$$

The solution to the SDE (3.7) is given by

$$\lambda_t = \lambda_0 \mathrm{e}^{-\alpha t} + \frac{k}{\alpha}\left(1 - \mathrm{e}^{-\alpha t}\right) + \sigma\int_0^t \mathrm{e}^{\alpha(u-t)}\mathrm{d}B_u. \tag{3.8}$$

Since as $t \to \infty$ $\mathbb{E}[\lambda_t] \to \frac{k}{\alpha}$, this quantity has to be understood as the long term mean, whereas α is the mean reversion speed. The exact discretisation of the process leads to the the Gaussian autoregressive process

$$\lambda_{t+1} = \mu + \beta\lambda_t + \epsilon_t, \tag{3.9}$$

where $\{\epsilon_t\}$ is a sequence of i.i.d. $N(0, \eta^2)$ random variables and

$$\beta = \mathrm{e}^{-\alpha} \tag{3.10a}$$

$$\mu = \frac{k}{\alpha}\left[1 - \mathrm{e}^{-\alpha}\right] \tag{3.10b}$$

$$\eta = \sigma\left[\frac{1 - \mathrm{e}^{-2\alpha}}{2\alpha}\right]^{\frac{1}{2}}. \tag{3.10c}$$

For the time series the sequence $\{\lambda_t\}_{t=1,\dots,T}$, the parameters in (3.9) can be estimated by ML yielding

$$\hat{\beta} = \frac{\text{Cov}\left[\lambda^{(1)}, \lambda^{(T)}\right]}{\text{Var}\left[\lambda^{(T)}\right]} = \frac{(T-1)\sum_{t=1}^{T-1}\lambda_t\lambda_{t+1} - \sum_{t=1}^{T-1}\lambda_t\sum_{t=1}^{T-1}\lambda_{t+1}}{(T-1)\sum_{t=1}^{T-1}\lambda_t^2 - \left(\sum_{t=1}^{T-1}\lambda_t\right)^2}$$

$$\hat{\mu} = \bar{\lambda}^{(1)} - \hat{\beta}\bar{\lambda}^{(T)} = \frac{\sum_{t=1}^{T-1}\lambda_{t+1}}{T-1} - \hat{\beta}\frac{\sum_{t=1}^{T-1}\lambda_t}{T-1}$$

$$\hat{\eta} = \sqrt{\frac{\sum_{t=1}^{T-1}\hat{\epsilon}_t^2}{T-1}}$$

with $\lambda^{(j)}$ and $\bar{\lambda}^{(j)}$ denoting the series of interest without the jth observation and its mean and $\hat{\epsilon}_t$ are the estimated residuals from (3.9). To obtain the estimates for the parameters of the original model (3.7) it is sufficient to plug this estimates in (3.10) and solve for α, k and σ.

It can be shown that under (3.7) an approximated formula for the expected value of the future spread is

$$\mathbb{E}_0[S(s_i, s_i + m)] \approx FS(s_i, s_i + m) + \frac{1}{2}\sigma^2 C_i[FS(s_i, s_i + m) - S(0, m)] \tag{3.11}$$

with

$$C_i = \frac{1 - e^{-\alpha s_i}}{k\alpha}.$$

Using the above results we can write (3.5) as

$$PR = \frac{S(0, T)}{\overline{FS}(0, T) + \frac{\sigma^2}{2}\frac{C(0, T)}{D(0, T)}} \tag{3.12}$$

where

$$D(0, T) = \sum_{i=1}^{n} d(s_i)\frac{1}{2}[SP(s_{i-1}) + SP(s_i)]$$

$$C(0, T) = \sum_{i=1}^{n} d(s_i)\frac{1}{2}[SP(s_{i-1}) + SP(s_i)]C_i[FS(s_{i-1}, s_{i-1} + m) - S(0, m)]$$

and $\overline{FS}(0, T)$ is a weighted average of the forward CDS spreads over the reset dates:

$$\overline{FS}(0, T) = \frac{\sum_{i=1}^{n} d(s_i)[SP(s_{i-1}) + SP(s_i)]FS(s_{i-1}, s_{i-1} + m)}{\sum_{i=1}^{n} d(s_i)[SP(s_{i-1}) + SP(s_i)]}.$$

Equation (3.12) suggests an approximated formula for the participation rate of a CMCDS with maturity T and with constant maturity tenor m:

$$PR = \frac{S(0, T)}{\frac{1}{n}\sum_{i=1}^{n} FS(s_{i-1}, s_{i-1} + m)}. \tag{3.13}$$

3.4   Bootstrapping Survival Probabilities

The most delicate issue related to CMCDS pricing is the derivation of the survival probabilities. In this section we present a number of techniques, both parametric and nonparametric, to infer survival probabilities from CDS market quotes.

3.4.1   Fitting the CDS Curve Using a OU Process for the Hazard Rate

It is well known that when the hazard rates are stochastic then the survival probability up to a time t is given by

$$SP(t) = \mathbb{E}_0\left[\exp\left(-\int_0^t \lambda_s ds\right)\right]. \tag{3.14}$$

When the hazard rate follows an OU process such as that in (3.8) we can calculate the expectation in closed form:

$$SP(t) = \exp[a(t) + b(t)\lambda_0] \tag{3.15}$$

with

$$a(t) = -\frac{(b(t) + t)(\alpha k - \frac{\sigma^2}{2})}{\alpha^2} - \frac{\sigma^2}{4\alpha}b(t)^2 \tag{3.16a}$$

$$b(t) = \frac{e^{-\alpha t} - 1}{\alpha}. \tag{3.16b}$$

We derive (3.15)–(3.16) in section A.2 (see also Vasicek, 1977; Luciano and Vigna, 2006).

Note that the above equation automatically satisfies the initial condition $SP(0) = 1$. There are four parameters to calibrate k, α, σ and λ_0.

This means that in order to calibrate the Vasicek model for the credit spreads we need only four points on the survival probability curve or, equivalently, four CDS spread values, for the same obligor.

Since there may be more than four values on the survival curve we propose to estimate the obligor individual parameters by minimising the residual error between the model implied SP values in (3.15) and the values obtained from the CDS market. This is done in subsection 3.4.4.

3.4.2   Piecewise Constant Hazard Rates

The survival probabilities can also be bootstrapped from

$$S(0, s_n = T) = \frac{q \sum_{i=1}^{n} DF(s_i)[SP(s_{i-1}) - SP(s_i)]}{\text{RiskyPV01}(0, s_n)} \tag{3.17}$$

considering the recovery rate $1 - q$ fixed. However, when there are less maturities for traded contracts than the entire set of time points for which survival probabilities must be calculated, this commonly used method does not work. A solution to this shortcoming can be bootstrapping the survival probabilities from the hazard rates curve as proposed by O'Kane and Turnbull (2003). This is done as follows. Suppose that for a given obligor the CDS contracts with maturities 6m, 1y, 2y, 3y, 5y, 7y and 10y are available (this is the typical situation). For $\tau = T - t_v$ the survival probability $SP(t_v, T) \equiv SP(\tau)$ is

$$SP(\tau) =$$

$$
= \begin{cases}
\exp[-\lambda_{0,0.5}\tau] & 0 < \tau \leq 0.5 \\
\exp[-0.5\lambda_{0,0.5} - \lambda_{0.5,1}(\tau - 0.5)] & 0.5 < \tau \leq 1 \\
\exp[-0.5\lambda_{0,0.5} - 0.5\lambda_{0.5,1} - \lambda_{1,2}(\tau - 1)] & 1 < \tau \leq 2 \\
\exp[-0.5\lambda_{0,0.5} - 0.5\lambda_{0.5,1} - \lambda_{1,2} - \lambda_{2,3}(\tau - 2)] & 2 < \tau \leq 3 \\
\exp[-0.5\lambda_{0,0.5} - 0.5\lambda_{0.5,1} - \lambda_{1,2} - \lambda_{2,3} - \lambda_{3,5}(\tau - 3)] & 3 < \tau \leq 5 \\
\exp[-0.5\lambda_{0,0.5} - 0.5\lambda_{0.5,1} - \lambda_{1,2} - \lambda_{2,3} - 2\lambda_{3,5} - \lambda_{5,7}(\tau - 5)] & 5 < \tau \leq 7 \\
\exp[-0.5\lambda_{0,0.5} - 0.5\lambda_{0.5,1} - \lambda_{1,2} - \lambda_{2,3} - 2\lambda_{3,5} - 2\lambda_{5,7} - \lambda_{7,10}(\tau - 7)] & \tau > 7
\end{cases}
$$

i.e. the hazard rate curve is assumed to be piecewise constant. We use a monthly time grid $0 = t_0 < t_1 < \ldots < t_n = T$ with $t_i = i/12$ for each $n \in \mathscr{M}$, $\mathscr{M}$ being the set of available maturities in months. For instance, with the above contracts we have $\mathscr{M} = \{6, 12, 24, 36, 60, 84, 120\}$.

The previous function can be easily generalized to the case in which we have CDS maturing at $T_1, \ldots, T_M$, $M > 1$. With the convention $\lambda_1 = \lambda_{0,T_1}$, $\lambda_i = \lambda_{T_{i-1},T_i}$, $i = 2, \ldots, M$, the function $SP(\tau)$ is such that

$$
-\log SP(\tau) = \lambda_1 \tau I_{[0,T_1)}(\tau) + \sum_{i=1}^{M-2} \left[\sum_{j=1}^{i} (\lambda_j - \lambda_{j+1})T_j + \lambda_{i+1}\tau \right] I_{[T_i, T_{i+1})}(\tau)
$$

$$
+ \left[\sum_{j=1}^{M-1} (\lambda_j - \lambda_{j+1})T_j + \lambda_{i+1}\tau \right] I_{[T_{M-1}, \infty)}(\tau). \tag{3.18}
$$

In the previous case $M = 7$, $T_1 = 0.5, \ldots, T_7 = 10$.

The function above can be easily implemented in Matlab as in Listing B.4.

Next, for each $n \in \mathcal{M}$, using a numerical searching algorithm, we solve iteratively the equations

$$\frac{1}{2}\frac{S(0,\frac{n}{12})}{q}\sum_{j=0}^{\frac{n}{3}-1}\Delta(t_{n-3j},t_{n-3(j+1)})DF(t_{n-3j})[SP(t_{n-3(j+1)})+SP(t_{n-3j})]$$

$$=\sum_{i=1}^{n}DF(t_i)[SP(t_{i-1})-SP(t_i)] \quad (3.19)$$

for λ_i, $i = 1,\ldots,M$. Note that the pro-rata payment due to settlement date outside quarterly market calendar is not included here. The Matlab code for this function is reported in Listing B.5.

For instance, when $T = 0.5$ ($n = 6$) we solve

$$\frac{1}{2}\frac{S(0,0.5)}{q}\sum_{j=0}^{1}\Delta(t_{n-3j},t_{n-3(j+1)})DF(t_{n-3j})[\mathrm{e}^{-2\lambda_{0,0.5}t_{n-3(j+1)}}+\mathrm{e}^{-2\lambda_{0,0.5}t_{n-3j}}]$$

$$=\sum_{i=1}^{6}DF(t_i)[\mathrm{e}^{-2\lambda_{0,0.5}t_{i-1}}-\mathrm{e}^{-2\lambda_{0,0.5}t_i}]$$

for $\lambda_{0,0.5}$.
Having done that, we move to $T = 1$ ($n = 12$) and solve for $\lambda_{0.5,1}$

$$\frac{1}{2}\frac{S(0,1)}{q}\sum_{j=0}^{1}\Delta(t_{n-3j},t_{n-3(j+1)})DF(t_{n-3j})[\mathrm{e}^{-\lambda_{0,0.5}-2\lambda_{0.5,1}(t_{n-3(j+1)}-0.5)}$$

$$+\mathrm{e}^{-\lambda_{0,0.5}-2\lambda_{0.5,1}(t_{n-3j}-0.5)}]$$

$$+\frac{1}{2}\frac{S(0,1)}{q}\sum_{j=2}^{3}\Delta(t_{n-3j},t_{n-3(j+1)})DF(t_{n-3j})[\mathrm{e}^{-2\lambda_{0,0.5}t_{n-3(j+1)}}+\mathrm{e}^{-2\lambda_{0,0.5}t_{n-3j}}]$$

$$=\sum_{i=1}^{6}DF(t_i)[\mathrm{e}^{-2\lambda_{0,0.5}t_{i-1}}-\mathrm{e}^{-2\lambda_{0,0.5}t_i}]$$

$$+\sum_{i=7}^{12}DF(t_i)[\mathrm{e}^{-\lambda_{0,0.5}-2\lambda_{0.5,1}(t_{i-1}-0.5)}-\mathrm{e}^{-\lambda_{0,0.5}-2\lambda_{0.5,1}(t_i-0.5)}]$$

Basically we use the CDS with maturity T_j to compute all the intensities up to λ_j. The Matlab code is reported in Listing B.6.

Finally we need to derive the RiskyPV01 in order to compute the forward spread. This is done taking into account a pro-rata payment that arises from the fact that $n \geq 3k = 3\left[\frac{n}{3}\right]$. The RiskyPV01 is given by the following formula

$$\mathrm{RiskyPV01}(0,t_n) = \frac{1bp}{2}\Big\{\Delta_0 \times DF(t_1) \times [1 + SP(t_1)]$$

$$+\sum_{j=0}^{k-1}\Delta(t_{n-3j},t_{n-3(j+1)})DF(t_{n-3j})[SP(t_{n-3(j+1)})+SP(t_{n-3j})]\Big\} \quad (3.20)$$

where

$$\Delta_0 = \left[\frac{t_1 - t_v}{t_1 - t_0}\right]_{\mathrm{ACT}/360}$$

and $[j/k]_{\mathrm{ACT}/360}$ is the date fraction j/k converted into an ACT/360 convention. The Matlab code can be found in Listing B.7.

Having done this it is straightforward to compute the theoretical CDS premia implied in the survival probabilities derived above (see Listing B.8).

3.4.3   Nelson-Siegel Interpolation

Many models assume that default is given by the arrival time of a Poisson process with the corresponding hazard rate constant over time.

If $SP(t)$ is the survival probability up to time t and $\lambda(\cdot)$ is the hazard rate then

$$SP(t) = \mathbb{E}_0 \left[\exp \left(- \int_0^t \lambda_s \mathrm{d}s \right) \right]$$

We assume here that the hazard rate is deterministic time-varying such that

$$\int_0^t \lambda(s)\mathrm{d}s = \Psi(t)t \tag{3.21}$$

The role of function $\Psi(t)$ is to capture any term structure variation.

One of the common choices for function $\Psi(t)$ is the Nelson-Siegel (see Nelson and Siegel, 1987) function[7]

$$\Psi(t) = \alpha_0 + (\alpha_1 + \alpha_2) \left(\frac{1 - \exp(-\frac{t}{\alpha_3})}{\frac{t}{\alpha_3}} \right) - \alpha_2 \exp \left(-\frac{t}{\alpha_3} \right) \tag{3.22}$$

This function can generate all sorts of curve shapes. The parameter α_0 is the long term mean of the default intensity. Parameter α_1 is the deviation from the mean, with $\alpha_1 > 0$ implying a downward sloping intensity and $\alpha_1 < 0$ implying an upward sloping term structure. In addition the reversion rate toward the long-term mean is negatively related to $\alpha_3 > 0$.

The parameter α_2 is responsible for generating humps when it is different than zero. In practice Bluhm et al. (2003) advocate not using humps as this may lead to overfitting problems. This means that here we shall assume that $\alpha_2 = 0$ and estimate $\alpha_0, \alpha_1, \alpha_3$ from CDS spread data. The survival function becomes

$$SP(t) = \exp \left[-\Big[\alpha_0 + \alpha_1 \left(\frac{1 - \exp(-\frac{t}{\alpha_3})}{\frac{t}{\alpha_3}} \right)\Big] t \right] \tag{3.23}$$

This function is implemented in Matlab as in Listing B.9.

The CDS pricing equation are then used to estimate the parameters from CDS spread data using a nonlinear optimization algorithm for a suitable minimization function such as sum of squared errors or sum of absolute errors.

3.4.4   Details on Parameter Estimation

OU Process

Given a set of CDS spreads with maturities $\{t_n\}_{n \in \mathcal{M}}$, in order to estimate the vector of parameters $\theta = (\lambda_0, \sigma, k, \alpha)'$, we first compute the theoretical CDS, $S(0, t_n; \theta)$ using (3.17) and then calculate[8]

$$\arg\min_\theta \sum_{n \in \mathcal{M}} [S(0, t_n) - S(0, t_n; \theta)]^2$$

or

$$\arg\min_\theta \sum_{n \in \mathcal{M}} |S(0, t_n) - S(0, t_n; \theta)|.$$

The optimization is done under the constraints

$$\theta > 0$$
$$SP'(T_M) < 0$$

---

[7]Markit Partners are using a similar approach based on Nelson-Siegel interpolation to produce theoretical credit curves in the situations where liquidity of data is very low

[8]Also the weighted objective functions of the form (3.28)–(3.29) have been considered.

where T_M is the last maturity (20yr) of the available CDS data and

$$SP'(t) = \frac{k}{\alpha}(1 - e^{-\alpha t}) + \lambda_0 e^{-\alpha t} - \frac{\sigma^2}{2\alpha^2}(1 - e^{-\alpha t})^2$$

Nelson-Siegel

Given $\boldsymbol{\alpha} = (\alpha_0, \alpha_1, \alpha_3)'$ and the parameter space $U_{\boldsymbol{\alpha}} \subset \mathbb{R}^3$, in the first case we compute

$$\tilde{\boldsymbol{\alpha}} = \arg \min_{\boldsymbol{\alpha} \in U_{\boldsymbol{\alpha}}} \sum_{n \in \mathcal{M}} [S(0, t_n) - S(0, t_n; \boldsymbol{\alpha})]^2 = \arg \min_{\boldsymbol{\alpha} \in U_{\boldsymbol{\alpha}}} \tilde{f}(\boldsymbol{\alpha})$$

and in the second

$$\breve{\boldsymbol{\alpha}} = \arg \min_{\boldsymbol{\alpha} \in U_{\boldsymbol{\alpha}}} \sum_{n \in \mathcal{M}} |S(0, t_n) - S(0, t_n; \boldsymbol{\alpha})| = \arg \min_{\boldsymbol{\alpha} \in U_{\boldsymbol{\alpha}}} \breve{f}(\boldsymbol{\alpha})$$

where $S(0, t_n; \boldsymbol{\alpha})$ denotes the theoretical CDS spread maturing at time t_n when using a Nelson-Siegel function with parameter $\boldsymbol{\alpha}$. The optimization should be done under the following constraints which identify $U_{\boldsymbol{\alpha}}$:

$$\alpha_3 \quad > \quad 0 \tag{3.24}$$
$$SP(t) - SP(t+1) \quad \geq \quad 0 \quad \forall t > 0. \tag{3.25}$$

Note that the formulation (3.23) automatically ensures $SP(0) = 1$.
Moreover condition (3.25) is equivalent to

$$\alpha_0 + \alpha_1 \exp\left(-\frac{t}{\alpha_3}\right) \geq 0 \tag{3.26}$$

which is obtained by imposing that the function $\Psi(t) \times t$ is not increasing.
However, we define $U_{\boldsymbol{\alpha}} \subset \mathbb{R}^3$ by the set of constraints:

$$\alpha_0 > 0 \tag{3.27a}$$
$$\alpha_1 \geq -\alpha_0 \exp\left(\frac{T_M}{\alpha_3}\right) \tag{3.27b}$$
$$\alpha_3 > 0 \tag{3.27c}$$

Note that (3.27a) can be obtained by letting $t \to \infty$ in (3.25), whereas condition (3.27b) implies

$$\alpha_0 + \alpha_1 \exp\left(-\frac{t}{\alpha_3}\right) \geq 0 \quad \forall t \leq T_M$$

As far as the choice of the function to be minimized, in practice we set $\hat{\boldsymbol{\alpha}} = \tilde{\boldsymbol{\alpha}}$ unless

$$\tilde{f}(\breve{\boldsymbol{\alpha}}) < \tilde{f}(\tilde{\boldsymbol{\alpha}})$$

or

$$\breve{f}(\breve{\boldsymbol{\alpha}}) < \breve{f}(\tilde{\boldsymbol{\alpha}}).$$

To reflect the fact that CDS contracts with different maturities may have different levels of liquidity (for instance, typically a 5 year contract is more traded than a 30 year contract on the same name) we use the "weighted" objective functions and compute:

$$\tilde{\boldsymbol{\alpha}}_w = \arg \min_{\boldsymbol{\alpha} \in U_{\boldsymbol{\alpha}}} \sum_{n \in \mathcal{M}} w_n [S(0, t_n) - S(0, t_n; \boldsymbol{\alpha})]^2 = \arg \min_{\boldsymbol{\alpha} \in U_{\boldsymbol{\alpha}}} \tilde{f}_w(\boldsymbol{\alpha}) \tag{3.28}$$

or

$$\breve{\boldsymbol{\alpha}}_w = \arg\min_{\boldsymbol{\alpha} \in U_{\boldsymbol{\alpha}}} \sum_{n \in \mathcal{M}} w_n \left| S(0, t_n) - S(0, t_n; \boldsymbol{\alpha}) \right| = \arg\min_{\boldsymbol{\alpha} \in U_{\boldsymbol{\alpha}}} \breve{f}_w(\boldsymbol{\alpha}). \qquad (3.29)$$

In particular, provided that the number, N, of contracts at some point time is bigger than 6 (which is the number of maturities with bigger weights) we attach to each observation the weights given by Table 3.1.

Table 3.1: Weights.

Maturity t_n (Months)	w_n
60	40%
36	30%
12	15%
84	6%
120	4%
24	3%
All the Others	$\frac{2}{N-6}$%

To implement the above we first need a code for the RiskyPV01 (Listing B.10).
This allows to compute the theoretical CDS premia $S(0, t; \hat{\boldsymbol{\alpha}})$, implemented in Listing B.11.
Now, the function to be minimized is reported in Listing B.12
Finally $\hat{\boldsymbol{\alpha}}$ is computed as in Listing B.13.

3.5   Implementation and Recap of the Algorithm

Suppose a Libor discount curve $\{DF(t)\}_{t \geq 0}$ and the sequence of survival probabilities $\{SP(t_i)\}_{i=0,1,\ldots}$ are given. The pricing time grid is described by n monthly periods given by $t_0 < t_1 < t_2 < \ldots < t_n = T$, with $t_i = i/12$ for all $i \in \{1, 2, \ldots, n\}$ and T denoting the maturity of the CMCDS contract. The contract is traded initially at time $t_v \in [t_0, t_3)$. The schedule of fixed payments is quarterly as this is the dominating market standard. The number of quarters fitting into the pricing time grid until maturity T is equal to $k = \left[\frac{n}{3}\right]$. It is evident that $k = \frac{n}{3}$ only if $t_v = t_0 \equiv 0$. The first premium is paid at time t_{n-3k+3} (which coincides with t_3 when n is a multiple of 3). A cash flow diagram is reported in Figure 3.1.

Figure 3.1: Cash flow diagram for constant maturity credit default swaps
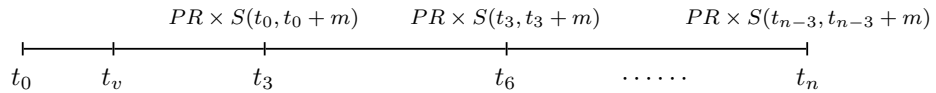
Hence the participation rate when $t_v \equiv 0$ and the reference CDS has maturity m months is

$$\text{PR} = \frac{S(0, T) \sum_{j=0}^{k-1} \Delta(t_{n-3j}, t_{n-3(j+1)}) DF(t_{n-3j})[SP(t_{n-3(j+1)}) + SP(t_{n-3j})]}{\sum_{j=0}^{k-1} \mathbb{E}_0[S(t_{n-3(j+1)}, t_{n-3(j+1)} + m)] \Delta(t_{n-3j}, t_{n-3(j+1)}) DF(t_{n-3j})[SP(t_{n-3(j+1)}) + SP(t_{n-3j})]}.$$

where the expectation, given by (3.11) depends on the forward CDS (3.6).

Table 3.2: Calculation of discount factors from USD Libor rates on 1st March 2004.

month	No days	Libor %	Discount factor
1	31	1.1	0.999053674
2	61	1.11	0.998122698
3	92	1.12	0.997145947
4	122	1.13	0.996185164
5	154	1.15	0.995136271
6	184	1.17	0.994055548
7	214	1.19375	0.99295382
8	245	1.22125	0.991757223
9	275	1.25	0.990541702
10	306	1.29	0.989153927
11	337	1.32875	0.987714241
12	365	1.3675	0.986324677

Some care is needed in calculating the first coupon. There are some differences caused by whether the current settling day is within a month of the day of the first calendar coupon day or not. Again, t_v is the current valuation day and t_0 is the calendar coupon payment day prior to t_v.

If $t_3 - t_v > 1mth$ then[9] the first coupon is equal to $S(0, T) \times \Delta(t_3, t_v)$ and it is going to be paid at t_3 so it is going to be discounted by $\frac{1}{2}DF(t_3)[1 + SP(t_3)]$.

If $t_3 - t_v \leq 1mth$ then the first payment is delayed until the next coupon date and is coupled with a full coupon for the period. Thus the first coupon is equal to $S(0, T) \times [1 + \Delta(t_3, t_v)]$ and it is going to be paid at t_6 so it is going to be discounted by $\frac{1}{2}DF(t_6)[SP(t_3) + SP(t_6)]$.

In order for this procedure to be implemented, a discount curve, the survival probability curve and the RiskyPV01 are needed as inputs.

The standard choice for the risk-free rate is the Libor rates up to one year and swap rates after one year up to 10 years. The first step is to compute the discount factors from the Libor/Swap rates.

3.5.1   Calculation of Discount Factors from Libor rates

The Libor rates over one week are spot rates. If L is the Libor rate for some maturity and there are exactly N days from today t_v until the maturity of the Libor contract then the discount factor is calculated as

$$DF = \frac{1}{1 + \frac{N}{360} \times L}. \tag{3.30}$$

This is easily implemented as in Listing B.1.

The following example (Table 3.2) shows how the discount rates are calculated from the USD Libor rates on March 1, 2004.

3.5.2   Calculation of Discount Factors from swap rates

The calculation of discount factors from swap rates is done via bootstrapping.

The swap rates are par rates so the discount factor could be bootstrapped from the term structure of swap rates. However, due to various factors such as frequency of payments semiannual or quarterly or monthly or even by the lack of market information for intermediary tenors, the system of equations is undetermined.

For exemplification consider the hypothetical market information in Table 3.3.

---

[9]Please note that Actual/360 is the day count convention and the number of days in one month may differ from month to month

Table 3.3: Hypothetical information about swap rates.

Maturity in years	swap rate %
1	3.20
2	3.65
3	4.15
5	4.55
10	4.85
15	5.00
20	5.10

Assuming that all swaps pay only annually we get the following system of equations

$$
\begin{aligned}
(100 + 3.20)DF(1) &= 100 \\
3.65DF(1) + (100 + 3.65)DF(2) &= 100 \\
4.15DF(1) + 4.15DF(2) + (100 + 4.15)DF(3) &= 100 \\
4.55DF(1) + 4.55DF(2) + 4.55DF(3) + 4.55DF(4) + (100 + 4.55)DF(5) &= 100 \\
4.85\sum_{i=1}^{10} DF(i) + 100DF(10) &= 100 \\
5.00\sum_{i=1}^{15} DF(i) + 100DF(15) &= 100 \\
5.10\sum_{i=1}^{20} DF(i) + 100DF(20) &= 100
\end{aligned}
$$

The first three equations can be solved easily. The system is clearly under-determined due to jumps in tenors. In order to determine the discount factors we must resort to constraints or assumptions.

Here are several methods used in the market to bootstrap the discount factors.

Linear Swap rates

This method is based on the assumption that each "missing" swap rate lies on a straight line between the given swap rates before and after the required tenor. So for example the 4 year swap rate would be calculated as the linear combination of the 3 year swap rate and the 5 year swap rate and it would be equal to 4.35%. Similarly the rates between 5 and 10 years would be 4.61%, 4.67%, 4.73%, 4.79%.

This method can produce occasionally unrealistic implied forward rate curves.

Constant Forward Rates

Another possibility is to assume that the one year forward rate is flat between the tenors where swap rates are missing. With our constructed example that would mean for example that the forward rate between 3 and 4 years $F_{3,4}$ is equal to the forward rate between 4 and 5 year $F_{4,5}$. Denoting for convenience this rate with F then $DF(4) = DF(3)/(1 + F)$ and $DF(5) = DF(3)/(1 + F)^2$. Replacing these into the third equation above gives

$$
4.55\frac{DF(3)}{1 + F} + 104.55\frac{DF(3)}{(1 + F)^2} = 100 - 4.55[DF(1) + DF(2) + DF(3)]
$$

and because $DF(3)$ is known this equation can be solved.

The procedure continues similarly for the other tenors.

Matlab code is reported in Listing B.2.

The discount factors corresponding to the previous example are reported in Table 3.4.

Table 3.4: Calculation of discount factors from Swap data with the Linear Swap rates (LSR) and the Constant Forward Rates (CFR) methods.

