RISK 2000

CREDIT ANALYSIS

by

Dr Philipp Schonbucher

(paper)

JUNE. 13. 2000

BOSTON

The Pricing of
Credit Risk and Credit Risk Derivatives

Philipp J. Schönbucher

University of Bonn, Department of Statistics

October 1997, this version February 2000

[1] Author's address:
University of Bonn, Faculty of Economics,
Department of Statistics
Adenauerallee 24-42, 53113 Bonn, Germany
Tel: +49-228-739264 , Fax: +49-228-735050
email: P.Schonbucher@finasto.uni-bonn.de
http://www.finasto.uni-bonn.de/ schonbuc/
Comments welcome.

Contents

Notation

Most of the notation is standard in mathematical finance. Defaultable parameters usually carry a dash (') or an overbar.

Default-Free Term Structure of Interest Rates

Bond Prices

$B(t,T)$	default-free zero coupon bond price
$\mu(t,T)$	drift of zero coupon bond price $B(t,T)$
$\eta(t,T)$	volatility of zero coupon bond price $B(t,T)$
$\widehat{B}$	default-free coupon bond price

Default-Free Interest Rates

r	default-free instantaneous short rate (default-free short rate)
β	default-free discount factor: $\beta(t) = \exp\{-\int_0^t r(s)ds\}$
$f(t,T)$	default-free continuously compounded instantaneous forward rate (default-free forward rate)
$\alpha(t,T)$	drift of the default-free forward rate $f(t,T)$
$\sigma(t,T)$	volatility of the default-free forward rate $f(t,T)$
$a(t,T)$	integral of forward rate volatility: $a(t,T) = -\int_t^T \sigma(t,s)ds$
$F(t,T_1,T_2)$	simply compounded default-free forward rate over $[T_1,T_2]$

Defaultable Term Structure of Interest Rates

Defaultable Bond Prices

	defaultable quantities usually carry an overbar
$\overline{B}(t,T)$	defaultable zero coupon bond price
$\overline{\mu}(t,T)$	drift of defaultable zero coupon bond price
$\overline{\eta}(t,T)$	volatility of defaultable zero coupon bond price
$\widetilde{B}(t,T)$	defaultable zero bond price adjusted for defaults before t $\widetilde{B}(t,T) = \overline{B}(t,T)/Q(t)$

Defaultable Interest Rates

$\overline{r}$	defaultable instantaneous short rate (defaultable short rate)
$\overline{\beta}$	defaultable discount factor: $\overline{\beta}(t) = \exp\{-\int_0^t \overline{r}(s)ds\}$
$\overline{f}(t,T)$	defaultable continuously compounded instantaneous forward rate (defaultable forward rate)
$\overline{\alpha}(t,T)$	drift of the defaultable forward rate
$\overline{\sigma}(t,T)$	volatility of the defaultable forward rate
$\overline{a}(t,T)$	integral of forward rate volatility: $\overline{a}(t,T) = -\int_t^T \overline{\sigma}(t,s)ds$
$\overline{F}(t,T_1,T_2)$	defaultable simply compounded forward rate over $[T_1,T_2]$

Spread

$h(t)$	instantaneous short spread (short spread)
$h(t,T)$	continuously compounded instantaneous forward rate spread (forward spread)
$\sigma^h(t,T)$	volatility of forward spread
$\alpha^h(t,T)$	drift of forward spread

Default Models

Default Time

N	point process whose first jump triggers the default
λ	intensity of N
Y_i	marker to T_i
$K(dY)$	distribution of Y
T_i, τ_i	time of the i-th jump of N
τ	time of default ($\tau = \tau_1$)
F	price of a general derivative security
X	payoff of F
F'	price of a derivative security like F, but defaultable
μ_λ	drift of λ
σ_λ	volatility of λ

Recovery Models

π	cash recovery model: recovery $= \pi$ in cash
c	equivalent recovery model: recovery $= c$ default -free bonds
q	fractional recovery model: payoff reduced by q at each default
Q	fractional recovery model: final payoff $Q(T) = \prod_{\tau_i \le T}(1 - q_{\tau_i})$

Probabilities

$P(t,T)$	probability of survival from t to T
$P^{\mathrm{def}}(t,T)$	probability of default in $[t,T]$
$\widetilde{P}(t,T)$	pseudo survival probability: $\widetilde{P}(t,T) = \overline{B}(t,T)/B(t,T)$

Firm's Value Model

V	firm's value
$\tilde{V}$	forward price of the firm's value
μ	drift of V
σ	volatility of V
$\bar{S}$	barrier triggering default
$\bar{D}$	total face value of debt outstanding
$\bar{S}$	total number of shares issued
c	bankruptcy costs
$\bar{c}$	bankruptcy costs as proportion of the price of a default-free bond
T	maturity of the debt
$\overline{B}(t,V,r,T)$	price at t of the defaultable bond with maturity T given firm's value V
$\tilde{\overline{B}}$	forward price of this bond. $B\tilde{\overline{B}} = \overline{B}$
$S(t,V)$	price at t of the firm's share given firm's value V
$B(t,r,T)$	price at t of the default-free bond with maturity T
r	the risk-free short rate
μ_r	drift of the risk-free short rate r
σ_r	volatility of the risk-free short rate
W	Brownian motion driving firm's value
$\tilde{W}$	Brownian motion driving interest rates
ρ	correlation between W and $\tilde{W}$

The CIR Model

x_i	i-th factor $(i = 1, \dots, n)$
w	factor weights of the default-free short rate: $r(t) = \sum_{i=1}^{n} w_i x_i(t)$
$\overline{w}$	factor weights of the intensity: $\lambda(t) = \sum_{i=1}^{n} \overline{w}_i x_i(t)$
η	weights of affine combination of noncentral chi-squared RVs

The Ratings Model

K	number of rating classes
$R(t)$	rating at time t
A	generator matrix
$P(s,t)$	transition probability matrix

MDM^{-1}	decomposition of P
MD_AM^{-1}	decomposition of A
F'	price of derivative conditioning on ratings
F	price of equivalent default-free derivative
F^*	payoff of F'

Preliminaries

In this section you will learn ...

- ... conditions for absence of arbitrage,
- ... risk-neutral pricing,
- ... Itô's lemma.

Risk-Neutral Pricing and Absence of Arbitrage

In every arbitrage-free securities market one can assign probabilities P such that the price of every security F is equal to the expectation of its discounted payoff, if one uses the probabilities P to form the expectation. Mathematically this means

$$F(t) = \mathrm{E}_P\left[\exp\{-\int_t^T r(s)ds\}F(T)\right], \tag{1}$$

where we discounted with the risk-free short rate r. The probabilities P are known as the *risk-neutral probabilities*, because once these probabilities are known we can use them to price *as if* we were risk-neutral, (even though we are not). The risk-neutral probabilities can also interpreted as *the market's opinion*, the risk-neutral probability of an event is how likely the market thinks this event is.

The risk-neutral probabilities can therefore usually be implied from market data, provided there are no arbitrage opportunities. Arbitrage opportunities would mean an inconsistent 'opinion' of the market, the market contradicts itself. Absence of arbitrage means absence of contradictions, hence a consistent set of pricing probabilities can be found.

In equation (1) the second time T was arbitrary, so we can choose T to be very close to t, i.e. $T = t + \Delta t$. Then we can look at the local properties of the *increments* of F:

$$\begin{aligned} F(t) &= \mathrm{E}_P\left[\exp\{-\int_t^{t+\Delta t} r(s)ds\}F(t+\Delta t)\right] \\ 0 &= \mathrm{E}_P\left[\exp\{-\int_t^{t+\Delta t} r(s)ds\}F(t+\Delta t) - F(t)\right] \end{aligned}$$

$$0 = \mathrm{E}_P\left[\exp\{-\int_t^{t+\Delta t} r(s)ds\}(F(t) + dF(t)) - F(t)\right]$$
$$0 = \mathrm{E}_P\left[\exp\{-r(t)\Delta t\}(F(t) + dF(t)) - F(t)\right]$$
$$0 = \mathrm{E}_P\left[\exp\{-r(t)\Delta t\}dF(t)\right] - (1 - \exp\{-r(t)\Delta t\})F(t).$$

Now we can use the approximation $e^x = 1 + x$ which is valid for small x to reach

$$0 = \mathrm{E}_P\left[(1 - r(t)\Delta t)dF(t)\right] - (1 - (1 - r(t)\Delta t)F(t)$$
$$0 = \mathrm{E}_P\left[(1 - r(t)\Delta t)dF(t)\right] - r(t)\Delta t F(t)$$
$$r(t)F(t)dt = \mathrm{E}_P\left[dF(t)\right]. \tag{2}$$

Equation (2) states that *the risk-neutral expectation of the rate of growth of a security price* is the short term interest rate r.

This result will drive the derivation of most of the pricing equations later on.

Itô's lemma

The main tool used in continuous time finance is Itô's lemma. It gives the increments of a smooth function $F(t, X)$ of time t and a n-dimensional diffusion process X given by

$$dX_i = \mu_i dt + \sum_{j=1}^{n} \sigma_{ij} dW^j,$$

where W is a n-dimensional Brownian motion and μ and σ satisfy some regularity conditions.

The increments of $F(t, X)$ are given by

$$dF = \frac{\partial}{\partial t}F\,dt + \sum_{i=1}^{n}\frac{\partial}{\partial x_i}F\,dX_i + \frac{1}{2}\sum_{i,j=1}^{n}\left(\sum_{k=1}^{n}\sigma_{ik}\sigma_{jk}\right)\frac{\partial^2}{\partial x_i\,\partial x_j}F\,dt. \tag{3}$$

For one dimension Itô's lemma has the simpler form

$$dF = \frac{\partial}{\partial t}F\,dt + \frac{\partial}{\partial x}F\,dX + \frac{1}{2}\sigma^2\frac{\partial^2}{\partial x^2}F\,dt. \tag{4}$$

The first two terms in the Itô lemma are what would be expected from a standard Taylor expansion of F, but now we have to add the second-derivative terms in X because the increments of X are of order $\sqrt{dt}$.

Chapter 1

Firm's Value Models

1.1 The Approach

1.1.1 In this section you will learn ...

- ... the modelling philosophy needed to understand firm's value models for defaultable bond pricing,
- ... how to directly transfer your equity derivatives knowledge to the pricing of defaultable bonds in a firm's value model,
- ... what ingredients and fundamentals you need to build your own firm's value model,
- ... the special characteristics of the mechanism triggering defaults in this class of models.

1.1.2 Modelling Philosophy

In firm's value models a fundamentalist's approach to valuing defaultable debt is taken: We assume there is a fundamental process V, usually interpreted as the total value of the assets of the firm that has issued the bonds in question. The value of the firm V is assumed to move around stochastically. It is the driving force behind the dynamics of the prices of all securities issued by the firm, all claims on the firm's value are modelled as *derivative securities with the firm's value as underlying.*

Default can be triggered in two ways: Either V is only used to pay off the debt at the maturity of the contract. A default occurs at maturity if V is insufficient to pay back the outstanding debt but during the lifetime of the contract a default can not be triggered.

Alternatively (and more realistically) one can assume that a default is already triggered as soon as the value of the collateral V falls below a barrier $\bar{S}$. This feature is exactly identical to a standard knockout barrier in equity options.

1.1.3 An Example

Assume firm ABC has issued zero coupon bonds of maturity $T = 2$ years with a total face value $\bar{D}$ of S 100 Mio. The value V of the firm's assets is currently S 150 Mio and follows a geometric Brownian motion

$$dV = \mu V dt + \sigma V dW.$$

We can observe the value of the firm's assets but are unable to intervene before the maturity of the debt. Both bonds and shares issued by the firm are actively traded. The firm has issued $\bar{S}$ shares. Interest rates are constant r.

- What should be the value of the firm's debt?
- What should be the value of the firm's shares?
- Can we hedge?

The state variable here is the firm's value V. We write the prices of both debt $\overline{B}(V,t)$ and shares $S(V,t)$ as functions of firm's value V and time t. For ease of notation S denotes the price of *all* shares and $\overline{B}$ of *all* bonds. For the prices of the individual securities we will have to divide these by the respective numbers $\bar{D}$ and $\bar{S}$.

The payoffs at time T are

$$\overline{B}(V,t) = \min(1, V) \tag{1.1}$$
$$S(V,t) = \max(V - \bar{D}, 0). \tag{1.2}$$

The payoff of the shares is exactly the payoff of an European Call option on the firm's value, the payoff of the bond is either its face value (if the firm's value is above $\bar{D}$ at T), or whatever is left of the firm's value V if it is below the face value of the debt. The payoffs are shown in figure 1.1.

This enables us to directly price the share: Its value is given by the Black-Scholes Formula for a European Call option:

$$S(V,t) = \mathrm{BSC}(V, t; T, \bar{D}, \sigma, r),$$

where $\mathrm{BSC}(V, t; T, \bar{D}, \sigma, r)$ denotes the Black-Scholes price[1] of a European call option with expiry date T and exercise price $\bar{D}$, where the underlying volatility is σ and the interest rate is r.

As shares and bonds together must give the total value of the firm, we get directly

$$V = \overline{B}(V,t) + S(V,t) \quad \Leftrightarrow \quad \overline{B}(V,t) = V - S(V,t).$$

Otherwise there would be an arbitrage opportunity. (Where?)

[1] I.e. $\mathrm{BSC}(V, t; T, \bar{D}, \sigma, r) = VN(d_1) - e^{-r(T-t)}\bar{D}N(d_2)$ where

$$d_1 = \frac{\ln V/\bar{D} + (r - \frac{1}{2}\sigma^2(T-t))}{\sigma\sqrt{T-t}}$$

and $d_2 = d_1 - \sigma\sqrt{T-t}$.

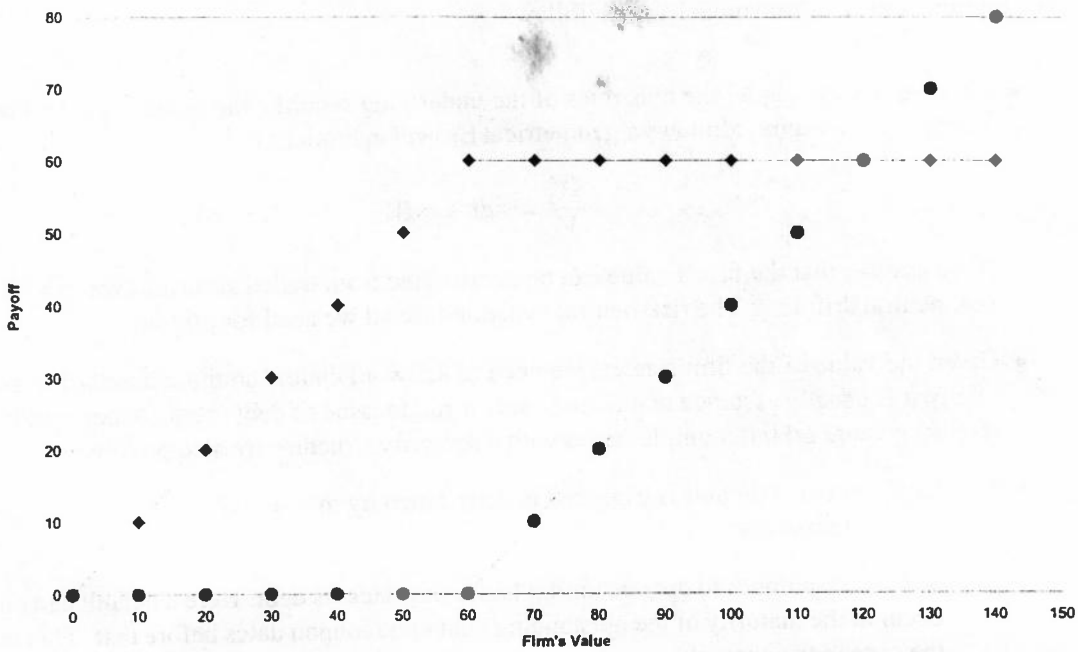

Figure 1.1: Payoffs of share and bond at $t = T$

As we have two securities with only one underlying source of uncertainty we are able to hedge the bond with the share or vice versa. Let's consider the case of hedging the bond with the share: We set up a portfolio Π consisting of one bond and Δ shares. Its value is:

$$\Pi = \overline{B}(V,t) + \Delta S(V,t)$$

and by Itô's lemma its change in value over a short interval is

$$d\Pi = d\overline{B} + \Delta dS \tag{1.3}$$

$$= \left(\frac{\partial \overline{B}}{\partial t} + \frac{1}{2}\frac{\partial^2 \overline{B}}{\partial V^2} + \Delta\frac{\partial S'}{\partial t} + \frac{1}{2}\Delta\frac{\partial^2 S'}{\partial V^2}\right) dt \tag{1.4}$$

$$+ \left(\frac{\partial \overline{B}}{\partial V} + \Delta\frac{\partial S}{\partial V}\right) dV. \tag{1.5}$$

To eliminate the stochastic dV-term from the portfolio we only have to choose

$$\Delta = -\frac{\frac{\partial \overline{B}}{\partial V}}{\frac{\partial S}{\partial V}},$$

then the portfolio is fully hedged and its return is predictable.

Alternatively, although hedging in the classical sense directly with V is not possible, we can use the relation $S + \overline{B} = V$ to synthesize V from a portfolio of one share and one bond to replicate the firm's value.

1.1.4 State Variables and Modelling

The input to a firm's value model are the following:

- First we have to model the dynamics of the underlying security, the value V of the firm's assets. Let's assume it follows a geometrical Brownian motion:

  $$\frac{dV}{V} = rdt + \sigma dW.$$

  If we assume that the firm's value can be constructed from traded securities we can set its risk-neutral drift to r: The risk-neutral dynamics are all we need for pricing.

- Given the value of the firm's assets we need to know all claims on these assets. For simplicity it is usually assumed that there is only a single issue of debt (zero coupon bonds of total face value $\bar{D}$) but multiple issues with a seniority structure are also possible.

- The way in which a default is triggered is determined by the capital structure. Here there are several alternatives:

  - The firm continues to operate until it has to pay back its debt. Here a default can only occur at the maturity of the outstanding debt or at coupon dates before that. This was the case in the example.
  - There are safety covenants in the issued debt that allow the creditors to close down and liquidate the firm if the value of its assets should fall below a certain level $\bar{S}$. Default occurs as soon as

    $$V_t \leq \bar{S}.$$

  - The level $\bar{S}$ is not constant but time-dependent. For example the covenant may be such that a default is only triggered if the firm's value falls below the *discounted* value of the assets outstanding, i.e. when the firm's value is worth less than a similar but default-free investment. Then default occurs as soon as

    $$V_t \leq \bar{S}B(t,T),$$

    where $B(t,T)$ is the value of a default-free bond of the same specification as the debt outstanding.

  - In a practical implementation one could go and exactly implement to the word the covenants of the debt contract and firm that one wants to model. This is not done very often because the accuracy gained here is swamped by other cruder approximations one has to make in other parts of the model. (See the critique of the model.)

  From now on we will assume the third case: Default occurs at the barrier $V_t = \bar{S}B(t,T)$. The state space (or its firm's value section) is shown in figure 1.2.

- The capital structure of the firm has another consequence, it determines the payoffs to the different securities:

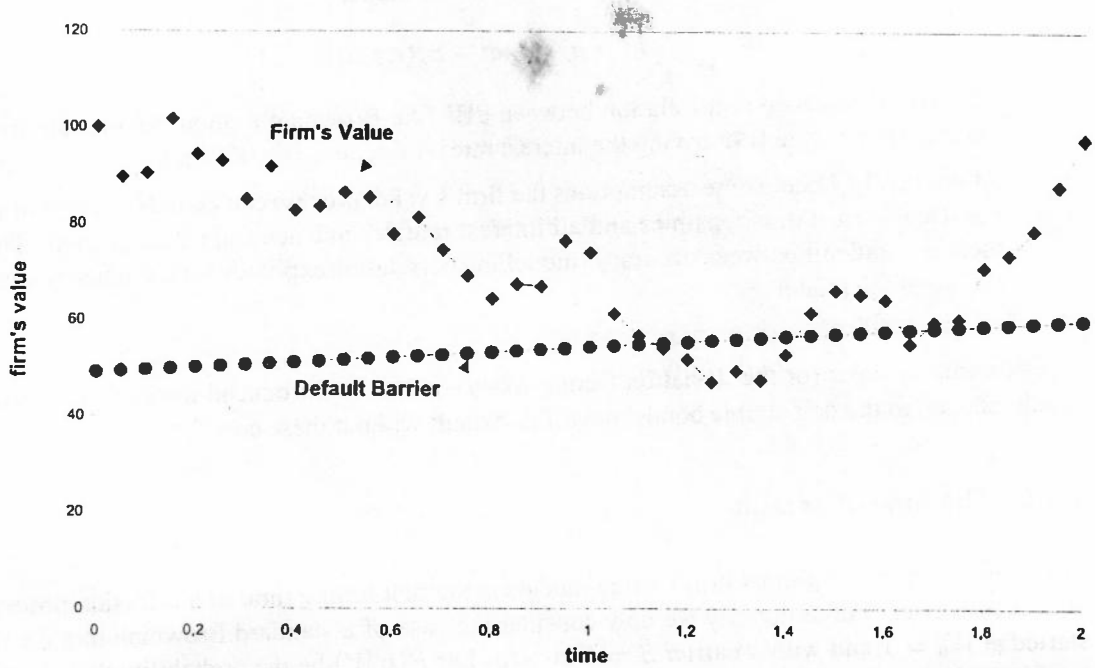

Figure 1.2: Default is triggered when the firm's value hits the barrier.

- The bonds pay off their face value $\bar{D}$ if there is no default, and their fraction of the value of the firm minus some bankruptcy costs c in the case of a default. The payoff function is:

$$\overline{B}(V, t = T) = \min\{\bar{D}, V\}$$

at the final payoff date, and

$$\overline{B}(V = \bar{S}, t) = V - c = \bar{S} - c$$

if a default is triggered in the meantime.

- The shares pay off $S(V, t = T) = (V - \bar{D})^+$ at maturity of the debt, and nothing if a default is triggered $S(V = \bar{S}, t) = 0$. This means we model strict absolute priority. Deviations from priority can be incorporated with a different payoff distribution at $V = \bar{S}$. For instance we could say that the bankruptcy costs c to the bondholders result from deviations from strict absolute priority[2]. Then $S(V = \bar{S}, t) = c$.
- If there is debt of different seniority classes this can be reflected in an appropriate modification of the payoffs at the default barrier.

- Finally we would like to incorporate interest rate uncertainty. Here we have two alternatives: If we assume correlation between the dynamics of the firm's value and the interest

[2]Empirical studies [1994] show, that in most bankruptcies and distressed reorganisations some deviation from absolute priority occurs.

rate dynamics, we will have to add another state variable to our model, the short term risk free interest rate r. Assuming a general one-factor model

$$dr = \mu_r(r, t)dt + \sigma_r(r, t)d\tilde{W},$$

we have instantaneous correlation between dW (the Brownian motion driving the firm's value) and $d\tilde{W}$ (the BM driving the interest rates) of ρ, i.e. $dWd\tilde{W} = \rho dt$.

Alternatively, under some assumptions the firm's value process can be independent of the risk-free interest rate dynamics and all interest rate dependence can be eliminated. Thus there is a tradeoff between accuracy (modelling correlation explicitly) and efficiency in the choice of the model.

The bankruptcy costs (or the deviation from priority) had to be introduced for a reason: What would happen to the defaultable bonds' payoff in default without these costs?

1.1.5 The time of default

The time of default τ in most firm's value models is the first hitting time of a diffusion process at a fixed barrier. For simplicity we now consider the case of a standard Brownian motion W started at $W_0 = 1$ and with a barrier $\bar{S} = 0$ at zero. Let $P(t, W)$ be the probability that W has *not* hit the barrier until time T, given that the Brownian motion is at W at time t and that it has not hit the barrier before t.

We know that $P(T, W) = 1$, at $t = T$ there is no time left to hit the barrier and the survival probability is one. We also know that $P(t, 0) = 0$, at $W = 0$ the barrier is already reached, the probability of evading it is zero. On the other hand, $P(t, W) \to 1$ as $W \to \infty$: as W becomes very large the survival probability approaches one.

In the meantime, for $0 < W < \infty$ and $t < T$ we know that P, being a function of W, has to satisfy Itô's lemma:

$$dP = (\frac{\partial P}{\partial t} + \frac{1}{2}\frac{\partial^2 P}{\partial W^2})dt + \frac{\partial P}{\partial W}dW.$$

But P cannot have a drift as a stochastic process. P is the best estimate of the survival indicator function $\mathbf{1}_{\{\tau > T\}}$, therefore *all* changes in P have to be purely stochastic and unpredictable. If there were a positive drift term in dP, we would know that the survival probability tomorrow will be higher than it is today, but this knowledge should already be incorporated in today's probability. Mathematically speaking:

$$P(t, W(t)) = \mathbf{E}\left[\, \mathbf{1}_{\{\tau > T\}} \,\right] = \mathbf{E}\left[\, P(t+s, W(t+s)) \,\right],$$

P has the property that it is its own future expectation, it is a martingale. Here we can use this to set the dt-term in P equal to zero:

$$0 = \frac{\partial P}{\partial t} + \frac{1}{2}\frac{\partial^2 P}{\partial W^2}.$$

P satisfies the forward heat equation with the final and boundary conditions given above. The solution to this equation is

$$P(t, W) = N(\frac{x}{\sqrt{t}}) - N(-\frac{x}{\sqrt{t}}).$$

Typical credit spreads that result from this survival probability are plotted in figure 1.3.

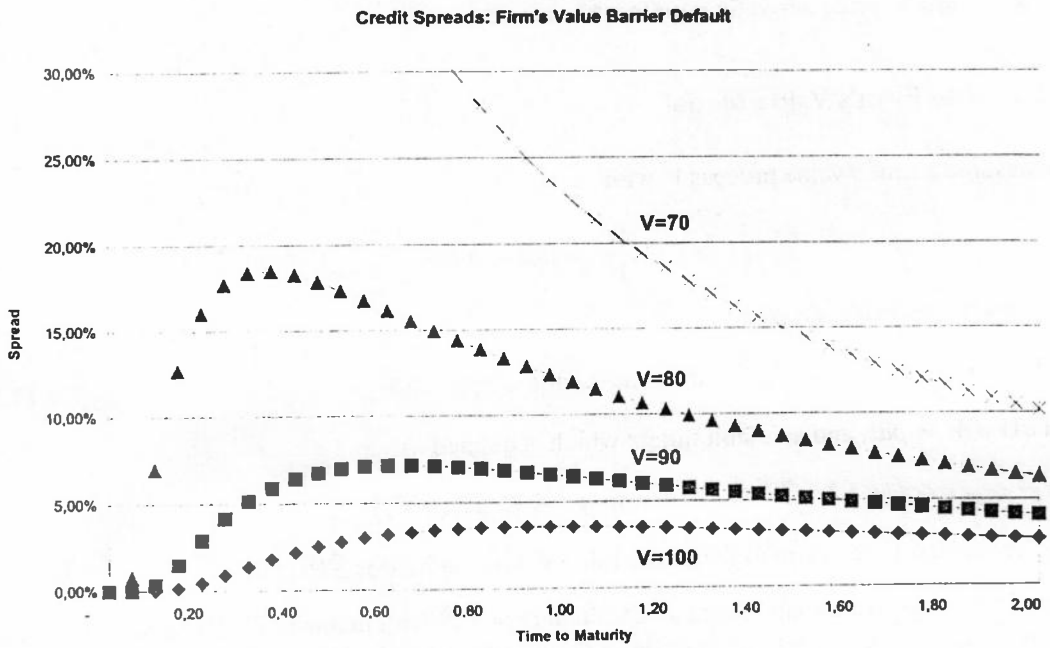

Figure 1.3: Survival Probabilities

We can see that the credit spreads become very low for short times to maturity. This is caused by the very flat shape of the survival probability P as $T - t \to 0$. (Remember that the spreads at the short end are determined by the slope of the survival probabilities.)

In firm's value models the probability of a default in a very short time interval $[t, t + \Delta t]$ from now is very small, it is smaller than of order Δt: Even if we divide the default probability by $\Delta t$the result goes to zero as $\Delta t \to 0$:

$$\frac{1}{\Delta t}(1 - P(t, t + \Delta t)) \to 0.$$

Intuitively speaking, we know, that the firm's value follows a diffusion, that it has continuous paths and cannot jump. So if we know that we are away from the knockout barrier we also know, that the firm's value cannot reach it at once – it cannot jump. Therefore the short term default probability goes to zero very quickly.

1.2 Pricing Equations

1.2.1 In this section you will learn ...

- ... how to derive a pricing equation in a firm's value model,
- ... how to price a variety of securities issued by a specific firm,
- ... how to price derivatives on these securities.

1.2.2 The Firm's Value Model

We assume a firm's value process V with

$$\frac{dV}{V} = \mu dt + \sigma dW, \tag{1.6}$$

an interest rate process r with

$$dr = \mu_r(r,t)dt + \sigma_r(r,t)d\tilde{W}, \tag{1.7}$$

and $d\tilde{W}dW = \rho dt$, and a default time τ which is defined as

$$\tau = \min\{t \mid V_t \leq \bar{S}B(t,T)\}$$

i.e. τ is the first time at which the firm's value V hits the barrier $\bar{S}B(t,T)$.

Traded securities are a share S and a defaultable bond $\overline{B}$ with maturity T. The total numbers of shares $\bar{S}$ and bonds $\bar{D}$, and the constant in the knockout barrier $\bar{S}$ are all normalized to 1. The share pays off

$$S(V,T) = (V-1)^+$$

at maturity and the bond pays off

$$\overline{B}(V,T) = 1 - (1-V)^+,$$

assuming there was no default before. If there is a default the defaultable bond pays off

$$\overline{B}(V = \bar{S}B(t,r,T),t) = V - c = B(t,r,T)(\bar{S} - \bar{c})$$

and the shares pay off

$$S(V = \bar{S}B(t,r,T),t) = B(t,r,T)\bar{c}$$

in default, assuming a deviation from absolute priority by $c = \bar{c}B(t,r,T)$ in favour of the shareholders[3].

Furthermore we have a full term structure of traded default-risk free bonds $B(t,T)$.

[3] Writing the deviation from absolute priority as a proportion $\bar{c}$ of the default-free bond price will faciliate the analysis later on.

1.2.3 The Pricing Equation

Like in the introductory example we now consider all securities as derivatives on the firm's value V. There is a knockout-barrier at $\bar{S}B(t,T) = B(t,T)$ where the default payoffs are triggered. Otherwise we only have the contractually specified payoffs.

The share and the bond added together have the firm's value as payoff in all states, thus we can consider the firm's value as a traded security. We know that under risk-neutral valuation traded securities have a drift term of rdt, thus we replace (1.6) with:

$$dV = rVdt + \sigma V dW. \tag{1.8}$$

We assume, that the risk-neutral dynamics of the interest rate is already given by (1.7) so that now everything is ready for the derivation of the pricing equation.

Bond and share are functions of firm's value V, interest rate r and time t. Itô's lemma gives us their dynamics:

$$\begin{aligned} d\overline{B} =& \left(\frac{\partial \overline{B}}{\partial t} + \frac{1}{2}\sigma^2 V^2 \frac{\partial^2 \overline{B}}{\partial V^2} + \rho\sigma\sigma_r V \frac{\partial^2 \overline{B}}{\partial V\, \partial r} + \frac{1}{2}\sigma_r^2 \frac{\partial^2 \overline{B}}{\partial r^2}\right) dt \\ &+ \frac{\partial \overline{B}}{\partial V} dV + \frac{\partial \overline{B}}{\partial r} dr, \end{aligned} \tag{1.9}$$

the share price dynamics dS are analogous. Now we apply again the fact that the risk-neutral drift must be $rBdt$ to reach

$$\begin{aligned} r\overline{B}dt =& \left(\frac{\partial \overline{B}}{\partial t} + \frac{1}{2}\sigma^2 V^2 \frac{\partial^2 \overline{B}}{\partial V^2} + \rho\sigma\sigma_r V \frac{\partial^2 \overline{B}}{\partial V\, \partial r} + \frac{1}{2}\sigma_r^2 \frac{\partial^2 \overline{B}}{\partial r^2}\right) dt \\ &+ rV \frac{\partial \overline{B}}{\partial V} dt + \mu_r \frac{\partial \overline{B}}{\partial r} dt \end{aligned}$$

and hence $\overline{B}$ has to satisfy the partial differential equation

$$\begin{aligned} 0 =& \frac{\partial \overline{B}}{\partial t} + \frac{1}{2}\sigma^2 V^2 \frac{\partial^2 \overline{B}}{\partial V^2} + \rho\sigma\sigma_r V \frac{\partial^2 \overline{B}}{\partial V\, \partial r} + \frac{1}{2}\sigma_r^2 \frac{\partial^2 \overline{B}}{\partial r^2} \\ &+ rV \frac{\partial \overline{B}}{\partial V} + \mu_r \frac{\partial \overline{B}}{\partial r} - r\overline{B}. \end{aligned} \tag{1.10}$$

Equation (1.10) is the fundamental partial differential equation every security has to satisfy. Bonds, shares and all other securities on the firm's value have to satisfy this partial differential equation, they only differ in the final- and boundary conditions that apply.

The final condition for the bond is its payoff given no default (as discussed before):

$$\overline{B}(T, V, r) = \min\{1, V\}, \tag{1.11}$$

and the boundary conditions are

$$\overline{B} = B(t,T)(1 - \bar{c}) \qquad \text{at} \qquad V = \bar{S}B(t,T) \tag{1.12}$$

$$
\begin{aligned}
\overline{B} &\to B(t,T) && \text{as} \quad V \to \infty \\
\overline{B} &\to 0 && \text{as} \quad r \to \infty \\
\overline{B} &< \infty && \text{at} \quad r = 0.
\end{aligned}
$$

Most of these conditions have an obvious intuitive interpretation except for the condition at $r = 0$ which had to be included to preclude a possible singularity of the solution.

1.2.4 Some other securities

As mentioned before, by applying different boundary- and final conditions to (1.10) we can price a variety of other securities:

1.2.4.1 Coupon Payments

For *Coupon Bonds* we need to add the coupon payments to the dynamics of V and the bond price $\overline{B}_c$. At a coupon date T_i the coupon C is paid from the firm's value. Thus the firm's value has to decrease by this amount: $V(T_i+) = V(T_i-) - C$, where $V(T_i-)$ is the firm's value shortly before T_i and $V(T_i+)$ shortly after the coupon payment at T_i.

The bond price has to make the same jump downwards, if a coupon payment is due at T_i, the price shortly *before* the payment i.e. $\overline{B}_c(T_i-, V(T_i-), r)$ is higher by the amount of the coupon C than the price *after* the coupon payment $\overline{B}_c(T_i+, V(T_i+), r)$. Again we see the similarity to equity options where we have a similar effect at dividend days with discrete dividends. We reach the condition

$$\overline{B}_c(T_i-, V, r) = \overline{B}_c(T_i+, V - C, r) + C. \tag{1.13}$$

Condition (1.13) is easily incorporated In a numerical time-stepping scheme to solve the pricing equation (1.10). The situation does not change for floating rate coupons except that then C is a function of r: $C = C(r)$.

1.2.4.2 Convertible Bonds

A particular strength of the firm's value model is the naturally arising connection between share- and bond prices. This can be exploited to price convertible bonds. A convertible bond can be converted into a certain number α of shares of the issuing firm. This conversion can take place either at certain fixed times or at any time. Whenever a conversion is allowed we must have

$$\overline{B}(t, V, r) \geq \alpha S(t, V, r),$$

because we can always convert the bond and secure a payoff of $\alpha S(t, V, r)$. Equality only holds in the above equation when conversion is optimal. Using the relation $V = \overline{B} + S$ this condition becomes

$$\overline{B}(t, V, r) \geq \frac{\alpha}{1 + \alpha} V. \tag{1.14}$$

This condition is similar to the early exercise conditions in American options. In the numerical solution, when stepping backwards in time, condition (1.14) has to be checked whenever conversion is allowed. If the price from the backward stepping scheme does not satisfy the inequality it has to be adjusted to satisfy (1.14) with equality. At these points one should convert[4]. The optimal exercise boundary is then at the values of the firm's value where conversion just became optimal.

Rule (1.14) takes *dilution* into account. If new shares are issued for the conversion the share price after conversion will be lower than before. After conversion of all bonds we have α additional shares, but no bonds left. Thus $1+\alpha$ shares own the firm's value V, one share is worth $S = \frac{V}{1+\alpha}$; thus, after conversion of the bonds you get $\frac{\alpha}{1+\alpha}V$, exactly as given in equation (1.14).

1.2.4.3 Callable Bonds

A callable bond can be bought back by the issuer at pre-specified times for a pre-specified price B^*. Again this feature has some similarity to American options, now the issuer has an American call on the bond, and the bondholders are short this call, they have to deliver if their bond is called. Thus there is another inequality the bond price has to satisfy at call-dates:

$$\overline{B}(t, V, r) \leq B^*$$

the bond price must be less than the call price (otherwise it will be called).

1.2.4.4 Derivatives on Defaultable Bonds

A derivative on a defaultable bond is priced in conjunction with the pricing of the underlying bond. Let's assume the derivative has a payoff at time T_1 which is given by the function

$$F(\overline{B}, V, r)$$

which has the price of the defaultable bond $\overline{B}$ with maturity $T_2 > T_1$, the firm's value and the risk-free interest rate as argument[5].

We could have an option on the defaultable bond, e.g. an exchange option to exchange the defaultable bond $\overline{B}$ for α default free bonds B of otherwise equivalent specification. Its payoff is

$$F(\overline{B}, V, r) = (\alpha B(T_1, r) - \overline{B}(T_1, V, r))^+.$$

[4]It may seem that by adjusting the price upward by a large amount an arbitrage opportunity is introduced. This is not so: If the price has to be adjusted by a large amount, this means that we have been going backwards in time from a region where conversion was not allowed to a region where conversion is allowed and optimal. Or forward looking: The conversion opportunity is about to end. When it ends and you still hold a bond you have missed the opportunity and lose the value of the conversion, but you should have converted already anyway. But you cannot profit from this drop in value as all bonds will have been converted and they would not be available to short them for the subsequent drop in value.

[5]Strictly speaking the defaultable bond's price $\overline{B}$ is not really needed as an argument of F because it can be represented in terms of V and r itself.

The yield spread of the defaultable bond $\overline{B}$ over the default-free bond B is given by

$$s(\overline{B}, V, r) = -\frac{\ln \overline{B}(T_1, V, r) - \ln B(T_1, r)}{T_2 - T_1}.$$

We can imagine a large variety of derivatives conditioning on the size of the spread. A caplet for instance would have a payoff like

$$F(\overline{B}, V, r) = L\Delta T(s - \bar{s})^+$$

where L is the principal and ΔT the time interval.

Besides the regular payoff at T_1, the expiry date of the derivative, we also have to specify the payoff the derivative will have if there is a default before T_1. This depends on the specification of the contract.

As derivative and underlying both have to satisfy (1.10) they can be priced simultaneously, at each step of the backward induction we now calculate *two* prices, the price of the underlying bond and the price of the derivative. Starting from the maturity T_2 of the underlying bond, equation (1.10) is solved for the price of the underlying until the expiry date T_1 of the derivative. Then the derivative's final condition F can be applied and backwards from there we can solve for both prices simultaneously.

1.2.5 Hedging

There is nothing special about Delta-hedging in this model. Given a price process $C(t, V, r)$ of the security to be hedged, and $\overline{B}(t, V, r)$ and $B(t, r)$ the prices of two hedge instruments, a defaultable bond and a default-risk free bond. We need two hedge instruments because we have two underlying sources of risk: V and r.

We are looking for values of Δ' and Δ that make the hedged portfolio

$$\Pi = C(t, V, r) + \Delta' \overline{B}(t, V, r) + \Delta B(t, r)$$

locally risk-free. Applying Itô's lemma we see that Δ' and Δ must satisfy

$$-\frac{\partial C}{\partial V} = \Delta' \frac{\partial \overline{B}}{\partial V} \tag{1.15}$$

$$-\frac{\partial C}{\partial r} = \Delta' \frac{\partial \overline{B}}{\partial r} + \Delta \frac{\partial B}{\partial r} \tag{1.16}$$

which solves to

$$\Delta' = -\frac{\frac{\partial C}{\partial V}}{\frac{\partial \overline{B}}{\partial V}} \tag{1.17}$$

$$\Delta = -\frac{1}{\frac{\partial B}{\partial r}} \left(\frac{\partial C}{\partial r} + \Delta' \frac{\partial \overline{B}}{\partial r} \right). \tag{1.18}$$

Obviously, close to the default barrier we can expect to encounter large Gammas, this problem is well-known from equity or FX barrier options, but in theory all securities in this model can be hedged perfectly.

1.3 Solutions to the Pricing Equation

1.3.1 In this section you will learn ...

- ... how to solve the pricing problem using probability theory,
- ... a closed-form solution for a defaultable zero-coupon bond in our model,
- ... the semi-closed form solutions of Longstaff and Schwartz.

The results of this section are a special case of the results of Briys and de Varenne [1997]. For details and extensions we refer to this reference.

We will assume a one-factor Gaussian term structure model of the Ho-Lee [1986] type for the initial term structure of defaultable bonds. This means, that the risk-neutral dynamics of the default-free zero-coupon bond with maturity T are given by

$$\frac{dB}{B} = rdt + \sigma_r(T-t)dW_1. \tag{1.19}$$

Furthermore we rewrite the dynamics of the firm's value process as

$$\frac{dV}{V} = rdt + \sigma_V(\rho dW_1 + \sqrt{1-\rho^2}\, dW_2). \tag{1.20}$$

In equation (1.20) we have explicitly represented the correlation between firm's value and interest rates with ρ and two orthogonal (=uncorrelated) Brownian motions W_1 and W_2.

1.3.2 The T-Forward Measure

The price of a defaultable zero coupon bond can be represented as

$$\overline{B} = B(1 - cP^T[\tau < T]) = B(1 - c(1 - P^T[\tau \geq T])), \tag{1.21}$$

where P^T denotes the probability under the T-forward measure and B is the price of a default-free zerobond with maturity T.

The event $\{\tau \geq T\}$ is equivalent to

$$\{\tau \geq T\} \quad \Leftrightarrow \quad \{\tilde{V} \geq \tilde{S} \quad \forall t \leq T\} \tag{1.22}$$

where $\tilde{V} = V/B$ has the following dynamics under P^T:

$$\begin{aligned}\frac{d\tilde{V}}{\tilde{V}} &= (\rho\sigma_V - \sigma_r(T-t))dW_1^T + \sigma_V\sqrt{1-\rho^2}dW_2^T \\ &=: \sigma(t)d\hat{W}\end{aligned} \tag{1.23}$$

where

$$\sigma(t) = [\sigma_V^2 - 2\rho\sigma_V\sigma_r(T-t) + \sigma_r^2(T-t)^2]^{1/2} \tag{1.24}$$

and $\hat{W}$ is a BM constructed from W_1^T and W_2^T. Now we have reduced the problem to calculating the probability of the event given in (1.22), the hitting probability of the process $\tilde{V}$ which follows a geometric BM (1.23) with time-dependent volatility (1.24).

1.3.3 Time Change

To eliminate the time-dependence in the volatility $\sigma(t)$ we perform the following time-change: Consider the process M given by

$$dM = \sigma(t)d\hat{W}.$$

Its quadratic variation is given by

$$\begin{aligned} < M >_t =: Q(t) &= \int_0^t \sigma(s)^2 ds \\ &= -\sigma_V^2(T-t) + \rho\sigma_V\sigma_r(T-t)^2 - \frac{1}{3}\sigma_r^2(T-t)^3 \\ &\quad + \sigma_V^2 T - \rho\sigma_V\sigma_r T^2 + \frac{1}{3}\sigma_r^2 T^3, \end{aligned}$$

and the value of this process M_t at time t can be represented as the value of a Brownian motion $W_{Q(t)}$ at time $Q(t)$, M is a time-changed BM (see Karatzas and Shreve [1991] theorem 3.4.6. p.174):

$$M_t = W_{Q(t)}.$$

Also (and equally important to us) we can define the same time change for $\tilde{V}$:

$$Y_{Q(t)} = \tilde{V}_t,$$

then (Karatzas and Shreve [1991] theorem 3.4.8. p.176)

$$\tilde{V}_t = \int_0^t \tilde{V}_v dM_v = \int_0^{Q(t)} Y_u dW_u = Y_{Q(t)} \tag{1.25}$$

The right hand side of (1.25) means nothing but that Y satisfies the stochastic differential equation

$$dY = YdW, \tag{1.26}$$

Y follows a lognormal random walk itself and it does not have any time dependent volatility.

1.3.4 The hitting probability

Next we look back at the event given in (1.22) whose probability we need to calculate:

$$\{\tilde{V} \geq \bar{S} \quad \forall t \leq T\}$$

$$\begin{aligned}&\Leftrightarrow\{Y \geq \bar{S} \quad \forall t \leq Q(T)\}\\&\Leftrightarrow\{\ln \frac{Y}{Y_0} \geq \ln \frac{\bar{S}}{Y_0} \quad \forall t \leq Q(T)\}.\end{aligned} \tag{1.27}$$

We know that $x := \ln Y$ follows the diffusion

$$dx = -\frac{1}{2}dt + dW,$$

and its barrier hitting probability is well known (see e.g. Musiela and Rutkowski [1997] corrolary B.3.4. p.340). The probability of the event given in (1.27) is

$$N\left(\frac{k + \frac{1}{2}Q(T)}{\sqrt{Q(T)}}\right) - e^{-2k}N\left(\frac{-k + \frac{1}{2}Q(T)}{\sqrt{Q(T)}}\right)$$

where $k = \ln Y_0 - \ln \bar{S}$ is the barrier.

1.3.5 Putting it together

Now we can substitute our results in equation (1.21):

$$\overline{B} = B(1 - c(1 - P)), \tag{1.28}$$

where P is given by

$$P = N\left(\frac{k + \frac{1}{2}Q(T)}{\sqrt{Q(T)}}\right) - e^{-2k}N\left(\frac{-k + \frac{1}{2}Q(T)}{\sqrt{Q(T)}}\right) \tag{1.29}$$

where k and $Q(T)$ are given by

$$k = \ln \frac{V_0}{B_0\bar{S}} \tag{1.30}$$

$$Q(T) = \sigma_V^2 T - \rho\sigma_V\sigma_r T^2 + \frac{1}{3}\sigma_r^2 T^3. \tag{1.31}$$

In this pricing formula we still have scope to fit the risk-free terms structure to an initial term structure.

1.3.6 The Longstaff-Schwartz Results

Longstaff and Schwartz [1995] use a slight modification of the model proposed above to reach semi-closed form solutions for defaultable zero coupon bonds and floating rate bonds.

In their model the risk-free interest rate follows the Vasicek process

$$dr = (\zeta - \beta r)dt + \sigma_r d\tilde{W}$$

with instantaneous correlation of ρ between W and $\tilde{W}$. The firm's value follows the same log-normal diffusion process as given in (1.6), and a default is triggered when V reaches a threshold value $\bar{S}$. In default the bondholders are paid $(1-c)$ default risk free bonds of the same specification. The partial differential equation they solve is exactly equivalent to (1.10). In this framework the price of a defaultable zero coupon bond is

$$\overline{B}(T, X, r) = B(r, T)(1 - cQ(X, r, T)) \tag{1.32}$$

where $X = V/\bar{S}$ and $B(r, t)$ is the value of a default-free zerobond of the same maturity. The function Q is given by

$$Q(X, r, T) = \lim_{n\to\infty} Q(X, r, T, n)$$

where

$$Q(X, r, T, n) = \sum_{i=1}^{n} q_i,$$

and the q_i are defined recursively by

$$q_1 = N(a_1)$$
$$q_i = N(a_i) - \sum_{j=1}^{i-1} q_j N(b_{ij}).$$

The parameters a_i and b_{ij} are now given by

$$a_i = \frac{-\ln X - M(iT/n, T)}{\sqrt{S(iT/n)}}$$
$$b_{ij} = \frac{M(jT/n, T) - M(iT/n, T)}{\sqrt{S(iT/n) - S(jT/n)}}.$$

Here we used the functions M and S which are

$$\begin{aligned} M(t, T) = &\left(\frac{\zeta - \rho\sigma\sigma_r}{\beta} - \frac{\eta^2}{\beta^2} - \frac{\sigma^2}{2}\right) t \\ &+ \left(\frac{\rho\sigma\sigma_r}{\beta^2} + \frac{\eta^2}{2\beta^3}\right) \exp(-\beta T)(\exp(\beta t) - 1) \\ &+ \left(\frac{r}{\beta} - \frac{\alpha}{\beta^2} + \frac{\sigma_r^2}{\beta^3}\right) (1 - \exp(-\beta t)) \\ &- \frac{\sigma_r^2}{2\beta^3} \exp(-\beta T)(1 - \exp(-\beta t)) \end{aligned}$$

and

$$S(t) = \left(\frac{\rho\sigma\sigma_r}{\beta} + \frac{\eta^2}{\beta^2} + \sigma^2\right) t$$

$$- \left(\frac{\rho\sigma\sigma_r}{\beta^2} + \frac{2\eta^2}{\beta^3} \right)(1 - \exp(-\beta t))$$
$$+ \frac{\sigma_r^2}{2\beta^3}(1 - \exp(-2\beta t)).$$

This is a rather long and involved expression which only qualifies as semi-closed form because of the limit $n \to \infty$ that has to be taken. In fact the expression is nothing but a numerical approximation scheme to an integral equation that is encountered when solving the problem. Longstaff and Schwartz propose using $n = 200$ as approximation to the infinite sum. This would imply $\frac{1}{2} \cdot 200(200 - 1) = 19,900$ evaluations of the cumulative standard normal distribution function, and the number of other time-intensive operations like square-roots, logarithms and exponentials is of the same order of magnitude. Thus the formula given by Longstaff and Schwartz will not necessarily be quicker to use than a fast finite-difference approximation to the pricing equation (1.10) which only uses fast multiplications and additions[6].

Another problem is that the model as it stands does not allow for fitting of the interest rate model to an initial term structure or any other modification of the model which makes it highly inflexible. Nevertheless the formula above has been successfully used in practice.

[6]Even if it is not directly apparent to the user, advanced functions (like square roots or exponentials) are several hundred times slower than basic multiplications or additions. The cumulative normal distribution function is again very much slower to evaluate.

1.4 Advantages and disadvantages

Most of the advantages and disadvantages of firm's value models for defaultable bonds have already been mentioned. These models are well suited if the relationship between the prices of different securities issued by the firm is of importance, as in convertible bond or callable bonds that can be converted into shares when called by the issuer. Furthermore the model allows to price defaultable bonds directly from fundamentals, from the firm's value.

The foundation on sound fundamentals makes models of this type also very well-suited for the analysis of questions from corporate finance like the relative powers of shareholders and creditors or questions of optimal capital structure design.

This strength, the orientation towards fundamentals, is also one of the model's weaknesses: Often it is hard to define a meaningful process for the firm's value, let alone observe it continuously. It can be very hard to calibrate such a firm's value process to market prices, and for some issuers, like souvereign debt, it may not exist at all. Furthermore the model may very quickly become too complex to analyse in a real-world application. If one were to model the full set of claims on the value of the assets of a medium sized corporation one may very well have to price twenty or more classes of claims: from banks, shareholders and private creditors down to workers' wages, taxes and suppliers demands. This obviously becomes quickly unfeasible. On the other hand it seems that firm's value models are tailor-made for collateralised loans with traded collateral.

For the above mentioned reasons, practical implementations of the firm's value approach tend not to stick too closely to the model but just take it as a rough guideline. For example, the model by KMV uses the firm's value model only to give an intuitive justification of the key summary statistic that is used in the model: the *distance to default*. Default predictions and risk classifications are then based on an extensive database of historical defaults, and not on any implications from the firm's value model.

An advantage of the firm's value approach which could not become apparent here, is that it provides an easy and intuitive way to incorporate correlations in a portfolio framework. The correlation structure of a multi-factor firm's value model is used by JPMorgan in their Credit Metrics model.

A weakness of the firm's value models is the unrealistic nature of the short term credit spreads implied by the model. As mentioned before these spreads are very low and tend towards zero as the maturity of the debt considered approaches zero.

Finally, for the pricing of credit risk derivatives one would like to have a model where the prices of defaultable bonds can be taken as *fundamentals* and do not have to be calculated (which then necessarily means a calibration process). The next chapter deals with intensity models which can remedy some of these problems.

1.5 Guide to the Literature

The firm's value approach is historically the oldest to the pricing of credit risky securities in modern continuous-time finance. It was first proposed by Black and Scholes in their pathbreaking article "The Pricing of Options *and Corporate Liabilities*" [1973] which already explicitly refers to corporate bond pricing in its title. Merton [1974] expands on this idea. In these models a default can only occur at maturity of the debt, the payoff is like an European option.

In Black and Cox [1976] this approach is extended to allow for intermediate defaults when the firm's value hits a lower boundary. Now the model has more similarity with a Barrier option model. Black and Cox show how to value a variety of corporate bonds and bond covenants in this framework. Further papers using this approach in a risk-free interest rate setup are Merton [1977], Geske [1977], Hull and White [1995], Nielsen et.al. [1993], Schönbucher [1996] and Zhou [1997].

Geske [1977] models defaultable coupon debt as a compound option on the firm's value, where defaults can occur at the coupon dates when the firm's value is insufficient to pay off the coupon. He gives a closed-form solution for a defaultable bond with one intermediate coupon. For a higher number of coupons the compound-options become of a too high order to be represented in simple integrals of the normal density.

Hull and White [1995] consider the problem of counterparty risk in derivatives transactions within the classical firm's value framework. They argue rightly that the interpretation as a firm's value is not necessary for the validity of the model, all that matters is that there is a process which can trigger a default. This allows them to avoid the difficulties in explaining bankruptcy costs consistently within the model.

In Nielsen et.al. [1993] the Black and Cox model is extended by the introduction of a stochastic default barrier. For pricing purposes it can be argued that this unnecessarily complicates the model because the only important quantity of the model the distance of the firm's value from the barrier. If the firm's value is stochastic already, introducing a stochastic barrier would not introduce any new quality.

In Schönbucher [1996] and Zhou [1997] the problem of the low credit spreads for short times to maturity is addressed. As the low credit spreads are caused because it is impossible for a continuous diffusion process to reach the default barrier in a very short time span, the obvious remedy is to introduce jumps into the process of the firm's value. Zhou [1997] proposes to solve the resulting equation with Monte-Carlo simulations while Schönbucher [1996] gives a numerical finite-difference algorithm to solve the resulting pricing equation.

Application of the firm's value approach to the pricing of credit derivatives can be found in Sanjiv Das [1995] and Pierides [1997]. Because of the inflexibility of the firm's value approach and the simplifying assumptions regarding the payoffs of the credit derivatives, these applications are only of limited practical value, their merit is more to give a benchmark for model comparison and calibration.

In the paper of Bensoussan et.al. [1995] the problem of implying the volatility of the firm's value from the volatility of the firm's equity (and vice versa) is adressed and the relationship between both volatilities is analysed in detail.

Longstaff and Schwarz [1995] have managed to reach semi-closed form solutions (an infinite series) for defaultable bonds in a firm's value model with stochastic interest rates that can be correlated with the firm's value process. They use the Vasicek [1977] model for the risk-free interest rates and have to assume a constant initial risk-free term structure. The closed-form solution for stochastic Gaussian interest rates correlated with the firm's value given here in section 1.3., is a special case of the results of Briys and de Varenne [1997]. Briys and de Varenne give closed-form solutions using the discounted default barrier for defaultable bonds within the Vasicek model for default-free interest rates.

Lehrbass [1999] extends the firm's value approach to the modelling of country risk. In this paper the country's stock price index (in international currency) is used as a proxy for the price of the country's 'assets'. Lehrbass tests the pricing and hedging performance of this model an DM Eurobonds of several emerging countries. In the same philosophy is the paper by Cumby and Evans [1995] who empirically fit a Black and Cox - model to Brady bonds of several Latin American countries (Mexico, Venezuela and Costa Rica).

Because of their more explainative approach firm's value models have been popular in more theoretical areas, too. E.g. in an approach initiated by Leland [1994], Leland and Toft [1996] and Mella-Barral and Perraudin [1997] the firm's value framework is used to analyse strategic interaction between debtors and creditors. The bankruptcy barrier is endogenously determined from the optimal behaviour of the debtors and with these models it is possible to satisfactorily explain bankruptcy costs at the default barrier.

Related to this more theoretical class of papers is a very innovative paper by Duffie and Lando [1997], who show in a setup with asymmetric information that there is a close link between firm's value models and the intensity models of the next chapter. They show that a firm's value model with barrier 'a la Black and Cox [1976] will appear as an intensity model to a creditor who is unable to accurately observe the firm's value process. With this (realistic) extension of the model the problem of the low short credit spread is also resolved without having to resort to jump-diffusions.

The best known commercial implementation of the firm's value models is certainly the KMV (Kealhover, McQuown, Vasicek) model. Almost all literature on this model is based upon the information released by KMV Corporation who are (understandably) reluctant to release all details of the model, as they are trying to sell it. Nevertheless a good picture of the principal ideas behind the model can be found in the papers by Crosbie (an employee of KMV) in the books by Francis et.al. [1999] or Das [1998]. As mentioned before, the KMV model uses the firm's value approach to define the key indicator for credit quality, the distance to default, but leaves the model framework after that to imply default likelihoods from historical data.

Unfortunately there are hardly any empirical studies of the performance of the firm's value models in pricing and hedging of defaultable bonds. Crosbie analyses in his papers the performance of the KMV model and finds that it gives better results than using the firm's ratings for default prediction. Kwan [1996] empirically investigates the relationship between stock returns and yield changes of bonds by the same firm (controlling for default-free interest rate changes). He finds negative correlation between the two, both on a pooled and on the individual level, with a stronger connection for low ratings. Further empirical studies are by Jones et.al. [1984], and Sarig and Warga [1989]. In general the result seems to be that there is a connection between share prices and defaultable bond prices, but it only exists for lower credit ratings and it is not stable enough to implement a successful hedging strategy based on it.

Chapter 2

Intensity Models

2.1 Poisson Processes

2.1.1 In this section you will learn ...

- ... how to get an intuition for the behaviour of Poisson processes,
- ... some important properties of Poisson processes,
- ... how to handle Poisson-type jump processes in financial modelling,
- ... about compound Poisson processes.

2.1.2 Intuitive construction of a Poisson process

A poisson process N_t is an increasing process in the integers $0, 1, 2, 3, \ldots$. More important than its unexciting set of values are the *times of the jumps* $T_1, T_2, T_3 \ldots$ and the probability of a jump in the next instant.

We assume that the probability of a jump in the next small time interval Δt is proportional to Δt:

$$P[N_{t+\Delta t} - N_t = 1] = \lambda \Delta t, \tag{2.1}$$

and that jumps by more than 1 do not occur. This means conversely, that the probability that the process remains constant is

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

and over the interval $[t, 2\Delta t]$ this probability is

$$\begin{aligned} P[N_{t+2\Delta t} - N_t = 0] &= P[N_{t+\Delta t} - N_t = 0] \cdot P[N_{t+2\Delta t} - N_{t+\Delta t} = 0] \\ &= (1 - \lambda \Delta t)^2. \end{aligned}$$

Now we can start to construct a Poisson process: We subdivide the interval $[t, s]$ into i subintervals of length $\Delta t = (s - t)/i$. In each of these subintervals the process N has a jump with probability $\Delta t\lambda$. Mathematically we conduct i independent binomial experiments each with a probability of a 'jump' outcome of $\Delta t\lambda$.

The probability of *no* jump at all in $[t, s]$ is given by

$$P[N_s = N_t] = (1 - \Delta t\lambda)^i = (1 - \frac{1}{i}(s-t)\lambda)^i$$

Because $(1 + x/i)^i \to e^x$ this converges to

$$P[N_s = N_t] \to \exp\{-(s-t)\lambda\}.$$

Next we look at the probability of exactly *one* jump in $[t, s]$: There are n possibilities of having exactly one jump, giving a total probability of

$$\begin{aligned} P[N_s - N_t = 1] &= i \cdot \Delta t\lambda(1 - \Delta t\lambda)^{i-1} \\ &= i \cdot \frac{(s-t)}{i}\lambda(1 - \frac{1}{i}(s-t)\lambda)^i/(1 - \frac{1}{i}(s-t)\lambda) \\ &= \frac{(s-t)\lambda}{1 - \frac{1}{i}(s-t)\lambda}(1 - \frac{1}{i}(s-t)\lambda)^i \\ &\to (s-t)\lambda\exp\{-(s-t)\lambda\}, \end{aligned}$$

again using the limit result for the exponential function. The term in the denominator converges to 1 and drops therefore out in the limit.

Similarly one can reach the limit probabilities of *two* jumps

$$P[N_s - N_t = 2] = \frac{1}{2}(s-t)^2\lambda^2\exp\{-(s-t)\lambda\},$$

or n jumps

$$P[N_s - N_t = n] = \frac{1}{n!}(s-t)^n\lambda^n\exp\{-(s-t)\lambda\}. \tag{2.2}$$

Equation (2.2) is now used to formally define a Poisson process:

A Poisson process with intensity λ is a non-decreasing, integer-valued process with initial value $N_0 = 0$ whose increments satisfy equation (2.2).

This discrete-time approximation of the Poisson process resembles the binomial approximation of a Brownian Motion. In both constructions we take a number of binomially distributed random variables and add them up to get the process: In the Poisson case above we take the individual jumps in the time interval Δt, for the Brownian Motion we take the 'up and 'down movements at the individual nodes.

The difference lies in what we do in the limit: For the Brownian motion we *decrease the jump size* (we keep it proportional to $1/\sqrt{i}$) and *keep the probabilities constant*.
Here for the Poisson process we do the opposite: We *keep the jump size constant* (at one) and

decrease the probability (we keep it proportional to $1/i$).
These two different ways to adjust the parameters when letting $i \to \infty$ result in two completely different processes.

Equation (2.1) points to a good way to check the plausibility of the results we will get later on: the *large portfolio approximation*. Instead of assuming that we have just one defaultable security to price that is driven by just one Poisson process, just assume that *we have a large portfolio of defaultable securities that are all driven by independent Poisson processes*. Then we can assume that Poisson events happen almost continuously at a rate of λdt to our whole portfolio. This trick allows one to transform the unwieldy discrete jumps to a continuous rate of events.

2.1.3 Properties of Poisson processes

Poisson processes are usually used to model either rare events (as for example in insurance mathematics) or discretely countable events (e.g. radioactive decay, the number of customers in a queue, the number of calls through a telephone exchange...). Both properties, also apply to defaults: They are rare (hopefully) and they are discrete. Usually one models the time of default of a firm as *the time of the first jump of a Poisson process*.

The parameter λ in the construction of the Poisson process is called the *intensity* of the process. We already saw that the Poisson process is discontinuous and the distribution of the jump heights in equation (2.2). Here are some further properties:

- The Poisson process has no memory. The probability of n jumps in $[t, t+s]$ is independent of N_t and the history of N before t. Specifically a jump is not more likely just because the last jump occurred a long time ago and "it's about time.
- The inter-arrival times of a Poisson process $(T_{n+1} - T_n)$ are exponentially distributed with density
$$P[(T_{n+1} - T_n) \in t\,dt] = \lambda e^{-\lambda t} dt.$$
(This would be another way to define a Poisson process.)
- Two or more jumps at exactly the same time have probability zero.

For financial modelling we need some tools how to handle Poisson-type jump processes in stochastic differential equations. From the construction we have obviously

$$\mathrm{E}[\,dN\,] = \lambda dt. \tag{2.3}$$

Furthermore:

$$\begin{aligned} dNdN &= dN \\ \mathrm{E}\left[\,dN^2\,\right] &= \lambda dt \\ \mathrm{E}[\,dNdW\,] &= 0. \end{aligned}$$

We need a modified version of Itô's lemma that allows us to handle jumps in processes. Let $dx = dx^c + \Delta x$ be the decomposition of the process x in a continuous part x^c and a discontinuous part Δx (e.g. caused by jumps) and f a twice continuously differentiable function. Then Itô's lemma has the form

$$\begin{aligned} df(t,x) &= \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial x}dx^c + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}d < x^c > \\ &+ (f(t, x+\Delta x) - f(t,x)). \end{aligned} \tag{2.4}$$

We only have to add a jump term $\Delta f = f(t, x+\Delta x) - f(t,x)$ to the familiar form of Itô's lemma which is shown in the first line. There is no additional complication like in the continuous case with the second derivative terms. For multidimensional processes we have the same situation

$$\begin{aligned} df(t,x) &= \frac{\partial f}{\partial t}dt + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}dx_i^c + \frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i\, \partial x_j}d < x_i^c, x_j^c > \\ &+ (f(t, x+\Delta x) - f(t,x)). \end{aligned} \tag{2.5}$$

2.1.4 Inhomogenous Poisson Processes

If the intensity λ of the Poisson process is a function of time $\lambda(t)$ we reach an *inhomogenous* Poisson process. Its properties are very similar to the properties of a homogenous Poisson process.

Starting from the local jump probability

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

we can calculate the probability of *no* jump in the interval $[t, T]$:

$$\begin{aligned} P[N_T - N_t = 0] &= \prod_{i=1}^{n}(1 - \lambda(t + i\Delta t)\Delta t) \\ \ln P[N_T - N_t = 0] &= \sum_{i=1}^{n} \ln(1 - \lambda(t + i\Delta t)\Delta t) \\ &to \sum_{i=1}^{n} -\lambda(t + i\Delta t)\Delta t \\ &to - \int_t^T \lambda(s)ds \\ P[N_T - N_t = 0] &\to \exp\{-\int_t^T \lambda(s)ds\}. \end{aligned}$$

We used that $\ln(1-x) \to -x$ for $x \to 0$. The constant $\lambda(T-t)$ was replaced by the integral $\int_t^T \lambda(s)ds$. Similarly one can derive the probability of n jumps in $[t, T]$

$$P[N_T - N_t = n] = \frac{1}{n!}\left(\int_t^T \lambda(s)ds\right)^n \exp\left\{-\int_t^T \lambda(s)ds\right\}. \tag{2.6}$$

Again the only difference to (2.2) is that $\lambda(T-t)$ has been replaced by the integral of $\lambda(s)$ over the respective time span.

2.1.5 Compound Poisson Processes

A Poisson arrival is used to model the arrival of a specific event. But what happens at this event? So far we can only model *when* something happens, but we obviously also want to model *what* will happen.

For example in default-risk modelling we do not only want to model when the default occurs, but also how large this default will be, or in a model of jumping stock prices one would have to model the size of the jump, not only the time of the jump.