Maturity (Years)	LSR	CFR
1	0.96899	0.96899
2	0.93066	0.93066
3	0.88446	0.88446
4	0.84225	0.84051
5	0.79866	0.79874
6	0.76093	0.75899
7	0.724	0.72123
8	0.68792	0.68534
9	0.6527	0.65124
10	0.61836	0.61884
11	0.58733	0.58688
12	0.55743	0.55657
13	0.52861	0.52783
14	0.50085	0.50057
15	0.47414	0.47472
16	0.44947	0.44946
17	0.42581	0.42555
18	0.40314	0.4029
19	0.3814	0.38147
20	0.36058	0.36117

In practice the swaps pay semiannually and thus the equations we look at are of the form

$$\frac{SW_j}{2} \sum_{k=2}^{2j-1} DF(k/2) + \left(100 + \frac{SW_j}{2}\right) DF(j) = 100 - \frac{SW_j}{2} DF(1/2) \tag{3.31}$$

where SW_j denotes the swap rate for maturity j years, and j takes values in some set, typically $j \in \{1, 2, \ldots, 5, 7, 10, 20, 30\}$. Note that in (3.31) the unknowns we solve for are $DF(k/2)$, $k = 2, 3, \ldots$ and $DF(1/2)$ is treated as given, as it is recovered from the Libor data.

3.5.3 Discount Factors for Intermediary Points between Tenors

The main problem in calculating the discount factors is the calculation at intermediary points between the tenors of Libor or swap rates. Suppose we know the value of the risk-free discount factors for $0 = \mathcal{T}_0 < \mathcal{T}_1 < \ldots \mathcal{T}_r$ A very popular way to obtain the intermediary discount factors is to use log-linear interpolation. The discount factor for $t \in [\mathcal{T}_j, \mathcal{T}_{j+1}]$, $DF(t)$ will be given by

$$\log \left(DF(t)\right) = \frac{\mathcal{T}_{j+1} - t}{\mathcal{T}_{j+1} - \mathcal{T}_j} \log \left(DF(\mathcal{T}_j)\right) + \frac{t - \mathcal{T}_j}{\mathcal{T}_{j+1} - \mathcal{T}_j} \log \left(DF(\mathcal{T}_{j+1})\right).$$

Matlab code is reported in Listing B.3.

In calculating the CMCDS premium the following steps are followed daily

1. Determine the Libor-swap discount curve

2. For each name build the survival probabilities with piecewise constant hazard rates or the ones from Nelson-Siegel interpolation

3. For each name calculate the numerical values of corresponding RiskyPV01 using (3.20)

4. Calculate the entire family of CDS forward curves using (3.6)

5. Calculate the participation rate PR using

$$
\text{PR} = \frac{S(0,T)\sum_{j=0}^{k-1}\Delta(t_{n-3j},t_{n-3(j+1)})DF(t_{n-3j})[SP(t_{n-3(j+1)}) + SP(t_{n-3j})]}{\sum_{j=0}^{k-1}FS(t_{n-3(j+1)},t_{n-3(j+1)}+m)\Delta(t_{n-3j},t_{n-3(j+1)})DF(t_{n-3j})[SP(t_{n-3(j+1)}) + SP(t_{n-3j})]}.
$$

The last two points can be implemented in Matlab using the codes Listing B.14–Listing B.17.

3.5.4   *Example 1: NS+Piecewise Constant HR*

In this example we consider the firms Abitibi Consol Inc and Tesco PLC on October 3rd 2005. As it is clear from Table 3.5, we have

a. a firm with big CDS premia and all maturities available

b. a firm with small CDS premia and few maturities available

c. a firm with small CDS premia and all maturities available

Table 3.5: CDS spreads – October 3rd 2005.

Maturity (Months)	Abitibi Consol Inc	Microsoft Corp	Tesco PLC
6	0.012841		0.000613
12	0.016526		0.000689
24	0.023347	0.000450015	0.001129
36	0.030028	0.00051501	0.001465
48	0.03499		0.001893
60	0.039373	0.000508335	0.00229
84	0.04251		0.003199
120	0.045291	0.000866685	0.004254
180	0.045355		0.00442
240	0.045828	0.001065644	0.004403
360	0.046868		0.0048
q	0.60714	0.6	0.606

We report the $\boldsymbol{\alpha}$ estimates in Table 3.6.

Table 3.6: $\boldsymbol{\alpha}$ Estimates. Panel A reports $\hat{\boldsymbol{\alpha}}$ and Panel B $\hat{\boldsymbol{\alpha}}_w$.

Panel A

	Abitibi Consol Inc	Microsoft Corp	Tesco PLC
$\hat{\alpha}_0$	0.007039	0.001153	0.084241
$\hat{\alpha}_1$	-0.02284	0	-0.1783
$\hat{\alpha}_3$	0.162935	0.153387	0.099982

Panel B

	Abitibi Consol Inc	Microsoft Corp	Tesco PLC
$\hat{\alpha}_0$	0.005082	—	0.079385
$\hat{\alpha}_1$	-0.01392	—	-0.14482
$\hat{\alpha}_3$	0.141648	—	0.078568

Figure 3.2: Bootstrapped Survival Probabilities and CDS spreads for Abitibi Consol Inc.
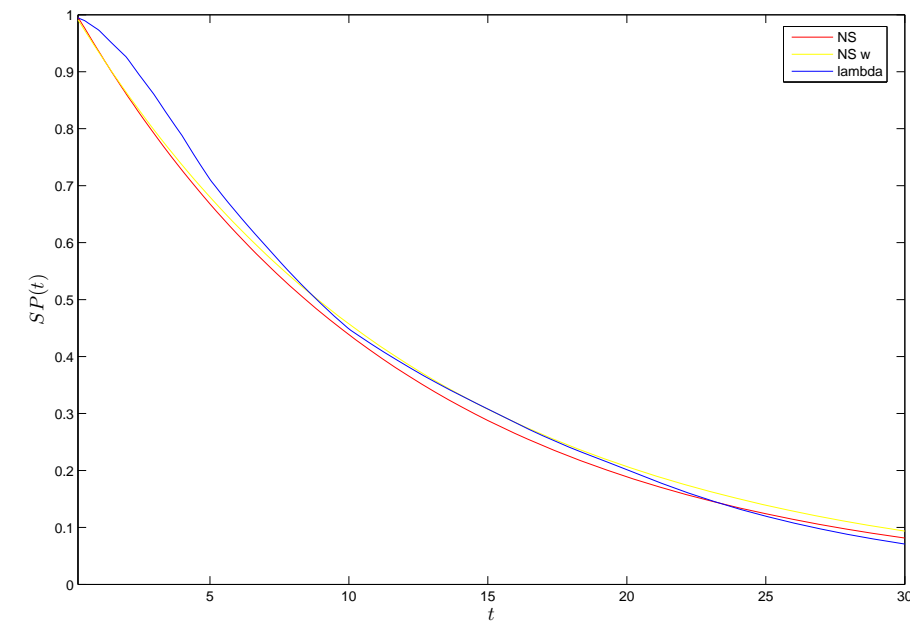
(a) Survival Probabilities
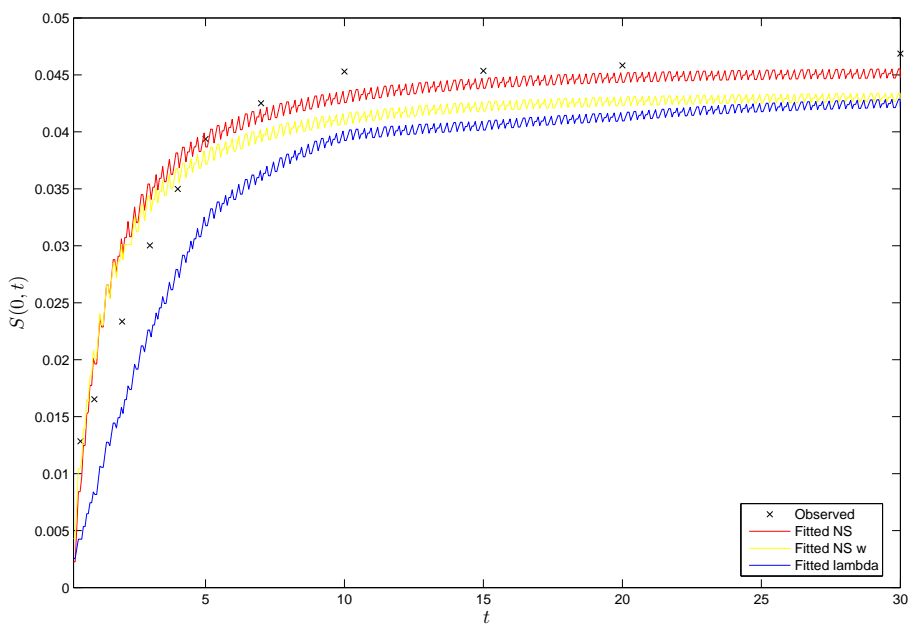
(b) CDS spreads

Figure 3.3: Bootstrapped Survival Probabilities and CDS spreads for Microsoft Corp.
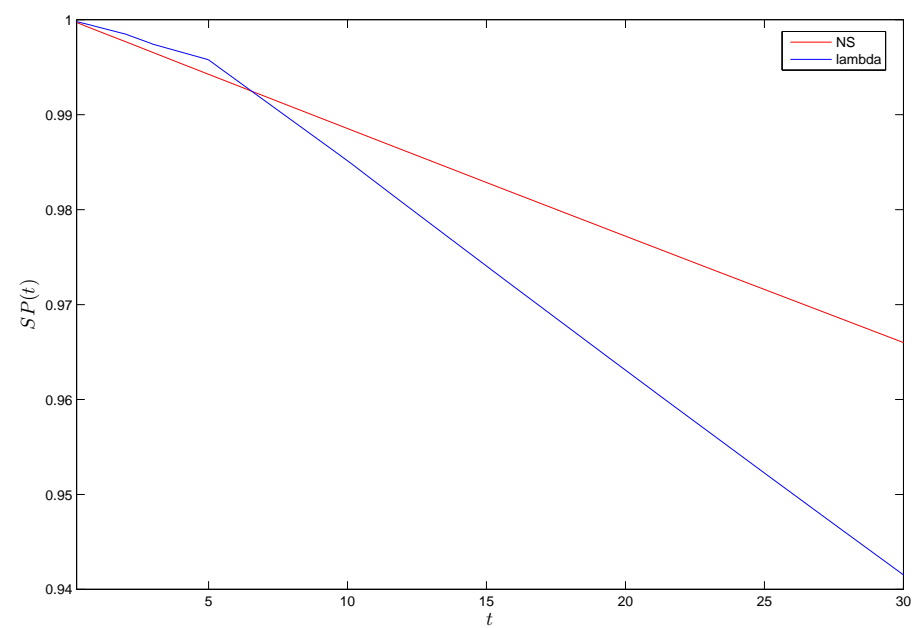
(a) Survival Probabilities
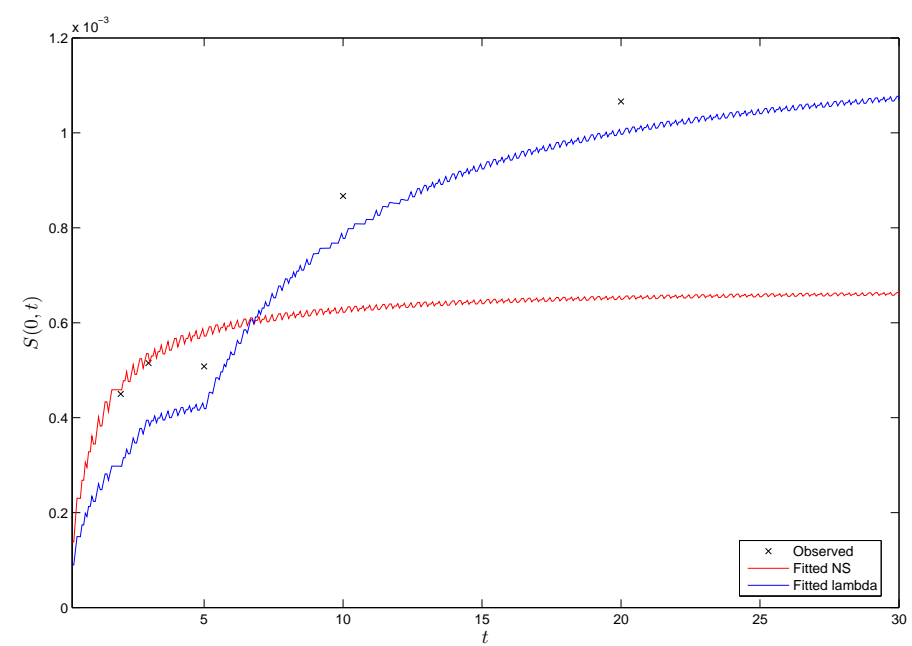
(b) CDS spreads

Figure 3.4: Bootstrapped Survival Probabilities and CDS spreads for Tesco PLC.
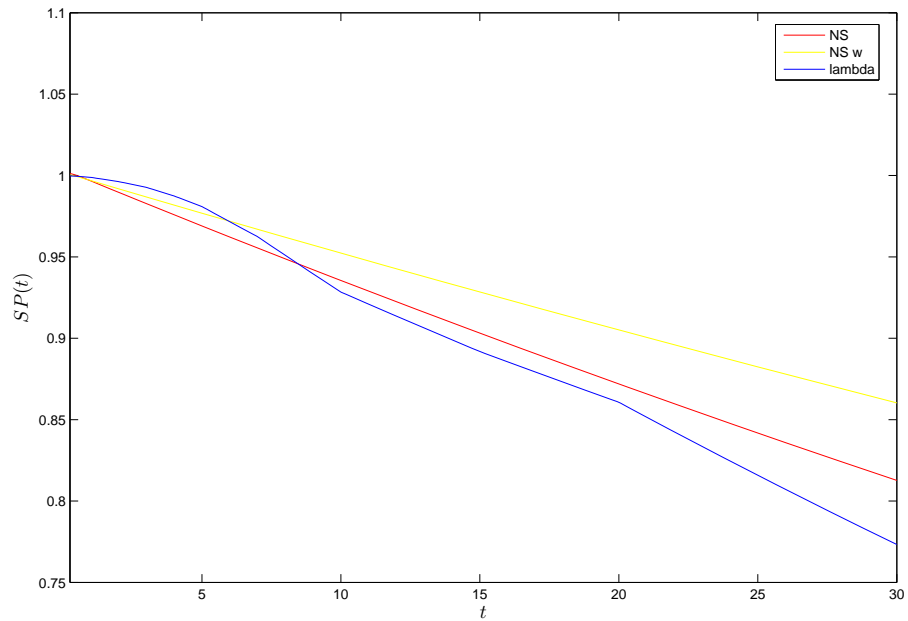
(a) Survival Probabilities
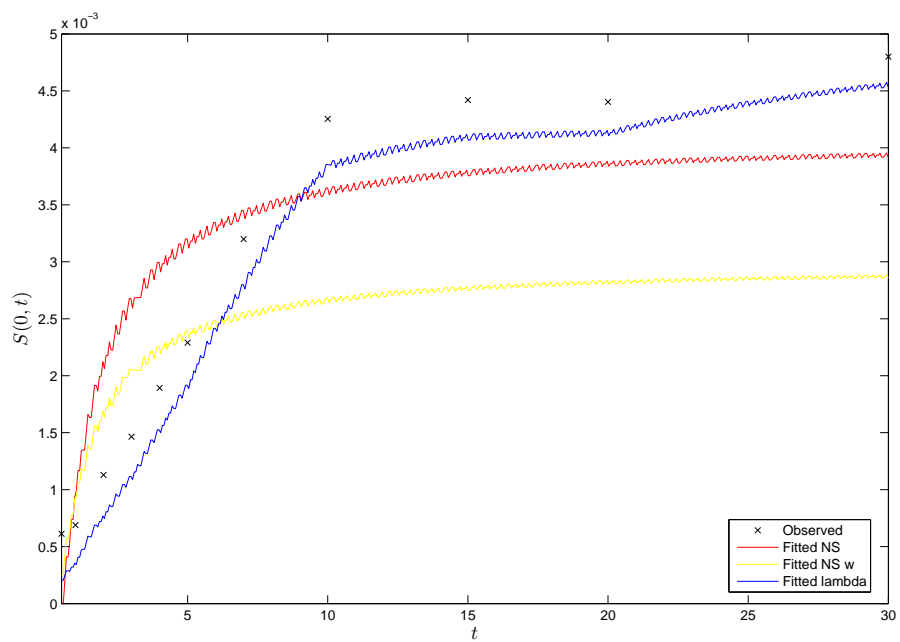
(b) CDS spreads

The corresponding survival probabilities and the ones bootstrapped as in the previous section are plotted in Figure 3.2–Figure 3.4 along with the theoretical and observed CDS spreads.

Finally we report in Table 3.7 the Participation Rates for a contract with maturity $T = 5$ years written against a reference spread with maturity $m = 5$ years.

Table 3.7: Participation Rates for a CMCDS with $T = m = 5$ using the bootstrapping procedure (row lambda), the Nelson-Siegel interpolation (NS) and the NS interpolation with weights in the objective function (NS w).

PR	Abitibi Consol Inc	Microsoft Corp	Tesco PLC
NS	0.780191	0.739188	0.544214
NS w	0.826868	—	0.752114
lambda	0.773298	0.589349	0.519317

3.5.5   Example 2: OU + Convexity adjustment

We use the same data from the previous example. Theoretical CDS and observed CDS are plotted in Figure 3.5–Figure 3.7 and parameters estimation is reported in Table 3.8. Notice that, given the limited number of observations for Microsoft, the fit is quite poor and this results in $\hat\sigma = 0$.

Participation rates are reported in Table 3.9.

Table 3.8: θ Estimates.

	Abitibi	Microsoft	Tesco
k	0.038	0.0037	0.004
α	0.3749	2.6806	0.2919
λ_0	0.0498	0	0.0001
σ	0.0087	0	0.026

Table 3.9: Participation Rates with and without convexity adjustment.

PR	Abitibi	Microsoft	Tesco
With C.A.	0.7415	0.6163	0.6493
Without C.A.	0.7398	0.6163	0.4998

3.6   Statistical Arbitrage Analysis

Having derived the time series of participation rates for a specific name and a reference CDS contract with maturity m years (typically 5 years) we can analyse the properties of the time series $\{y_t\}_{t=t_v,t_v+1,...}$, where

$$y_t = \mathrm{PR}^{tv} \times S(t, t+m) - S(t_v, t_v + T),$$

t_v is the date we are interested in and PR^{tv} is the participation rate on that day.

We express each y_t in basis points.

As a first example we consider the issuer AT&T, $m = 5$ years and the analysis is performed for the day January, 02 2003. In Table 3.10 we report the descriptive statistics for the series y computed using the three approaches described before.

Figure 3.5: CDS spreads for Abitibi using OU for the HR.
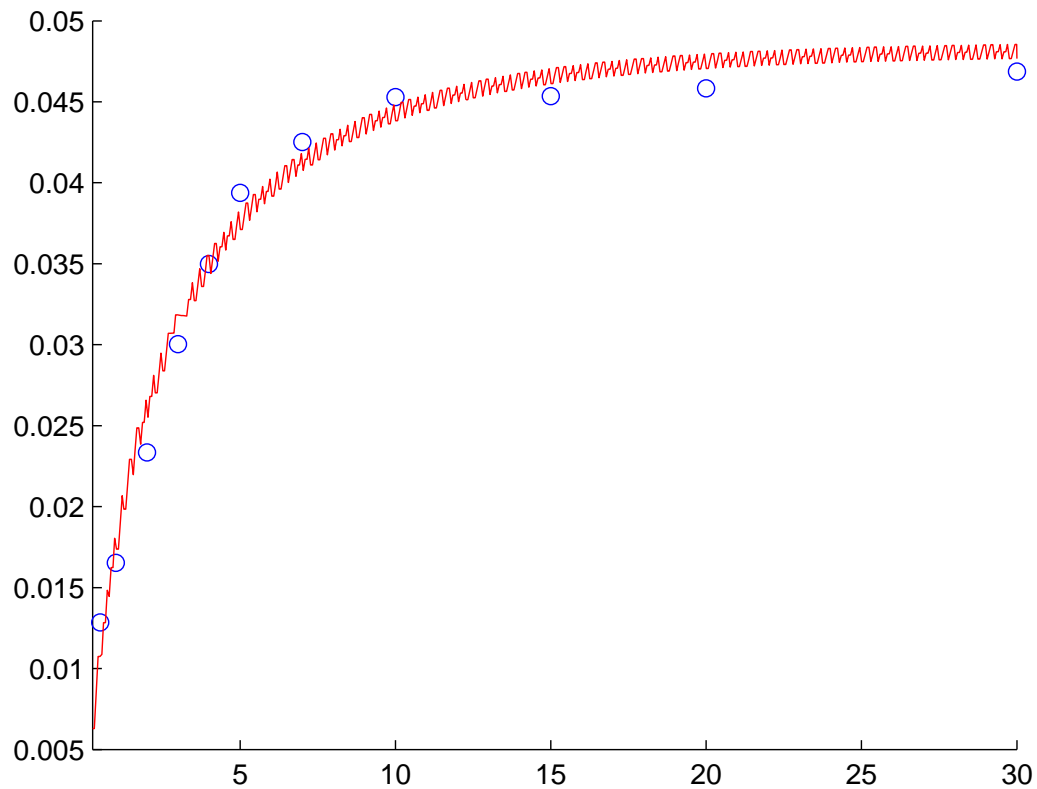
Table 3.10: AT&T – January, 02 2003: descriptive statistics for the time series y (basis points).

	mean	median	std	min	max	95% Range	99% Range
NS	-37.2535	-39.6678	9.5306	-46.517	4.4133	39.3184	45.6049
NSw	-33.4328	-35.6085	8.2444	-42.7188	0.5009	34.4122	39.9939
lambda	-37.031	-39.1492	9.023	-46.3233	-2.5512	38.6349	40.9442

Figure 3.6: CDS spreads for Microsoft using OU for the HR.
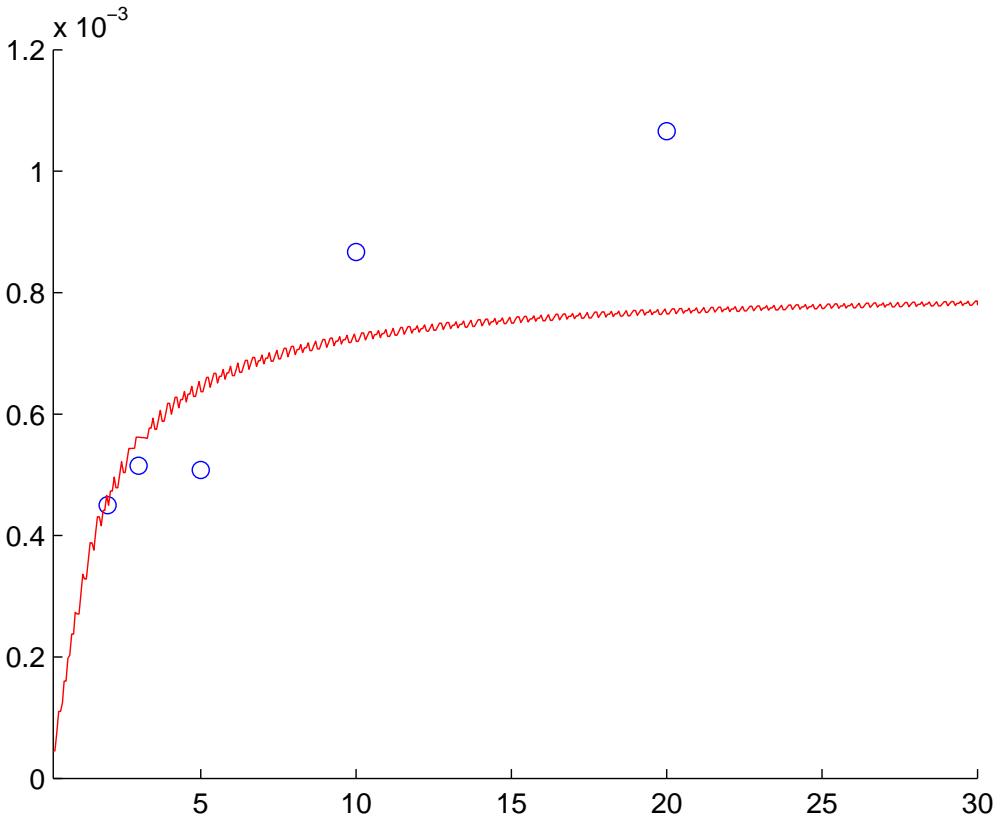
Figure 3.7: CDS spreads for Tesco PLC using OU for the HR.
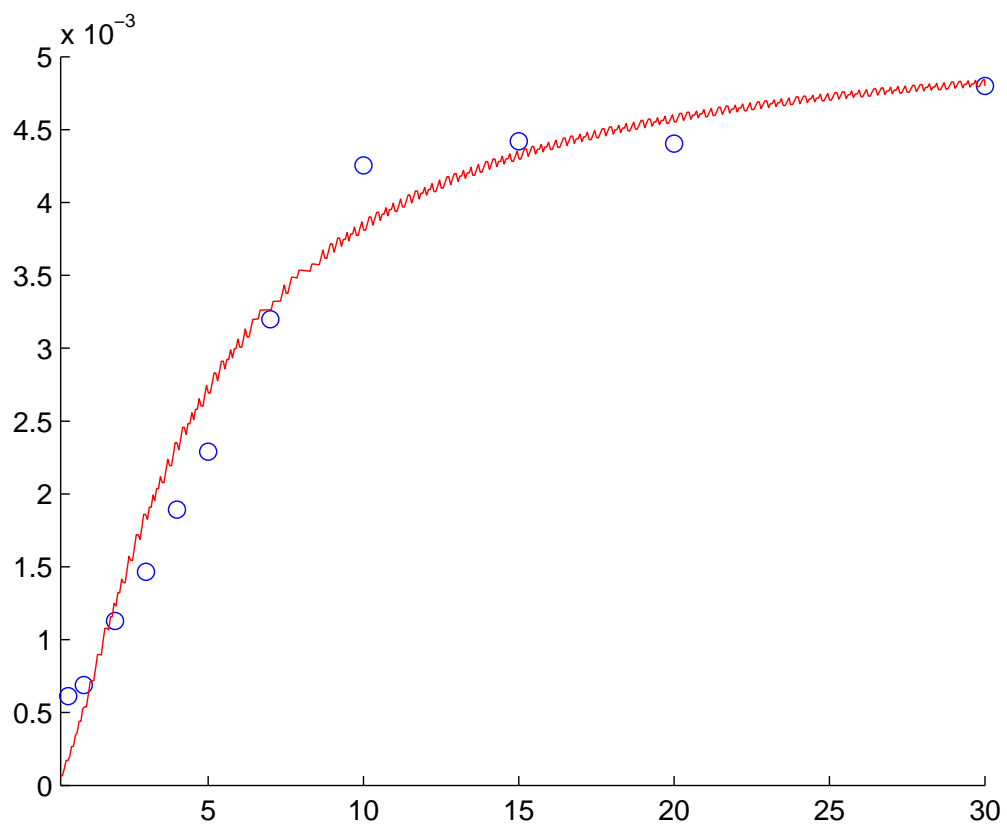
Figure 3.8: Time series and Empirical Density for AT&T – January, 02 2003.

(a) Time series y

(b) Empirical Density
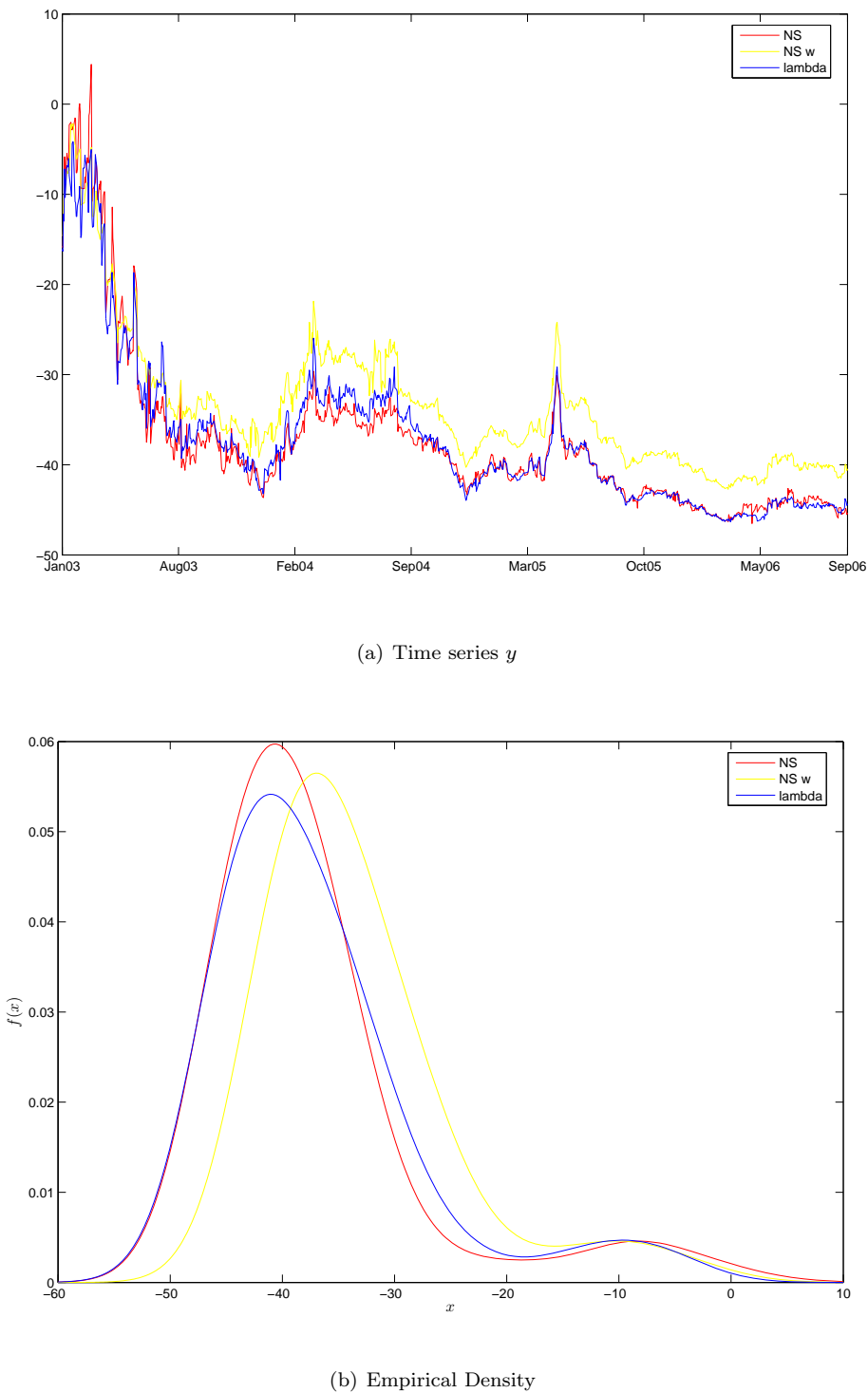
As a second example we consider the issuer Goldman Sachs Gp Inc, $m = 5$ years and the analysis is performed for the day January, 02 2003.