In a *compound Poisson process* at each time T_i of a jump of the Poisson process a random variable Y_i is drawn from a distribution $K(dY)$. Y_i is often called the *marker* to the point of jump T_i, the whole set $\{(T_i, Y_i)\}_{i\in\mathbb{N}}$ of points in time and markers is called a *marked point process*.

Let's assume we consider the cumulative sum of the Y_i:

$$X_t = \sum_{T_i \le t} Y_i.$$

A function $f(X)$ then has the following properties

$$\begin{aligned}
dX &= \Delta X = Y dN \\
df &= \Delta f = (f(X+Y) - f(X))dN \\
\mathrm{E}[\, dX \,] &= \int yK(dy)\lambda dt = y^e \lambda dt \\
\mathrm{E}[\, df(X) \,] &= \int (f(X+y) - f(X))K(dy)\lambda dt,
\end{aligned} \tag{2.7}$$

where y^e is the local expectation of Y. Note that $\mathrm{E}[\, df(X) \,] \neq f(\mathrm{E}[\, dX \,])$. Itô's lemma remains unchanged from (2.5), equation (2.5) is already Itô's lemma in its most general form.

2.1.6 Monte Carlo simulation

For a Monte Carlo simulation of a Poisson process equation (2.2) is better suited than the discrete approximation (2.1).

For the pricing of defaultable securities one is often only interested in the first jump of a Poisson process. The probability of at least one jump in Δt is exactly

$$P[N_{t+\Delta t} - N_t > 0] = 1 - \exp\{-\Delta t\lambda\}.$$

This probability is exact for all Δt. For the simulation it is usually sufficient to assume that all jumps are of size one (unless $\Delta t\lambda$ becomes large).

To Monte-Carlo simulate a Poisson process we draw a random number from the unit interval $[0, 1]$ and check whether it is lower than than $1 - \exp\{-\Delta t \lambda\}$. If it is lower, we assume a jump has occurred and increment the Poisson process N, otherwise it remains constant.

Especially for default-risk with Poisson-type jumps with low probabilities the Monte Carlo method is often very inaccurate and needs a very large number of runs. Defaults with low probability but significant losses may very well not occur in the simulation unless the number of runs is very large. On the other hand a default may occur in an early run and gain a weight that is not justified by its probability. A better way to price defaultable securities would be to try and eliminate the explicit reference to defaults before applying Monte Carlo, e.g. by using the large portfolio approximation mentioned above.

2.2 Pricing with Poisson Processes

2.2.1 In this section you will learn ...

- ...how to price defaultable securities when the default is triggered by a Poisson jump,
- ...how to incorporate positive recovery rates in the model,
- ...how to model and price fractional recovery and multiple defaults,
- ...how to model and price random recovery rates,
- ...to what extent hedging is still possible.

2.2.2 Zero recovery security pricing

In the easiest specification the defaultable bond has zero recovery and a default is triggered by the first jump of the Poisson process N with intensity λ.

To derive the pricing equation we have to look at the Itô-representation of the price process $\overline{B}$ of the defaultable bond

$$d\overline{B}(t,r,N) = \frac{\partial \overline{B}}{\partial t}dt + \frac{1}{2}\frac{\partial^2 \overline{B}}{\partial r^2}dt + \frac{\partial \overline{B}}{\partial r}dr - \overline{B}dN.$$

If there is a jump $dN = 1$ in the Poisson process this means that $d\overline{B} = -\overline{B}$ (the other terms are of lower order), the defaultable bond price jumps down to zero.

To derive the pricing equation we have to set $\mathbf{E}\left[\, d\overline{B} \,\right] = r\overline{B}dt$, i.e.

$$r\overline{B}dt = \frac{\partial \overline{B}}{\partial t}dt + \frac{1}{2}\frac{\partial^2 \overline{B}}{\partial r^2}dt + \frac{\partial \overline{B}}{\partial r}\mu_r dt - \overline{B}\lambda dt.$$

Here we have used $\mathbf{E}\left[\, dN \,\right] = \lambda dt$, and the pricing equation becomes

$$0 = \frac{\partial \overline{B}}{\partial t} + \frac{1}{2}\frac{\partial^2 \overline{B}}{\partial r^2} + \frac{\partial \overline{B}}{\partial r}\mu_r - \overline{B}(\lambda + r). \tag{2.8}$$

It is instructive to compare the pricing equation (2.8) for a defaultable bond with the pricing equation for a default free security:

$$0 = \frac{\partial B}{\partial t} + \frac{1}{2}\frac{\partial^2 B}{\partial r^2} + \frac{\partial B}{\partial r}\mu_r - Br.$$

The change by the presence of default risk is the addition of λ to the default-free interest rate in the final discounting term. Default risk here amounts to discounting the defaultable payments with a larger discount rate $\overline{r} = r + \lambda$. This holds for general defaultable securities provided the recovery in default is zero.

If $\overline{B}$ is a zero-coupon bond with maturity T, the solution to the pricing equation (2.8) is

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

where B is the price of a default-risk free bond. With this specification obviously $\frac{\partial \overline{B}}{\partial r} = \frac{\partial B}{\partial r}$ and $\frac{\partial^2 \overline{B}}{\partial r^2} = \frac{\partial^2 B}{\partial r^2}$ and $\frac{\partial \overline{B}}{\partial t} = \frac{\partial B}{\partial r} + \lambda \overline{B}$. Thus, if the default-free security B satisfies (2.8) without default ($\lambda = 0$) then $\overline{B}$ as given above satisfies (2.8) *with* default.

The solution derived above gives directly the yield spread of the defaultable bond over the default-free bond: It is

$$s(T,t) = \frac{1}{T-t}(\ln B - \ln \overline{B}) = \lambda,$$

the yield spread of a defaultable bond over the equivalent default-free bond is exactly the intensity λ of the default process N.

We did not use the fact that the defaultable security is a bond, the analysis above also carries through for other, more general payoff functions.

2.2.3 Pricing with positive recovery

To make the model more realistic we have to include positive recovery in case of default. One common assumption is that the defaultable bond pays off a constant amount of $(1-c)$ at default. c represents bankruptcy costs. Then $d\overline{B}$ is

$$d\overline{B}(t,r,N) = \frac{\partial \overline{B}}{\partial t}dt + \frac{1}{2}\frac{\partial^2 \overline{B}}{\partial r^2}dt + \frac{\partial \overline{B}}{\partial r}dr - \overline{B}dN + (1-c)dN.$$

In default we now lose the value of the bond $(-\overline{B}dN)$, but gain the recovery rate via $(1-c)dN$. To have $\mathrm{E}\left[\, d\overline{B} \,\right] = 0$ we reach the following modification to (2.8)

$$0 = \frac{\partial \overline{B}}{\partial t} + \frac{1}{2}\frac{\partial^2 \overline{B}}{\partial r^2} + \frac{\partial \overline{B}}{\partial r}\mu_r - \overline{B}(\lambda + r) + (1-c)\lambda. \tag{2.9}$$

For pricing purposes equation (2.9) is equivalent to the pricing equation we have to solve if $\overline{B}$ paid a coupon at the rate $(1-c)\lambda$ per time. Positive recovery of a constant amount thus boils down to having a coupon (or dividend) yield on the bond we are pricing. The solution is

$$\overline{B}(t,r) = B(t,r)e^{-\lambda(T-t)} + (1-c)\lambda B_c(t,r),$$

the defaultable bond price $\overline{B}$ is the price of a defaultable bond with zero recovery $Be^{-\lambda(T-t)}$ plus the price $B_c(r,t)$ of $(1-c)\lambda$ default-free *coupon* bonds paying a continuous coupon of $1dt$ per time.

The intuition behind this becomes obvious with the large portfolio approximation: In a large portfolio of defaultable bonds we would lose a proportion of λdt bonds to default in each time step. Therefore the additional discounting with $e^{-\lambda(T-t)}$: we have exponential decay in our portfolio, exactly as in radioactive decay. The defaults lead to an immediate stream of recovery payments of $(1-c)$ times the number of defaults: $(1-c)\lambda dt$. These payments are also discounted by r and λ therefore the decay in these payments is included.

The pricing equation (2.9) would not be changed if the payoff in default $(1-c)$ was a function of r and t itself, for example we could say that in default the defaultable bond $\overline{B}$ pays off a fraction d of the equivalent default-free bond. Then $(1-c) = dB$ has to be substituted in (2.9).

2.2.4 Multiple defaults and fractional recovery

The dividend or coupon stream that was introduced in the previous section has some disadvantages: Firstly it can complicate the pricing if one previously had an European-style final payoff. Secondly it is often unrealistic to assume a cash payoff in default.

Usually when a bond defaults, a reorganisation takes place and the bondholders lose a fraction q of the face value of their claims, but the claims continue to live and the issuer continues to operate. After the first default and reorganisation subsequent defaults are not impossible, we can have multiple defaults.

For pricing purposes we look again at the dynamics of the defaultable bond price:

$$d\overline{B}(t,r) = \frac{\partial \overline{B}}{\partial t}dt + \frac{1}{2}\frac{\partial^2 \overline{B}}{\partial r^2}dt + \frac{\partial \overline{B}}{\partial r}dr - q\overline{B}dN.$$

We made the additional assumption that the bond price is independent of the number N of defaults that have already occurred (provided we have already accounted for the loss in face value). The only influence of a default is the direct loss in face value. Setting $\mathrm{E}\left[\, d\overline{B}\,\right] = r\overline{B}dt$ we reach

$$0 = \frac{\partial \overline{B}}{\partial t} + \frac{1}{2}\frac{\partial^2 \overline{B}}{\partial r^2} + \frac{\partial \overline{B}}{\partial r}\mu_r - \overline{B}(q\lambda + r). \tag{2.10}$$

Equation (2.10) is exactly the pricing equation for zero recovery (2.8) but with $q\lambda$ instead of λ as additional discount factor. Thus equation (2.10) represents a more realistic model with a simpler pricing equation, unless there are good reasons to do otherwise, it should be preferred.

The solution of (2.10) is

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

which is closely related to the no-recovery case, and the yield spread is

$$s(t,T) = q\lambda. \tag{2.11}$$

Now the spread is the product of the intensity of default and the loss quota.

As the model with multiple defaults and fractional recovery is both more realistic and analytically easier to handle, we will use this model from now on.

2.2.5 Pricing using probability theory

The probability theory approach is often used in the literature and sometimes yields closed form results very quickly. First we price a defaultable security with zero recovery and a payoff of X at T. The promised payoff X is random and independent of the default event. Then today's price is given by the expectation of the discounted payoff under the risk-neutral probabilities:

$$F' = \mathbf{E}\left[\mathbf{1}_{\{\tau > T\}} \exp\{-\int_0^T r_s ds\} X \right].$$

Because of the independence of X and the default we can factor out the expectation and get

$$\begin{aligned} F' &= \mathbf{E}\left[\mathbf{1}_{\{\tau > T\}} \right] \mathbf{E}\left[\exp\{-\int_0^T r_s ds\} X \right] \\ &= P[N_T - N_0 = 0] F \\ &= e^{-\lambda(T-0)} F. \end{aligned}$$

The price F' of the defaultable derivative is the price F of the equivalent default-free derivative times the probability of survival.

In the case of multiple defaults the situation becomes slightly more complicated. If there are n defaults each with a loss quota of q the final payoff will be

$$(1-q)^n X.$$

At each default the payoff is reduced to $(1-q)$ times its previous value. Then

$$\begin{aligned} F' &= \mathbf{E}\left[(1-q)^{N_T} \exp\{-\int_0^T r_s ds\} X \right] \\ &= \mathbf{E}\left[(1-q)^{N_T} \right] F, \end{aligned}$$

and

$$\begin{aligned} \mathbf{E}\left[(1-q)^{N_T} \right] &= \sum_{n=0}^{\infty} (1-q)^n P[N_T = n] \\ &= \sum_{n=0}^{\infty} (1-q)^n \frac{1}{n!} \lambda^n T^n e^{-\lambda T} \\ &= e^{-\lambda T} \sum_{n=0}^{\infty} \frac{1}{n!} [(1-q)\lambda T]^n \\ &= e^{-\lambda T} e^{(1-q)\lambda T} = e^{-q\lambda T}. \end{aligned}$$

We used that

$$e^x = \sum_{n=0}^{\infty} \frac{1}{n!} x^n.$$

All calculations in this section have been derived under the assumption of independence of the default process N and the promised payoff X or the risk-free term interest rates. We will see later on how far these results can be generalized.

2.2.6 Random recovery and seniority

The next step in making the model more realistic is to allow for a random recovery rate $1 - c$ which is drawn at default from a certain distribution. We now have to deal with a compound Poisson process where the marker is the recovery rate. The derivation of (2.9) remains almost unchanged, only

$$\mathrm{E}[\,(1-c)dN\,] := (1-c^e)\lambda dt$$

has to be substituted, yielding a pricing equation with c replaced by its local expectation c^e. The same happens in equation (2.10): Here q has to be replaced with its expected value q^e.

With random recovery we can implement seniority between different bonds: At each default T_i a random variable Y_i is drawn. The payoffs of the different securities now depend on Y, a junior bond would have bankruptcy costs (losses) of $c^j(Y)$ and a senior bond $c^s(Y)$ with $c^s(Y) < c^j(Y)$. Similarly in the multiple default model a senior defaultable bond would have a loss quota $q^s(Y) < q^j(Y)$ which is smaller than the writedown on a junior bond.

If the recovery rates or the bankruptcy costs are deterministic, seniority can be implemented by choosing suitable numbers $c^j > c^s$ for the bankruptcy costs or $q^j > q^s$ for the writedown quotas.

2.2.7 Hedging

As opposed to the increments of the Brownian motions we had to hedge before we now have to deal with jumps of a Poisson process. These are large and unpredictable, they come without warning.

In the firm's value models the default came not as a complete surprise, we could observe the firm's value approaching the knockout barrier and could hedge against all movements. A Poisson jump on the other hand can occur at any time and must therefore always be hedged.

Let us consider a portfolio Π consisting of one defaultable bond $\overline{B}_1$ (the security to hedge) and α_1 default-free bonds B and α_2 of another defaultable bond $\overline{B}_2$, the hedge instruments. Then

$$\begin{aligned} d\Pi &= d\overline{B}_1^c + \alpha_1 dB + \alpha_2 d\overline{B}_2^c \\ &\quad + (\Delta\overline{B}_1 + \alpha_2\Delta\overline{B}_2)dN. \end{aligned}$$

The continuous changes in the price $d\overline{B}_1^c$ can be hedged as usual, we have to set the weighted sum of the first derivatives with respect to the continuous factors (here only the risk-free short rate r) to zero:

$$0 = \frac{\partial\overline{B}_1^c}{\partial r} + \alpha_1\frac{\partial B}{\partial r} + \alpha_2\frac{\partial\overline{B}_2^c}{\partial r}.$$

To hedge the jumps is more problematic now. Ideally we would like to have

$$\begin{aligned} 0 &= \Delta\overline{B}_1 + \alpha_2\Delta\overline{B}_2 \\ 0 &= (1-c) - \overline{B}_1 + \alpha_2((1-c) - \overline{B}_2). \end{aligned}$$

If $(1-c)$ is known *before* the jump dN then the above identity can be achieved by choosing α_2 in a suitable fashion. If c is a (continuously distributed) random variable like in the compound Poisson process example before, then there is no possibility to achieve a hedge by choosing the weights α of a finite number of hedge instruments. The best we can hope to do is to try and minimize some remaining risk, i.e. with the mean-square error as criterion

$$\int((1-c)-\overline{B}_1+\alpha_2((1-c)-\overline{B}_2))^2K(dc)\to\min_{\alpha_2}$$

would at least minimize the remaining risk. Alternatively we could try and minimize the local variance of the portfolio

$$\min_{\alpha}\mathrm{E}\left[\,d\Pi^2\,\right]$$

which would mean to minimize

$$\begin{aligned}\min_{\alpha}\sigma_r^2&\left(\frac{\partial\overline{B}_1}{\partial r}+\alpha_1\frac{\partial B}{\partial r}+\alpha_2\frac{\partial\overline{B}_2}{\partial r}\right)^2\\&+\lambda\mathrm{E}\left[\,\left(\Delta\overline{B}_1+\alpha_2\Delta\overline{B}_2\right)^2\,\right].\end{aligned}\tag{2.12}$$

[1]

In the fractional recovery / multiple default case we also have to ensure that

$$\begin{aligned}0&=\Delta\overline{B}_1+\alpha_2\Delta\overline{B}_2\\0&=-q_1\overline{B}_1-q_2\alpha_2\overline{B}_2.\end{aligned}$$

If $\overline{B}_1$ and $\overline{B}_2$ have different seniority then $q_1\neq q_2$ and a hedge ratio α_2 can be found. If both bonds have the same seniority, i.e. $q_1=q_2=q$ then the only choice of α_2 for a perfect hedge is to choose it such that the total value of defaultable instruments in the portfolio is zero. This is obviously not always a sensible solution.

[1]This is the idea underlying the local risk minimisation by Föllmer, Sondermann [1985].

2.3 Stochastic Intensity

2.3.1 In this section you will learn ...

- ... the properties of Poisson processes with stochastic intensities,
- ... how to derive the pricing equation for defaultable securities in the Cox-process framework,
- ... how to apply your knowledge from risk-free interest rate models to default risk pricing,
- ... the pricing equations for derivatives on the credit spread of defaultable bonds,
- ... the steps to building a credit risk model.

2.3.2 Cox Processes

In (2.11) we saw that the spread of a defaultable zero coupon bond is exactly $q\lambda$, a constant. Constant credit spreads are obviously not what we observe in the market, we need a more flexible model to be able to model *dynamics* of defaultable bond prices and credit spreads. This can be achieved by generalizing the Poisson process N to a *Cox processes.*

Roughly, Cox processes are Poisson processes with stochastic intensity. Going back to the discrete-time approximation of the Poisson process in section 2.1.2 we now introduce stochastic dynamics for λ:

$$d\lambda = \mu_\lambda dt + \sigma_\lambda dW_2. \tag{2.13}$$

Again we conduct i Binomial experiments, one for each time interval Δt, but now the probability of a jump in $[t, t + \Delta t]$ is

$$P[N_{t+\Delta t} - N_t = 1] = \lambda_t \Delta t,$$

proportional to the value of λ *at time t.*

The jumps are now correlated via the path taken by λ: If there was a jump already that means that λ is probably large which in turn means that a jump now is more likely, too. But the correlation is only indirect. There is no connection like 'the probability of the next jump is proportional to the number of jumps so far, which would constitute a direct connection.

If we knew the realization of λ in advance we would have an inhomogenous Poisson process like in section 2.1.4. This fact can be used to formally define a Cox process:

A *Cox process with intensity process* λ is an integer-valued non-decreasing process N with the property that, conditional on the realisation $(\lambda_t)_{t\in[0,T]}$ of λ, N is an inhomogenous Poisson process with intensity λ_t.

We can use this definition to directly calculate some properties of Cox processes. From the Poisson-type construction we still have

$$\mathrm{E}_t [\, dN \,] = \lambda_t dt,$$

and the other local properties of N also carry through. New are the jump probabilities. Given the realisation of λ we know by (2.6) that the probability of n jumps for an inhomogenous Poisson process is

$$P[N_T - N_t = n] = \frac{1}{n!}\left(\int_t^T \lambda(s)ds\right)^n \exp\left\{-\int_t^T \lambda(s)ds\right\}.$$

For a Cox process we have to take the expectation over the path λ will take:

$$\begin{aligned} P[N_T - N_t = n] &= \mathrm{E}[\, P[N_T - N_t = n|\lambda]\,] \\ &= \mathrm{E}\left[\frac{1}{n!}\left(\int_t^T \lambda(s)ds\right)^n \exp\left\{-\int_t^T \lambda(s)ds\right\}\right]. \end{aligned}$$

We know how to handle the pricing problems given the path of λ. So all we have to do is: First, solve the problem for a given λ, second take expectations over all λ. In the first step one is usually able to eliminate all direct reference to the jump process N and replace them with terms in λ and the other state variables. All these state variables follow continuous paths so we are back in the known fields of continuous-time finance.

2.3.3 The pricing equation

We have defined the following stochastic processes:

$$dr = \mu_r dt + \sigma_r dW_1$$

for the risk-free interest rate, and

$$d\lambda = \mu_\lambda dt + \sigma_\lambda(\rho dW_1 + \sqrt{1-\rho^2}\, dW_2)$$

for the intensity of the Cox process. We have included the possibility of correlation-ρ between λ and r, and all dynamics are already risk-neutral.

Given the additional state variable λ the price of a defaultable security will depend on it: $\overline{B}(t, r, \lambda)$. To derive the partial differential equation satisfied by $\overline{B}$ we proceed as usual. First we calculate $d\overline{B}$ by Itô's lemma:

$$\begin{aligned} d\overline{B} = &\frac{\partial \overline{B}}{\partial t}dt + \frac{\partial \overline{B}}{\partial r}dr + \frac{1}{2}\sigma_r^2\frac{\partial^2 \overline{B}}{\partial r^2}dt \\ &+ \frac{\partial \overline{B}}{\partial \lambda}d\lambda + \frac{1}{2}\sigma_\lambda^2\frac{\partial^2 \overline{B}}{\partial \lambda^2}dt \\ &+ \rho\sigma_r\sigma_\lambda\frac{\partial^2 \overline{B}}{\partial \lambda\, \partial r}dt - q\overline{B}dN. \end{aligned}$$

We used the multiple default model with fractional losses. Next we set $\mathrm{E}\left[\, d\overline{B}\,\right] = r\overline{B}dt$ which yields the pricing equation

$$0 = \frac{\partial \overline{B}}{\partial t} + \mu_r\frac{\partial \overline{B}}{\partial r} + \mu_\lambda\frac{\partial \overline{B}}{\partial \lambda}$$

$$+\frac{1}{2}\sigma_r^2\frac{\partial^2\overline{B}}{\partial r^2}+\rho\sigma_r\sigma_\lambda\frac{\partial^2\overline{B}}{\partial\lambda\,\partial r}+\frac{1}{2}\sigma_\lambda^2\frac{\partial^2\overline{B}}{\partial\lambda^2}$$
$$-(r+\lambda q)\overline{B}. \tag{2.14}$$

Again note the term in the last line where the risk-free interest rate r has been replaced with $r+q\lambda$. The final condition for a defaultable zero coupon bond with maturity T is

$$\overline{B}(T,r,\lambda)=1,$$

and the boundary conditions are $\overline{B}\to 0$ as $r,\lambda\to\infty$, and $\overline{B}<\infty$ as $r,\lambda\to 0$. The solution of the equation (2.14) then depends on the exact specification of the stochastic processes for r and λ, but some results can be reached with a more general approach in the next section.

2.3.4 Defaultable bonds and yield spreads

To see the general properties of the prices of defaultable bonds in this modelling framework we go back to a very general setup. We just say that r and λ follow some (continuous) stochastic processes and that N is a Cox process with intensity λ. The general representation of the price of a default-free zero coupon bond B with maturity T is

$$B(t)=\mathrm{E}_t\left[\exp\left\{-\int_t^T r(s)ds\right\}\right]. \tag{2.15}$$

This representation has to hold in every arbitrage-free model[2]. Similarly, any security with (random) final payoff X has the price process

$$F(t)=\mathrm{E}_t\left[\exp\left\{-\int_t^T r(s)ds\right\}X\right]. \tag{2.16}$$

The price is the expectation of the discounted value of the final payoff.

The situation for defaultable bonds is not different: Consider a defaultable zero coupon bond:

$$\overline{B}(t)=\mathrm{E}_t\left[\exp\left\{-\int_t^T r(s)ds\right\}(1-q)^{N_T}\right], \tag{2.17}$$

its price process is given by the discounted value of the final payoff. Now we can use the Cox process properties of N and expand $\mathrm{E}_t[\,\cdot\,]$ into

$$\mathrm{E}_t[\,\cdot\,]=\mathrm{E}_t[\,\mathrm{E}_t[\,\cdot\,|(\lambda(s))_{s\leq T}\,]\,],$$

because the expectation of a conditional expectation is again the expectation itself. We do not have the information about the path $(\lambda(s))_{s\leq T}$ of λ, but we first calculate what we would get *if* we knew it (the inner, conditional expectation); we get a different result for every possible path

[2] From now on all expectations are taken with respect to the risk-neutral probabilities unless otherwise stated.

$(\lambda(s))_{s\leq T}$. Then we take the expectation over these results (the outer expectation). Thus we are looking at

$$\overline{B}(t) = \mathrm{E}_t\left[\mathrm{E}_t\left[\exp\left\{-\int_t^T r(s)ds\right\}(1-q)^{N_T}|(\lambda(s))_{s\leq T}\right]\right]. \tag{2.18}$$

We already have calculated the inner expectation for constant intensity in section 2.2.5. For the inhomogenous Poisson processes here the result is similar:

$$\begin{aligned}
&\mathrm{E}_t\left[\exp\left\{-\int_t^T r(s)ds\right\}(1-q)^{N_T}|(\lambda(s))_{s\leq T}\right] \\
&= \mathrm{E}_t\left[\exp\left\{-\int_t^T r(s)ds\right\}\exp\left\{-\int_t^T q\lambda(s)ds\right\}|(\lambda(s))_{s\leq T}\right] \\
&= \mathrm{E}_t\left[\exp\left\{-\int_t^T r(s)+q\lambda(s)ds\right\}|(\lambda(s))_{s\leq T}\right].
\end{aligned}$$

Combining this with the outer expectation we get

$$\overline{B}(t) = \mathrm{E}_t\left[\exp\left\{-\int_t^T r(s)+q\lambda(s)ds\right\}\right]. \tag{2.19}$$

The price of a defaultable bond of maturity T is the price a default-free bond of the same maturity would have in a world where the risk-free short rate is

$$\overline{r} = r + q\lambda. \tag{2.20}$$

A similar result holds for the defaultable security F'. Its final payoff is $(1-q)^{N_T}X$ and the price has the representation

$$\begin{aligned}
F'(t) &= \mathrm{E}_t\left[\exp\left\{-\int_t^T r(s)ds\right\}(1-q)^{N_T}X\right], \\
&= \mathrm{E}_t\left[\exp\left\{-\int_t^T r(s)+q\lambda(s)ds\right\}X\right].
\end{aligned} \tag{2.21}$$

All default-affected payoffs can be priced by discounting them with the short rate process $\overline{r} = r + q\lambda$.

The result of equations (2.20) and (2.19) mean that we can directly transfer the whole machinery of risk-free interest rate modelling to defaultable bond pricing, for pricing purposes the analogy between default-free interest rate models and defaultable bond models is complete.

2.3.5 Intensity models in a Heath, Jarrow, Morton framework

This section is a summary of the results of the next chapter (chapter 3).

Equation (2.21) shows clearly the analogy between derivatives pricing in a default-free stochastic interest rate environment and in the multiple default stochastic intensity model. If we define the defaultable short rate $\bar{r} = r + q\lambda$ then the pricing equation (2.21) is indistinguishable from the equivalent pricing equation in an environment with $\bar{r}$ as interest rate.

This analogy suggests to use one of the standard short rate models to model the dynamics of $q\lambda$. The full set of risk-free interest rate models can be applied, and the dynamics of $q\lambda$ can be fitted to a given term structure of defaultable bond prices.

This also carries through to a Heath- Jarrow- Morton [1992] framework of forward rates.

First, we need to define the defaultable and the default-free forward rates:

- The *instantaneous risk-free forward rate* at time t for date $T > t$ is defined as

$$f(t,T) = -\frac{\partial}{\partial T} \ln B(t,T). \tag{2.22}$$

- The *instantaneous defaultable forward rate* at time t for date $T > t$ is defined as

$$g(t,T) = -\frac{\partial}{\partial T} \ln \overline{B}(t,T). \tag{2.23}$$

- The dynamics of the risky forward rates are given by

$$dg(t,T) = \alpha'(t,T)\,dt + \sum_{i=1}^{n} \sigma_i'(t,T)\,dW^i(t). \tag{2.24}$$

- The dynamics of the default risk free forward rates are given by

$$df(t,T) = \alpha(t,T)\,dt + \sum_{i=1}^{n} \sigma_i(t,T)\,dW^i(t). \tag{2.25}$$

Given the dynamics of the defaultable forward rates and the default-free forward rates we can derive the resulting dynamics of the zero coupon bonds. For the defaultable case:

$$\begin{aligned} \frac{d\overline{B}(t,T)}{\overline{B}(t-,T)} &= \left[-\hat{b}(t,T) + \bar{r}(t) + \frac{1}{2}\sum_{i=1}^{n} a_i^2(t,T)\right] dt \\ &\quad + \sum_{i=1}^{n} a_i(t,T)\,dW^i(t) - q\,dN(t). \end{aligned} \tag{2.26}$$

And for the default-free case:

$$\begin{aligned} \frac{dB(t,T)}{B(t-,T)} &= \left[-\hat{b}(t,T) + r(t) + \frac{1}{2}\sum_{i=1}^{n} a_i^2(t,T)\right] dt \\ &\quad + \sum_{i=1}^{n} a_i(t,T)\,dW^i(t) - q\,dN(t). \end{aligned} \tag{2.27}$$

where the processes a_i, a'_i, $\hat{b}$ and $\hat{b}'$ are given by

$$a'_i(t,T) := \int_t^T \sigma'_i(t,v)\, dv \tag{2.28}$$

$$\hat{b}'(t,T) := \int_t^T \alpha'(t,v)\, dv. \tag{2.29}$$

$$a_i(t,T) := \int_t^T \sigma_i(t,v)\, dv \tag{2.30}$$

$$\hat{b}(t,T) := \int_t^T \alpha(t,v)\, dv. \tag{2.31}$$

Now we can take the expectation of $d\overline{B}$ and dB and equate it to $\overline{B}rdt$ and $Brdt$ respectively. This yields the following conditions for absence of arbitrage under the risk-neutral probabilities:

(i) The short interest rate spread is q times the intensity of the default process. It is positive.

$$q\lambda(t) = \overline{r}(t) - r(t) > 0. \tag{2.32}$$

(ii) The drift coefficients of the defaultable forward rates satisfy

$$\alpha'(t,T) = \sum_{i=1}^{n} \sigma'_i(t,T) \int_t^T \sigma'_i(t,v)dv. \tag{2.33}$$

(iii) The drift coefficients of the risk-free forward rates satisfy

$$\alpha(t,T) = \sum_{i=1}^{n} \sigma_i(t,T) \int_t^T \sigma_i(t,v)dv. \tag{2.34}$$

The three conditions (i), (ii) and (iii) are all we have to observe when we are building a HJM model of defaultable short rates. The most important result of this section is equation (2.33), the defaultable-bond equivalent of the well-known Heath- Jarrow- Morton drift-restriction. This restriction has been derived for default risk-free bonds in HJM [1992], and, as we see here, it is also an important part of the modelling of the defaultable bonds' dynamics.

A similar restriction can also be derived for the spreads between the forward rates: The *forward rate spread* $h(t,T)$ is defined as the difference between the defaultable forward rate and the default-free forward rate:

$$h(t,T) = g(t,T) - f(t,T). \tag{2.35}$$

Under the risk-neutral probabilities we have

$$q\lambda(t) = h(t,t).$$

Now one has to find a model for $h(t,T)$ which is compatible with conditions (i), (ii) and (iii). The advantage of modelling $h(t,T)$ instead of $g(t,T)$ is that (2.32) reduces to the well known problem of ensuring that $h(t,t) > 0$. and we can hope to use some of the extensive literature on interest rate models with positive short rates. We use the following dynamics for h:

$$h(t,T) - h(0,T) = \int_0^t \alpha^h(v,T)\,dv + \sum_{i=1}^{n} \sigma_i^h(v,T)dW^i(v). \tag{2.36}$$

In place of the drift restriction (2.33) we reach:

$$\begin{aligned}\alpha^h(t,T) = \sum_{i=1}^{n} \Bigg[& \sigma_i(t,T) \int_t^T \sigma_i^h(t,v)\,dv \\ & + \sigma_i^h(t,T) \int_t^T \sigma_i(t,v)\,dv \\ & + \sigma_i^h(t,T) \int_t^T \sigma_i^h(t,v)\,dv \Bigg].\end{aligned} \tag{2.37}$$

For independence between spread and risk-free interest rate dynamics this simplifies to

$$\alpha^h(t,T) = \sum_{i=1}^{n} \sigma_i^h(t,T) \int_t^T \sigma_i^h(t,v)dv. \tag{2.38}$$

Satisfying the positivity requirement (2.32) on $h(t,t)$ becomes very easy in the setup of credit spread modelling: One can use any interest rate model for $h(t,T)$ that is known to generate positive short rates, e.g. the square root model of Cox, Ingersoll and Ross [1985, 1985] or the model with lognormal interest rates by Sandmann and Sondermann [1994].

2.3.6 Model building strategy

The results from the previous subsections suggest the following proceeding when building a model for credit risk pricing:

- First of all we need a default-free interest rate model for the default free bond prices.
- Then a process for $q\lambda$ has to be chosen. It is advisable to choose it such that with $\bar{r} = r + q\lambda$ and equation (2.19) there are closed form solutions for the defaultable bond prices. Note that we just need the joint process $q\lambda$ and not q and λ individually which may be difficult to separate.
- After that, this model is fitted to the market. If there are several liquid defaultable bonds by the same issuer this market data can be used to calibrate the model.
- Finally the model can be used to price either further bonds issued by the same issuer, or to price derivatives on the defaultable bonds, or to price credit risk derivatives.

2.4 Intensity Models: Guide to the Literature

In the intensity models the time of default is modeled directly as the time of the first jump of a Poisson process with random intensity (a Cox process), or – more generally – as a totally inaccessible stopping time with an intensity. In this group of models a striking similarity to default-free interest rate modelling is found.

The first models of this type were developed by Jarrow and Turnbull [1995], Madan and Unal [1998] and Duffie and Singleton [1997]. Jarrow and Turnbull consider the simplest case where the default is driven by a Poisson process with constant intensity with known payoff at default. This is changed in the Madan and Unal model where the intensity of the default is driven by an underlying stochastic process that is interpreted as firm's value process, and the payoff in default is a random variable drawn at default, it is not predictable before default. Madan and Unal estimate the parameters of their process using rates for certificates of deposit in the Savings and Loan Industry.

Duffie and Singleton [1997] developed a similar model where the payoff in default is also cash, but denoted as a fraction $(1 - q)$ of the value of the defaultable security just before default. This model was applied to a variety of problems including swap credit risk, estimation, and two-sided credit risk, by a group around Duffie (Duffie and Singleton [1997], Duffie, Schroder and Skiadas [1994], Duffie and Huang [1994] and Duffie [1994]).

Lando [1998] developed the Cox-process methodology with the iterated conditional expectations which will be used in the section on the pricing of credit derivatives later on. His model has a default payoff in terms a certain number of default-free bonds and he applies his results to a Markov chain model of credit ratings transitions.

In the Schönbucher [1996, 1997] model multiple defaults can occur and instead of liquidation with cash payoffs a restructuring with random recovery rate takes place. The model is set in a Heath- Jarrow- Morton framework and a rich variety of credit spread dynamics is allowed. For many pricing purposes the model can be reduced to a similar form of Duffie and Singleton, and in Schönbucher [1997] it is applied to the pricing of several credit risk derivatives.

There is a variety of other models that fall into the class of intensity based models, we only mention Flesaker et.al. [1994], Artzner and Delbaen [1992, 1994] and Jarrow and Turnbull [1997]. On the empirical side the papers by Duffee [1995], Duffie and Singleton [1997] and Düllmann et.al. [1999] have to be mentioned. In these papers the authors estimate the parameters for the stochastic process of the credit spread for the Duffie-Singleton model.

The intensity models have also been implemented in a commercial software package. The model is called *Credit Risk+* and it was developed by Credit Suisse Financial Products as a tool for the portfolio management of credit risk. In this model a default is triggered by the jump of a Poisson process whose intensity is randomly drawn for each debtor class.

Chapter 3

An Intensity Model in Detail: Extending HJM to Default Risk

3.1 In this section you will learn ...

a detailed analysis of an intensity based default risk model: The defaultable Heath-Jarrow-Morton model. In this chapter some important techniques for the advanced analysis of credit risk models are introduced, amongst them

- the analysis of marked point processes,
- the change of measure technique for marked point processes,
- defaultable bond pricing for the multiple defaults model
- a model which combines rating transitions, stochastic credit spreads within rating classes which can be fully fitted to a given credit spread structure for each rating class.

Because the contents of this chapter are of an advanced nature, it can be skipped at first reading.

For the readers who have remained with us or who have returned here on a second reading we now give a more detailed account of the contents of this chapter. In this chapter[1] we present a model of the development of the term structure of defaultable interest rates that is based on a multiple-defaults model. Instead of modelling a cash payoff in default we assume that defaulted debt is restructured and continues to be traded.
We use the Heath-Jarrow-Morton (HJM) [1992] approach to represent the term structure of defaultable bond prices in terms of forward rates. The focus of the chapter lies on the modelling the development of this term structure of defaultable bond prices and we give conditions under which these dynamics are arbitrage-free. These conditions are a drift restriction that is closely related to the HJM drift restriction for default-free bonds, and the restriction that the defaultable short rate must always be not below the default-free short rate.

[1] This first parts of this chapter are based upon Schönbucher [1998].

Similar restrictions are derived for two extensions of the model setup, the first one is in a marked point process framework and allows for jumps in the defaultable forward rates at times of default, and the second one is a general ratings transition framework which can incorporate stochastic dynamics for the credit spreads in all ratings classes and also stochastic transition intensities.

3.2 Introduction

Most bankruptcy codes provide several alternative procedures to deal with defaulted debt and the debtors. The most obvious option is to liquidate the debtor's remaining assets and distribute the proceeds amongst the creditors, but a often more popular alternative is to reorganize the defaulted issuer and keep the issuer in operation. The latter alternative has the advantage of preserving the value of the debtor's business as a going concern and it avoids inefficient liquidation sales. Frequently there is no alternative to reorganisation, either because a liquidation of an issuer is impossible (e.g. for sovereign debtors) or because it is undesirable (if a liquidation would have a large macroeconomic effect).

In their empirical study Franks and Torous [1994] found the following:
A default of a bond does not mean that this bond becomes worthless, usually there is a positive recovery rate between 40 and 80 percent. This recovery rate varies significantly between firms. The majority of firms in financial distress are reorganized and re-floated, they are not liquidated. On average, most of the compensation payments (about two thirds) are in terms of securities of the reorganized firm, not in cash.
If a firm is reorganized, and the debtors are paid in terms of newly issued debt, then a second default of this firm on its (newly issued) debt is possible. In principle there could be a sequence of any number of defaults each with a subsequent restructuring of the defaulted firm's debt.

In this chapter we present a model of defaultable bond prices in which a defaulted issuer is not liquidated but reorganized at default. Multiple defaults can occur and the magnitude of the losses in default is not predictable.

We use the Heath-Jarrow-Morton (HJM) [1992] framework to represent the term structure of defaultable bond prices in terms of forward rates and give conditions under which these dynamics are arbitrage-free. These conditions are a drift restriction that is closely related to the HJM drift restriction for default-free bonds, and the restriction that the defaultable short rate must always be not below the default-free short rate.

The model in this chapter is based on the intensity-approach, an approach in which the time of default is a totally inaccessible stopping time which has an intensity process. Amongst others this approach is followed by Artzner and Delbaen [1992, 1994], Jarrow and Turnbull [1995], Lando [1994, 1998], Jarrow, Lando and Turnbull [1997], Madan and Unal [1998], Flesaker et.al. [1994], Duffie and Singleton [1997, 1999], Duffie, Schroder and Skiadas [1994], Duffie and Huang [1996], and Duffie [1994]. In all these models a cash payoff (or a payoff in default-free bonds) is specified in default. Therefore only one default is allowed, and after default the firm that had issued the debt is liquidated. This excludes reorganisation of defaulted debt as well as multiple defaults. Usually (except Madan and Unal [1998]) the magnitude of the payoff in default is predictable, too. In this model there is the possibility of multiple defaults and the magnitude of the recovery need not be predictable. It should furthermore be pointed out that many of the results (e.g. the HJM drift restriction on the defaultable bond prices or the relationship between the short credit spread and the default intensity) are not restricted to the intensity-based framework, but remain valid also for other default models.

We start with a model in which the value of a defaultable bond drops to zero upon default. While this case has already been extensively studied in the literature it is a good introduction to more general models. After deriving the interconnections between the dynamics of the defaultable interest rates and the defaultable bond prices, we derive the key relationship that *under the martingale measure the difference between the defaultable short rate and the default-free short rate is the intensity of the default process* with an argument using the savings accounts. This result drives the conditions for the absence of arbitrage that are derived subsequently, and the arbitrage-free dynamics of the defaultable bond prices. We find a very strong similarity between the defaultable and the default free interest rate dynamics and drift restrictions, as both have to satisfy the HJM drift restrictions. (Similar results have been shown by Duffie in [1994] and Lando [1998].)

Next we explore how a model of the spread of the defaultable forward rates over the default-free forward rates may be used to add a default-risk module to an existing model of default-free interest rates. Surprisingly, for forward rates this spread can be negative although there has to be a positive spread for the short rates. An example is given demonstrating that this is only possible under strong correlation between spreads and default-free interest rates.

In the following sections we propose a model that includes positive recovery rates, reorganisations of the defaulted firms with the possibility of multiple defaults and uncertainty about the magnitude of the default. Even though it may seem that this will make the model far more complicated the restrictions for absence of arbitrage and the price dynamics remain unchanged. This model is closely related to the fractional recovery model proposed by Duffie and Singleton [1994, 1997, 1999].

Instead of modelling the defaultable forward rates, previous intensity models concentrated on modelling the default intensity (i.e. the short credit spread) directly. These two are connected and the link is shown in the next section. We show that the results of the classical intensity models can be recovered from the defaultable forward model directly. Particularly the representation of the defaultable bond prices as expectation of a defaultable discount factor follows much more easily than in most intensity models. It is also shown that the only restriction for absence of arbitrage is to ensure a positive spread between defaultable and default-free short rate.

In the following section the defaultable forward rates (and thus the bond prices) are allowed to change discontinously at default times if there are multiple defaults. Here we use the methods of Björk, Kabanov and Runggaldier (BKR) [1996] and give the drift- and no-arbitrage conditions for this more general version of the model.

The chapter is concluded with another extension of the model which incorporates ratings transitions. In this section we model the defaultable forward rates for all rating classes simultaneously and analyse the conditions that ensure absence of arbitrage for a given ratings transition intensity matrix and a given volatility structure of the credit spreads within each ratings class. Again, these conditions turn out to be closely related to the classical HJM conditions which makes the model accessible to a numerical implementation. Instead of interpreting the model as a ratings transition model it can also be viewed as a model including a *credit crisis* with stochastic transition from the 'normal' state to the 'crisis' state which is characterised by much higher spreads, volatility and default risk. A regime shift of this kind can not be replicated in most other credit risk models, but it is a very important feature of real debt markets.

3.3 Setup and Notation

For ease of exposition we first introduce the simplest setup which will be generalised in the following sections to include positive recovery rates, multiple defaults and jumps in the defaultable term structure.

The model is set in a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{(t\geq 0)}, P)$ where P is some subjective probability measure. We assume the filtration $(\mathcal{F}_t)_{(t\geq 0)}$ satisfies the usual conditions[2] and the initial filtration $\mathcal{F}_0$ is trivial. We also assume a finite time horizon $\bar{T}$ with $\mathcal{F} = \mathcal{F}_{\bar{T}}$, all definitions and statements are understood to be only valid until this time horizon $\bar{T}$.

The time of default is defined as follows:

Definition 1
The time of default *is a stopping time* τ. *We denote with* $N(t) := \mathbf{1}_{\{\tau\leq t\}}$ *the* default indicator function *and* $A(t)$ *the predictable compensator of* $N(t)$, *thus*

$$M(t) := N(t) - A(t)$$

is a (purely discontinuous) martingale. A is nondecreasing (because N is), predictable and of finite variation. Frequently we will assume that N has an intensity $\lambda(s)$ which means that A can be represented as

$$A(t) = \int_0^t \lambda(s)\,ds. \tag{3.1}$$

The filtration $(\mathcal{F}_t)_{(t\geq 0)}$ is generated[3] by n independent Brownian motions W^i, $i = 1, \dots, n$ and the default indicator $N(t)$.

For the default-free bond markets we use the HJM setup:

Definition 2

1. *At any time t there are* default-risk free zero coupon bonds *of all maturities $T > t$. The time-t price of the bond with maturity T is denoted by $B(t,T)$.*

[2] See Jacod and Shiryaev [1988].

[3] This assumption will be relaxed later on to include a marked point process in the case of multiple defaults.

2. *The* continuously compounded default-free forward rate *over the period* $[T_1, T_2]$ *contracted at time* t *is defined (for* $T_2 > T_1 \geq t$*)*

$$f(t, T_1, T_2) = \frac{1}{T_2 - T_1} \left(\ln B(t, T_1) - \ln B(t, T_2) \right). \tag{3.2}$$

3. *If the* T*-derivative of* $B(t, T)$ *exists, the* continuously compounded instantaneous default-free forward rate *at time* t *for date* $T > t$ *is defined as*

$$f(t, T) = -\frac{\partial}{\partial T} \ln B(t, T). \tag{3.3}$$

4. *The* instantaneous default-free short rate $r(t)$,
the default-free discount factor $\beta(t)$
and the default-free bank account $b(t)$ *are defined by*

$$r(t) := f(t, t), \qquad \beta(t) := \exp\{-\int_0^t r(s)ds\}, \quad b(t) := 1/\beta(t). \tag{3.4}$$

We use similar notation to describe the term structure of the defaultable bonds:

Definition 3

1. *At any time* t *there are* defaultable zero coupon bonds *of all maturities* T *(where* $T > t$*). The time-*t *price of the bond with maturity* T *is denoted by* $\overline{B}(t, T)$*. The payoff at time* T *of this bond is* $\mathbf{1}_{\{\tau > T\}} = 1 - N(t)$*: one unit of account if the default has not occurred until* T*, and nothing otherwise.*

2. *The* continuously compounded defaultable forward rate *over the period* $[T_1, T_2]$ *contracted at time* t *is defined (for* $T_2 > T_1 \geq t$*)*

$$\overline{f}(t, T_1, T_2) = \frac{1}{T_2 - T_1} \left(\ln \overline{B}(t, T_1) - \ln \overline{B}(t, T_2) \right). \tag{3.5}$$

3. *If the* T*-derivative of* $\overline{B}(t, T)$ *exists, the* continuously compounded instantaneous defaultable forward rate *at time* t *for date* $T > t$ *is defined as*

$$\overline{f}(t, T) = -\frac{\partial}{\partial T} \ln \overline{B}(t, T). \tag{3.6}$$

4. *The* instantaneous defaultable short rate $\overline{r}(t)$,
the defaultable discount factor $\overline{\beta}(t)$
and the defaultable bank account $c(t)$ *are defined by*

$$\overline{r}(t) := \overline{f}(t, t), \quad \overline{\beta}(t) := \exp\{-\int_0^t \overline{r}(s)ds\}, \quad c(t) := \mathbf{1}_{\{t < \tau\}} \frac{1}{\overline{\beta}(t)}. \tag{3.7}$$

All definitions of defaultable interest rates are only valid for times $t < \tau$ *before default.*

To shorten notation the reference to the continuous compounding frequency of the interest rates or yields is often omitted, and we only refer to default-free and defaultable *forward rates* (=continuously compounded instantaneous forward rates) and *short rates* (=instantaneous short rates). We will also often use *default-free* in place of *default-free* to denote non-defaultable quantities.

The defaultable forward rate $\overline{f}(t, T_1, T_2)$ as it is defined above is *not* the value of a T_1-forward contract on a defaultable bond with maturity T_2, but the *promised* yield of the following portfolio:

short	one	defaultable bond	$\overline{B}(t, T_1)$
long	$\overline{B}(t, T_1)/\overline{B}(t, T_2)$	defaultable bonds	$\overline{B}(t, T_2)$.

A forward contract on the defaultable bond T_2 would involve a short position in the default free bond $B(t, T_1)$. See also section 3.5.2 for some consequences of this definition.

The defaultable bank account $c(t)$ is the value of S 1 invested at $t = 0$ in a defaultable zero coupon bond of very short maturity and rolled over until t, given there has been no default until t. It will play a similar role to the default-free bank account $b(t)$ in default-free interest rate modelling.

In the definition of the defaultable forward rates — to avoid taking logarithms of defaultable bond prices that are zero — we assume that a future default cannot be predicted with certainty. At any time $t < \tau$ strictly before default, and for every finite prediction-horizon T $(t < T < \infty)$ the probability of a default until T is not one: $P[\tau \leq T \mid \mathcal{F}_t] < 1$. This can be achieved by setting the default time to be the first time τ' at which a future default can be predicted with certainty: $\tau' := \inf\{t \geq 0 \mid \exists\ T < \infty \text{ s.t. } P[\tau \leq T \mid \mathcal{F}_t] = 1\}$, effectively moving the time of default forward in time[4].

We assume τ has been defined as above. This assumption is in keeping with the real-world legal provisions that a bankruptcy must be filed as soon as the fact of the bankruptcy is known. Furthermore it does not change any of the qualitative features of the model. In addition to this we assume that all (forward) interest rates have continuous paths and that the instantaneous forward rates are well-defined.

3.4 Pricing with Zero Recovery

3.4.1 Dynamics: The defaultable Forward Rates

Given the above definitions we can start to explore the connections between the dynamics of the defaultable bond prices and the defaultable forward rates. We assume the following representation as stochastic integrals for the dynamics of the defaultable forward rates $\overline{f}(t, T)$ and the defaultable bonds $\overline{B}(t, T)$:

[4] By definition $\tau' \leq \tau$, but $\tau' = \infty$ is possible.

Assumption 1

1. The dynamics of the defaultable forward rates are given by

$$d\overline{f}(t,T) = \overline{\alpha}(t,T)\,dt + \sum_{i=1}^{n} \overline{\sigma}_i(t,T)\,dW^i(t). \tag{3.8}$$

2. The dynamics of the defaultable bond prices are [5]

$$\frac{d\overline{B}(t,T)}{\overline{B}(t-,T)} = \overline{\mu}(t,T)dt + \sum_{i=1}^{n} \overline{\eta}_i(t,T)\,dW^i(t) - dN(t). \tag{3.9}$$

3. The integrands $\overline{\alpha}(t,T), \overline{\sigma}_i(t,T), \overline{\mu}(t,T)$ and $\overline{\eta}_i(t,T)$ are predictable processes that are regular enough to allow

- *differentiation under the integral sign*
- *interchange of the order of integration*
- *partial derivatives with respect to the T-variable*
- *bounded prices $\overline{B}(t,\cdot)$ for almost all $\omega \in \Omega$.*

We start by analysing the consequences of the specification (3.8) of the defaultable forward rates. The *dynamics of the defaultable spot rate process* are [6]

$$\begin{aligned} \overline{r}(t) &= \overline{f}(t,t) = \overline{f}(0,t) + \int_0^t \overline{\alpha}(s,t)ds \\ &+ \sum_{i=1}^{n} \int_0^t \overline{\sigma}_i(s,t)dW^i(s). \end{aligned} \tag{3.10}$$

From definition (3.6) of the defaultable forward rates and definition 3 of the defaultable bonds the price of a defaultable zero coupon bond is given by

$$\overline{B}(t,T) = (1 - N(t)) \exp\left\{ -\int_t^T \overline{f}(t,s)\,ds \right\}. \tag{3.11}$$

The factor of $(1 - N(t))$ follows from the default condition $\overline{B}(t,T) = 0$ for $t \geq \tau$. Writing $G(t,T) := \int_t^T \overline{f}(t,s)\,ds$ this yields for $t \leq \tau$ using Itô's lemma

$$d\overline{B}(t,T)/\overline{B}(t-,T) = -dG(t,T) + \frac{1}{2}d < G,G > -dN, \tag{3.12}$$

where we have used that G is continuous. For the process $G(t,T)$ we have (see HJM [1992])

$$G(t,T) - G(0,T) = \int_t^T \left[\overline{f}(t,s) - \overline{f}(0,s)\right]\,ds - \int_0^t \overline{f}(0,s)\,ds$$

[5] The notation $dY(t)/Y(t-) = dX(t)$ is a shorthand for $dY(t)/Y(t-) = dX(t)$ for $Y(t-) > 0$ and $dY(t) = 0$ for $Y(t-) = 0$.

[6] It is understood that dynamics of defaultable interest rates are always the dynamics before default $t < \tau$.

$$
\begin{aligned}
&= \int_t^T \int_0^t \overline{\alpha}(u,s)\,du\,ds + \sum_{i=1}^n \int_t^T \int_0^t \overline{\sigma}_i(u,s)\,dW^i(u)\,ds \\
&\quad - \int_0^t \overline{f}(0,s)\,ds \\
&= \int_0^t \int_t^T \overline{\alpha}(u,s)\,ds\,du + \sum_{i=1}^n \int_0^t \int_t^T \overline{\sigma}_i(u,s)\,ds\,dW^i(u) \\
&\quad - \int_0^t \overline{f}(0,s)\,ds \\
&= \int_0^t \int_u^T \overline{\alpha}(u,s)\,ds\,du + \sum_{i=1}^n \int_0^t \int_u^T \overline{\sigma}_i(u,s)\,ds\,dW^i(u) \\
&\quad - \int_0^t \overline{f}(0,s)\,ds - \int_0^t \int_u^t \overline{\alpha}(u,s)\,ds\,du \\
&\quad - \sum_{i=1}^n \int_0^t \int_u^t \overline{\sigma}_i(u,s)\,ds\,dW^i(u) \\
&= \int_0^t \overline{\gamma}(u,T)\,du - \sum_{i=1}^n \int_0^t \overline{a}_i(u,T)\,dW^i(u) \\
&\quad - \int_0^t \overline{f}(0,s)\,ds - \int_0^t \int_0^s \overline{\alpha}(u,s)\,du\,ds \\
&\quad - \sum_{i=1}^n \int_0^t \int_0^s \overline{\sigma}_i(u,s)\,dW^i(u)\,ds \\
&= \int_0^t (\overline{\gamma}(u,T) - \overline{r}(u))\,du - \sum_{i=1}^n \int_0^t \overline{a}_i(u,T)\,dW^i(u)
\end{aligned}
$$

where

$$
\overline{a}_i(t,T) := -\int_t^T \overline{\sigma}_i(t,v)\,dv \tag{3.13}
$$

$$
\overline{\gamma}(t,T) := \int_t^T \overline{\alpha}(t,v)\,dv. \tag{3.14}
$$

The main tool in the equations above is Fubini's theorem and Fubini's theorem for stochastic integrals (see e.g. HJM [1992] and Protter [1990]).

With this result we reach the dynamics of the defaultable zero coupon bond prices

$$
\begin{aligned}
\frac{d\overline{B}(t,T)}{\overline{B}(t-,T)} &= \left[-\overline{\gamma}(t,T) + \overline{r}(t) + \frac{1}{2}\sum_{i=1}^n \overline{a}_i^2(t,T)\right] dt \\
&\quad + \sum_{i=1}^n \overline{a}_i(t,T)\,dW^i(t) - dN(t).
\end{aligned} \tag{3.15}
$$

The final condition $\overline{B}(T,T) = 0$ for $\tau < T$ is automatically satisfied by the functional specification of $\overline{B}$.

The above derivation of the dynamics of $G(t,T)$ follows the derivation of the dynamics of the default-free bond prices in HJM [1992]. Here the only addition is the jump term $-dN(t)$ which is introduced by the default process. Summing up:

Proposition 1

1. Given the dynamics of the defaultable forward rates (3.8)

(i) the dynamics of the defaultable bond prices are given by

$$\frac{d\overline{B}(t,T)}{\overline{B}(t-,T)} = \left[-\overline{\gamma}(t,T) + \overline{r}(t) + \frac{1}{2}\sum_{i=1}^{n}\overline{a}_i^2(t,T)\right] dt + \sum_{i=1}^{n}\overline{a}_i(t,T)\, dW^i(t) - dN(t). \tag{3.16}$$

where $\overline{a}_i(t,T)$ and $\overline{\gamma}(t,T)$ are defined by (3.13) *and* (3.14) *resp..*

(ii) the dynamics of the defaultable short rate are given by

$$\overline{r}(t) = \overline{f}(t,t) = \overline{f}(0,t) + \int_0^t \overline{\alpha}(s,t)ds + \sum_{i=1}^{n}\int_0^t \overline{\sigma}_i(s,t)dW^i(s). \tag{3.17}$$

2. Given the dynamics (3.9) of the defaultable bond prices the dynamics of the defaultable forward rates are (for $t \le \tau$) given by (3.8) with

$$\overline{\alpha}(t,T) = \sum_{i=1}^{n}\overline{\eta}_i(t,T)\frac{\partial}{\partial T}\overline{\eta}_i(t,T) - \frac{\partial}{\partial T}\overline{\mu}(t,T) \tag{3.18}$$

$$\overline{\sigma}_i(t,T) = -\frac{\partial}{\partial T}\overline{\eta}_i(t,T). \tag{3.19}$$

Proof: 1.) has been derived above, 2.) follows from Itô's lemma on $\ln \overline{B}(t,T)$ and taking the partial derivative w.r.t. T.

□

These relationships are well-known in the case of the default-risk free term structure. Assume the following dynamics of the default-free forward rates $f(t,T)$ and the default-free bond prices $B(t,T)$:

Assumption 2

1. The dynamics of the default risk free forward rates are given by

$$df(t,T) = \alpha(t,T)\, dt + \sum_{i=1}^{n}\sigma_i(t,T)\, dW^i(t). \tag{3.20}$$

2. *The dynamics of the default risk free bond prices are*

$$\frac{dB(t,T)}{B(t-,T)} = \mu(t,T)dt + \sum_{i=1}^{n} \eta_i(t,T)\, dW^i(t). \tag{3.21}$$

3. *The integrands $\alpha(t,T), \sigma_i(t,T), \mu(t,T)$ and $\eta_i(t,T)$ are predictable processes that are regular enough to allow*
   - *differentiation under the integral sign*
   - *interchange of the order of integration*
   - *partial derivatives with respect to the T-variable*
   - *bounded prices $B(t,\cdot)$ for almost all $\omega \in \Omega$.*

The dynamics of the default-free term structure do not contain any jumps at τ. Volatilities and drifts may change at τ but the direct impact of the default is only on the defaultable bonds.

Given these dynamics the following proposition is a well-known result by Heath, Jarrow and Morton [1992].

Proposition 2

1. Given the dynamics of the risk free forward rates (3.20)

(i) the dynamics of the risk free bond prices are given by

$$\begin{aligned} \frac{dB(t,T)}{B(t-,T)} &= \left[-\gamma(t,T) + r(t) + \frac{1}{2}\sum_{i=1}^{n} a_i^2(t,T)\right] dt \\ &\quad + \sum_{i=1}^{n} a_i(t,T)\, dW^i(t). \end{aligned} \tag{3.22}$$

where $a_i(t,T)$ and $\gamma(t,T)$ are defined by

$$a_i(t,T) := -\int_t^T \sigma_i(t,v)\, dv \tag{3.23}$$

$$\gamma(t,T) := \int_t^T \alpha(t,v)\, dv. \tag{3.24}$$

(ii) the dynamics of the risk free short rate are given by

$$\begin{aligned} r(t) &= f(t,t) = f(0,t) + \int_0^t \alpha(s,t)ds \\ &\quad + \sum_{i=1}^{n} \int_0^t \sigma_i(s,t)dW^i(s). \end{aligned} \tag{3.25}$$

2. Given the dynamics (3.21) of the risk free bond prices the dynamics of the defaultable forward rates are given by (3.20) with

$$\alpha(t,T) = \sum_{i=1}^{n} \eta_i(t,T)\frac{\partial}{\partial T}\eta_i(t,T) - \frac{\partial}{\partial T}\mu(t,T) \tag{3.26}$$

$$\sigma_i(t,T) = -\frac{\partial}{\partial T}\eta_i(t,T). \tag{3.27}$$

3.4.2 Change of Measure

Now that the connections between the dynamics of the defaultable zero coupon bonds and the forward rates are clarified, we can start analysing the conditions for absence of arbitrage opportunities in this model. We use the following standard definition:

Definition 4
There are no arbitrage opportunities if and only if there is a probability measure Q equivalent to P under which the discounted security price processes become local martingales. This measure Q is called the martingale measure, *and for any security price process $X(t)$ the discounted price process is defined as $\beta(t)X(t)$.*

The main tool to classify all to P equivalent probability measures is the following version of Girsanov's Theorem (see Jacod and Shiryaev [1988] III.3 and III.5 and BKR [1996]):

Theorem 3
Assume that the default process has an intensity. Let θ be a n-dimensional predictable processes $\theta_1(t), \dots, \theta_n(t)$ and $\phi(t)$ a strictly positive predictable process with

$$\int_0^t \|\theta(s)\|^2 ds < \infty, \qquad \int_0^t |\phi(s)|\lambda(s) ds < \infty$$

for finite t. Define the process L by $L(0) = 1$ and

$$\frac{dL(t)}{L(t-)} = \sum_{i=1}^{n} \theta_i(t) dW^i(t) + (\phi(t) - 1) dM(t).$$

Assume that $\mathrm{E}[\, L(t) \,] = 1$ for finite t.
Then there is a probability measure Q equivalent to P with

$$dQ = L_T dP \qquad \text{and} \qquad dQ_t = L_t dP_t \tag{3.28}$$

where $Q_t := Q|\,\mathcal{F}_t$ and $P_t := P|\,\mathcal{F}_t$ are the restrictions of Q and P on $\mathcal{F}_t$, such that

$$dW(t) - \theta(t)dt = d\widetilde{W}(t) \tag{3.29}$$

defines $\widetilde{W}$ as Q-Brownian motion and

$$\lambda_Q(t) = \phi(t)\lambda(t) \tag{3.30}$$

is the intensity of the default indicator process under Q.
Furthermore every probability measure that is equivalent to P can be represented in the way given above.

In the financial context here the processes θ_i are the *market prices of diffusion risk*, and the process ϕ represents a *market premium on jump risk* (in terms of a multiplicative factor per unit of jump intensity). To ensure absence of arbitrage the financial requirement of a well-defined set of market prices of risk with validity for all securities translates into the mathematical requirement of having a well-defined intensity process for the change of measure.

Given the defaultable bond price dynamics (3.9) the change of measure to the martingale measure leaves the volatilities of the defaultable bond prices unaffected, the same is true of the integral with respect to dN (the *compensator* of this integral has changed, though), the only effect is a change of drift in the defaultable bond price process.

From now on we will assume that the change of measure to the martingale measure has already been performed. The results of the preceding section on the dynamics remain valid if the underlying measure is the martingale measure. Therefore we simplify notation such that all specifications in section 3.4.1 are already with respect to Q. [7]

3.4.3 Absence of Arbitrage

By Itô's lemma we require under the martingale measure for absence of arbitrage that for all $t > T$

$$\mathrm{E}\left[\frac{d\overline{B}(t,T)}{\overline{B}(t-,T)}\right] = r(t)\,dt. \tag{3.31}$$

This means using (3.16)

$$\begin{aligned} r(t)\,dt &= \mathrm{E}\left[\frac{d\overline{B}(t,T)}{\overline{B}(t-,T)}\right] \\ &= \mathrm{E}\left[\left[-\overline{\gamma}(t,T) + \overline{r}(t) + \frac{1}{2}\sum_{i=1}^{n} a_i^2(t,T)\right] dt\right] \\ &\quad +\mathrm{E}\left[\sum_{i=1}^{n} \overline{a}_i(t,T)\,dW^i(t) - dN(t)\right] \\ r(t) &= -\overline{\gamma}(t,T) + \overline{r}(t) + \frac{1}{2}\sum_{i=1}^{n} a_i^2(t,T) - \lambda(t) \end{aligned} \tag{3.32}$$

Now we have to take a closer look at the compensator $A(t)$ of the default indicator process $N(t)$. $A(t)$ is increasing and we assumed that $A(t)$ is also continuous, therefore $A(t)$ is differentiable almost everywhere on $\mathbb{R}^+$ and thus $N(t)$ has an intensity $dA(t) = \lambda(t)dt$. Then

$$-M(t) = -N(t) + A(t) = -N(t) + \int_0^t \lambda(s)\,ds \tag{3.33}$$

is a martingale by the definition of the predictable compensator. Now consider the value process of the *defaultable* bank account $c(t)$, i.e. the development of \$ 1 invested at time 0 at the defaultable short rate and rolled over from then on. By definition its value at time t is

$$c(t) = 1_{\{\tau > t\}} \exp\{\int_0^t \overline{r}(s)ds\}. \tag{3.34}$$

[7] If $P = Q$ then $\theta \equiv 0$ and $\phi \equiv 1$.

Under the martingale measure the discounted (discounting with the *default-free* interest rate) value process of c

$$\bar{c}(t) := \frac{c(t)}{b(t)} = \mathbf{1}_{\{\tau > t\}} \exp\{\int_0^t \bar{r}(s) - r(s)\, ds\} \tag{3.35}$$

must be a martingale. This is the Doleans-Dade exponential of

$$\hat{M}(t) := -N(t) + \int_0^{t \wedge \tau} \bar{r}(s) - r(s)\, ds, \tag{3.36}$$

which in turn must also be a martingale. (The martingale property can also be seen from $\hat{M}(t) = \int_0^t \frac{1}{\bar{c}(s-)} d\bar{c}(s)$ and the uniqueness of the Doleans-Dade exponential up to τ.) We use the freedom we had in the specification of $\bar{r}(t)$ for $t \geq \tau$ and set $\bar{r}(t) := r(t)$ for $t \geq \tau$.

Taking the difference of (3.33) and (3.36)

$$M(t) - \hat{M}(t) = \int_0^t \lambda(s) - \bar{r}(s) + r(s)\, ds \tag{3.37}$$

one sees that – while the l.h.s. is a martingale – the r.h.s. is predictable, the only predictable martingales are constant, thus we have for (almost) all s

$$\lambda(s) = \bar{r}(s) - r(s). \tag{3.38}$$

The hazard rate $\lambda(s)$ of the default is exactly the short interest rate spread. Note that this relationship can also be inverted to define the defaultable short rate as $\bar{r}(s) := r(s) + \lambda(s)$.

Equation (3.38) is the key relation that yields, substituted in (3.32), as necessary condition for the absence of arbitrage:

$$\bar{\gamma}(t,T) = \frac{1}{2} \sum_{i=1}^{n} a_i^2(t,T). \tag{3.39}$$

Substituting the definition of $\bar{\gamma}$ in this condition yields the results of the following theorem.

Theorem 4
The following are equivalent:

1. *The measure under which the dynamics are specified is a martingale measure.*
2. *(i) The short interest rate spread is the intensity of the default process. It is nonnegative.*

$$\lambda(t) = \bar{r}(t) - r(t) \geq 0. \tag{3.40}$$

(ii) The drift coefficients of the defaultable forward rates satisfy for all $t \leq T$, $t < \tau$

$$\int_t^T \bar{\alpha}(t,v)dv = \frac{1}{2} \sum_{i=1}^{n} \left(\int_t^T \bar{\sigma}_i(t,v)dv \right)^2 \tag{3.41}$$

or, differentiated,

$$\bar{\alpha}(t,T) = \sum_{i=1}^{n} \bar{\sigma}_i(t,T) \int_t^T \bar{\sigma}_i(t,v)dv. \tag{3.42}$$

(iii) The drift coefficients of the default-free forward rates satisfy for all $t \leq T$

$$\alpha(t,T) = \sum_{i=1}^{n} \sigma_i(t,T) \int_t^T \sigma_i(t,v)dv. \tag{3.43}$$

3. *(i)* $\overline{r}(t) - r(t) = \lambda(t) > 0$.

(ii) The dynamics of the defaultable bond prices are given by

$$\frac{d\overline{B}(t,T)}{\overline{B}(t-,T)} = \overline{r}(t)dt + \sum_{i=1}^{n} \overline{a}_i(t,T)dW^i(t) - dN(t) \tag{3.44}$$

or, solving the s.d.e.

$$\begin{aligned}\overline{B}(t,T) &= \mathbf{1}_{\{\tau>t\}}\overline{B}(0,T)\exp\left\{\int_0^t \overline{r}(s)\,ds - \frac{1}{2}\sum_{i=1}^{n}\int_0^t a_i^2(s,T)\,ds\right. \\ &\quad \left. + \sum_{i=1}^{n}\int_0^t \overline{a}_i(s,T)\,dW^i(s)\right\}. \end{aligned} \tag{3.45}$$

(iii) The dynamics of the default-free bonds satisfy under the martingale measure

$$\frac{dB(t,T)}{B(t-,T)} = r(t)dt + \sum_{i=1}^{n} a_i(t,T)dW^i(t). \tag{3.46}$$

Proof: 1.) $\Rightarrow$ 2.) : (i) and (ii) have been derived above, (iii) has been shown in HJM [1992].
2.) $\Rightarrow$ 3.) : 2.(i) and 3.(i) coincide, 3.(ii) follows from 2.(ii) and (i) by substituting in proposition 1, again (iii) is by HJM [1992].
3.) $\Rightarrow$ 1.): follows from the definition of the martingale measure.

□

The most important result of this section is equation (3.42), the defaultable-bond equivalent of the well-known Heath-Jarrow-Morton drift-restriction. This restriction has been derived for default-free bonds in HJM [1992], and, as we see here, it is also an important part of the modelling of the defaultable bonds' dynamics. This was first noted by Duffie [1994] and Lando [1998].

Another important insight is that precise knowledge of the nature of the default process N and its compensator A is not necessary for setting up an arbitrage-free model of the term structure of defaultable bonds. With the restrictions 2.) of theorem 4 one can set up a model of defaultable bonds that uses readily observable market data (the term structure of the defaultable forward rates) as input, without having to try and find out about the precise nature of N.

We assumed that the default process has an intensity: $dA(t) = \lambda(t)dt$. This (and the fact that $\mathcal{F}_0$ is trivial and thus $M(0) = 0$ a.s.) implies that the time of default is a totally inaccessible stopping time. Dropping this assumption (to allow discontinuities in A) one sees readily from the derivation of equation (3.38) that the defaultable spot rate $\overline{r}$ cannot be finite at jumps of A. One would have to specify the defaultable term structure in a more general way by defining a process $R^d(t) := \int_0^t \bar{r}(s)ds$ which will be well-defined and can account for the jumps in A. Similar definitions will be needed for the forward rates. Then (3.38) translates into $dR^d(t) = r(t)dt + dA(t)$. With this specification we can also drop the initial assumption that a default cannot be predicted with certainty.

It is important to note that the default-free term structure and the defaultable term structure must satisfy the conditions simultaneously. This will become clearer in the following version of theorem 4 that is set under the *statistical* measure P:

Theorem 5
If the dynamics are given under a subjective probability measure P the following are equivalent:

1.) The dynamics are arbitrage-free.
2.) There are predictable processes $\theta_1(t), \dots, \theta_n(t)$ and a strictly positive predictable process $\phi(t)$ that satisfy the regularity conditions of theorem 3 such that for all $t < T$:
(i) The difference between default-free and defaultable short rate is ϕ times the hazard rate:

$$\bar{r}(t) - r(t) = \phi(t)\lambda(t). \tag{3.47}$$

(ii) The defaultable and the default-free forward rates satisfy

$$-\bar{\alpha}(t,T) + \sum_{i=1}^{n} \bar{\sigma}_i(t,T) \int_t^T \bar{\sigma}_i(t,v)\,dv = \sum_{i=1}^{n} \bar{\sigma}_i(t,T)\theta_i(t) \tag{3.48}$$

$$-\alpha(t,T) + \sum_{i=1}^{n} \sigma_i(t,T) \int_t^T \sigma_i(t,v)\,dv = \sum_{i=1}^{n} \sigma_i(t,T)\theta_i(t) \tag{3.49}$$

a.s. for all $t < T$.

Proof: 1.) $\Leftrightarrow$ There is an equivalent martingale measure Q
$\Leftrightarrow$ (Theorem 3) There are predictable processes $\theta_1(t), \dots, \theta_n(t)$ and a strictly positive predictable process $\phi(t)$ that satisfy the regularity conditions of theorem 3 such that (using theorem 4):
(i) $\lambda^Q = \phi\lambda$ and $h^Q(t) = \bar{r}(t) - r(t) = \phi(t)\lambda(t)$.
(ii) $dW_i^Q = dW_i - \theta_i dt$ and

$$d\bar{f}(t,T) = \sum_{i=1}^{n} \left(\bar{\sigma}_i(t,T) \int_t^T \bar{\sigma}_i(t,v)\,dv \right) dt + \sum_{i=1}^{n} \bar{\sigma}_i(t,T) dW_i^Q(t)$$

$$df(t,T) = \sum_{i=1}^{n} \left(\sigma_i(t,T) \int_t^T \sigma_i(t,v)\,dv \right) dt + \sum_{i=1}^{n} \sigma_i(t,T) dW_i^Q(t).$$

By substituting the P-dynamics of $\bar{f}(t,T)$ and $f(t,T)$ and equating coefficients the proof is concluded.

□

From theorem 5 one sees directly that there is only one set of market prices of risk for both the defaultable and the default-free term structure. This follows from the fact that there is only one set of underlying Brownian motions that drive the market. The market price of jump risk ϕ is uniquely determined by equation (3.47) which can be used as defining relationship for ϕ. Even with defaultable zero coupon bonds one can set up portfolios that are hedged against default risk (by making sure that the total value of the portfolio is zero), but still carry exposure to the Brownian motions. These portfolios must be related to the default-free term structure in their dynamics, and this relation is given in the two theorems above.

3.5 Modelling the Spread between the Forward Rates

When trying to connect the dynamics of the defaultable term structure and the default-free term-structure the most important relationship is (3.40):

$$0 \leq \lambda(t) = \overline{r}(t) - r(t).$$

In many cases a model of the default-free interest rates and forward rates will already be in place and the task is to find a specification of a model of the defaultable term structure that does not violate (3.40). If one directly estimated and implemented a model for the defaultable term structure $\overline{f}(t,T)$ without reference to the existing model of the default-free term structure, situations where $\overline{r} < r$ are bound to arise and the (combined) model will not be arbitrage-free.

A way around this problem is not to model the forward rates but the difference between these as proposed in Duffie [1994]:

Definition 5
The continuously compounded instantaneous forward rate spread $h(t,T)$ *is defined as the difference between the defaultable forward rate and the default-free forward rate:*

$$h(t,T) = \overline{f}(t,T) - f(t,T). \tag{3.50}$$

Under the martingale measure we have $\lambda(t) = h(t,t)$.

Now one has to find a model for $h(t,T)$ which is compatible with theorem 4. The advantage of modelling $h(t,T)$ instead of $\overline{f}(t,T)$ is that (3.40) reduces to the well known problem of ensuring that $h(t,t) > 0$. and we can hope to use some of the extensive literature on interest rate models with positive short rates. We use the following dynamics for h:

Assumption 3
The dynamics of h are given by

$$h(t,T) - h(0,T) = \int_0^t \alpha^h(v,T)\,dv + \sum_{i=1}^n \int_0^t \sigma_i^h(v,T)dW^i(v). \tag{3.51}$$

Then

$$\begin{aligned} \overline{\alpha}(v,t) &= \alpha(v,t) + \alpha^h(v,t) &(3.52)\\ \overline{\sigma}_i(v,t) &= \sigma_i(v,t) + \sigma_i^h(v,t). &(3.53) \end{aligned}$$

In place of the drift restriction (3.42) we reach:

Corollary 6
Let the default-free interest rates satisfy the HJM drift restriction (3.43). A model for the defaultable forward rates based on the forward rate spread $h(t,T)$ must imply under the martingale measure

$$\begin{aligned}\alpha^h(t,T) &= \sum_{i=1}^{n}\Big[\sigma_i(t,T)\int_t^T \sigma_i^h(t,v)\,dv \\ &\quad +\sigma_i^h(t,T)\int_t^T \sigma_i(t,v)\,dv \\ &\quad +\sigma_i^h(t,T)\int_t^T \sigma_i^h(t,v)\,dv\Big]. \end{aligned} \tag{3.54}$$

Proof: Substitute (3.52) and (3.53) in (3.42).

□

Again – as in the original HJM model – the drift of the spread is given in terms of the volatilities of the interest rates and spreads. Given these drift specifications one has to require that the process $h(t,t)$ is nonnegative. This will enable us to add a defaultable interest rate model to a given model of the default-free interest rate in a modular fashion.

If one chooses a specification of the dynamics of h that has nonnegative $h(t,t)$ a.s. under the subjective measure, this will ensure that $h(t,t)$ will be nonnegative a.s. under the martingale measure, too.

It will be interesting to analyse some possible specifications and the problems that may arise when modeling the spread structure.

3.5.1 Independence of Spreads and default-free Rates

The easiest way to specify the spreads is to avoid the cross-variation terms with the default-free term-structure in (3.54). Assume that every factor W_i either influences f or h but never both. Then $\forall\, i = 1,\ldots,n$

$$\begin{aligned}\sigma_i^h(t,T)\neq 0 \quad &\Rightarrow \quad \int_t^T \sigma_i(t,v)\,dv = 0 \\ \sigma_i(t,T)\neq 0 \quad &\Rightarrow \quad \int_t^T \sigma_i^h(t,v)\,dv = 0\end{aligned} \tag{3.55}$$

and the drift restriction for the spreads becomes the usual HJM restriction:

Corollary 7
(i) If σ_i^h and σ_i satisfy (3.55) then (under the martingale measure)

$$\alpha^h(t,T) = \sum_{i=1}^{n}\sigma_i^h(t,T)\int_t^T \sigma_i^h(t,v)dv \tag{3.56}$$

and $h(t,t) > 0$ a.s. are necessary and sufficient for absence of arbitrage.

(ii) *Equation (3.55) is satisfied if $h(t,T_1)$ and $\overline{f}(t,T_2)$ are independent for all $t \le T_1$, $t \le T_2$, i.e. the term structure of the spreads and the term structure of the default-free forward rates are independent.*

Proof: The first part follows directly by substituting the assumptions (3.55) in (3.54). For the last part observe that independence of the term structures of spreads and default-free rates implies that

$$\sigma_i(t,T_1)\sigma_i^h(t,T_2) = 0$$

for (almost) all $t \le T_1, T_2$ which in turn implies (3.55) directly.

□

Note that strict independence of h and g is not needed. One might imagine a model where the term structure of the spreads is driven by an additional Brownian motion alone (this will ensure (3.55)), but the volatility of the spread might still depend on the level of the default-free interest rates.

Satisfying the positivity requirement (3.40) on $h(t,t)$ becomes very easy in the setup of corollary 7: One can use any interest rate model for $h(t,T)$ that is known to generate positive short rates, e.g. the square root model of Cox, Ingersoll and Ross [1985] or the model with lognormal interest rates by Sandmann and Sondermann [1997].

3.5.2 Negative Forward Spreads

There is one additional caveat when using the nonnegative rate model for the forward rate spreads: Even though we require that the 'short' spread $h(t,t) > 0$ is greater than zero, a 'forward' spread $h(t,T)$ $(T > t)$ might still become negative.

As an example how this can arise we consider the two-period economy with points in time $t = 0, 1, 2$ from figure 1. There are three states $\omega_1, \omega_2, \omega_3$ and the filtration is $\mathcal{F}_0 = \{\{\omega_1,\omega_2,\omega_3\},\emptyset\}$; $\mathcal{F}_1 = \sigma(\{\omega_1,\omega_2\},\omega_3)$; $\mathcal{F}_2 = \sigma(\omega_1,\omega_2,\omega_3)$. The states have risk-neutral probabilities $P(\omega_1) = pq$; $P(\omega_2) = p(1-q)$; $P(\omega_3) = (1-p)$. There are default-free bonds $B(0,1), B(0,2)$ and defaultable bonds $\overline{B}(0,1), \overline{B}(0,2)$. In state ω_1 both defaultable bonds survive, in ω_2 only the defaultable bond with maturity 2 defaults and in ω_3 both bonds default (with zero recovery).

The initial default-free term structure is given by $B(0,1) = \beta_0$ and $B(0,2) = \beta_0(p\beta_u + (1-p)\beta_d)$. Thus the bond prices are

$$\begin{aligned} B(0,1) &= \beta_0 \\ B(0,2) &= \beta_0(p\beta_u + (1-p)\beta_d) \\ \overline{B}(0,1) &= p\,\beta_0 \\ \overline{B}(0,2) &= p\,q\,\beta_0\beta_u \end{aligned}$$

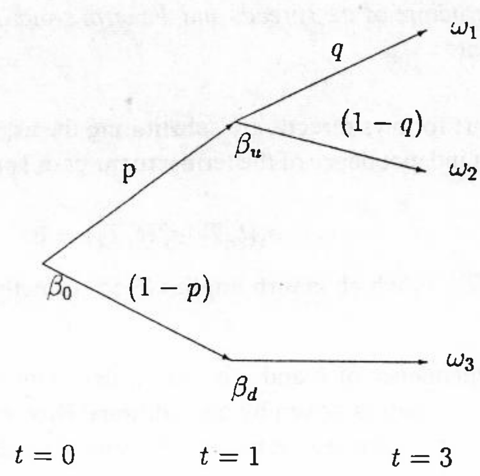

Figure 3.1: Example for negative forward spreads

The forward rates are

$$
\begin{aligned}
f(0,1,2) &= -\ln(p\,\beta_u + (1-p)\,\beta_d) \\
\overline{f}(0,1,2) &= -\ln(q\,\beta_u).
\end{aligned}
$$

We have a negative forward spread $h(0,1,2) < 0$ or $\overline{f}(0,1,2) < f(0,1,2)$ if

$$
p\,\beta_u + (1-p)\,\beta_d < q\,\beta_u \tag{3.57}
$$

holds. This is equivalent to

$$
q > p + (1-p)\frac{\beta_d}{\beta_u}, \tag{3.58}
$$

so a necessary condition for negative forward spreads is that $\beta_d < \beta_u$. Choose for instance $p = 0.9$, $\beta_d/\beta_u = 0.9$, $q = 0.995$.

This example can be regarded as a 'snapshot' from a continuous-time model in which the relevant prices and probabilities have been aggregated to the two-period example.

The condition $\beta_d < \beta_u$ means that $r_d > r_u$. For negative forward spreads to arise we need
– either q and p are of the same order of magnitude, then the ratio β_d/β_u must be very small, $r_u << r_d$,
– or q is much larger than p, then β_u can be of the same order as β_d. In practice $q >> p$ only

occurs if $T_1 >> T_2 - T_1$. But then $\beta_d/\beta_u \sim 1$ because of the short horizon $T_2 - T_1$ which is very far in the future, as well.

The occurrence of negative forward spreads is due to the special way in which we defined the defaultable forward rates. It is not possible to exploit the negative forward spread as an arbitrage-opportunity because the portfolio one would typically use for that will be destroyed by an early default:

In the default-free bonds one can set up a portfolio that replicates the payoff of a default free forward contract, but set up in defaultable bonds this portfolio disappears in the case of an early default. If one had gone long a default-free forward contract and short a *replicating portfolio* of a *defaultable* forward contract, an early default (which eliminates the replicating portfolio for the defaultable forward contract) leaves one with the default-risk free half of the portfolio, which now is exposed to changes in the default-free term structure. If the subsequent default-free interest rates are high the remainder of the portfolio will generate a loss.

Summarizing, negative forward spreads can only occur if there is a strong correlation between early default (event ω_3) and high interest rates (β_d small), and a strong correlation between early survivial (events $\{\omega_1, \omega_2\}$) and low interest rates (β_u large).

3.6 Positive Recovery and Restructuring

In the preceding sections we assumed that a defaultable zero coupon bond has a payoff of zero upon default. This assumption is unnecessarily restrictive and does not agree with market experience.

As mentioned in the introduction, real-world defaults often have the following features.
Positive recovery: Franks and Torous [1994] find recovery rates between 40 and 80 percent.
Reorganisation: The majority of firms in financial distress are reorganised and re-floated, they are not liquidated.
Compensation in terms of new securities: On average about one third of the compensation to the holders of defaulted bonds is in cash , two thirds are in terms of new securities of the defaulted and restructured firm.
Multiple defaults: A reorganised firm can default again. We have the possibility of multiple defaults with in principle any number of defaults (each with subsequent restructuring of the defaulted firm's debt).

The main results of the preceding sections are still valid if the recovery of the bond is positive and not zero. We choose the following setup including the possibility of multiple defaults:

If a default occurs a *restructuring* of the debts occurs. Holders of the old debt loose a fraction of q of their claims, where $q \in [0, 1]$ is possibly unpredictable, but known at default.

A pre-default claim of \$ 1 face value becomes a claim of \$ $(1 - q)$ face value after the default. The maturity of the claim remains unchanged.

This model mimicks the effect of a rescue plan as it is described in many bankruptcy codes: The old claimants have to give up some of their claims in order to allow for rescue capital to be invested in the defaulted firm. They are *not* paid out in cash [8] (this would drain the defaulted firm of valuable liquidity) but in 'new' defaultable bonds of the same maturity. As the vast majority of defaulted debtors continue to operate after default, a representation of the loss of a defaulted bond in terms of a reduction in face value is possible even if the actual payoff procedure is different.

For ease of modelling we use the convention that the *defaultable forward rates are quoted with respect to a bond of face value 1 \$.*

The reduction in the face value of a bond in default is *not* reflected in the forward rates. This convention enables us to separate the effects of changes in interest rates (representing expectations on future defaults) and the direct effect of the default.

This modelling approach has much in common with the *fractional recovery* introduced by Duffie and Singleton [1997, 1999] and Duffie et.al. [1994, 1996, 1994].

Duffie specifies the payoff in default to be a predictable fraction $(1-q)$ of the value of a 'non-defaulted but otherwise equivalent security'. This is inspired by the default procedures in swap contracts. In Duffie's mathematical model the value V_τ of the defaulted security directly after default (i.e. the payoff in default) is specified as $V_\tau := (1-q)V_{\tau-}$, the fraction $(1-q)$ times the value of the same security directly before default. [9]

Furthermore we include *magnitude risk* in our setup. The magnitude of the default is uncertain and the actual realisation of the loss q_i need not be predictable, it can be considered as a random draw at τ_i from the distribution $K(dq)$. This distribution may itself be stochastic.

The analysis of the pricing of defaultable zero coupon bonds goes along the lines of the zero-recovery case, where we increasingly have to introduce the theory of marked point processes. Standard references are Jacod and Shiryaev [1988], and Bremaud [1981] for the mathematical theory and Jarrow and Madan [1995] and Björk, Kabanov and Runggaldier [1996] and Björk, Di Masi, Kabanov and Runggaldier [1997] for the application to interest rate theory[10].

3.6.1 The Model Setup

Mathematically the setup is as follows:

Assumption 4

(i) *Defaults occur at the stopping times $\tau_1 < \tau_2 < \ldots$.*

(ii) *At each time τ_i of default a loss quota $q_i \in E$ is drawn from a measurable space $(E, \mathcal{E})$, $E \subset$ $\mathbb{R}$,* the mark space. *(Usually $E = [0,1]$ with the Borel sets.)*

---

[8] The holder of a defaulted bond is free to sell this bond on the market, though.

[9] The model presented here can be extended to include the possibility of predictable times of default. (See the remarks at the end of section 5.) In addition to this we avoid the backward recursive stochastic integral equations that are necessary in the Duffie model.

[10] For other financial applications see also Merton [1976].

(*iii*) *The double sequence* $(\tau_i, q_i), i \in I\!N_+$ *defines a marked point process* [11] *with defining measure*

$$\mu(\omega, t; dq, dt) \tag{3.59}$$

and predictable compensator

$$\nu(\omega, t)(dq, dt) = K(\omega, t)(dq)\ \lambda(t)\ dt \tag{3.60}$$

(*iv*) *Consider the defaultable zero coupon bond* $\overline{B}(0, T)$. *At time* T, *the maturity of this bond, it pays out*

$$Q(T) := \prod_{\tau_i \leq T} (1 - q_i), \tag{3.61}$$

the remainders after all fractional default losses. [12] $Q(t)$ *can be represented as a Doleans-Dade exponential:* $Q(0) = 1$ *and*

$$\frac{dQ(t)}{Q(t-)} = -\int_0^1 q\,\mu(dq, dt). \tag{3.62}$$

(*v*) *The filtration* $(\mathcal{F}_t)_{(t \geq 0)}$ *is generated by the Brownian motions* W^i *and the marked point process* μ.

(*vi*) *We assume sufficient regularity on the marked point process* μ *to justify all subsequent manipulations.*
– *The sequence of default times is nonexplosive.*
– μ *is a multivariate point process (see [1988]).*
– $\int_0^t \int_E K(dq)\lambda(s)\,ds < \infty$ *for all* $t < \infty$.
– *The processes introduced in definition 6 are square-integrable.*
– *The resulting bond prices are bounded.*

For the subsequent analysis we need to define the following processes:

Definition 6
The loss summation function $N'(t)$, *the instantaneous expected loss rate* $q(t)$, *the default compensator* $A'(t)$ *and the default martingale* $M'(t)$ *are defined as:*

$$N'(t) := \int_0^t \int_0^1 q\,\mu(dq, ds) \tag{3.63}$$

$$q(t) := \int_0^1 q\,K(dq) \tag{3.64}$$

$$A'(t) := \int_0^t \int_0^1 q\,K(dq)\lambda(s)ds = \int_0^t q(s)\lambda(s)\,ds \tag{3.65}$$

$$M'(t) := N'(t) - A'(t). \tag{3.66}$$

[11] For a general reference on marked point processes see Jacod and Shiryaev [1988] and Bremaud [1981].

[12] It will be clear from the context whether Q denotes the martingale measure or the accumulated fractional default losses. The latter will usually be the case from now on.

Note that A' is the predictable compensator of N', and M' is a martingale, and $dQ(t)/Q(t-) = -dN'(t)$.

Assumption 5
In the presence of multiple defaults:

(i) *The dynamics of the defaultable rates are given by assumption 1.*
(ii) *The dynamics of the defaultable bond prices are as in assumption 1, with N' replacing N:*

$$\frac{d\overline{B}(t,T)}{\overline{B}(t-,T)} = \overline{\mu}(t,T)dt + \sum_{i=1}^{n} \overline{\eta}_i(t,T)dW^i(t) - dN'.$$

(iii) *The defaultable bank account is $c(t) = Q(t)\exp\{\int_0^t \overline{r}(s)\,ds\}$.*
(iv) *The dynamics of the default risk free rates and bond prices are given by assumption 2.*
(v) *There is no total loss: $q_i < 1$ a.s.*

Assumption 5 implies that all forward rates and the process $G(t,T) = \int_t^T \overline{f}(t,s)\,ds$ are continuous at times of default.

For the default-free rates this is justifiable in most cases (unless a large sovereign debtor is concerned), but the defaultable rates should be modelled by explicitly allowing for dependence on the defaults μ. A default will usually discontinuously change the market's estimation of the future likelihood of defaults and thus the defaultable forward rates. This effect will be included in a later section, here we assume that the only direct effect of a default is the reduction of the face value of the defaultable debt. Nevertheless we allow the default to influence the diffusion parameters of the forward rates which distinguishes this setup from the literature on Cox processes (see Lando [1998]).

3.6.2 Change of Measure

First, Girsanov's theorem (theorem 3) takes the following form: [13]

Theorem 8
Let θ be a n-dimensional predictable processes $\theta_1(t),\dots,\theta_n(t)$ and $\Phi(t,q)$ a strictly positive predictable function [14] with

$$\int_0^t \|\theta(s)\|^2 ds < \infty, \qquad \int_0^t\int_E |\Phi(s,q)|\, K(dq)\,\lambda(s)ds < \infty$$

for finite t. Define the process L by $L(0) = 1$ and

$$\frac{dL(t)}{L(t-)} = \sum_{i=1}^{n} \theta_i(t)dW^i(t) + \int_E (\Phi(t,q) - 1)(\mu(dq,dt) - \nu(dq,dt)).$$

[13] See Jacod and Shiryaev [1988] and Björk, Kabanov and Runggaldier [1996].

[14] In functions of the marker q (like Φ here) *predictability* means measurable with respect to the σ-algebra $\tilde{\mathcal{P}} := \mathcal{P}\otimes\mathcal{E}$. Here $\mathcal{P}$ is the σ-algebra of the predictable processes. See Jacod and Shiryaev [1988] for details.

Assume that $\mathrm{E}[\, L(t)\,] = 1$ *for finite* t.
Then there is a probability measure Q *equivalent to* P *with*

$$dQ_t = L_t dP_t \tag{3.67}$$

such that

$$dW(t) - \theta(t)dt = d\widetilde{W}(t) \tag{3.68}$$

defines $\widetilde{W}$ *as* Q*-Brownian motion and*

$$\nu_Q(dq, dt) = \Phi(t, q)\nu(dq, dt) \tag{3.69}$$

is the predictable compensator of μ *under* Q.
Every probability measure that is equivalent to P *can be represented in the way given above.*

Proof: See BKR [1996].

□

The only change to theorem 3 is the new predictable compensator of the marked point process which has the form $\Phi(t, q)\nu(dq, dt)$. Instead of a single market price of risk for the jump risk we now have a market price of risk for each subset $e \in \mathcal{E}$ of the marker space. The market price of risk of a default with loss $q \in e$ is then $\int_e \Phi(q, t)K(dq) / \int_e K(dq)$ per unit of probability.

Note that now we have a much larger class of potential martingale measures, as for every (t, ω) a *function* $\Phi(t, q)$ has to be chosen and not just the value of the process $\phi(t)$. Typically we will have incomplete markets in this situation which poses entirely new problems for the hedging of contingent claims. See Björk, Kabanov and Runggaldier [1996] for a detailed analysis of trading strategies, hedging and completeness in bond markets with marked point processes.

As before, to save notation, we will assume that all dynamics are already specified with respect to the martingale measure.

3.6.3 Dynamics and Absence of Arbitrage

We start from the representation of the defaultable bond prices as (using the notation and results of the preceding sections)

$$\overline{B}(t, T) = \exp\{-G(t, T)\}\, Q(t).$$

The dynamics of $\overline{B}(t, T)$ are then

$$\begin{aligned} \frac{d\overline{B}(t, T)}{\overline{B}(t-, T)} &= -dG(t, T) + \frac{1}{2}d < G, G > -dN' \\ &= \left[-\overline{\gamma}(t, T) + \overline{r}(t) + \frac{1}{2}\sum_{i=1}^{n} \overline{a}_i^2(t, T) \right] dt \\ &\quad + \sum_{i=1}^{n} \overline{a}_i(t, T) dW^i(t) - dN'(t) \end{aligned}$$

$$= \left[-\overline{\gamma}(t,T) + \overline{r}(t) - q(t)\lambda(t) + \frac{1}{2}\sum_{i=1}^{n} \overline{a}_i^2(t,T) \right] dt$$

$$+ \sum_{i=1}^{n} \overline{a}_i(t,T) dW^i(t) - dM'(t),$$

where we used that G is continuous. Absence of arbitrage is here equivalent to

$$r(t) = -\overline{\gamma}(t,T) + \overline{r}(t) - q(t)\lambda(t) + \frac{1}{2}\sum_{i=1}^{n} \overline{a}_i^2(t,T). \tag{3.70}$$

To show that the results of the preceding sections remain valid we only need to show that

$$\lambda(t)q(t) = \overline{r}(t) - r(t). \tag{3.71}$$

The argument goes exactly as before: The Doleans-Dade exponential of the martingale $M'(t) = N'(t) - \int_0^t q(s)\lambda(s)\, ds$ is

$$\exp\{-\int_0^t q(s)\lambda(s)ds\} \prod_{T_i \le t} (1 - q_i), \tag{3.72}$$

while the discounted value of the defaultable bank account is the Q-martingale

$$\overline{c}(t) := \frac{c(t)}{b(t)} = \exp\{\int_0^t \overline{r}(s) - r(s)\, ds\} \prod_{T_i \le t} (1 - q_i). \tag{3.73}$$

This is the Doleans-Dade exponential of

$$\hat{M}(t) := -N'(t) + \int_0^t \overline{r}(s) - r(s)\, ds. \tag{3.74}$$

Because $q_i < 1$ a.s. we have that $\overline{c}(t) > 0$ a.s. and therefore $\hat{M}$ is unique and well-defined as $\hat{M}(t) = \int_0^t \frac{d\overline{c}(s)}{\overline{c}(s-)}$. Again, we see that

$$M'(t) + \hat{M}(t) = \int_0^t \overline{r}(s) - r(s) - q(s)\lambda(s)\, ds \equiv 0 \tag{3.75}$$

(being a predictable martingale with initial value zero) must be constant and equal to zero. Therefore

$$\lambda(t)q(t) = \overline{r}(t) - r(t) > 0. \tag{3.76}$$

Equation (3.76) is the equivalent of equation (3.38), the key relationship which allowed for the derivation of conditions for the absence of arbitrage. These conditions are exactly the same as for zero recovery, the proof is the same as for theorem 4.

Theorem 9
The following are equivalent:

1. *The measure under which the dynamics are specified is a martingale measure.*
2. (i) *The short interest rate spread is the intensity of the default process multiplied with the locally expected loss quota. It is positive (for $q(t) > 0$).*

$$q(t)\lambda(t) = \overline{r}(t) - r(t) > 0. \tag{3.77}$$

   (ii) *The drift coefficients of the defaultable forward rates satisfy for all $t \leq T$ equations (3.41) and (3.42).*

   (iii) *The drift coefficients of the default-free forward rates satisfy for all $t \leq T$ equation (3.43).*
3. (i) $\overline{r}(t) - r(t) = \lambda(t)q(t) > 0$.

   (ii) *The dynamics of the defaultable bond prices are given by*

$$\frac{d\overline{B}(t,T)}{\overline{B}(t-,T)} = \overline{r}(t)dt + \sum_{i=1}^{n} \overline{a}_i(t,T)dW^i(t) - dN'(t) \tag{3.78}$$

   or, solving the s.d.e.

$$\begin{aligned} \overline{B}(t,T) &= \overline{B}(0,T)\,Q(t)\,\exp\left\{\int_0^t \overline{r}(s)\,ds - \frac{1}{2}\sum_{i=1}^{n}\int_0^t \overline{a}_i^2(s,T)\,ds \right. \\ &\left. + \sum_{i=1}^{n}\int_0^t \overline{a}_i(s,T)\,dW^i(s)\right\}. \end{aligned} \tag{3.79}$$

   (iii) *The dynamics of the default-free bonds satisfy (3.46).*

All no-arbitrage restrictions on the dynamics of the interest rates are exactly identical to the restrictions in theorem 4, although theorem 4 only concerned the situation with zero recovery. Thus theorem 9 allows us to directly transfer all results on the modelling of the spreads. The drift restrictions of the corollaries 6 and 7 and of theorem 5 are also valid in the present setup.

For the modelling of arbitrage-free dynamics of the defaultable interest rates $\overline{f}(t,T)$ one need not be concerned with the specification of the recovery rates, it is sufficient to just model the interest rates subject to the positive spread restriction (3.40) or (3.77) and the drift restrictions (3.42) and (3.43). Again we see that the hard task of modelling an unobservable quantity (like the distribution of the loss quota q) can be replaced with a suitable model of the defaultable forward rates which are much more easily observed.

If one allows q to take on negative values, negative spot spreads $\overline{r} - r$ are possible. A negative q means that the defaultable bond gains in value upon default. Of course such an event is very rare but in some cases there might be an early (and full) repayment of the debt which will result in $q < 0$, for instance if the proceeds of a liquidation are greater than the outstanding debt or if the default event is caused by a strategic default or a takeover. The advantage of negative q is to allow a wider class of models to be used for $q\lambda$, e.g. particularly the Gaussian models like the models of of Vasicek [1977] and Ho and Lee [1986].

3.6.4 Seniority

Bonds of different seniority have different payoffs in default, the ones with higher seniority have a higher payoff than the ones with lower seniority. Strict seniority – junior debt has a positive payoff if and only if senior debt has full payoff – is rather rare in practice which is due to the various legal bankruptcy procedures, but in general senior debt has a higher payoff in default than junior debt.

With a loss quota of junior debt q^j that is higher than the loss of senior debt q^s, the defaultable instantaneous short rate of junior debt is greater than the short rate for senior debt: $\overline{r}^j > \overline{r}^s$.

For modelling junior and senior debt another stage is added to the usual defaultable debt modelling. First model the default-free term-structure. Then model the spread to the senior bonds. Then (this is the new step) model the spread between junior and senior debt using the senior debt as 'default-free' debt in the drift restrictions. The modelling restrictions we derived above still hold in this setup.

3.7 Instantaneous Short Rate Modelling

Going back to the representation (3.79) of the dynamics of the defaultable bond prices

$$\begin{aligned}\overline{B}(t,T) &= \overline{B}(0,T)\cdot\prod_{\tau_i\le t}(1-q_i)\exp\left\{\int_0^t \overline{r}(s)\,ds\right\}\\ &\quad\cdot\exp\left\{-\frac{1}{2}\sum_{i=1}^n\int_0^t \overline{a}_i^2(s,T)\,ds+\sum_{i=1}^n\int_0^t \overline{a}_i(s,T)\,dW^i(s)\right\},\end{aligned}$$

one can evaluate this expression at $t = T$ and use the final condition $\overline{B}(T,T) = \prod_{\tau_i\le T}(1-q_i)$ to reach

$$\begin{aligned}&\prod_{\tau_i\le T}(1-q_i)\exp\left\{-\int_0^T \overline{r}(s)\,ds\right\}\\ =\;&\overline{B}(0,T)\exp\left\{-\frac{1}{2}\sum_{i=1}^n\int_0^T \overline{a}_i^2(s,T)\,ds+\sum_{i=1}^n\int_0^T \overline{a}_i(s,T)\,dW^i(s)\right\}\\ &\cdot\prod_{\tau_i\le T}(1-q_i).\end{aligned}\tag{3.80}$$

If $q < 1$ a.s. we may divide both sides by $\prod_{\tau_i\le T}(1-q_i)$ and take expectations of both sides to reach

Corollary 10
If there is no total loss on the defaultable bond (i.e. $q_i < 1$), we have the following representation of the price of defaultable zero coupon bonds:

$$\overline{B}(0,T) = \mathrm{E}\left[\exp\left\{-\int_0^T \overline{r}(s)\,ds\right\}\mid\mathcal{F}_0\right].\tag{3.81}$$

Here we used that the second exponential is a stochastic exponential of the martingale $\sum_i \int \bar{a}_i(s,T)dW^i(s)$. Thus it is again a martingale with initial value 1.

In the first sections with zero recovery we were not able to derive this representation as $\bar{r}$ was not defined for times after the default. The above representation of the prices of defaultable bonds is the exact analogue to the representation of the prices of default-free bonds $B(t,T) = \mathrm{E}\left[\exp\{-\int_t^T r(s)ds\}\right]$ as discounted expected value of the final payoff 1. This representation is the starting point of all models of the term-structure of interest rates that are based on a model of the short rate[15]. A result of this type has been proved by Duffie, Schroder and Skiadas [1994] (but in their valuation formula an additional jump term occurs) and by Lando [1998] for the case of a default that is triggered by the first jump of a Cox process. Here the Cox process assumption is not needed, the default process can have an intensity that conditions on previous defaults.

Alternatively to the modeling of defaultable interest rates in the HJM- framework of assumption 1 one can model the short rates directly. With any arbitrage-free short rate model for the default-free short rate r and a positive short rate model for the spread h one can specify an arbitrage-free model framework. Because the model for the defaultable short rate will necessarily be at least a two-factor model, the calibration of this model might become difficult and the HJM approach may be preferable. On the other hand the short rate models need not worry about possibly negative forward spreads and are better suited for analysis with partial differential equations.

3.8 Jumps in the Defaultable Rates

In the presence of multiple defaults (with ensuing restructuring) it is more realistic to allow the defaultable rates to change discontinuously at times of default. the defaultable term structure must be allowed to change its *shape* at these events.

These jumps in the defaultable rates are not to be confused with the fractional loss at default. There are two distinct effects at a time of default which both cause a discrete change in the value process of the holders of defaultable bonds:
First, there is the direct loss caused by the rescue plan and the reduction of the claims of the defaulted bondholders. This is modeled by the marker process q.
Second, the market's valuation of the defaultable bonds may change due to the default. This is reflected in a discrete change in the yield curve but need not mean a irrecoverable loss. If there is no further default until maturity this jump in the value process is compensated[16].

[15]Popular models of the short rate are by (among others): Vasicek [1977], Cox, Ingersoll and Ross [1985], Ho and Lee [1986], Black, Derman, Toy [1990], Hull and White [1993] and Sandmann and Sondermann [1997].

[16]The general methodology of modelling interest rates in the presence of marked point processes is taken from Björk, Kabanov and Runggaldier [1996], who give an excellent account on default-free interest rate modelling with marked point processes.

3.8.1 Dynamics

To reach the most general setup we use the marked point process $\mu(dq, dt)$. At every default there are jumps in the defaultable term structure $\overline{f}(t, T)$ and the defaultable bond prices $\overline{B}(t, T)$.

We replace assumption 1 with:

Assumption 6

(i) The dynamics of the defaultable forward rates are:

$$d\overline{f}(t,T) = \overline{\alpha}(t,T)dt + \sum_{i=1}^{n} \overline{\sigma}_i(t,T)\, dW^i(t) + \int_E \delta(q;t,T)\, \mu(dq, dt) \tag{3.82}$$

(ii) The dynamics of $\widetilde{B}(t,T) := \exp\{-\int_t^T \overline{f}(t,s)ds\}$ are:

$$\begin{aligned} \frac{d\widetilde{B}(t,T)}{\widetilde{B}(t-,T)} &= m(t,T)\, dt + \sum_{i=1}^{n} \overline{a}_i(t,T)\, dW^i(t) \\ &\quad + \int_E \theta(q;t,T)\, \mu(dq, dt) \end{aligned} \tag{3.83}$$

(iii) The dynamics of the defaultable bond prices $\overline{B}(t,T) = Q(t)\widetilde{B}(t,T)$ are:

$$\begin{aligned} \frac{d\overline{B}(t,T)}{\overline{B}(t-,T)} &= m(t,T)\, dt + \sum_{i=1}^{n} \overline{a}_i(t,T)\, dW^i(t) \\ &\quad + \int_E (1-q)\theta(q;t,T)\, \mu(dq, dt) - \int_E q\, \mu(dq, dt) \end{aligned} \tag{3.84}$$

(iv) We assume sufficient regularity on the parameters to allow:
– differentiation (w.r.t. T) under the integral,
– interchange of order of integration,
– finite prices $\overline{B}(t,T)$ almost surely.

We distinguish between a 'pseudo' bond price $\widetilde{B}(t,T)$ in which the influence of previous defaults has been removed, and the 'real' bond price $\overline{B}(t,T) = Q(t)\widetilde{B}(t,T)$. The dynamics of $\overline{B}(t,T)$ in (3.84) follow directly from Itô's lemma.

These dynamics are interdependent due to the following result by BKR [1996]:

Proposition 11
Given the dynamics (3.82) of $\overline{f}(t,T)$
(i) the dynamics of the defaultable short rate $\overline{r}(t)$ are

$$\begin{aligned} d\overline{r}(t) &= \Big[\frac{\partial}{\partial T}\Big|_{T=t} \overline{f}(t,T) + \overline{\alpha}(t,t)\Big]dt + \sum_{i=1}^{n} \overline{\sigma}_i(t,t)dW^i(t) \\ &\quad + \int_E \delta(q;t,t)\mu(dq, dt). \end{aligned} \tag{3.85}$$

(ii) the dynamics of $\widetilde{B}(t,T)$ are

$$\begin{aligned} m(t,T) &= \overline{r}(t) - \int_t^T \overline{\alpha}(s,T)\,ds + \frac{1}{2}\sum_{i=1}^{n}\Big(\int_t^T \overline{\sigma}_i(s,T)\,ds\Big)^2 &(3.86)\\ \overline{a}_i(t,T) &= -\int_t^T \overline{\sigma}_i(s,T)\,ds &(3.87)\\ \theta(q;t,T) &= \exp\left\{-\int_t^T \delta(q;s,T)\,ds\right\} - 1. &(3.88) \end{aligned}$$

(iii) The dynamics of the defaultable bond prices are given by assumption 6 (iii) with the specification of (ii) above.

Proof: See BKR [1996] for (i) and (ii), point (iii) follows directly.

□

3.8.2 Absence of Arbitrage

The change of measure to the martingale measure is done according to theorem 8. The analysis leading to the key relation (3.77)

$$\lambda(t)q(t) = \overline{r}(t) - r(t) > 0$$

in section 3.6.3 is still valid, because the only defaultable security needed there is the defaultable bank account $c(t)$ which has no jump component in its development except the direct losses of q_i at default.

As usual we need for absence of arbitrage

$$\begin{aligned} r(t)\,dt &= \mathbf{E}\left[\frac{d\overline{B}(t,T)}{\overline{B}(t-,T)}\right]\\ &= \overline{r}(t)dt - \overline{\gamma}(t,T)\,dt + \frac{1}{2}\sum_{i=1}^{n}\overline{a}_i^2(t,T)\,dt\\ &\quad + \int_E \theta(q;t,T)(1-q)\,K(dq)\lambda(t)\,dt - q(t)\lambda(t)dt. \end{aligned}$$

Substituting (3.77) and (3.88) yields:

Proposition 12
Under the martingale measure

(i) The short rate spread is given by

$$\overline{r}(t) - r(t) = \lambda(t)q(t). \tag{3.89}$$

(ii) The drift of the defaultable forward rates is restricted by

$$\overline{\gamma}(t,T) = \frac{1}{2}\sum_{i=1}^{n}\overline{a}_i^2(t,T) + \int_E \left(\exp\left\{ -\int_t^T \delta(q;t,v)\,dv \right\} - 1 \right)(1-q)\,K(dq)\lambda(t). \tag{3.90}$$

or, differentiated,

$$\begin{aligned}\overline{\alpha}(t,T) &= \sum_{i=1}^{n}\overline{\sigma}_i(t,T)\int_t^T \overline{\sigma}_i(t,v)\,dv \\ &- \int_E \delta(q;t,T)\exp\left\{ -\int_t^T \delta(q;t,v)\,dv \right\}(1-q)\,K(dq)\lambda(t).\end{aligned} \tag{3.91}$$

(iii) The dynamics of the defaultable bond prices under the martingale measure are

$$\begin{aligned}\frac{d\overline{B}(t,T)}{\overline{B}(t-,T)} &= \overline{r}(t,T)\,dt + \sum_{i=1}^{n}\overline{a}_i(t,T)\,dW^i(t) \\ &+ \int_E (1-q)\theta(q;t,T)\left(\mu(dq,dt) - K(dq)\,\lambda(t)\,dt\right) \\ &- dN'(t)\end{aligned} \tag{3.92}$$

where $\theta(q;t,T)$ is defined as in (3.88).

Obviously the drift restriction (3.90) cannot be handled as easily as the other restrictions in theorems 4 and 9 before because of the integral over the jumps of the forward rates. As the defaults now have a jump-influence on the defaultable forward rates, the parameters of the default process do not disappear any more.

BKR [1996] and Jarrow and Madan [1995] reach a quite similar restriction to (3.91) for the modelling of default-free interest rates in the presence of marked point processes. The restriction of BKR is

$$\begin{aligned}\alpha(t,T) &= \sum_{i=1}^{n}\sigma_i(t,T)\int_t^T \sigma_i(t,v)\,dv \\ &- \int_E \delta'(q;t,T)\exp\left\{ -\int_t^T \delta'(q;t,v)\,dv \right\} K(dq)\lambda(t),\end{aligned}$$

and applies to the default-free interest rates. In this setup we assumed that the default-free interest rates do not jump (i.e. $\delta' = 0$) which reduces the restriction to the usual HJM-restriction (3.43).

The s.d.e. of the defaultable bond prices is of the usual type: There is a drift component of $\overline{r}$ and the default influence $-dN'(t)$. The other parts of the dynamics of the defaultable bond prices are local martingales.

3.9 Ratings Transitions

In the previous section we only allowed jumps in the term structure of defaultable bond prices at defaults. But in an environment where changes in credit ratings can cause large, sudden changes in the prices of defaultable bonds, one would like to be able to incorporate jumps in the term structure of defaultable bond prices that are not caused by defaults but only by rating transitions. This aim is easy to accomplish in the framework that has been laid in the previous sections.

First we have to define defaultable bond prices and forward rates for all rating classes:

Definition 7
There are K different rating classes for defaultable bonds. ('Default' is not counted as a rating class.) For each rating class $k = 1, \dots, K$ there is a term structure of defaultable (pseudo) bond prices $\widetilde{B}_k(t,T)$ and defaultable forward rates $\overline{f}_k(t,T)$

$$\widetilde{B}_k(t,T) = \exp\{-\int_t^T \overline{f}_k(t,s)ds\}. \tag{3.93}$$

The defaultable (pseudo) bond prices $\widetilde{B}_k(t,T)$ and forward rates $\overline{f}_k(t,T)$ for all classes can be observed at all times.
To every defaultable bond $\overline{B}(t,T)$ there is a ratings process $R(t)$ which gives the rating of the bond at time t. Thus its price is given by

$$\overline{B}(t,T) = Q(t)\widetilde{B}_{R(t)}(t,T).$$

Note that there is a difference between the price of *one given* defaultable bond $\overline{B}(t,T)$ and the respective 'pseudo' bond price $\widetilde{B}_k(t,T)$ of the ratings class $k = R(t)$ in which the defaultable bond happens to be at this time. First, the pseudo-bond price never changes its rating (it is only used to define the bond prices in this rating class), and second, the pseudo bond price is not affected by defaults.

Next we describe the rating transition dynamics of the defaultable bond prices and the dynamics of the pseudo defaultable bond prices for all rating classes. Note that in contrast to most of the other literature we do not include 'default' as $K+1$st rating class, but prefer to model defaults using the multiple default model of the previous section. Assumption 6 is replaced with:

Assumption 7
(i) *The rating transitions of a defaultable bond are driven by a marked point process $\mu_R(l, dt)$. The marker space of this process is $E = \{1, \dots K\}$ the set of possible rating class transitions. We write the predictable compensator of μ_R with a transition matrix $A = (a_{kl})_{1\le k,l\le K}$ as*

$$\nu_R(l, dt) = \sum_{k=1}^{K} \mathbf{1}_{\{R(t)=k\}} a_{kl} dt = a_{R(t),k} dt.$$

The elements a_{kl} of A can be predictable stochastic processes.

(ii) *The dynamics of the defaultable forward rates within class k are:*

$$d\overline{f}_k(t,T) = \overline{\alpha}_k(t,T)dt + \sum_{i=1}^{n} \overline{\sigma}_{i,k}(t,T)\, dW^i(t) \tag{3.94}$$

(iii) *The dynamics of a defaultable pseudo bond price $\widetilde{B}_k(t,T) := \exp\{-\int_t^T \overline{f}_k(t,s)ds\}$ in rating class k are:*

$$\frac{d\widetilde{B}_k(t,T)}{\widetilde{B}_k(t-,T)} = m_k(t,T)\,dt + \sum_{i=1}^{n} \overline{a}_{i,k}(t,T)\, dW^i(t) \tag{3.95}$$

(iv) *Defaults of the defaultable bond are driven by the multiple default model as in assumption 4, definition 6 and assumption 5. Thus the dynamics of the defaultable bond prices $\overline{B}(t,T) = Q(t)\widetilde{B}_{R(t)}(t,T)$ are:*

$$\begin{aligned}\frac{d\overline{B}(t,T)}{\overline{B}(t-,T)} =& m(t,T)\,dt + \sum_{i=1}^{n} \overline{a}_i(t,T)\, dW^i(t)\\ & - q(dN(t) - \lambda(t)dt) + \sum_{k\neq R(t)} \left(\frac{\widetilde{B}_k(t,T)}{\widetilde{B}_{R(t)}(t,T)} - 1\right) \left(\mu_R(k;dt) - a_{R(t),k}dt\right).\end{aligned} \tag{3.96}$$

(v) *Defaults and rating transitions never happen at the same time:*

$$\int_0^t dN_s\mu_R(k,ds) \equiv 0 \qquad \forall k \in \{1,\dots,K\},\ t \geq 0.$$

(vi) *We assume sufficient regularity on the parameters to allow:*
- *differentiation (w.r.t. T) under the integral,*
- *interchange of order of integration,*
- *finite prices $\overline{B}(t,T)$, $\widetilde{B}_k(t,T)$ almost surely.*

The ratings transition marked point process $\mu_R(k,dt)$ places Dirac-delta weights of 1 in the space $\{1,\dots K\} \times I\!R_+$ at points (R^*,t^*), where t^* is the time of a rating transition and R^* is the new rating class.

The compensator matrix A of μ_R contains the intensities of the different rating transitions: a_{kl} is the intensity of the process triggering a transition from class k to class l. A can be stochastic. The matrix A is called the *generator matrix* of the rating transitions. In the theory of Markov chains it is customary to set $a_{kk} := -\sum_{l\neq k} a_{kl}$, here this would not be consistent with the definition of A as predictable compensator of μ_R.

To make the transition intensities a_{kl} stochastic *and* dependent on the original rating class k is redundant, but it makes the simplification to the important case of constant a_{kl} trivial. For a constant rating transition matrix A the ratings transitions form a time-homogeneous Markov chain.

At a rating transition the defaultable bond price jumps to the equivalent (same maturity and face value) defaultable bond price of the new rating class. Similarly, a default leads to a reduction in face value of the defaultable bond by q of the previous face value. The recovery mechanism remains unchanged. Therefore, in equation (3.96) the predictable compensators have already been subtracted from the respective point processes to ensure that the resulting stochastic integrals are local martingales and that $m(t,T)$ contains the full (predictable) drift coefficients. Excluding joint rating transitions and defaults is done mainly to simplify the analysis. This assumption can be relaxed along the lines of the previous section, and it may be desirable to be able to place the bond in a different rating class after a default and reorganisation.

The following proposition gives the connection between the defaultable bond price dynamics, the defaultable pseudo bond price dynamics and the defaultable forward rates for each rating class.

Proposition 13
For every given rating class $k = 1,\dots,K$ and the dynamics (3.94) of $\overline{f}_k(t,T)$

(i) the dynamics of the defaultable short rate $\overline{r}_k(t) = \overline{f}_k(t,t)$ in rating class k are

$$d\overline{r}_k(t) = [\frac{\partial}{\partial T}\overline{f}_k(t,t) + \overline{\alpha}_k(t,t)]dt + \sum_{i=1}^{n}\overline{\sigma}_{i,k}(t,t)dW^i(t). \tag{3.97}$$

(ii) the dynamics of $\widetilde{B}_k(t,T)$ are

$$m_k(t,T) = \overline{r}_k(t) - \int_t^T \overline{\alpha}_k(s,T)\,ds + \frac{1}{2}\sum_{i=1}^{n}(\int_t^T \overline{\sigma}_{i,k}(s,T)\,ds)^2 \tag{3.98}$$

$$\overline{a}_{i,k}(t,T) = -\int_t^T \overline{\sigma}_{i,k}(s,T)\,ds. \tag{3.99}$$

(iii) The dynamics of the defaultable bond prices $\overline{B}(t,T) = \widetilde{B}_{R(t)}(t,T)Q(t)$ and the dynamics of the rating $R(t)$ of this bond are given by assumption 7 iv with the following specification:

$$m(t,T) = m_{R(t)}(t,T) - q(t)\lambda(t) + \sum_{k\neq R(t)}\left(\frac{\widetilde{B}_k(t,T)}{\widetilde{B}_{R(t)}(t,T)} - 1\right) a_{R(t),k}. \tag{3.100}$$

$$\overline{a}_i(t,T) = \overline{a}_{i,R(t)}(t,T) \tag{3.101}$$

$$dR(t) = \sum_{k=1}^{K}(R(t) - k)\mu_R(k,dt). \tag{3.102}$$

Proof: Within a given rating class $k = 1,\dots,K$ the specification is a standard HJM term structure model of interest rates. Therefore (i) and (ii) follow directly from HJM [1992]. Point (iii) follows directly from substituting the respective dynamics into $\overline{B}(t,T) = Q(t)\widetilde{B}_{R(t)}(t,T)$ and using Itô's lemma on point processes.

□

The change of measure to the martingale measure is done according to theorem 8. The analysis leading to the key relation (3.77)

$$\lambda(t)q(t) = \overline{r}(t) - r(t) > 0$$

in section 3.6.3 is still valid, but has to be done for each rating class individually.

Consider the rating class k and the investment in the defaultable bank account $c_k(t)$ of this rating class. As the investment in a defaultable bank account always matures at the next instant (and is then rolled over), it is not affected by rating transitions: If our short-term investment had a rating change to class k' (but no default), then it would still have the full payoff in the next instant. This payoff is then taken and invested in a short-term investment in the original rating class k. Thus it is possible to continuously roll over the defaultable bank account $c_k(t)$ in rating class k. It has no jump component in its development except the direct losses of q at default. Therefore (using exactly the same argument as in theorem 9) we have

$$\lambda_k(t)q(t) = \overline{r}_k(t) - r(t) > 0, \tag{3.103}$$

the default intensity $\lambda_k(t)$ in rating class k times the loss fraction q equals the short spread in this rating class. The default intensity of the defaultable bond is thus dependent on the current rating class $R(t)$ of the bond.

Equation (3.103) gives a second reason why we did not specify 'default' as $K+1$st rating class in definition 7: Firstly the transition intensity to 'default' is already specified by the term structure of the defaultable bond prices in this class, and secondly, in the fractional recovery model defaulted bonds are reorganised and survive, thus the classical approach (see e.g. Lando [1994, 1998]) of making default an absorbing state is not suitable here.

As usual we need for absence of arbitrage

$$\begin{aligned} r(t)\,dt &= \mathrm{E}\left[\frac{d\overline{B}(t,T)}{\overline{B}(t-,T)} \right] \\ r(t) &= m(t,T). \end{aligned}$$

Substituting (3.103) and the dynamics from proposition 13 yields now the following proposition which gives the drift restrictions that are to be imposed on the martingale-measure dynamics of the defaultable forward rates of each rating class to ensure absence of arbitrage of the full model.

Proposition 14
Under the martingale measure

(i) *The short rate spread in rating class k ($k \in \{1, \ldots K\}$) is given by*

$$\lambda_k(t)q(t) = \overline{r}_k(t) - r(t). \tag{3.104}$$

(ii) *The drift of the defaultable forward rates is restricted by*

$$\int_t^T \overline{\alpha}_k(t,s)ds = \frac{1}{2}\sum_{i=1}^{n}(\int_t^T \overline{\sigma}_{i,k}(t,s)\,ds)^2$$

$$+\sum_{l=1}^{K}\left(\frac{\widetilde{B}_l(t,T)}{\widetilde{B}_k(t,T)}-1\right)a_{k,l} \tag{3.105}$$

or differentiated:

$$\overline{\alpha}_k(t,T)=\sum_{i=1}^{n}\overline{\sigma}_{i,k}(t,T)(\int_t^T \overline{\sigma}_{i,k}(t,s)\,ds)$$
$$+\sum_{l=1}^{K}\frac{\widetilde{B}_l(t,T)}{\widetilde{B}_k(t,T)}(\overline{f}_k(t,T)-\overline{f}_l(t,T))\,a_{k,l}. \tag{3.106}$$

If the defaultable forward rates satisfy (3.105) *or* (3.106), *discounted defaultable bond prices are local martingales.*

This result holds without assuming Markov chain dynamics for the rating transition process, and it can easily be generalised to stochastic recovery rates and joint jumps of ratings and defaults. The drift restrictions (3.105) and (3.106) are versions of the drift restriction found by BKR in the context of default-free term structures of interest rates. To close the model, the default-free forward rates have to satisfy the classical HJM-drift restrictions, too.

Using proposition 14 it is now possible to set up a simulation model of defaultable bond price dynamics for a full portfolio of defaultable bonds of different rating classes and maturities. Correlations in the movements of the spreads and yields in all classes and in the dynamics of the default-free term structure of interest rates can be captured directly via the forward rate volatility functions $\overline{\sigma}_k(t,T)$.

The only remaining problem is the specification of the rating transition intensities a_{kl} under the *martingale measure*. There is a wealth of published data concerning *historical* rating transitions but there seems to be a large risk-premium in the market attached to downgrades below investment-grade and to defaults. This problem is left for further research.

3.10 Conclusion

In this chapter we presented a new approach to the modelling of the price processes of defaultable bonds that was inspired by the Heath-Jarrow-Morton [1992] model of the term structure of interest rates. This model avoids a precise specification of the mechanism that leads to default but rather gives necessary and sufficient conditions on the term structure of defaultable interest rates to ensure absence of arbitrage.

These restrictions show a striking similarity to the restrictions that are already well known from default-free term structure models. Specifically, the defaultable interest rates have to satisfy a drift restriction that is analogous to the HJM drift restriction, and a positive short spread restriction. Given these restrictions the model is arbitrage-free. In the implementation of the model one can therefore use the extensive machinery of default-free interest rate modelling.

As an alternative to the HJM-modelling approach it is shown that sufficient for the absence of arbitrage in a short rate model is a positive spread between the defaultable and the default-free short rate. Furthermore it is discussed how a defaultable bond model can be added to an existing model of default-free bonds while keeping the combined model arbitrage-free, and the key result of the direct intensity approach was recovered, viz the representation of the defaultable bond prices as the expectation of the defaultable discount factor.

In addition to this, a new *multiple default model* was introduced in this chapter in which defaulted firms are reorganised and their bonds continue to trade after a default. It was shown that this modelling approach leads to a very intuitive representation of the spread of the defaultable short rate over the default-free short rate as the product of the loss quota in default and the default intensity.

Finally, in the last two sections the approach was extended in two directions: First, jumps were allowed in the term structure of defaultable interest rates at times of default, and second, the model was extended to a full rating transition model. The rating transition model extended the existing literature in several ways: It is the first default model to allow: fully independent stochastic dynamics of the credit spreads in all rating classes, a perfect fit to the term structures of credit spreads in all rating classes, and stochastic rating transition matrices.

Chapter 4

Pricing Credit Risk derivatives

4.0.1 In this section you will learn ...

- ... what points to consider when specifying a credit risk derivatives,
- ... the different possibilities in the treatment of default events,
- ... how to price some typical credit derivatives.

In this chapter the pricing of several credit risk derivatives is discussed in an intensity-based framework with both default-free and defaultable interest rates stochastic and possibly correlated. The differences to standard interest rate derivatives are analysed. Where possible closed-form solutions are given.

4.1 Introduction

Credit derivatives are derivative securities whose payoff depends on the credit quality of a certain issuer. This credit quality can be 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. In this chapter we will concentrate on the latter case.

Credit risk derivatives can make large and important risks tradeable. They form an important step towards market completion and efficient risk allocation, they can help bridge the traditional market segmentation between corporate loan and bond markets.

Despite their large potential practical importance, there have been very few works specifically on the pricing and hedging of credit risk *derivatives*. The majority of papers is concerned with the pricing of defaultable *bonds* which is a necessary prerequisite for credit derivatives pricing. In this chapter we are going to use mainly the fractional recovery / multiple defaults model (see Duffie and Singelton [1997, 1999] and also Schönbucher [1996, 1998]) which we will compare to an alternative model of the recovery of defaulted bonds: recovery in terms of otherwise equivalent default-free bonds (here termed *equivalent recovery*, see e.g. Lando [1994, 1998]). This will enable us to analyse the relative advantages and disadvantages of the two modelling approaches.

Apart from this comparison, the second main point of the chapter is to demonstrate the methodology that has to be used in an intensity-based credit spread model to price credit contingent payoffs. This methodology is independent of the actual specification of the model and it is easy to transfer most of the results to other intensity-based models like e.g. the models by Jarrow and Turnbull [1995], Lando [1994, 1998], Madan and Unal [1998] or Artzner and Delbaen [1992]. Closely related to this point is the third aim of the chapter which is to analyse and identify the relative importance of model input parameters like recovery rates or the correlation between interest rate and credit spreads.

In the following section the structure and applications of the most popular credit derivatives will be presented. As credit derivatives are over-the-counter derivatives, no standard specification has evolved yet. Therefore we chose an wide array of common and/or natural specifications to exemplify typical specifications and applications, and the more exotic specifications will not be treated in the section on pricing later on. The discussed securities include *total rate of return swaps*, *default swaps*, straight *put options* on defaultable bonds, *credit spread forwards* and *credit spread puts*.

This is followed by a recapitulation of the fractional recovery and the equivalent recovery models and a short discussion of the relative advantages and disadvantages of both approaches. Using these models we begin the analysis of the pricing of those credit derivatives that can be priced directly off the term structures of interest rates and credit spreads, without needing to explicitly specify the dynamics of these term structures.

In the next section we present the specification of the stochastic processes for interest rates and credit spreads/default intensities. We chose two alternative setups for the credit spread and for the default-free short rate, a multifactor Gaussian model and a multifactor Cox Ingersoll Ross (CIR) [1985] square-root diffusion model.

The Gaussian setup can allow for arbitrary correlation between credit spreads and the default-free term structure of interest rates but has the disadvantage of allowing credit spreads and intensities to become negative. With the CIR setup on the other hand, only positive correlation can be reached but it ensures positive interest rates and credit spreads. Obviously these are not the only possible specifications, another promising specification would be a market model with lognormal LIBOR rates. (See Miltersen, Sandmann and Sondermann [1997] for the default-free case and Lotz and Schloegl [1999] for an analysis of mutual counterparty risk in a market model.)

The next section treats the pricing of the most important of the the credit risk derivatives that were introduced before. To keep the results as flexible as possible we used the specifications only in the last step of the calculation of the prices, and we also give the expectations that have to be calculated if another model is used or a numerical implementation is desired. In the CIR case the pricing formulae are expressed in terms of the multivariate noncentral chi-squared distribution function using results from Jamshidian [1996], in the Gaussian case we reach expressions in terms of the cumulative normal distribution function. We analyse the influence of interest-rate / credit spread correlation and the differences to standard interest rate contingent contracts.

The conclusion sums up the main results of the chapter and points out the consequences that these results have for the further development of credit spread models.

4.2 Credit Derivatives: Structures and Applications

This section contains an overview over the most common credit derivatives. First we have to clarify the common features of most credit derivatives.

Definition 8
A credit derivative *is a derivative security that has a payoff which is conditioned on the occurrence of a* credit event. *The credit event is defined with respect to a* reference credit *(or several reference credits), and the* reference credit asset(s) *issued by the reference credit. If the credit event has occurred, the* default payment *has to be made by one of the counterparties.*
Besides the default payment a credit derivative can have further payoffs that are not default contingent.

Most credit derivatives have a default-insurance feature. In naming the counterparties we will use the convention that counterparty **A** will be the insured counterparty (i.e. the counterparty that receives a payoff if a default happens or the party that is long the credit derivative), and counterparty **B** will be the insurer (who has to pay in default). Party C will be the reference credit.

4.2.1 Terminology

One of the attractions of credit derivatives is the large degree of flexibility in their specification. Key terms of most credit derivatives are:

Reference Credit: One (or several) issuer(s) whose defaults trigger the credit event.
This can be one (typical) or several (a basket structure) defaultable issuers.

Reference Credit Asset: A set of assets issued by the reference credit. They are needed for the determination of the credit event and for the calculation of the recovery rate (which is used to calculate the default payment).
The definition can range from 'any financial obligation of the reference credit' to a specific list of just a few bonds issued by the reference credit. Loans and liquidly traded bonds of the reference credit are a common choice. Frequently, different assets are used for the determination of the credit event and the recovery rate[1].

Credit Event: A precisely defined default event, which is usually defined with respect to the reference credit(s) and reference credit assets.
Possible definitions include:

[1] Assume a bank has a large loan exposure to C. To hedge this, a credit derivative could use a missed payment on the *loan* as credit event trigger and the post-default market price of a *bond* issued by C to determine the recovery rate. As the loan is not traded, its recovery rate cannot be determined from market prices.

- payment default (typically a certain materiality threshold must be exceeded)
- bankruptcy or insolvency,
- protection filing,
- ratings downgrade below given threshold (ratings triggered credit derivatives),
- changes in the credit spread
- payment moratorium

but the definition can include events that go as far as 'armed hostilities', 'social unrest' or earthquakes (for souvereigns) or 'merger or takeover' (for corporates).

Default Payment: The payments which have to be made if a credit event has happened. The default payment is the defining feature of most credit derivatives.
This is the defining feature of most credit derivatives, and we will consider possible alternatives when discussing the individual credit derivatives.

Example 1:
Default digital swap on the United States of Brazil:
Counterparty B (the insurer) agrees to pay USD 1 Mio to counterparty A if and when Brazil misses a coupon or principal payment on one of its Eurobonds. Here

- the *reference credit* are the United States of Brazil
- the *reference credit assets* are the Eurobonds issued by Brazil (in the credit derivative contract there would be an explicit list of these bonds),
- the *credit event* is a missed coupon or principal payment on one of the reference assets,
- the *default payment* is USD 1 Mio.

In return for this, counterparty A pays a fee to B.

So far no standard has evolved in the details of the specification of credit derivatives and there are many unresolved problems in this area. Despite the large degree of flexibility it is impossible to cover every contingency in the definition of the credit event, and – apart from digital payoffs – there will also be problems to match the default payment exactly to the exposure that is to be hedged. Even with physical delivery there may be problems if the reference asset is very illiquid or not traded at all (e.g. a loan).

Ignoring the problems in the details of the specification, some credit derivatives have become quasi-standard credit derivative structures. In the following sections we will examine these structures and their applications in further detail.