In Table 3.11 we report the descriptive statistics for the series y computed using the three approaches described before.

Table 3.11: Goldman Sachs Gp Inc – January, 02 2003: descriptive statistics for the time series y (basis points).

	mean	median	std	min	max	95% Range	99% Range
NS	-39.2847	-40.6095	7.227	-58.6667	-15.0316	27.2412	32.3140
NS w	-33.7759	-35.1025	7.7681	-58.6667	-7.744	29.7924	34.9852
lambda	-37.9413	-39.6047	8.1287	-58.6667	-11.2436	31.4027	35.1944

Another time series we are interested in is[10]

$$z_{t_v} = \sum_{j=0}^{k-1} \Delta(t_{n-3j}, t_{n-3(j+1)}) \left[\mathrm{PR}^{t_v} \times S(t_{n-3(j+1)}, t_{n-3(j+1)} + m) - S(t_v, t_v + T) \right].$$

In the previous formula we adopt the convention $t_0 \equiv t_v$.

In Table 3.12–Table 3.13 we report the descriptive statistics for the series z computed using the three different approaches for $m = T = 5$.

Since we have data from 02/01/2001 to 01/11/2006 the last CDS contract we can follow for 5 years is the one initiated on 19/12/2001 (first payment is due on 20/12/2001, the last payment is due on 20/09/2006).

Thus for both AT&T and Goldman Sachs Gp Inc we have 249 observation for the time series z.

Table 3.12: AT&T: descriptive statistics for the time series z (basis points).

	mean	median	std	min	max	95% Range	99% Range
NS	-101.3401	-100.8064	42.1469	-238.0150	38.4999	180.8113	228.5471
NSw	-83.2894	-84.6771	45.1470	-222.2393	9.8727	200.0312	219.0970
lambda	-101.8800	-95.4455	39.5045	-232.8395	-46.2014	174.9253	185.1914

Table 3.13: Goldman Sachs Gp Inc: descriptive statistics for the time series z (basis points).

	mean	median	std	min	max	95% Range	99% Range
NS	-133.1578	-121.4141	59.9921	-282.7775	-29.8505	225.0086	252.8016
NSw	-101.3201	-76.2454	56.2825	-250.3429	-25.0121	198.5655	224.8415
lambda	-137.3474	-122.5757	52.1769	-274.8442	-60.2887	191.4644	213.0837

The empirical densities are plotted in Figure 3.10.

---

[10]Actually we compute the first term of the summation as follows:

$$\Delta(t_3, t_v) I_{\{t_3 - t_v > 1mth\}} \left[\mathrm{PR}_{t_v} \times S(t_3, t_3 + m) - S(t_v, t_v + T) \right]$$
$$+ \Delta(t_3, t_v) I_{\{t_3 - t_v \leq 1mth\}} \left[\mathrm{PR}_{t_v} \times S(t_6, t_6 + m) - S(t_v, t_v + T) \right]$$
$$= \Delta(t_3, t_v) \left[\mathrm{PR}_{t_v} \left(S(t_3, t_3 + m) I_{\{t_3 - t_v > 1mth\}} + S(t_6, t_6 + m) I_{\{t_3 - t_v \leq 1mth\}} \right) - S(t_v, t_v + T) \right]$$

to take into account the different behavior of the first coupon.

Figure 3.9: Time series and Empirical Density for Goldman Sachs Gp Inc – January, 02 2003.
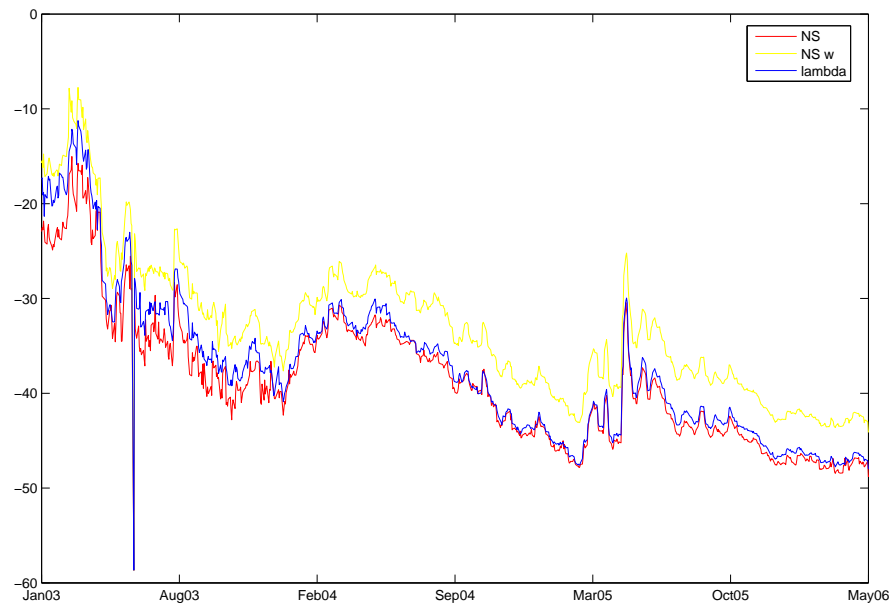
(a) Time series y
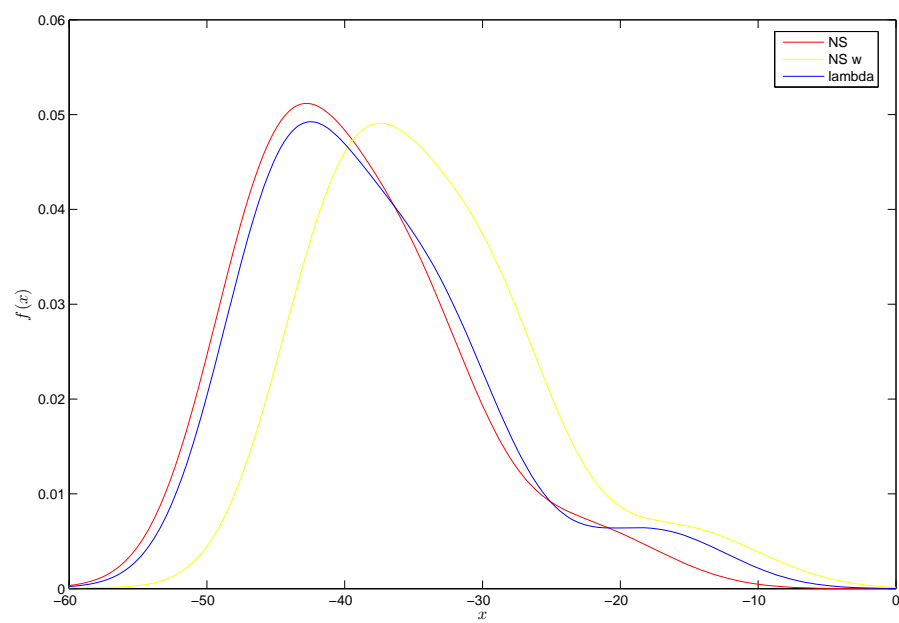
(b) Empirical Density

Figure 3.10: Time series z: Empirical Densities
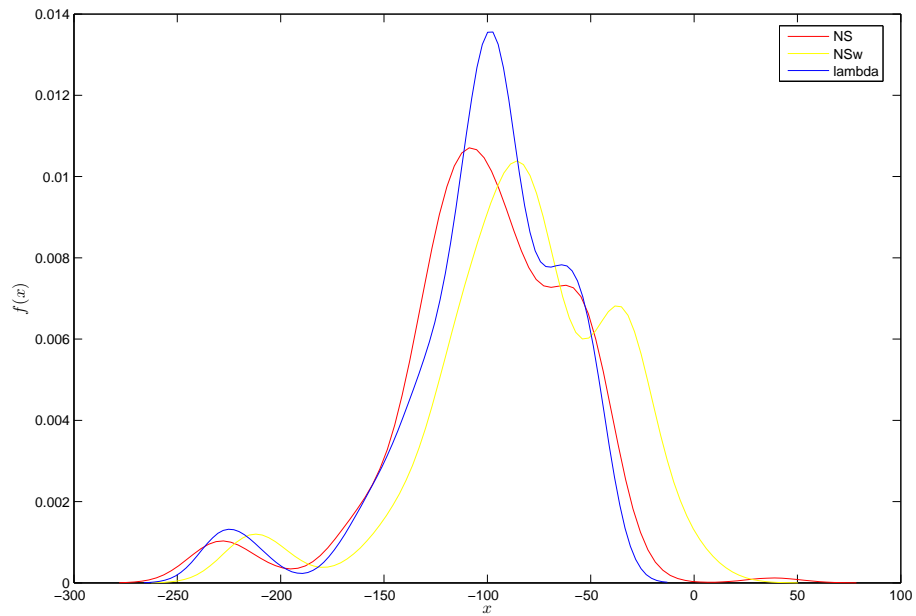
(a) Time series z: Empirical Density for AT&T
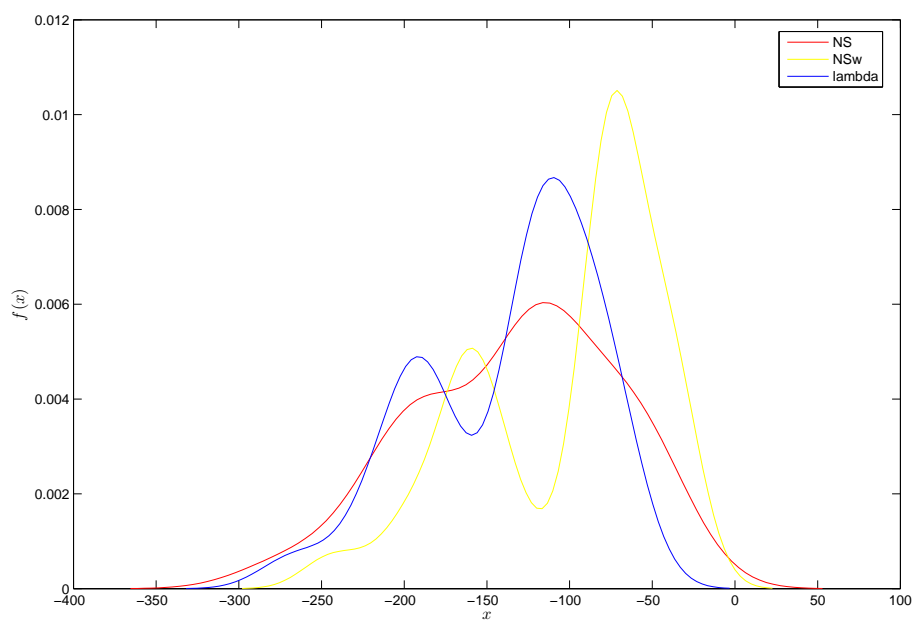
(b) Time series z: Empirical Density for Goldman Sachs

3.6.1   Data Description

Our dataset consists of daily single-name composite spreads for the period January 2001–November 2006 with maturities 6m, 1y, 2y, 3y, 4y, 5y, 7y, 10y, 15y, 20y, 30y.

The composite spread is the average spread for an instrument provided to Markit by its contributors[11] after prices and spreads failing the data quality tests have been removed from the sample set. The cleaning process includes testing for Stale, Flat curves and Outlying data[12]. There are 2250 companies in the dataset. In some cases there are missing values, especially for not very liquid maturities. For each day and for each name also a recovery rate is reported. Additional informations like sector, rating and country are reported as well.

The number of companies in each sector and rating category are displayed in Table 3.14.

Table 3.14:   Number of companies available for each sector and rating. NA means that the rating is not available.

Sector	No of Companies
Basic Materials	156
Consumer Goods	229
Consumer Services	336
Financials	556
Government	66
Health Care	88
Industrials	279
Oil & Gas	131
Technology	80
Telecommunications	110
Utilities	212

Rating	No of Companies
AAA	59
AA	178
A	533
BBB	636
BB	285
B	234
CCC	56
D	4
NA	258

Summary statistics regarding the whole sample are reported in Table 3.15. Statistics regarding the different sectors and ratings are given in Table 3.16 and Table 3.17.

In Figure 3.11 we report for each payment date (the 20th of March, June, September and December) the number of companies for which we have the 5-years CDS spread and the recovery rate.

As far as the construction of the discount factor is concerned, we use Libor rates with maturities 1 month to 11 months and swap rates with maturities 1y, 2y, 3y, 4y, 5y, 7y, 10y, 20y, 30y. Libor-Swap data spans the same interval we have CDS data for.

---

[11]Markit only builds composites when there are at least three contributors to each composite.

[12]On average Markit rejects approximately 45% of the CDS data received due to failure under any combination of the three criteria above.

Table 3.15: Summary statistics for CDS Data. The table display summary statistics for CDS spreads for maturities 1y, 3y, 5y, 7y, and 10y.

Mean CDS Spreads

	mean	median	std	min	max	95% Range	99% Range
1y	0.009847	0.002218	0.041555	9.17E-05	0.850294	0.047849	0.241539
3y	0.012076	0.003799	0.036753	0.000117	0.71018	0.053921	0.236742
5y	0.014637	0.005607	0.035076	0.000132	0.709107	0.061784	0.236483
7y	0.014311	0.006127	0.03074	0.000192	0.608353	0.059071	0.198261
10y	0.014702	0.006992	0.027667	0.000272	0.619683	0.058565	0.178252

Skewness in CDS Spreads

	mean	median	std	min	max	95% Range	99% Range
1y	1.180305	1.101156	1.604095	-8.74103	20.49202	6.096927	12.51488
3y	0.97082	0.93535	1.704507	-8.57648	16.78911	6.13819	14.82796
5y	0.880433	0.863844	1.865777	-10.8887	16.79208	7.579692	16.50355
7y	0.889122	0.818754	1.820572	-9.37944	21.05178	6.365396	15.99307
10y	0.815531	0.729948	1.793077	-9.37363	22.80157	6.354377	15.22473

Kurtosis in CDS Spreads

	mean	median	std	min	max	95% Range	99% Range
1y	6.564576	4.163205	8.70691	1.066845	111.2795	24.24266	63.55712
3y	5.915394	3.951006	8.182648	1.339658	123.9829	18.40068	54.23267
5y	5.936898	3.840767	11.72973	1.095111	340.5836	20.72231	42.03766
7y	6.078371	3.757899	16.5301	1.086447	509.6383	17.33197	73.8173
10y	6.138454	3.881449	18.02118	1.16189	570.8451	17.06132	61.71578

Table 3.16: Summary statistics for CDS Data. The table display summary statistics for CDS spreads for different sectors.

Mean CDS Spreads

	mean	median	std	min	max	95% Range	99% Range
Basic Materials	0.013115565	0.005847	0.01769	0.000778	0.129974	0.053576	0.116799
Consumer Goods	0.017914838	0.005195	0.064818	0.00033	0.850294	0.098179	0.548034
Consumer Services	0.021290725	0.008618	0.041802	0.000546	0.35549	0.152033	0.292399
Financials	0.004670969	0.003398	0.00508	0.000265	0.054019	0.018962	0.029933
Government	0.001231071	0.000721	0.001137	9.17E-05	0.004672	0.00408	0.004527
Health Care	0.007821398	0.00418	0.010616	0.00025	0.059477	0.047891	0.056071
Industrials	0.010791964	0.005321	0.01593	0.000478	0.164012	0.040733	0.146708
Oil & Gas	0.008422083	0.004349	0.014015	0.000417	0.100294	0.047868	0.091808
Technology	0.017950102	0.012335	0.018573	0.000362	0.101037	0.0646	0.099862
Telecommunications	0.015466806	0.007201	0.023648	0.000976	0.149589	0.070958	0.147704
Utilities	0.017529436	0.004858	0.06227	0.000492	0.619683	0.139286	0.556189

Skewness in CDS Spreads

	mean	median	std	min	max	95% Range	99% Range
Basic Materials	1.115199503	1.029712	0.829437	-1.09909	4.629681	3.233034	5.161346
Consumer Goods	1.199036878	1.060946	0.980106	-1.37137	7.583861	3.652261	7.907668
Consumer Services	1.226283353	1.128451	0.94509	-0.89613	8.128643	3.921055	5.341681
Financials	1.403494921	1.327875	1.435624	-3.93968	22.80157	4.577992	9.011403
Government	0.695333971	0.740212	1.001448	-3.39659	6.958991	3.950653	9.287309
Health Care	1.317245699	1.093989	1.213208	-2.71774	6.741141	4.935727	7.877284
Industrials	1.314711934	1.16284	1.065581	-3.2334	10.18853	3.900957	8.151431
Oil & Gas	1.187319361	1.180794	0.836157	-1.43266	5.802308	3.191153	5.038656
Technology	1.386277668	1.384586	1.103867	-5.01219	5.111549	4.257138	7.644472
Telecommunications	1.778714737	1.986945	1.103029	-0.32964	6.529731	3.892549	6.118368
Utilities	1.755516246	1.471207	1.336556	-1.75104	8.444239	5.204834	8.890933

Kurtosis in CDS Spreads

|	mean	median	std	min	max	95% Range	99% Range
Basic Materials	4.725374464	3.433504	4.712496	1.633169	66.23901	13.0316	23.50079
Consumer Goods	5.219657413	3.504127	6.329019	1.095111	68.43056	17.1359	52.72311
Consumer Services	5.180144076	3.638552	6.020843	1.37034	111.2795	15.56479	27.96076
Financials	7.261781756	4.419179	24.23604	1.339658	570.8451	22.51721	72.77808
Government	4.143783346	3.00651	5.582496	1.297846	64.39082	10.61637	52.15812
Health Care	6.399099549	3.867713	7.946202	1.066845	60.0065	30.17454	55.54095
Industrials	5.733085643	3.796261	7.739467	1.166231	129.3136	18.29561	51.70547
Oil & Gas	4.610600866	3.800453	3.248152	1.522167	38.73359	9.662182	17.92033
Technology	6.1141031	4.846896	5.464049	1.212769	45.98889	21.85722	37.8039
Telecommunications	7.363434391	6.601662	5.67587	1.356503	47.47271	17.217	38.66052
Utilities	8.515696087	4.423587	12.74694	1.630686	123.9829	33.55957	107.619

Table 3.17: Summary statistics for CDS Data. The table display summary statistics for CDS spreads for different rating.

Mean CDS Spreads

	mean	median	std	min	max	95% Range	99% Range
AAA	0.001396	0.000876	0.00132	9.17E-05	0.00649	0.004714	0.006061
AA	0.002159	0.001925	0.001404	0.000171	0.010347	0.005657	0.008021
A	0.003646	0.003271	0.002566	0.00047	0.031592	0.00816	0.020505
BBB	0.006941	0.005796	0.005136	0.000634	0.058662	0.019564	0.032704
BB	0.018396	0.017343	0.010343	0.001091	0.05519	0.039997	0.052501
B	0.03816	0.032265	0.025126	0.005806	0.178413	0.114188	0.170583
CCC	0.137578	0.129972	0.125269	0.007735	0.619683	0.562641	0.611948
D	0.372192	0.3057	0.16753	0.163037	0.850294	0.687257	0.687257
NA	0.013966	0.007319	0.01518	0.00033	0.077251	0.054285	0.07308

Skewness in CDS Spreads

	mean	median	std	min	max	95% Range	99% Range
AAA	1.085749	1.075072	1.159891	-3.39659	6.958991	4.730811	9.184179
AA	1.030061	0.983449	1.045906	-2.34701	8.215932	3.929237	8.266868
A	1.371482	1.244432	1.304575	-3.93968	22.80157	3.852406	8.143102
BBB	1.475064	1.333242	1.057831	-2.71774	8.444239	4.098936	7.465631
BB	1.312732	1.160984	1.069793	-0.73072	10.18853	3.617585	7.786574
B	1.055297	0.877724	1.137551	-5.01219	6.735526	4.232044	6.698706
CCC	1.307536	0.899755	1.730187	-1.75104	7.583861	7.610114	9.334898
D	1.871435	1.490808	1.654872	-0.65624	6.356648	7.012892	7.012892
NA	1.1715	1.084081	1.064639	-3.2334	7.269395	4.282904	6.94283

Kurtosis in CDS Spreads

	mean	median	std	min	max	95% Range	99% Range
AAA	5.863425	3.679848	7.639128	1.066845	64.39082	31.44679	55.25506
AA	5.119389	3.60366	8.014238	1.297846	107.4008	14.34381	75.1934
A	6.608426	3.999403	22.00189	1.375873	570.8451	16.95701	53.79945
BBB	6.238616	4.232175	7.070783	1.334403	109.6464	20.25037	50.68356
BB	5.724678	3.870099	7.988599	1.2595	129.3136	17.16115	55.38361
B	5.276405	3.358172	6.561089	1.207794	75.41828	21.73929	50.79833
CCC	8.101289	3.160812	13.65959	1.552613	68.43056	64.83355	66.87795
D	11.66408	4.812727	13.93133	1.095111	57.73114	56.63602	56.63602
NA	5.755597	3.778785	8.624497	1.166231	110.9138	20.43657	67.45183

Figure 3.11: Number of companies for which we have the 5-years CDS spread and the recovery rate in each payment date (the 20th of March, June, September and December).
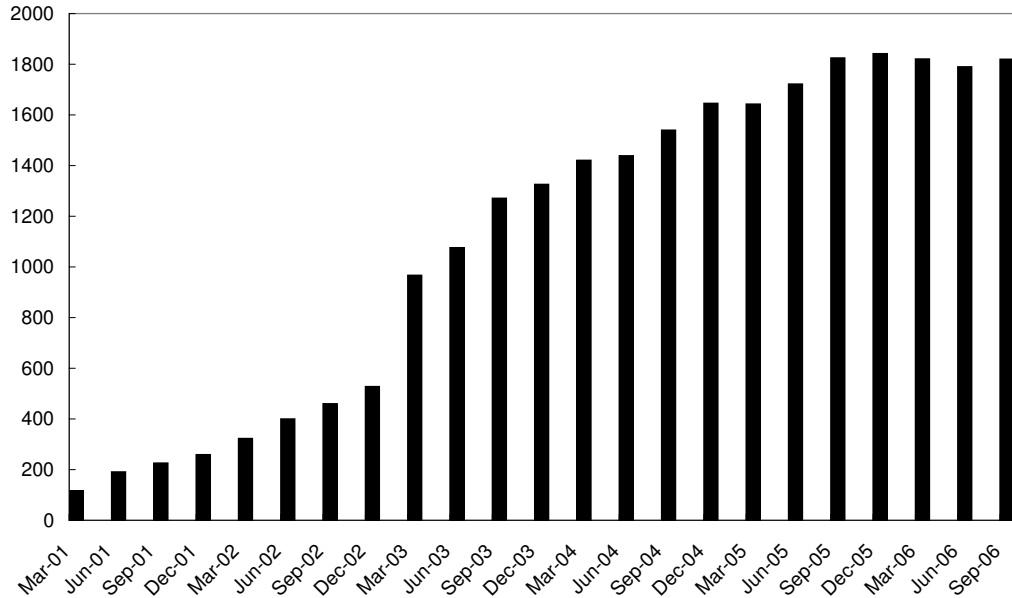
3.6.2   Analysis 1

In this first analysis we compute for all the companies for which we have the required data the profit (or loss) that an investor would have realized being long a CMCDS with maturity 5 years (the maturity of the reference CDS is 5 years as well) and short a CDS with both the contracts initiated on 20/09/2001. With this choice we can perform this analysis for a number of companies between 200 and 204, depending on the method used to compute the participation rate. In other words for $j = 1, 2, \ldots$ we compute the vector $\mathbf{z}$ with components

$$z_j = \sum_{i=0}^{k-1} \Delta(t_{n-3i}, t_{n-3(i+1)}) \left[\mathrm{PR}_j^{t_v} \times S_j(t_{n-3(i+1)}, t_{n-3(i+1)} + m) - S_j(t_v, t_v + T) \right].$$

where, as usual t_i denotes a payment date, t_v is 20/09/2001 and $S_j(u, u + m)$ denotes the CDS spread at time u with maturity m for company j and $\mathrm{PR}_j^{t_v}$ is the participation rate[13] for company j at time t_v. An illustration is given in Figure 3.12.

We first compute the participation rate using the Nelson-Siegel interpolation. Descriptive statistics are reported in Table 3.18 and a stem plot of vector $\mathbf{z}$ is reported in Figure 3.13.
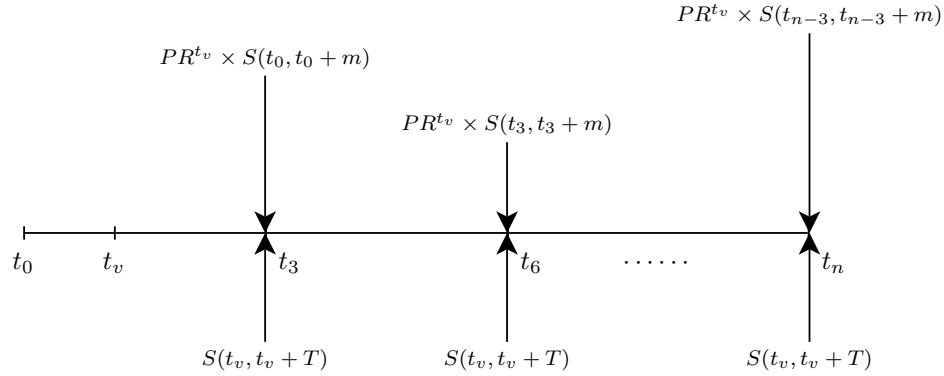
Using this method 173 observations out of 203 are negative (85.22%).

Table 3.18: Descriptive statistics for the vector of observations $\mathbf{z}$ (Nelson Siegel).

Observations	mean	median	std	min	max	95% Range	99% Range
203	-172.344	-146.141	410.7125	-1450.21	1221.589	2102.138	2658.712

Next, we repeat the same analysis in the case in which the participation rate is calculated using the method which employs piecewise constant hazard rates.

---

[13]To be precise this should read $\mathrm{PR}_j^{t_v}(m, T)$.

Figure 3.12: Illustration of the statistical arbitrage analysis.

Using this method 181 observations out of 204 are negative (88.7%).

Table 3.19: Descriptive statistics for the vector of observations **z** (Piecewise Constant Hazard Rates).

Observations	mean	median	std	min	max	95% Range	99% Range
204	-181.013	-156.096	382.8685	-1485.07	1121.221	2035.425	2515.842

When the participation rate is calculated assuming a OU process for the evolution of the hazard rate, we obtain for the vector of observations **z** the descriptive statistics reported in Table 3.20. A plot of vector **z** is reported in Figure 3.15. It is interesting to notice that the vector **z** has 181 negative components out of 200 (90.5%).

Table 3.20: Descriptive statistics for the vector of observations **z** (OU process for Hazard Rates).

Observations	mean	median	std	min	max	95% Range	99% Range
200	-184.272	-157.282	408.5472	-1476.73	1366.107	2149.537	2816.082

Using this method we can also evaluate the impact of the convexity adjustment. Table 3.21 reports the descriptive statistics for the vector of observations **z** when a OU process for Hazard Rates is used and the convexity adjustment is taken into account, while Figure 3.16 gives a graphical representation. In this case the vector **z** has 171 negative components out of 200 (85.5%).
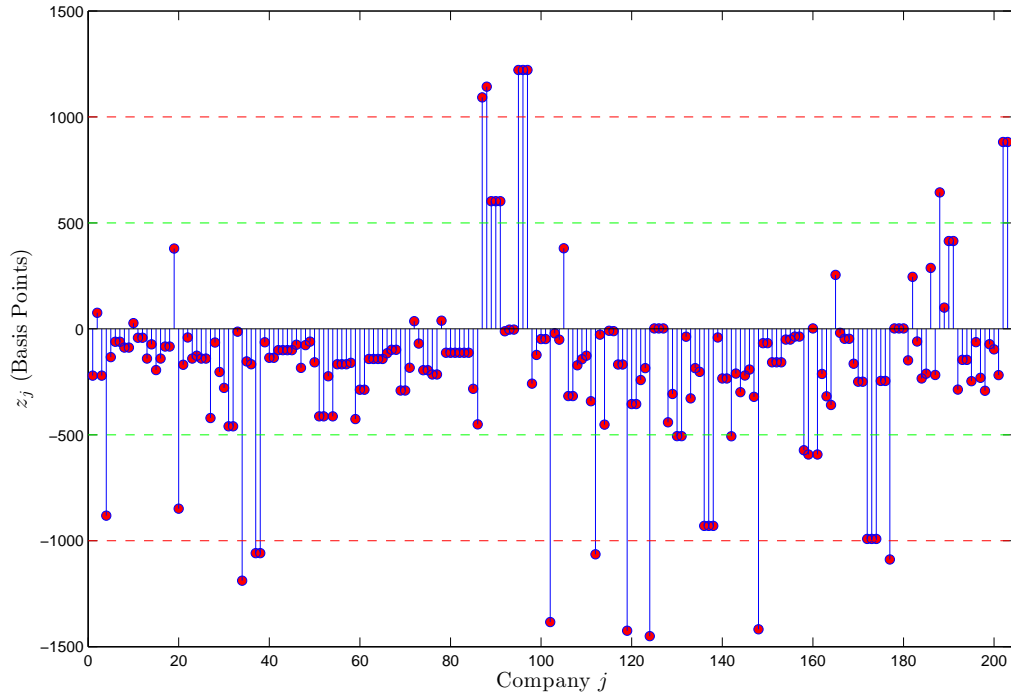
To summarize the results of this first analysis, we report (Table 3.22) for each method how many companies have a z_j positive (negative), bigger than 500 bp (smaller than -500 bp) and bigger than 1000 bp (smaller than -1000 bp).