4.2.2 Asset Swap Packages

An *asset swap package* is a combination of a defaultable fixed coupon bond (the asset) with a fixed-for-floating interest rate swap whose fixed leg is chosen such that the value of the whole package is the par value of the defaultable bond.

Example 2:
The payoffs of the asset swap package are:
B sells to **A** for 1 (the nominal value of the C-bond):

- a fixed coupon bond issued by C with coupon c payable at coupon dates t_i, $i = 1, \dots, N$,
- a fixed for floating swap (as below).

The payments of the swap: At each coupon date t_i, $i \leq N$ of the bond

- **A** pays to **B**: c, the amount of the fixed coupon of the bond,
- **B** pays to **A**: LIBOR + a.

a is called the *asset swap spread* and is adjusted to ensure that the initial asset swap package has indeed the value of 1.

The asset swap is not a credit derivative in the strict sense, because the swap is unaffected by any credit events. Its main purpose is to transform the payoff streams of different defaultable bonds into the same form: *LIBOR + asset swap spread* (given that no default occurs). **A** still bears the full default risk and if a default should happen, the swap would still have to be serviced.

To ensure that the value of the asset swap package (asset swap plus bond) to **A** is at par at time t_0 we require:

$$C_0 + (s_0 + a_0 - c)A_0 = 1 \tag{4.1}$$

where C_0 is the initial price of the bond, s_0 is the fixed-for-floating swap rate for the same maturity and payment dates t_i, and A_0 is the value of an annuity paying 1 at all times t_i, $i = 1, \dots, N$. All these quantities can be readily observed in the market at time t_0. To ensure that the value of the asset-swap package is indeed one, the asset swap rate must be chosen as

$$a_0 = \frac{1}{A_0}(1 - C_0) + c - s_0.$$

Note that the asset swap rate would explode at a default of C, because then $(1 - C_i)$ would change from being very small to a large number. Using the definition of the fixed-for-floating swap rate: $s_0 A_0 = 1 - B(t_0, t_N)$ this can be rearranged to yield:

$$A_0 a_0 = \underbrace{B(t_0, t_N) + cA_0}_{\text{def. free bond}} - \underbrace{C_0}_{\text{defaultable bond}}, \tag{4.2}$$

the asset swap rate a_0 is the price difference between the defaultable bond C_0 and an equivalent default free coupon bond (with the same coupon c, it has the price $B(t_0, t_N) + cA_0$) in the swap-measure numeraire asset A_0.

Asset swap packages are very popular and liquid instruments in the defaultable bonds market, sometimes their market is even more liquid than the market for the underlying defaultable bond alone. They also serve frequently as underlying assets for options on asset swaps, so called *asset swaptions*. An asset swaption gives **A** the right to enter an asset-swap package at some future date T_1 at a pre-determined asset swap spread a.

4.2.3 Total Rate of Return Swaps

In a *total rate of return swap* (or *total-return swap*) A and B agree to *exchange all cash flows* that arise from two different investments, usually one of these two investments is a defaultable investment, and the other is a default-free LIBOR investment. This structure allows an exchange of the assets' payoff profiles without legally transferring ownership in the asset.

Example 3:
Payoffs of a total rate of return swap: Counterparty **A** pays to counterparty **B** at regular intervals:

- the coupon of the bond C (if there was one)
- the price appreciation of bond C since the last payment
- the principal repayment of bond C (at the final payment date)
- the recovery value of the bond (if there was a default)

B pays at the same intervals

- a regular fee of LIBOR $+x$
- the price depreciation of the bond C (if there was any)
- the par value of the bond (if there was a default in the meantime)

These payments are netted.

In the example above the two investments whose payoff streams are exchanged are: (a) an investment of a dollar amount of the face value of the C bond at LIBOR and (b) the investment in the C bond, adjusted by a spread on the LIBOR investment. The reference credit is C, the reference asset is the C bond, credit event is a default on the reference asset and the payoff in default is specified above.

B has almost the same payoff stream as if he had invested in the bond C directly and funded this investment at LIBOR $+x$. The only difference is that in the total rate of return swap is marked to market at regular intervals. Price changes in the bond C become cash flows for the TRORS immediately, while for a direct investment in the bond they would only become cash flows when the bond matures or the position is unwound. This makes the TRORS similar to a futures contract on the C bond, while the direct investment is more similar to the forward. The TRORS is not exactly equivalent to a futures contract because it is marked to market using the *spot* price of the underlying security, and not the *futures* price. This difference can be adjusted using the spread x on the floating payment of **B**.

The reference asset should be liquidly traded to ensure objective market prices for the marking to market. If this is not the case (e.g. for bank loans), the total return swap cannot be marked to market. Then its term must match the term of the underlying loan or it must be terminated by physical delivery.

Total rate of return swaps are among the most popular credit derivatives. They has several advantages to both counterparties:

- Counterparty **B** is long the reference asset without having to fund the investment up front. This allows counterparty **B** to leverage his position much higher than he would otherwise be able to.
- If the reference asset is a loan and **B** is not a bank then this may be the only way in which **B** can invest in the reference asset.
- Counterparty **A** has hedged his exposure to the reference credit if he owns the reference asset (but he still retains some counterparty risk).
- If **A** does not own the reference asset he has a created short position in the asset. (Directly shorting defaultable bonds or loans is often impossible.)
- The transaction can be effected without the consent or the knowledge of the reference credit C. A is still the lender to C and keeps the bank-customer relationship.

4.2.4 Default Swap

In a *default swap* (also known as *credit swap*) **B** agrees to pay the default payment to **A** *if a default has happened.* If there is no default of the reference security until the maturity of the default swap, counterparty **B** pays nothing.

A pays a fee for the default protection. The fee can be either a lump-sum fee up front (default put) or a regular fee at intervals until default or maturity (default swap).

An example of a default swap with a fixed repayment at default is given in example 1 (default digital swap on Brazil). Default swaps mainly differ in the specification of the default payment. Common alternatives are

- notional minus post-default market value[2] of the reference asset (cash settlement),
- physical delivery of one or several of the reference assets against repayment at par,
- a pre-agreed fixed percentage of the notional amount (default digital swap).

Sometimes substitute securities may be delivered, or an exotic payoff may be specified (e.g. to hedge counterparty exposure in derivatives transactions).

The default swap allows the separation of the credit risk component of a defaultable bond from its non-credit driven market risk components. The protection buyer (**A**) retains the market risk but is hedged against the credit risk of C, while the protection seller (**B**) can assume the credit risk alone.

The price of a default swap is closely related to the price of a *defaultable floating rate note (FRN)*. Although the defaultable FRN is not a derivative, its pricing is nevertheless nontrivial because it does not always have to trade at par like default-free FRNs. This relationship will be explained in the section on the pricing of credit derivatives.

[2] The post-default market value could be determined from dealer bid and ask quotes, averaged over a certain period of time and over several dealers and reference assets.

4.2.5 Credit Spread Products

Some credit derivatives have payoffs that condition on the credit spread of the reference credit asset over an equivalent default-free bond. Here the credit event is a change in the credit spread and not necessarily a default. To isolate the reference on credit spreads some credit spread derivatives are even knocked out if a default happens on the reference asset. The reference asset must be a liquidly traded bond to allow a meaningful definition of the credit spread.

Credit spread structures are mainly used for trading. They allow the counterparties to separately trade the credit risk component of a defaultable bond and contribute thus to more efficiency in these markets. A credit spread forward can be used to make a position in the reference asset neutral against credit spread movements or defaults, and a credit spread put can be used to limit the downside risk due to credit spread movements. With a very high strike spread a credit spread put can also be used as a substitute for a default swap which can have advantages for regulatory purposes.

4.2.5.1 Credit Spread Forward and Credit Spread Swap

In a *credit spread forward*, counterparty A pays at time T a pre-agreed fixed payment and receives the credit spread of the reference asset at time T. Conversely, counterparty **B** receives the fee and pays the credit spread. The fixed payment is chosen at time $t < T$ to set the initial value of the credit spread forward to zero.

The credit spread forward can also be structured around the *relative* credit spread between two different defaultable bonds. This case can be decomposed into two credit spread forwards on the *absolute* credit spread of the reference assets. The credit spread forward is similar to a forward rate agreement with the only difference that the credit spread is referenced and not an interest rate. Several credit spread forwards can be combined to a *credit spread swap*.

4.2.5.2 Credit Spread Options

A *credit spread put* gives counterparty A the right to sell the reference asset to counterparty **B** at a pre-specified strike spread over default-free interest rates. A credit spread put can be viewed as an exchange option that gives A the right to exchange one defaultable bond for a certain number (< 1) of default free bonds. Frequently the underlying security is not a defaultable bond but an asset swap package on the defaultable bond.

- the *reference credit asset* is the referenced bond or asset swap
- the *credit event* has occured if the credit spread of the reference asset is above the strike spread at maturity of the option
- the *default payment* is the price that the reference asset would have at the strike spread minus the market price of the reference asset.

Many funds and insurance companies are restricted by their statutes or regulation to investment-grade investments. A credit spread put option would allow these investors a switch out of the defaultable bond investment if the credit quality of the reference credit decreases. For this application the credit event could also be defined as a downgrade of the reference credit to a rating class below investment grade, and the default payment can be an exchange of the reference asset against an index of investment-grade debt.

A second application is the exposure management of committed lines of credit. A committed line of credit is similar to an overdraft on a bank account. The debtor **C** has the right to enter a pre-agreed loan contract at any time he chooses, he can draw his line of credit. For this **C** pays a regular fee to the committed bank **A**. If bank **A** does not want any additional exposure to **C**, it can still commit the line of credit, but hedge with a credit spread option (American style) which enables bank **A** to put the loan to **B** as soon as the line of credit is drawn[3]. In this context the option is also called a *synthetic lending facility*.

4.2.6 Options on Defaultable Bonds

A plain *put option* on a defaultable bond gives **A** the right to sell the defaultable bond to **B** at the strike price at the exercise date. It is different from a classical bond option in that the price of the underlying (the defaultable bond) may jump downwards at defaults. It offers protection against rising default-free interest rates, rising credit spreads and (if the option survives) against defaults.

4.2.7 Basket Structures

In a *basket default swap* we have several reference credits $C_1, \dots, C_N$ with their respective reference credit assets. The credit event can be triggered either by the first default in the basket (first-to-default) or if a certain loss level is exceeded. A first-to-default basket default swap would have the following payoffs:

- **A** pays a regular fee.
- The credit event is the first default of one of the reference credits.
- If the credit event has happened, **B** pays to **A** the default payment.
- The default payment is either par minus the recovery value of the defaulted security, or a binary payment, or another default payment.
- After the first default the basket default swap is terminated.

With a first-to-default swap the credit quality of a portfolio can be enhanced significantly (provided the defaults in the portfolio are largely uncorrelated). **A**'s motivation here is to gain default protection at a lower price while accepting the risk of more than one default in the basket. For counterparty **B** a basket structure can have regulatory capital advantages[4]. Furthermore **B** can

[3] A priori this would not be an optimal early exercise policy, but given bank **A**'s operating constraints it may still be optimal to them.

[4] By the BIS regulatory capital rules **B** only has to provide regulatory capital for *one* defaultable bond, although he bears most of the credit risk of the whole basket.

reach a higher (promised) return on its investment than he could by investing in one of the reference credits alone.

Often, the combination with a basket default swap can enhance the credit quality of a loan portfolio to 'investment grade' rating and thus make it suitable for a new class of investors. The first-to-default structure is most useful for baskets with a small number (less than six) of defaultable bonds, for large portfolios it is more appropriate to specify a certain loss level that has to be exceeded.

4.2.8 Credit Linked Notes

A *credit linked note* is combination of a credit derivative (usually a default swap or a basket default swap) with a bond issued by counterparty **A**. A default-swap linked note has the following structure

- **B** buys the note.
- At the coupon dates **A** pays the coupon of the note (provided **C** has not defaulted).
- If **C** defaults, the note is terminated and **A** pays the recovery rate of the reference asset (a bond issued by **C**) to **B**.
- If **C** has not defaulted until maturity of the note, **A** repays the principal amount of the note.

In general, after a default of **C**, **A** pays to **B** the nominal value of the note *minus* the default payment of the credit derivative that is combined with the note. To counterparty **B** this structure is (almost[5]) equivalent to an investment in a bond issued by **A** and a short position in the credit derivative with **C** as reference credit.

In such a structure, **A** is often a bank that has given a loan to **C**. **A**'s motives are the following:

- **A** receives funding for the loan to **C**.
- **A** has laid off the credit risk to **C**.
- There is no counterparty risk to **A** because **A**'s claims are fully collateralised. (On the other hand **B** does bear counterparty risk to **A**.)
- The credit protection can be bought more easily: Counterparty risk is eliminated and the note can be sold in small denominations to more than one investors.

To **B** this credit linked note offers customised exposure to **C**'s credit risk even if **B** himself would not qualify as a default swap counterparty because of his credit standing. If **A** has a very high credit rating, the credit linked note is equivalent to debt issued by **C** directly. Total return credit linked notes are also a popular vehicle for the securitisation of large pools of small claims in form of *collateralised loan obligations (CLO)*.

[5]The only difference is that after a default of C the credit linked note is terminated while a bond issued by A would survive.

4.2.9 Applications of Credit Derivatives

Some of the applications of credit derivatives have already been mentioned when they were specific to the credit derivative discussed. General fields of application common to most credit derivatives are:

- Applications in the management of credit exposures:
  These include the reduction of credit concentration (through basket structures), easier diversification of credit risk and the direct hedging of default risk.
- In trading, credit derivatives can be used for the arbitrage of mispricing in defaultable bonds (through the possibility of short positions on credit risk) and the general possibility to trade a view on the credit quality of a reference credit (usually through credit spread products). Default digital products also allow the trading of views on the recovery rate of defaulted debt.
- The largest group of credit derivative users are banks who use credit derivatives to free up or manage credit lines, manage loan exposure without needing the consent of the debtor, manage (or arbitrage) regulatory capital or exploit comparative advantages in costs of funding. Another important application here is the securitisation of loan portfolios in form of CLOs.
- The specification of the credit derivatives can be adjusted to the needs of the counterparties: Denomination, currency, form of coupon, maturity or even the general payoff need not match the reference asset. This is especially useful for the management of counterparty exposures from derivatives transactions.

4.3 Defaultable Bond Pricing with Cox Processes

4.3.1 Model Setup and Notation

As in the previous chapter, the model is set up in a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{(t\geq 0)}, P)$ where P is a pre-specified martingale measure. We assume the filtration $(\mathcal{F}_t)_{(t\geq 0)}$ satisfies the usual conditions[6] and the initial filtration $\mathcal{F}_0$ is trivial. We also assume a finite time horizon $\bar{T}$ with $\mathcal{F} = \mathcal{F}_{\bar{T}}$, all definitions and statements are understood to be only valid until this time horizon $\bar{T}$.

The notation used is:

- $B(t, T)$: default free zero coupon bond price,
- $r(t)$: default free short rate,
- $\beta_{t,T}$: discount factor over $[t, T]$,
- $\overline{B}(t, T)$: defaultable bond price,
- $P(t, T)$: survival probability for $[t, T]$.

[6] See Jacod and Shiryaev [1988].

4.3.2 The Time of Default

Although we are going to use two different models to model the *recovery* of defaulted bonds, the model for the *time* of the default(s) is the same for both:

We assume that the times of default τ_i are generated by a Cox process. Intuitively, a Cox Process is defined as a Poisson process with stochastic intensity λ (see Lando [1998], p.101). Formally the definition is:

Definition 9
N is called a Cox process, if there is a nonnegative adapted stochastic process $\lambda(t)$ (called the intensity of the Cox process) with $\int_0^t \lambda(s)ds < \infty \quad \forall\, t > 0$, and conditional on the realization $\{\lambda(t)\}_{\{t>0\}}$ of the intensity, $N(t)$ is a time-inhomogeneous Poisson process with intensity $\lambda(t)$.

This definition follows Lando [1998] and differs from the usual definition of a Cox process where the intensity process $\lambda(t)$ is $\mathcal{F}_0$-measurable (see e.g. Brémaud [1981]). It is not necessary to reveal directly *all* information about the future development of the intensity, and for the valuation of some derivatives this modelling approach would even introduce pricing errors[7].

Assumption 8
(i) The default counting process

$$N(t) := \max\{i|\tau_i \le t\} = \sum_{i=1}^{\infty} \mathbf{1}_{\{\tau_i \le t\}} \tag{4.3}$$

is a Cox process with intensity process $\lambda(t)$.

(ii) In the equivalent recovery model the time of default is the time of the first jump of N. To simplify notation the time of the first default will be referred to with $\tau := \tau_1$.

(iii) In the fractional recovery model the times of default are the times of the jumps of N.

Remark 1
Given the realisation of λ, the probability of having exactly n jumps is

$$\mathrm{P}\left[\, N(T) - N(t) = n \mid \{\lambda(s)\}_{\{T \ge s \ge t\}} \,\right] = \frac{1}{n!}\left(\int_t^T \lambda(s)ds\right)^n \exp\{-\int_t^T \lambda(s)ds\}. \tag{4.4}$$

The probability of having n jumps (without knowledge of the realisation of λ) is found by conditioning on the realisation of λ *within* an outer expectation operator:

$$\begin{aligned} \mathrm{P}\left[\, N(T) - N(t) = n \mid \mathcal{F}_t \,\right] &= \mathrm{E}\left[\ \mathrm{P}\left[\, N(T) - N(t) = n \mid \{\lambda(s)\}_{\{T \ge s \ge t\}} \,\right] \mid \mathcal{F}_t \right] \\ &= \mathrm{E}\left[\frac{1}{n!}\left(\int_t^T \lambda(s)ds\right)^n \exp\{-\int_t^T \lambda(s)ds\} \,\middle|\, \mathcal{F}_t \right], \end{aligned} \tag{4.5}$$

[7] Consider e.g. an American Put option on a defaultable bond in a world with constant zero default-free interest rates. If all information about $\lambda(t)$ is revealed at $t = 0$ this would enable the investor to condition his optimal exercise policy on the future development of λ which is not realistic.

Define the process $P(t,T)$

$$P(t,T) = \mathbf{E}\left[e^{-\int_t^T \lambda(s)ds} \,\middle|\, \mathcal{F}_t \right]. \tag{4.6}$$

For $\tau > t$ $P(t,T)$ can be interpreted as the *survival probability* from time t until time T. In general,

$$\mathbf{1}_{\{\tau>t\}} P(t,T) = \mathbf{E}\left[\mathbf{1}_{\{\tau>T\}} \,\middle|\, \mathcal{F}_t \right]. \tag{4.7}$$

Given $\tau > t$ the density of the time of the first default as seen from t is for $T > t$

$$p(t,T) = \mathbf{E}\left[\lambda(T) \exp\{-\int_t^T \lambda(s)ds\} \,\middle|\, \mathcal{F}_t \right], \tag{4.8}$$

and $p(t,T) = 0$ for $T \leq t$. [8]

The law of iterated expectations will prove extremely useful later on, it was first used with Cox processes in a credit risk context by Lando [1998].

The specification of the default trigger process as a Cox process precludes a dependence of the default intensity on previous defaults[9] and also ensures totally inaccessible stopping times τ_i as times of default. Apart from this it allows rich dynamics of the intensity process, specifically, we can reach stochastic credit spreads. If only the time of the *first* jump of N is of interest, the Cox-process specification is completely without loss of generality within the totally inaccessible stopping times.

Equations (4.4), (4.5) and (4.8) will be used frequently later on. In the following sections we will consider time t as 'today', and assume that no default has happened so far $\tau > t$. (The statements for $\tau < t$ are trivial.)

4.3.3 The Fractional Recovery Model

The model used here is an extension of the Duffie-Singleton [1999] model to multiple defaults. More details to the model can be found in Schönbucher [1996, 1998] and in section 3.6 in the previous chapter. The new feature of this model is that a default does not lead to a liquidation but a reorganisation of the issuer: defaulted bonds lose a fraction q of their face value and continue to trade. This feature enables us to consider European-type payoffs in our derivatives without necessarily needing to specify a payoff of the derivative at default (although we will consider this case, too). The next assumption (a repetition of assumption 4) summarises the fractional recovery model:

Assumption 9
There is an increasing sequence of stopping times $\{\tau_i\}_{i\in\mathbb{N}}$ that define the times of default (given in definition 9 and assumption 8). At each default τ_i the defaultable bond's face value is reduced

[8] If a default has already happened, $p(t,T) = \epsilon_\tau$ the density of the first default reduces to the Dirac measure at τ.
[9] It is therefore not possible to specify an intensity that jumps at defaults.

by a factor q_i, where q_i may be a random variable itself. A defaultable zero coupon bond's final payoff is the product

$$Q(T) := \prod_{\tau_i \leq T} (1 - q_i) \tag{4.9}$$

of the face value reductions after all defaults until the maturity T of the defaultable bond. The loss quotas q_i can be random variables drawn from a distribution $K(dq)$ at time τ_i, but for the first calculations we will assume $q_i = q$ to be constant.

It is now easily seen[10] that in this setup the price of a defaultable zero coupon bond is given by

$$\overline{B}(t,T) = Q(t)\mathbf{E}\left[e^{-\int_t^T \overline{r}(s)ds} \;\middle|\; \mathcal{F}_t \right]. \tag{4.10}$$

The process $\overline{r}$ is called the defaultable short rate $\overline{r}$ and it is defined by

$$\overline{r} = r + \lambda q. \tag{4.11}$$

Here r is the default-free short rate, λ the hazard rate of the defaults and q is the loss quota in default. If q is stochastic then q has to replaced by its (local) expectation $q_t^e = \int qK_t(dq)$ in equation (4.11).

It is convenient to decompose the defaultable bond price $\overline{B}$ as follows:

$$\overline{B}(t,T) = Q(t)\widetilde{B}(t,T) = Q(t)B(t,T)\widetilde{P}(t,T). \tag{4.12}$$

Here $Q(t)$ represents the face-value reduction due to previous defaults (before time t). Frequently we will be able to set $t = 0$ and thus $Q(t) = 1$, but for the analysis at intermediate times it is important to be clear about the notation[11]. The defaultable bond price $\overline{B}(t,T)$ is thus the product of $Q(t)$, the influence of previous defaults, and the product of the default-free bond price $B(t,T)$ and the third factor $\widetilde{P}(t,T)$ which is uniquely defined by equation (4.12), or equivalently:

$$\widetilde{P}(t,T) = \frac{1}{Q(t)} \frac{\overline{B}(t,T)}{B(t,T)}. \tag{4.13}$$

Remark 2
$\widetilde{P}(t,T)$ is related to the *survival probability* $P(t,T)$ of the defaultable bond: If r and λ are independent and there is a total loss ($q = 1$) at default then $\widetilde{P}(t,T)$ is the probability (under the martingale measure) that there is no default in $[t,T]$.
If r and λ are not independent, $\widetilde{P}(t,T)$ is the survival probability under the T-forward measure P^T

$$\widetilde{P}(t,T) = \mathbf{E}^{P^T}\left[e^{-\int_t^T q\lambda(s)ds} \;\middle|\; \mathcal{F}_t \right]. \tag{4.14}$$

(Under independence $\mathbf{E}^T\left[e^{-\int_t^T q\lambda(s)ds} \;\middle|\; \mathcal{F}_t \right]$ and $\mathbf{E}\left[e^{-\int_t^T q\lambda(s)ds} \;\middle|\; \mathcal{F}_t \right]$ coincide.)
If there is positive recovery ($q < 1$) then $\widetilde{P}(t,T)$ is the *expected final payoff* under the T-forward measure, but the implied survival probability cannot be recovered.

[10] Using the iterated expectations, see equation (3.81) and also Duffie and Singleton [1999] and Schönbucher [1996, 1998] for a more general proof.

[11] For example, at the expiry date of an option we would like to separate previous defaults and credit spreads in the price of the underlying.

4.3.4 The Equivalent Recovery Model

The equivalent recovery model has been proposed by several authors, amongst them Jarrow and Turnbull [1995] Lando [1998] and Madan and Unal [1998]. Here the recovery of defaulted debt is treated as follows:

Assumption 10
At the time of default τ one defaultable bond $\overline{B}(\tau, T)$ with maturity T has the payoff of c default free bonds $B(\tau, T)$ of the same maturity and face value, where c may be random, too.

Under the equivalent recovery model (with constant c and given no default so far $\tau > t$) the price of a defaultable bond can be decomposed into c default-free bonds and $(1-c)$ defaultable bonds with zero recovery

$$\begin{aligned}\overline{B}(t,T) &= \mathbf{E}\left[\,\beta_{t,T}\mathbf{1}_{\{\tau>T\}} + c\beta_{t,\tau}B(\tau,T)\mathbf{1}_{\{\tau\leq T\}} \mid \mathcal{F}_t\,\right] \\ &= \mathbf{E}\left[\,\beta_{t,T}\mathbf{1}_{\{\tau>T\}} \mid \mathcal{F}_t\,\right] + c\mathbf{E}\left[\,\beta_{t,T} \mid \mathcal{F}_t\,\right] - c\mathbf{E}\left[\,\beta_{t,T}\mathbf{1}_{\{\tau>T\}} \mid \mathcal{F}_t\,\right] \\ &= (1-c)\overline{B}_0(t,T) + cB(t,T), \end{aligned} \tag{4.15}$$

where $\overline{B}_0(t,T)$ is the price of a defaultable bond under zero recovery:

$$\mathbf{E}\left[\,\beta_{t,T}\mathbf{1}_{\{\tau>T\}} \mid \mathcal{F}_t\,\right] = \mathbf{1}_{\{\tau>t\}}\mathbf{E}\left[\,e^{-\int_t^T r(s)+\lambda(s)ds} \;\middle|\; \mathcal{F}_t\,\right].$$

It should be pointed out that the equivalent recovery model is not able to fit all term structures of credit spreads with a given fixed common recovery rate c. Assume $\tau > t$ and the term structure of credit spreads is at a constant credit spread h for all maturities T. Then

$$\frac{\overline{B}(t,T)}{B(t,T)} = e^{-h(T-t)}$$

and for large enough $T-t$ (such that $T-t > -(\ln c)/h$),

$$\tilde{P}(t,T) = \frac{1}{1-c}\left(\frac{\overline{B}(t,T)}{B(t,T)} - c\right) = \frac{1}{1-c}\left(e^{-h(T-t)} - c\right) < 0.$$

the survival probability (see below) that can be implied from the zero-recovery bond $\overline{B}_0(t,T)$ would become negative, which is obviously not sensible. In the equivalent recovery model there is a lower bound on the ratio of defaultable bond prices to default-free bond prices and this bound is the recovery rate c. Therefore the zero coupon yield spread must satisfy

$$\overline{y}(t,T) - y(t,T) < -\frac{\ln c}{T-t},$$

which may not be satisfied by market prices for longer times to maturity $T-t$ and high credit spreads. E.g. for a recovery rate of $c = 50\%$ and a time to maturity of $T - t = 10$ years the maximal (continuously compounded) credit spread is $h = 6.93\%$.

Despite these different properties of the two modelling approaches, with a suitable choice of (time dependent or stochastic) parameters, both models can be transformed into each other: The value of the security in default is only expressed in different numeraires, once in terms of defaultable bonds and once in terms of default-free bonds. Both approaches are therefore equivalent and one should use the specification that is best suited for the issue at hand.

4.3.5 Implied Survival Probabilities

In the equivalent recovery model it is easy to recover *implied survival probabilities* from a given term structure of defaultable bond prices and a given value for c. From equation (4.15) we have

$$\widetilde{P}(t,T) = \frac{\overline{B}_0(t,T)}{B(t,T)} = \frac{1}{1-c}\left(\frac{\overline{B}(t,T)}{B(t,T)} - c\right). \tag{4.16}$$

As before $\widetilde{P}(t,T)$ is the probability of survival from t to T under the T-forward measure (and also under the spot martingale measure for independence of credit spreads and interest rates).

This survival probability and the prices of defaultable zero coupon bonds $\overline{B}_0(t,T)$ under zero recovery are very useful to value survival contingent payoffs. For many pricing applications knowledge of $\overline{B}_0(t,T)$ is already sufficient. It is a great advantage of the equivalent recovery model that it allows to derive the value of a survival contingent payoff just from the defaultable and default-free term structures and an assumption about recovery rates c.

In the fractional recovery model it is not possible to derive the value of a zero-recovery defaultable bond just from knowledge of the recovery rate q, the defaultable bond price and the default-free bond prices unless the recovery rate is zero. Here a full specification of the dynamics of r and λ is needed.

Given independence of interest rates and the default intensity, the implied survival probability is the ratio of the zero coupon bond prices:

$$P(t,T) = \frac{\overline{B}_0(t,T)}{B(t,T)}$$

Typically the survival probability $P(t,T)$ will change over time because of two effects: First, if there was no default in $[t, t+\Delta t]$ this reduces the possible default times, information has arrived via the (non)-occurrence of the default. Secondly, additional default-relevant information could have arrived in the meantime.

For the analysis of the local default probability in some future time interval it is instructive to consider the *conditional* probability of survival. The probability of survival in $[T_1, T_2]$, given that there was no default until T_1 and given the information at time t is:

$$P(t, T_1, T_2) = \frac{P(t,T_2)}{P(t,T_1)} = \frac{\overline{B}_0(t,T_2)}{B(t,T_2)}\frac{B(t,T_1)}{\overline{B}_0(t,T_1)}.$$

This is a simple consequence of Bayes' rule. The probability of survival until T is the probability of survival until $s < T$ times the conditional probability of survival from s until T:

$$P(t,T) = P(t,s)P(t,s,T).$$

There is a close connection between forward rates and conditional survival / default probabilities.

Definition 10
The default-free simply compounded forward rate *over the period* $[T_1, T_2]$ *as seen from t is:*

$$F(t, T_1, T_2) = \frac{B(t, T_1)/B(t, T_2) - 1}{T_2 - T_1}$$

The zero-recovery defaultable simply compounded forward rate *over the period* $[T_1, T_2]$ *as seen from t is:*

$$\overline{F}(t, T_1, T_2) = \frac{\overline{B}_0(t, T_1)/\overline{B}_0(t, T_2) - 1}{T_2 - T_1}$$

Proposition 15
The conditional probability of default *over* $[T_1, T_2]$ *is given by:*

$$\frac{P^{\text{def}}(t, T_1, T_2)}{T_2 - T_1} = \frac{\overline{F}(t, T_1, T_2) - F(t, T_1, T_2)}{1 + (T_2 - T_1)\overline{F}(t, T_1, T_2)}.$$

The marginal probability of default *at time T is the spread of the continuously compounded defaultable forward rate over the default-free forward rate:*

$$\lim_{\Delta t \searrow 0} \frac{P^{\text{def}}(t, T, T + \Delta t)}{\Delta t} = \overline{f}(t, T) - f(t, T).$$

Proof: (dropping the t-index)

$$\begin{aligned}
P^{\text{def}}(T_1, T_2) &= 1 - P(T_1, T_2) \\
&= 1 - \frac{\overline{B}_0(T_2) B(T_1)}{B(T_2)\overline{B}_0(T_1)} \\
&= \frac{B(T_2)\overline{B}_0(T_1) - \overline{B}_0(T_2) B(T_1)}{B(T_2)\overline{B}_0(T_1)} \\
&= \frac{B(T_2)[\overline{B}_0(T_1) - \overline{B}_0(T_2)] - \overline{B}_0(T_2)[B(T_1) - B(T_2)]}{B(T_2)\overline{B}_0(T_1)} \\
&= \frac{\overline{B}_0(T_2)}{\overline{B}_0(T_1)} \frac{\overline{B}_0(T_1) - \overline{B}_0(T_2)}{\overline{B}_0(T_2)} - \frac{\overline{B}_0(T_2)}{\overline{B}_0(T_1)} \frac{B(T_1) - B(T_2)}{B(T_2)}
\end{aligned}$$

therefore

$$\frac{P^{\text{def}}(T_1, T_2)}{T_2 - T_1} = \frac{\overline{B}_0(T_2)}{\overline{B}_0(T_1)} \Big(\overline{F}(T_1, T_2) - F(T_1, T_2)\Big),$$

and from definition 10 follows that

$$\frac{\overline{B}_0(T_1)}{\overline{B}_0(T_2)} = 1 + (T_2 - T_1)\overline{F}(T_1, T_2).$$

The result for the marginal default probability follows directly from taking the limit. □

The default probability over the interval $[T_1, T_2]$ equals *the length of the interval* times *the spread of the simply compounded forward rates over the interval* times *discounting with the defaultable forward rates*.

For small time intervals, the probability of default in $[T, T + \Delta t]$ is approximately *proportional* to the length of the interval with proportionality factor $(\bar{f}(t,T) - f(t,T))$.

These results highlight two points. First, there is an intimate connection between default probabilities and credit spreads. A full term structure of credit spreads contains a wealth of information about the market's perception of the likelihood of default at each point in time. The equivalent recovery model has the advantage of making this information more easily accessible than the fractional recovery model. Unfortunately, to reach this information in a practical application, a recovery rate c is needed, and the assumption of independence of defaults and default-free term structure of interest rates must be made. There is a large degree of uncertainty about recovery rates with variation between 20% and 80%. The independence assumption will have a smaller effect on the results. This assumption will be relaxed in the next sections.

The second observation is the reason why processes like Poisson or Cox processes are so well suited for credit-spread based default modelling. These processes have intensities, and the probability of jump of a point process with an intensity is approximately proportional to the length of the time interval considered (for small intervals). The proportionality factor is the intensity at that point. This property is exactly equivalent to the second equation in proposition 15, and it also gives a link to the model of forward credit spreads in the previous chapter. But proposition 15 is also valid for default models that are not based on an intensity model.

4.4 Direct Valuation of Credit Risk Derivatives

In this section we begin the pricing of credit derivatives in a framework that is independent of the specification of the model.

We first concentrate on credit derivatives that can be priced directly off the term structures of interest rates and defaultable bond prices. Among these, the default digital put option and the default digital swap will be analysed first, as these products are the simplest and the pricing of the other products can often be reduced to the pricing of default digital puts. Next are defaultable floating rate notes[12] and the default swap. Finally, intermediate valuation formulae are derived for credit spread options and options on defaultable bonds.

Assumption 11
The following data is needed for all maturities $T > 0$:

- *the default-free term structure of bond prices*

$$B(0,T) = \mathrm{E}\left[e^{-\int_0^T r(s)ds} \right]$$

[12] Strictly speaking the defaultable floating rate note is not a credit derivative. It is included here because its valuation is closely linked to the valuation of the default swap.

- *the defaultable bond prices*

$$\overline{B}(0,T)$$

- *the defaultable bond prices under zero recovery*

$$\overline{B}_0(0,T) = \mathbf{E}\left[e^{-\int_0^T r(s)+\lambda(s)ds} \right].$$

At some points we will assume that the term structures of interest rates and credit spreads are independent.

The assumptions about the available data are open to criticism. One of the largest difficulties in credit risk modelling is the scarcity of useful data, and only rarely (maybe for some sovereign issuers) a full term structure of defaultable bond prices is available, the same applies to survival probabilities or recovery rates. These data problems will have to be addressed separately, here we just point out that all of the defaultable inputs (recovery rates, $\overline{B}_0(0,T)$ or $P(0,T)$) can also stem from other sources than market prices, e.g. fundamental analysis, a different credit risk model or historical data of the same rating class or industry. The only requirement is that this information is given under the martingale measure.

The independence of credit spreads and interest rates will yield the survival probabilities

$$P(0,T) = \mathbf{E}\left[e^{-\int_0^T \lambda(s)ds} \right] = \overline{B}_0(0,T)/B(0,T).$$

If this assumption is dropped, we need knowledge of the full dynamics of the term structures of credit spreads, defaults and interest rates. This case will be treated in the following sections.

Alternatively to giving the defaultable bond prices under zero recovery we can also assume that an expected recovery rate is given for the equivalent recovery model. As mentioned in the previous section, it is then straightforward to imply zero-recovery defaultable bond prices from bond prices that are given under positive recovery.

4.4.1 Forms of Payment for Default Protection

There are two alternative ways in which prices for credit derivatives can be quoted: Either the price for the default protection is expressed in terms of an *up-front* fee D that is payable at $t = 0$, or a default protection fee S has to be paid *in regular intervals* until a default has happened. The latter alternative is very popular because of its similarity to swap contracts, the only difference is that the regular payments end at default. The difference is the numeraire in which the default protection is priced.

To convert the two representations we observe the following:
The value of receiving S at the times $T_i, i = 1, \dots N$ until a default is at time 0:

$$S \sum_{i=0}^{N} \overline{B}_0(0,T_i),$$

and for continuous payment (receiving sdt from T_0 until default or T_N) the value at time 0 is

$$s \int_{T_0}^{T_N} \overline{B}_0(0,t)dt. \tag{4.17}$$

These equations remain valid under correlation of r and λ.

To find the regular payments S or s that correspond to a given up-front price D we therefore only have to calculate

$$S = \frac{D}{\sum_{i=0}^{N} \overline{B}_0(0,T_i)} \qquad \text{or} \qquad s = \frac{D}{\int_{T_0}^{T_N} \overline{B}_0(0,t)dt}. \tag{4.18}$$

In the following sections we will therefore only calculate the up-front prices D of the credit derivatives.

4.4.2 Default Digital Payoffs

The default digital put option has the payoff 1 *at the time of default*. The timing is important because it determines until when the payoff has to be discounted, but we will consider a simpler specification in a first step:

4.4.2.1 Payoff at Maturity

As a first example consider an European-style default digital put which pays off 1 *at* T iff there has been a default at some time before (or including) T. Its value is easily found (τ denotes the time of the first default):

$$\begin{aligned}
D =& \mathrm{E}\left[e^{-\int_0^T r(s)ds} \mathbf{1}_{\{\tau<T\}} \right] \\
=& \mathrm{E}\left[\mathrm{E}\left[e^{-\int_0^T r(s)ds} \mathbf{1}_{\{\tau<T\}} \,\middle|\, \{\lambda(s)\}_{s\le T} \right] \right] \\
=& \mathrm{E}\left[e^{-\int_0^T r(s)ds} (1 - e^{-\int_0^T \lambda(s)ds}) \right] \\
=& B(0,T) - \mathrm{E}\left[e^{-\int_0^T r(s)+\lambda(s)ds} \right] \\
=& B(0,T) - \overline{B}_0(0,T).
\end{aligned} \tag{4.19}$$

This follows also directly from the fact that this default digital put replicates the payoff of a portfolio that is long one default-free bond and short one zero recovery defaultable bond.

4.4.2.2 Payoff at Default

In the case when the payoff takes place *at* default, the expectation to calculate is

$$D = \mathrm{E}\left[e^{-\int_0^\tau r(s)ds} \mathbf{1}_{\{\tau<T\}} \right].$$

Conditioning on the realization of λ yields

$$D = \mathbf{E}\left[e^{-\int_0^T r(s)ds} 1_{\{\tau<T\}} \right] = \mathbf{E}\left[\mathbf{E}\left[e^{-\int_0^T r(s)ds} 1_{\{\tau<T\}} \,\middle|\, \lambda \right] \right]. \tag{4.20}$$

Now note that the probability distribution of τ (given λ) is

$$\mathbf{P}[\,\tau \leq T\,] = 1 - \mathbf{P}[\,\tau > T\,] = 1 - \exp\{-\int_0^T \lambda(s)ds\},$$

so the density of τ is

$$\lambda(t)\exp\{-\int_0^t \lambda(s)ds\}. \tag{4.21}$$

Substituting equation (4.21) into (4.20) yields for the price of the default digital put:

$$\begin{aligned} D &= \mathbf{E}\left[\mathbf{E}\left[e^{-\int_0^T r(s)ds} 1_{\{\tau<T\}} \,\middle|\, \lambda \right] \right] \\ &= \mathbf{E}\left[\int_0^T \lambda(t) e^{-\int_0^t \lambda(s)ds} e^{-\int_0^t r(s)ds} dt \right] \\ &= \int_0^T \mathbf{E}\left[\lambda(t) e^{-\int_0^t \lambda(s)ds} e^{-\int_0^t r(s)ds} \right] dt, \end{aligned} \tag{4.22}$$

where we assume sufficient regularity to allow the interchange of expectation and integration. Under independence of τ and λ the expectation in the integral in equation (4.22) can be further simplified to

$$\mathbf{E}\left[\lambda(t) e^{-\int_0^t \lambda(s)ds} e^{-\int_0^t r(s)ds} \right] = B(0,t)P(0,t)h(0,t) = \overline{B}_0(0,t)h(0,t). \tag{4.23}$$

Here $P(0,t)$ is the survival probability, and $h(0,t)$ is the associated 'forward rate' of the spreads of the zero recovery bonds:

$$\begin{aligned} P(0,t) &= \overline{B}_0(0,t)/B(0,t) = \mathbf{E}\left[e^{-\int_0^t \lambda(s)ds} \right] \\ h(0,t) &= -\frac{\partial}{\partial T} \ln P(0,T)\bigg|_{T=t}. \end{aligned}$$

Equation (4.23) follows from the well-known fact[13] in default-free interest rate theory that the expectation of the short rate $r(T)$ at time T under the T-forward measure is the T-forward rate $f(0,T)$:

$$\mathbf{E}\left[r(t) e^{-\int_0^t r(s)ds} \right] = B(0,t)f(0,t).$$

Here we have a mathematically equivalent situation where the notation is changed such that $P(0,t)$ are our bond prices, P_t is the forward measure, $\lambda(t)$ is the short rate at t and $h(0,t)$ is the t-forward rate. Using the survival probability $P(0,t)$ as numeraire in a change of measure it therefore follows that

$$\mathbf{E}\left[\lambda(t) e^{-\int_0^t \lambda(s)ds} \right] = P(0,t)\mathbf{E}^{P_t}[\,\lambda(t) \mid \mathcal{F}_t\,] = P(0,t)h(0,t).$$

[13] See e.g. Sandmann and Sondermann [1997].

Under the equivalent recovery model we know the zero-recovery bond prices $\overline{B}_0(t,T) = \frac{1}{1-c}(\overline{B}(t,T) - cB(t,T))$, hence under independence

$$h(t,T) = -\frac{B(t,T)}{\overline{B}(t,T) - cB(t,T)} \frac{\partial}{\partial T}\left(\frac{\overline{B}(t,T)}{B(t,T)}\right) \tag{4.24}$$

$$\overline{B}_0(t,T)h(t,T) = \frac{1}{1-c}\overline{B}(t,T)(\overline{f}(t,T) - f(t,T)) \tag{4.25}$$

hold, where the defaultable forward rate $\overline{f}(t,T)$ is defined using the defaultable bonds with positive recovery: $\overline{f}(t,T) = -\frac{\partial}{\partial T}\ln\overline{B}(t,T)$.

Putting all together we gain the following representation for the value of a default digital put with maturity T:

Proposition 16

Consider a default digital put with maturity T and notional value 1. All prices are as seen from time 0.

(i) *If the default digital put is settled at T, its value is*

$$D = B(0,T) - \overline{B}_0(0,T).$$

(ii) *If the default digital put is settled at default, its value is*

$$D = \int_0^T \mathbf{E}\left[\lambda(t)e^{-\int_0^t \lambda(s)ds}e^{-\int_0^t r(s)ds}\right]dt. \tag{4.26}$$

(iii) *Assume that r and λ are independent. Then the value of the default digital put with settlement at default is*

$$D = \int_0^T \overline{B}_0(0,t)h(0,t)dt. \tag{4.27}$$

(iv) *Assume that r and λ are independent, and the equivalent recovery rate is c. Then the price of a default digital put is*

$$D = \frac{1}{1-c}\int_0^T \overline{B}(0,t)(\overline{f}(0,t) - f(0,t))dt. \tag{4.28}$$

The expectation in equation (4.26) will be calculated later on.

4.4.2.3 The Default Digital Swap

For the default digital payoff, the swap specification is as follows:

> Party **A** pays 1 at default if there is a default. Party **B** pays a regular fee $\dot{s}(t)dt$ until default.

If $s(t)$ can be stochastic, the fair fee is the local default intensity $\lambda(t)$. The intuitive reason is that $\lambda(t)$ is the local default probability at time t, because for small Δt

$$\mathbf{E}\left[\, \mathbf{1}_{\{\tau \in [t,t+\Delta t]\}} \;\middle|\; \mathcal{F}_t \right] \approx \lambda(t)\Delta t.$$

If the default protection is re-negotiated for every small time-interval $[t, t + dt]$, then the fee should be $\lambda(t)dt$ for this interval.

Mathematically this follows from the fact that $\int_0^t \lambda(s)\mathbf{1}_{\{s\leq\tau\}}ds$ is the predictable compensator of $\mathbf{1}_{\{\tau\leq t\}}$, thus the expected values of stochastic integrals w.r.t. either of them are equal:

$$\mathbf{E}\left[\int_0^T \beta_{0,t} d\mathbf{1}_{\{\tau\leq t\}} \;\middle|\; \mathcal{F}_t \right] = \mathbf{E}\left[\int_0^T \beta_{0,t}\lambda(t)\mathbf{1}_{\{t\leq\tau\}}dt \;\middle|\; \mathcal{F}_t \right].$$

The l.h.s. of the equation is the discounted payoff of the default-insurance side of the swap (receiving 1 at default), the r.h.s. is the premium side of the swap. The expected discounted values of both sides of the swap are equal. This result is independent of the maturity of the swap and also valid for correlation between interest rates and defaults.

The previous result required a continuous adaptation of the swap rate. Typically the swap rate s is a constant that is determined at $t = 0$ for all future times $t \leq T$. Then from equation (4.18) and proposition 16 follows the fair swap rate for a default digital swap:

Proposition 17
Assume that r and λ are independent.

(i) *The fair swap rate for a default digital swap is*

$$s = \frac{\int_0^T \overline{B}_0(0,t)h(0,t)dt}{\int_0^T \overline{B}_0(0,t)dt}. \tag{4.29}$$

(ii) *Assume that the equivalent recovery rate is c. Then the fair rate of a default digital swap is*

$$s = \frac{\int_0^T \overline{B}(0,t)(\overline{f}(t,T) - f(t,T))dt}{\int_0^T \overline{B}(0,t) - cB(0,t)dt}. \tag{4.30}$$

Remark 3
Compare the fair default digital swap rate to the plain vanilla fixed-for-floating swap rate in default-free interest rate theory:

$$s_r = \frac{1 - B(0,T)}{\int_0^T B(0,t)dt} = \frac{\int_0^T B(0,t)f(0,t)dt}{\int_0^T B(0,t)dt}.$$

(The equality follows from $\frac{\partial}{\partial t}B(0,t) = -f(0,t)B(0,t)$ and integrating.) We see that the equation for the default digital swap rate has a structure that is very similar to the structure of the equation for plain fixed-for-floating swap rates. The forward rate $f(0,t)$ in the default-free case has been replaced by the credit spread $h(0,t)$ and the default-free bond prices $B(0,t)$ are replaced by their defaultable counterparts $\overline{B}_0(0,t)$. In both cases the numerator denotes the value of the floating leg of the swap, and the denominator the fixed leg of the swap, where the floating leg for the default digital swap is the default contingent payment.

The fact that the default digital swap is terminated at default introduces another difference to default-free swaps: The fixed leg has to be discounted using *defaultable* bond prices ($\overline{B}_0$, the defaultable bond prices for zero recovery). If the swap was not killed at default, the fixed side would have to be discounted with default-free rates.

4.4.3 Default Swaps

As opposed to a default digital swap which only pays a lump sum at default, a default swap covers the full loss in default of a defaultable bond. There are two possible specifications:

(a) Replacement of *the difference to par:* The payoff at the time τ of default is (using the equivalent recovery model)

$$1 - cB(\tau, T).$$

A default swap on a defaultable *coupon* bond will pay off the difference between the defaulted coupon bond price and the par value of this bond. (Thus only the principal is protected.)

(b) Replacement of *the difference to an equivalent default-free bond:* The payoff is

$$(1 - c)B(\tau, T),$$

The first variant (replacement of the difference to par) is the more common specification. Both variations (and the case of a default swap on a defaultable *coupon* bond) can be priced using the results of the previous section using portfolio arguments. For simplicity, we will first only derive the spot prices for the default insurance components the fair swap rates follow then directly from equation (4.18).

4.4.3.1 Difference to Par

Consider the portfolio consisting of

- 1 defaultable bond $\overline{B}(0, T)$
- 1 default put.

This portfolio has the following payoffs:

1. if the bond survives until T:
   1 at time T;
2. if the bond defaults before T:
   1 at the default time τ (combined payoff of defaultable bond and default put).

This payoff profile can be replicated using a portfolio consisting of

- 1 defaultable bond with zero recovery $\overline{B}_0(0,T)$
- 1 default digital put.

Thus, the price of a default put equals:

$$D^{\text{def put}} = \overline{B}_0(0,T) - \overline{B}(0,T) + D^{\text{def digital put}}.$$

The zero-recovery bond price $\overline{B}_0(0,T)$ was given by assumption, the price of the default digital put $D^{\text{def digital put}}$ was derived in the previous section.

For a default put on a defaultable *coupon* bond $\widehat{\overline{B}}(0,T)$ we only need to replace the zero coupon bonds in the example above with the respective coupon bonds. The first portfolio becomes

- 1 defaultable coupon bond $\widehat{\overline{B}}(0,T)$
- 1 default put on this bond,

and the second portfolio will be

- 1 defaultable coupon bond with zero recovery $\widehat{\overline{B}}_0(0,T)$
- 1 default digital put on the principal amount of this bond.

Again both portfolios will have identical payoff streams. The price of the zero-recovery defaultable coupon bond can be derived from the term structure of zero-recovery defaultable *zero-coupon* bonds. Note that the default digital put in the second portfolio is only on the principal of the coupon bond.

$$D^{\text{def put}} = \widehat{\overline{B}}_0(0,T) - \widehat{\overline{B}}(0,T) + D^{\text{def digital put}}.$$

4.4.3.2 Difference to default-free

If the default put pays off the price difference of the defaulted bond to an equivalent default-free bond, then the combination of this default put with a defaulted bond yields a default-free bond. Therefore its price must be the difference between the default-free bond and the defaultable bond:

$$D^{\text{def put}} = B(0,T) - \overline{B}(0,T).$$

This result is independent of correlation between credit spreads and interest rates or any assumptions about the recovery rates in default.

4.4.4 Defaultable FRNs and Default Swaps

A default swap on a defaultable floating rate note (FRN) can be used to set up a perfect hedge even for stochastic recovery rates and correlation between spreads and interest rates. We use this property to give a characterisation of the default swap rate in terms of the credit spread of defaultable FRNs.

Assume the following payoffs: The defaultable FRN pays the floating rate $r(t)$ plus a constant spread $\bar{s}$ per time. The default swap rate is s and the default swap pays the loss to par in default.

	defaultable FRN	default swap	default-free FRN
Price at $t = 0$	$\overline{F}(0)$	0	1
periodic payments	$(r(t) + \bar{s})dt$	$-sdt$	$r(t)dt$
final payoff at $t = T$	$1 + (r(t) + \bar{s})dt$	$-sdt$	$1 + r(t)dt$
value at default	$1 - c$	c	1

Table 4.1: The Payoffs of defaultable FRN, default swap and default-free FRN.

The default-free FRN pays the floating rate r. Both floaters have a final payoff of 1. The payoffs are shown in table 1.

If the defaultable FRN trades at par at $t = 0$, i.e. $\overline{F}(0) = 1$ then initially (at $t = 0$), at defaults and at $t = T$ we have

defaultable FRN + default swap = default-free FRN.

As both the spread $\bar{s}$ of the defaultable FRN and the default swap rate s are constant they must coincide. Therefore

$$\bar{s} = s,$$

Assume there is a defaultable FRN that trades at par and that has a coupon spread of $\bar{s}$ over the default-free rate r. Then $s = \bar{s}$ is the fair swap rate for a default swap of the same maturity.

This argument only uses a simple comparison of payoff schedules, it does not use any assumptions about the dynamics and distribution of default-free interest rates or credit spreads or about the recovery rates. If the FRNs only pay coupons at discrete time interval this relationship is only approximately valid, with an exact fit at the coupon dates because only then the default-free FRN is worth exactly 1.

4.4.5 Credit Spread Forwards

A *credit spread forward* is a contract to exchange at a future time T_2 against a fixed credit spread $\bar{s}$, the credit spread that a defaultable bond $\overline{B}(T_1, T_2)$ had over an equivalent default-free bond $B(T_1, T_2)$ at time T_1. Thus the spread is fixed at T_1 and exchanged at T_2. The payoff function is

$$\frac{1/\overline{B}(T_1, T_2) - 1}{T_2 - T_1} - \frac{1/B(T_1, T_2) - 1}{T_2 - T_1} - \bar{s} \qquad \text{at } T_2. \tag{4.31}$$

The credit spread forward is the adaptation of a classical forward rate agreement (FRA) to credit spreads. Here the defaultable simply compounded interest rate is defined using defaultable bond prices with positive recovery (and not zero recovery as in definition 10). If a default has happened before T_1 and the defaulted bond is used for the calculation, the defaultable rate in equation (4.31) will become very large thus yielding a very high payoff to the receiver of the credit spread.

It should be noted that the classical replication portfolio for a FRA cannot be used to replicate the defaultable rate $\frac{1}{T_2-T_1}(1/\overline{B}(T_1,T_2)-1)$. The classical replication portfolio is:

- Invest \$ $B(t,T_1)$ in $B(t,T_1)$ at t (to get 1 of these bonds).
- Payoff at T_1 is \$ 1.
- Invest \$1 in $B(T_1,T_2)$ at T_1 (to get $1/B(T_1,T_2)$ of these bonds).
- Payoff at T_2 is $1/B(T_1,T_2)$.

This strategy will replicate $1/B(T_1,T_2)$ at T_2. To replicate -1 at T_2 we need to short one zero bond $B(t,T_2)$ with maturity T_2, and for the factor $1/(T_2-T_1)$ the whole strategy has to be done $1/(T_2-T_1)$ times. If this strategy is attempted with defaultable bonds $\overline{B}(t,T)$ it will fail if defaults happen before T_1.

The payoff $1/\overline{B}(T_1,T_2)$ at T_2 is unaffected by defaults in $[T_1,T_2]$, thus one needs to invest in $1/\overline{B}(T_1,T_2)$ default free bonds $B(T_1,T_2)$ at T_1 to replicate $1/\overline{B}(T_1,T_2)$ at T_2. This will cost $B(T_1,T_2)/\overline{B}(T_1,T_2)$ at T_1. The payoff of the credit spread forward is thus equivalent to

$$\frac{1}{T_2-T_1}\left(\frac{B(T_1,T_2)}{\overline{B}(T_1,T_2)}-1\right)-\bar{s}B(T_1,T_2) \qquad \text{at } T_1. \tag{4.32}$$

The key term for the pricing of the credit spread forward is

$$\mathrm{E}\left[\beta_{T_1}\frac{B(T_1,T_2)}{\overline{B}(T_1,T_2)}\right]=\mathrm{E}\left[\beta_{T_1}\frac{B(T_1,T_2)}{\overline{B}(T_1,T_2)}1_{\{\tau>T_1\}}+\beta_{T_1}\frac{1}{c}1_{\{\tau\le T_1\}}\right], \tag{4.33}$$

under the equivalent recovery model. Using the Cox process properties of the default event equation (4.33) is readily simplified to

$$\mathrm{E}\left[\beta_{T_1}\frac{B(T_1,T_2)}{\overline{B}(T_1,T_2)}\right]=\mathrm{E}\left[e^{-\int_t^{T_1}r(s)+\lambda(s)ds}\frac{B(T_1,T_2)}{\overline{B}(T_1,T_2)}\right]+\frac{1}{c}(B(t,T_1)-\overline{B}_0(t,T_1)). \tag{4.34}$$

Summing up, under the equivalent recovery model the value of a credit spread forward agreement can be written as

$$D^{\mathrm{CSF}}=\frac{1}{T_2-T_1}\mathrm{E}\left[e^{-\int_t^{T_1}r(s)+\lambda(s)ds}\frac{B(T_1,T_2)}{\overline{B}(T_1,T_2)}\right] \tag{4.35}$$

$$+\frac{1}{T_2-T_1}\frac{1}{c^2}((1-c^2)B(t,T_1)-\overline{B}(t,T_1))-\bar{s}B(t,T_2). \tag{4.36}$$

The fair forward credit spread rate $\bar{s}$ for the credit spread forward is thus

$$\bar{s}=\frac{1}{B(t,T_2)(T_2-T_1)}\left(\mathrm{E}\left[e^{-\int_t^{T_1}r(s)+\lambda(s)ds}\frac{B(T_1,T_2)}{\overline{B}(T_1,T_2)}\right]\right.$$

$$+\frac{1}{c^2}((1-c^2)B(t,T_1)-\overline{B}(t,T_1))).$$

Under the fractional recovery model the key term (4.33) for the pricing of the credit spread forward resolves to

$$\begin{aligned}
\mathrm{E}\left[\beta_{T_1}\frac{B(T_1,T_2)}{\overline{B}(T_1,T_2)}\right] &= \mathrm{E}\left[\beta_{T_1}\frac{B(T_1,T_2)}{Q(T_1)\widetilde{B}(T_1,T_2)}\right] \\
&= \mathrm{E}\left[e^{\int_0^{T_1}\frac{q}{1-q}\lambda(s)-r(s)ds}\frac{B(T_1,T_2)}{\widetilde{B}(T_1,T_2)}\right]. \qquad (4.37)
\end{aligned}$$

In the step to equation (4.37) we used the following derivation of the expectation of $Q(T_1)^{-1}$, which is also an application of the Cox process properties of N for multiple jumps:

$$\begin{aligned}
\mathrm{E}\left[\frac{1}{Q(T)}\,\middle|\,\lambda(t)\,t\geq 0\right] &= \sum_{n=1}^{\infty}(1-q)^{-n}\mathrm{P}\left[\,N(T)=n\mid\lambda(t)\,t\geq 0\right] \\
&= \sum_{n=1}^{\infty}(1-q)^{-n}\frac{1}{n!}\left(\int_0^T\lambda(s)ds\right)^n e^{-\int_0^T\lambda(s)ds} \\
&= e^{-\int_0^T\lambda(s)ds}\sum_{n=1}^{\infty}\frac{1}{n!}\left(\frac{1}{1-q}\int_0^T\lambda(s)ds\right)^n \\
&= e^{\frac{q}{1-q}\int_0^T\lambda(s)ds}.
\end{aligned}$$

4.4.6 Credit Spread Put Options

Definition 11

A Credit Spread Put *on a defaultable bond $\overline{B}(t,T_2)$ with maturity $T_1<T_2$ and strike spread $\bar{s}$ gives the holder the right to sell the defaultable bond at time T_1 at a price that corresponds to a yield spread of $\bar{s}$ above the yield of an (otherwise identical) default-free bond $B(T_1,T_2)$.*

Define the exchange ratio $\bar{S}:=e^{-\bar{s}(T_2-T_1)}$. Then the credit spread put entitles the holder to exchange one defaultable bond $\overline{B}(T_1,T_2)$ for $\bar{S}$ default-free bonds $B(T_1,T_2)$ at time T_1. The payoff function is:

$$(\bar{S}B(T_1,T_2)-\overline{B}(T_1,T_2))^+. \qquad (4.38)$$

Depending on the specification the option can either survive a default or be knocked out by it. In the first case the payoff is conditioned on the full defaultable bond price $\overline{B}(T_1,T_2)$, and the default risk is borne by the writer of the security. The holder of the security is protected by the contract against any losses worse than a final credit spread of $\bar{s}$ and against all defaults but not against interest rate risk. The exchange option is therefore well suited to lay off most of the credit exposure in the underlying instrument while keeping some limited exposure to the credit spread of the underlying. We assume that the strike spread $\bar{s}$ is small enough (or the loss in default large enough) to ensure exercise of the option after a default.

Lemma 18
The price of the credit spread put is given by the following expectation:

$$\begin{aligned} D_t = & \mathbf{E}\left[e^{-\int_t^{T_1} r(s)+\lambda(s)ds} B(T_1, T_2)(\bar{S} - \tilde{P}(T_1, T_2))^+ \;\middle|\; \mathcal{F}_t \right] \\ & + \bar{S}B(t, T_2) - \overline{B}(t, T_2) \\ & - \mathbf{E}\left[e^{-\int_t^{T_1} r(s)+\lambda(s)ds} B(T_1, T_2)(\bar{S} - \tilde{P}(T_1, T_2)) \;\middle|\; \mathcal{F}_t \right]. \end{aligned} \tag{4.39}$$

The price of the credit spread put with knockout at default is given by

$$D_t = \mathbf{E}\left[e^{-\int_t^{T_1} r(s)+\lambda(s)ds} B(T_1, T_2)(\bar{S} - \tilde{P}(T_1, T_2))^+ \;\middle|\; \mathcal{F}_t \right]. \tag{4.40}$$

Although equations (4.39) and (4.40) do not represent a closed-form solution yet, they have several uses: Firstly, the problem is now accessible to the standard techniques of interest-rate theory. In the next section we will apply these techniques to derive full closed form solutions. Secondly, a numerical solution of equations (4.39) and (4.40) is feasible, again with standard techniques. In a Monte-Carlo simulation the elimination of the indicator functions $\mathbf{1}_{\{\tau > T_1\}}$ will speed up convergence significantly[14]. Finally, the analytical techniques can be also used as a pre-processing step for more complex problems (e.g. coupon bond options), before a numerical solution is attempted.

The credit spread put option price in equation (4.39) can be decomposed in two parts: A spread protection part (the first line) and a default protection part (the second and third lines). If the option is knocked out at default only the spread protection part remains.

Proof: The expectation to evaluate is (assuming no defaults so far $Q(t) = 1$)

$$\begin{aligned} D_t = & \mathbf{E}\left[\beta_{t,T_1}(\bar{S}B(T_1, T_2) - \overline{B}(T_1, T_2))^+ \mid \mathcal{F}_t \right] \\ = & \mathbf{E}\left[\mathbf{1}_{\{\tau > T_1\}} \beta_{t,T_1}(B(T_1, T_2)\bar{S} - \tilde{B}(T_1, T_2))^+ \;\middle|\; \mathcal{F}_t \right] \\ & + \mathbf{E}\left[\mathbf{1}_{\{\tau \le T_1\}} \beta_{t,T_1}(B(T_1, T_2)\bar{S} - Q(T_1)\tilde{B}(T_1, T_2))^+ \;\middle|\; \mathcal{F}_t \right], \end{aligned}$$

where we split up the payoff in the case without intermediate default ($\mathbf{1}_{\{\tau > T_1\}}$) and the case with intermediate default ($\mathbf{1}_{\{\tau \le T_1\}}$). If there has been an intermediate default, we assumed that the option will be exercised and we omit the maximum sign for that event:

$$\begin{aligned} = & \mathbf{E}\left[\mathbf{1}_{\{\tau > T_1\}} \beta_{t,T_1}(B(T_1, T_2)\bar{S} - \tilde{B}(T_1, T_2))^+ \;\middle|\; \mathcal{F}_t \right] \\ & + \mathbf{E}\left[\mathbf{1}_{\{\tau \le T_1\}} \beta_{t,T_1}(B(T_1, T_2)\bar{S} - Q(T_1)\tilde{B}(T_1, T_2)) \;\middle|\; \mathcal{F}_t \right], \end{aligned}$$

and then substitute $\mathbf{1}_{\{\tau \le T_1\}} = 1 - \mathbf{1}_{\{\tau > T_1\}}$ to reach

$$= \mathbf{E}\left[\mathbf{1}_{\{\tau > T_1\}} \beta_{t,T_1}(B(T_1, T_2)\bar{S} - \tilde{B}(T_1, T_2))^+ \;\middle|\; \mathcal{F}_t \right]$$

[14] The speed of convergence without default indicators is about 100 times faster than direct simulation.

$$
\begin{aligned}
&+ \mathrm{E}\left[\beta_{t,T_1}(B(T_1,T_2)\bar{S} - Q(T_1)\tilde{B}(T_1,T_2)) \;\Big|\; \mathcal{F}_t \right] \\
&- \mathrm{E}\left[1_{\{\tau>T_1\}}\beta_{t,T_1}(B(T_1,T_2)\bar{S} - Q(T_1)\tilde{B}(T_1,T_2)) \;\Big|\; \mathcal{F}_t \right] \\
=&\mathrm{E}\left[1_{\{\tau>T_1\}}\beta_{t,T_1}(B(T_1,T_2)\bar{S} - \tilde{B}(T_1,T_2))^+ \;\Big|\; \mathcal{F}_t \right] \\
&+ B(t,T_2)\bar{S} - \overline{B}(T_1,T_2) \\
&- \mathrm{E}\left[1_{\{\tau>T_1\}}\beta_{t,T_1}(B(T_1,T_2)\bar{S} - \tilde{B}(T_1,T_2)) \;\Big|\; \mathcal{F}_t \right].
\end{aligned}
$$

In the last line we used the fact that $1_{\{\tau>T_1\}}Q(T_1) = 1_{\{\tau>T_1\}}$, the indicator function of 'no default until T_1' switches off the 'influences $Q(T_1)$ of default until T_1'.

Next we have to eliminate the indicator functions from the expectations. This is done by conditioning on the realisation of λ and using the Cox process properties of the default triggering process. Using $\tilde{B}(T_1,T_2) = B(T_1,T_2)\tilde{P}(T_1,T_2)$ this yields:

$$
\begin{aligned}
&\mathrm{E}\left[1_{\{\tau>T_1\}}\beta_{t,T_1}B(T_1,T_2)(\bar{S} - \tilde{P}(T_1,T_2))^+ \;\Big|\; \mathcal{F}_t \right] \\
=&\mathrm{E}\left[\mathrm{E}\left[1_{\{\tau>T_1\}}\beta_{t,T_1}B(T_1,T_2)(\bar{S} - \tilde{P}(T_1,T_2))^+ \;\Big|\; \lambda \right] \;\Big|\; \mathcal{F}_t \right] \\
=&\mathrm{E}\left[e^{-\int_t^{T_1}\lambda(s)ds}\beta_{t,T_1}B(T_1,T_2)(\bar{S} - \tilde{P}(T_1,T_2))^+ \;\Big|\; \mathcal{F}_t \right].
\end{aligned}
$$

The second expectation is treated similarly to reach equation (4.39).

□

4.4.7 Put Options on Defaultable Bonds

The credit risk derivatives considered so far conditioned on the *spread* of the defaultable bond over a default-free bond, but the reference to a default-free bond is not necessary, we can also consider derivatives on the defaultable bond alone, e.g. a Put or call on a defaultable bond. Again we consider the cases when the Put covers default losses, and when it is killed at default.

Definition 12
A European put with maturity T_1 and strike K on a defaultable bond $\overline{B}(t,T_2)$ pays off

$$(K - \overline{B}(T_1,T_2))^+. \tag{4.41}$$

If the put is knocked out at default the payoff is

$$1_{\{\tau>T_1\}}(K - \overline{B}(T_1,T_2))^+. \tag{4.42}$$

This derivative protects the buyer against all risks: Interest rates, credit spreads or defaults that could move the defaultable bond price below K. Assuming exercise of the option in default we reach:

Lemma 19
The price of the Put option on a defaultable bond is

$$\begin{aligned} D_t =& \mathrm{E}\left[e^{-\int_t^{T_1} r(s)+\lambda(s)ds}(K - \widetilde{B}(T_1,T_2))^+ \;\middle|\; \mathcal{F}_t \right] \\ &+ B(t,T_1)K - \overline{B}(t,T_2) \\ &- \mathrm{E}\left[e^{-\int_t^{T_1} r(s)+\lambda(s)ds}(K - \widetilde{B}(T_1,T_2)) \;\middle|\; \mathcal{F}_t \right]. \end{aligned} \tag{4.43}$$

If the option is knocked out at default its price is

$$D_t = \mathrm{E}\left[e^{-\int_t^{T_1} r(s)+\lambda(s)ds}(K - \widetilde{B}(T_1,T_2))^+ \;\middle|\; \mathcal{F}_t \right]. \tag{4.44}$$

Proof: Replace $\bar{S}B(T_1,T_2)$ with $KB(T_1,T_1) = K$ in lemma 18. Then the proof carries through exactly as the proof to lemma 18.

□

Like the credit spread put option the plain put option can be decomposed in a default protection component and a price protection component. If the option is knocked out at default it has only the value of the price protection component.

It is important to note that the value of this option is *not* the value of an European Put on a bond in a world where the short rate is $\overline{r} = r + q\lambda$. (One might be lead to this view by the simple representation (4.10) of the defaultable bond prices.) This view would ignore the large jumps the bond price makes at defaults and only price the spread component.

To illustrate this point, we introduce a new security with the following payoff function under the fractional recovery model:

$$Q(T_1)(K - \widetilde{B}(T_1,T_2))^+. \tag{4.45}$$

At each default the face value of the defaultable bond is reduced by a factor $1 - q$. This payoff function is also reduced by the same factor, such that this option always covers exactly one defaultable bond, the size of the protection is adjusted to the 'size' (=face value) of the security. Again very similarly to (4.39) we reach the expectation

$$\mathrm{E}\left[e^{-\int_t^{T_1} r(s)+q\lambda(s)ds}(K - \widetilde{B}(T_1,T_2))^+ \;\middle|\; \mathcal{F}_t \right],$$

which is exactly identical to the expectation that has to be evaluated to price a put option on a zero coupon bond in a world where $\overline{r} = r + q\lambda$ is the riskless short rate. Unfortunately the specification of this payoff function heavily uses the fractional recovery model as input, specifically it conditions on loss quota (Q) and loss-quota adjusted price $\widetilde{B}$ separately. In real-world applications it would be impossible to separate these two quantities in a meaningful way, therefore the payoff function (4.45) remains an academic example.

4.5 Models

In the previous section we could derive some pricing formulae for credit derivatives without having to specify the dynamics we assume for credit spreads and interest rates. This was done by assuming independence of the realisations of r and λ.

For the pricing of credit derivatives with option features and to allow for correlation between r and λ we need a full specification of the dynamics of the default-free interest rate r and the intensity of the default process λ. A suitable specification should have the following properties:

- Both r and λ are stochastic. Stochastic default-free interest rates are indispensable for fixed-income analysis, and a stochastic default intensity is required to reach stochastic credit spreads which is necessary for meaningful prices for credit spread options.
- The dynamics of r and λ are rich enough to allow for a realistic description of the real-world prices. Duffie and Singleton [1997] and Duffee [1995] come to the conclusion that in many cases a multifactor model for the credit spreads is necessary.
- There should be scope to include correlation between credit spreads and default-free interest rates.
- It is desirable to have interest rates and credit spreads that remain positive at all times. Although negative credit spreads or interest rates represent an arbitrage opportunity, relaxing this requirement in favour of a Gaussian specification is still acceptable because of the analytical tractability that is gained. The Gaussian specification should then be viewed as a local approximation to the real-word dynamics rather than as a fully closed model. Furthermore, many important effects are more easily understood in the Gaussian setup.

Therefore we chose two alternative setups:

1. A *multifactor Gaussian* setup. Here there is the possibility of reaching negative credit spreads and interest rates with positive probability, but a high degree of analytical tractability is retained and the full term structures of bond prices and volatilities can be specified.
2. A *multifactor Cox Ingersoll Ross (CIR) [1985]* setup, following mainly Jamshidian [1996].[15] This model setup gives us the required properties while still retaining a large degree of analytical tractability. Furthermore, models of credit spreads of the CIR square-root type have been estimated by Duffie and Singleton [1997] and Duffee [1995].

Both specifications use the fractional recovery model, but most the results can be transferred to the equivalent recovery model: By assuming full loss in default $q = 1$, the specification can be viewed as a specification of the dynamics of zero-recovery defaultable bond prices $\overline{B}_0(t, T)$. Combining these dynamics of the zero-recovery defaultable bond prices with an assumption on the equivalent recovery rate c will yield a specification in terms of the equivalent recovery model where defaultable bond prices are given by $\overline{B}(t, T) = cB(t, T) + (1 - c)\overline{B}(t, T)$.

[15] Related models can also be found in Jamshidian [1987, 1995], Duffie and Kan [1996], Chen and Scott [1995] and Longstaff and Schwartz [1992].

4.6 The Multifactor Gaussian Model

In this section we present the setup of the Gaussian specification of the fractional recovery model. This specification is problematic as it allows credit spreads and interest rates to become negative with positive probability. This problem is well known from default-free Gaussian interest-rate models and it is partly compensated by the analytical tractability gained, provided the probability of these events remains low. This setup should therefore be viewed as an approximation to more realistic models of credit spread and interest rates, and despite this drawback it will yield some valuable insights.

Assumption 12
The following asset price dynamics under the martingale measure are given:

- *default-free bond prices* $B(t,T)$
$$\frac{dB(t,T)}{B(t,T)} = r(t)dt + a(t,T)dW,$$
- *and defaultable bond prices* $\overline{B}(t,T) = Q(t)\widetilde{B}(t,T)$
$$dQ(t) = -Q(t-)q dN(t)$$
$$\frac{d\widetilde{B}(t,T)}{\widetilde{B}(t,T)} = (r(t) + q\lambda(t))dt + \overline{a}(t,T)dW.$$

The Brownian motion dW *is d-dimensional, the volatilities* $a(t,T)$ *and* $\overline{a}(t,T)$ *are d-dimensional deterministic functions of time t and time to maturity T only, and at* $t = 0$ *there has been no default:* $\tau > 0$ *and* $Q(0) = 1$ *and* $\widetilde{B}(0,T) = \overline{B}(0,T)$.

This specification is equivalent to a Gaussian HJM model with the following forward rate volatilities for the bond prices:

$$\sigma(t,T) = -\frac{\partial}{\partial T}a(t,T) \qquad \overline{\sigma}(t,T) = -\frac{\partial}{\partial T}\overline{a}(t,T)$$
$$\sigma^h(t,T) = -\frac{\partial}{\partial T}(\overline{a}(t,T) - a(t,T)).$$

This setup can be used to model any degree of correlation and de-correlation that is desired. Multiplication of the volatilities is defined by the scalar product in $\mathbb{R}^d$: $a\overline{a} = \sum_{i=1}^{d} a_i\overline{a}_i$.

4.7 Credit Derivatives in the Gaussian Model

4.7.1 Implied Survival Probabilities

In the Gaussian setup we can derive the implied survival probability from t to T in closed form. Given $\tau > t$ the survival probability is defined as:

$$P(t,T) = \mathrm{E}\left[e^{-\int_t^T \lambda(s)ds} \;\middle|\; \mathcal{F}_t \right].$$

Lemma 20
The dynamics of the survival probabilities are

$$\frac{dP(t,T)}{P(t,T)} = \lambda(t)dt + a^p(t,T)dW \tag{4.46}$$

$$a^p(t,T) = \frac{1}{q}\left(\overline{a}(t,T) - a(t,T)\right), \tag{4.47}$$

and the initial values $P(0,T)$ are given by:

$$P(0,T) = \left(\frac{\overline{B}(0,T)}{B(0,T)}\right)^{\frac{1}{q}}$$
$$\exp\left\{-\frac{1}{2q^2}\int_0^T (\overline{a}(s,T) - a(s,T))[(1+q)a(s,T) - (1-q)\overline{a}(s,T))]ds\right\}. \tag{4.48}$$

or

$$P(0,T) = \left(\frac{\overline{B}(0,T)}{B(0,T)}\right)^{\frac{1}{q}}$$
$$\exp\left\{-\frac{1}{2}\int_0^T a^p(s,T)[a(s,T) + \overline{a}(s,T) - a^p(s,T)]ds\right\}. \tag{4.49}$$

Proof: See appendix A.1 on page 153.

□

For zero recovery (full loss $q = 1$ in default) equation (4.48) reduces to

$$P(0,T) = \frac{\overline{B}(0,T)}{B(0,T)}\exp\{-\int_0^T a^p(s,T)a(s,T)ds\}. \tag{4.50}$$

Under independence of credit spreads and interest rates the implied survival probability would further reduce to $\frac{\overline{B}(0,T)}{B(0,T)}$ as it was given in section 4.3.5. Hence the factor

$$e^{-\int_0^T a^p(s,T)a(s,T)ds}$$

represents the influence of correlation between spreads and interest rates on the implied default probabilities. There is an intuitive explanation of the direction of the effect:

If interest rates and credit spreads are positively correlated ($a^p a > 0$) this means that defaults are slightly more likely in states of nature when interest rates are high. Because of the higher interest rates these states are discounted more strongly when they enter the price of the defaultable bond, and conversely states with low interest rates enter with less discounting and simultaneously fewer defaults. To reach a *given* price for a defaultable bond, the absolute default likelihood must therefore be higher. This implies a lower survival probability which is also what equation (4.50) yields for $a^p a > 0$. The argument runs conversely for negative correlation $a^p a < 0$.

4.7.2 The Survival Contingent Measure

The following lemma describes the change of measure that is necessary to value survival contingent payoffs in this framework.

Lemma 21
Let $X \geq 0$ be a $\mathcal{F}_T$-measurable random variable and $\tau > t$. In the Gaussian model framework, the time t value of receiving X at T given no default has happened before T is

$$\mathbf{E}^Q\left[\,\beta_{t,T}\mathbf{1}_{\{\tau>T\}}X \mid \mathcal{F}_t\,\right] = \overline{B}_0(t,T)\mathbf{E}^{P_S}\left[\,X \mid \mathcal{F}_t\,\right], \tag{4.51}$$

where

$$\overline{B}_0(t,T) = B(t,T)P(t,T)\exp\{\int_t^T a^p(s,T)a(s,T)ds\} \tag{4.52}$$

$$= \left(\frac{\overline{B}(t,T)}{B(t,T)^{1-q}}\right)^{\frac{1}{q}} \exp\left\{\frac{1-q}{2q}\int_t^T (\overline{a}(s,T)-a(s,T))^2 ds\right\}. \tag{4.53}$$

The measure P_S is called the survival contingent measure. *It is defined by the Radon-Nikodym density*

$$dP_{ST} = M_T dQ_T \tag{4.54}$$

where $dP_{ST} = dP_S|\,\mathcal{F}_T$ and $dQ_T = dQ|\,\mathcal{F}_T$, and

$$M_u = \mathcal{E}\left(\int_t^u (a^p(s,T)+a(s,T))dW_s\right), \qquad u \geq t. \tag{4.55}$$

The Q-Brownian motion W_s is transformed into a P_S-Brownian motion via

$$dW_s = dW_s^{P_S} + (a^p(s,T)+a(s,T))ds. \tag{4.56}$$

Proof: See appendix A.2 on page 154.

□

If the survival contingent measure is known, the valuation of the payoff X can be decoupled from the default valuation which is represented by $\overline{B}_0(t,T)$. The numeraire used in this change of measure is

$$\mathbf{E}\left[e^{-\int_t^T r(s)+\lambda(s)ds} \;\middle|\; \mathcal{F}_t\right]$$

which is almost exactly the price of a zero-recovery defaultable bond:

$$\overline{B}_0(t,T) = \mathbf{1}_{\{\tau>t\}}\mathbf{E}\left[e^{-\int_t^T r(s)+\lambda(s)ds} \;\middle|\; \mathcal{F}_t\right].$$

For no previous defaults ($\tau > t$) both expressions coincide. In equations (4.52) and (4.53) the lemma also gives the value of a zero-recovery defaultable bond with maturity T in the fractional recovery setup.

Using lemma 20 we can now derive the values of some credit derivatives that were introduced in section 4.2.

4.7.3 Default Digital Payoffs

The price for a default digital put is according to equation (4.22)

$$D = \int_0^T \mathrm{E}\left[\lambda(t) e^{-\int_0^t \lambda(s)ds} e^{-\int_0^t r(s)ds} \right] dt.$$

The resulting price for the default digital put is:

Proposition 22
The price of a default digital put with maturity T and payoff 1 at default is in the multifactor Gaussian model framework

$$D = \int_0^T P(0,t)B(0,t)e^{\int_0^t a(s,t)a^P(s,t)ds} \left[\lambda(0,t) + \int_0^t a(s,t)\sigma^P(s,t)ds\right] dt \qquad (4.57)$$

where $\lambda(0,t) := -\frac{\partial}{\partial t}\ln P(0,t)$ *and* $\sigma^P(s,t) := -\frac{\partial}{\partial t}a^P(s,t)$ *are given in lemma 20.*

Proof: See appendix A.3 on page 155.

□

4.7.4 The Credit Spread Put

Proposition 23
The price of the Credit Spread Component of the Credit Spread Put is given by:

$$\begin{aligned} C^{CSP} =& N(d_1)\ \bar{S}\, B(t,T_2)\, P(t,T_1) \exp\{\int_t^{T_1} a(s,T_2)a^P(s,T_1)ds\} \\ &- N(d_2)\ \overline{B}(t,T_2)\, (P(t,T_1))^{1-q} \exp\{\int_t^{T_1} (1-q)a^P(s,T_1)(\overline{a}(s,T_2) - \frac{1}{2}qa^P(s,T_1))ds\} \end{aligned}$$

where

$$d_1 = \frac{K + F + \frac{1}{2}V}{\sqrt{V}} \qquad \text{and} \qquad d_2 = d_1 - \sqrt{V}.$$

The other parameters are:

$$\begin{aligned} K &= \ln \bar{S} + \ln \frac{\overline{B}(t,T_1)}{B(t,T_1)} - \ln \frac{\overline{B}(t,T_2)}{B(t,T_2)} \\ F &= \int_t^{T_1} qa^P(s,T_1)\left[a(s,T_2) - a(s,T_1) - (1-q)(a^P(t,T_2) - a^P(t,T_1))\right] ds \\ V &= \int_t^{T_1} q^2(a^P(s,T_2) - a^P(s,T_1))^2\, ds. \end{aligned}$$

Proof: See appendix A.4 on page 156.

□

The parameters of this pricing formula have a number of similarities to classical Black-Scholes option prices.
The parameter K is similar to the logarithm of *'share/strike'* as it is found in the exchange option. Here we have the ratio of 'forward spread' and strike spread instead but the function is the same. The parameter V gives the 'volatility' of the payoff. Here it is the volatility of the forward credit spread over the interval $[T_1, T_2]$.
F is a correlation parameter. Under zero recovery ($q = 1$) it reduces to zero if the volatilities of spreads a^p and default-free bond prices a generate independent dynamics.
Finally, the form of the representation of the points d_1, d_2, at which the normal distribution is evaluated, are similar to the classical Black-Scholes parameters.

The similarities to the classical Black-Scholes formula are not surprising if one bears in mind that – given that no default happens – under the multifactor Gaussian setup defaultable and default free bond prices are lognormally distributed.

For completeness we also give the price of the default-protection component of the credit spread put:

Proposition 24
The price of the default protection component of the credit spread put is given by

$$C^{def} = \bar{S}\ B(t,T_2)(1 - \ P(t,T_1)e^H) - \overline{B}(t,T_2)(1 - \left(P(t,T_1)e^H\right)^{1-q})$$

where

$$e^H = \exp\{\int_t^{T_1} a(s,T_2)a^P(s,T_1)ds\}.$$

The default protection component of the credit spread put gives the holder the right to put a *defaulted* bond $\overline{B}(t,T_2)$ to the seller of the option at the price of $\bar{S}$ default-free bonds $B(t,T_2)$. If $\bar{S}$ is close to 1 this protection can be approximated by a default swap.

4.7.5 The Put on a Defaultable Bond

Proposition 25
The price of a European put option with strike $\bar{S}$, expiry T_1 and knockout at default on a defaultable bond $\overline{B}(t,T_2)$ is (under the fractional recovery model)

$$\begin{aligned} C^{put} &= \bar{S}\overline{B}_0(t,T_1)N(d_1) \\ &- \overline{B}(t,T_2)P(t,T_1)^{1-q}\exp\{(1-q)\int_t^{T_1} a^P(s,T_1)(\bar{a}(s,T_2) - \frac{1}{2}qa^P(s,T_1))ds\}N(d_2) \end{aligned}$$

where

$$d_1 = \frac{K + F_1 + \frac{1}{2}V}{\sqrt{V}} \qquad d_2 = \frac{K + F_2 - \frac{1}{2}V}{\sqrt{V}}$$

$$K = \ln \frac{\bar{S}\overline{B}(t,T_1)}{\overline{B}(t,T_2)}$$

$$F_1 = -\int_t^{T_1} (\bar{a}(s,T_2) + a(s,T_1) + a^P(s,T_1))(\bar{a}(s,T_2) - \bar{a}(s,T_1))ds$$

$$F_2 = F_1 + q\int_t^{T_1} (\bar{a}(s,T_2) - \bar{a}(s,T_1))a^P(s,T_1)ds$$

$$V = \int_t^{T_1} (\bar{a}(s,T_2) - \bar{a}(s,T_1))^2 ds.$$

Proof: See appendix A.5 on page 159.

□

Under the equivalent recovery model the put option on a defaultable bond is equivalent to a put option on a *portfolio* consisting of c default-free bonds $B(t,T_2)$ and $(1-c)$ defaultable bonds with zero recovery $\overline{B}_0(t,T_2)$. In this case the valuation of the option will become much more difficult and either approximations or numerical methods will have to be employed.

4.8 The Multifactor CIR Model

The multifactor CIR model is set up as follows:

Assumption 13
Interest rates and default intensities are driven by n independent factors x_i, $i = 1,\dots,n$ with dynamics of the CIR square-root type:

$$dx_i = (\alpha_i - \beta_i x_i)dt + \sigma_i\sqrt{x_i}dW_i. \tag{4.58}$$

The coefficients satisfy $\alpha_i > \frac{1}{2}\sigma_i^2$ to ensure strict positivity of the factors.

The default-free short rate r and the default intensity λ are positive linear combinations of the factors x_i with weights $w_i \geq 0$ and $\overline{w}_i$, $(0 \leq i \leq n)$ respectively:

$$r(t) = \sum_{i=1}^{n} w_i x_i(t). \tag{4.59}$$

$$\lambda(t) = \sum_{i=1}^{n} \overline{w}_i x_i(t). \tag{4.60}$$

With all coefficients $w_i \geq 0$ and $\overline{w}_i \geq 0$ nonnegative we have ensured that $r > 0$ and $\lambda > 0$ almost surely. Unfortunately this specification can only generate *positive* correlation between r

and λ. If negative correlation is needed one could define modified factors x_i' that are negatively correlated to the x_i by $dx_i' = (\alpha_i - \beta_i x_i')dt - \sigma_i\sqrt{x_i'}dW_i$ (note the minus in front of the Brownian motion). This would complicate the analysis. Alternatively one could restrict the specification to a squared Gaussian model.

Typically only the first $m < n$ factors would describe the default-free term structure (i.e. $w_i = 0$ for $i > m$), and the full set of state variables would be used to describe the spreads. The additional $n - m$ factors for the spreads ensure that the dynamics of the credit spreads have components that are independent of the default-free interest rate dynamics. The simplest example would be independence between r and λ, where $r = x_1$ and $\lambda = x_2$: one factor driving each rate,

$$\begin{aligned} n &= 2 & m &= 1 \\ w_1 &= 1 & w_2 &= 0 \\ \overline{w}_1 &= 0 & \overline{w}_2 &= 1. \end{aligned}$$

4.8.1 Bond Prices

For a linear multiple cx_i ($c > 0$ is a positive constant) of the factor x_i the following equation gives the corresponding 'bond price' (see CIR [1985]):

$$\mathrm{E}\left[\exp\{-\int_t^T cx_i(s)ds\} \,\middle|\, \mathcal{F}_t\right] = H_{1i}(T-t,c)e^{-H_{2i}(T-t,c)cx_i} \tag{4.61}$$

where

$$H_{1i}(T-t,c) = \left[\frac{2\gamma_i e^{\frac{1}{2}(\gamma_i+\beta_i)(T-t)}}{(\gamma_i+\beta_i)(e^{\gamma_i(T-t)}-1)+2\gamma_i}\right]^{2\alpha_i/\sigma_i^2} \tag{4.62}$$

$$H_{2i}(T-t,c) = \frac{2(e^{\gamma_i(T-t)}-1)}{(\gamma_i+\beta_i)(e^{\gamma_i(T-t)}-1)+2\gamma_i} \tag{4.63}$$

$$\gamma_i = \sqrt{\beta_i^2 + 2c\sigma_i^2}. \tag{4.64}$$

The default-free bond prices are given by

$$\begin{aligned} B(t,T) &= \mathrm{E}\left[e^{-\int_t^T r(s)ds} \,\middle|\, \mathcal{F}_t\right] = \mathrm{E}\left[e^{-\sum_i \int_t^T w_i x_i(s)ds} \,\middle|\, \mathcal{F}_t\right] \\ &= \mathrm{E}\left[\prod_i e^{-\int_t^T w_i x_i(s)ds} \,\middle|\, \mathcal{F}_t\right] = \prod_i \mathrm{E}\left[e^{-\int_t^T w_i x_i(s)ds} \,\middle|\, \mathcal{F}_t\right] \end{aligned}$$

because the factors are independent. The bond price is thus a product of one-factor bond prices

$$B(t,T) = \prod_{i=1}^{n} H_{1i}(T-t,w_i)e^{-H_{2i}(T-t,w_i)w_i x_i(t)}. \tag{4.65}$$

Similarly, the defaultable bond prices under fractional recovery are given by:

$$\overline{B}(t,T) = Q(t)\prod_{i=1}^{n} H_{1i}(T-t,w_i+q\overline{w}_i)e^{-H_{2i}(T-t,w_i+q\overline{w}_i)(w_i+q\overline{w}_i)x_i}, \tag{4.66}$$

because of $\overline{B}(t,T) = Q(t)\mathrm{E}\left[\exp\{-\int_t^T r(s) + q\lambda(s)ds\} \mid \mathcal{F}_t\right]$, zero recovery bond prices can be derived from (4.66) by setting $q = 1$.

4.8.2 Affine Combinations of Independent Non-Central Chi-Squared Distributed Random Variables

The mathematical tools for the analysis of the model have been provided by Jamshidian [1996], who used them to price interest-rate derivatives in a default-free interest rate environment. Of these tools we need the expressions for the evaluation of the expectations that will arise in the pricing equations, and the methodology of a change of measure to remove discount factors from these expectations.

First the distribution function of an affine combination of noncentral chi-squared random variables is presented. Most of the following expressions are based upon this distribution function. Then we consider the expressions for the expectations of noncentral chi-squared random variables, and finally the change-of-measure technique that has to be applied in this context.

Definition 13

Let z_i, $1 \le i \le n$ be n independent, noncentral chi-square distributed random variables with ν_i degrees of freedom and noncentrality parameter[16] $\bar{\lambda}_i$.

Let Y be an affine combination of the random variables z_i, $1 \le i \le n$ with weights η_i and offset ϵ:

$$Y = \epsilon + \sum_{i=1}^{n} \eta_i z_i. \tag{4.67}$$

We call Y a Affine combination of Noncentral Chi-squared random variables (ANC) *and denote its distribution function with χ_n^2:*

$$\mathrm{P}[\, Y \le y \,] =: \chi_n^2(\, y\,;\, \nu, \bar{\lambda}, \eta, \epsilon),$$

where η, ν and $\bar{\lambda}$ are vectors giving weight, degrees of freedom and noncentrality of the z_i, and for $\epsilon = 0$ we will omit the last argument[17].

The evaluation of this distribution function can be efficiently implemented via a fast Fourier transform of its characteristic function (see e.g. Chen and Scott (1992)). Alternatively, Jamshidian gives the following formula (which is basically the transform integral of the characteristic function for $\eta > -\frac{1}{2}$):

$$\chi^2(y;\, \nu, \bar{\lambda}, \eta, 0) = \frac{1}{2} + \frac{1}{\pi}\int_0^{\infty} \Psi(\frac{\nu}{2}, \bar{\lambda}, 2\xi^2\eta^2)\, \sin(\xi y - \Phi(\nu, \bar{\lambda}, \xi\eta))\frac{d\xi}{\xi}, \tag{4.68}$$

where $\Psi(\nu, \bar{\lambda}, \eta) = \mathrm{E}\left[\, e^{-Y}\,\right]$ will be given in equation (4.72) and

$$\Phi(\nu, \bar{\lambda}, \gamma) = \sum_{i=1}^{n}\left(\frac{\nu_i}{2}\arctan(2\gamma_i) + \frac{\gamma_i\bar{\lambda}_i}{1+4\gamma_i^2}\right).$$

[16] The noncentrality parameter $\bar{\lambda}$ is not to be confused with the intensity λ of the defaults.

[17] This is justified by equation (4.71).

The integral in (4.68) is numerically well-behaved in that the limit of the integrand as $\xi \to 0$ is finite and the integrand is absolutely integrable. Using equation (4.71) this can be extended for the case of $\epsilon \neq 0$.

One advantage of this setup is that – although we are working with a n-factor model – most valuation problems can be reduced to the evaluation of this one-dimensional integral. The evaluation of the noncentral chi-squared distribution function which one encounters in the one-factor case, also requires a numerical approximation, the implementation effort of the multifactor model therefore does not seem to be very much higher than the effort required for a model with one independent factor each for interest rates and spreads.

The following lemma by Jamshidian [1996] provides most of the expressions we need:

Lemma 26
Let Y and Y' be ANC distributed with the same z_i, but Y with weights η and offset ϵ, and Y' ANC with weights η' and offset ϵ'.

$$Y \sim \chi_n^2(\,\cdot\,;\,\nu,\,\tilde{\lambda},\,\eta,\,\epsilon), \qquad\qquad Y' \sim \chi_n^2(\,\cdot\,;\,\nu,\,\tilde{\lambda},\,\eta',\,\epsilon'). \tag{4.69}$$

Let y be a constant. Then
the expectation of Y is:

$$\mathrm{E}[\,Y\,] = \epsilon + \sum_{i=1}^{n} \eta_i(\tilde{\lambda}_i + \nu_i), \tag{4.70}$$

the distribution of Y is:

$$\mathrm{P}[\,Y \leq y\,] = \chi_n^2(y - \epsilon;\ \nu, \tilde{\lambda}, \eta), \tag{4.71}$$

the expectation of e^{-Y} is:

$$\mathrm{E}\left[\,e^{-Y}\,\right] = e^{-\epsilon} \prod_{i=1}^{n} \frac{1}{(1+2\eta_i)^{\nu_i/2}} \exp\left\{ -\frac{\eta_i \tilde{\lambda}_i}{1+2\eta_i} \right\} \tag{4.72}$$

the value of a call option on e^{-Y} with strike e^{-y} is:

$$\begin{aligned} \mathrm{E}\left[\,(e^{-Y} - e^{-y})^+\,\right] &= \mathrm{E}\left[\,e^{-Y}\,\right]\ \chi_n^2\left(y - \epsilon; \nu, \frac{\tilde{\lambda}}{1+2\eta}, \frac{\eta}{1+2\eta} \right) \\ &\quad - e^{-y}\ \chi_n^2(y - \epsilon; \nu, \tilde{\lambda}, \eta) \end{aligned} \tag{4.73}$$

the value of an exchange option on e^{-Y} and $e^{-Y'}$ is:

$$\begin{aligned} \mathrm{E}\left[\,(e^{-Y} - e^{-Y'})^+\,\right] &= \mathrm{E}\left[\,e^{-Y}\,\right]\ \chi_n^2\left(\epsilon' - \epsilon; \nu, \frac{\tilde{\lambda}}{1+2\eta}, \frac{\eta - \eta'}{1+2\eta} \right) \\ &\quad - \mathrm{E}\left[\,e^{-Y'}\,\right]\ \chi_n^2\left(\epsilon' - \epsilon; \nu, \frac{\tilde{\lambda}}{1-2\eta'}, \frac{\eta - \eta'}{1+2\eta'} \right). \end{aligned} \tag{4.74}$$

Proof: The expectation (4.70) and the moment generating function (4.72) are well-known (see e.g. Johnson and Kotz [1970]), the statement (4.71) follows directly from the definition of the probability distribution function. For a proof of (4.73) and (4.74) see Jamshidian [1996].

□

4.8.3 Factor Distributions

As observed by CIR [1985, 1985], the square-root dynamics of the factors x_i give rise to non-central chi-square distributed final values.

Lemma 27
Let x be given by

$$dx = (\alpha - \beta x)dt + \sigma\sqrt{x}dW. \tag{4.75}$$

Then $x(T)$ given $x(t)$ is ANC distributed with weight

$$\eta = \frac{\sigma^2}{4\beta}(1 - e^{-\beta(T-t)})$$

and noncentrality $\tilde{\lambda}$ and degrees of freedom ν

$$\tilde{\lambda} = x(t)\frac{4\beta e^{-\beta(T-t)}}{\sigma^2(1 - e^{-\beta(T-t)})} \qquad \nu = \frac{4\alpha}{\sigma^2}.$$

The distribution of the factors $x_i(T)$ remains of the ANC type even under a change of measure to a T-forward measure. This change of measure will be necessary to eliminate discounting with the factors later on, it is defined in lemma 28 which is Girsanov's theorem (points (i) and (ii)) combined with a slight extension of results by Jamshidian [1987, 1996] (poins (iii) and (iv)).

Lemma 28
Let x follow a CIR-type square-root process under the measure P:

$$dx = (\alpha - \beta x)dt + \sigma\sqrt{x}\, dW, \tag{4.76}$$

such that $x = 0$ is an unattainable boundary ($\alpha > \frac{1}{2}\sigma^2$) and let $c > 0$ be a positive real number.

(i) Then there is an equivalent probability measure $\tilde{P}_c$, whose restriction on $\mathcal{F}_t$ has the following Radon-Nikodym density w.r.t. P

$$\frac{d\tilde{P}_{c_t}}{dP_t} := Z(t) = \mathcal{E}\left(-\int_0^t \sigma(s)c\sqrt{x(s)}H_{2x}(T-s,c)dW(s)\right).$$

and under which the process $\tilde{W}_t^c$

$$dW_t = d\tilde{W}_t^c - H_{2x}(T-t,c)c\sigma\sqrt{x}dt.$$

is a $\tilde{P}_c$-Brownian motion.

(ii) Expectations under P are transformed to expectations under $\tilde{P}_c$ via

$$\mathrm{E}^P\left[e^{-\int_t^T cx(s)ds} F(x(T)) \,\Big|\, \mathcal{F}_t \right] = G(x(t), t, T, c)\mathrm{E}^{\tilde{P}_c}\left[F(x(T)) \mid \mathcal{F}_t \right] \tag{4.77}$$

where

$$G(x,t,T,c) = \mathrm{E}\left[e^{-\int_t^T cx(s)ds} \,\Big|\, \mathcal{F}_t \right] = H_{1x}(T-t,c)e^{-H_{2x}(T-t,c)cx}. \tag{4.78}$$

(iii) Under $\tilde{P}_c$ the process x has the dynamics

$$dx = [\alpha - (\beta + H_{2x}(T-t,c)c\sigma^2)x]dt + \sigma\sqrt{x}\, d\tilde{W}^c, \tag{4.79}$$

and $x(T)$ given $x(t)$ is ANC distributed under $\tilde{P}_c$ with weight η_T

$$\eta_T = \frac{c\sigma^2}{4} H_{2x}(T-t,c) \tag{4.80}$$

and ν_T degrees of freedom and noncentrality parameter $\tilde{\lambda}_T$:

$$\nu_T = \frac{4\alpha}{\sigma^2} \tag{4.81}$$

$$\tilde{\lambda}_T = \frac{4}{\sigma^2} \frac{\frac{\partial}{\partial T} H_{2x}(T-t,c)}{H_{2x}(T-t,c)} x(t). \tag{4.82}$$

(iv) The dynamics of the other factors are unaffected.

Proof: See appendix B.1

□

The lemma also holds for time-dependent parameters with $\alpha(t)/\sigma^2(t) > 1/2$, but here we only need the time-independent case.

The change of measure to $\tilde{P}_c^Z$ removes the discounting with $e^{-\int_0^T cx(t)dt}$, while changing the distribution of $x(T)$ to a ANC distribution with parameters given in equations (4.80) to (4.82).

4.9 Credit Derivatives in the CIR Model

Using the results from section 4.8 we are now able to give solutions for the general pricing formulae for credit derivatives that were derived in section 4.4.

4.9.1 Default Digital Payoffs

In section 4.4.2 the price for a default digital put was given in equation (4.22) as

$$D = \int_0^T \mathrm{E}\left[\lambda(t)e^{-\int_0^t \lambda(s)ds} e^{-\int_0^t r(s)ds} \right] dt.$$

The resulting price for the default digital put is:

Proposition 29
The price of a default digital put with maturity T and payoff 1 at default is in the CIR model framework

$$D = \int_0^T \mathrm{E}\left[\lambda(t)e^{-\int_0^t \lambda(s)ds}e^{-\int_0^t r(s)ds}\right]dt \tag{4.83}$$

$$= \int_0^T \left(\sum_{i=1}^n \overline{w}_i(w_i+\overline{w}_i)(\alpha_i H_{2i}(t,w_i+\overline{w}_i) + \frac{\partial H_{2i}(t,w_i+\overline{w}_i)}{\partial t}x_i(0))\right)\prod_{j=1}^n \overline{B}_{j0}(0,t)dt. \tag{4.84}$$

where for $1 \le j \le n$

$$\overline{B}_{j0}(0,t) = H_{1j}(t,w_j+\overline{w}_j)e^{-H_{2j}(t,w_j+\overline{w}_j)(w_j+\overline{w}_j)x_j}. \tag{4.85}$$

Proof: See appendix B.2 on page 164.

□

4.9.2 Credit Derivatives with Option Features

Using the results of lemma 18 we can derive the price of a credit spread put option in the CIR framework. Here, according to equation (4.40), we have to evaluate

$$D_t = \mathrm{E}\left[e^{-\int_t^{T_1} r(s)+\lambda(s)ds}B(T_1,T_2)(\bar{S}-\bar{P}(T_1,T_2))^+ \;\middle|\; \mathcal{F}_t\right].$$

The result is given in the following proposition:

Proposition 30

(i) The price of a credit spread option with maturity T_1, strike spread $\bar{s}$ and knockout at default on a defaultable zero coupon bond with maturity T_2 is (in the CIR model specification)

$$D^{CSP} = \overline{B}_0(t,T_1)\left\{\mathrm{E}\left[e^{-Y}\right]\chi^2_n\left(\epsilon'-\epsilon;\nu,\frac{\bar{\lambda}}{1+2\eta},\frac{\eta-\eta'}{1+2\eta}\right)\right.$$
$$\left.-\mathrm{E}\left[e^{-Y'}\right]\chi^2_n\left(\epsilon'-\epsilon;\nu,\frac{\bar{\lambda}}{1+2\eta'},\frac{\eta-\eta'}{1+2\eta'}\right)\right\}. \tag{4.86}$$

(ii) The price of a put option with maturity T_1, strike $\bar{S} = e^{-y}$ and knockout at default on a defaultable zero coupon bond with maturity T_2 is (in the CIR model specification)

$$D^{Put} = \overline{B}_0(t,T_1)\left\{e^{-y}\chi^2_n\left(\epsilon'-y;\nu,\bar{\lambda},-\eta'\right)\right.$$
$$\left.-\mathrm{E}\left[e^{-Y'}\right]\chi^2_n\left(\epsilon'-y;\nu,\frac{\bar{\lambda}}{1+2\eta'},-\frac{\eta'}{1+2\eta'}\right)\right\}. \tag{4.87}$$

The variables in these formulae are:

$$Y \sim \chi_n^2(\cdot; \nu, \bar{\lambda}; \eta, \epsilon) \tag{4.88}$$

$$Y' \sim \chi_n^2(\cdot; \nu, \bar{\lambda}; \eta', \epsilon') \tag{4.89}$$

$$\nu_i = \frac{4\alpha_i}{\sigma_i^2} \tag{4.90}$$

$$\bar{\lambda}_i = \frac{4\frac{\partial}{\partial T}H_{2i}(T_1 - t, w_i + \overline{w}_i)}{\sigma_i^2 H_{2i}(T_1 - t, w_i + \overline{w}_i)} x_i(t) \tag{4.91}$$

$$\epsilon = -\ln \bar{S} - \sum_{i=1}^{n} \ln H_{1i}(T_2 - T_1, w_i) \tag{4.92}$$

$$\eta_i = \frac{\sigma_i^2}{4} w_i(w_i + \overline{w}_i) H_{2i}(T_2 - T_1, w_i) H_{2i}(T_1 - t, w_i + \overline{w}_i) \tag{4.93}$$

$$\epsilon' = -\sum_{i=1}^{n} \ln H_{1i}(T_2 - T_1, w_i + q\overline{w}_i) \tag{4.94}$$

$$\eta_i' = \frac{\sigma_i^2}{4}(w_i + \overline{w}_i)(w_i + q\overline{w}_i) H_{2i}(T_2 - T_1, w_i + q\overline{w}_i) H_{2i}(T_1 - t, w_i + \overline{w}_i) \tag{4.95}$$

and $\mathrm{E}\left[e^{-Y} \right]$ *is defined in equation (4.72).*

Proof: See appendix B.3 on page 165.

□

In these pricing formulae there are negative signs in front of the last arguments of the ANC distribution function $\chi_n^2(\cdot)$. This is not a problem because $\chi_n^2(\cdot)$ is well-defined as long as these parameters are still larger than $-\frac{1}{2}$. For practical applications this is usually the case, because the η_i and η_i' contain a factor in σ_i^2 (which is very small) and the other factors in η_i and η_i' are not large for a realistic specification.

4.10 Conclusion

In this chapter we demonstrated the Cox process approach to the pricing of credit risk derivatives, and some closed form solutions were given for the case of a multifactor CIR-specification of the credit spread and interest rate dynamics.

The Cox process modelling approach derives its tractability from the fact that the pricing of derivatives can be done in stages:

First, by conditioning on the realisation of the intensity process λ the pricing problem can be reduced to the fairly straightforward case of inhomogeneous Poisson processes, for which most expected values are easily calculated. Thus the explicit reference to default events can be eliminated and replaced by expressions in the intensity of the default process. After this stage the methods of default-free interest rate theory can be applied. because the reference to jump processes has disappeared.

We used specifically the change of measure technique in several dimensions to remove discount factors of the form $\exp - \int_0^T g(s)ds$ where g is an 'interest-rate like' process, and then some well-known properties of the final distributions of the factor processes. In most cases we needed to change the measure to remove the discount factor

$$e^{-\int_0^T r(t)+\lambda(t)dt}$$

which typically arose from expectations of the form

$$\mathrm{E}\left[e^{-\int_0^T r(t)dt} 1_{\{\tau>T\}} X \right] = \mathrm{E}\left[e^{-\int_0^T r(t)+\lambda(t)dt} X \right],$$

i.e. survival contingent payoffs. The resulting measure can be termed a *survival contingent measure*.

The change of measure *after* the elimination of the default indicator function makes sure that the new measure remains equivalent to the original martingale measure. Simply taking a defaultable bond as numeraire would mean that this property is lost, and the change of measure would become much more complicated because the price path of a defaultable bond can be discontinuous.

Even when no closed form solutions are available the conditioning technique is still very useful as pre-processing procedure for a subsequent numerical solution of the pricing equation. Especially Monte Carlo methods are notoriously slow to converge for low-probability events (like defaults), the results can be speeded up significantly if the default events have been removed. P.d.e. methods (like finite differences solvers) are usually based on a Feynman-Kac representation of the price-expectation as an expectation of a stochastic integral and a final payoff in terms of *diffusion processes*. Therefore they cannot handle discrete default events directly[18], a pre-processing to remove the jumps is necessary here, too. The only case when this does not make sense is if one would like to recover the full distribution of the payoffs from the numerical scheme, including default events.

It was seen furthermore that the analogy to the default-free interest rate world does not carry through completely. There is the simple representation (4.10) of a defaultable bond price as 'bond price expectation' with defaultable interest rates given by $\bar{r} = r + q\lambda$ but this does *not* mean that options on defaultable bonds can also be treated as if there were no defaults but just a new interest rate $\bar{r}$.

Most of the results of this chapter are independent of the specification used for the interest-rate and credit spread processes. The transfer of the techniques demonstrated for the Gaussian and CIR specification to another specification should not be too hard.

Another point to mention is the amount of data and information that is needed to price credit risk derivatives. One of the advantages of the Duffie-Singelton setup appeared to be the fact, that default intensities λ and default loss rates q only appeared linked as a product in the bond pricing equation, and the product $q\lambda$ could therefore be estimated directly from the credit spreads — without any further need for separate default and recovery information. This breaks down when

[18] It is possible to incorporate jump processes in a finite-difference scheme, but at the cost of computing speed and having to solve matrix inversion problems for full matrices.

credit risk derivatives are considered, the default intensity λ *does* appear separated from the default losses q, both pieces of information are needed separately.

In a practical implementation much of this data will not be available and the financial engineer will have to resort to historical data, market data from other issuers of the same industry, region and rating class, or to fundamental analysis. On the other hand, every derivative needs an underlying, and this underlying should be fairly immune to market manipulation[19]. This means that products with option features are viable if and only if there is some degree of liquidity in the market for the underlying defaultable bond. Liquidity can usually only be expected for large issuers and this means better quality on historical price data and probably liquid markets for further defaultable bonds by the same issuer that can be used to build a credit spread curve. Therefore CRDs will only be viable in markets and for issuers where the data problems are less severe anyway. Nevertheless, data availability and quality is one of the most pressing problem in the field at the moment.

Apart from data problems, market incompleteness is a further large problem that has to be addressed in future research. With uncertainty about the recovery rates the markets for defaultable bonds are incomplete, and perfect hedges are possible only in exceptional circumstances. The pricing formulae that were derived here depend on the input of a pre-specified martingale measure that should be the output of a thorough analysis of the possibility of hedging and risk-management in credit markets. On the other hand, while being a problem for the pricing and hedging of credit derivatives, market incompleteness has also been one of the major reasons for the success of these instruments: credit derivatives are an important step towards the completion of the credit markets.

4.11 Credit Derivatives Literature

Despite a number of articles that have been written on the application and uses of credit derivatives, there is very little literature on the direct *pricing* of credit derivatives. Among the exceptions here are the articles of Duffie [1999] and Longstaff and Schwartz [1995]. Das [1998] gives a simple discretisation of the HJM- approach to credit spreads, and Pierides [1997] and Das [1995] use a firm's value approach to value credit derivatives.

Basket default swaps and default correlation are the topic of Duffie [1998], Duffie et.al. [1999], Li [1999] [1999], Zhou [1997], Lucas et.al. [1999] and Kijima [1999] [1999].

Good but not very rigorous introductions to the applications and uses of credit derivatives are the books by Mathieu and d'Herouville [1998], Tavakoli [1998], Das [1998], Francis et.al. [1999] and Nelken [1999].

[19]See Schönbucher and Wilmott [1993, 1999?] for a detailed discussion of option pricing and hedging in illiquid markets.

Chapter 5

Rating Transitions

5.1 Markov Chains

5.1.1 In this section you will learn ...

- ... the intuition behind models of credit rating migration,
- ... how this is modeled using Markov chains,
- ... how to handle Markov chains mathematically,
- ... how to derive a Markov model from empirical data.

5.1.2 An example

Firm ABC is currently rated A, but obviously this rating can change. We assume that there are only three possible ratings, A and B, and D, the rating for defaulted debt. For simplicity we set the recovery rate of defaulted debt to zero.

The rating agency publishes the following rating migration data:

	A	B	D
A	$p_{AA} = 0.80$	$p_{AB} = 0.15$	$p_{AD} = 0.05$
B	$p_{BA} = 0.10$	$p_{BB} = 0.80$	$p_{BD} = 0.10$
D	$p_{DA} = 0.00$	$p_{DB} = 0.00$	$p_{DD} = 1.00$

For example, of the companies rated A at the beginning of a year,

- $p_{AA} = 80$ percent were still rated A after one year,
- $p_{AB} = 15$ percent were rated B after that year, and

- $p_{AD} = 5$ percent were rated D, they had defaulted within the year.

We assumed that no firm can recover from default, the state D is called *absorbing*.

What is the probability of a default of the A rated ABC-debt within the next 2 years? There is an obvious, but wrong answer:

One could say, we have two events of default each with a probability of $p_{AD} = 0.05$ or a survival probability of $(1 - p_{AD} = 0.95$, giving a total survival probability of $(1 - p_{AD})^2 = 0.9025$ and a default probability of 0.0975.

This answer does not take *rating migration* into account: In the next two years a default can also occur via a transition to the B rating. Default can be reached via the following transitions:

$$\begin{array}{lll} A \to A \to D & \text{with a probability of} & p_{AA}p_{AD} = 0.85 \cdot 0.05 = 0.0425 \\ A \to B \to D & \text{with a probability of} & p_{AB}p_{BD} = 0.15 \cdot 0.10 = 0.015 \\ A \to D(\to D) & \text{with a probability of} & p_{AD}p_{DD} = 0.05 \cdot 1 = 0.05, \end{array}$$

which gives a total default probability of 0.1075 which is a full percentage point larger than before. This effect is even stronger with real-world ratings. Here the credit risk for investment-grade bonds lies mainly in the risk of downgrading (with subsequently very much higher risk of default), and not in the risk of direct default.

To sum up, the two-period probability of default given initial rating A, i.e. the *two-period transition probability from A to D* is

$$p_{AD}^{(2)} = p_{AA}p_{AD} + p_{AB}p_{BD} + p_{AD}p_{DD} = \begin{pmatrix} p_{AA} & p_{AB} & p_{AD} \end{pmatrix} \begin{pmatrix} p_{AD} \\ p_{BD} \\ p_{DD} \end{pmatrix}.$$

If one takes the transition matrix

$$A = \begin{pmatrix} p_{AA} & p_{AB} & p_{AD} \\ p_{BA} & p_{BB} & p_{BD} \\ p_{DA} & p_{DB} & p_{DD} \end{pmatrix}$$

then it is easily seen that the 2-period transition probability $p_{AD}^{(2)}$ is exactly the (A, D)-component of the square A^2 of A. This also holds for the other two period transition probabilities, we reach the two-period transition probability matrix as

$$A^{(2)} = A \cdot A = A^2.$$

5.1.3 Markov chains

The mathematical instrument to correctly model rating transitions is a continuous-time *Markov chain*. Let us assume we have K different states (=ratings) from $1 = AAA$ to $K = D$. Similarly to the construction of the Poisson process we assume that the transition probability from state k to state l in the small time interval $\Delta t$is proportional to Δt:

$$P[R(t+\Delta t) = l | R(t) = k] = a_{kl}\Delta t.$$

The probability of a transition of the rating $R(t)$ of the firm to $R(t+\Delta t) = l$ at time $t+\Delta t$, given that it was k at time t, is $a_{kl}\Delta t$. This means that the diagonal elements of A are given by

$$P[R(t+\Delta t) = k | R(t) = k] = 1 - \sum_{l \neq k} a_{kl}\Delta t =: 1 + a_{kk}\Delta t,$$

$a_{kk} = -\sum_{l \neq k} a_{kl}$ (This is a slight change from the definition of A in the previous section.)

Then the matrix of transition probabilities for the time interval $[t, t+\Delta t]$ is

$$P(t, t+\Delta t) = I + \Delta t A,$$

where I is the unit matrix. Now we can go and derive the matrix of transition probabilities for a larger interval $[t, s]$. Exactly as in the Poisson case we subdivide $[t, s]$ into i subintervals of length Δt. From the previous section we know that the two-period transition matrix is reached by multiplying the one-period transition matrix with itself:

$$\begin{aligned} P(t, t+2\Delta t) &= P(t, t+\Delta t)P(t+\Delta t, t+2\Delta t) \\ &= (I + \Delta t A)(I + \Delta t A) = (I + \Delta t A)^2. \end{aligned}$$

Similarly we have for the i-period transition matrix that

$$P(t, t+i\Delta t) = P(t, s) = (I + \Delta t A)^i = (I + \frac{(s-t)}{i} A)^i, \tag{5.1}$$

and again we have in the limit

$$P(t, s) = \exp\{(s-t)A\}, \tag{5.2}$$

where the exponential function for matrices is defined by the series expansion of the exponential function:

$$\exp\{(s-t)A\} = \sum_{n=0}^{\infty} \frac{((s-t)A)^n}{n!}.$$

Besides this representation of the transition probabilities the limiting process has another consequence: If the transition matrix from t to s is given by $P(t, s)$ then the transition matrix from t to $s + \Delta t$ is

$$P(t, s+\Delta t) = P(t, s)(I + \Delta t A) = P(t, s) + \Delta t P(t, s) A,$$

or, rearranging

$$\frac{1}{\Delta t}(P(t, s+\Delta t) - P(t, s)) = \frac{\partial}{\partial s} P(t, s) = P(t, s) A. \tag{5.3}$$

The differential equations (5.3) are known as the *Kolmogorov forward differential equations*.

A Markov chain does not behave any different from a compound Poisson process, it can in fact be viewed as a collection of compound Poisson processes, one for each rating class. The process for class k has intensity $\lambda_k = -a_{kk} = \sum_{l\neq k} a_{kl}$, and whenever there is a jump in this process a marker m is drawn which tells us, to which rating class m the firm in class k migrates. The relative probabilities of the different rating classes are given by the relative magnitudes of the a_{kl}: The probability of $m = l$ is given by $K(l+) - K(l) = a_{kl} \sum_{n\neq k} a_{kn}$.

Alternatively the Markov chain can be considered as a collection of K Poisson processes: The k-th Poisson process then drives the transition to rating class k. The intensity of this process would then be given by a_{nk} if we are now in rating class n. (Obviously we have to exclude the process driving the 'transition to the current class.)

We can now define for each firm a rating process $R(t)$ that gives the rating of the firm at time t. Infinitesimally the credit rating process $R(t)$ has the following properties

$$\mathrm{E}[\,dR\,] = (\sum_{k\neq R}(k - R)a_{Rk})dt = \sum_{k=1}^{K} k a_{Rk} dt \tag{5.4}$$

$$P(dR = k - R) = a_{Rk}dt. \tag{5.5}$$

A function $f(R)$ of the credit rating R of a firm has the increments:

$$\mathrm{E}[\,df(R)\,] = (\sum_{l\neq R}(f(l) - f(R))a_{Rl})dt = \sum_{l=1}^{K} f(l)a_{Rl}dt, \tag{5.6}$$

$$P(df(R) = f(k) - f(R)) = a_{Rk}dt. \tag{5.7}$$

We see that the treatment of f is exactly as in the case of a compound Poisson process.

5.1.4 Deriving the Generating Matrix

Usually rating agencies do not publish the generating matrix A of the rating migration process, but only the transition probabilities P for a given time period, usually one year. From equation (5.2) we know that[1]

$$P(t) = e^{At}.$$

If A were a scalar and not a matrix we would simply solve this to

$$A = \frac{1}{t}\ln P.$$

This is in fact the solution of (5.2) for A, when we define the logarithm of a matrix by its power series:

$$\ln(I + X) = X - X^2 + X^3 \dots$$

[1] If the transition probabilities do not depend on calendar time the Markov process is called *time homogeneous*. In this case we write the transition probabilities $P(t)$ only in dependence of the length t of the time interval considered.

In praxis we do not have to calculate this infinite series, though. We assume that we can decompose the matrix P in the following form: [2]

$$P = MDM^{-1}, \tag{5.8}$$

where D is a diagonal matrix and M a square matrix. (This decomposition can be done in a standard spreadsheet program.) Then we have

$$\begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1K} \\ p_{21} & p_{22} & \cdots & p_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ p_{K1} & p_{K2} & \cdots & p_{KK} \end{pmatrix} =$$

$$\begin{pmatrix} m_{11} & m_{12} & \cdots & m_{1K} \\ m_{21} & m_{22} & \cdots & m_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ m_{K1} & m_{K2} & \cdots & m_{KK} \end{pmatrix} \begin{pmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_K \end{pmatrix} \begin{pmatrix} m_{11} & m_{12} & \cdots & m_{1K} \\ m_{21} & m_{22} & \cdots & m_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ m_{K1} & m_{K2} & \cdots & m_{KK} \end{pmatrix}^{-1}.$$

The columns of M are called the *eigenvectors* of P, and the diagonal elements of D are called the *eigenvalues*.

Having achieved this decomposition of P we can take exponentials or logarithms of P by taking these operations of the diagonal elements[3] in D:

$$\ln \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1K} \\ p_{21} & p_{22} & \cdots & p_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ p_{K1} & p_{K2} & \cdots & p_{KK} \end{pmatrix} =$$

$$\begin{pmatrix} m_{11} & m_{12} & \cdots & m_{1K} \\ m_{21} & m_{22} & \cdots & m_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ m_{K1} & m_{K2} & \cdots & m_{KK} \end{pmatrix} \begin{pmatrix} \ln d_1 & 0 & \cdots & 0 \\ 0 & \ln d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \ln d_K \end{pmatrix} \begin{pmatrix} m_{11} & m_{12} & \cdots & m_{1K} \\ m_{21} & m_{22} & \cdots & m_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ m_{K1} & m_{K2} & \cdots & m_{KK} \end{pmatrix}^{-1}$$

$$= \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1K} \\ a_{21} & a_{22} & \cdots & a_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ a_{K1} & a_{K2} & \cdots & a_{KK}. \end{pmatrix}$$

The advantage of this method is, that we now directly have the decomposition of A:

$$\ln P = A = M(\ln D)M^{-1} = MD_AM^{-1},$$

[2] Not every matrix can be decomposed like (5.8), but the decomposition is usually possible for the empirical rating transition matrices we are interested in.

[3] We obviously need that the eigenvalues of P are positive.

where $D_A = \ln D$, without any additional calculations. This makes the calculation of the transition probabilities for other times very easy:

$$P(s) = \exp\{sA\} = M \exp\{sD_A\} M^{-1} = M \exp\{s \ln D\} M^{-1} = M D^s M^{-1},$$

where D^s is calculated by taking the diagonal elements of D to the s-th power.

One final warning: Exponentials, logarithms, square roots etc... can all be taken of (suitable) matrices, but they have to be decomposed in the form (5.8) and then the operations can be applied to the diagonal matrix. If the operations are applied to the elements of the matrix directly the results are wrong[4].

[4]Note that even for the diagonal matrices D and D_A we only applied the operations to the *diagonal* elements, not to the off-diagonal zeroes.

5.2 Pricing rating transitions

5.2.1 In this section you will learn ...

- ... how to price zero coupon bonds in a rating-transition model,
- ... the spreads implied by rating-transition models,
- ... how to derive the pricing equations for derivatives conditioning on rating transitions,
- ... how to derive hitting probabilities and how to price securities with downgrade or upgrade triggers,
- ... about hedging in rating transition models,
- ... how to reach realistic dynamics for the credit spreads in these models.

5.2.2 Pricing Zero Coupon Bonds

The price of a defaultable zero coupon bond will now be a function of the issuer's credit rating $R(t)$ at time t:

$$\overline{B} = \overline{B}(t, r, R),$$

and it will also depend on time and the risk-free interest rate r. By construction of the model it is impossible to price a bond for a single rating class without simultaneously pricing the bonds for the other rating classes, too: As transitions are possible at any time we always need to know the price the bond will have after a transition, i.e. we need to know the price of the bond for all other ratings.

Because R can only take K different values it is convenient to write the defaultable bond price as a vector

$$\overline{B}(t, r) = \begin{pmatrix} \overline{B}(t, r, R = 1) \\ \overline{B}(t, r, R = 2) \\ \vdots \\ \overline{B}(t, r, R = K) \end{pmatrix}$$

with the k-th component $\overline{B}_k$ denoting the price of the defaultable bond if the rating is $R = k$. The price for a given rating can then be read directly from $\overline{B}$.

If we assume zero recovery, the payoffs at maturity are

$$\overline{B}(T) = (1, 1, \ldots, 1, 0)'.$$

The payoff in maturity is one for all rating classes, except in the last class (the default) where it is zero. The price of the defaultable bond is the risk-neutral expectation of its discounted payoff, given an initial rating of $R(0)$:

$$\overline{B}(0) = \mathrm{E}\left[\exp\{-\int_0^T r(s)ds\}\overline{B}_{R(T)}(T)\right]$$

$$
\begin{aligned}
&= \mathrm{E}\left[\exp\{-\int_0^T r(s)ds\}\right]\mathrm{E}\left[\,\overline{B}_{R(T)}(T)\,\right] \\
&= B(0,r)\mathrm{E}\left[\,\overline{B}_{R(T)}(T)\,\right] \\
&= B(0,r)(P(T)\overline{B}(T))_{R(0)} \\
&= B(0,r)\sum_{k=1}^{K-1} P_{R(0)k}(T) \\
&= B(0,r)(1-P_{R(0)K}(T)),
\end{aligned}
$$

where $P(T)$ is the transition probability matrix until T. In this simple case the defaultable bond prices are given by the default-free bond prices times the survival probability (i.e. one minus the transition probability to default).

There are several observations to be made here: First, for a given credit class and maturity the transition probability to default is $P(T)$, where T is the time to maturity. This remains constant over time, and therefore the credit spreads remain constant within every credit class, too.
Second, we assumed zero recovery at default. This can be relaxed to: At default there is a cash payoff of $(1-c)$ default-free bonds, i.e. $(1-c)B(t)$. Then the pricing formula becomes

$$\overline{B}(0) = B(0,r)(1-cP_{\cdot K}(T)),$$

it is just modified by the factor c in front of the transition probability to default. This is still not satisfying as we still have static credit spreads.
Third, we assumed that the generator matrix that was used already reflected the risk-neutral probabilities of rating transitions and defaults. This is usually not the case either. The empirical transition matrices published by the rating agencies only justify credit spreads that are significantly lower than the ones observed in the market, even if one assumes zero recovery. There is a risk-premium on credit risk in the market that makes the pricing probabilities of default larger than the corresponding empirical probabilities. Here we assume that A and P already reflect the risk-neutral probabilities.

5.2.3 Pricing Derivatives on the Credit Rating

Again we use vector notation for the price of a derivative security whose payoff depends on the credit rating of an underlying bond. Let

$$F'(t,r) = (F_1'(t,r), F_2'(t,r), \ldots, F_K'(t,r))'$$

be the price of this derivative.

5.2.3.1 European-style payoffs

If F has an European payoff at T, which is given by

$$F'(T,r) = (F_1^*(r), \ldots, F_K^*(r))'$$

we can apply the methods we used for the defaultable zero-coupon bonds to this situation. Again we have that $F'(t)$ is the risk-neutral expectation of its discounted payoff. Assuming that $F'(T)$ does not depend on r we have for the expectation of the final payoff (without discounting and given an initial rating of R)

$$\mathrm{E}[\, F'_R(T)\,] = F'_1(T)P_{R1} + F'_2(T)P_{R2} + \ldots + F'_K(T)P_{RK}$$

the sum of the payoffs weighted with the transition probabilities. In vector notation this is

$$\mathrm{E}[\, F'(T)\,] = P(T)F^*.$$

With interest rate dependence and discounting this extends to (remember the definition of the discount factor $\beta_{tT} = \exp\{-\int_0^T r(s)ds\}$)

$$\mathrm{E}[\, F'(r,T)\,] = P(T)\mathrm{E}[\, \beta_{tT}F^*(T)\,] = P(T)F(t),$$

where $F_k(t)$ is the (default-free) price of a security paying off $F^*_k(r)$ at time T for sure.

This can be seen if the equation is written for the R-th rating class and conditioning on the final class is used:

$$\begin{aligned}
F'_R(t) &= \mathrm{E}\left[\, \beta_{t,T}F'_{R(T)}(T,r)\,\right] \\
&= \mathrm{E}\left[\, \beta_{t,T}F'_{R(T)}(T,r) \mid R(T) = 1\,\right] P_{R1}(T) \\
&\quad + \mathrm{E}\left[\, \beta_{t,T}F'_{R(T)}(T,r) \mid R(T) = 2\,\right] P_{R2}(T) \\
&\quad + \ldots + \mathrm{E}\left[\, \beta_{t,T}F'_{R(T)}(T,r) \mid R(T) = K\,\right] P_{RK}(T) \\
&= \mathrm{E}[\, \beta_{t,T}F^*_1(r)\,] P_{R1} + \mathrm{E}[\, \beta_{t,T}F^*_2(r)\,] P_{R2} + \ldots + \mathrm{E}[\, \beta_{t,T}F^*_K(r)\,] P_{RK} \\
&= F_1(t,r)P_{R1} + F_2(t,r)P_{R2} + \ldots + F_K(t,r)P_{RK} \\
&= \sum_{k=1}^{K} F_K(t,r)P_{Rk}(T).
\end{aligned}$$

The value of the European-style payoff is just the value of the individual risk-free payoffs times the transition probabilities of actually recieving this payoff.

5.2.4 General Payoffs

To price more general payoffs we again need an Itô-Lemma representation of the price process of the derivative:

$$\begin{aligned}
dF'_R(t,r) &= \frac{\partial}{\partial t}F'_R dt + \frac{1}{2}\sigma_r^2 \frac{\partial^2}{\partial r^2}F'_R dt \\
&\quad + \frac{\partial}{\partial r}F'_R dr + (F'_{R+\Delta R}(t,r) - F'_R(t,r)).
\end{aligned}$$

Here the last term represents the changes in the value of the derivative due to the change ΔR in the credit rating of the underlying instrument.[5] The expectation of dF' is given by

$$\begin{aligned} \mathrm{E}\,[\,dF'_R(t,r)\,] = {} & \frac{\partial}{\partial t}F'_R dt + \frac{1}{2}\sigma_r^2\frac{\partial^2}{\partial r^2}F'_R dt \\ & + \mu_r\frac{\partial}{\partial r}F'_R dt + \sum_{k=1}^{K} A_{Rk}F'_k dt. \end{aligned}$$

As usual, this must equal $rF'dt$ which yields the pricing equation

$$\begin{aligned} 0 = {} & \frac{\partial}{\partial t}F'_R + \frac{1}{2}\sigma_r^2\frac{\partial^2}{\partial r^2}F'_R \\ & + \mu_r\frac{\partial}{\partial r}F'_R + \sum_{k=1}^{K} A_{Rk}F'_k. \end{aligned}$$

or in compact matrix notation

$$\begin{aligned} 0 = {} & \frac{\partial}{\partial t}F' + \frac{1}{2}\sigma_r^2\frac{\partial^2}{\partial r^2}F' \\ & + \mu_r\frac{\partial}{\partial r}F' + AF' - rF'. \end{aligned} \tag{5.9}$$

Without any further knowledge about the structure of the problem we have to stop here and solve the pricing equation numerically subject to the appropriate boundary and final conditions.

The solution is complicated by the fact that we are dealing with K coupled partial differential equations, one for each rating class. Because the prices in the different classes will interact in general, we have to solve these equations simultaneously.[6]

5.2.5 Downgrade triggers

A common specification of credit derivatives triggers some payment or a knockout at a specific change of rating. As an example we consider here a downgrade-insurance, which would have a payoff of

$$F' = \overline{B}_A - \overline{B}$$

if the rating of the bond $\overline{B}$ with maturity T_2 drops to less than A. Here $\overline{B}_A$ is the price of an equivalent A-rated defaultable bond, so as soon as the bond $\overline{B}$ is downgraded the buyer of the insurance can exchange it for an equivalent A-rated bond[7]. We assume that the insurance is valid until time $T_1 \leq T_2$.

For the pricing of F' this means that we already know F'_k for all $k > 3$ (where $k = 3$ represents the rating A and $k > 3$ represents ratings worse than A):

$$F'_k(r,t) = \overline{B}_3(r,t) - \overline{B}_k(r,t)$$

[5] All partial derivatives apply to every component if applied to a vector function.

[6] This situation is not worse than having an additional dimension in the partial differential equation.

[7] This payoff is obviously always positive.

$$= B(t, T_2, r)[1 - cP_{3K}(t, T_2) - (1 - cP_{kK}(t, T_2))]$$
$$= cB(t, T_2, r)[-P_{3K}(t, T_2) + P_{kK}(t, T_2)]$$

where we used the bond-pricing formula from the previous section. Note that the default transition probabilities until T_2 in the bond prices. For clarity we included T_2 in the arguments where appropriate.

The final payoff for F' at T_1 is zero. (If the insurance has a payoff, it has it earlier.) This is sufficient information to go and solve equation (5.9) numerically for the three remaining components F_1', F_2' and F_3'.

Often it is interesting to calculate *hitting probabilities* for the rating transition process: What is the probability of a transition to a rating of k in the period $[0, T]$? This is a different problem from asking for the probability of being at k at T, if we only condition on the final moment we ignore the possibility that the rating may have changed to k and then away from it. We would underestimate the hitting probability.

To frame this problem as a final state problem we only have to make the state k an *absorbing* state. The k-th row of the generating matrix A contained the intensities of the Poisson processes triggering the transitions to other classes. a_{kl} for instance is the intensity of a transition from k to l. If we thus set the k-th row of A to zero we make transitions away from k impossible: If the process hits k once, it will remain there. The matrix A is thus changed from

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & \cdots & \cdots & a_{1K} \\ a_{21} & a_{22} & \cdots & \cdots & \cdots & a_{2K} \\ \vdots & \vdots & \ddots & & & \vdots \\ a_{k1} & a_{k2} & \cdots & \cdots & \cdots & a_{kK} \\ \vdots & \vdots & & & \ddots & \vdots \\ a_{K1} & a_{K2} & \cdots & \cdots & \cdots & a_{KK} \end{pmatrix}$$

to

$$A_k = \begin{pmatrix} a_{11} & a_{12} & \cdots & \cdots & \cdots & a_{1K} \\ a_{21} & a_{22} & \cdots & \cdots & \cdots & a_{2K} \\ \vdots & \vdots & \ddots & & & \vdots \\ 0 & 0 & \cdots & \cdots & \cdots & 0 \\ \vdots & \vdots & & & \ddots & \vdots \\ a_{K1} & a_{K2} & \cdots & \cdots & \cdots & a_{KK} \end{pmatrix}.$$

With this change of the generating matrix we can the probability of hitting k until T is equal to the probability of being at k at T, the transition probability to k. We already have derived the expression for the transition probabilities, it is

$$P_k(t) = e^{A_k t}. \tag{5.10}$$

P_k will have zeros in the k-th row except on the diagonal where $p_{kk} = 1$. There are two types of information we can gain from P_k:

For $j \neq k$ the (ij)-th element p_{ij} of P_k denotes the *probability of a transition from class i to class j without going through k in the meantime.*
And in the k-th column we have the hitting probabilities of class k: The element p_{ik} is the probability of hitting k if the rating process started at i.