It is interesting to sort the companies according to their z_j. First of all we report (Table 3.23) a list of the companies with $z_j < -500$ bp (Panel A) and those with $z_j > 500$ bp (Panel B) for *all* the methods used. In Table 3.24 we do the same for the -1000 and 1000 bp bounds.

In Table 3.25 we report for each method the five companies with the most negative z_j (Panel A) and the first five companies with the biggest z_j (Panel B).

According to three methods out of four (the exception is the Nelson-Siegel method) Hasbro Inc is the company for which the loss ($-z_j$) that an investor would have realized being long a CMCDS with maturity 5 years and short a CDS with both the contracts initiated on 20/09/2001 is maximum.

On the other hand using three methods (the exception being the one with piecewise constant hazard rates) Global Marine Inc is the company with maximum profit z_j. It is interesting to notice that every method include General Motors and Ford among the five companies with the biggest z_j.

Figure 3.13: Vector of observations **z** (Nelson Siegel): Stem plot.

Table 3.21: Descriptive statistics for the vector of observations **z** (OU process for Hazard Rates with the convexity adjustment).

Observations	mean	median	std	min	max	95% Range	99% Range
200	-165.075	-152.346	422.681	-1474.86	1366.107	2073.687	2815.146

Table 3.22: This table reports for each method how many companies have a z_j positive (negative), bigger than 500 bp (smaller than -500 bp) and bigger than 1000 bp (smaller than -1000 bp).

	NS	lambda	OU	OU conv
pos	30	23	19	29
neg	173	181	181	171
> 500 bp	11	10	10	11
< −500 bp	23	23	23	23
> 1000 bp	5	4	4	4
< −1000 bp	9	7	8	8

Figure 3.14: Vector of observations **z** (Piecewise Constant Hazard Rates): Stem plot.
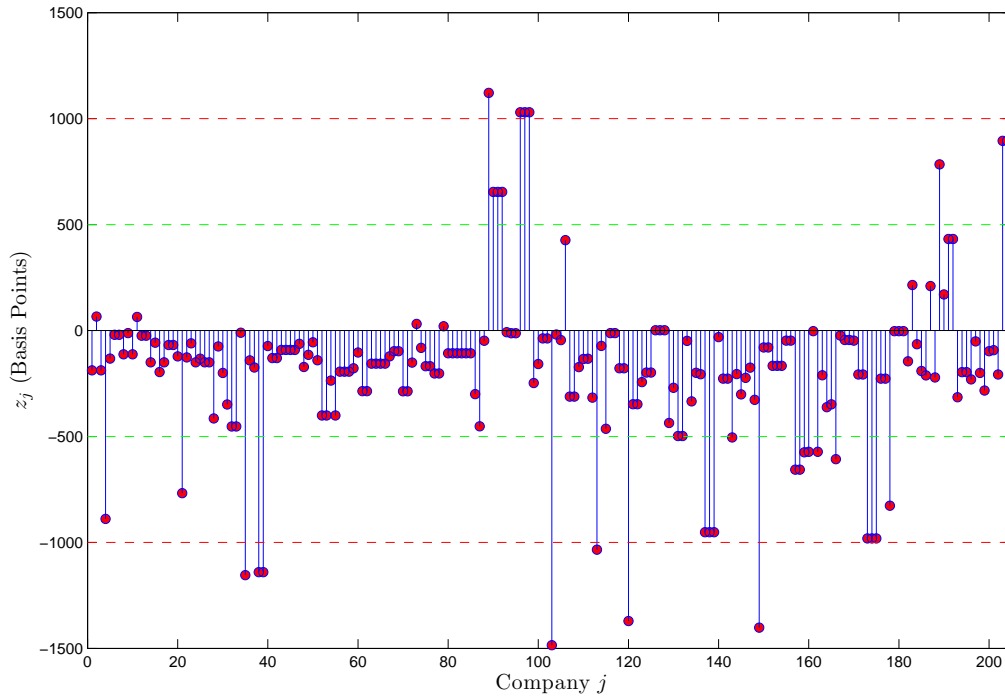
Table 3.23: Panel A reports for each method the companies for which $z_j < -500$ bp according to all methods. Panel B reports for each method the companies for which $z_j > 500$ bp according to all methods.

Panel A	Panel B
Aetna Inc.	Ford Mtr Co
Arrow Electrs Inc	Ford Mtr Cr Co
CNA Finl Corp	GA Pac Corp
Cap One Bk	GATX Finl Corp
Cap One Finl Corp	Gen Mtrs Corp
Hasbro Inc	Gillette Co
J C Penney Co Inc	Global Marine Inc
LA Pac Corp	Toys R Us Inc
Motorola Inc	Williams Cos Inc
NOVA Chems Corp	Wyeth
Nabors Inds Inc	
Nordstrom Inc	
Pennzoil Quaker St Co	
Raytheon Co	
Reebok Intl Ltd	
Roche Hldgs Inc	
ServiceMaster Co	
Shaw Comms Inc	
Sherwin Williams Co	

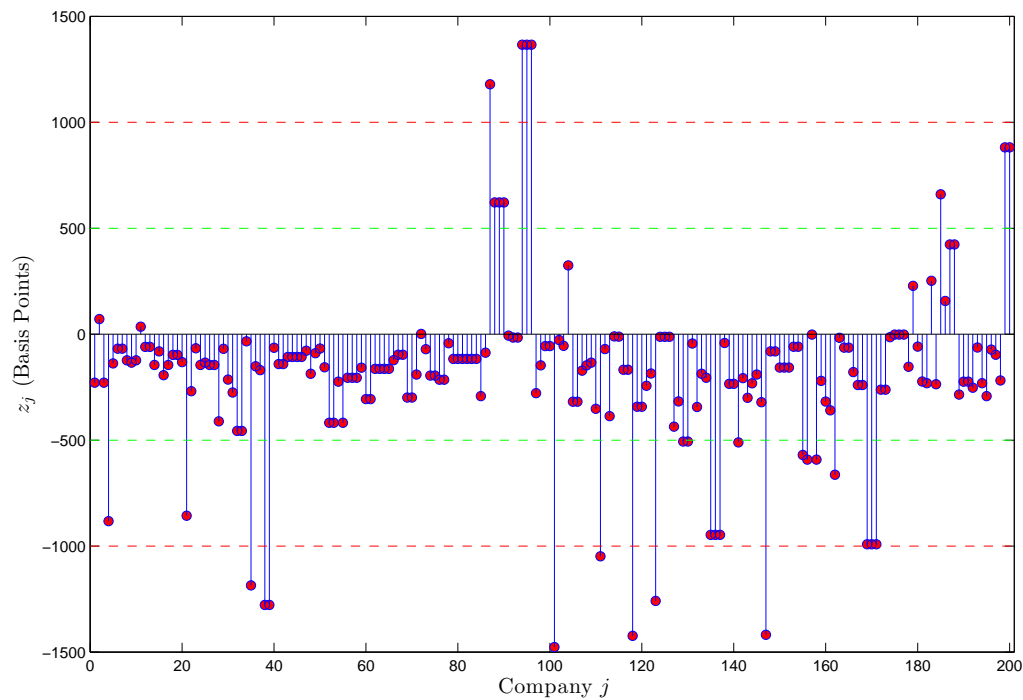**Figure 3.15:** Vector of observations **z** (OU process for Hazard Rates): Stem plot.

Table 3.24: Panel A reports for each method the companies for which $z_j < -1000$ bp according to all methods. Panel B reports for each method the companies for which $z_j > 1000$ bp according to all methods.

Panel A	Panel B
CNA Finl Corp	Ford Mtr Co
Cap One Bk	Gen Mtrs Corp
Cap One Finl Corp	Gillette Co
Hasbro Inc	Global Marine Inc
J C Penney Co Inc	
LA Pac Corp	
Pennzoil Quaker St Co	

Figure 3.16: Vector of observations **z** (OU process for Hazard Rates with the convexity adjustment): Stem plot.
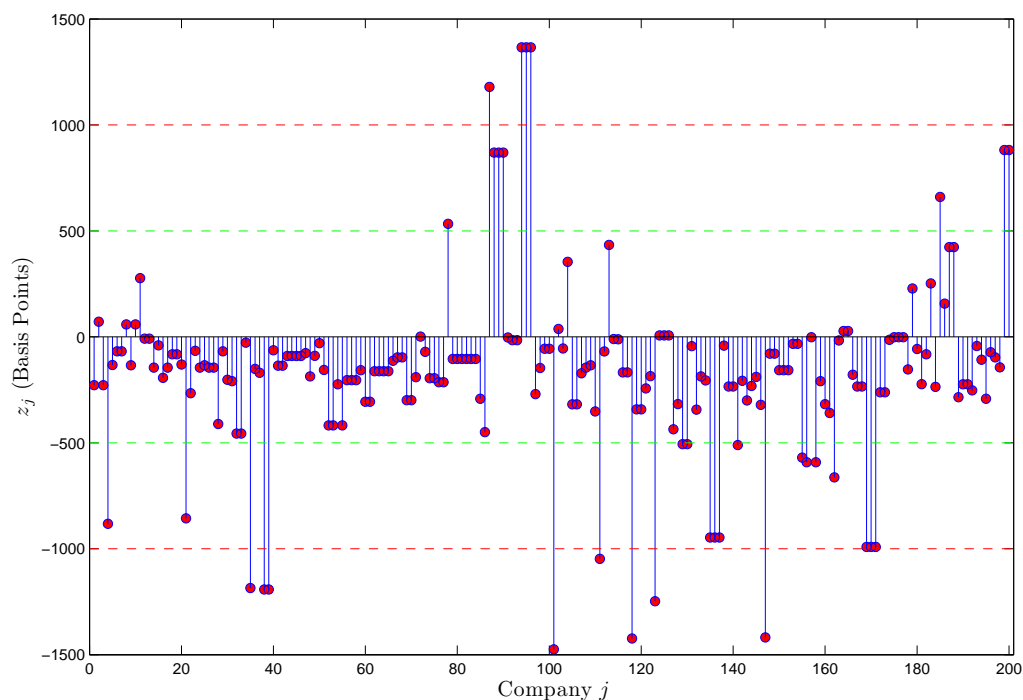

Table 3.25: Panel A reports for each method the five companies with the most negative z_j. Panel B reports for each method the five companies with the biggest z_j.

Panel A

NS	lambda	OU	OU conv
Lucent Tech Inc	Hasbro Inc	Hasbro Inc	Hasbro Inc
LA Pac Corp	Pennzoil Quaker St Co	LA Pac Corp	LA Pac Corp
Pennzoil Quaker St Co	LA Pac Corp	Pennzoil Quaker St Co	Pennzoil Quaker St Co
Hasbro Inc	CNA Finl Corp	Cap One Bk	Lucent Tech Inc
CNA Finl Corp	Cap One Bk	Cap One Finl Corp	Cap One Bk

Panel B

NS	lambda	OU	OU conv
Finl Sec Assurn Inc	Wyeth	Wyeth	Wyeth
Ford Mtr Co	Gen Mtrs Corp	Ford Mtr Co	Ford Mtr Co
Gen Mtrs Corp	Gillette Co	Gen Mtrs Corp	Gen Mtrs Corp
Gillette Co	Global Marine Inc	Gillette Co	Gillette Co
Global Marine Inc	Ford Mtr Co	Global Marine Inc	Global Marine Inc

Note that we also applied a cap on the floating payment and computed

$$z_j^{t_v,\text{cap}} = \sum_{i=0}^{k-1} \Delta(t_{n-3i}, t_{n-3(i+1)}) \Big[\min \big\{ 800bp, \text{PR}_j^{t_v} \times S_j(t_{n-3(i+1)}, t_{n-3(i+1)} + m) \big\}$$
$$-S_j(t_v, t_v + T) \Big]$$

but the results are exactly the same.

3.6.3   Analysis 2

In the analysis that follows for each company j we look at

$$z_j^{t_v} = \sum_{i=0}^{k-1} \Delta(t_{n-3i}, t_{n-3(i+1)}) \left[\text{PR}_j^{t_v} \times S_j(t_{n-3(i+1)}, t_{n-3(i+1)} + m) - S_j(t_v, t_v + T) \right]$$

for each day t_v between 02/01/2001 and 19/12/2001 for which we have the data required. Given company j, we denote by n_j the number of days t_v for which we can derive $z_j^{t_v}$.

The output of the present analysis is the vector $\bar{\mathbf{z}}$ with elements

$$\bar{z}_j = \frac{1}{n_j} \sum_{t_v=1}^{n_j} z_j^{t_v}.$$

Using the Nelson-Siegel interpolation to compute the participation rate, we obtain the descriptive statistics reported in Table 3.26 and a stem plot reported in Figure 3.17 for the vector $\bar{\mathbf{z}}$.

Using this method 160 observations out of 213 are negative (75.12%).

Table 3.26: Descriptive statistics for the vector of observations $\bar{\mathbf{z}}$ (Nelson Siegel).

Observations	mean	median	std	min	max	95% Range	99% Range
213	-109.54	-112.407	298.6551	-1494.8	1129.924	1222.977	2349.756

When we use the method of the piecewise constant hazard rates 183 observations out of 213 are negative (85.92%).

Descriptive statistics for the vector $\bar{\mathbf{z}}$ are reported in Table 3.27 and a stem plot reported in Figure 3.18

Table 3.27: Descriptive statistics for the vector of observations $\bar{\mathbf{z}}$ (Piecewise Constant Hazard Rates).

Observations	mean	median	std	min	max	95% Range	99% Range
213	-153.432	-146.417	274.2414	-1090.6	1100.494	1239.63	2131.705

Next, we compute the participation rate assuming a OU process for the evolution of the hazard rate. For the vector of observations we obtain $\bar{\mathbf{z}}$ the descriptive statistics reported in Table 3.28. A plot of vector $\mathbf{z}$ is reported in Figure 3.19. The negative components in the vector $\bar{\mathbf{z}}$ are 187 out of 213 (87.79%).

Table 3.28: Descriptive statistics for the vector of observations $\bar{\mathbf{z}}$ (OU process for Hazard Rates).

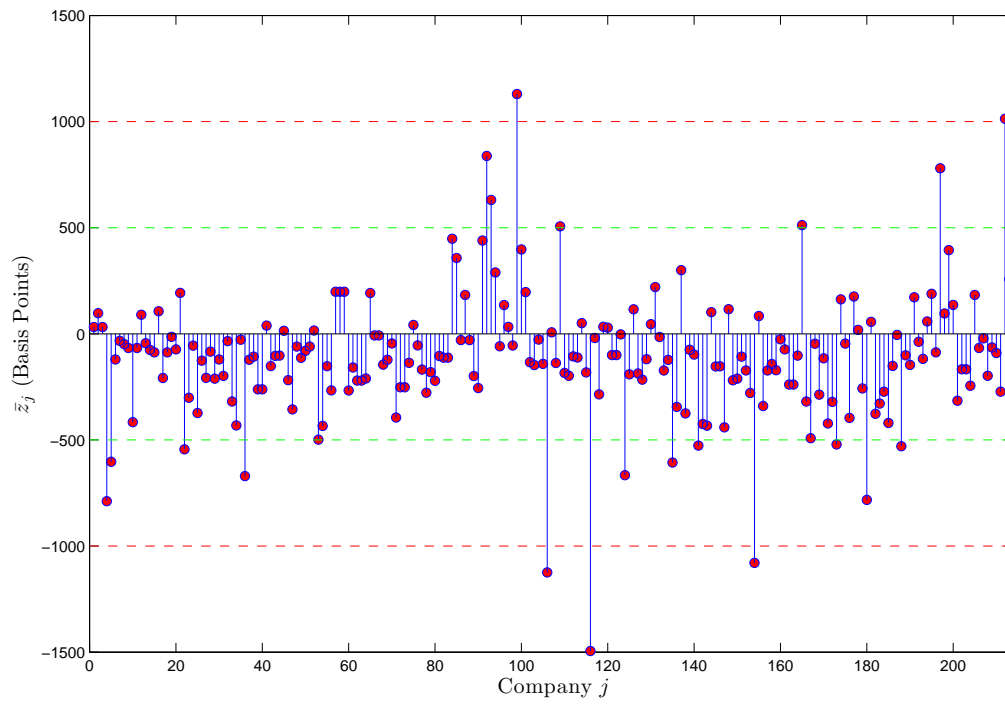| Observations	mean	median	std	min	max	95% Range	99% Range
213	-169.898	-157.106	277.3286	-1475.95	1048.465	1191.47	2310.164

Figure 3.17: Vector of observations $\bar{\mathbf{z}}$ (Nelson Siegel): Stem plot.

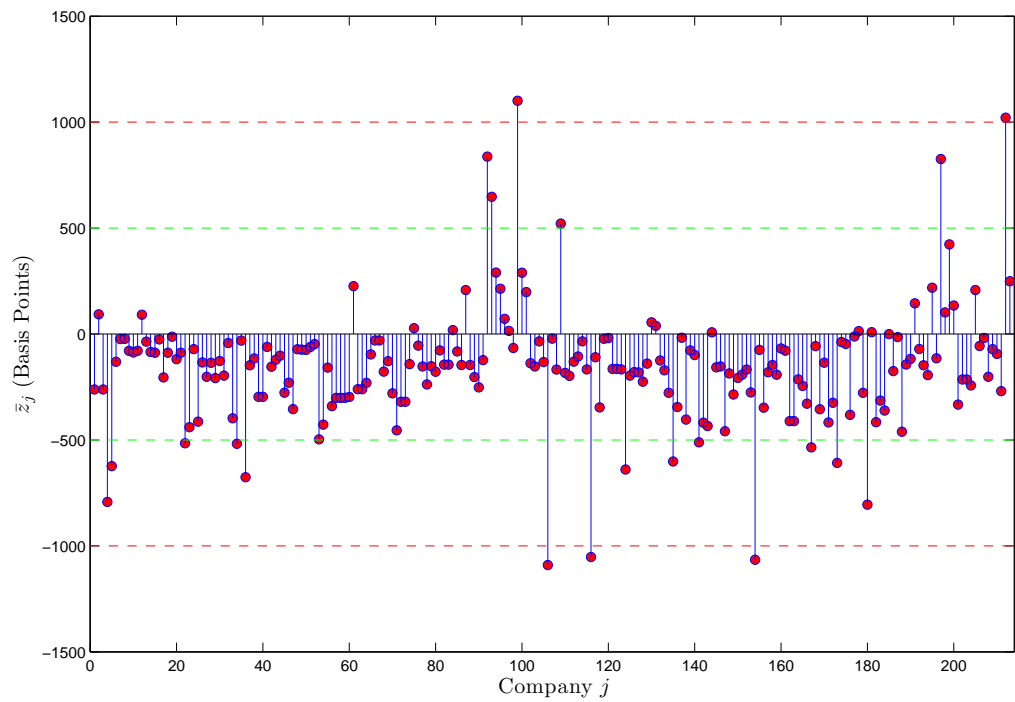**Figure 3.18:** Vector of observations $\bar{\mathbf{z}}$ (Piecewise Constant Hazard Rates): Stem plot.

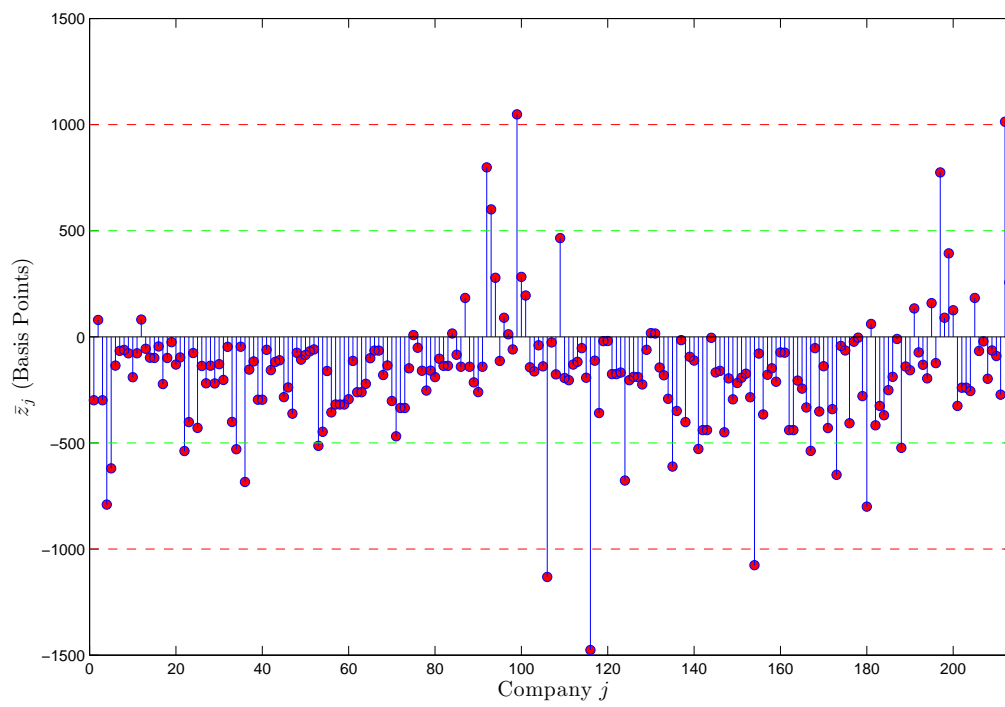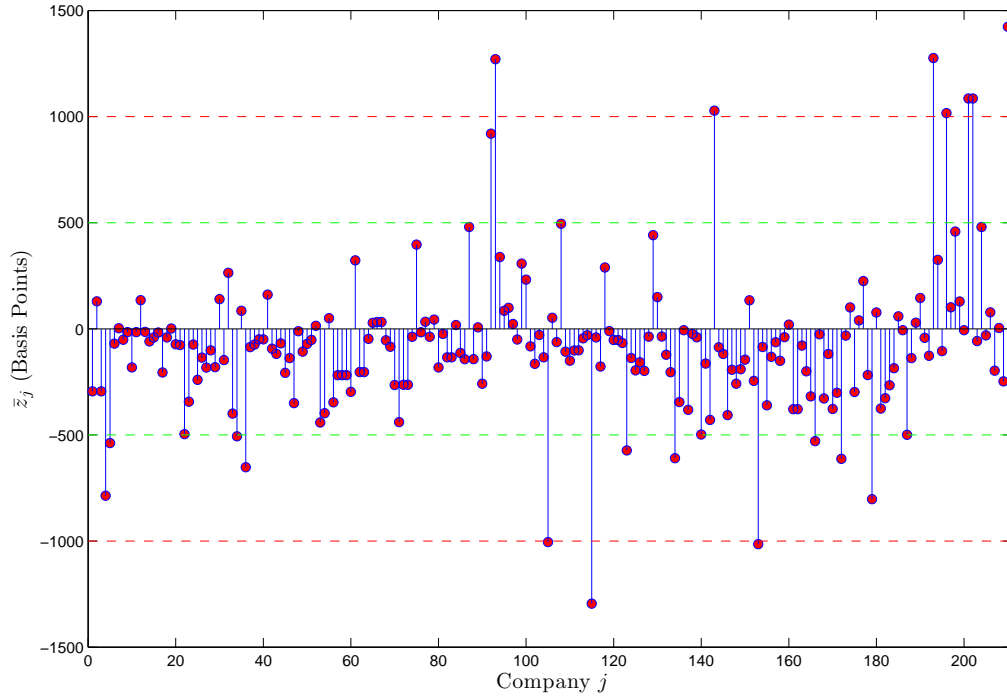**Figure 3.19:** Vector of observations $\bar{\mathbf{z}}$ (OU process for Hazard Rates): Stem plot.

Figure 3.20: Vector of observations **z** (OU process for Hazard Rates with the convexity adjustment): Stem plot.

Finally, using this method we take into account the impact of the convexity adjustment. Table 3.29 reports the descriptive statistics for the vector of observations $\bar{\mathbf{z}}$ while Figure 3.20 gives a graphical representation. In this case the vector $\bar{\mathbf{z}}$ has 155 negative components out of 213 (72.77%).

Table 3.29: Descriptive statistics for the vector of observations $\bar{\mathbf{z}}$ (OU process for Hazard Rates with the convexity adjustment).

Observations	mean	median	std	min	max	95% Range	99% Range
213	-40.6761	-76.3784	445.3035	-1294.75	2592.18	1947.012	3505.057

As a summary of the results of this analysis we include Table 3.30, in which for each method are reported how many companies have a $\bar{z}_j$ positive (negative), bigger than 500 bp (smaller than -500 bp) and bigger than 1000 bp (smaller than -1000 bp).

In Table 3.31 we report a list of the companies with $\bar{z}_j < -500$ bp (Panel A) and those with $\bar{z}_j > 500$ bp (Panel B) for *all* the methods used. In Table 3.32 we do the same for the -1000 and 1000 bp bounds.

In Table 3.33 we report for each method the five companies with the most negative $\bar{z}_j$ (Panel A) and the first five companies with the biggest $\bar{z}_j$ (Panel B).

3.7   Conclusions

In this chapter a large database of single-name CDS premia has been used to produce the corresponding CM-CDS prices. In order to derive the participation rate needed to calculate the CMCDS prices, we implemented

Table 3.30: This table reports for each method how many companies have a $\bar{z}_j$ positive (negative), bigger than 500 bp (smaller than -500 bp) and bigger than 1000 bp (smaller than -1000 bp).

	NS	lambda	OU	OU conv
pos	53	30	26	58
neg	160	183	187	155
> 500 bp	7	6	5	11
< -500 bp	13	14	16	12
> 1000 bp	2	2	2	10
< -1000 bp	3	3	3	3

Table 3.31: Panel A reports for each method the companies for which $\bar{z}_j < -500$ bp according to all methods. Panel B reports for each method the companies for which $\bar{z}_j > 500$ bp according to all methods.

Panel A	Panel B
Aetna Inc.	Ford Mtr Co
Agrium Inc	Ford Mtr Cr Co
CNA Finl Corp	Gen Mtrs Corp
Hasbro Inc	Toys R Us Inc
J C Penney Co Inc	Williams Cos Inc
LA Pac Corp	
Mattel Inc	
Pennzoil Quaker St Co	
SUPERVALU INC	
ServiceMaster Co	

Table 3.32: Panel A reports for each method the companies for which $\bar{z}_j < -1000$ bp according to all methods. Panel B reports for each method the companies for which $\bar{z}_j > 1000$ bp according to all methods.

Panel A	Panel B
Hasbro Inc	Gen Mtrs Corp
J C Penney Co Inc	Williams Cos Inc
Pennzoil Quaker St Co	

Table 3.33: Panel A reports for each method the five companies with the most negative $\bar{z}_j$. Panel B reports for each method the five companies with the biggest $\bar{z}_j$.

Panel A

NS	lambda	OU	OU conv
J C Penney Co Inc	Hasbro Inc	J C Penney Co Inc	J C Penney Co Inc
Hasbro Inc	Pennzoil Quaker St Co	Hasbro Inc	Pennzoil Quaker St Co
Pennzoil Quaker St Co	J C Penney Co Inc	Pennzoil Quaker St Co	Hasbro Inc
Aetna Inc.	ServiceMaster Co	ServiceMaster Co	ServiceMaster Co
ServiceMaster Co	Aetna Inc.	Aetna Inc.	Aetna Inc.

Panel B

NS	lambda	OU	OU conv
Ford Mtr Cr Co	Ford Mtr Cr Co	Ford Mtr Cr Co	Textron Inc
Toys R Us Inc	Toys R Us Inc	Toys R Us Inc	Wyeth
Ford Mtr Co	Ford Mtr Co	Ford Mtr Co	Gen Mtrs Corp
Williams Cos Inc	Williams Cos Inc	Williams Cos Inc	Williams Cos Inc
Gen Mtrs Corp	Gen Mtrs Corp	Gen Mtrs Corp	Wells Fargo & Co

both parametric (the Nelson-Siegel interpolation and the hazard rates described by an Ornstein-Uhlenbeck process) and nonparametric methods (piecewise constant hazard rates). For each day and for each name all these methods utilize the term structure of single-name CDS prices along with information regarding the recovery rate and the discount factors bootstrapped from the Libor/Swap rates to return the corresponding participation rate. This allowed us to build a database of single-name CMCDS premia that was used to identify possible imbalances that may exist in the credit markets when pairing CDS and CMCDS on the same name. The general idea is to form a swap type of trading strategy whereby a fixed premium payment is netted against a floating one, both representing protection premia against default. This strategy has the advantage that default risk is eliminated and only counterparty risk is taken. We have then computed for all the companies for which we have the required data the profit (or loss) that an investor would have realized being long a CMCDS with maturity 5 years and short a CDS with both the contracts initiated on 20/09/2001. Then we have done the same analysis for a contract initiated between the beginning of our sample and 01/11/2001. In both cases and for each method implemented we have reported for how many companies the above strategy produces a gain or a loss bigger than 500 or 1000 bp. It appears that, in general, it would have been more profitable to sell CDS and to buy CMCDS. Considering all the methods implemented, at least 85% of the names analysed had a negative cumulative net trading profit/loss over the 5 years period considered. Also when a convexity adjustment has been taken into account, we reached substantially the same conclusion. The method which involves convexity adjustment seems only to reduce the loss that an investor would have incurred in by buying CDS and selling CMCDS over the 5 years period considered. The percentage of names for which the above strategy would have led to a loss, however, does not change dramatically. Also a cap on the floating payment was introduced, but this did not change the results. We have also reported the names for which the strategy gives a gain or a loss bigger than 500 or 1000 bp no matter what the method used to compute the participation rate. Interestingly the strategy involving the company "Ford Mtr Co" leads to a profit bigger than 1000 bp in the first analysis and between 500 bp and 1000 bp for the second analysis. On the other hand, for both the analysis considered, the trading strategy involving the company "Gen Mtrs Corp" leads to a profit bigger than 1000 bp for all the methods implemented. In general, these two companies appear among those with the biggest profit, in both the first and the second statistical arbitrage analysis.

4
True vs. Spurious Long Memory in Credit Default Swaps and Credit Spreads

4.1 Introduction

As pointed out by Dolado et al. (2005) and by Mikosch and Stărică (2004) for the case of multiple breaks, the long memory feature of financial time series could be a spurious effect caused by the presence of one or more structural breaks. The first study presents a simple example to illustrate the source of confusion between a long memory process and a short memory one subject to structural breaks. Consider the sequence $\{Y_t\}_{t=1,\dots,T}$ with the following data generating process

$$Y_t = \alpha_1 + (\alpha_2 - \alpha_1)DU_t(\lambda) + u_t = \begin{cases} \alpha_1 + u_t & t \le T_B \\ \alpha_2 + u_t & t > T_B \end{cases} \tag{4.1}$$

where u_t is a zero-mean $I(0)$ process with autocovariances $\gamma_u(j)$ and $DU_t(\lambda) = I_{(t>T_B)} \equiv I_{(t>\lambda T)}$ with λ the fraction of the sample where the break occurs and T_B the date of the break.

The sample ACF of an $I(d)$ and an $I(0)$ with a structural break processes can be found in Figure 4.1.

A straightforward application of the ergodic theorem allows us to derive the asymptotic behavior of the sample autocovariance of $\{Y_t\}$

$$\hat{\gamma}_Y(j) = \frac{1}{T} \sum_{t=j+1}^{T} (Y_t - Y_{t-j}) - (\bar{y}_T)^2 \to \gamma_u(j) + \lambda(1-\lambda)(\alpha_2 - \alpha_1)^2 \quad \text{a.s.}$$

and thus, even though $\gamma_u(j)$ approaches 0 as $j \uparrow \infty$ because u_t is $I(0)$,

$$\lim_{j \uparrow \infty} \hat{\gamma}_y(j) = \lambda(1-\lambda)(\alpha_2 - \alpha_1)^2.$$

As long as $\alpha_2 \neq \alpha_1$, the limit is not zero and thus the autocovariance function of the process mimics a slow hyperbolic decay characteristic of a long memory process. To verify empirically the validity of this result we present a small Monte Carlo study in which 5000 series of sample size $T = 2500$ are generated according to (4.1). We set $\alpha_1 = 0$, $\lambda = 0.5$ (so that the break comes in the middle of the sample) and $u_t \sim n.i.d.(0,1)$. Three different break sizes are considered: $\alpha_2 - \alpha_1 = 0.0$ (no break), $\alpha_2 - \alpha_1 = 0.3$ (small break) and $\alpha_2 - \alpha_1 = 0.6$ (large break). In Table 4.1 the estimates of the long memory parameter d for different value of he bandwidth parameter and for different sizes of the break are reported. From the table we can see that the estimates of d increase monotonically with the size of the shift in the mean, and, when the break is present, they seem to increase as the bandwidth parameter decreases[1].

---

[1]The possibility of confusing spurious long memory with genuine long memory aggravates even more when the DGP contains a break in the trend. A Monte Carlo experiment in Dolado et al. (2005) using the GDP

$$Y_t = \alpha_1 + \beta_1(t - T_B)I_{(T_B+1 \le t \le T)} + u_t$$

with $\beta_1 = 0.1$, yields estimates of d in the range $(1.008, 1.0310)$.

Figure 4.1: Sample ACF of an $I(d)$ and an $I(0)$ with a structural break processes.
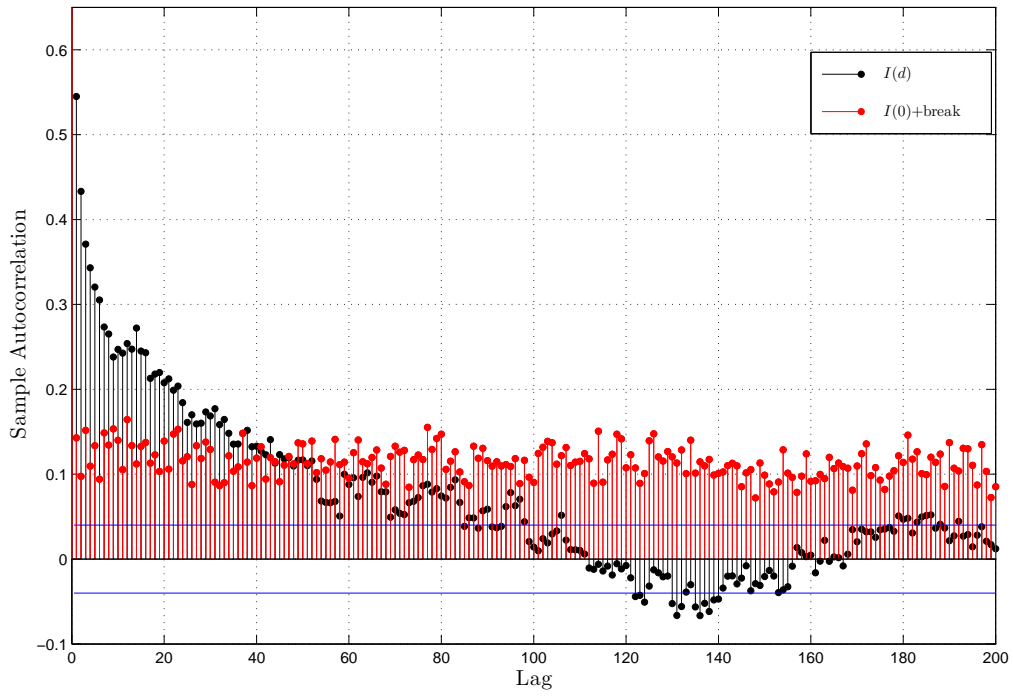
Table 4.1: GPH estimates of the parameter d in (4.1) for different values of the break. One asterisk denotes significance at 5% level and two asterisks denote significance at 1% level for the null hypothesis $d = 0$.

	$T^{0.5}$	$T^{0.45}$	$T^{0.4}$	$T^{0.35}$
$\alpha_2 - \alpha_1 = 0.0$	-0.0015	-0.0013	-0.0033	-0.0027
$\alpha_2 - \alpha_1 = 0.3$	0.199*	0.2873*	0.3788**	0.4811**
$\alpha_2 - \alpha_1 = 0.6$	0.3226**	0.4388**	0.5482**	0.6589**

The fact that long memory and structural change are easily confused is pointed out in many papers in the econometric literature. Mayoral (2006) develops a time-domain test of $I(d)$ versus $I(0)$ plus trends and/or breaks, and reports an empirical application that analyzes US inflation. The conclusions are that the null of $I(d)$ cannot be rejected. Diebold and Inoue (2001) provide both theoretical justification and Monte Carlo evidence to support the claim that a time series with structural breaks can induce a strong persistence in the autocorrelation function and hence generate spurious long memory. In Granger and Hyung (2004) two time series models, an occasional-break model and an $I(d)$ model to analyze S&P 500 absolute stock returns are compared. In their empirical analysis, the authors find that both the models can equally well explain the absolute stock returns series and that long memory models have better out-of-sample forecasting performance than the occasional break models, but the evidence is statistically insignificant. Choi and Zivot (2007) estimate the long memory parameter d monthly forward discount series for five G7 countries after adjusting for breaks in their mean. They find that multiple breaks in the mean are present and that after adjusting for structural breaks, the persistence in the forward discount is considerably reduced. However, they find that when one allows for structural breaks there is still evidence of long memory. Ohanissian et al. (2007) develop a test to distinguish between true long memory and spurious long memory based on invariance of the long memory parameter for temporal aggregates of the process under the null of true long memory and find that the long memory property in exchange rate volatility is generated by a true long memory process. Dolado et al. (2005) proposes a time-domain test of a process being $I(d)$ under the null, against the alternative of being $I(0)$ with deterministic components subject to structural breaks at known or unknown dates. The proposed test is used for two empirical applications. The first one deals with the log of the time series of U.S. real GNP for the period 1869 to 2001. The authors conclude that for this data the null of $I(d)$ cannot be rejected at the conventional significance levels. The second empirical application employs the test to study the long memory properties of returns from the S&P 500 composite stock index over the period 1953–1977. They conclude that the null hypothesis of $I(d)$ cannot be rejected for moderate values of d when either the absolute values or squares of the returns are considered. Two simple tests for true versus spurious long memory are proposed also in Shimotsu (2006). The tests are applied to the daily realized volatility of the S&P 500 index to show some evidence of infrequent structural breaks in the data. The results, however, don't provide strong evidence against the hypothesis of true long memory.

In this chapter several tests recently proposed in the econometric literature to detect if a time series is truly long memory are illustrated and then applied to the data of chapter 2 and to the CDS data used in chapter 3.

4.2   Statistical Tests

4.2.1   A Test Based on Temporal Aggregation

A direct implication of scaling and self-similarity properties of the fractional Brownian motion (see Mandelbrot and Van Ness, 1968) is that the process will have the same memory at all levels of sampling frequency. This result is confirmed by Beran and Ocker (2000) who study Gaussian ARFIMA(p, d, q) models. They show that aggregation acts to modify only the short memory properties whereas the long memory parameter does not change across temporal aggregation. The same ideas are employed in Andersen and Bollerslev (1997) and Andersen et al. (2001) in the study of realized volatility to conclude that their estimated long memory parameters, using different frequency data don't change too much, implying true long memory. A formal test procedure which exploits the invariance of the long memory parameter for temporal aggregates of the process under the null of true long memory is proposed by Ohanissian et al. (2007). Their test is applicable to a stationary mean zero Gaussian long-memory time series $\{Y_t\}$. As usual, the spectral density is assumed to be $f(\lambda) = |1 - e^{-i\lambda}|^{-2d} f^*(\lambda)$, $0 < d < 0.5$, where $f^*(\cdot)$ denotes the spectral density of the short memory component of $\{Y_t\}_{t=1,...,T}$ and is assumed to be continuous, bounded above, bounded away from zero, twice differentiable with the second derivative bounded in a neighborhood of zero. This assumption allows the use of the GPH estimator which is used in order to avoid many potential mis-specification issues involving the short memory component. Given the integer n such that T/n is integer, the n-period non-overlapping

aggregates of the time series Y_t are defined as

$$Z_s^{(n)} = \sum_{\tau=1}^{n} Y_{n(s-1)+\tau} \quad \text{for } 1 \le s \le \frac{T}{n}.$$

With the convention that a superscript (n) on any statistic represents the corresponding statistic for the n-temporally aggregated data, let

$$I_j^{(n)} = \frac{n}{2\pi T} \left| \sum_{s=1}^{T/n} Z_s^{(n)} \exp\left(\mathrm{i} s 2\pi \frac{jn}{T} \right) \right|^2.$$

denote the periodogram of an n-temporally aggregated series. Note that this can be rewritten as

$$I_j^{(n)} = \frac{n}{2\pi T} \left| \sum_{s=1}^{T} Y_s \exp\left(\mathrm{i} \left[\frac{s}{n} \right] 2\pi \frac{jn}{T} \right) \right|^2$$

$$= \frac{n}{2\pi T} \sum_{s=1}^{T} \sum_{t=1}^{T} Y_s Y_t \cos\left(2\pi \frac{jn}{T} \left(\left[\frac{s}{n} \right] - \left[\frac{t}{n} \right] \right) \right).$$

This means that $I_j^{(n)}$ can be written in terms of the periodogram of the entire series, $I_j^{(1)}$:

$$I_j^{(n)} = n I_j^{(1)} + Y' B_j Y$$

where $Y = (Y_1, \ldots, Y_T)'$, $B_j = [b_j(s,t)]_{1 \le s,t \le T}$ and

$$b_j(s,t) = \frac{n}{2\pi T} \left(\cos\left(2\pi \frac{jn}{T} \left(\left[\frac{s}{n} \right] - \left[\frac{t}{n} \right] \right) \right) - \cos\left(2\pi \frac{j}{T}(s-t) \right) \right).$$

Given the bandwidth parameter $m^{(n)}$, the GPH estimate of the long memory paramenter for the n-temporally aggregated data is

$$\hat{d}^{(n)} = -\frac{1}{2 \sum_{j=1}^{m^{(n)}} \left(a_j^{(n)} \right)^2} \sum_{j=1}^{m^{(n)}} a_j^{(n)} \log I_j^{(n)},$$

with

$$a_j^{(n)} = \log \left| 2 \sin\left(\pi \frac{jn}{T} \right) \right| - \frac{1}{m^{(n)}} \sum_{j=1}^{m^{(n)}} \log \left| 2 \sin\left(\pi \frac{jn}{T} \right) \right|.$$

Note that number of ordinates used in the GPH estimator depends on the length of the series and since temporal aggregation decreases the length of the series, we have

$$m^{(n_1)} > m^{(n_2)} \quad \text{for } n_1 < n_2$$

and also the frequencies are note the same:

$$\lambda_j^{(n_1)} \ne \lambda_j^{(n_2)} \quad \text{for } n_1 \ne n_2$$

and $\lambda_j^{(kn)} = \lambda_{kj}^{(n)}$.

    Let N denote the fixed, but arbitrarily large, number of aggregation levels and $(n_1, n_2, \ldots, n_N)$ denote the fixed, but arbitrarily large, aggregation levels for the N series such that $n_1 < n_2 < \ldots < n_N$. In Ohanissian et al. (2007) it is shown that, if for any fixed, but arbitrarily large, aggregation level n, the growth rate of

the number of ordinates used for the GPH estimation is such that $m^{(n)} = o(T^{\frac{2-4d}{3}})$ as $T \to \infty$, then the joint distribution of the GPH estimates obtained using the temporally aggregated series is asymptotically normal. Furthermore, they prove that, asymptotically, the covariance between any two GPH estimates obtained using temporally aggregated series equals the variance of the lesser aggregated series:

$$4m^{(n_i)} \left(\text{Cov}(\hat{d}^{(n_i)}, \hat{d}^{(n_j)}) - \text{Var}(\hat{d}^{(n_i)}) \right) = o(1) \quad \text{as } T \to \infty \quad \text{for } 1 \le i < j \le N. \tag{4.2}$$

Given that the GPH is known to have an asymptotic normal distribution:

$$\sqrt{4m^{(n_i)}}(\hat{d}^{(n_i)} - d) \to_d N\left(0, \frac{\pi^2}{6}\right)$$

eq. (4.2) allows us to compute the theoretical covariance matrix. When the time series of interest is relatively small, as confirmed by a simulation study in Deo and Hurvich (2001), the approximation suggested by Geweke and Porter-Hudak (1983) should be used:

$$\text{Var}_{\text{approx.}}(\hat{d}^{(n_i)} - d) = \left[24 \sum_{j=1}^{m^{(n_i)}} \left(a_j^{(n_i)}\right)^2 \right]^{-1} \pi^2.$$

Now, let $\widehat{\mathbf{d}}_N = (\hat{d}^{(n_1)}, \hat{d}^{(n_2)}, \ldots, \hat{d}^{(n_N)})'$ be the N-dimensional vector of our estimated long memory parameters, and $\mathbf{d}_N = (d^{(n_1)}, d^{(n_2)}, \ldots, d^{(n_N)})'$ be the constant N-dimensional vector of the actual long memory parameters. The null hypothesis is

$$\text{H}_0 : d^{(n_1)} = d^{(n_2)} = \ldots = d^{(n_N)} = d$$

and it can be tested by considering the quadratic form

$$\left(\widehat{\mathbf{d}}_N - \mathbf{d}_N\right)' \Lambda^{-1} \left(\widehat{\mathbf{d}}_N - \mathbf{d}_N\right)$$

where Λ is the asymptotic covariance matrix whose structure was discussed above. Note that in order for the asymptotic covariance matrix to be invertible the asymptotic variance of each individual GPH estimate must be different. This can be achieved by using a different number of ordinates for the estimation of each temporally aggregated series. The test statistic above is asymptotically distributed as $\chi^2(N)$. However, in practice, $\mathbf{d}$ will not be known so the mean value of the estimates is used. Hence the test statistic to implemented is

$$\widehat{W}_N = (P\widehat{\mathbf{d}}_N)'(P\Lambda P')^{-1}(P\widehat{\mathbf{d}}_N)$$

where

$$P = \begin{pmatrix} 1 - \frac{1}{N} & -\frac{1}{N} & -\frac{1}{N} & \cdots & -\frac{1}{N} & -\frac{1}{N} \\ -\frac{1}{N} & 1 - \frac{1}{N} & -\frac{1}{N} & \cdots & -\frac{1}{N} & -\frac{1}{N} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ -\frac{1}{N} & -\frac{1}{N} & -\frac{1}{N} & \cdots & 1 - \frac{1}{N} & -\frac{1}{N} \end{pmatrix}.$$

Since $P\widehat{\mathbf{d}}$ is $N(0, P\Lambda P')$, the test statistic $\widehat{W}$ has an asymptotic $\chi^2(N-1)$ distribution under H_0. In a simulation study Ohanissian et al. (2007) show that the distribution of $\widehat{\mathbf{d}}$ does not depend on the value of the long memory parameter and that there is significant size distortion when using the theoretical variances for short series. Thus, they suggest to use simulated and approximated variances instead of theoretical variances.

4.2.2   *The Structural Break–Fractional Dickey–Fuller Test*

The Structural Break–Fractional Dickey–Fuller (SB–FDF) Test is developed in Dolado et al. (2005). The null is that the process of interest is $I(d)$, $d \in (0,1)$. Under the alternative the process is $I(0)$ with one structural break. To account for structural breaks, the process is assumed to be described by the following equation:

$$Y_t = A_B(t) + \frac{a_t I_{(t>0)}}{\Delta^d - \phi L} \tag{4.3}$$

where $A_B(t)$ is a linear deterministic trend function that may contain breaks at the unknown date T_B, a_t is a stationary $I(0)$ process, L the lag operator and $\Delta = 1 - L$. The null now corresponds to the case $\phi = 0$, whereas $\phi < 0$ means that the process is $I(0)$ and it is subject to the regime shifts defined by $A_B(t)$. Rearranging (4.3) yields:

$$\Delta^d Y_t = \phi Y_{t-1} + \Delta^d A_B(t) - \phi A_B(t-1) + \varepsilon_t \tag{4.4}$$

where $\varepsilon_t = a_t I_{(t>0)}$. The SB–FDF test of $I(d)$ vs. $I(0)$ in the presence of structural breaks is simply given by the t-ratio of the coefficient ϕ in the regression model (4.4). The definition of $A_B(t)$ we consider is[2]

$$A_B(t) = \mu_0 + (\mu_1 - \mu_0)DU_t(\lambda)$$

which corresponds to the so-called *crash* hypothesis. Dolado et al. (2005) prove that for a process generated by (4.3) with $a_t \sim i.i.d(0, \sigma^2)$ the asymptotic distribution of the test statistic under the null of $\phi = 0$, when ϕ is estimated by OLS, is given by:

$$t_{\hat{\phi}(\lambda)} \to_d \begin{cases} N(0,1) & \text{if } d \in (0, 0.5) \\ \dfrac{\int_0^1 B_d^*(\lambda, r) \mathrm{d}B(r)}{\left(\int_0^1 B_d^*(\lambda, r)^2 \mathrm{d}r\right)^{1/2}} & \text{if } d \in (0.5, 1) \end{cases}$$

where $B_d^*(\lambda, r)$ is the L_2 projection residual from the corresponding continuous time regressions associated to model (4.4):

$$B_d(r) = \hat{\alpha}_0 + \hat{\alpha}_1 du(\lambda, r) + \hat{\alpha}_2 r^{-d} + \hat{\alpha}_3 r^{-d} du(\lambda, r) B_d^*(\lambda, r)$$

where $B_d(r)$ is Type–I fBM as defined in Marinucci and Robinson (1999) and $du(\lambda, r) = 1$ if $r > \lambda$ and 0 otherwise.

When, more realistically, the break date T_B is unknown, the idea is to choose the break point that gives the least favorable result for the null hypothesis of $I(d)$ using the SB-FDF test in (4.4). Therefore, using the same range of Andrews (1993) the t-statistic of the coefficient ϕ is computed for the values of $\lambda \in (0.15, 0.85)$ and the most negative value is chosen:

$$\widehat{t}_{\hat{\phi}} = \inf_{\lambda \in (0.15, 0.85)} t_{\hat{\phi}(\lambda)}. \tag{4.5}$$

Again, Dolado et al. (2005) show that under the null and when $\mu_0 = \mu_1$ so that $\Delta^d(y_t - \mu_0) = a_t$,

$$\widehat{t}_{\hat{\phi}} \to_d \begin{cases} N(0,1) & \text{if } d \in (0, 0.5) \\ \inf_{\lambda \in (0.15, 0.85)} \dfrac{\int_0^1 B_d^*(\lambda, r) \mathrm{d}B(r)}{\left(\int_0^1 B_d^*(\lambda, r)^2 \mathrm{d}r\right)^{1/2}} & \text{if } d \in (0.5, 1) \end{cases}.$$

The simulated critical values are provided by Dolado et al. (2005) in their Appendix B. The test reject the null of genuine long memory when

$$\widehat{t}_{\hat{\phi}} < k_{\inf, \alpha},$$

where $k_{\inf, \alpha}$ is the simulated critical value at the significance level α.

4.2.3   A Test Based on Sample Splitting

A formal statistical test based on sample splitting is proposed by Shimotsu (2006). Under the null of true long memory each subsample of the time series also follows an $I(d)$ process with the same value of the long memory parameter d. Hence a simple test can be constructed by splitting the sample into b subsamples,

---

[2] In Dolado et al. (2005) also the *changing growth* hypothesis

$$A_B(t) = \mu_0 + \beta_0 t + (\beta_1 - \beta_0)(t - T_B) I_{(T_B+1 \leq t \leq T)}$$

and a combination of the crash and changing growth hypothesis

$$A_B(t) = \mu_0 + \beta_0 t + (\mu_1 - \mu_0)DU_t(\lambda) + (\beta_1 - \beta_0) I_{(T_B+1 \leq t \leq T)}$$

are considered.

estimating d for each subsample and computing how these estimates differ from the full sample estimate. Assuming that both b and T/b are integer, the sample $\{Y_s\}_{s=1,\ldots,T}$ is split into b blocks, so that each block has T/b observations. The author suggests to compute the local Whittle estimator for each block of observations using m/b periodogram ordinates, m being the number of the periodogram ordinates used for the estimation of d in the full sample. More formally, let $\tilde{d}$ be the local Whittle estimator of $d \in (-0.5, 0.5)$ computed from the whole sample and $\tilde{d}^{(a)}$, $a = 1, \ldots, b$, be the local Whittle estimator of d computed from the ath block of the observations, $\{Y_t : t = (a-1)T/b + 1, \ldots, aT/b\}$:

$$\tilde{d}^{(a)} = \arg \min_{-\frac{1}{2} < d < \frac{1}{2}} \left\{ \log \left[\frac{b}{m} \sum_{j=1}^{m/b} \varpi_j^{2d} I_j^{(a)} \right] - 2d \frac{b}{m} \sum_{j=1}^{m/b} \log \varpi_j \right\} \quad a = 1, \ldots, b$$

$$\varpi_j = 2\pi \frac{jb}{T} \quad j = 1, \ldots, T/b$$

$$I_j^{(a)} = \frac{b}{2\pi T} \left| \sum_{s=(a-1)T/b+1}^{aT/b} Y_s \exp\left(\mathrm{i}s\varpi_j\right) \right|^2.$$

The test statistic for the null of genuine long memory is then given by

$$\widetilde{W}_b = (A\widehat{\mathbf{d}}_b)'(A\Upsilon A')^{-1}(A\widehat{\mathbf{d}}_b)$$

where $\widetilde{\mathbf{d}}_b$ is a the $b+1$ vector

$$\widetilde{\mathbf{d}}_b = \begin{pmatrix} \tilde{d} - d \\ \tilde{d}^{(1)} - d \\ \vdots \\ \tilde{d}^{(b)} - d \end{pmatrix}$$

Υ is the covariance matrix of $\widetilde{\mathbf{d}}_b$ and A is the $b \times (b+1)$ matrix

$$A = \begin{pmatrix} 1 & -1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & \cdots & -1 \end{pmatrix}.$$

The test statistic $\widetilde{W}_b$ has a chi-squared limiting distribution with $b-1$ degrees of freedom.

The covariance matrix Υ is derived in a simulation exercise in which 2000 series of dimension $T = 2500$ following an $I(4)$ process[3] have been generated.

In Table 4.2–Table 4.3 we report the simulated covariance matrices $A\Upsilon A'$ corresponding to different levels of the bandwidth parameter m and for $b = 2$ or $b = 4$.

4.2.4 Test using dth Differencing

This test developed in Shimotsu (2006) exploits the fact that if an $I(d)$ process is differenced d times, then the resulting time series is trivially an $I(0)$ process but this is not the case when spurious long-memory processes are taken into account. To implement the test, thus, it suffices to demean the data and apply the Phillips-Perron (PP) unit root test or the KPSS test to its $\hat{d}$th difference, $\hat{d}$ being a consistent estimate of d. Some care is needed in demeaning the data. Assuming that Y_t follows a truncated $I(d)$ process with initialization at $t = 0$:

$$Y_t - \mu = (1-L)^{-d} u_t I_{(t \geq 1)}$$

---

[3]This exercise has also shown that the particular choice of d in the $I(d)$ process used in the simulations does not change the covariance matrix of $\widetilde{\mathbf{d}}_b$ significantly.

Table 4.2: Covariance matrices $A\Upsilon A'$ corresponding to different values of the bandwidth parameter m for $b = 2$.

$m = T^{0.5}$	$m = T^{0.45}$
$\begin{pmatrix} 0.010603 & -0.00693 \\ -0.00693 & 0.010701 \end{pmatrix}$	$\begin{pmatrix} 0.01854 & -0.01086 \\ -0.01086 & 0.019642 \end{pmatrix}$

$m = T^{0.4}$	$m = T^{0.35}$
$\begin{pmatrix} 0.034114 & -0.01774 \\ -0.01774 & 0.037391 \end{pmatrix}$	$\begin{pmatrix} 0.078889 & -0.03246 \\ -0.03246 & 0.085082 \end{pmatrix}$

Table 4.3: Covariance matrices $A\Upsilon A'$ corresponding to different values of the bandwidth parameter m for $b = 4$.

$m = T^{0.5}$

$$\begin{pmatrix} 0.042856 & -0.00649 & -0.00757 & -0.00644 \\ -0.00649 & 0.039942 & -0.00732 & -0.00691 \\ -0.00757 & -0.00732 & 0.042047 & -0.00641 \\ -0.00644 & -0.00691 & -0.00641 & 0.04403 \end{pmatrix}$$

$m = T^{0.45}$

$$\begin{pmatrix} 0.082302 & -0.01039 & -0.00948 & -0.01082 \\ -0.01039 & 0.078555 & -0.00894 & -0.01124 \\ -0.00948 & -0.00894 & 0.0821 & -0.01181 \\ -0.01082 & -0.01124 & -0.01181 & 0.078604 \end{pmatrix}$$

$m = T^{0.4}$

$$\begin{pmatrix} 0.196986 & -0.02074 & -0.01317 & -0.01896 \\ -0.02074 & 0.198046 & -0.01792 & -0.03186 \\ -0.01317 & -0.01792 & 0.204205 & -0.01246 \\ -0.01896 & -0.03186 & -0.01246 & 0.187643 \end{pmatrix}$$

$m = T^{0.35}$

$$\begin{pmatrix} 0.628313 & -0.02831 & -0.04251 & -0.02825 \\ -0.02831 & 0.610207 & -0.04966 & -0.0399 \\ -0.04251 & -0.04966 & 0.688979 & -0.03042 \\ -0.02825 & -0.0399 & -0.03042 & 0.577297 \end{pmatrix}$$

where μ is the mean Y_t when $d < 1/2$, we have $T^{-1} \sum_{t=1}^{T} Y_t - \mu = O_P(T^{d-1/2})$ and as shown by Shimotsu (2006)

$$(1 - L)^{-d} \left(Y_t - T^{-1} \sum_{t=1}^{T} Y_t \right) = u_t + O_P(T^{d-1/2} t^{-d}).$$

Hence, If $d \geq 1$, the second term on the right has a nonnegligible effect on the sample statistics of the dth differenced demeaned data.