Obviously we can also modify several rows of the matrix A, for instance if we wanted the probability of a downgrade to k or worse we would set the last rows of A, starting from row k, to zero. The rating transition process is then stopped as soon as one of the classes that have been set to zero, is reached.

In all these cases we have to re-calculate the diagonal decomposition of the modified matrix A_k.

5.2.6 Hedging rating transitions

Assuming we have K rating classes we need K different hedge instruments F_k to hedge the rating transitions. The hedge weights α_k to hedge a security V have to satisfy

$$V(r, R) - V(r, R_0) = \sum_{k=1}^{K} \alpha_k (F_k(r, R) - F_k(r, R_0)) \qquad \forall 1 \leq R \leq K,$$

where R_0 is the rating class we are starting from. This equation means, that for every possible rating class R we need that the change of the value of the hedge portfolio (on the right-hand side) has to equal the change of the value of the security to hedge (on the left hand side), if there is a transition to this class R. This yields $(K - 1)$ equations[8] that are to be satisfied by the K hedge instruments. The K-th equation to make the solution unique comes from the hedging of the continuous interest rate risk: We need that

$$\frac{\partial}{\partial r} V(r, R_0) = \sum_{k=1}^{K} \alpha_k \frac{\partial}{\partial r} F_k(r, R_0).$$

With these equations the hedging problem has a unique solution, provided that the hedge instruments F_k are chosen suitably.

5.2.7 Stochastic Spreads

So far the credit spreads in this model only changed with a rating transition but remained constant within a given credit class. This is obviously an unrealistic feature of this model. Lando [1993, 1994, 1994] proposes to multiply the generating matrix A with a scalar factor h,

$$A_h = hA$$

where h follows some stochastic process

$$dh = \mu_h dt + \sigma_h dW.$$

[8] Note that the equation for $R = R_0$ drops out trivially.

If h is a deterministic function it can be used to fit an initial term structure for one class of bonds. Then the decomposition of A becomes

$$A_h = MD_hM^{-1} = MhDM^{-1},$$

the diagonal matrix with the eigenvalues of A is just multiplied with h. This means for the transition probabilities, that for a given realisation $h(s)$ of h we have

$$P(t) = \exp\{\int_0^t A_{h(s)}ds\} = \exp\{A\int_0^t h(s)ds\}.$$

For stochastic h we just have to take expectations over the realisations of h to get:

$$P_h(t) = \mathrm{E}\left[\exp\left\{A\int_0^t h(s)ds\right\}\right] = M\mathrm{E}\left[\exp\left\{D\int_0^t h(s)ds\right\}\right]M^{-1}. \tag{5.11}$$

This specification will indeed generate stochastic credit spreads, but unfortunately these spreads will still be perfectly correlated over different classes. The multiplication with h changed the *speed* at which the Markov chain changes, a large h increases the likelihood of all jumps, while a small h decreases it.

Das and Tufano [1994] take a slightly different approach: they keep the generator matrix A constant, but make the *bankruptcy costs* c stochastic. This enables them to individually assign a specific cost c_k to the rating class k, and thus to individually model the spreads for each class. Care has to be taken though, to keep c in the range $c \in [0, 1]$.

Both these specifications allow for dynamics in the implied credit spreads and have the additional advantage that the new parameters can be used to fit the model to observed market data.

5.3 Markov Models: Guide to the Literature

The first continuous-time model of credit-risk pricing in a rating-transition framework is due to Lando [1994] and Jarrow, Lando and Turnbull [1993]. This model only incorporates the Markov-chain dynamics of the ratings without allowing for stochastic spread dynamics.

Das and Tufano [1994] extend the Jarrow- Lando- Turnbull model to incorporate stochastic recovery rates and thus to have stochastic dynamics of the credit spreads within the individual classes. Their model is set up as a discrete-time approximation to a continuous-time model.

Finally, Lando [1994] extends his model to have stochastic credit spreads by incorporating a stochastic multiplier in front of the transition generator matrix.

Appendix

Appendix A

Calculations to the Gaussian Model

A.1 Proof of Lemma 20

Proof: (of Lemma 20)
The survival probability is defined as

$$P(0,T) = \mathrm{E}\left[\, 1_{\{\tau > T\}} \,\right]. \tag{A.1}$$

Using iterated expectations and the Cox process properties this can be expanded to

$$\begin{aligned} P(0,T) &= \mathrm{E}\left[\, \mathrm{E}\left[\, 1_{\{\tau>T\}} \,\middle|\, \lambda(t),\ t \geq 0 \right] \right] \\ &= \mathrm{E}\left[\exp\{-\int_0^T \lambda(t)dt\} \right]. \end{aligned} \tag{A.2}$$

From assumption 12 we know the dynamics of the defaultable and default-free bond prices:

$$\begin{aligned} \frac{dB(t,T)}{B(t,T)} &= r(t)dt + a(t,T)dW \\ \frac{d\overline{B}(t,T)}{\overline{B}(t,T)} &= (r(t) + q\lambda(t))dt + \overline{a}(t,T)dW. \end{aligned}$$

Thus the bond prices satisfy for all $t \leq T_1 \leq T$ (conditional on survival):

$$B(T_1,T) = B(t,T)\exp\left\{\int_t^{T_1} r(s)ds - \frac{1}{2}\int_t^{T_1} a^2(s,T)ds + \int_t^{T_1} a(s,T)dW_s\right\} \tag{A.3}$$

$$\overline{B}(T_1,T) = \overline{B}(t,T)\exp\left\{\int_t^{T_1} r(s) + q\lambda(s)ds - \frac{1}{2}\int_t^{T_1} \overline{a}(s,T)^2 ds + \int_t^{T_1} \overline{a}(s,T)dW_s\right\}. \tag{A.4}$$

Using the previous equations with $T_1 = T$, $t = 0$ and $B(T,T) = \overline{B}(T,T) = 1$ we reach

$$1 = B(0,T)\exp\left\{\int_0^T r(s)ds - \frac{1}{2}\int_0^T a^2(s,T)ds + \int_0^T a(s,T)dW_s\right\}$$

$$1 = \overline{B}(0,T)\exp\left\{\int_0^T r(s) + q\lambda(s)ds - \frac{1}{2}\int_0^T \overline{a}(s,T)^2 ds + \int_0^T \overline{a}(s,T)dW_s\right\},$$

which can be solved for the integral of λ to yield

$$\exp\{-\int_0^T \lambda(s)ds\} = \left(\frac{\overline{B}(0,T)}{B(0,T)}\right)^{\frac{1}{q}} \mathcal{E}\left(\int_0^T \frac{1}{q}(\overline{a}(s,T) - a(s,T))dW_s\right)$$
$$\exp\left\{-\frac{1}{2q^2}\int_0^T (\overline{a}(s,T) - a(s,T))[(1+q)a(s,T) - (1-q)\overline{a}(s,T))]ds\right\}. \quad \text{(A.5)}$$

Taking expectations of (A.5) yields equation (4.48).

It remains to show the dynamics (4.46) and (4.47) of the survival probability. From representation (A.2) follows, that

$$M_t := P(t,T)e^{-\int_0^t \lambda(s)ds}$$

is a martingale. Thus $\mathbf{E}[\,dM\,] = 0$ and combined with Itô's lemma this means that

$$\mathbf{E}[\,dP(t,T)\,] = \lambda(t)P(t,T)dt. \quad \text{(A.6)}$$

The volatility of $P(t,T)$ can be derived from (4.48) by substituting for the bond prices $B(t,T)$ and $\overline{B}(t,T)$ from equations (A.3) and (A.4) respectively, and using Itô's lemma again. Finally, similar to equations (A.3) and (A.4), the survival probability can be represented as follows for all $t \leq T_1 \leq T$:

$$P(T_1,T) = P(t,T)\exp\left\{\int_t^{T_1} \lambda(s)ds - \frac{1}{2}\int_t^{T_1} a^P(s,T)^2 ds + \int_t^{T_1} a^P(s,T)dW_s\right\}. \quad \text{(A.7)}$$

□

A.2 Proof of Lemma 21

The proof of this lemma is an application of Girsanov's theorem.

Proof: (of Lemma 21)
First, the expectation has to be converted using iterated expectations and the Cox process properties of the default process:

$$\mathbf{E}^Q\left[\,\beta_{t,T}\mathbf{1}_{\{\tau>T\}}X\,\right] = \mathbf{E}^Q\left[\mathbf{E}^Q\left[e^{-\int_t^T r(s)ds}\mathbf{1}_{\{\tau>T\}}X \,\middle|\, \lambda(s)\ s \geq 0\right]\right]$$
$$= \mathbf{E}^Q\left[e^{-\int_t^T r(s)+\lambda(s)ds}X\right].$$

Now we use the representation of the default-free bond price (A.3) and the survival probability in (A.7) to substitute for $e^{\int r(s)ds}$ and $e^{\int \lambda(s)ds}$:

$$\mathbf{E}^Q\left[\,\beta_{t,T}\mathbf{1}_{\{\tau>T\}}X\,\right] = B(t,T)P(t,T)\exp\{-\frac{1}{2}\int_t^T a(s,T)^2 + a^P(s,T)^2 ds\}$$

$$
\begin{aligned}
&\mathrm{E}^Q\left[\exp\{\int_t^T (a(s,T)+a^p(s,T))dW_s\} X\right] \\
=&B(t,T)P(t,T)\exp\{\int_t^T a(s,T)a^p(s,T)ds\} \\
&\mathrm{E}^Q\left[\mathcal{E}\left(\int_t^T (a(s,T)+a^p(s,T))dW_s\right) X\right]
\end{aligned}
$$

The rest of the lemma follows directly from Girsanov's theorem. The representation of the zero-recovery defaultable bond price follows by setting $X = 1$.

□

A.3 Proof of Proposition 22

Proof: (of Proposition 22)
We have to calculate

$$x(t) := \mathrm{E}\left[\lambda(t)e^{-\int_0^t \lambda(s)ds}e^{-\int_0^t r(s)ds}\right].$$

The change of measure to the survival contingent measure P_S (see lemma 21, the measure is contingent on survival until t) reduces the problem to finding

$$x(t) = B(0,t)P(0,t)\exp\{\int_0^t a(s,t)a^p(s,t)ds\}\,\mathrm{E}^{P_S}[\lambda(t)] \tag{A.8}$$

where under P_S

$$dW_s^{P_S} = dW_s - (a(s,t)+a^p(s,t))ds$$

is a P_S-Brownian motion.

To evaluate equation (A.8) we need to find the dynamics of λ. Define the forward default intensity

$$\lambda(t,T) := -\frac{\partial}{\partial T}\ln P(t,T).$$

and denote its dynamics with

$$d\lambda(t,T) = \alpha^p(t,T)dt + \sigma^p(t,T)dW_t.$$

The spot default intensity is $\lambda(t) = \lambda(t,t)$. The volatility $\sigma^p(t,T)$ and the drift $\alpha^p(t,T)$ of $\lambda(t,T)$ can be derived directly from equation (A.7):

$$
\begin{aligned}
\sigma^p(t,T) &= -\frac{\partial}{\partial T}a^p(t,T) \\
\alpha^p(t,T) &= -\sigma^p(t,T)a^p(t,T).
\end{aligned}
$$

This yields

$$\lambda(t) = \lambda(t,t) = \lambda(0,t) + \int_0^t \alpha^p(s,t)ds + \int_0^t \sigma^p(s,t)dW_s$$

$$
\begin{aligned}
&= \lambda(0,t) + \int_0^t \alpha^P(s,t)ds + \int_0^t \sigma^P(s,t)dW_s^{Ps} \\
&+ \int_0^t (a(s,t) + a^P(s,t))\sigma^P(s,t)ds \\
&= \lambda(0,t) + \int_0^t a(s,t)\sigma^P(s,t)ds + \int_0^t \sigma^P(s,t)dW_s^{Ps}
\end{aligned}
$$

and therefore

$$
\mathbf{E}^{Ps}[\,\lambda(t)\,] = \lambda(0,t) + \int_0^t a(s,t)\sigma^P(s,t)ds.
$$

□

A.4 Proof of Proposition 23

Proof: (of Proposition 23)
The proof takes place in three steps: First, we analyse the 'in-the-money' event, then the two parts of the payoff are valued using standard option pricing methods.

Let A be the event that the option is in the money, i.e. $\bar{S}B(T_1,T_2) > \overline{B}(T_1,T_2)$. Using the bond price representations (A.3), (A.4) and (A.7) this event is equivalent to

$$
\begin{aligned}
&\ln\frac{\bar{S}B(t,T_2)}{\overline{B}(t,T_2)} + \frac{1}{2}\int_t^{T_1} \bar{a}(s,T_2)^2 - a(s,T_2)^2 ds \\
&> q\int_t^{T_1} \lambda(s)ds + \int_t^{T_1} \bar{a}(s,T_2) - a(s,T_2)dW_s
\end{aligned}
$$

or

$$
\begin{aligned}
&\ln\frac{\bar{S}B(t,T_2)}{\overline{B}(t,T_2)} + q\ln P(t,T_1) + \frac{1}{2}\int_t^{T_1} \bar{a}(s,T_2)^2 - a(s,T_2)^2 - qa^P(s,T_1)^2 ds \\
&> q\int_t^{T_1} a^P(s,T_2) - a^P(s,T_1)dW_s,
\end{aligned}
$$

which is equivalent to

$$
\begin{aligned}
&\ln\bar{S} + \ln\frac{\overline{B}(t,T_1)}{B(t,T_1)} - \ln\frac{\overline{B}(t,T_2)}{B(t,T_2)} \\
&- \frac{1}{2}\int_t^{T_1} \bar{a}(s,T_1)^2 - a(s,T_1)^2\,ds + \frac{1}{2}\int_t^{T_1} \bar{a}(s,T_2)^2 - a(s,T_2)^2\,ds \\
&> q\int_t^{T_1} (a^P(s,T_2) - a^P(s,T_1))\,dW_s.
\end{aligned}
$$

Using the 'In the Money' event A the option's payoff can be decomposed:

$$\begin{aligned}&\mathbf{E}\left[e^{-\int_t^{T_1} r(s)+\lambda(s)ds}(\bar{S}B(T_1,T_2)-\overline{B}(T_1,T_2))^+ \,\middle|\, \mathcal{F}_t\right]\\ =&\bar{S}\mathbf{E}\left[e^{-\int_t^{T_1} r(s)+\lambda(s)ds}B(T_1,T_2)\mathbf{1}_{\{A\}} \,\middle|\, \mathcal{F}_t\right]-\mathbf{E}\left[e^{-\int_t^{T_1} r(s)+\lambda(s)ds}\overline{B}(T_1,T_2)\mathbf{1}_{\{A\}} \,\middle|\, \mathcal{F}_t\right]\end{aligned}$$

We now calculate the first term in this expression.

$$\begin{aligned}I_1 :=& e^{-\int_t^{T_1} r(s)+\lambda(s)ds}B(T_1,T_2)\mathbf{1}_{\{A\}}\\ =&B(t,T_2)P(t,T_1)\times\exp\{\int_t^{T_1} a(s,T_2)a^p(s,T_1)ds\}\\ &\times\exp\{-\frac{1}{2}\int_t^{T_1}(a(s,T_2)+a^p(s,T_1))^2ds+\int_t^{T_1} a(s,T_2)+a^p(s,T_1)dW_s\}\mathbf{1}_{\{A\}}\\ =&B(t,T_2)P(t,T_1)\times\exp\{\int_t^{T_1} a(s,T_2)a^p(s,T_1)ds\}\\ &\times\mathcal{E}(\int_t^{T_1} a(s,T_2)+a^p(s,T_1)dW_s)\mathbf{1}_{\{A\}}\end{aligned}$$

As the volatility functions are deterministic in the Gaussian setup all we need to evaluate now is $\mathbf{E}\left[\mathcal{E}(\int_t^{T_1} a(s,T_2)+a^p(s,T_1)dW_s)\mathbf{1}_{\{A\}} \,\middle|\, \mathcal{F}_t\right]$. From the Girsanov theorem we know that there is a measure $\tilde{P}$ such that

$$\widetilde{W}_t = W_t-\int_0^t a(s,T_2)+a^p(s,T_1)ds$$

is a Brownian Motion under $\tilde{P}$, and

$$d\tilde{P}=\mathcal{E}(\int_t^{T_1} a(s,T_2)+a^p(s,T_1)dW_s)dP.$$

Thus

$$\begin{aligned}&\mathbf{E}^P\left[\mathcal{E}(\int_t^{T_1} a(s,T_2)+a^p(s,T_1)dW_s)\mathbf{1}_{\{A\}} \,\middle|\, \mathcal{F}_t\right]\\ =&\mathbf{E}^{\tilde{P}}\left[\mathbf{1}_{\{A\}}\mid\mathcal{F}_t\right]=\tilde{\mathbf{P}}\left[A\right].\end{aligned}$$

The event A is defined as

$$\begin{aligned}&\ln\bar{S}+\ln\frac{\overline{B}(t,T_1)}{B(t,T_1)}-\ln\frac{\overline{B}(t,T_2)}{B(t,T_2)}\\ &-\frac{1}{2}\int_t^{T_1}\bar{a}(s,T_1)^2-a(s,T_1)^2\,ds+\frac{1}{2}\int_t^{T_1}\bar{a}(s,T_2)^2-a(s,T_2)^2\,ds\\ >&q\int_t^{T_1}(a^p(s,T_2)-a^p(s,T_1))\,dW_s\Leftrightarrow\end{aligned}$$

$$\begin{aligned}
&\ln \bar{S} + \ln \frac{\overline{B}(t,T_1)}{B(t,T_1)} - \ln \frac{\overline{B}(t,T_2)}{B(t,T_2)} \\
&+ \int_t^{T_1} -\frac{1}{2}\bar{a}(s,T_1)^2 + \frac{1}{2}a(s,T_1)^2 + \frac{1}{2}\bar{a}(s,T_2)^2 - \frac{1}{2}a(s,T_2)^2 \\
&- q(a(s,T_2) + a^p(s,T_1))(a^p(s,T_2) - a^p(s,T_1))\,ds \\
>& q\int_t^{T_1} a^p(s,T_2) - a^p(s,T_1) d\tilde{W}_s
\end{aligned}$$

Simple calculations now yield that under the measure $\tilde{P}$, the event A has the probability

$$N(d_1)$$

where

$$\begin{aligned}
d_1 &= \frac{K + F + \frac{1}{2}V}{\sqrt{V}} \\
K &= \ln \bar{S} + \ln \frac{\overline{B}(t,T_1)}{B(t,T_1)} - \ln \frac{\overline{B}(t,T_2)}{B(t,T_2)} \\
F &= \int_t^{T_1} qa^p(s,T_1)\big[a(s,T_2) - a(s,T_1) \\
&\quad - (1-q)(a^p(t,T_2) - a^p(t,T_1))\big]\,ds \\
V &= \int_t^{T_1} q^2(a^p(s,T_2) - a^p(s,T_1))^2\,ds
\end{aligned}$$

The second term:
The second term in the decomposed payoff is valued along the same lines. First we prepare the Girsanov transformation:

$$\begin{aligned}
I_2 &:= e^{-\int_t^{T_1} r(s)+\lambda(s)ds}\overline{B}(T_1,T_2)\mathbf{1}_{\{A\}} && \text{(A.9)} \\
&= \mathbf{1}_{\{A\}}\overline{B}(t,T_2)e^{-(1-q)\int_t^{T_1}\lambda(s)ds}\exp\{-\frac{1}{2}\int_t^{T_1}\bar{a}(s,T_2)^2ds + \int_t^{T_1}\bar{a}(s,T_2)dW_s\} \\
&= \mathbf{1}_{\{A\}}\overline{B}(t,T_2)P(t,T_1)^{1-q} && \text{(A.10)} \\
&\exp\{-\frac{1}{2}\int_t^{T_1}\bar{a}(s,T_2)^2 + (1-q)a^p(s,T_1)^2ds && \text{(A.11)} \\
&+ \int_t^{T_1}\bar{a}(s,T_2) + (1-q)a^p(s,T_1)dW_s\} \\
&= \mathbf{1}_{\{A\}}\overline{B}(t,T_2)P(t,T_1)^{1-q} && \text{(A.12)} \\
&\exp\{\int_t^{T_1}(1-q)a^p(s,T_1)\big[\bar{a}(s,T_2) - \frac{1}{2}q\bar{a}(s,T_1)\big]ds\} && \text{(A.13)} \\
&+ \mathcal{E}\Big(\int_t^{T_1}\bar{a}(s,T_2) + (1-q)a^p(s,T_1)dW_s\Big) && \text{(A.14)}
\end{aligned}$$

Girsanov: There is a measure $\bar{P}$ such that

$$\widetilde{W}_t = W_t - \int_0^t \bar{a}(s,T_2) + (1-q)a^p(s,T_1)ds$$

is a Brownian Motion under $\bar{P}$, and

$$d\bar{P} = \mathcal{E}(\int_t^{T_1} \bar{a}(s,T_2) + (1-q)a^p(s,T_1)dW_s)\, dP.$$

Thus

$$\begin{aligned}
&\mathbf{E}^P\left[\mathcal{E}(\int_t^{T_1} \bar{a}(s,T_2) + (1-q)a^p(s,T_1)dW_s)\mathbf{1}_{\{A\}} \,\middle|\, \mathcal{F}_t\right] \\
=&\mathbf{E}^{\bar{P}}\left[\,\mathbf{1}_{\{A\}} \mid \mathcal{F}_t\right] = \bar{\mathbf{P}}[\,A\,].
\end{aligned}$$

The event A is defined as

$$\begin{aligned}
&\ln \bar{S} + \ln \frac{\overline{B}(t,T_1)}{B(t,T_1)} - \ln \frac{\overline{B}(t,T_2)}{B(t,T_2)} \\
&-\frac{1}{2}\int_t^{T_1} \bar{a}(s,T_1)^2 - a(s,T_1)^2\, ds + \frac{1}{2}\int_t^{T_1} \bar{a}(s,T_2)^2 - a(s,T_2)^2\, ds \\
>&q\int_t^{T_1} (a^p(s,T_2) - a^p(s,T_1))\, dW_s \\
\Leftrightarrow& \\
&\ln \bar{S} + \ln \frac{\overline{B}(t,T_1)}{B(t,T_1)} - \ln \frac{\overline{B}(t,T_2)}{B(t,T_2)} \\
&+\int_t^{T_1} -\frac{1}{2}\bar{a}(s,T_1)^2 + \frac{1}{2}a(s,T_1)^2 + \frac{1}{2}\bar{a}(s,T_2)^2 - \frac{1}{2}a(s,T_2)^2 \\
&- q(\bar{a}(s,T_2) + (1-q)a^p(s,T_1))(a^p(s,T_2) - a^p(s,T_1))\, ds \\
>&q\int_t^{T_1} a^p(s,T_2) - a^p(s,T_1)d\bar{W}_s.
\end{aligned}$$

After short calculations we reach the probability $\bar{\mathbf{P}}[\,A\,]$ as $N(d_2)$, where d_2 is given by

$$d_2 = d_1 - \sqrt{V}.$$

Combination of the results yields the claim of the proposition.

□

A.5 Proof of Proposition 25

The pricing of the plain put option on a defaultable bond is very similar to the pricing of the credit spread put.

Proof: (of Proposition 25)
Step 1: The in-the-money event A:

$$A = \{\omega \in \Omega \,|\, \bar{S} > \overline{B}(T_1, T_2)\}$$

A is equivalent to

$$\int_t^{T_1} (\bar{a}(s, T_2) - \bar{a}(s, T_1))dW_s < \ln \frac{\bar{S}\overline{B}(s, T_1)}{\overline{B}(t, T_2)} + \int_t^{T_1} -\frac{1}{2}\bar{a}(s, T_2)^2 + \frac{1}{2}\bar{a}(s, T_1)^2 ds.$$

Step 2: Decomposing the payoff:

$$\begin{aligned}
&\mathbf{E}\left[e^{-\int_t^{T_1} r(s)+\lambda(s)ds}(\bar{S} - \overline{B}(T_1, T_2))^+ \,\middle|\, \mathcal{F}_t \right] \\
=&\bar{S}\mathbf{E}\left[e^{-\int_t^{T_1} r(s)+\lambda(s)ds}\mathbf{1}_{\{A\}} \,\middle|\, \mathcal{F}_t \right] - \mathbf{E}\left[e^{-\int_t^{T_1} r(s)+\lambda(s)ds}\overline{B}(T_1, T_2)\mathbf{1}_{\{A\}} \,\middle|\, \mathcal{F}_t \right]
\end{aligned}$$

Step 3: The first expectation:

Simplification:

$$I_1 := \bar{S}e^{-\int_t^{T_1} r(s)+\lambda(s)ds}\mathbf{1}_{\{A\}} = \bar{S}\overline{B}_0(t, T_1)\mathcal{E}(\int_t^{T_1} a(s, T_1) + a^p(s, T_1)dW_s)\mathbf{1}_{\{A\}}$$

For this term it remains to evaluate

$$\mathbf{E}\left[\mathcal{E}(\int_t^{T_1} a(s, T_1) + a^p(s, T_1)dW_s)\mathbf{1}_{\{A\}} \,\middle|\, \mathcal{F}_t \right].$$

Girsanov:
This is done using a change of measure to the survival contingent measure introduced in lemma 21. There is a measure $\tilde{P}$ such that

$$\widetilde{W}_t = W_t - \int_0^t a(s, T_1) + a^p(s, T_1)ds$$

is a Brownian Motion under $\tilde{P}$, and

$$d\tilde{P} = \mathcal{E}(\int_t^{T_1} a(s, T_1) + a^p(s, T_1)dW_s)dP.$$

Thus

$$\mathbf{E}^P\left[\mathcal{E}(\int_t^{T_1} a(s, T_1) + a^p(s, T_1)dW_s)\mathbf{1}_{\{A\}} \,\middle|\, \mathcal{F}_t \right]$$

$$=\mathbf{E}^{\tilde{P}}\left[\,1_{\{A\}}\mid\mathcal{F}_t\right]=\tilde{\mathbf{P}}[\,A\,].$$

Probability of A under $\tilde{P}$:

$$\begin{aligned}
&\int_t^{T_1}(\overline{a}(s,T_2)-\overline{a}(s,T_1)))dW_s\\
&=\int_t^{T_1}(\overline{a}(s,T_2)-\overline{a}(s,T_1)))d\widetilde{W}_s+\int_t^{T_1}(\overline{a}(s,T_2)-\overline{a}(s,T_1))(a(s,T_1)+a^p(s,T_1))ds\\
&<\ln\frac{\overline{S}\overline{B}(t,T_1)}{\overline{B}(t,T_2)}+\int_t^{T_1}-\frac{1}{2}\overline{a}(s,T_2)^2+\frac{1}{2}\overline{a}(s,T_1)^2ds.
\end{aligned}$$

This event has the probability $N(d_1)$ where d_1 is given by

$$d_1=\frac{K+F_1+\frac{1}{2}V}{\sqrt{V}}$$

where

$$\begin{aligned}
K&=\ln\frac{\overline{S}\overline{B}(t,T_1)}{\overline{B}(t,T_2)}\\
F_1&=-\int_t^{T_1}(\overline{a}(s,T_2)+a(s,T_1)+a^p(s,T_1))(\overline{a}(s,T_2)-\overline{a}(s,T_1))ds\\
V&=\int_t^{T_1}(\overline{a}(s,T_2)-\overline{a}(s,T_1))^2ds.
\end{aligned}$$

Step 4: The second expectation:

Simplification:

I_2 here is of the same form as I_2 in the proof of the credit spread put option. Therefore its value is given by expression (A.14):

$$\begin{aligned}
I_2:=&e^{-\int_t^{T_1}r(s)+\lambda(s)ds}\overline{B}(T_1,T_2)\mathbf{1}_{\{A\}}\\
=&\overline{B}(t,T_2)P(t,T_1)^{1-q}\exp\{(1-q)\int_t^{T_1}a^p(s,T_1)(\overline{a}(s,T_2)-\frac{1}{2}qa^p(s,T_1))ds\}\\
&\mathcal{E}(\int_t^{T_1}\overline{a}(s,T_2)+(1-q)a^p(s,T_1)dW_s)\mathbf{1}_{\{A\}}
\end{aligned}$$

Girsanov:

The appropriate change of measure also carries through from the credit spread put option. There is a measure $\tilde{P}$ such that

$$\widetilde{W}_t=W_t-\int_0^t\overline{a}(s,T_2)+(1-q)a^p(s,T_1)ds$$

is a Brownian Motion under $\tilde{P}$, and

$$d\tilde{P} = \mathcal{E}\left(\int_t^{T_1} \bar{a}(s,T_2) + (1-q)a^p(s,T_1)dW_s\right)dP.$$

Thus

$$\begin{aligned}&\mathbf{E}^P\left[\mathcal{E}\left(\int_t^{T_1} \bar{a}(s,T_2) + (1-q)a^p(s,T_1)dW_s\right)\mathbf{1}_{\{A\}} \,\middle|\, \mathcal{F}_t\right]\\ =&\mathbf{E}^{\tilde{P}}\left[\,\mathbf{1}_{\{A\}} \mid \mathcal{F}_t\right] = \tilde{\mathbf{P}}[\,A\,].\end{aligned}$$

After short calculations the event A is found to have under $\tilde{P}$ the probability $N(d_2)$, where d_2 is given by

$$d_2 = \frac{K + F_2 - \frac{1}{2}V}{\sqrt{V}}$$

where

$$\begin{aligned}K &= \ln\frac{S\overline{B}(t,T_1)}{\overline{B}(t,T_2)}\\ F_2 &= F_1 + q\int_t^{T_1}(\bar{a}(s,T_2) - \bar{a}(s,T_1))a^p(s,T_1)ds\\ V &= \int_t^{T_1}(\bar{a}(s,T_2) - \bar{a}(s,T_1))^2 ds.\end{aligned}$$

Combining these results yields the claim of the proposition.

□

Appendix B

Calculations to the CIR Model

B.1 Proof of Lemma 28

Proof: (of Lemma 28)
Consider the process x with the P-dynamics

$$dx = (\alpha - \beta x)dt + \sigma\sqrt{x}\, dW$$

and the process $y = cx$. Note that y has the dynamics

$$\begin{aligned} dy = d(cx) &= (c\alpha - \beta y)dt + \sigma\sqrt{c}\,\sqrt{y}\, dW \\ &=: (\hat{\alpha} - \hat{\beta} y)dt + \hat{\sigma}\sqrt{y}\, dW, \end{aligned} \tag{B.1}$$

which is still of the CIR square-root form.

(i) and (ii)
The proof of points (i) and (ii) of lemma 28 is an application of the Girsanov theorem and the change of measure technique: Define

$$g(0,t) = e^{-\int_0^t y(s)ds}. \tag{B.2}$$

Then

$$Z(t) := \frac{1}{G(y(0), 0, T)} g(0,t) G(y(t), t, T) \tag{B.3}$$

is a positive martingale with initial value $Z(0) = 1$ and can therefore be used as a Radon-Nikodym density for a change of measure from P to $\bar{P}_c$ defined by $d\bar{P}_c/dP = Z$, and

$$Z(t) = \mathcal{E}\left(-\int_0^t \sigma c\sqrt{x(s)} H_{2x}(T-s, c) dW(s)\right).$$

Then

$$Z(t)\mathrm{E}^{\bar{P}_c}\left[\, F(x(T)) \mid \mathcal{F}_t \right] = \mathrm{E}^{P}\left[\, Z(T) F(x(T)) \mid \mathcal{F}_t \right] = \mathrm{E}^{P}\left[e^{-\int_t^T cx(s)ds} F(x(T)) \,\middle|\, \mathcal{F}_t \right]$$

(iii)
The P dynamics of x are

$$dx = (\alpha - \beta x)dt + \sigma\sqrt{x}\,dW$$

where W is a P-Brownian motion. By Girsanov's theorem

$$dW = d\tilde{W}^c - H_{2x}(T-t,c)c\sigma\sqrt{x}dt$$

where $\tilde{W}^c$ is a $\tilde{P}_c$-Brownian motion. Thus

$$dx = (\alpha - \beta x)dt - H_{2x}(T-t,c)c\sigma^2 x dt + \sigma\sqrt{x}\,d\tilde{W}^c. \tag{B.4}$$

The distribution of $x(T)$ under $\tilde{P}_c$ can be found e.g. in Jamshidian [1996] or Schlögl [1997]. There the distribution is only given for $c = 1$, but by applying these results to y equations (4.79)-(4.82) follow directly.

(iv)
The dynamics of the other factors remain unchanged because the change of measure here does not affect them. (Z does not depend on any of the other factors.)

□

B.2 Proof of Proposition 29

Proof: (of Proposition 29)
We have to evaluate

$$\mathbf{E}\left[\lambda(t)e^{-\int_0^t \lambda(s)ds}e^{-\int_0^t r(s)ds}\right].$$

First, simplify the expression in the expectation operator:

$$\begin{aligned}
&\lambda(t)e^{-\int_0^t \lambda(s)+r(s)ds}\\
&=\left(\sum_{i=1}^{n}\overline{w}_i x_i(t)\right)e^{\sum_{j=1}^{n}-\int_0^t(w_j+\overline{w}_j)x_j(s)ds}\\
&=\sum_{i=1}^{n}\left(\overline{w}_i x_i(t)e^{\sum_{j=1}^{n}-\int_0^t(w_j+\overline{w}_j)x_j(s)ds}\right).
\end{aligned}$$

Looking at the i-th summand:

$$\begin{aligned}
&\mathbf{E}\left[\overline{w}_i x_i(t)e^{\sum_{j=1}^{n}-\int_0^t(w_j+\overline{w}_j)x_j(s)ds}\right]\\
=&\mathbf{E}\left[\overline{w}_i x_i(t)e^{-\int_0^t(w_i+\overline{w}_i)x_i(s)ds}\right]\mathbf{E}\left[e^{\sum_{j\neq i}^{n}-\int_0^t(w_j+\overline{w}_j)x_j(s)ds}\right]\\
=&\mathbf{E}\left[\overline{w}_i x_i(t)e^{-\int_0^t(w_i+\overline{w}_i)x_i(s)ds}\right]\prod_{j\neq i}\mathbf{E}\left[e^{-\int_0^t(w_j+\overline{w}_j)x_j(s)ds}\right]
\end{aligned}$$

$$=\overline{w}_i \mathbf{E}\left[x_i(t) e^{-\int_0^t (w_i+\overline{w}_i) x_i(s)ds} \right] \prod_{j\neq i} \overline{B}_{j0}(0,t),$$

where $\overline{B}_{j0}(0,t)$ is defined as in the proposition.
It remains to evaluate $\mathbf{E}\left[\overline{w}_i x_i(t) e^{-\int_0^t (w_i+\overline{w}_i) x_i(s)ds} \right]$. For this we change the measure according to lemma 28 choosing $c = w_i + \overline{w}_i$. Then

$$\mathbf{E}\left[x_i(t) e^{-\int_0^t c x_i(s)ds} \right] = \overline{B}_{i0}(0,t)\, \mathbf{E}^{\tilde{P}_c} [\, x_i(t)\,].$$

By lemma 28 (iii) and lemma 26 equation (4.70) the expectation of $x_i(t)$ under $\tilde{P}_c$ is

$$\begin{aligned}
\mathbf{E}^{\tilde{P}_c} [\, x_i(t)\,] &= \eta_t(\nu_t + \tilde{\lambda}_t) \\
&= c\frac{\sigma_i^2}{4} H_{2i}(t,c)\left(\frac{4\alpha_i}{\sigma_i^2} + \frac{4}{\sigma_i^2}\frac{\frac{\partial}{\partial T} H_{2i}(t,c)}{H_{2i}(t,c)} x_i(0)\right) \\
&= c\alpha_i H_{2i}(t,c) + c\frac{\partial}{\partial T} H_{2i}(t,c) x_i(0).
\end{aligned}$$

Thus, combining all yields

$$\begin{aligned}
&\mathbf{E}\left[\lambda(t) e^{-\int_0^t \lambda(s)ds} e^{-\int_0^t r(s)ds} \right] \\
&= \left(\sum_{i=1}^n \overline{w}_i(\overline{w}_i + w_i)(\alpha_i H_{2i}(t, w_i + \overline{w}_i) + \frac{\partial}{\partial T} H_{2i}(t, w_i + \overline{w}_i) x_i(0)) \right) \prod_{j=1}^n \overline{B}_{j0}(0,t).
\end{aligned}$$

□

B.3 Proof of Proposition 30

Proof: (of Proposition 30)
The expectation, that has to be calculated for the credit spread put, is

$$\begin{aligned}
D^{CSP} &= \mathbf{E}\left[e^{-\int_t^{T_1} \sum_{i=1}^n (w_i+\overline{w}_i) x_i(s)ds} (\bar{S}B(T_1,T_2) - \overline{B}(T_1,T_2))^+ \right] \\
&= \overline{B}_0(t,T_1)\, \mathbf{E}^{\tilde{P}} [\, (\bar{S}B(T_1,T_2) - \overline{B}(T_1,T_2))^+\,]
\end{aligned}$$

where the bond prices are

$$\begin{aligned}
B(T_1,T_2) &= \prod_{i=1}^n H_{1i}(T_2 - T_1, w_i) e^{-H_{2i}(T_2-T_1,w_i) w_i x_i(T_1)} \\
\overline{B}(T_1,T_2) &= \prod_{i=1}^n H_{1i}(T_2 - T_1, w_i + q\overline{w}_i) e^{-H_{2i}(T_2-T_1,w_i+q\overline{w}_i)(w_i+q\overline{w}_i) x_i(T_1)}.
\end{aligned}$$

The measure $\tilde{P}$ is the T_1-survival measure which is reached by changing the measure for each component x_i according to lemma 28 with $c_i = w_i + \overline{w}_i$. Under $\tilde{P}$ the component $x_i(T_1)$ is $\chi_1^2(\nu_i, \lambda_i; \bar{\eta}_i)$ ANC distributed with ν_i and λ_i given in the proposition and

$$\bar{\eta}_i = (w_i + \overline{w}_i)\frac{\sigma_i^2}{4}H_{2i}(T_1 - t, w_i + \overline{w}_i).$$

Then the variable Y in the proposition is defined s.t. e^{-Y} has the same distribution as $\bar{S}B(T_1, T_2)$ under the T_1-survival measure.
The variable Y' is defined to have the distribution of $\overline{B}(T_1, T_2)$ under the T_1-survival measure.

The claim of the proposition now follows directly from equations (4.73) and (4.74) in lemma 26. The formula for the Put-option is reached by setting $\eta_i = 0$ and $\epsilon = y$ in the formula for the credit spread put option.

□

Bibliography

[1] P. Artzner and F. Delbaen. Credit risk and prepayment option. *Astin Bulletin*, 22:81–96, 1992.

[2] P. Artzner and F. Delbaen. Default risk insurance and incomplete markets. Working paper, Université Luis Pasteur et C.N.R.S., Strasbourg, June 1994.

[3] A. Bensoussan, M. Crouhy, and D. Galai. Stochastic equity volatility and the capital structure of the firm. In S. Howison, F. Kelly, and P. Wilmott, editors, *Mathematical Models in Finance*, pages 81–92. The Royal Society, Chapman and Hall, 1995.

[4] T. Björk, G. Di Masi, Y. Kabanov, and W. Runggaldier. Towards a general theory of bond markets. *Finance and Stochastics*, 1(2):141–174, 1997.

[5] T. Björk, Y. Kabanov, and W. Runggaldier. Bond market structure in the presence of marked point process. Working paper, Department of Finance, Stockholm School of Economics, S-113 83 Stockholm, Sweden, July 1996.

[6] F. Black and J. Cox. Valuing corporate securities: Some effects of bond indenture provisions. *Journal of Finance*, pages 351–367, 1976.

[7] F. Black, E. Derman, and W. Toy. A one factore model of interest rates and its application to treasury bond options. *Financial Analysts Journal*, pages 33–39, 1990.

[8] F. Black and M. Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81:637–54, 1973.

[9] P. Brémaud. *Point Processes and Queues*. Springer, Berlin, Heidelberg, New York, 1981.

[10] E. Briys and F. de Varenne. Valuing risky fixed rate debt: An extension. *Journal of Financial and Quantitative Analysis*, 32(2):239–248, June 1997.

[11] R.-R. Chen and L. Scott. Interest rate options in multifactor cox-ingersoll-ross models of the term structure. *The Journal of Derivatives*, 3(2):52–72, 6 1995.

[12] J. Cox, J. E. Ingersoll, and S. A. Ross. An intertemporal general equilibrium model of asset prices. *Econometrica*, 53:363–384, 1985.

[13] J. Cox, J. E. Ingersoll, and S. A. Ross. A theory of the term structure of interest rates. *Econometrica*, 53:385–407, 1985.

[14] R. E. Cumby and M. D. D. Evans. The term structure of credit risk: Estimates and specification. Discussion Paper 219, Financial Markets Group, London School of Economics, 1995.

[15] S. Das, editor. *Credit Derivatives: Trading and Management of Credit and Default Risk*. Wiley Frontiers in Finance. John Wiley & Sons, New York, Chichester, Singapore, 1998.

[16] S. R. Das. Credit risk derivatives. *Journal of Derivatives*, 2(3):7–23, Spring 1995.

[17] S. R. Das and R. K. Sundaram. A direct approach to arbitrage-free pricing of credit derivatives. Working Paper 6635, National Bureau of Economic Research, July 1998.

[18] S. R. Das and P. Tufano. Pricing credit-sensitive debt when interest rates, credit ratings and credit spreads are stochastic. Working paper, Harvard Business School, December 1994.

[19] G. R. Duffee. Estimating the price of default risk. Working paper, Federal Reserve Board, Washington DC, September 1995.

[20] D. Duffie. *Security Markets: Stochastic Models*. Academic Press, San Diego, 1988.

[21] D. Duffie. Forward rate curves with default risk. Working paper, Graduate School of Business, Stanford University, December 1994.

[22] D. Duffie. *Dynamic Asset Pricing Theory*. Princeton University Press, Princeton, 1996.

[23] D. Duffie. First-to-default valuation. Working paper, Graduate School of Business, Stanford University, 1998.

[24] D. Duffie. Credit swap valuation. Working paper, Graduate School of Business, Stanford University, 1999.

[25] D. Duffie and N. Garleanu. Risk and valuation of collateralised debt obligations. working paper, Graduate School of Business, Stanford University, August 1999.

[26] D. Duffie and M. Huang. Swap rates and credit quality. Working paper, Graduate School of Business, Stanford University, October 1994.

[27] D. Duffie and M. Huang. Swap rates and credit quality. *Journal of Finance*, 51:921–949, 1996.

[28] D. Duffie and M. Huang. Swap rates and credit quality. *Journal of Finance*, 51(3):921–949, 7 1996.

[29] D. Duffie and R. Kan. A yield-factor model of interest rates. *Mathematical Finance*, 6: 379–406, 1996.

[30] D. Duffie and D. Lando. Term structures of credit spreads with incomplete accounting information. Working paper, Graduate School of Business, Stanford University, June 1997.

[31] D. Duffie, M. Schroder, and C. Skiadas. Recursive valuation of defaultable securities and the timing of resolution of uncertainty. Working Paper 195, Kellogg Graduate School of Management, Department of Finance, Northwestern University, November 1994. revised March 1995.

[32] D. Duffie and K. J. Singleton. An econometric model of the term structure of interest-rate swap yields. *The Journal of Finance*, 52(4):1287–1322, 1997.

[33] D. Duffie and K. J. Singleton. Modeling term structures of defaultable bonds. *The Review of Financial Studies*, 12(4):687–720, 1999.

[34] K. Düllmann, M. Uhrig-Homburg, and M. Windfuhr. Risk structure of interest rates: an empirical analysis for Deutschemark-denominated bonds. *to appear in: European Financial Management*, 1999.

[35] N. El Karoui and V. Lacoste. Multifactor models of the term structure of interest rates. Working paper, University of Paris VI, July 1992.

[36] B. Flesaker, L. Houghston, L. Schreiber, and L. Sprung. Taking all the credit. *Risk Magazine*, 7:105–108, 1994.

[37] H. Föllmer and D. Sondermann. Hedging of non-redundant contingent claims. Working Paper B-3, Rheinische Friedrich-Wilhelms-Universität Bonn, Sonderforschungsbereich 303: 'Information und die Koordination wirtschaftlicher Aktivitäten', Projektbereich B, May 1985.

[38] J. C. Francis, J. A. Frost, and J. G. Whittaker, editors. *Handbook of Credit Derivatives*. McGraw-Hill, New York, 1999.

[39] J. R. Franks and W. N. Torous. A comparison of financial recontracting in distressed exchanges and Chapter 11 reorganizations. *Journal of Financial Economics*, 35:349–370, 1994.

[40] R. Geske. The valuation of corporate liabilities as compound options. *Journal of Financial and Quantitative Analysis*, 12:541–552, 1977.

[41] D. Heath, R. Jarrow, and A. Morton. Bond pricing and the term structure of interest rates: A new methodology for contingent claims valuation. *Econometrica*, 60:77–105, 1992.

[42] T. S. Y. Ho and S.-B. Lee. Term structure movements and pricing interest rate contingent claims. *Journal of Finance*, 41:1011–1029, 1986.

[43] G. T. Hradsky and R. D. Long. High-yield default losses and the return performance of bankrupt debt. *Financial Analysts Journal*, 38:38–49, 1989.

[44] J. Hull. *Options, Futures and Other Derivative Securities*. Prentice Hall, 1989.

[45] J. Hull and A. White. One factor interest rate models and the valuation of interest rate derivative securities. *Journal of Financial and Quantitative Analysis*, 28:235–254, 1993.

[46] J. Hull and A. White. The impact of default risk on the prices of options and other derivative securities. *Journal of Banking and Finance*, 19(2):299–322, 1995.

[47] J. Jacod and A. N. Shiryaev. *Limit Theorems for Stochastic Processes*. Springer, Berlin, Heidelberg, New York, 1988.

[48] F. Jamshidian. Pricing of contingent claims in the one-factor term structure model. Working paper, Sakura Global Capital, July 1987.

[49] F. Jamshidian. A simple class of square-root interest-rate models. *Applied Mathematical Finance*, 2:61–72, 1995.

[50] F. Jamshidian. Bond futures and option evaluation in the quadratic interest rate model. *Applied Mathematical Finance*, 3:93–115, 1996.

[51] R. Jarrow and D. Madan. Option pricing using the term structure of interest rates to hedge systematic discontinuities in asset returns. *Mathematical Finance*, 5(4):311–336, 10 1995.

[52] R. A. Jarrow, D. Lando, and S. M. Turnbull. A Markov model for the term structure of credit risk spreads. Working paper, Johnson Graduate School of Management, Cornell University, December 1993. revised January 1994.

[53] R. A. Jarrow, D. Lando, and S. M. Turnbull. A markov model for the term structure of credit risk spreads. *The Review of Financial Studies*, 10(2):481–523, 1997.

[54] R. A. Jarrow and S. M. Turnbull. Pricing derivatives on financial securities subject to credit risk. *Journal of Finance*, 50:53–85, 1995.

[55] N. L. Johnson and S. Kotz. *Continuous Univariate Distributions*, volume 2. Houghton Mifflin Co., New York, 1970.

[56] E. P. Jones, S. P. Mason, and E. Rosenfeld. Contingent claims analysis of corporate capital structures: An empirical investigation. *Journal of Finance*, 39:611–627, 1984.

[57] I. Karatzas and S. E. Shreve. *Brownian Motion and Stochastic Calculus*. Springer, Berlin, Heidelberg, New York, 1991.

[58] S. Karlin and H. M. Taylor. *A First Course in Stochastic Processes*. Academic Press, 1981.

[59] S. Karlin and H. M. Taylor. *A Second Course in Stochastic Processes*. Academic Press, 1981.

[60] M. Kijima. Valuation of a credit swap of basket type. working paper, Faculty of Economics, Tokyo Metropolitan University, April 1999.

[61] M. Kijima and Y. Muromachi. Credit events and the valuation of credit derivatives of basket type. working paper, Faculty of Economics, Tokyo Metropolitan University, April 1999.

[62] S. H. Kwan. Firm-specific information and the correlation between individual stocks and bonds. *Journal of Financial Economics*, 40:63–80, 1996.

[63] D. Lando. On Cox processes and credit risky bonds. Working paper, Institute of Mathematical Statistics, University of Copenhagen, March 1994. revised December 1994.

[64] D. Lando. *Three Essays on Contingent Claims Pricing*. PhD thesis, Graduate School of Management, Cornell University, 1994.

[65] D. Lando. On Cox processes and credit risky securities. *Review of Derivatives Research*, 2(2/3):99–120, 1998.

[66] F. B. Lehrbass. A simple approach to country risk. working paper, Zentrales Kreditmanagement, WestLB AG, November 1999.

[67] H. E. Leland. Risky debt, bond covenants and optimal capital structure. *Journal of Finance*, 49:1213–1252, 1994.

[68] H. E. Leland and K. B. Toft. Optimal capital structure, endogenous bankrupcy and the term structure of credit spreads. *Journal of Finance*, 51(3):987–1019, 1996.

[69] D. X. Li. On default correlation: A copula function approach. working paper, Risk Metrics Group, September 1999.

[70] D. X. Li. The valuation of basket credit derivatives. *Credit Metrics Monitor*, 2:34–50, April 1999.

[71] F. Longstaff and E. Schwartz. Interest rate volatility and the term structure: A two factor general equilibrium model. *The Journal of Finance*, 47:1259–82, 1992.

[72] F. Longstaff and E. Schwartz. The pricing of credit risk derivatives. *Journal of Fixed Income*, 5(1):6–14, June 1995.

[73] F. A. Longstaff and E. S. Schwartz. A simple approach to valuing risky fixed and floating rate debt. *The Journal of Finance*, 50(3):789–819, 9 1995.

[74] C. Lotz and L. Schlögl. Default risk in a market model. *to appear in: Journal of Banking and Finance*, 1999.

[75] A. Lucas, P. Klaassen, P. Spreij, and S. Straetmans. An analytic approach to credit risk of large corporate bond and loan portfolios. Research Memorandum 1999-18, Vrije Universiteit Amsterdam, February 1999.

[76] D. Madan and H. Unal. Pricing the risks of default. *Review of Derivatives Research*, 2 (2/3):121–160, 1998.

[77] P. Mathieu and P. d'Herouville. *Les Dérivés de Crédit*. Politique générale, Finance et Marketing. Economica, Paris, collection gestion edition, 1998.

[78] P. Mella-Barral and W. R. M. Perraudin. Strategic debt service. *Journal of Finance*, 52: 531–556, 1997.

[79] R. C. Merton. On the pricing of corporate debt: The risk structure of interest rates. *Journal of Finance*, 29:449–470, 1974.

[80] R. C. Merton. Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, 3:125–144, 1976.

[81] R. C. Merton. On the pricing of contingent claims and the Modigliani-Miller theorem. *Journal of Financial Economics*, 5:241–249, 1977.

[82] R. C. Merton. *Continuous Time Finance*. Blackwell Publishers, Oxford, 1992.

[83] K. R. Miltersen, K. Sandmann, and D. Sondermann. Closed form solutions for term structure derivatives with log-normal interest rates. *The Journal of Finance*, 52(1):409–430, 3 1997.

[84] K. Morton and D. Mayers. Numerical solution of partial differential equations. Oxford University Computing Laboratory, mimeo, 1993.

[85] M. Musiela and M. Rutkowski. *Martingale Methods in Financial Modelling*, volume 36 of *Applications of Mathematics*. Springer, Berlin, Heidelberg, New York, 1997.

[86] I. Nelken. *Implementing Credit Derivatives*. Irwin Library of Investment and Finance. McGraw-Hill, New York, London, Sydney, 1999.

[87] L. Nielsen, J. Saá-Requejo, and P. Santa-Clara. Default risk and interest rate risk: The term structure of default spreads. Working paper, INSEAD, 1993.

[88] Y. A. Pierides. The pricing of credit risk derivatives. *Journal of Economic Dynamics and Control*, 21:1579–1611, 1997.

[89] P. Protter. *Stochastic Integration and Differential Equations*. Springer, Berlin, Heidelberg, New York, 1990.

[90] K. Sandmann and D. Sondermann. On the stability of lognormal interest rate models. Working Paper B-263, University of Bonn, Department of Economics, October 1994.

[91] K. Sandmann and D. Sondermann. A note on the stability of lognormal interest rate models and the pricing of eurodollar futures. *Mathematical Finance*, 7(2):119–126, 1997.

[92] O. Sarig and A. Warga. Some empirical estimates of the risk structure of interest rates. *Journal of Finance*, 44(5):1351–1360, 1989.

[93] E. Schlögl. *Interest Rate Factor Models: Term Structure Dynamics and Derivatives Pricing*. PhD thesis, University of Bonn, Faculty of Economics, 1997.

[94] P. J. Schönbucher. Option pricing and hedging in finitely liquid markets. Master's thesis, University of Oxford, Mathematical Institute, September 1993.

[95] P. J. Schönbucher. The term structure of defaultable bond prices. Discussion Paper B-384, University of Bonn, SFB 303, August 1996.

[96] P. J. Schönbucher. Valuation of securities subject to credit risk. Working paper, University of Bonn, Department of Statistics, February 1996.

[97] P. J. Schönbucher. Modelling defaultable bond prices. Working paper, London School of Economics, Financial Markets Group, July 1997.

[98] P. J. Schönbucher. Pricing credit risk derivatives. Working Paper wp-10, London School of Economics, Financial Markets Group, July 1997.

[99] P. J. Schönbucher. Term structure modelling of defaultable bonds. *The Review of Derivatives Studies, Special Issue: Credit Risk*, 2(2/3):161–192, 1998.

[100] P. J. Schönbucher, D. Epstein, and R. Haber. Pricing Parisian options. *ECMI Newsletter*, 22:22–24, 1997.

[101] P. J. Schönbucher, D. Epstein, N. Mayor, E. Whalley, and P. Wilmott. The valuation of a firm advertising optimally. Working paper, Oxford University, Mathematical Institute, February 1997. To appear in: The Quarterly Review of Economics and Finance.

[102] P. J. Schönbucher, D. Epstein, N. Mayor, E. Whalley, and P. Wilmott. The value of market research when a firm is learning: Real option pricing and optimal filtering. Working paper, Oxford University, Mathematical Institute, February 1997.

[103] P. J. Schönbucher, R. Haber, and P. Wilmott. An American in Paris. Working paper, Oxford University, Mathematical Institute, May 1997.

[104] P. J. Schönbucher and E. Schlögl. Credit derivatives and competition in the loan market. Working paper, University of Bonn, Department of Statistics, November 1996.

[105] P. J. Schönbucher and P. Wilmott. The feedback effect of hedging in illiquid markets. Working paper, Oxford University, Mathematical Institute, October 1993.

[106] P. J. Schönbucher and P. Wilmott. Hedging in illiquid markets: Nonlinear effects. In O. Mahrenholtz, K. Marti, and R. Mennicken, editors, *ICIAM/GAMM 95: Proceedings of the Third International Congress on Industrial and Applied Mathematics, Hamburg, July 3-7 1995: Issue 3: Applied Stochastics and Optimization*, pages 81–84. Special Issue: Zeitschrift für Angewandte Mathematik und Mechanik (ZAMM), 1996.

[107] P. J. Schönbucher and P. Wilmott. The feedback effect of hedging in illiquid markets. *to appear in: SIAM Journal of Applied Mathematics*, 1999.

[108] J. M. Tavakoli. *Credit Derivatives: A Guide to Instruments and Applications*. Wiley Series in Financial Engineering. John Wiley & Sons, New York, Chichester, Singapore, 1998.

[109] O. Vasicek. An equilibrium characterisation of the term structure. *Journal of Financial Economics*, 5:177–188, 1977.

[110] P. Wilmott, S. Howison, and J. Dewynne. *Option Pricing: Mathematical Models and Computation*. Oxford Financial Press, Oxford, 1993.

[111] C. Zhou. Default correlation: An analytical result. technical report, Federal Reserve Board, 1997.

[112] C. Zhou. A jump-diffusion approach to modeling credit risk and valuing defaultable securities. Finance and Economics Discussion Paper Series 1997/15, Board of Governors of the Federal Reserve System, March 1997.

A LIBOR MARKET MODEL WITH DEFAULT RISK

PHILIPP J. SCHONBUCHER

Department of Statistics, Bonn University

December 1999,
this version April 2000

ABSTRACT. In this paper a discrete-tenor model for default risk is developed along the lines of the Libor Market Models by Miltersen / Sandmann / Sondermann (1997) and Brace / Gatarek / Musiela (1997). The effective forward rates and effective forward credit spreads are modelled as diffusion processes with a lognormal volatility structure, recovery is modelled as a fraction of the par value of the defaulted coupon bond. No-arbitrage dynamics of the forward rates and forward spreads are derived, as well as closed-form solutions for defaultable coupon bonds, default swap rates and asset swap rates, and approximate solutions are given for options on default swaps, which can be made exact in a modified modelling framework. Furthermore, the numerical implementation of the model is discussed.

1. INTRODUCTION

In this paper a discrete-tenor model for default risk is developed along the lines of the Libor Market Models by Miltersen / Sandmann / Sondermann (1997) and Brace / Gatarek / Musiela (1997). The effective forward rates and effective forward credit spreads are modelled as diffusion processes with a lognormal volatility structure, recovery is modelled as a fraction of the par value of the defaulted coupon bond. No-arbitrage dynamics of the forward rates and forward spreads are derived, as well as closed-form solutions for defaultable coupon bonds, default swap rates and asset swap rates, and approximate solutions are given for default swaptions. Furthermore, the implementation of the model is discussed.

JEL Classification. G 13.

Key words and phrases. Default Risk, Libor Market Model, Credit Derivatives, Default Swap, Asset Swap.

The author would like to thank David Lando, Lutz Schlögl and Markus Fulmek for helpful discussions. He also thanks seminar participants at the Department of Mathematics of Copenhagen University, the Association for Modern Risk Management at Vienna University, and the Department of Mathematics at Bonn University for comments and suggestions. Financial support by the Deutsche Forschungsgemeinschaft is gratefully acknowledged.

This is a preliminary and incomplete version. Comments and suggestions are welcome. All errors are my own.

Interest-rate models of the market-model class (either Libor-based or Swap-based) are amongst the most widely used interest rate models in practice. Their popularity is due to several factors, one of which is the ease with which these models can be implemented and calibrated to market data. Instead of using a short rate process as fundamental variable, directly observed money market rates are modelled, and because Caplets and Swaptions can be priced with the Black formula, the volatility parameters of the model can be directly calibrated to market prices.

The literature on market models has grown substantially in recent years, and it is impossible to give a full list. Apart from the standard references given above, the mathematical methods in E. Schlögl's (1999) multicurrency extension of the Libor market model are related to the methods in this paper. Schlögl also analyses the problems that arise when several numeraires and martingale measures have to be used in parallel, although his work concentrates on the foreign exchange sector while in this paper we have the additional complication of default risk. The paper by C. Lotz and L. Schlögl (2000) treats the valuation of money market instruments under *counterparty* default risk, but they do not use the market-model framework to describe defaultable term structures of interest rates. References to techniques for the numerical implementation of market models are given in the section on implementation.

If the literature on market models is large, the literature on credit risk modelling is even larger. For a survey of the different modelling approaches we refer the reader to Schönbucher (1997). This paper is in the tradition of the intensity-based modelling approach, where it quickly turned out that there is a close relationship between default-free interest rate models and intensity-based default risk models. Representatives of this approach are Jarrow and Turnbull (1995), Madan and Unal (1998), Duffie and Singleton (1997; 1999), Lando (1998) and Schönbucher (1998; 1999).

The rest of the paper is structured as follows: After the introduction of some notation in the next section, a description of the no-arbitrage conditions in the continuous-time setup following Heath / Jarrow / Morton (1992) for the default-free term structure of interest rates and Schönbucher (1998) is given. This is followed by the introduction and discussion of the probability measures which we are going to use, notably the T-forward measures and the T-survival measures. The changes of drift that are associated with the respective changes of measure are also given. This is used to derive the dynamics of the forward rates and credit spreads under the respective measures.

In the next step, positive recovery is introduced. The recovery model used here is based upon the fractional recovery of par model by Duffie (1998) which has the advantage of closely adhering to real-world recovery proceedings and of recognizing the importance of the distinction between principal and coupon claims. The value of a principal claim under this model is derived.

In the following section some important payoffs are valued: defaultable floating coupon claims, default swaps and asset swap packages. The next section considers the pricing of options on default swaps. Here, we need to introduce the *default swap measure* under which the default swap rates become martingales. Using this probability measure we are able to derive option price formulae similar to the well-known Black-formula. These pricing formulae can either be used in a direct default-swap based model, or after some approximating assumptions in the Libor-based approach of the previous sections.

The paper is concluded with a discussion of the strategy to numerically implement this model for the pricing of more exotic credit derivatives.

2. Notation and Model Setup

As usual the model is set in a filtered probability space $(\Omega, (\mathcal{F}_t)_{(t\geq 0)}, Q)$ where the filtration satisfies the usual conditions, and Q is the spot martingale measure. For convenience we assume a large but finite time-horizon $\overline{\overline{T}}$. Usually, quantities that refer to defaultable bond prices or interest rates carry an overbar.

2.1. **The Default Model.** In this model only a single defaultable issuer is considered. The default time is given by the stopping time τ. It is assumed that default is triggered by the first jump of a Cox process $N(t)$ which has an intensity process $\lambda(t)$. Particularly, the survival probability from t to T is given by

$$\mathrm{E}\left[e^{-\int_t^T \lambda(s)ds} \right], \tag{1}$$

and $N(t) - \int_0^t \lambda(s)ds$ is a martingale. For more details on Cox processes in default-risk modelling see e.g. Lando (1998). For now we restrict the attention to the case of zero recovery, positive recovery will be introduced later on. The survival indicator function is denoted $I(t) := 1_{\{\tau > t\}}$.

2.2. **Bond Prices and Basic Rates.**
The tenor structure: We mainly consider payoffs that only occur on a discrete set

$$T_0, T_1, \ldots, T_K$$

of points in time. This could be coupon and repayment dates for bonds or loans, or reset and settlement dates for derivatives. Consequently, the attention is concentrated on the defaultable and default free bonds that mature at these dates.

The function $\kappa(t) = \min\{k \mid T_k > t\}$ gives the next date in the tenor structure after t. Thus $T_{\kappa(t)-1} \leq t < T_{\kappa(t)}$.

Bond prices: Default-free zero coupon bond prices at time t with maturity T_k are denoted by:

$$B(t, T_k) = B_k(t).$$

Defaultable zero coupon bond prices at time t with maturity T_k are:

$$I(t)\overline{B}(t, T_k) = I(t)\overline{B}_k(t).$$

Note that the influence of the defaults (I) and the pre-default price $\overline{B}$ are separated. The default-risk factor at time t for maturity T_k is:

$$D(t, T_k) = D_k(t) = \frac{\overline{B}_k(t)}{B_k(t)}.$$

The default-risk factors D allow to separate the influence of default risk on the defaultable bond prices from the standard discounting with default-free interest rates.

Forward Rates:
We introduce the following forward rates:

The continuously compounded default-free forward rate

$$f(t, T) = -\frac{\partial}{\partial T} \ln B(t, T)$$

The default-free simply compounded forward rate over $[T_k, T_{k+1}]$ as seen from time t:

$$F(t, T_k, T_{k+1}) = F_k(t) = \frac{1}{\delta}\left(\frac{B_k(t)}{B_{k+1}(t)} - 1\right).$$

The continuously compounded defaultable forward rate

$$\overline{f}(t, T) = -\frac{\partial}{\partial T} \ln \overline{B}(t, T)$$

The defaultable simply compounded forward rate over $[T_k, T_{k+1}]$ as seen from time t:

$$\overline{F}(t, T_k, T_{k+1}) = \overline{F}_k(t) = \frac{1}{\delta}\left(\frac{\overline{B}_k(t)}{\overline{B}_{k+1}(t)} - 1\right).$$

The forward credit spread over $[T_k, T_{k+1}]$ as seen from time t:

$$S(t, T_k, T_{k+1}) = S_k(t) = \overline{F}_k(t) - F_k(t).$$

The discrete-tenor forward default intensity over $[T_k, T_{k+1}]$ as seen from time t:

$$H(t, T_k, T_{k+1}) = H_k(t) = \frac{1}{\delta}\left(\frac{D_k(t)}{D_{k+1}(t)} - 1\right).$$

From these definitions follow the following identities:

$$\frac{B_k}{B_{k+1}} = 1 + \delta F_k \qquad B_k = B_1 \prod_{j=1}^{k-1} (1 + \delta F_j)^{-1}$$

$$\frac{\overline{B}_k}{\overline{B}_{k+1}} = 1 + \delta \overline{F}_k \qquad \overline{B}_k = \overline{B}_1 \prod_{j=1}^{k-1} (1 + \delta \overline{F}_j)^{-1}$$

$$\frac{D_k}{D_{k+1}} = 1 + \delta H_k \qquad D_k = D_1 \prod_{j=1}^{k-1} (1 + \delta H_j)^{-1}$$

Furthermore

$$S_k = H_k(1 + \delta F_k) \tag{2}$$

$$\overline{B}_k = D_k B_k = D_1 \cdot \prod_{j=1}^{k-1}(1 - \delta H_j)^{-1} \cdot B_1 \prod_{j=1}^{k-1}(1 + \delta F_j)^{-1}. \tag{3}$$

2.3. **Dynamics.** In this subsection the volatility structure of the forward rate processes is specified. The Brownian motion W is a N-dimensional standard Q-Brownian motion, and all volatility processes are N-dimensional vector processes.

We want F_k and S_k to have a lognormal volatility structure

$$\frac{dF_k}{F_k} = \mu_k^F dt + \sigma_k^F dW \tag{4}$$

$$\frac{dS_k}{S_k} = \mu_k^S dt + \sigma_k^S dW, \tag{5}$$

where σ_k^F and σ_k^S are constant vectors. The drifts μ_k^F and μ_k^S are more complicated and their full form under the respective martingale measures will be derived later on. Alternatively, we could also model the discrete default intensities H_k to have a lognormal volatility structure:

$$\frac{dH_k}{H_k} = \mu_k^H dt + \sigma_k^H dW. \tag{6}$$

Both alternatives (lognormal S_k and lognormal H_k) have their advantages which are discussed in the section on implementation. We should also point out that the lognormal volatility structure ensures positive forward credit spreads. This may seem a desirable feature of the model but there are (extreme) examples where negative forward spreads may occur (see e.g. Schönbucher (1998)).[1] For the next tenor date the credit spread must be positive, though.

For convenience we also define

$$\frac{d\overline{F}_k}{\overline{F}_k} = \mu_k^{\overline{F}} dt + \sigma_k^{\overline{F}} dW \tag{7}$$

(8)

but here the volatilities are not assumed to be constant. The relationships between the volatilities are:

$$\sigma_k^H = \sigma_k^S - \frac{\delta_k F_k}{1 - \delta_k F_k}\sigma_k^F \tag{9}$$

$$\overline{F}_k \sigma_k^{\overline{F}} = \sigma_k^F F_k + \sigma_k^S S_k = (1 + \delta_k F_k) H_k \sigma_k^H + (1 + \delta_k H_k) F_k \sigma_k^F. \tag{10}$$

Default-free lognormal forward rate volatilities are now market standard, in this case the Black Caplet volatilities can be used directly to calibrate the model (for a thorough

[1] These examples involve extreme correlations between interest-rate movements and defaults. In practice, except for possibly short squeezes in extremely illiquid corporate bond markets, negative credit spreads do not occur.

dicussion of the issues in the calibration of market models see Rebonato (1998) and in particular (1999b)). Directly prescribing the dynamics of the defaultable forward rates $\overline{F}$ on the other hand is problematic because then positive credit spreads cannot be ensured any more, and the no-arbitrage condition (14) may not be satisfied. Therefore we choose either H or S to have a lognormal volatility structure, thus absence of arbitrage is ensured.

3. Drift Restrictions for the Continuous Tenor Case

As a starting point of the model we take the Heath / Jarrow / Morton framework. In this framework the conditions for absence of arbitrage are well-known and the proofs can be found in Heath / Jarrow / Morton (1992) and Schönbucher (1998). The main results are the following:

To ensure absence of arbitrage, the dynamics of the default-free continuously compounded forward rates must satisfy under the spot martingale measure Q

$$df(t,T) = \sigma^f(t,T)\left(\int_t^T \sigma^f(t,s)ds\right)dt + \sigma^f(t,T)dW_Q, \tag{11}$$

or, equivalently, the drift of the continuously compounded T-forward rate at time t must be

$$\sigma^f(t,T)\left(\int_t^T \sigma^f(t,s)ds\right). \tag{12}$$

The defaultable continuously compounded forward rates must satisfy simultaneously (again under Q)

$$d\overline{f}(t,T) = \sigma^{\overline{f}}(t,T)\left(\int_t^T \sigma^{\overline{f}}(t,s)ds\right)dt + \sigma^{\overline{f}}(t,T)dW_Q, \tag{13}$$

and

$$\overline{f}(t,t) = \lambda(t) + f(t,t). \tag{14}$$

The drift restriction on the defaultable forward rates takes the same form as (12)

$$\sigma^{\overline{f}}(t,T)\left(\int_t^T \sigma^{\overline{f}}(t,s)ds\right). \tag{15}$$

Under the spot martingale measure Q, conditions (12), (14) and (15) are sufficient to ensure absence of arbitrage in the market. The resulting dynamics of the bond prices are:

$$\frac{dB(t,T)}{B(t,T)} = r(t)dt - \alpha(t,T)dW_Q(t) \tag{16}$$

$$\frac{d\overline{B}(t,T)}{\overline{B}(t,T)} = (\lambda(t) + r(t))dt - \overline{\alpha}(t,T)dW_Q(t) \tag{17}$$

$$\alpha(t,T) = \int_t^T \sigma^f(t,s)ds \quad \text{and} \quad \overline{\alpha}(t,T) = \int_t^T \sigma^{\overline{f}}(t,s)ds. \tag{18}$$

The solutions to the bond price stochastic differential equations are

(19)

$$B(t,T) = B(0,T)\exp\left\{\int_0^t r(s)ds - \frac{1}{2}\int_0^t \alpha^2(s,T)ds - \int_0^t \alpha(s,T)dW_Q(s)\right\}$$

(20)

$$\overline{B}(t,T) = \overline{B}(0,T)\exp\left\{\int_0^t \lambda(s) + r(s)ds - \frac{1}{2}\int_0^t \overline{\alpha}^2(s,T)ds - \int_0^t \overline{\alpha}(s,T)dW_Q(s)\right\}.$$

4. The Forward- and Survival Measures

There are two new types of probability measures which are particularly well suited for the analysis of this model: the T_k-*forward measure* and the T_k-*survival measure*. To each of these probability measures there is an associated *numeraire*, and probabilities under the respective measure may also be regarded as *state prices* expressed in units of the numeraire.

The default-free probability measures of this section are well-known. The spot-martingale measure is described in most textbooks on quantitative finance, and the T_k forward measure is a standard tool in models of the term structure of interest rates particularly in Gaussian term structure models and in the market models. The introduction of the T_k-forward measure goes back to Jamshidian (1987).

4.1. **Girsanov's Theorem:** Girsanov's theorem[2] describes how the Radon-Nikodym density L of a change of measure from a measure Q to an equivalent measure P determines which processes are Brownian motions or compensated jump processes under the new measure. Here we give a general form of this theorem which is valid for probability spaces that support marked point processes and diffusions.

Theorem 1 (Girsanov Theorem: Marked Point Processes). *Let $(\Omega, (\mathcal{F}_t)_{(t\geq 0)}, Q)$ be a filtered probability space which supports a n-dimensional Q-Brownian motion $W_Q(t)$ and a marked point process $\mu(dq; dt)$ with marker q from the mark space $(E, \mathcal{E})$ whose compensator process takes the form $\nu_Q(dq, dt) = K_Q(dq)\lambda_Q(t)dt$ under Q.*
Let θ be a n-dimensional predictable process and $\Phi(t, q)$ a nonnegative predictable

[2]See Jacod and Shiryaev (1988) and Björk, Kabanov and Runggaldier (1996).

function [3] *with*

$$\int_0^t \|\theta(s)\|^2 ds < \infty, \qquad \int_0^t \int_E |\Phi(s,q)| \, K_Q(dq) \, \lambda_Q(s) ds < \infty$$

for finite t. Define the process L by $L(0) = 1$ *and*

$$\frac{dL(t)}{L(t-)} = \theta(t) dW_Q(t) + \int_E (\Phi(t,q) - 1)(\mu(dq, dt) - \nu_Q(dq, dt)).$$

Assume that $\mathbf{E}^Q[\, L(t) \,] = 1$ *for finite t.*
Then for the probability measure P with

$$dP(t) = L(t) dQ(t) \tag{21}$$

it holds that

$$dW_Q(t) - \theta(t) dt = dW_P(t) \tag{22}$$

defines W_P *as P-Brownian motion and*

$$\nu_P(dq, dt) = \Phi(t,q)\nu_Q(dq, dt) \tag{23}$$

is the predictable compensator of μ *under P.*

The marker q of the point process can be used to model uncertainty in the recovery rate. Without the marker, the theorem becomes:

Theorem 2 (Girsanov Theorem: Point Processes). *Under the assumptions of theorem 1, let* $N(t)$ *be a point process with Q-intensity* $\lambda_Q(t)$. *Let* $\phi(t)$ *be a nonnegative predictable process with* $\int_0^t \phi(s)\lambda_Q(s) ds < \infty$ *for finite t. Define the process L by* $L(0) = 1$ *and*

$$\frac{dL(t)}{L(t-)} = \theta(t) dW_Q(t) + (\phi(t) - 1)(dN(t) - \lambda_Q(t) dt).$$

Assume that $\mathbf{E}^Q[\, L(t) \,] = 1$ *for finite t.*
Then for the probability measure P with $dP(t) = L(t) dQ(t)$ *it holds that*

$$dW_Q(t) - \theta(t) dt = dW_P(t) \tag{24}$$

defines W_P *as P-Brownian motion and*

$$\lambda_P(t) = \phi(t)\lambda_Q(t) \tag{25}$$

is the intensity of N *under P.*

---

[3] In functions of the marker q (like Φ here) *predictability* means measurable with respect to the σ-algebra $\tilde{\mathcal{P}} := \mathcal{P} \otimes \mathcal{E}$. Here $\mathcal{P}$ is the σ-algebra of the predictable processes. See Jacod and Shiryaev (1988) for details.

4.2. **The Subjective Measure P.** The subjective (or historical) probability measure P gives the 'real' probabilities of the events. Because it does not take risk premia into account it cannot be used for pricing. A detailed account of the change of measure from the historical probability measure to the spot martingale measure in the case of credit risk models can be found in Schönbucher (1998), it is therefore not repeated here. Apart from the usual change of drift in the Brownian motions, this change of measure typically results in a significantly higher default intensity λ_Q under Q for investment-grade credits, for lower quality credits the default-risk premium is less pronounced.

4.3. **The Spot Martingale Measure Q.** The spot martingale measure Q is the probability measure, under which the discounted security price processes become martingales. The numeraire to the spot-martingale measure is the continuously compounded savings account. Its inverse is known as the discount factor

$$\beta(t) = e^{-\int_0^t r(s)ds}. \tag{26}$$

Under the spot-martingale measure Q, the value of a random payoff X at time T_k is

$$p = \mathrm{E}^Q[\, \beta(T_k)X \,], \tag{27}$$

and the price of this payoff at time $t \leq T_k$ is

$$p(t) = \mathrm{E}^Q\left[\frac{\beta(T_k)}{\beta(t)} X \;\middle|\; \mathcal{F}_t \right]. \tag{28}$$

Thus $\beta(t)p(t)$, i.e. the price $p(t)$ normalized with the Q-numeraire $1/\beta(t)$, is a Q-martingale, as claimed.

4.4. **The T_k-Forward Measure P_k.** The T_k-forward measure is used to price payoffs that occur at time T_k, and the associated numeraire is $B(t, T_k) = B_k(t)$ the default-free bond that matures at T_k. Equation (27) for the price of payoff X at T_k can be transformed to

$$p = B_k(0)\, \mathrm{E}^Q\left[\frac{\beta(T_k)B_k(T_k)}{B_k(0)} X \right], \tag{29}$$

where we used that $B_k(T_k) = 1$. Now let us consider the Radon-Nikodym density process

$$L_k(t) := \frac{\beta(t)B_k(t)}{B_k(0)} =: \left.\frac{dP_k}{dQ}\right|_{\mathcal{F}_t}. \tag{30}$$

This process is a positive martingale with initial value one

$$L_k(0) = 1, \quad L_k(t) > 0 \quad \forall\, t \leq T_k, \quad \mathrm{E}^Q[\, L_k(t) \,] = L_k(s) \quad \forall\, s \leq t \leq T_k, \tag{31}$$

which justifies its use as a Radon-Nikodym density for a change of measure from Q to a new measure P_k. Thus

$$\begin{aligned} p &= B_k(0)\, \mathrm{E}^Q[\, X L_k(T_k) \,] = B_k(0) \int_\Omega X L_k(T_k)\, dQ \\ &= B_k(0) \int_\Omega X \frac{dP_k}{dQ}\, dQ = B_k(0)\, \mathrm{E}^{P_k}[\, X \,]. \end{aligned} \tag{32}$$

By the change of measure the discount factor $\beta(T_k)$ was removed from the expectation in equation (32). This is often a crucial step in the valuation of a derivative security. We can also see why P_k can be considered to be a set of state prices: Consider a state security p_A for state $A \in \mathcal{F}_{T_k}$. Then by equation (32), $P_k[A]$ is the price of p_A expressed in units of $B_k(0)$.

This change of measure technique would not be useful if there was no way to calculate the new expectation $\mathrm{E}^{P_k}[\,X\,]$ in (32) other than going back and evaluate $\mathrm{E}^{Q}[\,L_k(T_k)X\,]$. Here, Girsanov's theorem allows us to derive the dynamics of the stochastic processes under the T_k-forward measure and thus to directly evaluate expectations under P_k. Analyzing the Radon-Nikodym density given in (31) yields the change of drift to reach the P_k-Brownian motion $dW_k(t)$ from the Q-Brownian Motion $dW_Q(t)$:

$$dW_k(t) := dW_Q(t) + \alpha_k(t)dt, \tag{33}$$

where for $t < T_k$ the process $\alpha_k(t)$ is minus the vector of the volatilities of the default-free zero-coupon bond $B_k(t)$

$$\alpha_k(t) = \int_t^{T_k} \sigma^f(t,s)ds. \tag{34}$$

Note that the default intensity is *not* affected by the change of measure, $\lambda_Q = \lambda_{P_k}$.
In the Libor market-model setup, the primitive quantities are the simply compounded forward rates $F_k(t)$. In terms of these rates the $\alpha_k(t)$ are recursively related to each other through

$$\alpha_{k+1}(t) = \alpha_k(t) + \frac{\delta_k F_k(t)}{1+\delta_k F_k(t)}\sigma_k^F(t) \tag{35}$$

(see e.g. Jamshidian (1997) or Brace, Gatarek, Musiela (1997)), and the change of measure from P_k to P_{k-1} can also be achieved using only forward rates by the following direct consequence of the definition of L_k:

$$\mathrm{E}^{P_k}[\,X\,] = \frac{1}{1+\delta_k F_k(0)}\,\mathrm{E}^{P_{k+1}}[\,(1+\delta_k F_k(T_k))X\,]. \tag{36}$$

4.5. **Default Probabilities under P_k:** Under the P_k-forward measure, D_k is the probability of survival until T_k:

$$D_k(0) = \frac{\overline{B}_k(0)}{B_k(0)} = \frac{1}{B_k(0)}\,\mathrm{E}^Q[\,\beta(T_k)I(T_k)\,] = \mathrm{E}^{P_k}[\,I(T_k)\,] = P_k[\tau > T_k].$$

This relationship holds for all times before default $t < \tau$. The general form of this relation is

$$I(t)D_k(t) = P_k[\tau > T_k \mid \mathcal{F}_t], \tag{37}$$

which also proves that $I(t)D_k(t)$ is a P_k-martingale. Consequently, given survival until T_k, at time T_k the P_{k-1} probability of survival until T_{k+1} is D_{k+1} and the probability of default is $1 - D_{k-1} = \delta_k D_{k-1} H_k$. For small time steps ($\delta_k \to 0$) the default probability per time becomes H_k:

$$\frac{1}{\delta_k}(1 - D_{k+1}) \to H_k \qquad \text{as } \delta_k \to 0.$$

The default probability for the next infinitesimal small time step is known as the *default intensity*, therefore H_k was called the discrete-tenor default intensity. Another interesting expression for the P_{k+1} default probability per time at T_k is

$$\frac{1}{\delta_k}\mathrm{E}^{P_k}\left[\, \tau \le T_{k+1} \mid \mathcal{F}_{T_k} \right] = \frac{1}{\delta_k}(1 - D_{k+1}) = \frac{H_k}{1 + \delta_k H_k} = \frac{S_k}{1 + \delta_k \overline{F}_k}, \tag{38}$$

the default probability per time is the credit spread discounted with the defaultable forward rate.

4.6. The T_k-Survival Measure $\overline{P}_k$:

4.6.1. *Definition:* In the same way that the T_k-forward measure is used to price default-free payoffs at T_k, the T_k-*survival measure* is used to price *defaultable* payoffs at T_k. Assume, the payoff in equation (27) is defaultable. Then

$$p = \mathrm{E}^Q\left[\, \beta(T_k) I(T_k) X \,\right] \tag{39}$$

$$= \overline{B}_k(0)\, \mathrm{E}^Q\left[\frac{\beta(T_k) I(T_k) \overline{B}_k(T_k)}{\overline{B}_k(0)}\, X \right], \tag{40}$$

where we used that $\overline{B}_k(T_k) = 1$. Again we can consider the Radon-Nikodym density process

$$\overline{L}_k(t) := \frac{\beta(t) I(t) \overline{B}_k(t)}{\overline{B}_k(0)} =: \left.\frac{d\overline{P}_k}{dQ}\right|_{\mathcal{F}_t}, \tag{41}$$

which is a Q-martingale, as $\beta(t)I(t)\overline{B}_k(t)$ is the discounted price of a defaultable zero coupon bond. This process is not strictly positive but jumps to zero at default ($I(\tau) = 0$). This means, that the measure $\overline{P}_k$ attaches a weight of *zero* to all events that involve default before T_k:

$$\overline{P}_k(\tau \le T_k) = \mathrm{E}^Q\left[\, \overline{L}_k(T_k) 1_{\{\tau \le T_k\}} \,\right] = 0. \tag{42}$$

because either $\overline{L}_k(T_k)$ or $1_{\{\tau \le T_k\}}$ are zero for all $\omega \in \Omega$. Because it only attaches probability to survival events, this measure is termed the T_k-*survival measure*. The survival measure $\overline{P}_k$ is not equivalent to Q any more, but it is absolutely continuous w.r.t. Q, which is all we need to apply Girsanov's theorem.

There is another intuitive interpretation of the T_k survival measure: It is the measure that is reached when the T_k-forward measure is *conditioned on survival* until T_k. Consider an event $A \in \mathcal{F}_{T_k}$:

$$
\begin{aligned}
\mathrm{E}^{P_k}[\,A \mid \tau > T_k\,] &= \frac{\mathrm{E}^{P_k}[\,AI(T_k)\,]}{\mathrm{E}^{P_k}[\,I(T_k)\,]} = \frac{\mathrm{E}^{P_k}[\,AI(T_k)\,]}{D_k(0)} \\
&= \mathrm{E}^{Q}\left[\,AI(T_k)\frac{1}{D_k(0)}\frac{\beta(T_k)}{B_k(0)}\,\right] = \mathrm{E}^{Q}\left[\,A\frac{\beta(T_k)I(T_k)}{\overline{B}_k(0)}\,\right] \\
&= \mathrm{E}^{\overline{P}_k}[\,A\,]. \qquad (43)
\end{aligned}
$$

The P_k-probability of A conditional on survival equals the probability of A under the $\overline{P}_k$-survival measure. This relationship will provide the basis of the simulation-implementation later on.

4.6.2. *Change of Drift:* Analysing $\overline{L}_k$ yields the components of the change of measure in theorem 2: The intensity factor is zero $\varphi(t) = 0$ (which shows that under $\overline{P}$ defaults have zero probability). The change of drift factors for the Brownian motions are (for $t < T_k$):

$$d\overline{W}_k(t) := dW_Q(t) + \overline{\alpha}_k(t)dt. \qquad (44)$$

where now $\overline{\alpha}_k(t)$ is minus the volatility vector of $\overline{B}_k(t)$

$$\overline{\alpha}_k(t) = \int_t^{T_k} \overline{\sigma}(t,s)ds. \qquad (45)$$

Again, the $\overline{\alpha}_k(t)$ are recursively related through

$$\overline{\alpha}_{k+1}(t) = \overline{\alpha}_k(t) + \frac{\delta_k\overline{F}_k(t)\overline{\sigma}_k(t)}{1+\delta_k\overline{F}_k(t)} \qquad (46)$$

but as the defaultable bond prices or forward rates are not the primitives of our model, it is more convenient to use the representation of $\overline{B}_k(t) = B_k(t)D_k(t)$. This yields directly the $\overline{B}_k$ volatility in (44) as sum of the factor volatilities:

$$\overline{\alpha}_k(t) = \alpha_k(t) + \alpha_k^D(t). \qquad (47)$$

where $\alpha_k^D(t)$ is the volatility vector of $D_k(t)$. The $\alpha_k^D(t)$ are recursively related through

$$
\begin{aligned}
\alpha_{k+1}^D(t) &= \alpha_k^D(t) + \frac{\delta_k H_k(t)\sigma_k^H(t)}{1+\delta_k H_k(t)} \qquad (48)\\
&= \alpha_k^D + \frac{\delta_k S_k}{1+\delta_k\overline{F}_k}\left(\sigma^S - \frac{\delta_k F_k}{1+\delta_k F_k}\sigma_k^F\right). \qquad (49)
\end{aligned}
$$

The following formula is a variant of equation (36), it describes the change from $\overline{P}_k$ to $\overline{P}_{k+1}$ and is a consequence of the form (41) of the density $\overline{L}_k(T_k)$ of the change of measure:

$$\mathrm{E}^{\overline{P}_k}[\,X\,] = \frac{1}{1+\delta_k\overline{F}_k(0)}\,\mathrm{E}^{\overline{P}_{k+1}}[\,(1+\delta_k\overline{F}_k(T_k))X\,]. \qquad (50)$$

4.7. **Change of Measure from Survival- to Forward Measure, the Measure $\overline{P}'_k$:** The change from the survival measure $\overline{P}_k$ to the forward measure P_k will be used frequently later on. This is achieved for $\mathcal{F}_t$-measurable X by

$$\mathrm{E}^{\overline{P}_k}[\,X\,] = \frac{1}{D_k(0)}\,\mathrm{E}^{P_k}[\,I(t)D_k(t)X\,] = \frac{B_k(0)}{\overline{B}_k(0)}\,\mathrm{E}^{P_k}\left[\,I(t)\frac{\overline{B}_k(t)}{B_k(t)}X\,\right]. \tag{51}$$

In particular, the relation between the Brownian motions under the T_k forward measure and the T_k survival measure is (for $t < T_k$)

$$d\overline{W}_k(t) = dW_k(t) - \alpha_k^D(t)dt. \tag{52}$$

If the Cox process properties of $N(t)$ are used and X does not contain any direct reference to defaults[4], then (51) can be further expanded to remove the survival indicator function $I(t)$

$$\mathrm{E}^{\overline{P}_k}[\,X\,] = \frac{B_k(0)}{\overline{B}_k(0)}\,\mathrm{E}^{P_k}\left[\,e^{-\int_0^t \lambda(s)ds}\frac{\overline{B}_k(t)}{B_k(t)}X\,\right], \tag{53}$$

and using equations (19) and (20) this becomes

$$\mathrm{E}^{\overline{P}_k}[\,X\,] = \frac{B_k(0)\overline{B}(0,t)}{\overline{B}_k(0)B(0,t)}\,\mathrm{E}^{P_k}\left[\,\mathcal{E}\left(\int_0^t \alpha^D(s,t)dW_k(s)\right)\frac{\overline{B}_k(t)}{B_k(t)}X\,\right]. \tag{54}$$

where $\mathcal{E}\left(\int_0^t \alpha^D(s,t)dW_k(s)\right) = \exp\{-\frac{1}{2}\int_0^t(\alpha^D(s,t))^2ds + \int_0^t \alpha^D(s,t)dW_k(s)\}$ is the Doleans-Dade exponential. For $t = T_k$ equation (54) simplifies to

$$\mathrm{E}^{\overline{P}_k}[\,X\,] = \mathrm{E}^{P_k}\left[\,\mathcal{E}\left(\int_0^{T_k} \alpha_k^D(s)dW_k(s)\right)X\,\right] =: \mathrm{E}^{P_k}[\,L_k^D(T_k)X\,]. \tag{55}$$

This in turn confirms the change of drift given in equation (52). In equation (55) we introduced the Radon-Nikodym density L_k^D of the change of measure from the T_k forward measure to a variant of the survival measure $\overline{P}_k$, the measure $\overline{P}'_k$. This survival measure is identical to the survival measure $\overline{P}_k$ for all X that do not contain direct references to default, and it is often more convenient to use $\overline{P}'_k$ instead of $\overline{P}_k$.

By equation (52) the density $L_k^D(t)$ has the alternative form

$$\frac{d\overline{P}'_k}{P_k}(t) = L_k^D(t) = e^{-\int_0^t \lambda(s)ds}\frac{\overline{B}_k(t)}{B_k(t)}\frac{B_k(t)}{\overline{B}_k(0)} =: \frac{\gamma(t)D_k(t)}{D_k(0)}, \tag{56}$$

where $\gamma(t) = e^{-\int_0^t \lambda(s)ds}$. In particular, $1/(\gamma(t)D_k(t))$ is a $\overline{P}'_k$-martingale and therefore it is also a $\overline{P}_k$-martingale.

For the practical evaluation of equation (55) it will be useful to note that $L_k^D(t)$ satisfies the stochastic differential equation $dL_k^D(t) = \alpha_k^D(t)dW_k(t)$ and that by equation (48)

$$\alpha_k^D(t) = \alpha_{\kappa(t)}^D(t) + \sum_{l=\kappa(t)}^{k-1} \frac{\delta_l H_l}{1+\delta_l H_l}\sigma_l^H. \tag{57}$$

[4] I.e. τ, $N(t)$) or $I(t)$ do not occur in X, but S or $\overline{B}$ may occur in X.

4.7.1. *Independence:* An important special case is the case of independence between the default-free bond prices B_k and defaults, i.e. D_k and τ. In this case the volatility vectors $\alpha_k(t)$ and $\alpha_k^D(t)$ drive independent processes. This means that the change of drift (47) of the survival measure works on the interest-rate components like the change of drift (33) of the T_k-forward measure. Note that independence of defaults and risk-free interest rates does not mean independence of F and S, but of F and H. If in equation (55) X *only* depends on the default-free term structure of interest rates then $\mathrm{E}^{\overline{P}_k}[\,X\,] = \mathrm{E}^{P_k}[\,X\,]$.

4.8. **The Spot Libor Measure Q' and the Survival Spot Libor Measure $\overline{Q}'$.** Jamshidian (1997) introduced the discrete spot martingale measure Q' under which $W_{\kappa(t)}(t)$ is a Brownian motion, and termed this measure the *spot Libor measure*. This measure is associated with the numeraire of a discretely-rolled zero coupon bond investment

$$B'(t) := B(t, T_{\kappa(t)}) \prod_{j=0}^{\kappa(t)-1} (1 + \delta_j F_j(T_j)),$$

which is achieved by investing 1 in $B(0, T_1)$ at $t = 0$ and then reinvesting the proceeds in B_{k+1} at T_k. The advantage of this measure is that it is ideally suited for simulation, basically we iteratively change the measure to the next forward measure. The change of measure from Q' to P_k has the following change of drift

$$dW_k(t) = dW_{Q'}(t) + \sum_{j=\kappa(t)}^{k-1} \frac{\delta_j F_j(t)}{1 + \delta_j F_j(t)} \sigma_j^F =: dW_{Q'}(t) + \alpha_k'(t). \tag{58}$$

Analogously to the default-free spot Libor measure, we can also introduce the defaultable spot Libor measure $\overline{Q}'$, under which $\overline{W}_{\kappa(t)}(t)$ is a Brownian motion. Here the numeraire is the rolled investment in the *defaultable* zero coupon bond with the next maturity, and the changes of drift all follow similar rules:

$$d\overline{W}_k(t) = dW_{\overline{Q}'}(t) + \sum_{j=\kappa(t)}^{k-1} \frac{\delta_j \overline{F}_j(t)}{1 + \delta_j \overline{F}_j(t)} \overline{\sigma}_j =: dW_{Q'}(t) + \overline{\alpha}_k'(t) \tag{59}$$

$$=: dW_{\overline{Q}'}(t) + \alpha_k'(t) + \alpha_k'^D(t). \tag{60}$$

The changes of drift α', $\overline{\alpha}'$ and α'^D are identical to their discrete counterparts when the forward rate volatilities $\sigma^f(t, T)$ and $\sigma^{\overline{f}}(t, T)$ are set to zero until the next maturity (i.e. for $T \leq T_{\kappa(t)}$). Finally, the change from $\overline{Q}'$ to Q' can be achieved by

$$dW_{\overline{Q}'}(t) = dW_{Q'}(t) + \alpha_{\kappa(t)}'^D(t). \tag{61}$$

and for zero forward rate volatilities until $T_{\kappa(t)}$ there is no change of drift between $\overline{Q}'$ and Q'.

5. Drift Restrictions for the Discrete Tenor Case

Using the results of the previous section we can now derive the dynamics of the defaultable and default-free forward rates under the new probability measures. We only give the dynamics of each process under one of the measures, the dynamics under the other measures follow from the change of drift formulae (33), (45) and (51).

5.1. **Default-Free Forward Rates:** For the default-free interest rates it is well-known that B_k/B_{k-1} is a Martingale under the T_{k+1}*-forward measure*. Hence

$$F_k = \frac{1}{\delta_k}(\frac{B_k}{B_{k+1}} - 1) \tag{62}$$

is a martingale under the T_{k+1}*-forward measure*. Its dynamics are therefore (according to the lognormal assumption)

$$dF_k(t) = F_k(t)\sigma_k^F dW_{k-1}(t). \tag{63}$$

The dynamics of $F_k(t)$ under the other measures can now be derived by transforming the P_{k+1}-Brownian motion W_{k+1} into the Brownian motion for the other measures using equations (35) and (48). In particular, under the T_{k+1} survival measure, the dynamics of F_k are

$$dF_k(t) = F_k(t)\sigma_k^F(d\overline{W}_{k+1}(t) - \alpha_{k+1}^D(t)dt). \tag{64}$$

5.2. **Defaultable Forward Rates:** Similar facts hold for the defaultable forward rates: $\overline{B}_k/\overline{B}_{k-1}$ is a martingale[5] under the T_k*-survival measure*, therefore

$$\overline{F}_k = \frac{1}{\delta_k}(\frac{\overline{B}_k}{\overline{B}_{k+1}} - 1) \tag{65}$$

is a martingale under the T_{k+1}*-survival measure*. Again, its dynamics are

$$d\overline{F}_k(t) = \overline{F}_k(t)\overline{\sigma}_k d\overline{W}_{k+1}(t). \tag{66}$$

Calculating the dynamics of $\overline{F}_k$ under the forward measures does not make much sense, as the defaultable forward rates are only meaningful in the survival events.

5.3. **Forward Spreads:** The dynamics of the forward spreads under the T_{k+1} survival measure are

$$dS_k = F_k\sigma_k^F\alpha_{k+1}^D dt + S_k\sigma_k^S d\overline{W}_{k+1}. \tag{67}$$

5.4. **Forward Intensities:** The forward discrete default intensities H_k have the following dynamics under $\overline{P}_{k+1}$:

$$dH_k = F_k\sigma_k^F\left(\frac{1+\delta_k H_k}{1+\delta_k F_k}\alpha_{k+1}^D - \delta_k H_k\sigma_k^H\right)dt + H_k\sigma_k^H d\overline{W}_{k+1}. \tag{68}$$

[5] Strictly speaking $\overline{B}_k/\overline{B}_{k+1}$ is only defined up to default. After default we consider the process stopped.

5.5. **Independence:** If default-free forward rates F and the discrete default intensities D are independent under Q many of the relationships above simplify significantly: Independence in this sense means that

$$\sigma_k^F \sigma_l^D = 0 \qquad \forall k, l \leq K. \tag{69}$$

In particular independence is given when the default-free forward rates F_k and the discrete-time default intensities H_k are driven by different components of the vector Brownian motion W_Q[6].

In this case we also have $\sigma^S \sigma^F = 0 = \alpha^D \sigma^F$, and the discrete default intensities H_k, the credit spreads S_k, the defaultable forward rates $\overline{F}_k$ and the default-free forward rates F_k, in short *all forward rates with fixing at T_k are martingales under the T_{k+1}-survival measure.* In particular, if H_k and F_k are independent under Q, then this independence is preserved and S_k and F_k are also independent under $\overline{P}_{k+1}$.

Even if independence does not hold, the drift of of the default intensities H_k and of the credit spreads S_k is of a small order of magnitude: risk-free interest rates times the covariation between credit spreads and the k-th risk-free forward rate. A good strategy for model calibration with correlation is to first calibrate the model to the closed-form solutions that are reached under the assumption of independence, and then to iteratively adjust the parameters to the case of correlation, which should be not too far away. For pricing purposes, closed-form solutions under independence can be used as control variates to increase the accuracy of simulations.

6. Positive Recovery of Par

6.1. **The Recovery Model:** Most recovery mechanisms in intensity-based models of default risk are specified in terms of recovery on defaultable zero coupon bonds. Recovery is specified for these defaultable zero coupon bonds and for pricing purposes all defaultable claims (coupon bonds) are decomposed into these defaultable zero coupon bonds.

One example of this approach is the *equivalent recovery model*[7] where one defaultable bond $\overline{B}(t, T)$ has a recovery of c equivalent *default-free* bonds $B(t, T)$ at the time of default. Another zero-bond recovery model is the *fractional recovery model* by Duffie / Singleton (1997; 1999). Here a defaultable bond pays off a fraction q of its *pre-default value.* The multiple-defaults model by Schönbucher (1998) yields the same results as the fractional recovery model.

Unfortunately these modelling approaches ignore a fundamental difference between principal and coupon claims in real-world default proceedings where the claim of a creditor on the defaulted debtor's assets is only determined by the outstanding principal

[6] By an orthogonal transformation of W this structure can always be achieved if (69) holds.

[7] This model was used in Jarrow / Turnbull (1995), Jarrow / Lando / Turnbull (1997), Lando (1998) and many others.

and accrued interest payments of the defaulted loan or bond, any future coupon payments do *not* enter the consideration. Thus, the recovery is exclusively determined by the outstanding principal, outstanding future coupons have effectively zero recovery.

This effect becomes important for defaultable bonds with high-coupons, for price differences between defaultable bonds with different coupons, or for defaultable bonds that trade far from their par value. Furthermore, in the case of the equivalent recovery model, not all observed credit spread structures can be explained within the model.

Modelling the recovery of a defaultable bond as a fraction of its par value was first suggested by Duffie (1998), who used this model in an affine term-structure setup but did not model recovery of accrued interest. For a discussion of the disadvantages of traditional recovery models see e.g. Schönbucher (1999).

Bearing this in mind, the defaultable coupon bonds and loans should be decomposed in two distinct classes of elementary claims in a realistic recovery model: zero-recovery claims $\overline{B}(t,T)$, and positive recovery claims $\overline{B}^p(t,T)$ which have a recovery of π times their face value in cash at default.

For example, a defaultable fixed coupon bond with N semiannual fixed coupons of c at T_i, $i = 1,\dots,N$ and a notional of 1 would then be decomposed at time $t < T_1$ into

$$\sum_{i=1}^{N-1} c\overline{B}(t,T_i) + (1+c)\overline{B}^p(t,T_N).$$

At a default there is positive recovery on the notional of 1 and the next outstanding coupon c, therefore we have $(1+c)$ positive recovery bonds $\overline{B}^p$ in the decomposition. To avoid counting this coupon twice we only took zero-recovery coupons up to time T_{N-1}. For defaultable *floating* coupon debt the decomposition would be slightly different but the fundamental idea remains: the coupons have zero recovery, and the principal recovers on the notional and the coupon that was outstanding at the time of default.

The recovery of par model in the discrete-tenor setup is as follows:

Assumption 1 (Recovery of Par). *If a defaultable coupon bond defaults in the time interval* $]T_k, T_{k+1}]$ *then its recovery is composed of the recovery rate* π *times the sum of the notional of the bond (here normalised to* 1*) and the accrued interest over* $]T_k, T_{k+1}]$. *The accrued interest can be*

(a) c*, a constant in the case of a fixed-coupon bond with coupon* c*,*
recovery is $\pi(1+c)$

(b) F_k *in the case of a floating rate bond*[8]*,*
recovery is $\pi(1+\delta_k F_k(T_k))$

The recovery payoffs occur in cash at $T_{\kappa(\tau)}$ *i.e. at the next tenor date* T_{k+1} *if a default was in* $]T_k, T_{k+1}]$.

[8]Defaultable floating rate notes usually pay Libor F plus a constant spread x. In this case recovery is $\pi(1+x+\delta_k F_k(T_k))$

We assume that all claims of the same seniority have the same recovery rate π at the time of default. The recovery rate π can be stochastic in $[0,1]$ but its distribution is assumed to be independent of the default-free interest rates, and time-invariant. For pricing purposes it is in most cases sufficient to work with the expected recovery rate which we will do from now on.

6.2. **Discrete-Tenor Defaults:** Restricting recovery payments to the next tenor date $T_{K(\tau)}$ is not a strong restriction for a number of reasons: First, most defaults do indeed occur on payment dates — at least, they become publicly *apparent* when a payment has to be made and cannot be made. Even if strictly speaking the default had happened between two payment dates, many debtors tend to hang on and hope for resurrection before the next payment is due. Some credit derivatives even define a default event as the event of a missed payment on one (or one of a number) defaultable bond. A missed payment can obviously only occur on a payment date. Second, if the tenor dates are spaced reasonably closely (i.e. quarterly or closer) the error is reduced, too. Third, given the large uncertainty that prevails about recovery rates, the error committed by restricting defaults on the tenor structure is of second order importance.

Finally, the effect of this assumption is a postponement of the default from somewhere in $]T_k, T_{k+1}]$ to T_{k+1}. There is an approximate correction to this effect by adjusting the recovery rate as follows: We assume that continuously compounded short rate r and default intensity λ are constant over $[T_k, T_{k+1}]$. Then, given $H = H_K(T_k)$, $F := F_k(T_k)$ and $\delta := T_{k+1} - T_k$, the default-intensity is $\lambda := \frac{1}{\delta}\ln(1+\delta H)$ and the continuously compounded short rate is $r := \frac{1}{\delta}\ln(1+\delta F)$. Given a default happens in $]T_k, T_{k+1}]$, the T_{k+1}-value of π received at default and invested at r until T_{k+1} is

$$\pi' := \frac{\lambda}{\lambda + r}\frac{F(1+\delta H)}{H(1+\delta F)}\pi \geq \pi. \tag{70}$$

Thus, as a correction we can use π' instead of π and work with recovery payoffs at the next tenor date T_{k+1}. Typically, this adjustment amounts to a factor π'/π of 1.005 to 1.02. it increases with high interest rates and long time steps δ_k, and is rather insensitive to changes in the default intensity λ. A similar adjustment can be constructed for the alternative case when only accrued interest until τ is taken into consideration for the recovery, i.e. for a default at $\tau \in]T_k, T_{k+1}]$ the recovery is $\pi(1+(\tau - T_k)c)$ where c is the coupon. In this case the adjustment will be even smaller.

6.3. **Valuation of the Recovery Payoffs under Independence.** The obvious disadvantage of the 'recovery of par' modelling approach is that we now have to do some work to reach even the price of a simple defaultable coupon bond. For fixed coupon bonds this was not necessary in the equivalent recovery model or the fractional recovery model. Nevertheless, this is not lost labour because we will see that the pricing of default swaps becomes very simple in this setup, and for defaultable *floating* coupon debt there are no simple formulae in the alternative recovery models either.

The indicator function of default in $]T_k, T_{k-1}]$ is $I(T_k) - I(T_{k+1})$. Thus the value of the recovery payoff for this interval only is:

(71)
$$\mathrm{E}^Q[\, \beta(T_{k-1})(I(T_k) - I(T_{k-1}))X_k\,] = \mathrm{E}^Q[\, \beta(T_{k-1})I(T_k)X_k\,] - \mathrm{E}^Q[\, \beta(T_{k+1})I(T_{k+1})X_k\,]$$

where X_k can be either of the alternative recovery payoffs. To lighten notation. the time index is dropped for $t = 0$. i.e. $\overline{B}_k$ stands for $\overline{B}_k(0)$ etc.

6.3.1. *The First Term in* (71). The first term in (71) is of the form

$$\mathrm{E}^Q[\, \beta(T_k)I(T_{k-1})X_k\,] = \overline{B}_k\,\mathrm{E}^{\overline{P}_k}\left[\frac{\beta(T_{k+1})}{\beta(T_k)}X_k\right]$$
$$= \overline{B}_k\,\mathrm{E}^{\overline{P}_k}[\, B_{k+1}(T_k)X_k\,] = \overline{B}_k\,\mathrm{E}^{\overline{P}_k}\left[\frac{1}{1+\delta_k F_k(T_k)}X_k\right],$$

because in all cases of assumption 1, X_k is $\mathcal{F}_{T_k}$-measurable. We now consider each case in turn:

In case (a) $X_k = 1$ and under independence the value is

(72)
$$\overline{B}_k\,\mathrm{E}^{\overline{P}_k}\left[\frac{1}{1+\delta_k F_k(T_k)}X_k\right] = \overline{B}_k\,\mathrm{E}^{\overline{P}_k}[\, B_{k+1}(T_k)\,] = \overline{B}_k\,\mathrm{E}^{P_k}[\, B_{k+1}(T_k)\,] = \overline{B}_k\frac{B_{k+1}}{B_k}.$$

Case (b) with $X_k = (1 + \delta_k F_k(T_k))$ yields (without having to assume independence)

$$\mathrm{E}^Q[\, \beta(T_k)I(T_{k+1})X_k\,] = \overline{B}_k. \tag{73}$$

6.3.2. *The Second Term of* (71). The second term of the difference in (71) expands to

(74)
$$\mathrm{E}^Q[\, \beta(T_{k+1})I(T_{k+1})X_k\,] = \overline{B}_{k+1}\,\mathrm{E}^{\overline{P}_{k+1}}[\, X_k\,] \;= \overline{B}_{k+1}\begin{cases}1 & \text{in case (a)}\\ 1+\delta_k F_k & \text{in case (b)}\end{cases}$$

where we used in case (b) that under independence the default free forward rates are martingales under $\overline{P}_{k+1}$. The solution for case (a) holds in general.

Combining the results, the value of the payoffs if a default happens in the interval $]T_k, T_{k+1}]$ is under independence

in case (a): (fixed coupon)

$$\pi\overline{B}_{k+1}\delta_k H_k \tag{75}$$

in case (b): (floating coupon)

$$\pi\overline{B}_{k+1}\delta_k S_k \tag{76}$$

Furthermore with the solution of equation (74) we also have the value of the floating coupon under zero recovery: A defaultable coupon of $\delta_k F_k$ paid at T_{k-1} with zero recovery at default has the value $\mathrm{E}^Q[\, \beta(T_{k+1})I(T_{k+1})F_k\,] = \overline{B}_{k-1}\delta_k S_k$.

6.4. **Valuation under Correlation of Defaults and Interest Rates.** Under correlation there are only approximative solutions for the different coupon bonds. Again we proceed step-by-step through the different cases:

6.4.1. *The First Term in* (71). **Case (a):** We have to evaluate $\mathrm{E}^{\overline{P}_k}\left[\frac{1}{1+\delta_k F_k(T_k)}\right]$. Changing to $\overline{P}_{k-1}$ yields

$$\mathrm{E}^{\overline{P}_k}\left[\frac{1}{1+\delta_k F_k(T_k)}\right] = \mathrm{E}^{\overline{P}_k}\left[\, B_{k-1}(T_k)\,\right] = \mathrm{E}^{\overline{P}_k}\left[\frac{1}{1+\delta_k F_k(T_k)}\right] \tag{77}$$

$$= \frac{1}{1+\delta_k \overline{F}_k}\,\mathrm{E}^{\overline{P}_{k-1}}\left[\frac{1+\delta_k \overline{F}_k(T_k)}{1+\delta_k F_k(T_k)}\right] = \frac{1}{1+\delta_k \overline{F}_k}\left(1+\delta_k\,\mathrm{E}^{\overline{P}_{k-1}}\left[\, H_k(T_k)\,\right]\right).$$

Case (b): carries through from equation (73)

$$\mathrm{E}^{Q}\left[\,\beta(T_k) I(T_{k+1}) X_k\,\right] = \overline{B}_k. \tag{78}$$

6.4.2. *The Second Term of* (71). For **case (a)** we have already derived $\overline{B}_{k-1}$ as the value in equation (74). The equation in **case (b)** is

$$\mathrm{E}^{\overline{P}_{k-1}}\left[\frac{1}{B_{k+1}(T_k)}\right] = 1+\delta_k\,\mathrm{E}^{\overline{P}_{k+1}}\left[\, F_k(T_k)\,\right]. \tag{79}$$

In equations (77) and (79) we have to evaluate the expectation of certain forward rates under the respective *survival* measures.

For equation (77) we change to the T_k-forward measure using equation (55) to reach

$$\mathrm{E}^{\overline{P}_k}\left[\, B_{k-1}(T_k)\,\right] = \mathrm{E}^{P_k}\left[\, L_k^D(T_k)\frac{B_{k+1}(T_k)}{B_k(T_k)}\right] = \mathrm{E}^{P_k}\left[\, L_k^D(T_k)\frac{1}{1+\delta_k F_k(T_k)}\right].$$

Both $\frac{1}{1+\delta_k F_k}$ and L_k^D are martingales under the T_k-forward measure, thus

$$= \frac{1}{1+\delta_k F_k} + \mathrm{cov}^{P_k}\left(L_k^D(T_k),\ \frac{1}{1+\delta_k F_k(T_k)}\right). \tag{80}$$

Similarly, for (79) we change to the T_{k+1}-forward measure:

$$\mathrm{E}^{\overline{P}_{k+1}}\left[\, F_k(T_k)\,\right] = \mathrm{E}^{P_{k+1}}\left[\, L_{k+1}^D(T_k) F_k(T_k)\,\right].$$

Again, both expressions F_k and L_{k+1}^D are martingales under the T_{k+1}-forward measure

$$= F_k + \mathrm{cov}^{P_{k+1}}\left(L_{k+1}^D(T_k),\ F_k(T_k)\right). \tag{81}$$

There are no closed-form expressions for the covariances in the previous expressions but a common approximation in such cases is to approximate both processes with lognormal processes by setting the stochastic components in the diffusion parameters equal to their values at time $t = 0$[9].

[9] For the default-free market models, Brace / Gatarek / Musiela (1997) interpret this approximation as a first-order chaos expansion. Rebonato (1998) reports very good results for similar approximations.

The volatility of L^D_{k-1} is

$$\alpha^D_{k+1}(t) = \int_t^{T_{k-1}} \sigma^{\bar{f}}(t,s) - \sigma^f(t,s)\, ds$$
$$= \sum_{l=\kappa(t)}^{k} \frac{\delta_l H_l(t)\sigma_l^H(t)}{1-\delta_l H_l(t)} - \int_t^{\kappa(t)} \sigma^{\bar{f}}(t,s) - \sigma^f(t,s)\, ds$$

and approximated

$$\approx \sum_{l=\kappa(t)}^{k} \frac{\delta_l H_l(0)\sigma_l^H(0)}{1+\delta_l H_l(0)}$$

finally,

$$\int_0^{T_m} \alpha^D_{k-1}(t)dt \approx \sum_{l=0}^{k} \frac{\delta_l H_l(0)\sigma_l^H(0)}{1+\delta_l H_l(0)} T_{l\wedge m} =: A^D_{k+1,m}.$$

In the approximation we set the $t=0$ values for H_l and σ_l^H, and ignored the volatilities at the short end before the next tenor time[10].
The dynamics of $X(t) := \frac{1}{1+\delta_k F_k(t)}$ are approximated via

$$dX(t) \approx X(t)(1-F_k(0))\sigma_k^F dW_k,$$

which is easily found by Itô's lemma. Because F_k has constant volatility σ_k^F under P_{k+1} we do not need an approximation for F_k itself.

For the covariance in equation (80) we find thus the approximative value

$$\text{(82)}\qquad \operatorname{cov}^{P_k}\left(L^D_k(T_k), \frac{1}{1+\delta_k F_k(T_k)}\right) \approx \frac{1}{1+\delta_k F_k(0)}\left(e^{A^D_{k,k}(1-F_k(0))\sigma_k^F} - 1\right)$$

and for the covariance in equation (81)

$$\text{(83)}\qquad \operatorname{cov}^{P_{k+1}}\left(L^D_{k+1}(T_k), F_k(T_k)\right) \approx F_k(0)\left(e^{A^D_{k+1,k}\sigma_k^F} - 1\right).$$

The error of this approximation should be very low for reasonable parameter values, it will certainly be an order of magnitude lower than the approximative correlation correction itself which in turn is of the order of a few basis points.

We can now sum up the results of this section:

Proposition 3. *(i) The value of* 1 *at* T_{k+1} *if a default occurs in* $]T_k, T_{k+1}]$ *is under independence*

$$\text{(84)}\qquad e_k := \overline{B}_{k+1}\delta_k H_k$$

[10] Setting the short volatility to zero does not introduce a large error, in fact we only need to set the *correlation with the default-free interest rates* to zero and not the volatility.

(ii) *The value of* $1 - \delta_k F_k(T_k)$ *at* T_{k+1} *if a default occurs in* $]T_k, T_{k+1}]$ *is under independence*

$$\overline{B}_{k+1}\delta_k S_k \tag{85}$$

The value of $F_k(T_k)$ *at* T_{k+1} *if no default occurs until* T_{k+1} *is* $\overline{B}_{k+1} F_k$.

(iii) *Under independence and recovery according to assumption 1 the value of a defaultable bond with notional 1, coupon dates* T_k, $k = 1, \ldots N$ *and*

(a) *with a fixed coupon of c is*

$$\sum_{k=0}^{N-1}(c + (1+c)\pi\delta_k H_k)\overline{B}_{k+1}(0) + \overline{B}_N(0). \tag{86}$$

(b) *with a floating coupon plus spread* x *of* $\delta_k(F_k + x)$ *at* T_{k+1} *is*

$$\sum_{k=0}^{N-1}(F_k + \pi(S_k + xH_k))\delta_k\overline{B}_{k+1}(0) + \overline{B}_N(0). \tag{87}$$

(iv) *Under correlation, the value of a payment of* 1 *at* T_{k+1} *if a default happens in* $]T_k, T_{k+1}]$ *is*

$$\begin{aligned} e_k &:= \delta_k H_k \overline{B}_{k+1} + \overline{B}_k \operatorname{cov}^{P_k}\left(L_k^D(T_k), \frac{1}{1+\delta_k F_k(T_k)}\right) \\ &\approx \delta_k H_k \overline{B}_{k+1} + \frac{\overline{B}_k}{1+\delta_k F_k}\left(e^{(1-F_k(0))A^D_{k,k}\sigma^F_k} - 1\right), \\ &= \delta_k H_k \overline{B}_{k+1} + \overline{B}_{k+1}(1+\delta_k H_k)\left(e^{(1-F_k(0))A^D_{k,k}\sigma^F_k} - 1\right). \end{aligned} \tag{88}$$

and the value of $1 + \delta_k F_k(T_k)$ *at* T_{k+1} *if a default occurs in* $]T_k, T_{k+1}]$ *is*

$$\begin{aligned} &\overline{B}_{k+1}\delta_k S_k - \delta_k \overline{B}_{k+1} \operatorname{cov}^{P_{k+1}}\left(L^D_{k+1}(T_k), F_k(T_k)\right) \\ &\qquad \approx \overline{B}_{k+1}\delta_k S_k - \overline{B}_{k+1}\delta_k F_k\left(e^{A^D_{k+1,k}\sigma^F_k} - 1\right). \end{aligned} \tag{89}$$

The value of $F_k(T_k)$ *at* T_{k+1} *if no default occurs until* T_{k+1} *is* $\overline{B}_{k+1}F_k e^{A^D_{k+1,k}\sigma^F_k}$.

7. Basic Credit Derivatives

In naming the counterparties for credit derivatives we will use the convention that counterparty A will be the insured counterparty (i.e. the counterparty that receives a payoff if a default happens or the party that is long the credit derivative), and counterparty B will be the insurer (who has to pay in default). Party C will be the reference credit.

7.1. Default Swap.

7.1.1. *Description:* In a *default swap* (also known as *credit swap*) **B** agrees to pay the default payment to **A** *if a default has happened.* If there is no default of the reference security until the maturity of the default swap, counterparty **B** pays nothing.

A pays a fee for the default protection. The fee can be either a lump-sum fee up front (default put) or – more commonly – a regular fee at intervals until default or maturity (default swap).

Default swaps mainly differ in the specification of the default payment. Here we only consider the standard default swap without going into the problems of the fine print of the specification of the default payment.

- A pays s at T_i until T_N or default (fee stream)
- B pays the difference between the post-default price of the reference asset (a bond issued by C) and its par value at default (default payment).

7.1.2. *The Fee.* The value of the fee stream can be directly determined as

$$s \sum_{k=1}^{N} \overline{B}_k(0) \tag{90}$$

This valuation is valid for all fee streams of credit derivatives that pay fees until default.

7.1.3. *The Default Payment.* As a preliminary step we consider the pricing of a *default digital put.* This security pays off 1 at T_{k+1} if a default occurred in $]T_k, T_{k+1}]$. From proposition 3 we see that the value of this payoff is

$$D^{\mathrm{DDP}} = \sum_{k=0}^{N-1} e_k, \tag{91}$$

where we have to substitute for e_k from equations (84) or (88) respectively.

Let $C(t)$ be the value of the reference asset and $C_0(t)$ the value of the reference asset under zero recovery[11]. To price the default payment consider the portfolio (a) consisting of the reference asset C and default put on this asset. This portfolio pays off 1 at default (from the default put) and the regular payments in survival.

Compare this to portfolio (b) which is composed of a zero-recovery reference asset C_0 and a default digital put. Both portfolios (a) and (b) have the same payoffs in default and survival, therefore their initial values must be equal, which means

$$C_0 + \sum_{k=0}^{N-1} e_k - C = D^{\mathrm{Def\,Put}}. \tag{92}$$

An important special case is the case of a defaultable coupon bond with coupon c as reference asset. In this case $C_0 = \overline{B}_N + c \sum_k \overline{B}_k$ and $C = C_0 + \pi(1+c) \sum_k e_k$ and

[11]For the case of a fixed-coupon bond the zero-recovery value is obvious, the zero-recovery value of a floating coupon bond follows from proposition 3.

the value of the default put is

$$D^{\text{Def Put}} = (1 - \bar{\pi} - c\bar{\pi}) \sum_{k=0}^{N-1} e_k. \tag{93}$$

7.1.4. *The Default Swap Rate.* The default swap rate is the level s of the fee payment that makes the default swap fairly priced. In the case of the coupon bond as reference asset we have

$$\bar{s} = (1 - \bar{\pi}(1 + c)) \frac{\sum_{k=0}^{N-1} e_k}{\sum_{k=0}^{N-1} \overline{B}_{k+1}} \tag{94}$$

$$= (1 - \bar{\pi}(1 + c)) \sum_{k=0}^{N-1} \overline{w}_k \delta_k H_k \qquad \text{(for independence), and} \tag{95}$$

$$\approx (1 - \bar{\pi}(1 + c)) \sum_{k=0}^{N-1} \overline{w}_k \left(\delta_k H_k + (1 + \delta_k H_k)(e^{(1-F_k(0))A^D_{k,k}\sigma^F_k} - 1)\right) \tag{96}$$

under correlation, where $\overline{w}_k := \overline{B}_{k+1} / \sum_{j=0}^{N-1} \overline{B}_{j+1}$ are the weights in the weighted average representation of the default-swap rate. The weighted-average representation of the default-swap rate is reminiscent of the equivalent representation of a plain vanilla fixed-for-floating interest rate swap rate:

$$s = \sum_{k=0}^{N-1} w_k \delta_k F_k$$

with weights $w_k := B_{k+1} / \sum_{j=0}^{N-1} B_{j+1}$. This property will be useful later on in the pricing of options on default swaps.

7.2. **Asset Swap Packages.**

7.2.1. *Description:* An *asset swap package* is a combination of a defaultable fixed coupon bond (the asset) with a fixed-for-floating interest rate swap whose fixed leg is chosen such that the value of the whole package is the par value of the defaultable bond.

The payoffs of the asset swap package are:
B sells to A for 1 (the nominal value of the C-bond):

- a fixed coupon bond issued by C with coupon c payable at coupon dates t_i, $i = 1, \dots, N$.
- a fixed for floating swap (as below).

The payments of the swap: At each coupon date t_i, $i \leq N$ of the bond

- A pays to B: c, the amount of the fixed coupon of the bond,
- B pays to A: LIBOR + a.

a is called the *asset swap spread* and is adjusted to ensure that the asset swap package has initially indeed the value of 1.

The asset swap is not a credit derivative in the strict sense, because the swap is unaffected by any credit events. Its main purpose is to transform the payoff streams of different defaultable bonds into the same form: *Libor + asset swap spread* (given that no default occurs). **A** still bears the full default risk and if a default should happen, the swap would still have to be serviced.

7.2.2. *Pricing:* To ensure that the value of the asset swap package (asset swap plus bond) to **A** is at par at time $t = 0$ we require:

$$C + (s + a - c)A = 1 \tag{97}$$

where C is the initial price of the bond, s is the fixed-for-floating swap rate for the same maturity and payment dates T_i, and A is the value of an annuity paying 1 at all times T_i, $i = 1, \ldots, N$. All these quantities can be readily observed in the market at time $t = 0$. To ensure that the value of the asset-swap package is indeed one, the asset swap rate must be chosen as

$$a = \frac{1}{A}(1 - C) + c - s.$$

Note that the asset swap rate would explode at a default of **C**, because then $(1 - C(t))$ would change from being very small to a large number. Using the definition of the fixed-for-floating swap rate: $sA = 1 - B_N$ this can be rearranged to yield:

$$Aa = \underbrace{B_N + cA}_{\text{def. free bond}} - \underbrace{C}_{\text{defaultable bond}}, \tag{98}$$

the asset swap rate a is the price difference between the defaultable bond C and an equivalent default free coupon bond (with the same coupon c, it has the price $B_N + cA$) in the numeraire asset A of the swap-measure introduced by Jamshidian (1997).

Asset swap packages are very popular and liquid instruments in the defaultable bonds market, sometimes their market is even more liquid than the market for the underlying defaultable bond alone. They also serve frequently as underlying assets for options on asset swaps, so called *asset swaptions*. An asset swaption gives **A** the right to enter an asset-swap package at some future date T at a pre-determined asset swap spread a.

8. Options on Default Swaps

The weighted-average representation of the default-swap rate in equation (95) and (96) enables us to derive an accurate approximation to the price of an option on a default swap which is based on the approximation given by Brace / Gatarek / Musiela (1997), Andersen and Andreasen (1998) and Zühlsdorff (1999) for the price of a swaption in the Libor market model. Here, we only consider the case of independence between H and F, and to remain in the Libor-modelling framework we need to make some approximations regarding the dynamics of the forward default swap rates, but these simplifications are *not* central to the derivation of the pricing formulae (110) and (112).

It will also be shown how these formulae can be derived from a direct specification of the volatility of the default swap rate alone.

8.1. **Description and Payoffs:** A call on a default swap (*default swaption*) gives the buyer A the right to enter a default swap at at pre-determined spread $\overline{s}^*$ at time T_K.

If a default has occurred before T_K, there are two alternatives: If the default swaption is still alife, A will certainly exercise the default swaption at T_K and immediately receive the payoff $(1-\pi)$. The value of this default-protection component of the default swaption is

$$(1-\pi)(B_K - \overline{B}_K). \tag{99}$$

Alternatively, the default swaption is knocked out at an earlier default[12]. In this case the payoff is zero. Either way, the complicated part of the valuation problem is the pricing of the option in the case when there has not been a default yet, and we will concentrate on the pricing of this option.

At time $t < T_K < T_N$, a default swap with maturity T_N that is entered at time T_K at a default swap rate of $\overline{s}^*$ and that is knocked out at defaults before T_K has the value

$$(\overline{s}(t, T_K, T_N) - \overline{s}^*) \sum_{j=K}^{N-1} \overline{B}_j(t). \tag{100}$$

where $\overline{s}(t, T_K, T_N)$ is the forward default swap rate. According to equation (95) the forward default swap rate is given by

$$\overline{s}(t, T_K, T_N) = \sum_{k=K}^{N-1} \overline{w}_k \delta_k H_k.$$

where now $\overline{w}_k = \overline{B}_{k-1} / \sum_{j=K}^{N-1} \overline{B}_{j-1}$.[13]

If no default has occurred before T_K the default swaption will only be exercised if it is in-the-money at T_K, i.e. if $\overline{s}^* < \overline{s}(T_K, T_K, T_N)$. Then the payoff function of the default swaption is

$$(\overline{s}(T_K) - \overline{s}^*)^+ \sum_{k=K}^{N-1} \overline{B}_k. \tag{101}$$

8.2. **Dynamics of the Forward Default Swap Rate:** To price an option on the default swap we need to know the dynamics of the default swap rate, and most importantly its volatility (the drift will follow from a no-arbitrage argument). Let $H^T := (H_K, H_{K+1}, \dots, H_{N-1})$ denote the vector of forward spreads, and $\overline{w}^T := (\overline{w}_K, \overline{w}_{K+1}, \dots, \overline{w}_{N-1})$ the vector of the weights of these rates in the forward default swap rate. Without loss

[12] This is the case if the option takes the form of an *option to extend* and existing default swap. If the existing default swap has already been triggered, it obviously cannot be extended any more.

[13] This holds under independence, in general the forward default swap rate will be defined similar to (94) and (96).

of generality we set the tenor distances $\delta_k = 1$ equal to one (for general distances the following orthogonality argument would become only slightly more complicated), and we ignore the constant introduced by the positive recovery and the coupon. We also write $\bar{s}$ for the *forward* default swap rate. Note, that

$$(102) \qquad \bar{s} = H^T \bar{w} \quad \text{and} \quad \sum_{k=K}^{N-1} \bar{w}_k = 1 = \mathbf{1}^T \bar{w},$$

where $\mathbf{1}^T = (1, 1, \ldots, 1)$ is a vector composed of ones. Then (given survival) the dynamics of $\bar{s}$ are given by

$$(103) \qquad d\bar{s} = \bar{w}^T dH + H^T d\bar{w} + d < \bar{w}, H >,$$

and the dynamics of H are

$$(104) \qquad dH_i = \ldots dt + H_i \sigma_i^H dW,$$

where σ_i^H are d-dimensional row vectors. For now we are only concerned with the volatility of $\bar{s}$, so we do not yet specify the measure under which dW is a Brownian motion.

In the next step we make two approximations[14]:

(i)

$$H^T d\bar{w} \approx 0$$

"The effect of the changes in the weights are negligible."
Note that by equation (102), we have $\mathbf{1}^T \bar{w} = 0$, so this approximation is exact when H is *flat*. The quality of this approximation therefore depends on the degree, to which
(a) the deviation of H from a flat structure is small
(b) the volatility of $\bar{w}$ itself is small.
Both conditions should be satisfied rather well in practice.

(ii)

$$H_i \sigma_i^H dW = H_i \sum_{j=1}^{d} \sigma_{ij}^H dW_j \approx H_i \sigma_0^H dW_0$$

"The H_i are only driven by parallel shifts."
Here W_0 is a Brownian motion that is reached by a suitable rotation of the other $W_1, \ldots, W_d$, such that the first column of the volatility matrix σ_{ij}^H equals $\sigma_0^H \mathbf{1}$, the other components are effectively ignored. The error of this approximation depends on the weight that higher order components have. Principal component decompositions of the variance/covariance matrix of interest-rates typically exhibit a strongly dominating first component which is almost flat. The larger this component is, the better the approximation will work.

[14]The approximation argument in this subsection is based upon Zühlsdorff (1999) and also Andersen and Andreasen (1998).

Similar approximations have proven to be very precise in the default-free market model setup, so we expect similarly good performance in the default-risk setting. After these approximations the dynamics of $\bar{s}$ become from (103)

$$\begin{aligned} d\bar{s} &= \ldots dt + \sum_{i=K}^{N-1} H_i \overline{w}_i \sigma_0^H dW_0 \\ &= \ldots dt + \bar{s}\sigma_0^H dW_0. \end{aligned} \tag{105}$$

Instead of going throught the approximations above, one could also *directly* specify the dynamics (105) for the forward default swap rate. This is advisable if the defaults and default probabilities for times after T_K do not have to be modelled. Then the H_k need not to be known for $k > K$, and specifying the forward default swap rate and its volatility is sufficient.

8.3. **Pricing the Option, the Default Swap Measure.** The key point to note for pricing is that $\bar{s}(t)\sum_{k=K}^{N-1} I(t)\overline{B}_k(t)$ is a traded asset in the market: It is the value of the default-protection component of the forward default swap over $[T_K, T_N]$. Thus, in analogy to the swap-market measure introduced by Jamshidian (1997) and to the introduction of the survival measure before, we can take

$$X(t) := I(t)\sum_{k=K}^{N-1} \overline{B}_k(t) \tag{106}$$

as numeraire for a new probability measure $\overline{P}^s$, the *default swap measure*. We do not go through the derivation of the Radon-Nikodym density of this measure with respect to the other martingale measures which is exactly analogous to the derivations in the previous sections. The measure $\overline{P}^s$ is associated with the Brownian motion $\overline{W}^s$, and under this measure prices of (defaultable) traded assets divided by the new numeraire $X(t)$ are martingales. The measure is again a *survival-based* measure, e.g. the probability of a default until T_K is zero under the default swap measure.

Starting from the dynamics in equation (105) we now know that $\bar{s}$ is a martingale under $\overline{P}^s$, thus its dynamics are given by the following simple lognormal Brownian motion without drift:

$$d\bar{s} = \bar{s}\sigma_0^H d\overline{W}^s. \tag{107}$$

As mentioned before, we could take the direct specification of the dynamics of $\bar{s}$ under $\overline{P}^s$ in equation (107) as starting point without having to go through the approximations in the previous subsection, and also without having to use independence of H and F.

Using the measure $\overline{P}^s$, the derivation of the price of the option on the default swap is now a simple exercise similar to using the Black formula. Starting from

$$C = \mathrm{E}^Q\left[\beta(T_K)I(T_K)\sum_{k=K}^{N-1} \overline{B}_k(T_K)(\bar{s}(T_K) - \bar{s}^*)^+ \right] \tag{108}$$

the change of measure to $\overline{P}^s$ yields

$$= \left(\sum_{k=K}^{N-1} \overline{B}_k(T_K) \right) \mathbf{E}^{\overline{P}^s} \left[(\overline{s}(T_K) - \overline{s}^*)^+ \right]. \tag{109}$$

Now we can use that $\overline{s}$ is lognormally distributed under $\overline{P}^s$:

$$= \left(\sum_{k=K}^{N-1} \overline{B}_k(T_K) \right) \{\overline{s}(0)N(d_1) - \overline{s}^* N(d_2)\}, \tag{110}$$

where d_1 and d_2 are given by

$$d_{1;2} = \frac{\ln(\overline{s}/\overline{s}^*) \pm (\sigma^H)^2 T_K}{\sigma^H \sqrt{T_K}}. \tag{111}$$

Equations (110) and (111) give the value of a European Call option to enter at time T_K a default swap with maturity T_N and strike default swap rate $\overline{s}^*$, which is knocked out at defaults before $\overline{s}$.

An European Put option to enter as protection *seller* at time T_K a default swap with maturity T_N and strike default swap rate $\overline{s}^*$ has the price

$$P = \left(\sum_{k=K}^{N-1} \overline{B}_k(T_K) \right) \{\overline{s}^* N(-d_2) - \overline{s}(0)N(-d_1)\}, \tag{112}$$

where we assumed that the Put would be out of the money at T_K if a default happened on or before T_K. Note that Put-Call parity between equations (110) and (112) only holds up to defaults before T_K.