To circumvent the problem, the author suggests to estimate μ using a linear combination of the sample mean and Y_1:

$$\hat{\mu}(d) = \frac{w(d)}{T} \sum_{t=1}^{T} Y_t + (1 - w(d))Y_1 \tag{4.6}$$

where $w(d)$ is a is a smooth weight function such that $w(d) = 1$ for $d \leq 1/2$ and $w(d) = 0$ for $d \geq 3/4$. One choice of the weight function could be

$$w(d) = \begin{cases} 1 & \text{for } d \leq 1/2 \\ 1/2 \left(1 + \cos(4\pi d)\right) & \text{for } d \in (1/2, 3/4) \\ 0 & \text{for } d \geq 3/4 \end{cases} .$$

Once the data has been demeaned using (4.6) to estimate the mean, it is possible to compute the $\hat{d}$th differenced series u_t:

$$U_t = (1 - L)^{\hat{d}}(Y_t - \hat{\mu}(\hat{d})) = \sum_{k=0}^{t-1} \frac{\Gamma(k - \hat{d})}{\Gamma(-\hat{d})k!}(Y_{t-k} - \hat{\mu}(\hat{d})).$$

Finally we can calculate the test statistics on the series U_t. The test regression for the PP tests is

$$\Delta U_t = c + \pi U_{t-1} + \xi_t$$

where ξ_t is $I(0)$ and may be heteroskedastic. The test statistic is constructed in a way that any serial correlation and heteroskedasticity in the errors ξ_t are corrected for. The test statistic is

$$Z_t = t_{\pi=0} \times \left(\frac{\hat{\sigma}^2}{\hat{\lambda}^2} \right)^{1/2} - \frac{\hat{\lambda}^2 - \hat{\sigma}^2}{2\hat{\lambda}^2} \frac{SE(\hat{\pi})}{\hat{\sigma}^2/T} \tag{4.7}$$

The quantities $\hat{\sigma}^2$ and $\hat{\lambda}^2$ are consistent estimates of the variance parameters

$$\sigma^2 = \lim_{T \to \infty} T^{-1} \sum_{t=1}^{T} \mathbb{E}(\xi_t^2)$$

$$\lambda^2 = \sum_{j=-\infty}^{+\infty} \mathbb{E}(\xi_t \xi_{t-j})$$

and are given by the sample variance of the least squares residual $\hat{\xi}_t$, and the Newey-West estimate of the long-run variance of ξ_t using $\hat{\xi}_t$ respectively:

$$\hat{\sigma}^2 = \frac{1}{T} \sum_{t=1}^{T} e_t^2$$

$$\hat{\lambda}^2 = \frac{1}{T} \sum_{t=1}^{T} e_t^2 + \frac{2}{T} \sum_{s=1}^{l} \left(1 - \frac{s}{l+1} \right) \sum_{t=s+1}^{T} e_t e_{t-s}$$

with

$$e_t = \hat{\xi}_t - \frac{1}{T}\sum_{t=1}^{T}\hat{\xi}_t.$$

Using the same notation, the KPSS test statistic is given by

$$\hat{\eta}_\mu = \frac{\sum_{t=1}^{T}S_t^2}{T^2\hat{\lambda}^2}, \qquad S_t = \sum_{k=1}^{t}e_k. \tag{4.8}$$

4.3   A Monte Carlo Study

In this section we compare the four different tests for the null of true long memory versus the alternative of spurious long memory induced by a structural break, using simulated data. We simulate 5000 series of dimension $T = 2500$ using the three data generating processes:

1. An $I(0.4)$ process

2. The process (4.1) with $\alpha_1 = 0$, $\alpha_2 = 0.3$ and $T_B = T/2$

3. The process (4.1) with $\alpha_1 = 0$, $\alpha_2 = 0.6$ and $T_B = T/2$

In brief, we consider a genuine long memory process, a process with a small break and one with a large break in the middle of the sample. As confirmed by Table 4.1, the last two are spurious long memory processes.

Table 4.4 reports the rejection frequencies for the four test using a 1% confidence level. Every time the GPH estimation is involved the bandwidth parameter used is $m_j = \left[\sqrt{T_j}\right]$, where T_j is the number of observations in the (sub)sample. For the "Temporal Aggregation" test of Ohanissian et al. (2007) the number of aggregation levels is set to $N = 4$ and the aggregation levels are $n_j = 2^j$ for $j = 0, \ldots, N-1$. For the "Sample Splitting" test of Shimotsu (2006) the number of subsamples considered is $b = 4$.

Table 4.4: Rejection frequencies at a 1% confidence level. For each test the null is true long memory. "Temporal Aggregation" denotes the test of Ohanissian et al. (2007) described in subsection 4.2.1. "SB–FDF" denotes the test developed in Dolado et al. (2005) and described in subsection 4.2.2. Finally the last two columns report the rejection frequencies for the two tests of Shimotsu (2006) (subsection 4.2.3 and subsection 4.2.4). In the latter case the KPSS test is used. Every time the GPH estimation is involved the bandwidth parameter used is $m_j = \left[\sqrt{T_j}\right]$, T_j being the number of observations in the (sub)sample.

	Temporal Aggregation	SB–FDF	Sample Splitting	dth Differencing
$I(0.4)$	0.0174	0.395	0.0438	0.0688
Small Break	0.0454	0.6002	0.7752	0.8766
Large Break	0.0554	0.873	0.9978	0.9484

The first thing to notice in the table is that for each method the number of rejections, correctly, increases moving from the top to the bottom of the table. From the table it is also clear that overall the test based on sample splitting has the best performance in terms of both power and size: in the experiment, in fact, it is the one with the smallest rejection frequency when the DGP is the one of a genuine long memory process and is the one with the rejection frequency closest to 99% when a process with a large break is considered. The test based on dth differencing seems to be second best. It performs worse than the test based on sample splitting when the true long memory process and the process with a large break are considered, but it does better with

the small break process. The test based on temporal aggregation appears to have excellent size properties but it is way too conservative when the processes with a structural break are considered. The rejection frequency is only about 4.5% for the data with a small break and only about 5.5% for the data with a large break. The explanation for such a poor performance could be due to the fact that the number of aggregation levels ($N = 4$ in the experiment) considered is not big enough. However, given that the number of observations is 2500, increasing N results in a number of observations in the aggregated series which would be too limited for the GPH estimator to be valid. This is the reason why in Ohanissian et al. (2007) the simulation study is conducted on $I(0.4)$ process with a number of observations which equals 610,304. Finally the structural break–fractional Dickey–Fuller test, although performing better than the one based on temporal aggregation under the alternative hypothesis of spurious long memory (second and third row of the table), rejects too often the null even when the null is true. When the $I(0.4)$ process is considered, the null is rejected 1975 times out of 5000 (39.5%). The number of rejection grows to 3001 (60.02%) for the small break case and to 4365 (87.3%) for the large break case. Even though the rejection frequency increases significantly moving from the small to the large break case, still the two frequencies are too low if compared to tests based on sample splitting and on dth differencing. The reason for this not so brilliant performance could be due to the fact that the test is not capable to detect the structural break unless this is sufficiently large. From the table it is possible to guess that considering a break size of $\alpha_2 - \alpha_1 = 0.8$ or $\alpha_2 - \alpha_1 = 0.9$ would lead to a correct rejection frequency of 99%.

4.4   Empirical Applications

In this section the tests of section 4.2, are applied firstly to the datasets used in chapter 2, for which some evidence of long memory was found. Next the long memory properties of CDS spreads from a large number of issuers are studied and the test are applied to those series for which evidence of long memory is found.

4.4.1   Application to Credit Spreads

Since form the analysis of chapter 2 credit spreads can be described by long memory nonstationary processes, we could use the test presented in the previous section to verify whether this property is true or whether it is induced by a structural break, which is not a remote possibility, given that we are dealing with credit data. The only tests we apply are the one based on dth differencing and the SB-FDF test.

The two datasets used are

a) 30-year Historical US Treasury Constant Maturity Yields and Moody's Aaa, Aa, A and Baa Long-Term Corporate Bond Yield Averages. The data covers the period from December 1992 to November 2003, for 2703 observations.

b) Lehman Brothers Eurodollar Aaa, Aa, A and Baa Indices and U.S. Global Treasury Index. The indices include primarily corporates bonds (even though they can include government-related and securitized bonds). The data covers the period from June 1996 to July 2006, for 2613 observations.

Spreads are calculated as the difference between corporate yields and Treasury yields, as well as between different corporate yields. Thus we have 15 series: Treasury yields (denoted by T), corporate yields (Aaa, Aa, A, and Baa), spreads over Treasury (sTAaa, sTAa, sTA, and sTBaa), spreads between corporate yields (sAaaAa, sAaaA, sAaaBaa, sAaA, sAaBaa, and sABaa).

Table 4.5 reports the results for the data from Moody's and Table 4.6 the results for the Lehman Brothers Eurodollar indeces.

As far as the first dataset is concerned, we can see that the yields are genuine long memory processes according to both the tests procedures. Further, the test based on dth differencing never rejects the null of true long memory also for the credit spreads series. The SB–FDF test, on the other hand, rejects the null for the series sAaaA, sAaaBaa, sAaA and sABaa at the 1% confidence level. In Table 4.7 the fraction of the sample where the break occurs is reported for the time series for which there is a rejection of the null using the SB–FDF test. The quantity λ^* is the $\lambda \in (0.15, 0.85)$ that achieves the minimum in (4.5).

Table 4.5: True vs. Spurious Long Memory tests for Moody's credit spreads data. The test statistics (4.8) and (4.5) are calculated for yields and spreads. One asterisk denotes significance at 5% level and two asterisks denote significance at 1% level.

Yields					
	T	Aaa	Aa	A	Baa
$\hat{\eta}_\mu$	0.043741	0.06237	0.070139	0.087256	0.094507
$\widehat{t}_{\hat{\phi}}$	-2.58427	-2.1995	-1.98075	-1.50977	-1.03776

Spreads									
	sAaa	sAa	sA	sBaa	sAaaAa	sAaaA	sAaaBaa	sAaA	sAaBaa	sABaa
$\hat{\eta}_\mu$	0.280007	0.302545	0.33116	0.235645	0.057505	0.056918	0.046218	0.044615	0.244622	0.035814
$\widehat{t}_{\hat{\phi}}$	-2.15794	-2.03361	-2.49965	-3.26767	-3.23292	-4.06801**	-4.62083**	-5.20612**	1.090709	-4.50789**

Table 4.6: True vs. Spurious Long Memory tests for Lehman Brothers' credit spreads data. The test statistics (4.8) and (4.5) are calculated for yields and spreads. One asterisk denotes significance at 5% level and two asterisks denote significance at 1% level.

		Yields			
	T	Aaa	Aa	A	Baa
$\hat{\eta}_\mu$	0.50502*	0.184451	0.210992	0.241603	0.107743
$\widehat{t}_{\hat{\phi}}$	-3.60927	-3.47558	-3.04849	-0.64383	-2.56709

				Spreads					
	sAaa	sAa	sA	sBaa	sAaaAa	sAaaA	sAaaBaa	sAaA	sAaBaa	sABaa
$\hat{\eta}_\mu$	0.231533	0.225268	0.287408	0.163766	0.134201	0.288276	0.207318	0.284206	0.190325	0.080241
$\widehat{t}_{\hat{\phi}}$	-2.01807	-1.85178	-1.06949	-2.59922	-2.51649	-0.57434	-1.0361	1.285724	-0.4855	-0.84174

Table 4.7: Moody's credit spreads data: fraction of the data in which the break is estimated to occur.

	sAaaA	sAaaBaa	sAaA	sABaa
λ^*	0.668341	0.782584	0.633002	0.81241

The results regarding the second dataset do not lend themselves to any difficulty in the interpretation. Apart from some limited evidence of spurious long memory for the U.S. Global Treasury Index (the null is rejected at the 5% level according to the test based on dth differencing, but it is not rejected by the SB–FDF test, if not at the 10% level), we can safely conclude that both yields and spreads are genuine long memory processes.

4.4.2   Application to CDS Data

The dataset consists of daily single-name composite spreads for the period January 2001–November 2006 with maturities 1y, 3y, 5y, 7y and 10y. The data used is the same as chapter 3 and was provided by Markit Partners. In the empirical application we use the first differences of the log-CDS quotes for the above maturities. Only obligors with at least 1000 consecutive daily observations are considered. For each maturity considered, Table 4.8 reports the number of companies left after those that don't satisfy the above requirement are filtered out.

Table 4.8: The table lists for each maturity the number of obligors for which there are more than 1000 consecutive daily observations.

Maturity (Years)	No of Obligors
5	638
3	545
7	530
1	511
10	484

When for a time series we have T_j observations, the bandwidth parameter m_j in the long memory parameter estimation is chosen to be $m_j = [T_j^\rho]$. The exponent ρ takes on values between 0.3 and 0.6 with increments of 0.05. The results of the GPH estimates are reported in Table 4.13–Table 4.13.

Table 4.9: GPH estimates for the first differences of 5 years CDS log-spreads: descriptive statistics. The last two columns report the number of obligors for which the estimate of the d parameter is statistically different from zero and bigger than zero respectively at the 1% confidence level.

ρ	mean	median	std	min	max	95% Range	99% Range	$\#(d \neq 0)$	$\#(d > 0)$
0.6	0.0025864	0.01558	0.12822	-0.50158	0.3657	0.51093	0.65755	125	54
0.55	-0.0012582	0.01073	0.13641	-0.50025	0.41607	0.54959	0.78029	79	32
0.5	0.0074436	0.01454	0.15672	-0.61333	0.36466	0.62284	0.7923	69	30
0.45	-0.0098929	0.002	0.1846	-0.60609	0.49867	0.73228	0.93106	51	17
0.4	-0.0062103	0.00958	0.22825	-0.82588	0.64538	0.9117	1.222	54	19
0.35	-0.06046	-0.06372	0.28406	-1.0983	0.779	1.1691	1.5725	58	18
0.3	-0.095057	-0.07628	0.3318	-1.2561	0.99806	1.3648	1.7742	43	6

The first thing to notice in the tables is that for all the maturities, the long memory parameter is, on average, very close to zero, at least when the most popular choices for the exponent determining the bandwidth parameter (i.e. $\rho = 0.6$ or $\rho = 0.5$) are considered. The only exception is given by the 1 year maturity, for which the average d is close to -0.09 and the median is approximately -0.09. Therefore, it seems that there is little evidence of long range dependence in CDS log-spreads. First differencing appears to

Table 4.10: GPH estimates for the first differences of 3 years CDS log-spreads: descriptive statistics. The last two columns report the number of obligors for which the estimate of the d parameter is statistically different from zero and bigger than zero respectively at the 1% confidence level.

ρ	mean	median	std	min	max	95% Range	99% Range	$\#(d \neq 0)$	$\#(d > 0)$
0.6	-0.025851	-0.00488	0.1324	-0.50129	0.2658	0.49676	0.67678	112	29
0.55	-0.019674	-0.00415	0.14058	-0.56885	0.31381	0.54213	0.66549	83	24
0.5	-0.0074618	0.01104	0.15452	-0.566	0.38478	0.59336	0.7859	51	14
0.45	-0.026191	-0.01251	0.17418	-0.527	0.43121	0.69404	0.89744	42	10
0.4	-0.023417	-0.00533	0.21642	-0.71501	0.67028	0.82197	1.1277	41	11
0.35	-0.076107	-0.05796	0.27057	-0.99449	0.89772	1.1021	1.5426	37	8
0.3	-0.10753	-0.11208	0.33698	-1.2794	0.90919	1.4073	1.7738	35	6

Table 4.11: GPH estimates for the first differences of 7 years CDS log-spreads: descriptive statistics. The last two columns report the number of obligors for which the estimate of the d parameter is statistically different from zero and bigger than zero respectively at the 1% confidence level.

ρ	mean	median	std	min	max	95% Range	99% Range	$\#(d \neq 0)$	$\#(d > 0)$
0.6	-2.89E-05	0.01376	0.1191	-0.37637	0.33527	0.47358	0.60237	83	34
0.55	0.0020043	0.00607	0.12562	-0.45423	0.38122	0.51109	0.66892	52	22
0.5	0.013071	0.02014	0.15101	-0.50453	0.456	0.60988	0.79348	53	27
0.45	-0.0049766	0.00731	0.1741	-0.47497	0.54431	0.72107	0.89854	44	17
0.4	0.0012121	0.01242	0.22113	-0.59375	0.61224	0.89971	1.1252	44	19
0.35	-0.048014	-0.04192	0.27917	-0.98171	0.82635	1.1167	1.4769	43	12
0.3	-0.069552	-0.05982	0.3479	-1.2598	1.2234	1.3877	2.0458	34	9

Table 4.12: GPH estimates for the first differences of 1 year CDS log-spreads: descriptive statistics. The last two columns report the number of obligors for which the estimate of the d parameter is statistically different from zero and bigger than zero respectively at the 1% confidence level.

ρ	mean	median	std	min	max	95% Range	99% Range	$\#(d \neq 0)$	$\#(d > 0)$
0.6	-1.35E-01	-0.12566	0.13243	-0.57798	0.19389	0.52411	0.71042	190	3
0.55	-0.11196	-0.10832	0.1427	-0.66407	0.28544	0.56542	0.74165	130	9
0.5	-0.094321	-0.08614	0.15937	-0.74935	0.32344	0.63186	0.78532	86	2
0.45	-0.099257	-0.10385	0.18666	-0.80756	0.44377	0.73638	0.9789	60	6
0.4	-0.088001	-0.08551	0.21874	-0.80299	0.59009	0.85193	1.1598	42	6
0.35	-0.12967	-0.11959	0.26403	-1.0033	0.87386	1.0467	1.5414	44	7
0.3	-0.14531	-0.11419	0.3424	-1.4286	0.73669	1.3502	1.8443	38	3

Table 4.13: GPH estimates for the first differences of 10 years CDS log-spreads: descriptive statistics. The last two columns report the number of obligors for which the estimate of the d parameter is statistically different from zero and bigger than zero respectively at the 1% confidence level.

ρ	mean	median	std	min	max	95% Range	99% Range	$\#(d \neq 0)$	$\#(d > 0)$
0.6	-2.50E-02	-0.01509	0.12723	-0.45313	0.31131	0.50036	0.69007	88	20
0.55	-0.018875	-0.01051	0.13511	-0.49988	0.39541	0.52831	0.77583	57	17
0.5	-0.0068585	-0.00076	0.15252	-0.56555	0.37933	0.60018	0.77261	44	14
0.45	-0.019492	-0.00088	0.17731	-0.79848	0.46821	0.68993	0.91074	38	12
0.4	-0.0075678	0.00776	0.21632	-0.73368	0.57988	0.86062	1.1038	35	7
0.35	-0.037314	-0.03285	0.26173	-1.0021	0.68582	1.0167	1.4242	25	3
0.3	-0.049611	-0.03266	0.3371	-1.6922	0.94317	1.3861	1.9316	32	10

be sufficient to achieve stationarity. Another interesting feature of the results is that, generally, the average mean decreases as the exponent ρ increases. Further, as one would expect, both the standard deviation and the 95% and the 99% range of the d estimates increase as the exponent ρ increases. Some comments are in order for the last two columns of Table 4.13–Table 4.13, in which, using a 1% confidence level, the number of obligors for which the estimate of the d parameter is statistically different from zero and bigger than zero, is reported. If we focus on the first column, the maturity with the biggest rejection frequency and also the biggest absolute number of rejections is 1 year for which 190 time series out of 511 (37.2%) have a parameter d statistically different from zero. For this particular maturity the difference between the number of rejections of the null $d = 0$ and the number of rejections of the null $d \leq 0$ is extremely large. In the latter case there are only 3 companies for which the test rejects the null. Therefore these results suggest that approximately a third of the time series of the log-CDS with maturity 1 year have associated a long memory parameter d falling in the interval $(0.75, 1)$. If one were to list the rejection frequencies of the null $d = 0$ for the different maturities, one would next find the 3 years one (112 rejections out of 545, 20.5%), the 5 years one (125 rejections out of 638, 19.6%), the 10 years maturity (88 rejections out of 484, 18.2%) and finally the 7 years maturity with 83 rejections out of 511 and a frequency of 15.6%. On the other hand the number of rejections in the test with null $d \leq 0$ is achieved for the most liquid maturity, i.e. 5 years for which there are 54 rejections (8.5%). The 5 years maturity is followed by the maturities 7 years (34 rejections–6.4%), 3 years (29 rejections–5.3%), 10 years (20 rejections–4.1%) and 1 year for which, as mentioned, the test rejects only 3 times (0.58%). This numbers pretty much confirm that log-CDS spreads, after first differencing, become stationary series for which there is few evidence of long memory.

Despite the most common choice for the bandwidth parameter is $m = [T^{0.5}]$, we now take into account the results for the case $\rho = 0.6$ and run the tests of section 4.2 for the obligors whose time series seem to be long memory with this choice of the bandwidth parameter. The reason for this choice are essentially two: we want to run the tests of true vs. spurious long memory on as many obligors as possible and the estimates using $\rho = 0.6$ are sufficiently reliable. In Table 4.14 are thus reported the descriptive statistics regarding the GPH estimates for the time series for which d is statistically bigger than zero for this choice of the bandwidth parameter.

Table 4.14: GPH estimates for the first differences of CDS log-spreads which exhibit long range dependence: descriptive statistics.

Maturity	Observations	mean	median	std	min	max	95% Range	99% Range
5Y	54	0.2103	0.1999	0.042539	0.15975	0.3657	0.16056	0.20595
3Y	29	0.19217	0.18944	0.02118	0.15955	0.2658	0.096228	0.10625
7Y	34	0.19963	0.19012	0.034471	0.1563	0.33527	0.14647	0.17897
1Y	3	0.18709	0.18998	0.008627	0.17738	0.19389	0.016507	0.016507
10Y	20	0.20665	0.20305	0.033185	0.17685	0.31131	0.13446	0.13446

The mean of the d parameter estimated only for the series that seem to display the long memory property is close to 0.2 for all maturities with a maximum of 0.2103 (5 years) and a minimum of 0.18709 (1 year). The latter number, however, is only based on a sample of three issuers. This makes also the 1 year maturity the one with the smallest standard deviation. The median is not far from the mean, but apparently closer to 0.19 than to 0.2.

Table 4.15 reports the results of the statistical tests of true vs. spurious long memory, run only on the time series for which d is statistically bigger than zero at the 1% confidence level, when d is estimated using the GPH method with bandwidth parameter $m_j = [T_j^{0.5}]$.

The table confirms once again that the test based on temporal aggregation is way too conservative. It never rejects the null of true long memory except in one case for the 10 years maturity. Of course one must be careful when commenting the results regarding the 1 year maturity because these are based only on 3 data points. This is the explanation for the fact that, for instance the SB–FDF jumps to a zero rejection frequency and the test based on dth differencing always rejects. If we look at the 5 years data, we can see

Table 4.15: Rejection frequencies at 1% confidence level. For each test the null is true long memory. "SB–FDF" denotes the test developed in Dolado et al. (2005) and described in subsection 4.2.2. "dth Differencing" is the test of Shimotsu (2006) (subsection 4.2.4). which uses the KPSS test statistic. Every time the GPH estimation is involved the bandwidth parameter used is $m_j = \left[\sqrt{T_j}\right]$, T_j being the number of observations in the (sub)sample.

	SB–FDF	dth Differencing
5Y	0.203704	0.277778
3Y	0.137931	0.551724
7Y	0.264706	0.470588
1Y	0	1
10Y	0.05	0.6

that the test based on sample splitting and the one based on dth differencing have exactly the same rejection frequency: 15 series out of 54 can be viewed as being spurious long memory whereas the remaining 39 can be viewed as true long memory. The number of rejections reduces to 11 when the SB–FDF is used. Only for the 5 years maturity the results of tests, apart from the one based on temporal aggregation, seem to be quite robust. For the other maturities, the difference in the rejection frequencies of two different test can be as large as 55%, like in the 10 years maturity case. For the 3 years maturity case, the rejection frequencies for the SB–FDF test and the test based on sample splitting are comparably (13.8% and 20.7%), but when the test based on dth differencing this frequency is equal to 55%. Also for the 7 years maturity, the first two test have rejection frequencies not too far away (26.5% and 20.6%) but the last one has a rejection frequency much bigger (47%). Therefore only for the 5 years maturity the results are unambiguous and we can conclude that among the 54 time series between 39 and 43 are true long memory processes. In the other cases the conclusions from different tests are not always consistent.

5

Conclusions and Further Developments

This thesis proposes a fractional version of two well-know structural models to credit risk pricing: the Merton and Black and Cox models. In the classical formulation of these models the firm value is assumed to be driven by a geometric brownian motion. A consequence is that the models do not take into account the dependency structure of credit spreads over time and thus fail to predict credit spreads sufficiently close to those observed in the market. To correct for this drawback, we propose the fractional version of the aforementioned models, in which the firm value is driven by a fractional geometric brownian motion. In Chapter 2 the sensitivity analysis performed in order to understand the properties of the proposed models has shown that the theoretical spreads predicted by the fractional models are bigger that those predicted by the classical structural models and closer the empirical ones. The validity of the models has been confirmed by the fact that credit spreads are long memory processes. The study of the long memory properties of yields and credit spreads has been done in Chapter 2, in which a number of estimators and their properties have been presented. Mainly semiparametric estimator are used, in particular the GPH and the Whittle estimator and their extensions. All these estimation procedures and the related statistical test show that credit spreads can be described by a non-stationary long memory process. Further, several tests recently proposed in the econometric literature to detect the presence of fractional cointegration, have been studied and implemented. We find evidence of fractional cointegration in the bivariate systems of yields. This implies that there still exists a long run equilibrium relationship between yields, and deviations from the fractionally cointegrating relationship are mean reverting, so that a shock to the system will eventually die out. Formal tests for the null of true long memory versus the alternative of long memory induced by one or more structural break, described in details in Chapter 4 and run on time series of credit spreads, tell us that, at least for the datasets used, the long memory characteristics of credit spreads are genuine. The result of the last chapter appear reasonably robust, given that four different tests lead to the same conclusion of true long memory, at least for the second database studied. The tests of Chapter 4 have been also applied to the CDS data of Chapter 3, after having determined how many obligors for the most liquid maturities display the property of long range dependence. The result show that, estimating the fractional differencing parameter with the GPH estimator for the first difference of CDS quotes, results in a rather small proportions of obligors having observations which are long memory. However the vast majority of obligors whose time series seem to display long memory features, have been shown to be genuine long memory by the four tests above. A possible extension of this chapter would be the study of the long memory properties of CMCDS contracts. The idea is to first construct a large database with the theoretical CMCDS prices, determined using the different methods implemented. Once the database is available it is possible to apply the estimators of the long memory parameter to the constructed quotes and carry on the statistical tests for the null of true versus spurious long memory. Further extensions include the study of the long memory properties in the volatility of CDS/CMCDS quotes. Since long memory processes have been used extensively in modeling the strong persistence observed in volatility of asset prices (see for instance Baillie et al., 1996; Andersen and Bollerslev, 1997; Bollerslev and Mikkelsen, 1996; Breidt et al., 1998), it could be interesting to verify whether this feature is common to this kind of credit data as well. In Chapter 3, single-name CDS spread data is used to replicate the *would have been* CMCDS spreads. These, in turn, are used to study paired trading strategy of going long CDS and shorting CMCDS and

analyse the profit-loss profile across the database of obligors considered. The idea was to investigate whether historically investors are better off trading in one type of instrument versus the other (CMCDS vs CDS). On average the aforementioned strategy would have led to a loss. In other words it would have been more profitable to sell CDS and to buy CMCDS. Considering all the methods implemented for the calculation of the participation rates, at least 85% of the names analysed had a negative cumulative net trading profit/loss over the 5 years period considered. The convexity adjustment is found to have some effect over the trading strategies considered, because it appears to reduce the loss that an investor would have incurred in by buying CDS and selling CMCDS over the 5 years period considered. However, the percentage of names for which the above strategy would have led to a loss is not influenced significantly when the convexity adjustment has been taken into account. Further, introducing a cap on the floating payment does not have a significant impact on the results. The results are also quite robust, because the average losses across the database corresponding to the different (parametric and non-parametric) methods implemented to compute the participation rates and hence the CMCDS theoretical quotes don't change too much. The next step of the research was to "sieve" the list of obligors in the database in order to find the names for which the strategy gives a gain or a loss bigger than 500 or 1000 bp no matter what the method used to compute the participation rate. Interestingly the strategies involving the companies "Ford Mtr Co" and "Gen Mtrs Corp" lead to a profits bigger than 500 bp and bigger than 1000 bp respectively for all the methods implemented.

An interesting research question would be: how bigger are the credit spreads from the fractional models than those predicted by the classical models? In other words it would be interesting to understand how close they are to the credit spreads observed in the market. To answer this question a calibration exercise is needed. The difficulties related to this exercise are the same that a researcher faces during the calibration of the classical Merton model, i.e. the non-observability of the firm value process V and the need to restrict the attention to firms with a particularly simple debt structure. This difficulties exacerbate in the case of fractional models, because for the estimation of the Hurst parameter a sufficient number of observation are needed in order to have a reliable estimate and usually it is not easy to find accounting informations with the desired frequency. Many extension of the work developed in Chapter 3 are possible. First of all, it could be very interesting to solve the problem of the convexity adjustment by evaluating in a lattice payment and loss leg and hence the participation rate. This method is briefly introduced in Section A.3, but still some work is needed to calibrate and implement the proposed method. This additional tool would be important because so far only one method which allows for the convexity adjustment has been identified and implemented. Another method will make the results more robust and the conclusions regarding the trading strategies stronger. Of course there are many issues to consider in these approach, for instance the choice of the process for the hazard rate, the calibration of the tree and the parameter estimation. For instance it is not clear at the moment if it is better to use for the intensity rate the one-factor Hull and White model in order to exploit the results in Grant and Vora (2002) as far as the calibration of the tree is concerned and make the computations faster, rather than the BK model as suggested by Pedersen and Sen (2004). Some other issues that could be investigated as an extension of this research include the effect of the different maturities of the CMCDS on the statistical arbitrage analysis, the impact of rating and sector category. Replicating the statistical arbitrage analysis studying a three years contract starting at the end of 2003 with a three years reference CDS, would be something worth trying. This would be a way to verify if also in this case the same pattern is observed, i.e. investors are better off selling CDS and buying CMCDS and if the obligors for which the strategy leads to a profit (loss) are the same that led to a profit (loss) for the 5 years contract indexed to the 5 years CDS. The importance of this further analysis is in the fact that there are much more companies in the dataset for which this can be done, making the conclusions more robust, even though the 3 years CDS contract is less liquid than the 5 years one. Furthermore, a possible extension could be the study of the impact of the convexity adjustment on different maturities. For instance it could be interesting to verify empirically if the convexity adjustment has a bigger impact for longer maturities, as one would expect.

A

Technical Proofs

A.1 Analytic Formula for Forward CDS Spreads

Consider a forward contract for default protection during the time period (t, T) The buyer will enter this contract at time t if the default has not occurred by that date. Once again, the pricing equation for forward CDS is written by equating the estimated values of the protection and premium leg of the contract as of starting date t assuming that the debt issuer survived until that time. If there is a credit event prior to t the forward CDS contract ceases to exist and, hence, the equality is trivial with no payments made on either side. If we do not consider accrual, $FS(0, t, T) = FS(t, T)$, satisfies

$$FS(t,T) \int_t^T \exp\left(-\int_t^v (r_u + \lambda_u) \mathrm{d}u\right) \mathrm{d}v = q \int_t^T \lambda_v \exp\left(-\int_t^v (r_u + \lambda_u)\mathrm{d}u\right)\mathrm{d}v.$$

Multiplying both sides by $\exp\left(-\int_0^t (r_u + \lambda_u)\mathrm{d}u\right)$, taking this under the integral and splitting the integral in $\int_t^T = \int_0^T - \int_0^t$, gives

$$FS(t,T)\left[\int_0^T \exp\left(-\int_0^v (r_u+\lambda_u)\mathrm{d}u\right)\mathrm{d}v - \int_0^t \exp\left(-\int_0^v (r_u+\lambda_u)\mathrm{d}u\right)\mathrm{d}v\right]$$

$$= q\left[\int_0^T \lambda_v \exp\left(-\int_0^v (r_u+\lambda_u)\mathrm{d}u\right)\mathrm{d}v - \int_0^t \lambda_v \exp\left(-\int_0^v (r_u+\lambda_u)\mathrm{d}u\right)\mathrm{d}v\right].$$

Denoting by $\delta(t,T) = \frac{\int_0^t \exp\left(-\int_0^v (r_u+\lambda_u)\mathrm{d}u\right)\mathrm{d}v}{\int_0^T \exp\left(-\int_0^v (r_u+\lambda_u)\mathrm{d}u\right)\mathrm{d}v} \equiv \frac{\text{RiskyPV01}(0,t)}{\text{RiskyPV01}(0,T)}$ and dividing both the left and right hand side by $\int_0^T \exp\left(-\int_0^v (r_u + \lambda_u)\mathrm{d}u\right)$ yields

$$FS(t,T)\left[1 - \underbrace{\frac{\int_0^t \exp\left(-\int_0^v (r_u+\lambda_u)\mathrm{d}u\right)\mathrm{d}v}{\int_0^T \exp\left(-\int_0^v (r_u+\lambda_u)\mathrm{d}u\right)\mathrm{d}v}}_{\delta(t,T)}\right] = q\underbrace{\frac{\int_0^T \lambda_v \exp\left(-\int_0^v (r_u+\lambda_u)\mathrm{d}u\right)\mathrm{d}v}{\int_0^T \exp\left(-\int_0^v (r_u+\lambda_u)\mathrm{d}u\right)}}_{S(0,T)}$$

$$- q\underbrace{\frac{\int_0^t \lambda_v \exp\left(-\int_0^v (r_u+\lambda_u)\mathrm{d}u\right)\mathrm{d}v}{\int_0^t \exp\left(-\int_0^v (r_u+\lambda_u)\mathrm{d}u\right)}}_{S(0,t)} \times \underbrace{\frac{\int_0^t \exp\left(-\int_0^v (r_u+\lambda_u)\mathrm{d}u\right)\mathrm{d}v}{\int_0^T \exp\left(-\int_0^v (r_u+\lambda_u)\mathrm{d}u\right)\mathrm{d}v}}_{\delta(t,T)}$$

and thus

$$FS(t,T) = \frac{S(0,T) - \delta(t,T)S(0,t)}{1 - \delta(t,T)}.$$

A.2 Survival Probabilities under a OU Process for the Hazard Rate

From (3.14) we have

$$
\begin{aligned}
SP(t) &= \mathbb{E}_0\left[\exp\left(-\int_0^t\left\{\lambda_0 e^{-\alpha s} + \frac{k}{\alpha}\left(1 - e^{-\alpha s}\right) + \sigma\int_0^s e^{\alpha(u-s)}\mathrm{d}B_u\right\}\mathrm{d}s\right)\right] \\
&= \exp\left[-\frac{kt}{\alpha} - \frac{1-e^{-\alpha t}}{\alpha}(\lambda_0 - \frac{k}{\alpha})\right]\mathbb{E}_0\left\{\exp\left[-\sigma\int_0^t\left(\int_0^s e^{\alpha(u-s)}\mathrm{d}B_u\right)\mathrm{d}s\right]\right\}
\end{aligned}
$$

Denoting by Y the term inside the first exponential under the expectation we have to calculate $\mathbb{E}_0(\exp(Y))$. Now Y is Gaussian distributed with mean 0 and variance equal to

$$
var(Y) = \sigma^2\int_0^t e^{2\alpha u}\left(\int_u^t e^{-\alpha s}\mathrm{d}s\right)^2\mathrm{d}u = -\sigma^2\frac{3 + e^{-2\alpha t} - 4e^{-\alpha t} - 2t\alpha}{2\alpha^3}
$$

The expectation is easy to calculate because it is the mean of a lognormal variable. Hence

$$
SP(t) = \exp\left[-\frac{kt}{\alpha} - \frac{1-e^{-\alpha t}}{\alpha}(\lambda_0 - \frac{k}{\alpha})\right]\exp[var(Y)/2]
$$

$$
= \exp\left[-\frac{kt}{\alpha} - \frac{1-e^{-\alpha t}}{\alpha}(\lambda_0 - \frac{k}{\alpha}) - \sigma^2\frac{3 + e^{-2\alpha t} - 4e^{-\alpha t} - 2t\alpha}{4\alpha^3}\right].
$$

Rearranging, we get (3.15)–(3.16).

Another way to derive this formula is to express the stochastic intensity λ as a function Λ of an affine process X whose dynamics is given by the SDE:

$$
\mathrm{d}X_t = f(X_t)\mathrm{d}t + g(X_t)\mathrm{d}\tilde{B}_t
$$

where $\tilde{B}$ is a multidimensional Brownian motion and the drift $f(X_t)$ and the covariance matrix $g(X_t)g(X_t)'$ have affine dependence on X_t (see Duffie et al., 2003). It is possible to show that that, under technical conditions (see Duffie and Singleton, 2003), for any $\omega \in \mathbb{R}$

$$
\mathbb{E}_t\left[e^{\int_t^T -\Lambda(X_u)\mathrm{d}u + \omega X_T}\right] = e^{a(T-t) + b(T-t)X_t} \tag{A.1}
$$

where the coefficients $a(\cdot)$ and $b(\cdot)$ satisfy generalized Riccati ODEs. If we assume that the intensity itself is an affine process like in (3.7), then we can apply (A.1) with $\omega = 0$ and $\Lambda(x) = x$. In this case the ODEs can be solved analytically yielding (3.16).

A.3 Valuation of CMCDS in a Lattice

As an alternative to the approach which assumes a OU process for the default intensity, we could construct a lattice (or recombining tree) for the hazard rate which can be used to value the denominator of (3.5).

In this respect, we can exploit the analogy to interest rate models. In interest rate models we specify a stochastic process for the instantaneous interest rate, r_s and value the price at time zero of a bond maturing at t as

$$
B(0,t) = \mathbb{E}_0\left[\exp\left(-\int_0^t r_s\mathrm{d}s\right)\right].
$$

If in the RHS we replace r_s with λ_s we get the formula for the survival probability (3.14).

Since we know how to implement and calibrate short-rate models in a lattice, it is possible to do the same for the hazard rate.

In particular, following Pedersen and Sen (2004), we implement the Black-Karasinski (BK) model (see Black and Karasinski, 1991) on a trinomial lattice with a constant hazard rate volatility. The BK process is lognormal with mean reversion.

First a trinomial lattice like the one for the Hull-White (HW) model for interest rates is considered. Then, a simple modification allows us to build the tree for hazard rates following a BK model.

The steps are:

1. Fit a piece-wise constant hazard rate model to a curve of CDS spreads and the chosen recovery-given-default (See section 3.4). Produce a curve of risk-neutral survival probabilities.

2. Estimate the hazard rate volatility parameter σ and the mean reversion speed a and build a tree for x following the process
$$\mathrm{d}x = -ax\mathrm{d}t + \sigma\mathrm{d}B. \qquad (A.2)$$

Suppose we are interested in an interval of length T. The length of the time step is $\Delta t = \frac{T}{n}$ and spacing between x is $\Delta x = \sigma\sqrt{3\Delta t}$. The value of x at the node (i,j), $i = 0, 1, \ldots, n$, where time is $i\Delta t$, is $x_{ij} = j\Delta x$ where j is a positive or negative integer.

The tree is not standard, because of the mean reversion. In particular we have special branching at the top and bottom nodes when $j > j_{\max}$ and $j < j_{\min}$ (See Figure A.1).

Following Hull and White (1994, 1996) we set

$$j_{\max} = \left\lceil \frac{0.184}{a\Delta t} \right\rceil, \quad j_{\min} = -j_{\max}.$$

Branch probabilities are chosen so that at all nodes the first two moments of these discrete model and first two moments of the continuous process (A.2) coincide. Define by p_u, p_m and p_d the probabilities of the highest, middle and lowest branches emanating from the node. Then at a node like the one in Figure A.1–(a) we have

$$p_u = \frac{1}{6} + \frac{a^2 j^2 \Delta t^2 - aj\Delta t}{2}$$
$$p_m = \frac{2}{3} - a^2 j^2 \Delta t^2$$
$$p_d = \frac{1}{6} + \frac{a^2 j^2 \Delta t^2 + aj\Delta t}{2}$$

If the branching is like the one in Figure A.1–(b)

$$p_u = \frac{7}{6} + \frac{a^2 j^2 \Delta t^2 - 3aj\Delta t}{2}$$
$$p_m = -\frac{1}{3} - a^2 j^2 \Delta t^2 + 2aj\Delta t$$
$$p_d = \frac{1}{6} + \frac{a^2 j^2 \Delta t^2 - aj\Delta t}{2}$$

Finally if the branching has the form shown in Figure A.1–(c)

$$p_u = \frac{1}{6} + \frac{a^2 j^2 \Delta t^2 + aj\Delta t}{2}$$
$$p_m = -\frac{1}{3} - a^2 j^2 \Delta t^2 - 2aj\Delta t$$
$$p_d = \frac{7}{6} + \frac{a^2 j^2 \Delta t^2 + 3aj\Delta t}{2}$$

Figure A.1: The three branching types of the Hull-White trinomial tree. (a) is the standard branching method at inner nodes of the tree, (b) is used at the upper edge of the tree ($j > j_{\max}$), and (c) is used at the lower edge of the tree ($j < j_{\min}$).

(a)   (b)   (c)
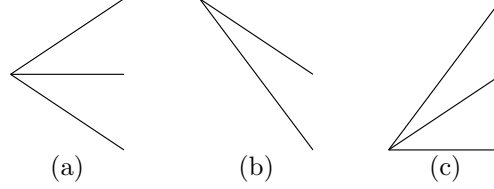
3. Create a new default node for each time slice in the lattice and an extra branch from each original node to the default state on the next time slice to obtain a tree like the one in Figure A.2. Switch from a tree for x to a tree for λ by setting the hazard rate at a given node to $\lambda_{ij} = \exp(x_{ij})$. The probability of going to the default state is set to $p = 1 - \exp(-\lambda_{ij}\Delta t)$ where λ_{ij} is the hazard rate at that node. For each node, the branch probabilities to the original nodes are multiplied by $1 - p$, where p is the default probability for that node.

Figure A.2: Tree for the process x. The layer above the tree is the default state.
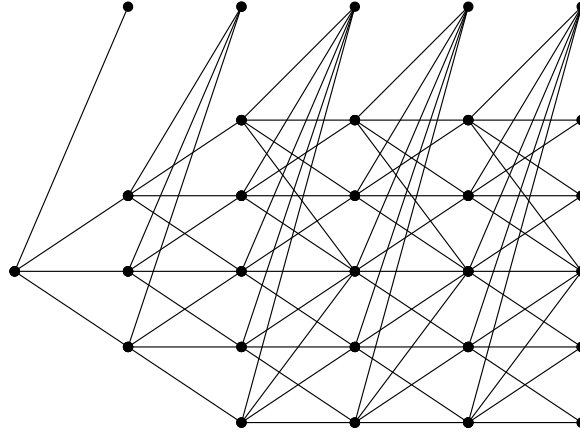
4. Calibrate the model to ensure it fits the curve of survival probabilities found in step 1 above. For each time slice in the lattice we shift the hazard rates at all nodes on that time slice by the same relative amount to match the default probabilities from the data. This is done iteratively. The hazard rate at node $(0,0)$ is

$$\lambda_{0,0}^* = -\frac{\log SP(\Delta t)}{\Delta t}$$

where $SP(\Delta t)$ is the survival probability from step 1.

If calibration has been done up to time $i-1$, we set the probability of having a default between $i\Delta t$ and $(i+1)\Delta t$ from the tree equal to the corresponding observed probability. This results in the following

equation:

$$\sum_{j=-\ell}^{\ell} p_{(i,j)}^* P((i,j) \to def) = 1 - SP((i+1)\Delta t | i\Delta t) = 1 - \frac{SP((i+1)\Delta t)}{SP(i\Delta t)} \tag{A.3}$$

where $\ell = \min(i, j_{\max})$, $1 - SP((i+1)\Delta t | i\Delta t)$ is the observed probability of having a default between $(i+1)\Delta t$ and $i\Delta t$, and $p_{(i,j)}^*$ is the already adjusted probability (as seen from the initial node) of reaching node j at time $i\Delta t$. For a given time slice these probabilities only depend on the already fitted hazard rates at previous time slices. Denoting by h_i the shift in the hazard rates at time $i\Delta t$, we write the probability of defaulting from node (i,j) as

$$P((i,j) \to def) = 1 - \exp\left[-(\lambda_{ij} + h_i)\Delta t\right]$$

so that we can solve (A.3) for h_i.

Equation (A.3) implies

$$e^{-h_i \Delta t} \sum_{j=-\ell}^{\ell} p_{(i,j)}^* e^{-\lambda_{ij}\Delta t} = \sum_{j=-\ell}^{\ell} p_{(i,j)}^* - 1 + \frac{SP((i+1)\Delta t)}{SP(i\Delta t)}.$$

Now, since

$$\sum_{j=-\ell}^{\ell} p_{(i,j)}^* = SP(i\Delta t)$$

we have for $i = 1, \ldots, n-1$

$$h_i = -\frac{1}{\Delta t} \log\left[\frac{SP(i\Delta t) - 1 + \frac{SP((i+1)\Delta t)}{SP(i\Delta t)}}{\sum_{j=-\ell}^{\ell} p_{(i,j)}^* e^{-\lambda_{ij}\Delta t}}\right]$$

Note that this equation holds true for $i = 0$ as well. The last h_i we compute is the one for time $T - \Delta t$ because we don't need the values of λ at the final time slice.

If instead of a BK model for the intensity rate, we use a one-factor Hull and White model (see Hull and White, 1990), we can exploit the results in Grant and Vora (2002) as far as the calibration of the tree is concerned. In this paper an analytical solution for the mean interest rate at each date in the tree for the Hull and White model is provided. We can apply the same results when we build the tree for the hazard rate.

Assume the hazard rate evolves according to

$$d\lambda_t = [\mu_t + \alpha(\gamma_t - \lambda_t)]\, dt + \sigma_t dB_t. \tag{A.4}$$

The discrete-time analogue of (A.4) is

$$\Delta\lambda_t = \lambda_{t+\Delta t} - \lambda_t = [\mu_t + (1-k)(\gamma_t - \lambda_t)]\,\Delta t + b\sigma_t \Delta B_t \tag{A.5}$$

where

$$k = e^{-\alpha\Delta t} \tag{A.6}$$

$$b = \left(\frac{1 - e^{-2\alpha\Delta t}}{2\alpha}\right)^{1/2} \tag{A.7}$$

$$\gamma_t = \gamma_0 + \sum_{j=0}^{t-1} \mu_j \tag{A.8}$$

and ΔB_t has a normal distribution with mean 0 and variance Δt.

When $\Delta t = 1$ from (A.5) and (A.8) it follows

$$\lambda_t = \gamma_t + k^t[\lambda_0 - \gamma_0] + b\sum_{j=0}^{t-1} k^{t-j-1}\sigma_j \Delta B_j, \tag{A.9}$$

i.e. λ_t has a normal distribution.

The calibration of the tree entails

$$SP(t) = \mathbb{E}_0\left[\exp\left(-\sum_{j=0}^{t-1}\lambda_j\right)\right].$$

Using the fact that if X is normally distributed, $N(\mu, \sigma^2)$, then

$$\mathbb{E}_0\left[\exp\left(-X\right)\right] = e^{-\mu + \frac{1}{2}\sigma^2}$$

we can conclude

$$\mathbb{E}_0\left[\lambda_t\right] = f(t) + \frac{1}{2}\text{Var}\left(\sum_{j=1}^{t}\lambda_j\right) - \frac{1}{2}\text{Var}\left(\sum_{j=1}^{t-1}\lambda_j\right)$$

with

$$\begin{cases} f(0) = \lambda_0 \\ f(t) = \log\frac{SP(t)}{SP(t+1)} \end{cases}.$$

Assuming a constant volatility σ we have

$$\text{Var}\left(\sum_{j=1}^{t}\lambda_j\right) = b^2\sigma^2\text{Var}\left(\sum_{j=1}^{t}\sum_{i=0}^{j-1}k^{j-i-1}\Delta B(i)\right)$$

$$= b^2\sigma^2\text{Var}\left(\sum_{i=0}^{t-1}\sum_{j=i+1}^{t}k^{j-i-1}\Delta B(i)\right) = b^2\sigma^2\text{Var}\left(\sum_{i=0}^{t-1}\Delta B(i)\sum_{j=i+1}^{t}k^{j-i-1}\right)$$

$$= b^2\sigma^2\sum_{i=0}^{t-1}\text{Var}(\Delta B(i))\left(\sum_{j=i+1}^{t}k^{j-i-1}\right)^2 = b^2\sigma^2\sum_{i=0}^{t-1}\left(\sum_{j=i+1}^{t}k^{j-i-1}\right)^2.$$

Extension to the case $\Delta t \neq 1$ are straightforward.

The knowledge of $\mathbb{E}_0\left[\lambda_t\right]$ for each time slice and the increment $\Delta\lambda = \sqrt{3\Delta t}$ allows us to build the calibrated tree for the hazard rate.

5. When the lattice has been calibrated, we have all the branch probabilities. We then find the one-step Libor discount factors between each pair of adjacent time slices. These are the forward discount factors from a curve fitted to current Libor/swap rates. The one step discount factors are used together with the branch probabilities when we value a security by calculating expected discounted values at each node backwards though time.

To value the floating premium leg on a CMDS, we consider a different lattice for each payment date. To value the payment set at date s_i (and paid at s_{i+1}) we construct the lattice out to $s_i + m$, where m is the constant maturity tenor (typically $m = 5$ years). The valuation is done by first discounting back in the lattice from $s_i + m$ to s_i to obtain at each node on the s_i time slice the values of RiskyPV01$(s_i, s_i + m)$,

RiskyPV01(s_i, s_{i+1}) and PVP01$(s_i, s_i + m)$, where PVP01$(s_i, s_i + m)$ is the value of the protection leg for the interval $(s_i, s_i + m)$ assuming a premium of 1bp.

We can therefore compute the value of the forward spread at each node as:

$$\frac{\text{PVP01}(s_i, s_i + m)}{\text{RiskyPV01}(s_i, s_i + m)}.$$

Since this is paid at time s_{i+1} we need to further multiply by RiskyPV01(s_i, s_{i+1}) to have the value at time s_i of the payment.

Note that the use of the lattice allows us also to include the accrual payment when valuing the present value of future payments.

The value at time zero of the payment set at date s_i is then obtained by discounting back in the lattice up to time zero.

Finally, the value of the floating premium leg is the sum of the present values of the individual payments.

B

Listings

Listing B.1: Calculation of Discount Factors from Libor rates: Matlab Code

```matlab
function df=rates2discount_curveL(rates,date)
% LiborSwap=rates2discount_curve(rates,date)
% Computes the discount curve
% from the Libor rates for  maturities 0:1:12 months
% on the day date with format 'dd/mm/yyyy'.

%Computing the vector N
%with the number of days
%to maturity
v=datevec(date,'dd/mm/yyyy');
temp=v;

for i=1:12
    temp(2)=temp(2)+1;
    s=datestr(temp);
    N(i)=datenum(s)-datenum(v);
end


df=zeros(12,1);
df(1)=1;
for i=1:12
    df(i+1)=1/(1+rates(i)/100*N(i)/360);
end
```

Listing B.2: Calculation of Discount Factors from Swap rates: Matlab Code

```matlab
function [dfLSR,dfCFR]=rates2discount_curveS(rates,Maturities)
% [dfLSR,dfCFR]=rates2discount_curveS(rates,Maturities)
% Computes the discount curve from the Swap rates for a vector of
% maturities, for instance [12;24;36;48;60;84;120;240;360] months.
% Both the "Linear Swap rates" and the "Constant Forward Rates" methods
% are used. The first maturity need be 12m.

Maturities=Maturities/12;
N=length(Maturities);

%Linear Swap rates
temp=[];%this will be the vector with
        %the missing value replaced by
        %interpolated swap rates.

for i=2:N
    inc=(rates(i)-rates(i-1))/(Maturities(i)-Maturities(i-1));
    temp2=[];
    for j=0:Maturities(i)-Maturities(i-1)-1
        temp2(j+1)=rates(i-1)+inc*j;
    end
    temp2=temp2';
    temp=[temp;temp2];
end
temp=[temp;rates(N)];

  %solving the linear system
A=diag(temp+100);
for i=1:Maturities(N)-1
 A(i+1:Maturities(N),i)=temp(i+1:Maturities(N));
end
dfLSR=A\(100*ones(Maturities(N),1));

%Constant Forward Rates
dfCFR(1)=100/(100+rates(1));
for i=2:N
    d=Maturities(i)-Maturities(i-1);
    if d==1
        temp=(100-rates(i)*sum(dfCFR))/(100+rates(i));
    else
        F_star=fzero(@(F) helperCFR(F,rates(i),dfCFR,d),1);
        temp=[];
        for j=1:d
          temp(j)=dfCFR(length(dfCFR))*(1+F_star)^(-j);
        end
        temp=temp';
    end
  dfCFR=[dfCFR;temp];
end
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
function f=helperCFR(F,r,df,d)

N=length(df);
f=-100+r*sum(df)+...
    r*df(N)*sum((1+F).^(-(1:d-1)))+...
    (100+r)*df(N)*(1+F)^(-d);
```

Listing B.3: Libor-swap Discount curve: Matlab Code

```
1  function DF=discount_curve(LiborSwap,Maturities)
2  % Computes the discount curve
3  % DF(i)=DF(t_i)=DF(i/12), interpolating
4  % from the vector LiborSwap, which contains
5  % the Libor or Swap rates for the maturities
6  % given by the vector Maturities.
7  % Maturities=[0;(1:12)';24;36;48;60;84;120;240;360]
8  % and, for instance LiborSwap=[1;0.997;0.995;...]
9
10 for i=2:length(Maturities)
11    for j=Maturities(i-1)+1:Maturities(i)
12     DF(j)=exp((Maturities(i)-j )/( Maturities(i)-Maturities(i-1))...
13     *log(LiborSwap(i-1))+...
14     (j-Maturities(i-1))/(Maturities(i)-Maturities(i-1))...
15     *log(LiborSwap(i)));
16    end
17 end
```

Listing B.4: Survival Probabilities with piecewise constant hazard rates: Matlab Code

```
1  function s=SP_lambda(tau,Tenors,lambda)
2  % SP_lambda(tau,Tenors,lambda)
3  % Survival probabilities with piecewise constant hazard rates lambda.
4  % Tau is time to maturity (in years), lambda is the vector of
5  % hazard rates and Tenors is the vector of maturities (months)
6  % available for CDS contracts. Tau can be a vector.
7
8  Tenors=Tenors/12;
9  M=length(Tenors);
10
11 for t=1:length(tau)
12
13    for i=1:M-1
14     innersums(i)=Tenors(1:i)'* ( lambda(1:i) - lambda(2:i+1))+...
15     lambda(i+1)*tau(t);
16     I(i)= (tau(t)>=Tenors(i)) *(tau(t)<Tenors(i+1));
17    end
18
19 s(t)=exp(-(lambda(1)*tau(t)*(tau(t)<Tenors(1))...
20 +innersums(1:M-1)*I(1:M-1)'...
21 +(innersums(M-1) )*(tau(t)>=Tenors(M))));%This is for tau>T_M
22
23 end
24 s=s';
```

Listing B.5: Objective Function used to bootstrap survival probabilities: Matlab Code

```
1  function f=obj_lambda(S,q,Tenors,n,DF,lambda)
2  % f=obj_lambda(S,q,Tenors,n,DF,lambda)
3  % This is the function whose
4  % zero we are looking for when the
5  % maturity of the CDS is n (months).
6  % n is in Tenors=[6;12;24;36;60;84;120].
7  % S is a length(Tenors)-vector with the available
8  % cds for the contracts maturing in n months.
9  % DF is the vector of discount factors.
10
11 k=(1:length(Tenors))*(n==Tenors);
12 % k is the element of S we are interested in
13 % for instance if n=60, k=5
14 Delta=1/4;
15 f=S(k)/(2*q)*Delta*sum(DF(n:-3:3).*...
16   (SP_lambda((n-3:-3:0)'/12,Tenors,lambda)+...
17   SP_lambda( (n:-3:3)' /12,Tenors,lambda)))-...
18   sum( DF(1:n).* ( SP_lambda((0:n-1)'/12,Tenors,lambda)-...
19   SP_lambda((1:n)'/12,Tenors,lambda)));
```

Listing B.6: Derivation of hazard rates: Matlab Code

```
1  function lambda=bootstrapped_lambda(S,q,DF,Tenors)
2  % lambda=bootstrapped_lambda(S,q,DF,Tenors)
3  % Returns the vector of hazard rates lambda
4  % given the vector of CDS premiums S with
5  % maturities (months) Tenors, the recovery rate 1-q
6  % and the discount factors vector DF.
7
8  M=length(Tenors);
9
10 lambda(1)=...
11 fzero(@(x) obj_lambda(S,q,Tenors,Tenors(1),DF,[x;zeros(M-1,1)]),0.5 );
12
13 for i=2:M-1
14     lambda(i)=fzero(@(x) obj_lambda(S,q,Tenors,Tenors(i),DF,...
15         [lambda;x;zeros(M-i,1)]),0.5 );
16     if i==2
17     lambda=lambda';
18     end
19 end
20
21 lambda(M)=fzero(@(x) obj_lambda(S,q,Tenors,Tenors(M),DF,[lambda;x]),0.5);
```

Listing B.7: RiskyPV01 when using Survival Probabilities with piecewise constant hazard rates: Matlab Code

```matlab
function RPV=RPV01_lambda(T,DF,Tenors,lambda)
% RPV=RPV01(T,DF,Tenors,lambda)
% T: Time in years
% DF is the vector with the discount curve
% Tenors,lambda are used to compute the
% survival probabilities SP

n=floor(T*12); %Time in months
k=floor(n/3);
Delta=1/4;
Delta0=1;

if T==0
    RPV=1;
else
RPV= .5*(Delta0*DF(1)*(1+SP_lambda(1/12,Tenors,lambda))+...
 Delta*sum(DF(n:-3:n-3*k+3).*...
 (SP_lambda((n:-3:n-3*k+3)'/12,Tenors,lambda)+...
 SP_lambda((n-3:-3:n-3*k)' /12,Tenors,lambda))));
end
```

Listing B.8: Theoretical CDS premiums when using Survival Probabilities with piecewise constant hazard rates: Matlab Code

```matlab
function S=CDS_lambda(T,DF,q,Tenors,lambda)
% S=CDS_NS(T,DF,alpha) computes
% the theoretical CDS spread for
% maturity T (years)
% using the bootstrapped
% survival probabilities.

for i=1:length(T)
    if T(i)==0
        S=1;
    else
 n=floor(T(i)*12); %Time in months
 S(i)=q/RPV01_lambda(T(i),DF,Tenors,lambda)*...
 sum(DF(1:n).*(SP_lambda((0:n-1)/12,Tenors,lambda)...
 -SP_lambda((1:n)/12,Tenors,lambda)));
    end

end
S=S';
```

Listing B.9: Survival Probabilities with the Nelson-Siegel function: Matlab Code

```
1   function s=SP_NS(t,alpha)
2   % s=SP_NS(t,alpha)
3   % Survival probabilities using
4   % the Nelson-Siegel  function.
5   % t is time to maturity (in years),
6   % alpha is the vector of parameters:
7   % alpha(1)=alpha_0;
8   % alpha(2)=alpha_1;
9   % alpha(3)=alpha_3.
10
11  for i=1:length(t)
12   s(i)=alpha(1)*t(i)+alpha(2)*alpha(3)*(1-exp(-t(i)/alpha(3)));
13   s(i)=exp(-s(i));
14  end
15  s=s';
```

Listing B.10: RiskyPV01 when using Survival Probabilities with the Nelson-Siegel function: Matlab Code

```
1   function RPV=RPV01_NS(T,DF,alpha)
2   % RPV=RPV01_NS(T,DF,alpha)
3   % Risky PV with the Nelson-Siegel  function.
4   % T: Time in years
5
6   n=floor(T*12); %Time in months
7   k=floor(n/3);
8   Delta=1/4;
9   Delta0=1;
10
11
12  if T==0
13      RPV=1;
14  else
15   RPV= .5*(Delta0*DF(1)*(1+SP_NS(1/12,alpha))+...
16      Delta*sum(  DF(n:-3:n-3*k+3).*...
17      (SP_NS( (n:-3:n-3*k+3)'/12,alpha)+...
18      SP_NS( (n-3:-3:n-3*k)' /12,alpha))));
19  end
```

Listing B.11: Theoretical CDS premiums when using Survival Probabilities with the Nelson-Siegel function: Matlab Code

```
1   function S=CDS_NS(T,DF,q,alpha)
2   % S=CDS_NS(T,DF,q,alpha) computes
3   % the theoretical CDS spread for
4   % maturity T (years)
5   % given a SP function
6   % with parameters alpha, a Discount curve
7   % and a loss rate q.
8
9   for i=1:length(T)
10      if T(i)==0
11          S=1;
12      else
13   n=floor(T(i)*12); %Time in months
14   S(i)=q/RPV01_NS(T(i),DF,alpha)*...
15   sum(DF(1:n).*(SP_NS((0:n-1)/12,alpha)-SP_NS((1:n)/12,alpha)));
16      end
17
18   end
19   S=S';
```

Listing B.12: Objective Function for the Nelson-Siefel interpolation: Matlab Code

```
1   function f=obj_NS(S,Tenors,DF,q,alpha)
2   % f=obj_NS(S,Tenors,DF,q,alpha)
3   % Computes the sum of squared residuals.
4   % Residuals are computed as the difference
5   % between thepretical and observed
6   % CDS spreads.
7   % S is the vector of observed spreads
8   % with maturity given by the vector Tenors.
9   % DF is the discount curve.
10  % q is the loss rate.
11  % The theoretical CDS spreads
12  % are computed with a NS function
13  % with parameters alpha.
14
15
16  theo=CDS_NS(Tenors/12,DF,q,alpha);
17  f=(theo-S)'*(theo-S);
```

Listing B.13: Derivation of $\hat{\alpha}$: Matlab Code

```matlab
1   function [alpha_star,fval]=alpha_NS(S,Tenors,DF,q)
2   % alpha_star=alpha_NS(S,Tenors,DF,q)
3   % computes the parameter alpha_star
4   % which minimizes the sum of squared
5   % residuals (res=theo CDS- obs CDS)
6   % or the sum of absolute residuals
7   % subject to the survival probabilites
8   % being non increasing and alpha_3>0.
9   % The theoretical CDS spreads are calculated
10  % using the Nelson-Siegel function.
11
12  warning off;
13  options=optimset('display','off');
14  [alpha_star,fval]=fmincon(@(alpha) obj_NS(S,Tenors,DF,q,alpha),...
15      .5*ones(3,1),[],[],[],[],zeros(3,1),[],[],options);
16
17  [b,fval2]=fmincon(@(alpha) obj_NS(S,Tenors,DF,q,alpha),...
18      .5*ones(3,1),[],[],[],[],[0;-Inf;0],[],@mycon,options);
19  if fval2<fval
20      alpha_star=b;fval=fval2;
21  end
22
23  [temp,fvalabs]=fmincon(@(alpha) obj_NSabs(S,Tenors,DF,q,alpha),...
24      .5*ones(3,1),[],[],[],[],zeros(3,1),[],[],options);
25
26  [b,fval2abs]=fmincon(@(alpha) obj_NSabs(S,Tenors,DF,q,alpha),...
27      .5*ones(3,1),[],[],[],[],[0;-Inf;0],[],@mycon,options);
28  if fval2abs<fvalabs
29      temp=b;fvalabs=fval2abs;
30  end
31
32  if  obj_NSabs(S,Tenors,DF,q,temp)<obj_NSabs(S,Tenors,DF,q,alpha_star)...
33          ||...
34      obj_NS(S,Tenors,DF,q,temp)<obj_NS(S,Tenors,DF,q,alpha_star)
35      alpha_star=temp;fval=fvalabs;
36  end
37  %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
38  function [c,ceq] = mycon(alpha)
39  % Constraint on alpha_1
40  c = -alpha(1)-alpha(2)*exp(-20/alpha(3));
41  ceq=[];
```

Listing B.14: Forward CDS when Survival Probabilities with piecewise constant hazard rates are used: Matlab Code

```
1  function FS=FS_lambda(t,T,DF,q,Tenors,lambda)
2  % FS=FS_lambda(t,T,DF,alpha)
3  % Forward CDS spreads
4  % using the bootstrapped
5  % survival probabilities.
6  % Time in years.
7
8  for i=1:length(t)
9      if t(i)==0
10         FS(i)=CDS_lambda(T(i),DF,q,Tenors,lambda);
11     else
12  delta=RPV01_lambda(t(i),DF,Tenors,lambda)/...
13      RPV01_lambda(T(i),DF,Tenors,lambda);
14  FS(i)=(CDS_lambda(T(i),DF,q,Tenors,lambda)-...
15      delta*CDS_lambda(t(i),DF,q,Tenors,lambda))/(1-delta);
16     end
17 end
18 FS=FS';
```

Listing B.15: Participation Rate when Survival Probabilities with piecewise constant hazard rates are used: Matlab Code

```
1  function PR=PR_lambda(m,T,S,Tenors,DF,q)
2  % PR=PR_lambda(m,T,S,Tenors,DF,q)
3  % Participation rate for
4  % a CMCDS contract with
5  % maturity T (years)
6  % with a reference CDS contract
7  % maturing in m years and a
8  % function for the SP with piecewise constant
9  % hazard rates.
10 % S is the vector of avaliable
11 % CDS spreads maturing at the
12 % time given by the vector Tenors (months).
13 % DF is the vector of discount factors
14 % and q the loss rate.
15
16 warning off;
17 lambda=bootstrapped_lambda(S,q,DF,Tenors);
18 n=floor(T*12); %Time in months
19 k=floor(n/3);
20 Delta=1/4;
21
22
23
24 j=(1:length(Tenors))*(m*12==Tenors);
25
26 NUM=S(j)*Delta*sum(  DF(n:-3:n-3*k+3).*...
27     (SP_lambda((n:-3:n-3*k+3)' /12,Tenors,lambda)+...
28     SP_lambda((n-3:-3:n-3*k)' /12,Tenors,lambda)));
29
30
31 DEN=Delta*sum(...
32     FS_lambda((n-3:-3:n-3*k)'/12,(n-3:-3:n-3*k)'/12+m,DF,q,Tenors,lambda).*...
33     DF(n:-3:n-3*k+3).* (SP_lambda((n:-3:n-3*k+3)' /12,Tenors,lambda)+...
34     SP_lambda((n-3:-3:n-3*k)' /12,Tenors,lambda)));
35 PR=NUM/DEN;
```

Listing B.16: Forward CDS when Nelson-Siegel interpolation is used: Matlab Code

```
1   function FS=FS_lambda(t,T,DF,q,Tenors,lambda)
2   % FS=FS_lambda(t,T,DF,q,Tenors,lambda)
3   % Forward CDS spreads
4   % using the bootstrapped
5   % survival probabilities.
6   % Time in years.
7
8   for i=1:length(t)
9       if t(i)==0
10          FS(i)=CDS_lambda(T(i),DF,q,Tenors,lambda);
11      else
12   delta=RPV01_lambda(t(i),DF,Tenors,lambda)/...
13       RPV01_lambda(T(i),DF,Tenors,lambda);
14   FS(i)=(CDS_lambda(T(i),DF,q,Tenors,lambda)-...
15       delta*CDS_lambda(t(i),DF,q,Tenors,lambda))/(1-delta);
16      end
17   end
18   FS=FS';
```

Listing B.17: Participation Rate when Nelson-Siegel interpolation is used: Matlab Code

```
1   function PR=PR_NS(m,T,S,Tenors,DF,q)
2   % PR=PR_NS(m,T,S,Tenors,DF,q)
3   % Participation rate for
4   % a CMCDS contract with
5   % maturity T (years)
6   % with a reference CDS contract
7   % maturing in m years and a
8   % Nelson-Siegel function for the SP.
9   % S is the vector of avaliable
10  % CDS spreads maturing at the
11  % time given by the vector Tenors (months).
12  % DF is the vector of discount factors
13  % and q the loss rate.
14
15  alpha_star=alpha_NS(S,Tenors,DF,q);
16  n=floor(T*12); %Time in months
17  k=floor(n/3);
18  Delta=1/4;
19
20
21
22  j=(1:length(Tenors))*(m*12==Tenors);
23
24  NUM=S(j)*Delta*sum(  DF(n:-3:n-3*k+3).*...
25      (SP_NS( (n:-3:n-3*k+3)' /12,alpha_star)+...
26      SP_NS( (n-3:-3:n-3*k)' /12,alpha_star)));
27
28
29  DEN=Delta*sum(...
30      FS_NS((n-3:-3:n-3*k)'/12,(n-3:-3:n-3*k)'/12+m,DF,q,alpha_star).*...
31      DF(n:-3:n-3*k+3).* (SP_NS( (n:-3:n-3*k+3)' /12,alpha_star)+...
32      SP_NS( (n-3:-3:n-3*k)' /12,alpha_star)));
33  PR=NUM/DEN;
```

Bibliography

Adenstedt, R. K. (1974). On large-sample estimation for the mean of a stationary random sequence. *Annals of Statistics 2*(6), 1095–1107.

Andersen, T. G. and T. Bollerslev (1997). Heterogeneous information arrivals and return volatility dynamics: Uncovering the long-run in high frequency returns. *Journal of Finance 52*(3), 975–1005.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2001). The distribution of realized exchanged rate volatility. *Journal of the American Stastistical Association 96*(453), 42–55.

Andrews, D. W. (1991). Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica 59*(3), 817–858.

Andrews, D. W. (1993). Test for parameter instability and structural break with unknown change point. *Econometrica 61*(4), 821–856.

Andrews, D. W. and P. Guggenberger (2003). A bias-reduced log-periodogram regression estimator for the long-memory parameter. *Econometrica 71*(2), 675–712.

Andrews, D. W. and Y. Sun (2004). Adaptive local polynomial Whittle estimation of long-range dependence. *Econometrica 72*(2), 569–614.

Arvanitis, A. and J. Gregory (2001). *Credit: The Complete Guide to Pricing, Hedging and Risk Management.* London: Risk Books.

Baillie, R. T., T. Bollerslev, and H. O. Mikkelsen (1996). Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics 74*(1), 3–30.

Ben Ameur, H., D. Brigo, and E. Errais (2006). A dynamic programming approach for pricing CDS and CDS options. Technical report, Stanford University.

Beran, J. (1995). Maximum likelihood estimation of the differencing parameter for invertible short and long memory autoregressive integrated moving average models. *Journal of the Royal Statistical Society, Series B 57*(4), 659–672.

Beran, J. and D. Ocker (2000). Temporal aggregation of stationary and nonstationary FARIMA$(p, d, 0)$ models. CoFE Discussion Paper 00-22, Center of Finance and Econometrics, University of Konstanz.

Berd, A. M. (2003, November). Forward CDS spreads. *Lehman Brothers Quantitative Credit Research Quarterly*, 40–44.

Berndt, A., R. Douglas, D. Duffie, M. Ferguson, and D. Schranz (2005). Measuring default risk premia from default swap rates and EDFs. Working paper, Stanford University.

Bianchi, S. (2005). Pathwise identification of the memory function of multifractional Brownian motion with application to finance. *International Journal of Theoretical and Applied Finance 8*(2), 1–27.

Bielecki, T. R., M. Jeanblanc, and M. Rutkowski (2005). Pricing and trading credit default swaps under deterministic intensity. Working paper, Isaac Newton Institute for Mathematical Sciences.

Bierens, H., J.-Z. Huang, and W. Kong (2005). Time-series estimation of aggregate corporate bond credit spreads. Working paper, Penn State University.

Black, F. and J. C. Cox (1976). Valuing corporate securities: Some effects of bond indenture provisions. *Journal of Financial and Quantitative Analysis 31*(2), 351–367.

Black, F. and P. Karasinski (1991, July-August). Bond and option pricing when short rates are lognormal. *Financial Analysts Journal*, 52–59.

Blanco, R., S. Brennan, and I. W. Marsh (2005). An empirical analysis of the dynamic relation between investment-grade bonds and credit default swaps. *Journal of Finance 60*(5), 2255–2282.

Bluhm, C., L. Overbeck, and C. Wagner (2003). *An introduction to credit risk modelling.* Boca Raton: Chapman & Hall/CRC.

Bollerslev, T. and H. O. Mikkelsen (1996). Modeling and pricing long memory in stock market volatility. *Journal of Econometrics 73*(1), 151–184.

Breidt, F. J., N. Crato, and P. de Lima (1998). The detection and estimation of long memory in stochastic volatility. *Journal of Econometrics 83*(1–2), 248–325.

Brigo, D. (2005). Constant maturity credit default swap pricing with market models. Technical report, Available at www.damianobrigo.it.

Brigo, D. and F. Mercurio (2006). *Interest Rate Models - Theory and Practice: With Smile, Inflation and Credit* (2 ed.). Berlin Heidelberg New York: Springer.

British Bankers' Association (2006). *BBA Credit Derivatives Report 2006.*

Carr, P., K. Ellis, and V. Gupta (1998). Static hedging of exotic options. *Journal of Finance 53*(3), 1165–1190.

Chan, K. C., G. A. Karolyi, F. A. Longstaff, and A. B. Sanders (1992). An empirical comparison of alternative models of the short-term interest rate. *Journal of Finance 47*(3), 1209–1227.

Cheng, T. C. K. (2004). Long memory features in the exchange rates of Asia-Pacific countries. Working paper, Department of Economics National University of Singapore.

Cherubini, U., E. Luciano, and W. Vecchiato (2004). *Copula Methods in Finance.* John Wiley & Sons.

Cheung, Y. and K. Lai (2001). Long memory and nonlinear mean reversion in Japanese Yen-based real exchange rates. *Journal of International Money and Finance 20*, 115–132.

Choi, K. and E. Zivot (2007). Occasional structural breaks and long memory with an application to the S&P 500 absolute stock returns. *Journal of International Money and Finance 26*(3), 342–363.

Chu, C. C. and Y. K. Kwok (2003, Spring). No-arbitrage approach to pricing credit spread derivatives. *Journal of Derivatives*, 51–64.

Collin-Dufresne, P. and R. S. Goldstein (2001). Do credit spreads reflect stationary leverage ratios? *Journal of Finance 56*(5), 1929–1957.

Cox, J. C., J. E. Ingersoll, and S. A. Ross (1985). A theory of the term structure of interest rates. *Econometrica 53*(2), 385–408.

Crato, N. and P. Rothman (1994). A reappraisal of parity reversion for UK real exchange rates. *Applied Economic Letters 1*, 139–141.

Das, S. and P. Tufano (1996). Pricing credit-sensitive debt when interest rates, credit ratings and credit spreads are stochastic. *Journal of Financial Engineering 5*(2), 161–198.

Das, S. R. and R. K. Sundaram (2005). A simple model for pricing securities with equity, interest-rate, and default risk. Working paper, Santa Clara University.

Della Ratta, L. and G. Urga (2005). Modelling credit spread: A fractional integration approach. Working paper, Centre for Econometric Analysis, Cass Business School.

Deo, R. S. and C. M. Hurvich (2001). On the log periodogram regression estimator of the memory parameter in long memory stochastic volatility models. *Econometric Theory 17*(4), 686–710.

Diebold, F. X. and A. Inoue (2001). Long memory and regime switching. *Journal of Econometrics 105*(1), 131–159.

Dittmann, I. (2004). Error correction models for fractionally cointegrated time series. *Journal of Time Series Analysis 25*, 27–32.

Dolado, J. J., J. Gonzalo, and L. Mayoral (2005). What is what?: A simple time-domain test of long-memory vs. structural breaks. Working paper, Universidad Carlos III de Madrid and Universidat Pompeu Fabra.

Doornik, J. A. and H. Hansen (1994). An omnibus test for univariate and multivariate normality. Discussion paper, Nuffield College.

Duffee, G. R. (1998). The relation between treasury yields and corporate bond yield spreads. *Journal of Finance 53*(6), 2225–2241.

Duffie, D. (1999). Credit swap valuation. *Financial Analysts Journal 55*(1), 73–87.

Duffie, D., D. Filipović, and W. Schachermayer (2003). Affine processes and applications in finance. *Annals of Applied Probability 13*(3), 984–1053.

Duffie, D. and K. J. Singleton (1997). An econometric model of the term structure of interest-rate swap yields. *Journal of Finance 52*(4), 1287–1321.

Duffie, D. and K. J. Singleton (1999). Modeling term structures of defaultable bonds. *Review of Financial Studies 12*(4), 687–720.

Duffie, D. and K. J. Singleton (2003). *Credit Risk*. Princeton: Princeton University Press.

Fox, R. and M. S. Taqqu (1986). Large-sample properties of parameter estimates for strongly dependent stationary Gaussian time series. *Annals of Statistics 14*(2), 517–532.

Geske, R. (1977). The valuation of corporate securities as compound options. *Journal of Finance 12*(4), 541–552.

Geweke, J. and S. Porter-Hudak (1983). The estimation and application of long memory time series models. *Journal of Time Series Analysis 4*, 221–238.

Granger, C. W. and N. Hyung (2004). Occasional structural breaks and long memory with an application to the S&P 500 absolute stock returns. *Journal of Empirical Finance 11*, 399–421.

Granger, C. W. J. (1980). Long memory relationships and the aggregation of dynamic models. *Journal of Econometrics 14*, 227–238.

Granger, C. W. J. (1986). Developments in the study of cointegrated economic variables. *Oxford Bulletin of Economics and Statistics 48*, 213–228.

Granger, C. W. J. and R. Joyeux (1980). An introduction to long-memory time series and fractional differencing. *Journal of Time Series Analysis 1*, 15–19.

Grant, D. and G. Vora (2002, Winter). The Hull and White model of the short rate: an alternative analytical representation. *Journal of Financial Research 25*(4), 463–476.

Greene, M. T. and B. D. Fielitz (1977). Long-term dependence in common stock returns. *Journal of Financial Economics 4*, 339–349.

Gripenberg, G. and I. Norros (1996). On the prediction of fractional brownian motion. *Journal of Applied Probability 33*, 400–410.

Haslett, J. and A. E. Raftery (1989). Space-time modelling with long-memory dependence: assessing irelands wind power resource. *Journal of the Royal Statistical Society, Series C 38*(1), 1–50.

Hassler, U., F. Marmol, and C. Velasco (2006). Residual log-periodogram inference for long-run relationships. *Journal of Econometrics 130*(1), 165–207.

Heath, D., R. Jarrow, and A. Morton (1992). Bond pricing and the term structure of interest rates: a new methodology for contingent claim valuation. *Econometrica 60*(1), 77–105.

Hiemstra, C. and J. D. Jones (1997). Another look at long memory in common stock returns. *Journal of Empirical Finance 4*(4), 373–401.

Hobijn, B., P. H. Franses, and M. Ooms (1998). Generalizations of the kpss–test for stationarity. Econometric institute report no. 9802/a, Erasmus University, Rotterdam.

Hu, Y. and B. Øksendal (2003). Fractional white noise calculus and applications to finance. *Infinite Dimensional Analysis, Quantum Probability and Related Topics 6*(1), 1–32.

Hull, J., I. Nelken, and A. White (2004). Merton's model, credit risk and volatility skews. *Journal of Credit Risk 1*(1), 53–60.

Hull, J., M. Predescu, and A. White (2004). The relationship between credit default swap spreads, bond yields, and credit rating announcements. *Journal of Banking and Finance 28*(11), 2789–2811.

Hull, J. and A. White (1990). Pricing interest rate derivative securities. *The Review of Financial Studies 3*(4), 573–592.

Hull, J. and A. White (1994, Fall). Numerical procedures for implementing term structure models I. *Journal of Derivatives*, 7–16.

Hull, J. and A. White (1996, Spring). Using Hull-White interest rate trees. *Journal of Derivatives*, 26–36.

Hull, J. and A. White (2000). Valuing credit default swaps I: No counterparty default risk. *Journal of Derivatives 8*(1), 29–40.

Hurst, H. E. (1951). Long-term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineers 116*, 770–799.

Jacobsen, B. (1996). Long term dependence in stock returns. *Journal of Empirical Finance 3*(4), 393–417.

Jamshidian, F. (2004). Valuation of credit default swaps and swaptions. *Finance and Stochastics 8*(3), 343–371.

Jarrow, R. and S. Turnbull (1995). Pricing derivatives on financial securities subject to credit risk. *Journal of Finance 50*(1), 53–86.

Jarrow, R., S. Turnbull, and D. Lando (1997). A markov model for the term structure of credit risk spreads. *The Review of Financial Studies 10*(2), 481–523.

Jeanblanc, M. and M. Rutkowski (2000). Modelling of default risk: Mathematical tools. Working paper, Department of Mathematics, Universite d'Evry.

Jones, E. P., S. P. Mason, and E. Rosenfeld (1984). Contingent claims analysis of corporate capital structures: An empirical investigation. *Journal of Finance 39*(3), 611–625.

Kiesel, R., W. Perraudin, and A. Taylor (2001). The structure of credit risk: spread volatility and ratings transitions. Working paper, Bank of England.

Kolmogorov, A. N. (1961). Local structure of turbulence in fluid for very large Reynolds numbers. In S. Friedlander and L. Topper (Eds.), *Turbulence*, pp. 151–155. Interscience Publishers.

Krekel, M. and J. Wenzel (2006). A unified approach to credit default swaption and constant maturity credit default swap valuation. Technical report, Fraunhofer-Institut für Techno- und Wirtschaftsmathematik.

Künsch, H. R. (1987). Statistical aspects of self-similar processes. In Y. Prohorov and V. Sazonov (Eds.), *Proceedings of the first World Congress of the Bernoulli Society*, pp. 67–74. Utrecht: VNU Science Press.

Lando, D. (1998). On Cox processes and credit risky securities. *Review of Derivatives Research 2*, 99–120.

Lando, D. (2004). *Credit Risk Modeling: Theory and Applications.* Princeton University Press.

Leland, H. E. (1994). Corporate debt value, bond covenants and optimal capital structure. *Journal of Finance 49*(4), 1213–1252.

Leland, H. E. and K. B. Toft (1996). Optimal capital structure, endogenous bankruptcy, and the term structure of credit spreads. *Journal of Finance 51*(3), 987–1019.

Lo, A. W. (1991). Long-term memory in stock market prices. *Econometrica 59*(5), 1279–1313.

Lobato, I. N. and N. E. Savin (1998). Real and spurious long memory properties of stock market data. *Journal of Business and Economic Statistics 16*, 261–267.

Longstaff, F. A., S. Mithal, and E. Neis (2005). Corporate Yield Spreads: Default Risk or Liquidity? New Evidence from the Credit Default Swap Market. *Journal of Finance 60*(5), 2213–2253.

Longstaff, F. A. and E. S. Schwartz (1995). A simple approach to valuing risky fixed and floating rate debt. *Journal of Finance 50*(3), 789–819.

Luciano, E. and E. Vigna (2006). Non mean reverting affine processes for stochastic mortality. Working paper, Collegio Carlo Alberto.

Madan, D. and H. Unal (1998). Pricing the risks of default. *Review of Derivatives Research 2*, 121–160.

Madan, D. and H. Unal (2000). A two-factor hazard rate model for pricing risky debt and the term structure of credit spreads. *Journal of Financial and Quantitative Analysis 35*(1), 43–65.

Mandelbrot, B. B. (1967). Forecasts of future prices, unbiased markets and martingale models. *Journal of Business 39*, 242–255.

Mandelbrot, B. B. and J. W. Van Ness (1968). Fractional Brownian motions, fractional noises and applications. *SIAM Review 10*(4), 422–437.

Mandelbrot, B. B. and J. R. Wallis (1969). Robustness of the rescaled range R/S in the measurement of noncyclic long run statistical dependence. *Water Resources Research 5*, 967–988.

Marinucci, D. and P. M. Robinson (1999). Alternative forms of Brownian motion. *Journal of Statistical Planning and Inference 80*, 11–122.

Mayoral, L. (2006). Testing for fractional integration versus short memory with trends and structural breaks. Working paper, Universidat Pompeu Fabra.

Merton, R. (1974). On the pricing of corporate debt: The risk structure of interest rates. *Journal of Finance 29*(2), 449–470.

Mikosch, T. and C. Stărică (2004). Nonstationarities in financial time series, the long-range dependence, and the IGARCH effects. *Review of Economics and Statistics 86*(1), 378–390.

Necula, C. (2002). Option pricing in a fractional brownian motion environment. Draft, Academy of Economic Studies, Bucharest, Romania.

Necula, C. (2003). Barrier options and a reflection principle of the fractional brownian motion. Draft, DOFIN, Academy of Economic Studies, Bucharest, Romania.

Nelson, C. R. and A. F. Siegel (1987). Parsimonious modelling of yield curves. *Journal of Business 60*(4), 473–489.

Nielsen, M. Ø. (2004). Optimal residual based tests for fractional cointegration and exchange rate dynamics. *Journal of Business and Economic Statistics 22*(3), 331–345.

Nielsen, M. Ø. (2005). Multivariate Lagrange multiplier tests for fractional integration. *Journal of Financial Econometrics 3*(3), 372–398.

Norden, L. and M. Weber (2004). Informational efficiency of credit default swap and stock markets: The impact of credit rating announcements. *Journal of Banking and Finance 28*, 2813–2843.

Ohanissian, A., J. R. Russell, and R. S. Tsay (2007). True or Spurious Long Memory? A New Test. Working paper, Graduate School of Business, University of Chicago.

O'Kane, D. and S. Turnbull (2003, April). Valuation of credit default swaps. *Lehman Brothers Quantitative Credit Research Quarterly*, 1–17.

Pan, J. and K. J. Singleton (2008). Default and recovery implicit in the term structure of sovereign CDS spreads. *Journal of Finance*. Forthcoming.

Pedersen, C. M. and S. Sen (2004, June). Valuation of constant maturity default swaps. *Lehman Brothers Quantitative Credit Research Quarterly*, 42–58.

Pedrosa, M. and R. Roll (1998). Systematic risk in corporate bond credit spreads. *Journal of Fixed Income 8*, 7–26.

Peltier, R. F. and J. Lévy Véhel (1995). Multifractional Brownian motion: definition and preliminary results. Rapport de recherche n.2645, INRIA, Le Chesnay Cedex.

Peters, E. E. (1994). *Fractal Market Analysis*. John Wiley & Sons.

Prigent, J.-L., O. Renault, and O. Scaillet (2001). An empirical investigation in credit spread indices. *Journal of Risk 3*, 27–55.

Rajan, A., G. McDermott, and R. Roy (2007). *The Structured Credit Handbook*. John Wiley & Sons.

Robinson, P. M. (1994). Semiparametric analysis of long-memory time series. *Annals of Statistics 22*(1), 221–238.

Robinson, P. M. (1995a). Gaussian semiparametric estimation of long range depedence. *Annals of Statistics 23*(5), 1630–1661.

Robinson, P. M. (1995b). Log-periodogram regression of time series with long-range dependence. *Annals of Statistics 23*(3), 515–539.

Robinson, P. M. (2005). The distance between rival nonstationary fractional processes. *Journal of Econometrics 128*(2), 283–300.

Robinson, P. M. and D. Marinucci (2001). Narrow-band analysis of nonstationary processes. *Annals of Statistics 29*(4), 947–986.

Schneider, P., L. Sögner, and T. Veža (2007). Jumps and recovery rates inferred from corporate CDS premia. Working paper, Vienna University of Economics and Business Administration.

Schönbucher, P. J. (2003). *Credit Derivatives Pricing Models: Models, Pricing and Implementation*. John Wiley & Sons.

Shimko, D., N. Tejima, and D. Van Deventer (1993). The pricing of risky debt when interest rates are stochastics. *Journal of Fixed Income 3*, 58–65.

Shimotsu, K. (2006). Simple (but effective) tests of long memory versus structural breaks. Working paper, Queen's University.

Shiryaev, A. N. (1999). *Essentials of Stochastic Finance: Facts, Models, Theory*. World Scientific.

Sottinen, T. (2001). Fractional Brownian motion, random walks and binary market models. *Finance and Stochastics 5*, 343–355.

Sottinen, T. and E. Valkeila (2001). Fractional Brownian motion as a model in finance. Working paper, University of Helsinki.

Sowell, F. (1992). Maximum likelihood estimation of stationary univariate fractionally integrated time series models. *Journal of Econometrics 53*, 165–188.

Stanton, R. (1997). A nonparametric model of term structure dynamics and the market price of interest rate risk. *Journal of Finance 52*(5), 1973–2002.

Taqqu, M. S. and V. Teverovsky (1996). Semi-parametric graphical estimation techniques for long-memory data. *Lecture Notes in Statistics 115*, 420–432.

Taqqu, M. S., V. Teverovsky, and W. Willinger (1995). Estimators for long-range dependence: an empirical study. *Fractals 3*, 785–788.

Tavakoli, J. M. (1998). *Credit Derivatives: A Guide to Instruments and Applications*. John Wiley & Sons.

Vasicek, O. (1977). An equilibrium characterization of the term structure. *Journal of Financial Economics 5*, 177–188.

Velasco, C. (1999a). Gaussian semiparametric estimation of non-stationary time series. *Journal of Time Series Analysis 20*, 87–127.

Velasco, C. (1999b). Non-stationary log-periodogram regression. *Journal of Econometrics 91*(2), 325–371.

Velasco, C. (2000). Non-Gaussian log-periodogram regression. *Econometric Theory 16*(1), 44–79.

Willinger, W., M. S. Taqqu, and V. Teverovsky (1999). Stock market prices and long-range dependence. *Finance ans Stochastics 3*(1), 1–13.