9. Numerical Implementation

Because of their great importance in practice there is a quickly growing literature on the implementation and calibration of Libor and Swap market models, and we cannot mention all contributions in this area. The question of calibration is addressed by Rebonato (1998; 1999a; 1999b), advanced techniques for Monte-Carlo simulation can be found e.g. in Glasserman and Zhao (2000) and the survey article by Broadie and Glasserman (1998), and there are numerous empirical studies and papers on the pricing of Bermudan interest-rate products. On the background of this large literature we restrict ourselves here to the details of the implementation that are specific to the case of credit risk modelling.

9.1. **Setup.** First, a choice has to be made whether to model the discrete-time default intensities H_k or the credit spreads S_k as lognormal. Given the scarcity of available data it is unlikely that a statistical test would be able to decide between the two specifications because their effects on the defaultable forward rates are very similar.

Spreads are more intuitive to work with and their drift modification under the survival measure is simpler, but the H_k and their volatilities appear more frequently in the pricing formulae and they are more closely associated with the numeraire of the survival measure and the change between survival and forward measure. It seems that the advantages of having a lognormal H_k outweigh the advantages of lognormal S_k particularly for the simplification in calibration, but this judgement depends on the security to price.

Next, the tenor structure has to be chosen such that all payoff relevant dates are covered and the distances between the dates are not too large. Then the number of driving Brownian motions for the combined model has to be determined. Usually, given the scarcity of data, only one Brownian motion is needed in addition to the Brownian motions that drive the default-free term structure of interest-rates.

9.2. **Calibration:** For details to the calibration of the default-free part of the model the reader is referred to Rebonato (1998; 1999a; 1999b). When calibrating the default-free part of the model one should bear in mind that really *default-free* interest-rates should be used. Libor and swap rates are not default free but carry the default-risk of a AA issuer, and have to be adjusted. A detailed study of the problem of counterparty risk in swaps and FRAs can be found e.g. in Lotz and Schlögl (2000).

Second, the volatility vectors σ_k^H for the H_k have to be specified. Typically, these will involve correlation with the first principal component ('level') of the default-free interest rates and the idiosyncratic movements of the credit spreads / intensities.

Given this information, the defaultable bond prices B_k in the model can be calibrated to observed defaultable bond prices, default swap rates and asset swap rates using the closed-form or approximate solutions given in the paper. If independence between H and F is assumed, then this fitting can be achieved without the need to refer to volatility input. In all cases the expected recovery rate π is needed as an input, too. If desired, the expected recovery rate can be made time-dependent and it can be adjusted for defaults between the tenor dates as described earlier.

9.3. **Simulation:** Monte-Carlo simulation of low-probability events like defaults is *extremely* slow. To avoid this problem we propose a hybrid Monte-Carlo / Default-Tree simulation technique: All diffusion dynamics are simulated with standard Monte-Carlo techniques, but the defaults are modelled in a tree. This combined technique does not require more computing effort than a full simulation because the tree (and the simulation) ends after a branch to default.

The basis for the simulation method is the following reasoning: Let $C(T_k)$ be the value of a credit derivative given $\tau > T_k$, and X_k its payoff at T_{k+1} if a default occurs in

$]T_k, T_{k+1}]$. X_k is assumed to be $\mathcal{F}_{T_k}$–measurable. Then

$$
\begin{aligned}
C(T_k) &= \mathrm{E}^Q\left[\frac{\beta(T_{k+1})}{\beta(T_k)} C(T_{k+1})\right] = B_{k+1}(T_k)\,\mathrm{E}^{P_{k+1}}[\,C(T_{k+1})\,]\\
&= B_{k+1}(T_k)\,\mathrm{E}^{P_{k+1}}[\,C(T_{k+1})I(T_{k+1})\,] + B_{k+1}(T_k)\,\mathrm{E}^{P_{k+1}}[\,C(T_{k+1})(1-I(T_{k+1}))\,]\\
&= B_{k+1}(T_k)\mathrm{E}^{P_{k+1}}[\,C(T_{k+1}) \mid \tau > T_{k+1}\,]\,P_{k+1}[\tau > T_{k+1}]\\
&\quad + B_{k+1}(T_k)\mathrm{E}^{P_{k+1}}[\,C(T_{k+1}) \mid \tau \le T_{k+1}\,]\,P_{k+1}[\tau \le T_{k+1}]\\
&= B_{k+1}(T_k)D_{k+1}(T_k)\,\mathrm{E}^{\overline{P}_{k+1}}[\,C(T_{k+1})\,] + B_{k+1}(T_k)(1-D_{k+1}(T_k))X_{k+1}\\
&= \overline{B}_{k+1}(T_k)\,\mathrm{E}^{\overline{P}_{k+1}}[\,C(T_{k+1})\,] + B_{k+1}(T_k)(1-D_{k+1}(T_k))X_{k+1}.
\end{aligned}
$$

The survival part of the value is evaluated under the T_{k+1}-survival measure, and weighted with the defaultable bond price $\overline{B}_{k+1}(T_k)$. The defaulted part of the value is weighted with the default probability $1 - D_{k+1}(T_k)$ and discounted with the default-free bond price $B_{k+1}(T_k)$.

Starting from T_k the simulation until T_{k+1} proceeds as follows:

- We work under the T_{k+1}-forward measure.
- The P_{k+1} survival probability until T_{k+1} is $D_{k+1}(T_k)$. This is the probability on the survival branch.
- The P_{k+1} default probability until T_{k+1} is $1 - D_{k+1}(T_k)$. This is the probability on the default branch.
- Specify the payoff to the credit derivative in default. Add the default-probability weighted (with $1 - D_{k+1}(T_k)$) discounted (with $B_{k+1}(T_k)$) payoff of the credit derivative to the list of final payoffs of the credit derivative.
- Add any discounted (with $\overline{B}_{k+1}(T_k)$) probability weighted (with the pathwise survival probability until T_k) survival payoffs at T_{k+1} to the list of the final payoffs of the credit derivative.
- The pathwise survival probability until T_{k+1} is the survival probability until T_k times D_k, the survival probability for this time interval.
- Simulate the development of the forward rates and credit spreads until T_{k+1} on the survival branch of the tree. This must be done using the dynamics *given survival*, i.e. under the T_{k+1} *survival* measure $\overline{P}_{k+1}$.

The relevant measure for the simulation is the survival measure $\overline{P}_{\kappa(t)}$ for the next tenor date $T_{\kappa(t)}$, i.e. the simulation takes place under the discrete spot survival measure $\overline{Q}'$, with adjustment for intermediate defaults.

10. Conclusion

In this paper we showed how default risk can be incorporated in the modelling framework of the market models for interest rates. The change of measure technique which already was very prominent for the default-free market models, now becomes the most important tool for analyzing the relationships between the different rates and for pricing and implementation of the model. In particular, a new class of probability measures, the *survival measures*, provides the appropriate tools for the pricing of default-dependent payoffs. These survival measures can be viewed as the probability measures that are reached, when a particular default-free probability measure is conditioned on survival until a certain point in time.

In the modelling of the recovery of defaultable bonds we chose to use the recovery of par modelling approach. This choice of recovery model is justified by its ability to represent real-world recovery rules, but it complicates the pricing of even a simple defaultable coupon bond. On the other hand, it turned out that this choice of recovery model greatly facilitates the pricing of default swaps, and that under independence closed-form solutions are still available. Nevertheless, the extension of the Libor market framework to other specifications of the recovery at default is straightforward: For recovery in equivalent default-free bonds all pricing problems can be reduced to the pricing of zero-recovery bonds, which is already solved here, and the extension to fractional recovery (Duffie / Singleton (1999) and Schönbucher (1998)) should not present any problems either.

Using this modelling framework, we then addressed the pricing of some popular credit derivatives. Most of the work for the pricing of default swaps had already been done in the analysis of the par recovery model, and the pricing formula for asset swap packages is entirely model-independent. To be able to price options on default swaps we again had to transfer and extend notions from the default-free market model world: The introduction of the *default swap measure* — the defaultable analogy of Jamshidian's (1997) swap market measure — enabled us to derive closed-form solutions for these second-generation instruments. As default swaps are becoming more and more liquid and standardised, a modelling approach based on the default swap measure making default swap rates to martingales has much potential for the future.

Finally, the numerical implementation of the model was discussed, and we proposed to use a combined simulation/tree methodology to avoid having to simulate the rare default events directly. Preliminary numerical investigations have shown that such an approach can greatly speed up the simulation.

REFERENCES

Andersen, Leif B.G. and Jesper Andreasen (1998). "Volatility skews and extensions of the libor market model." Working paper, General Re Financial Products.

Björk, T., Y. Kabanov, and W. Runggaldier (1996). "Bond market structure in the presence of marked point process." Working paper, Department of Finance, Stockholm School of Economics, S-113 83 Stockholm, Sweden.

Brace, Alan, Dariusz Gatarek, and Marek Musiela (1997). "The market model of interest rate dynamics." *Mathematical Finance*, 7(2), 127–155.

Broadie, Mark and Paul Glasserman (1998). "Simulation for option pricing and risk management." In Alexander, Carol (ed.), "Risk Management and Analysis: Measuring and Modelling Financial Risk," vol. 1 of *Wiley Series in Financial Engineering*.

chap. 5, 173–207. John Wiley and Sons.

Duffie, Darrel and Kenneth J. Singleton (1997). "An econometric model of the term structure of interest rate swap yields." *Journal of Finance*, 52(4), 1287–1321.

Duffie, Darrell (1998). "Defaultable term structure models with fractional recovery of par." Working paper, Graduate School of Business, Stanford University.

Duffie, Darrell and Kenneth J. Singleton (1999). "Modeling term structures of defaultable bonds." *The Review of Financial Studies*, 12(4), 687–720.

Glasserman, Paul and Xiaoliang Zhao (2000). "Arbitrage-free discretization of lognormal forward libor and swap rate models." *Finance and Stochastics*, 4(1).

Heath, David, Robert Jarrow, and Andrew Morton (1992). "Bond pricing and the term structure of interest rates: A new methodology for contingent claims valuation." *Econometrica*, 60, 77–105.

Jacod, Jean and Albert N. Shiryaev (1988). *Limit Theorems for Stochastic Processes*. Springer, Berlin, Heidelberg, New York.

Jamshidian, Farshid (1987). "Pricing of contingent claims in the one-factor term structure model." Working paper, Sakura Global Capital.

Jamshidian, Farshid (1997). "LIBOR and swap market models and measures." *Finance and Stochastics*, 1(4), 293–330.

Jarrow, Robert A., David Lando, and Stuart M. Turnbull (1997). "A markov model for the term structure of credit risk spreads." *The Review of Financial Studies*, 10(2), 481–523.

Jarrow, Robert A. and Stuart M. Turnbull (1995). "Pricing derivatives on financial securities subject to credit risk." *Journal of Finance*, 50, 53–85.

Lando, David (1998). "On Cox processes and credit risky securities." *Review of Derivatives Research*, 2(2/3), 99–120.

Lotz, Christopher and Lutz Schlögl (2000). "Default risk in a market model." *Journal of Banking and Finance*, 24(1), 301–327.

Madan, Dilip and Haluk Unal (1998). "Pricing the risks of default." *Review of Derivatives Research*, 2(2/3), 121–160.

Miltersen, Kristian R., Klaus Sandmann, and Dieter Sondermann (1997). "Closed form solutions for term structure derivatives with log-normal interest rates." *The Journal of Finance*, 52(1), 409–430.

Rebonato, Riccardo (1998). *Interest-Rate option models*. Wiley Series in Financial Engineering. John Wiley & Sons, Chichester, New York, Weinheim, 2 edn.

Rebonato, Riccardo (1999a). "Calibrating the BGM model." *Risk*, 74–79.

Rebonato, Riccardo (1999b). *Volatility and correlation in the pricing of equity, FX and interest-rate options*. Wiley Series in Financial Engineering. John Wiley & Sons, Chichester, New York, Weinheim.

Schlögl, Erik (1999). "A multicurrency extension of the lognormal interest rate market models." Working paper, School of Mathematical Sciences, University of Technology, Sydney.

Schönbucher, Philipp J. (1997). "The valuation of credit risk and credit derivatives." Working paper, Department of Statistics, Bonn University.

Schönbucher, Philipp J. (1998). "Term structure modelling of defaultable bonds." *Review of Derivatives Research*, 2(2/3), 161–192.

Schönbucher, Philipp J. (1999). "A tree implementation of a credit spread model for credit derivatives." Working paper, Department of Statistics, Bonn University.

Zühlsdorff, Christian (1999). "Extended market models with affine and quadratic volatility." Working paper, University of Bonn, Department of Statistics.

Author's address:
Philipp J. Schönbucher
Department of Statistics, Faculty of Economics, Bonn University
Adenauerallee 24–42, 53113 Bonn, Germany,
Tel: +49 - 228 - 73 92 64, Fax: +49 - 228 - 73 50 50

E-mail address: P.Schonbucher@finasto.uni-bonn.de

URL: http://www.finasto.uni-bonn.de/~schonbuc/

Valuation of Exotic Options Under Shortselling Constraints

Uwe Schmock *
Department Mathematik
ETH Zentrum
CH-8092 Zürich
SWITZERLAND
schmock@math.ethz.ch
http://www.math.ethz.ch/~schmock

Steven E. Shreve †
Department of Mathematical Sciences
Carnegie Mellon University
Pittsburgh, PA 15213
USA
shreve@andrew.cmu.edu

Uwe Wystup
Commerzbank AG
Neue Mainzer Strasse 32–36
60261 Frankfurt am Main
GERMANY
wystup@mathfinance.de
http://www.mathfinance.de

December 17, 1999

Abstract. Options with discontinuous payoffs are generally traded above their theoretical Black-Scholes prices because of the hedging difficulties created by their large delta and gamma values. A theoretical method for pricing these options is to constrain the hedging portfolio and incorporate this constraint into the pricing by computing the smallest initial capital which permits super-replication of the option. We develop this idea for exotic options, in which case the pricing problem becomes one of stochastic control. Our motivating example is a call which knocks out in the money, and explicit formulas for this and other instruments are provided.

Key words: Exotic options, super-replication, stochastic control.
JEL classification: G13
Mathematics Subject Classification (1991): 90A09, 60H30, 60G44

* Financial support by Crédit Suisse Group is gratefully acknowledged.
† Work supported by the National Science Foundation under grants DMS-95-00626 and DMS-98-02464. Part of this paper was written while this author was a visitor at the Institute for Mathematics Research, ETH Zürich, whose support is gratefully acknowledged.

1 In-the-money knock-out call

The results reported in this paper were motivated by the problem of pricing and hedging a particular exotic option, a call which knocks out in the money. More specifically, we assume a geometric Brownian motion model

$$dS(t) = rS(t)\,dt + \sigma S(t)\,dW(t), \quad S(0) > 0, \tag{1.1}$$

for an underlying asset, henceforth called a *stock*, even though it is often some other instrument, such as an exchange rate. The *interest rate* $r \in \mathbb{R}$ and *volatility* $\sigma > 0$ are assumed to be constant. The process $(W(t); 0 \le t \le T)$ is a Brownian motion under a probability measure $\mathbb{P}$ which is *risk-neutral,* i.e., is chosen so that the stock has mean rate of return r. The planning horizon T is a fixed positive number throughout.

We introduce a *knock-out call option* whose payoff at expiration date T is

$$(S(T) - K)^+ I_{\{\max_{0 \le t \le T} S(t) < B\}}, \tag{1.2}$$

where the *strike price* K and the *knock-out barrier* B satisfy $0 < K < B$ and I_A denotes the indicator of the generic event A. This call "knocks out" in the money, which makes the implementation of the Black-Scholes hedging strategy difficult, as we now explain.

If $0 \le t \le T$, $S(t) = x > 0$, and the call has not knocked out prior to time t, then the value of the call at time t is given by the risk-neutral pricing formula

$$v(t,x) \triangleq \mathbb{E}\big[e^{-r(T-t)}(S(T) - K)^+ I_{\{\max_{t \le u \le T} S(u) < B\}} \;\big|\; S(t) = x\big]. \tag{1.3}$$

The joint distribution of the drifted Brownian increment $Y = W(T) - W(t) + \theta(T-t)$ and its maximum $Z = \max_{t \le s \le T}(W(s) - W(t) + \theta(s-t))$ over the interval $[t,T]$ can be derived using Girsanov's theorem (see formula 1.1.8 of [1] or [12], Section 3.5) and is

$$\begin{aligned}
&\mathbb{P}\{Y \in dy,\, Z \in dz\} \\
&\quad = \frac{2(2z-y)}{\tau\sqrt{2\pi\tau}} \exp\left\{-\frac{(2z-y)^2}{2\tau} + \theta y - \frac{1}{2}\theta^2\tau\right\} dy\,dz, \quad z > 0,\ y < z,
\end{aligned} \tag{1.4}$$

where $\tau \triangleq T - t$. Let N denote the standard normal distribution function. Using formula (1.4), $v(t,x)$ can be computed explicitly: For $t \in [0,T)$ and $x \in (0,B]$,

$$\begin{aligned}
v(t,x) = {} & x\big[N(b-\theta_+) - N(k-\theta_+)\big] \\
& + xe^{2b\theta_+}\big[N(b+\theta_+) - N(2b-k+\theta_+)\big] \\
& - Ke^{-r\tau}\big[N(b-\theta_-) - N(k-\theta_-)\big] \\
& - Ke^{-r\tau+2b\theta_-}\big[N(b+\theta_-) - N(2b-k+\theta_-)\big],
\end{aligned} \tag{1.5}$$

where $b \triangleq \frac{1}{\sigma\sqrt{\tau}} \log\frac{B}{x}$, $k \triangleq \frac{1}{\sigma\sqrt{\tau}} \log\frac{K}{x}$ and $\theta_\pm = (\frac{r}{\sigma} \pm \frac{\sigma}{2})\sqrt{\tau}$.

Definition (1.3) implies that $v(t,B) = 0$ for $0 \le t \le T$. For $0 < x \le B$, as $t \uparrow T$, we obtain from (1.5) that $v(t,x)$ approaches the discontinuous limit $v(T,x) = (x-K)^+ I_{\{x<B\}}$. Consequently, for x near B and t near T, the

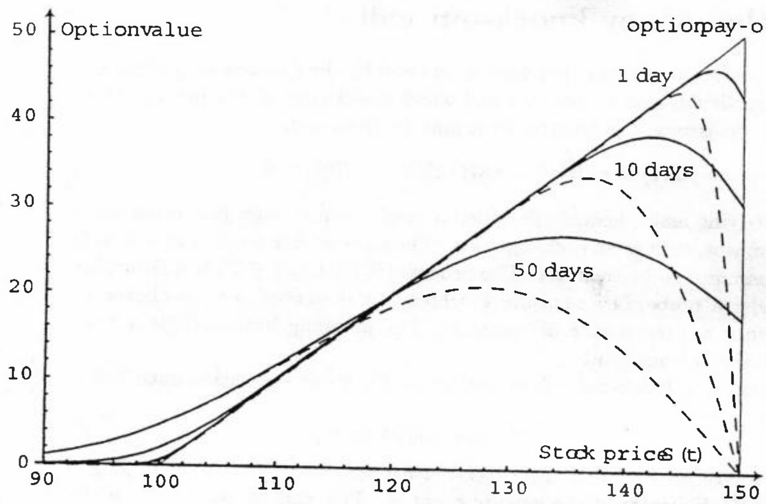

Figure 1: Upper hedging prices $v^*(0, S_0; \alpha)$ given by (4.9) for call options with strike $K = 100$, knock-out barrier $B = 150$ and maturities $T = 50/365$, $T = 10/365$ and $T = 1/365$. We used the interest rate $r = 5\%$, the volatility $\sigma = 30\%$ and the hedge-portfolio constraint $\alpha = 10$. The dashed lines show the corresponding prices given by (1.5) without the hedge-portfolio constraint (2.2).

"delta" $v_x(t,x)$ and "gamma" $v_{xx}(t,x)$ of this option become large in absolute value. The slope and the curvature of the dashed lines in Figure 1 illustrate this. As a result, a trader seeking to hedge a short position in this option will find himself taking a large short position in the underlying stock and making large adjustments to this position frequently. As a practical matter, this naive implementation of the delta hedging strategy is not possible.

A common pricing practice for such options is to *move the barrier*. If one prices and hedges the option as if the barrier were some number $B' > B$, then the dangerous region of high delta and gamma can be moved above B, and the option will knock out before the stock reaches this region. Of course, the computed price of the option increases with increasing B', and there is no clear procedure for choosing an appropriate value for B'. Furthermore, this practice necessarily prices one option at a time, rather than pricing a book of options by taking into account offsetting exposures in the book.

We propose in this paper an alternative to moving the barrier, namely, *constraining the hedging portfolio*. In Section 4 we show that this can be interpreted in terms of the transaction cost associated with liquidating a large short position (Remark 4.4), and also provides a first-order approximation to the price obtained by moving the barrier (Remark 4.3). Furthermore, the theory applies to a book of options as well as to individual options, although the computational issues for a book can be substantial. We work out a simple case of a book of two barrier options in Example 6.7.

2 Model formulation

Throughout this paper, we work within the context of the canonical probability space for Brownian motion. In particular, we take Ω to be the set of continuous functions from $[0,T]$ to $\mathbb{R}$ taking the value zero at zero, we take $\mathbb{P}$ to be Wiener measure, and we take $W(t,\omega) = \omega(t)$ for all $t \in [0,T]$ and all $\omega \in \Omega$. For $0 \le t \le T$, we denote by $\mathcal{F}^W(t)$ the σ-algebra generated by $(W(s); 0 \le s \le t)$. The σ-algebra $\mathcal{F}(T)$ is the $\mathbb{P}$-completion of $\mathcal{F}^W(T)$, and for $0 \le t \le T$, $\mathcal{F}(t)$ is the augmentation of $\mathcal{F}^W(t)$ by the $\mathbb{P}$-null sets of $\mathcal{F}(T)$. A random variable X is $\mathcal{F}(t)$-measurable if and only if there exists an $\mathcal{F}^W(t)$-measurable random variable Y with $\{X \neq Y\} \in \mathcal{F}(T)$ and $\mathbb{P}(X \neq Y) = 0$.

We introduce a contingent claim whose payoff at expiration date T is $g(S(\cdot))$. Let $C_+[0,T]$ denote the space of nonnegative continuous functions on $[0,T]$. We assume that the nonnegative function $g\colon C_+[0,T] \to [0,\infty)$ is lower semicontinuous in the supremum norm topology. The argument of g is the path of the stock price process S from date 0 to date T, and because this path is random, $g(S(\cdot))$ is a random variable on $(\Omega, \mathcal{F}(T), \mathbb{P})$.

The problem of super-replication of a short position in this option can be posed as follows. Let $X(0) > 0$ be a given nonrandom *initial wealth*, and choose an $(\mathcal{F}(t); 0 \le t \le T)$-adapted *portfolio process* $(\pi(t); 0 \le t \le T)$ and *cumulative consumption process* $(C(t); 0 \le t \le T)$. We interpret $\pi(t)$ as the proportion of wealth invested in the stock at time t (sometimes called the *gearing*). The remaining wealth is invested at interest rate r, and $C(t)$ is the amount of wealth consumed up to time t. In particular, $C(t)$ is nondecreasing, right-continuous with left limits (RCLL), and $C(0) = 0$. This leads us to model the differential of wealth as

$$\begin{aligned} dX(t) &= \pi(t)X(t)\frac{dS(t)}{S(t)} + rX(t)(1-\pi(t))\,dt - dC(t) \\ &= rX(t)\,dt + \sigma\pi(t)X(t)\,dW(t) - dC(t). \end{aligned} \tag{2.1}$$

If $X(T) \ge g(S(\cdot))$ almost surely, we say that (π, C) ***super-replicates*** $g(S(\cdot))$ beginning with initial wealth $X(0)$.

We next impose the *portfolio constraint*

$$\pi(t) \ge -\alpha, \quad 0 \le t \le T, \text{ a.s.,} \tag{2.2}$$

where $\alpha \in [0,\infty)$ is some fixed number. The point of this constraint, in the context of the knock-out call of the previous section, is to avoid short positions which are too large relative to the value of the contingent claim being hedged. The parameter α must be chosen by the person pricing the contingent claim; in the case of the knock-out call, we interpret α in terms of a transaction cost in Remark 4.4, and this provides a guide to choosing it. If $\alpha = 0$, then short positions in the underlying are prohibited.

The *upper hedging price* of the contingent claim $g(S(\cdot))$ is defined to be

$$v(0,S(0);\alpha) \triangleq \inf\left\{ X(0) \;\middle|\; \begin{array}{l} \text{there exists a portfolio process } \pi \\ \text{satisfying (2.2) and there exists} \\ \text{a cumulative consumption process } C \\ \text{such that } X(T) \geq g(S(\cdot)) \text{ almost surely} \end{array} \right\}. \tag{2.3}$$

Cvitanić & Karatzas [5] have shown that when $v(0,S(0);\alpha)$ is finite, there exists an $X(0)$, denoted $\widehat{X}(0)$, and corresponding portfolio and consumption processes $\hat{\pi}$ and $\widehat{C}$ attaining the infimum in (2.3). We denote the corresponding wealth process by $\widehat{X}(t)$, $0 \leq t \leq T$. For $0 \leq t < T$, we define the *upper hedging price at time t* of the contingent claim $g(S(\cdot))$ to be $\widehat{X}(t)$. The upper hedging price $\widehat{X}(t)$ generally exceeds the risk-neutral price $\mathbb{E}[e^{-r(T-t)}g(S(\cdot))\,|\,\mathcal{F}(t)]$ because the upper hedging price includes a "reserve" to offset the portfolio constraint. During the evolution of the process, some part of this reserve might be revealed to be unnecessary. The process $\widehat{C}$ is included in the formulation of the upper hedging price so that unnecessary reserve can be removed and thus no longer included in the upper hedging price.

Cvitanić & Karatzas [5] and El Karoui & Quenez [8] have shown that the problem of computing the upper hedging price, which is a minimization problem, can be transformed to a dual maximization problem. These results apply to path-dependent contingent claims written on multiple assets whose models may have random, time-varying volatilities, and they require only that π be constrained to lie in a closed, convex set. The dual problem is one of maximization over changes of probability measure, and in its full generality is not easy to solve. In our model, the dual problem takes the form of the right-hand side of (2.4) below.

Theorem 2.1 (Cvitanić & Karatzas, El Karoui & Quenez) *The upper hedging price of (2.3) satisfies*

$$v(0,S(0);\alpha) = \sup_{\lambda} \mathbb{E}_{\lambda}\left[e^{-rT-\alpha\lambda(T)}g(S(\cdot))\right], \tag{2.4}$$

where the supremum is over all adapted, nondecreasing, processes which are Lipschitz continuous in t, uniformly in ω, and satisfy $\lambda(0) = 0$. Here $\mathbb{E}_{\lambda}$ denotes expectation under the probability measures $\mathbb{P}_{\lambda}$ whose Radon-Nikodým derivative with respect to $\mathbb{P}$ is

$$\frac{d\mathbb{P}_{\lambda}}{d\mathbb{P}} = \exp\left\{-\frac{1}{\sigma}\int_0^T \lambda'(t)\,dW(t) - \frac{1}{2\sigma^2}\int_0^T (\lambda'(t))^2 dt\right\}. \tag{2.5}$$

The supremum in (2.4) over Lipschitz continuous processes is often not attained, and Lipschitz continuity is not easily relaxed in Theorem 2.1 because of the need to define $\mathbb{P}_{\lambda}$ by (2.5). In this paper we shall formulate the dual problem in such a way that no change of measure is required, and we can then extend the class of processes over which the supremum in the dual problem is computed.

Broadie, Cvitanić & Soner [4] showed that in the case of a contingent claim whose payoff at expiration is a function of the final value of a single, geometric

Brownian motion, the dual problem can be solved in two steps. One first computes a certain transform, which we call the *face-lift*, of the payoff function (see (2.6) below). One next prices the contingent claim whose payoff at the final time is the face-lifted version of the original payoff. One does this using the usual risk-neutral pricing formula, i.e., without regard to the portfolio constraint. For the model of this section, the result of [4] takes the form of the next theorem. A presentation of the results of both [5] and [4] in full generality may be found in [13].

Theorem 2.2 (Broadie, Cvitanić & Soner) *Let $\varphi: [0,\infty) \to [0,\infty)$ be lower semicontinuous, and suppose the contingent claim $g(S(\cdot))$ is given by $g(S(\cdot)) = \varphi(S(T))$. Define*

$$\widehat{\varphi}_\alpha(x) \triangleq \sup_{\lambda \geq 0} e^{-\alpha\lambda}\varphi(xe^{-\lambda}), \quad x \geq 0. \tag{2.6}$$

Then the upper hedging price under hedge-portfolio constraint (2.2) is given by

$$v(0, S(0); \alpha) = \mathbb{E}\big[e^{-rT}\widehat{\varphi}_\alpha(S(T))\big]. \tag{2.7}$$

The goal of this paper is to extend Theorem 2.2 to the case of path-dependent options. The main result is that in place of the face-lifting procedure (2.6), one must solve a singular stochastic control problem. This problem can sometimes be solved by inspection, and in particular, such a solution is possible for the knock-out call of the previous section. The solution of the stochastic control problem leads directly to a formula for the upper hedging price, in the spirit of (2.7).

The present paper is more general than [4] in that it allows path-dependent options, but more special in that the only portfolio constraint considered here is (2.2), whereas [4] permits a general convex constraint on π. There appears to be no insurmountable obstacle to working out a theory along the lines of the present paper for the more general constraint.

The role of upper hedging prices in the presence of stochastic volatility and/or transaction costs is studied in [2], [6], [7], [15]. Gamma constraints are treated in [14]. Lower hedging prices are introduced in [10], and [11] treats perpetual American options using similar methodology. Classical Black-Scholes prices for a large number of exotic options are provided by Zhang [17].

3 The dual problem as singular stochastic control

We wish to convert the computation of the supremum on the right-hand side of (2.4) to a singular stochastic control problem. Toward this end, we let $(W(t), \mathcal{F}(t); 0 \leq t \leq T)$ be the canonical Brownian motion defined on the canonical probability space $(\Omega, \mathcal{F}(T), \mathbb{P})$ of the previous section, and we denote

$$\begin{aligned} \mathcal{C} \triangleq \big\{\lambda; \lambda \text{ is an } \{\mathcal{F}(t); 0 \leq t \leq T\}\text{-adapted,} \\ \text{nondecreasing, continuous process with } \lambda(0) = 0\big\}. \end{aligned} \tag{3.1}$$

One result of this paper is the following.

Theorem 3.1 *Let g be a nonnegative, lower-semicontinuous function defined on $C_+[0,T]$. The upper hedging price for the contingent claim with payoff $g(S)$ at expiration date T and hedge-portfolio constraint (2.2) is*

$$v(0,S(0);\alpha) = \sup_{\lambda\in C} \mathbb{E}\big[e^{-rT-\alpha\lambda(T)}g(Se^{-\lambda})\big], \tag{3.2}$$

where

$$S(t) = S(0)\exp(\sigma W(t) + \mu(t)), \quad 0 \le t \le T, \tag{3.3}$$

with $\mu(t) \triangleq (r - \frac{1}{2}\sigma^2)t$ is the solution of (1.1).

The problem of maximizing $\mathbb{E}[e^{-rT-\alpha\lambda(T)}g(Se^{-\lambda})]$ over all $\lambda \in C$ is one of stochastic control. In the examples we shall see that there is often a sequence of processes $\{\lambda_n\}_{n=1}^{\infty}$ in C for which

$$\lim_{n\to\infty} \mathbb{E}\big[e^{-rT-\alpha\lambda_n(T)}g(Se^{-\lambda_n})\big] = v(0,S(0);\alpha)$$

and the limit λ of the sequence $\{\lambda_n\}_{n=1}^{\infty}$ is a singularly continuous process; hence the characterization of the right-hand side of (3.2) as a singular stochastic control problem. However, the limiting λ can fail to obtain the supremum in (3.2) because g is lower semicontinuous rather than upper semicontinuous; lower semicontinuity is needed for the proof of Theorem 3.1 (see Lemma 7.2).

The difference between Theorems 2.1 and 3.1 is that whereas the former requires a maximization over changes of measure, the latter allows one to maximize over processes $\lambda \in C$, always computing expectations using the same operator $\mathbb{E}$. Of course, one can use the Radon-Nikodým derivative $d\mathbb{P}_\lambda/d\mathbb{P}$ to rewrite the right-hand side of (2.4) as an expectation under the expectation operator corresponding to $\mathbb{P}$, but the presence of the Radon-Nikodým derivative in the resulting stochastic control problem complicates it considerably. As we shall see in the examples, the stochastic control problem of (3.2) can often be solved by inspection. The proof of Theorem 3.1 is given in Section 7.

4 Constrained in-the-money knock-out call

For the in-the-money knock-out call of Section 1, the function g is

$$g(y) \triangleq \big(y(T)-K\big)^+ I_{\{\max_{0\le t\le T} y(t)<B\}}, \quad y \in C_+[0,T]. \tag{4.1}$$

We have chosen to write the set $\{\max_{0\le t\le T} y(t) < B\}$ with the strict inequality so that g will be lower semicontinuous. For geometric Brownian motion (3.3), the probability of reaching a barrier is the same as the probability of crossing the same barrier, so the contingent claim defined by

$$g^*(y) \triangleq \big(y(T)-K\big)^+ I_{\{\max_{0\le t\le T} y(t)\le B\}}, \quad y \in C_+[0,T], \tag{4.2}$$

has the same upper hedging price.

We consider the problem of maximization of

$$\mathbb{E}\big[e^{-rT-\alpha\lambda(T)}\big(S_\lambda(T)-K\big)^+ I_{\{M_\lambda(T)<B\}}\big], \tag{4.3}$$

where

$$S_\lambda(t) \triangleq S(t)e^{-\lambda(t)}, \quad M_\lambda(t) \triangleq \max_{0\le u\le t} S_\lambda(u),$$

and $0 < S(0) < B$. The maximization is over processes $\lambda \in \mathcal{C}$. To find the maximal value of (4.3) it is clear that one should choose the nondecreasing process λ so that $M_\lambda(T)$ is strictly less than B. On the other hand, one should not have λ be any larger than necessary because λ appears in both the discount term $e^{-rT-\alpha\lambda(T)}$ and as a discount in the formula for S_λ. If g were given by (4.2), the maximizing λ would be that nondecreasing process which causes reflection of S_λ at the barrier B, i.e.,

$$\lambda^*(t) \triangleq \max_{0\le u\le t}\big(\log S(u) - \log B\big)^+. \tag{4.4}$$

Since g is dominated by g^*, we have

$$v(0,S(0);\alpha) \le \mathbb{E}\big[e^{-rT-\alpha\lambda^*(T)}\big(S_{\lambda^*}(T)-K\big)^+\big]. \tag{4.5}$$

But with g given by (4.1), it is still possible to choose a sequence of barriers $\{B_n\}_{n=1}^\infty$ converging up to B but always strictly less than B and then take the sequence of processes $\{\lambda_n\}_{n=1}^\infty$ for which λ_n causes reflection at B_n. Then $\lambda_n(T) \downarrow \lambda^*(T)$ and therefore $S_{\lambda_n}(T) \uparrow S_{\lambda^*}(T)$ as $n \to \infty$. By the bounded convergence theorem,

$$\begin{aligned} v(0,S(0);\alpha) &\ge \limsup_{n\to\infty} \mathbb{E}\big[e^{-rT-\alpha\lambda_n(T)}\big(S_{\lambda_n}(T)-K\big)^+\big] \\ &= \mathbb{E}\big[e^{-rT-\alpha\lambda^*(T)}\big(S_{\lambda^*}(T)-K\big)^+\big]. \end{aligned} \tag{4.6}$$

These considerations have led us to the following corollary of Theorem 3.1.

Corollary 4.1 *For $0 \le t \le T$ and $0 < x \le B$, define*

$$v^*(t,x;\alpha) \triangleq \mathbb{E}\big[e^{-r(T-t)-\alpha(\lambda^*(T)-\lambda^*(t))}\big(S_{\lambda^*}(T)-K\big)^+ \mid S_{\lambda^*}(t) = x\big]. \tag{4.7}$$

Let $t \in [0,T]$ be given, and assume that $S(t) = x$. Then the upper hedging price at time t of the in-the-money knock-out call of Section 1 is

$$v(t,x;\alpha) = v^*(t,x;\alpha) I_{\{x<B\}}, \tag{4.8}$$

and for $t \in [0,T)$ the function $v^*(t,x;\alpha)$ can be computed (with removable sin-

gularities at $\alpha = 2r/\sigma^2$ and $\alpha = -1 + 2r/\sigma^2$ in the case $2r \geq \sigma^2$) to be

$$\begin{aligned}
&x\Big[N(b-\theta_+) - N(k-\theta_+) \\
&\quad + e^{\frac{1}{2}s(s-2\theta_+)}\left\{e^{sb}N(-b+\theta_+-s) - e^{sk}N(-k+\theta_+-s)\right\}\Big] \\
&+ \frac{sxe^{2b\theta_+}}{s-2\theta_+}\Big[N(b+\theta_+) - N(\ell+\theta_+) \\
&\quad + e^{\frac{1}{2}s(s-2\theta_+)}\left\{e^{(s-2\theta_+)b}N(-b+\theta_+-s) - e^{(s-2\theta_+)\ell}N(-\ell+\theta_+-s)\right\}\Big] \\
&- Ke^{-r\tau}\Big[N(b-\theta_-) - N(k-\theta_-) \\
&\quad + e^{\frac{1}{2}\bar{s}(\bar{s}-2\theta_-)}\left\{e^{\bar{s}b}N(-b+\theta_--\bar{s}) - e^{\bar{s}k}N(-k+\theta_--\bar{s})\right\}\Big] \\
&- \frac{\bar{s}Ke^{-r\tau+2b\theta_-}}{\bar{s}-2\theta_-}\Big[N(b+\theta_-) - N(\ell+\theta_-) \\
&\quad + e^{\frac{1}{2}\bar{s}(\bar{s}-2\theta_-)}\left\{e^{(\bar{s}-2\theta_-)b}N(-b+\theta_--\bar{s}) - e^{(\bar{s}-2\theta_-)\ell}N(-\ell+\theta_--\bar{s})\right\}\Big],
\end{aligned} \tag{4.9}$$

where we have used the abbreviations $\tau = T - t$ and

$$\begin{aligned}
b &= \frac{1}{\sigma\sqrt{\tau}}\log\frac{B}{x}, & k &= \frac{1}{\sigma\sqrt{\tau}}\log\frac{K}{x}, & \theta_\pm &= \left(\frac{r}{\sigma} \pm \frac{\sigma}{2}\right)\sqrt{\tau}, \\
\ell &= 2b - k, & s &= (1+\alpha)\sigma\sqrt{\tau}, & \bar{s} &= \alpha\sigma\sqrt{\tau}.
\end{aligned}$$

PROOF: Theorem 3.1 and the argument preceding the statement of the corollary show that $v^*(0, S(0); \alpha)$ is the upper hedging price of the knock-out call for $0 < S(0) < B$. For $S(0) = B$, the call is knocked out at inception, and hence has upper hedging price zero. This establishes (4.8) when $t = 0$. For other values of t, one can verify the formula by a translation of the initial condition. Equation (4.9) is obtained by direct calculation using (1.4). $\diamond$

It is instructive to construct the short-position hedge. Formula (4.9) with $v^*(T, x; \alpha) = (x - K)^+$ shows that $v^*(t, x; \alpha)$ is continuous on $[0, T] \times (0, B]$ and smooth on $[0, T) \times (0, B]$.

Let S be the underlying geometric Brownian motion given by (3.3). Then S_{λ^*} is a Markov process and

$$dS_{\lambda^*}(t) = S_{\lambda^*}(t)[r\,dt + \sigma\,dW(t) - d\lambda^*(t)]. \tag{4.10}$$

Moreover,

$$e^{-rt-\alpha\lambda^*(t)}v^*(t, S_{\lambda^*}(t); \alpha) = \mathbb{E}\left[e^{-rT-\alpha\lambda^*(T)}\left(S_{\lambda^*}(T) - K\right)^+ \mid \mathcal{F}(t)\right]. \tag{4.11}$$

We compute the differential using Itô's formula

$$\begin{aligned}
&d\left(e^{-rt-\alpha\lambda^*(t)}v^*(t, S_{\lambda^*}(t); \alpha)\right) \\
&\quad = e^{-rt-\alpha\lambda^*(t)}\Big[-(\alpha v^* + S_{\lambda^*}v_x^*)\,d\lambda^* \\
&\qquad\qquad + \left(-rv^* + v_t^* + rS_{\lambda^*}v_x^* + \frac{1}{2}\sigma^2 S_{\lambda^*}^2 v_{xx}^*\right)dt + \sigma S_{\lambda^*}v_x^*\,dW\Big].
\end{aligned}$$

But the right-hand side of (4.11) is a martingale, which implies

$$\left[\alpha v^*(t, S_{\lambda^*}(t);\alpha) + S_{\lambda^*}(t) v^*_x(t, S_{\lambda^*}(t);\alpha)\right] d\lambda^*(t) = 0,$$

$$\Big[-rv^*(t, S_{\lambda^*}(t);\alpha) + v^*_t(t, S_{\lambda^*}(t);\alpha) + rS_{\lambda^*}(t) v^*_x(t, S_{\lambda^*}(t);\alpha) + \frac{1}{2}\sigma^2 S^2_{\lambda^*}(t) v^*_{xx}(t, S_{\lambda^*}(t);\alpha)\Big] dt = 0,$$

i.e., for $0 \le t < T$ and $0 < x \le B$,

$$\alpha v^*(t, B;\alpha) + Bv^*_x(t, B;\alpha) = 0, \tag{4.12}$$

$$v^*_t(t, x;\alpha) + rxv^*_x(t, x;\alpha) + \frac{1}{2}\sigma^2 x^2 v^*_{xx}(t, x;\alpha) = rv^*(t, x;\alpha). \tag{4.13}$$

One can also obtain (4.12) and (4.13) by direct, albeit tedious, computation beginning with (4.9).

It can be verified by direct computation that because v^* satisfies the Black-Scholes partial differential equation (4.13) in the region $[0,T) \times (0,B]$, the function xv^*_x does also; just differentiate (4.13) with respect to x and identify the resulting terms as the partial derivatives of xv^*_x. It follows that $\alpha v^* + xv^*_x$ satisfies the Black-Scholes partial differential equation. But $\alpha v^*(t,x;\alpha) + xv^*_x(t,x;\alpha)$ is nonnegative for $t = T$ and $0 < x < B$, $x \ne K$, and this function is zero on the upper barrier $x = B$ for $0 \le t < T$ (see (4.12)). It follows from the maximum principle (or by regarding $\alpha v^* + xv^*_x$ as the price of a knock-out option with nonnegative payoff upon expiration) that

$$\alpha v^*(t,x;\alpha) + xv^*_x(t,x;\alpha) \ge 0, \quad 0 \le t \le T,\ 0 < x \le B. \tag{4.14}$$

Now suppose $0 < S(0) < B$, and define

$$\Theta \triangleq \inf\{t \ge 0;\ S(t) = B\}$$

to be the time of knock-out; we allow the possibility that $\Theta > T$, i.e., knock-out does not occur before expiration. Let us begin with initial capital $\widehat{X}(0) = v^*(0, S(0);\alpha)$ and use the portfolio process

$$\hat{\pi}(t) \triangleq \frac{S(t)v^*_x(t,S(t);\alpha)}{v^*(t,S(t);\alpha)} I_{\{t \le T\wedge\Theta\}}, \quad 0 \le t \le T, \tag{4.15}$$

(see Figure 2) and consumption process

$$\widehat{C}(t) \triangleq v^*(\Theta, B;\alpha) I_{\{\Theta \le t \le T\}}, \quad 0 \le t \le T. \tag{4.16}$$

Relation (4.14) shows that $\hat{\pi}(t) \ge -\alpha$ for all $t \in [0,T]$. The cumulative consumption process $\widehat{C}$ is identically zero until the option knocks out, at which time it has a positive jump; see Figure 1 for the jump size.

Equation (4.13) shows that for $0 \le t < \Theta \wedge T$,

$$dv^*(t, S(t);\alpha) = rv^*(t,S(t);\alpha)\,dt + \sigma\hat{\pi}(t)v^*(t,S(t);\alpha)\,dW(t).$$

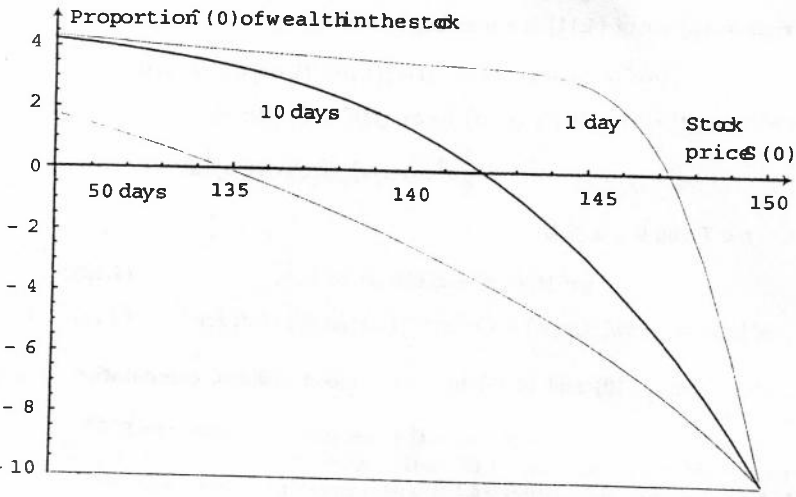

Figure 2: Proportions $\hat{\pi}(0)$ of the wealth in the underlying stock, calculated with (4.15) and (4.9), to super-replicate the in-the-money knock-out call options with parameters as in Figure 1. Note that the constraint $\pi_0 \geq -\alpha$ with $\alpha = 10$ is satisfied. The three dashed curves show the corresponding proportions $S_0 v_x(0, S(0))/v(0, S_0)$ with $v(0, S(0))$ given by (1.5) without the hedge-portfolio constraint (2.2). Note that these proportions are not bounded from below.

Comparison to (2.1) shows that for $t < \Theta \wedge T$, $v^*(t, S(t); \alpha) = \widehat{X}(t)$, the wealth process corresponding to $\widehat{X}(0)$, $\hat{\pi}$ and $\widehat{C}$. If $\Theta \leq T$, then $\lim_{t \uparrow \Theta} \widehat{X}(t) = v^*(\Theta, B; \alpha)$ and $\widehat{X}(\Theta) = \lim_{t \uparrow \Theta} \widehat{X}(t) - C(\Theta) = 0$. For $\Theta < t \leq T$, we also have $\widehat{X}(t) = 0$. In general,

$$\widehat{X}(t) = v(t \wedge \Theta, S(t \wedge \Theta); \alpha), \quad 0 \leq t \leq T, \tag{4.17}$$

where v is defined by (4.8). In particular, $\widehat{X}(T) = v(T \wedge \Theta, S(T \wedge \Theta); \alpha) = (S(T) - K)^+ I_{\{\Theta > T\}}$, i.e., we have hedged a short position in the option in a manner which respects the portfolio constraint $\hat{\pi} \geq -\alpha$.

Remark 4.2 If the knock-out call payoff were given by g^* of (4.2) rather than g of (4.1), the maximum of the quantity analogous to (4.3) would be attained by λ^* of (4.4). For $0 < S(0) < B$, this maximum would be the upper hedging price and making this replacement would simplify the discussion preceding Corollary 4.1. However, for $S(0) = B$, this would not give the upper hedging price. If $S(0) = B$ the option is certain to knock out and the upper hedging price is zero, as is the maximum of the quantity in (4.3). However, if we replace g by g^*, the maximum of the quantity analogous to (4.3) is strictly positive, and in fact is $v^*(0, B; \alpha)$.

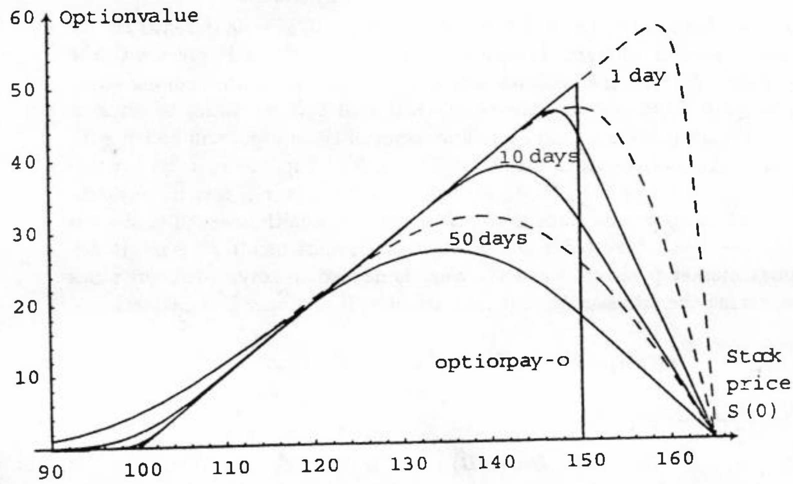

Figure 3: Upper hedging prices $v^*(0, S(0), \alpha)$ of the in-the-money knock-out call options from Figure 1. Prices are extrapolated linearly and continuously differentiably beyond the barrier $B = 150$ using (4.18). The dashed curves show the prices calculated via (1.5) without hedge-portfolio constraint (2.2) but a barrier moved to $B' = B(1+1/\alpha) = 165$. For applications, only the prices for $S(0) < B = 150$ are relevant.

Remark 4.3 (Interpretation as moving the barrier) A common practical method for dealing with up-and-out call options which knock out in the money is to price and hedge the option as if the barrier were at some level B' strictly greater than the contractual barrier B. The resulting pricing function is continuous on $[0,T] \times (0,B']$, satisfies the Black-Scholes partial differential equation on $[0,T) \times (0,B']$, is zero at the barrier B', and agrees with the call payoff $(x-K)^+$ at the expiration time T. Our function $v^*(t,x;\alpha)$ is strictly positive at $x = B$. For $\alpha > 0$ we may extrapolate it linearly above this point so that it is continuously differentiable by the formula

$$v^*(t,B;\alpha) + (x-B)v_x^*(t,B;\alpha), \quad x \geq B. \tag{4.18}$$

Because of (4.12), this linear extrapolation takes the value zero at $x = \left(1+\frac{1}{\alpha}\right)B$, independently of t. Consequently, $v^*(t,x;\alpha)$ may be regarded as an approximation to the option price obtained by moving the barrier to $B' = \left(1+\frac{1}{\alpha}\right)B$.

Remark 4.4 (Interpretation as transaction cost) The Black-Scholes formula is based on the assumption that the bid-ask spread does not play a significant role in option hedging. A trader who hedges a short position in the knock-out option of this section can be left with a large short position in the underlying stock when the option knocks out, and covering this position can entail a significant transaction cost. Let us suppose the trader prices the option according to a function $v(t,x)$ which is continuous in $[0,T]\times(0,B]$, satisfies the Black-Scholes partial differential equation in $[0,T)\times(0,B]$, and agrees with the call payoff $(x-K)^+$ at the expiration time T. Using the "delta-hedging strategy" to hedge a short position, the trader will hold $v_x(t,x)$ shares of stock at time t if the stock price is x, and upon knock-out of the option, will be left with a position $v_x(t,B)$ in the stock valued at $|Bv_x(t,B)|$. Suppose $v_x(t,B)$ is negative and it requires $-\left(1+\frac{1}{\alpha}\right)Bv_x(t,B)$ with $\alpha>0$ to cover this short position. The total hedging portfolio value is $v(t,B)$, and since wealth invested in stock is $Bv_x(t,B)$, the wealth invested in the money market must be $v(t,B)-Bv_x(t,B)$. The money market position is exactly what is needed to cover the short stock position, taking the transaction cost into account, if and only if the equation

$$v(t,B)-Bv_x(t,B)=-\Big(1+\frac{1}{\alpha}\Big)Bv_x(t,B)$$

holds. This is equivalent to

$$\alpha v(t,B)+Bv_x(t,B)=0,\quad 0\le t<T,$$

which is condition (4.12) satisfied by $v^*(t,x;\alpha)$. Together with the conditions already specified on v, this uniquely determines v, and in fact ensures that $v(t,x)=v^*(t,x;\alpha)$.

5 The dual problem as impulsive control

Let $0<t_1<t_2<\cdots<t_I\le T$ be a fixed set of dates. For the examples of Section 6, it is helpful to generalize Theorem 3.1 to allow λ to be right-continuous with possible jumps at these dates, and to be continuous at all other times. We denote by $R[0,T]$ the set of nondecreasing functions λ defined on $[0,T]$, continuous on $[0,T]\setminus\{t_1,\dots,t_I\}$, right-continuous at $t_1,\dots,t_I$, and with $\lambda(0)=0$. We then define

$$\mathcal{R}\triangleq\{\lambda;\lambda \text{ is an } \{\mathcal{F}(t);0\le t\le T\}\text{-adapted processes with paths in } R[0,T]\}. \tag{5.1}$$

A function in $R[0,T]$ can be regarded as the cumulative distribution function of a measure on $[0,T]$. The measures corresponding to a sequence $\{\lambda_n\}_{n=1}^\infty$ in $R[0,T]$ converge weakly to the measure with cumulative distribution function $\lambda\in R[0,T]$ if and only if $\lambda_n(t)\to\lambda(t)$ at every continuity point t of λ and for $t=T$. Because the weak topology on measures can be metrized (see [16]), there is a metric d_w on $R[0,T]$ satisfying

$$\lambda_n(t)\to\lambda(t) \text{ for every continuity point } t \text{ of } \lambda \text{ and } t=T \iff d_w(\lambda_n,\lambda)\to 0.$$

We may define a stronger metric d on $R[0,T]$ by

$$d(\eta,\lambda)=d_w(\eta,\lambda)+\sum_{i=1}^{I}|\eta(t_i)-\lambda(t_i)|.$$

Then for every sequence $\{\lambda_n\}_{n=1}^{\infty}$ in $R[0,T]$ and $\lambda \in R[0,T]$,

$$\lambda_n(t) \to \lambda(t) \; \forall t \in [0,T] \iff d(\lambda_n, \lambda) \to 0, \tag{5.2}$$

i.e., d metrizes pointwise convergence in $R[0,T]$.

Remark 5.1 If λ in (5.2) is continuous, pointwise convergence of λ_n to λ implies uniform convergence. Given $\varepsilon > 0$, choose $\delta > 0$ such that $|t-s| \le \delta$ implies $|\lambda(t) - \lambda(s)| \le \varepsilon$. Choose $0 = s_0 < s_1 < \cdots < s_K = T$ such that $s_{k+1} - s_k \le \delta$ for all k. If $\lambda_n \to \lambda$ pointwise, we may choose N so that $|\lambda_n(s_k) - \lambda(s_k)| \le \varepsilon$ for every $n \ge N$ and every k. Given $t \in [0,T]$, we choose k so that $s_k \le t \le s_{k+1}$, and then for all $n \ge N$,

$$\begin{aligned}
|\lambda_n(t) - \lambda(t)| &\le \lambda_n(t) - \lambda_n(s_k) + |\lambda_n(s_k) - \lambda(s_k)| + |\lambda(s_k) - \lambda(t)| \\
&\le \lambda_n(s_{k+1}) - \lambda_n(s_k) + 2\varepsilon \\
&\le |\lambda_n(s_{k+1}) - \lambda(s_{k+1})| + |\lambda(s_{k+1}) - \lambda(s_k)| \\
&\quad + |\lambda(s_k) - \lambda_n(s_k)| + 2\varepsilon \\
&\le 5\varepsilon.
\end{aligned}$$

More generally, suppose $\lambda \in R[0,T]$ is not necessarily continuous. Let $\{\lambda_n\}_{n=1}^{\infty}$ be a sequence in $R[0,T]$ converging pointwise to λ, and let $\delta > 0$ be given. Then $\{\lambda_n\}_{n=1}^{\infty}$ converges uniformly to λ on $[0,T] \setminus \bigcup_{i=1}^{I}(t_i - \delta, t_i)$. Indeed, the argument in the previous paragraph shows uniform convergence on each connected component of this set, and since there are only finitely many such components, the convergence is uniform on the whole set. ◇

The function g is defined on $C_+[0,T]$, the space of nonnegative continuous functions on $[0,T]$. We extend it to $C_+[0,T] \times R[0,T]$ by the definition

$$g_*(y,\lambda) \triangleq \inf\left\{ \liminf_{n\to\infty} g\left(ye^{-\lambda_n}\right) \middle| \begin{array}{l} \{\lambda_n\}_{n=1}^{\infty} \text{ is a sequence in } R[0,T] \\ \cap\, C_+[0,T] \text{ converging pointwise to } \lambda \end{array} \right\};$$

$$y \in C_+[0,T],\ \lambda \in R[0,T]. \tag{5.3}$$

Proposition 5.2 *Suppose* $g\colon C_+[0,T] \to [0,\infty)$ *is of the form*

$$g(y) = \varphi(y(t_1), \ldots, y(t_I), m(y), M(y), A(y)),$$

where

$$m(y) \triangleq \inf_{0\le t\le T} y(t), \quad M(y) \triangleq \sup_{0\le t\le T} y(t), \quad A(y) \triangleq \frac{1}{T}\int_0^T y(t)\,dt,$$

and $\varphi\colon \mathbb{R}^{I+3} \to [0,\infty)$ *is a lower-semicontinuous function which is jointly left-continuous in its last three arguments. Then for* $y \in C_+[0,T]$ *and* $\lambda \in R[0,T]$,

$$g_*(y,\lambda) = \varphi\Big(y(t_1)e^{-\lambda(t_1)}, \ldots, y(t_I)e^{-\lambda(t_I)}, m(ye^{-\lambda}), M(ye^{-\lambda}), A(ye^{-\lambda})\Big).$$

PROOF: We claim that for fixed $y \in C_+[0,T]$, the mappings $\lambda \mapsto m(ye^{-\lambda})$, $\lambda \mapsto M(ye^{-\lambda})$ and $\lambda \mapsto A(ye^{-\lambda})$ are continuous from $R[0,T]$ to $[0,\infty)$. Indeed, fix $y \in C_+[0,T]$ and suppose $\{\lambda_n\}_{n=1}^{\infty}$ converges pointwise to λ. Let $\varepsilon > 0$ be given and choose $\delta > 0$ so that $|y(s) - y(t)| \le \varepsilon$ whenever $|s-t| \le \delta$. For the sake of notational simplicity we define $t_0 \triangleq 0$ and assume that $t_I = T$. We may assume without loss of generality that $\delta < \min_{1\le i\le I}(t_i - t_{i-1})$. Because of Remark 5.1, we may choose $N(\varepsilon)$ so large that whenever $n \ge N(\varepsilon)$, we have

$$|\lambda_n(t) - \lambda(t)| \le \varepsilon \quad \forall t \in \bigcup_{i=1}^{I}[t_{i-1}, t_i - \delta].$$

For these $t \in \bigcup_{i=1}^{I}[t_{i-1}, t_i - \delta]$, we have

$$\begin{aligned} \left|y(t)e^{-\lambda_n(t)} - y(t)e^{-\lambda(t)}\right| &\le M(y)e^{-\lambda(t)}\left|e^{\lambda(t)-\lambda_n(t)} - 1\right| \\ &\le M(y)(e^{\varepsilon} - 1). \end{aligned} \tag{5.4}$$

On the other hand, for $i \in \{1,\dots,I\}$ and $t \in [t_i - \delta, t_i]$, we have

$$\begin{aligned} y(t)e^{-\lambda_n(t)} &\le y(t)e^{-\lambda_n(t_i-\delta)} \\ &\le |y(t) - y(t_i-\delta)|\, e^{-\lambda_n(t_i-\delta)} + y(t_i-\delta)e^{-\lambda_n(t_i-\delta)} \\ &\le \varepsilon + y(t_i-\delta)e^{-\lambda(t_i-\delta)} + y(t_i-\delta)\left|e^{-\lambda_n(t_i-\delta)} - e^{-\lambda(t_i-\delta)}\right| \\ &\le \varepsilon + M\big(ye^{-\lambda}\big) + M(y)(e^{\varepsilon}-1). \end{aligned}$$

Combining this inequality with (5.4), we see that

$$M\big(ye^{-\lambda_n}\big) \le \varepsilon + M\big(ye^{-\lambda}\big) + M(y)(e^{\varepsilon}-1). \tag{5.5}$$

Similarly,

$$\begin{aligned} y(t)e^{-\lambda(t)} &\le y(t)e^{-\lambda(t_i-\delta)} \\ &\le |y(t) - y(t_i-\delta)|\, e^{-\lambda(t_i-\delta)} + y(t_i-\delta)e^{-\lambda(t_i-\delta)} \\ &\le \varepsilon + y(t_i-\delta)e^{-\lambda_n(t_i-\delta)} + y(t_i-\delta)\left|e^{-\lambda(t_i-\delta)} - e^{-\lambda_n(t_i-\delta)}\right| \\ &\le \varepsilon + M\big(ye^{-\lambda_n}\big) + M(y)(e^{\varepsilon}-1). \end{aligned}$$

We may combine this with (5.4) to obtain

$$M\big(ye^{-\lambda}\big) \le \varepsilon + M\big(ye^{-\lambda_n}\big) + M(y)(e^{\varepsilon}-1). \tag{5.6}$$

We conclude that for $n \ge N(\varepsilon)$, the inequality

$$\left|M\big(ye^{-\lambda_n}\big) - M\big(ye^{-\lambda}\big)\right| \le \varepsilon + M(y)(e^{\varepsilon}-1)$$

holds, and hence the mapping $\lambda \mapsto M\big(ye^{-\lambda}\big)$ is continuous. Similar arguments show the continuity of $\lambda \mapsto m\big(ye^{-\lambda}\big)$. The continuity of $\lambda \mapsto A\big(ye^{-\lambda}\big)$ follows from the dominated convergence theorem.

Now let $y \in C_+[0,T]$ and $\lambda \in R[0,T]$ be given, and let $\{\lambda_n\}_{n=1}^{\infty}$ be a sequence in $R[0,T] \cap C_+[0,T]$ converging pointwise to λ. Because $\lim_{n\to\infty} m\big(ye^{\lambda_n}\big) =$

$m(ye^{-\lambda})$, $\lim_{n\to\infty} M(ye^{\lambda_n}) = M(ye^{-\lambda})$, $\lim_{n\to\infty} A(ye^{\lambda_n}) = A(ye^{-\lambda})$ and φ is lower semicontinuous, we have

$$\begin{aligned}
&\varphi\Big(y(t_1)e^{-\lambda(t_1)},\ldots,y(t_I)e^{-\lambda(t_I)},m(ye^{-\lambda}),M(ye^{-\lambda}),A(ye^{-\lambda})\Big)\\
&\quad\le \liminf_{n\to\infty}\varphi\Big(y(t_1)e^{-\lambda_n(t_1)},\ldots,y(t_I)e^{-\lambda_n(t_I)},m(ye^{-\lambda_n}),M(ye^{-\lambda_n}),A(ye^{-\lambda_n})\Big)\\
&\quad= \liminf_{n\to\infty} g(ye^{-\lambda_n}).
\end{aligned}$$

Taking the infimum of the right-hand side over sequences $\{\lambda_n\}_{n=1}^{\infty}$ converging pointwise to λ, we conclude that

$$\varphi\Big(y(t_1)e^{-\lambda(t_1)},\ldots,y(t_I)e^{-\lambda(t_I)},m(ye^{-\lambda}),M(ye^{-\lambda}),A(ye^{-\lambda})\Big)\le g_*(y,\lambda).$$

To obtain the reverse inequality, we choose $\lambda_n \in R[0,T]\cap C_+[0,T]$ so that $\lambda_n(t_i)=\lambda(t_i)$ for $i=1,\ldots,I$ and $\lambda_n(t)\downarrow\lambda(t)$ for every $t\in[0,T]$ as $n\to\infty$. Then $ye^{-\lambda_n}\uparrow ye^{-\lambda}$ pointwise. The joint left-continuity of φ in its last three variables implies

$$\begin{aligned}
&\varphi\Big(y(t_1)e^{-\lambda(t_1)},\ldots,y(t_I)e^{-\lambda(t_I)},m(ye^{-\lambda}),M(ye^{-\lambda}),A(ye^{-\lambda})\Big)\\
&\quad= \lim_{n\to\infty}\varphi\Big(y(t_1)e^{-\lambda_n(t_1)},\ldots,y(t_I)e^{-\lambda_n(t_I)},m(ye^{-\lambda_n}),M(ye^{-\lambda_n}),A(ye^{-\lambda_n})\Big)\\
&\quad= \lim_{n\to\infty} g(ye^{-\lambda_n})\\
&\quad\ge g_*(y,\lambda).
\end{aligned}$$

◇

Theorem 5.3 *Let g be a nonnegative, lower-semicontinuous function defined on $C_+[0,T]$. The upper hedging price for the contingent claim with payoff $g(S)$ at expiration date T and hedge-portfolio constraint* (2.2) *is*

$$v(0,S(0);\alpha)=\sup_{\lambda\in\mathcal{R}}\mathbb{E}\big[e^{-rT-\alpha\lambda(T)}g_*(S,\lambda)\big],\tag{5.7}$$

where the geometric Brownian motion S is given by (3.3).

The proof of Theorem 5.3 is given in Section 8.

Remark 5.4 Theorem 5.3 leads immediately to an alternate proof of Theorem 2.2 (Broadie, Cvitanić & Soner). Let $\varphi\colon [0,\infty)\to[0,\infty)$ be lower semicontinuous, and define $g\colon C_+[0,T]\to[0,\infty)$ by $g(y)=\varphi(y(T))$. In the definition of $\mathcal{R}$, take $I=1$ and $t_1=T$, i.e., the only allowed discontinuity for processes in $\mathcal{R}$ is at time T. Proposition 5.2 implies that $g_*(y,\lambda)=\varphi\big(y(T)e^{-\lambda(T)}\big)$. According to Theorem 5.3, the upper hedging price is

$$v(0,S(0);\alpha)=\sup_{\lambda\in\mathcal{R}}\mathbb{E}\big[e^{-rT-\alpha\lambda(T)}\varphi\big(S(T)e^{-\lambda(T)}\big)\big],\tag{5.8}$$

which is obviously bounded above by $\mathbb{E}\big[e^{-rT}\widehat{\varphi}_\alpha(S(T))\big]$ (see (2.6) for notation). On the other hand, a selection theorem due to Freedman [9] (see, e.g., [3], Proposition 7.34) asserts that for each $\varepsilon > 0$ there is a Borel measurable function $\psi_\varepsilon : [0,\infty) \to [0,\infty)$ satisfying

$$e^{-\alpha\psi_\varepsilon(x)}\varphi\big(xe^{-\psi_\varepsilon(x)}\big) \geq \begin{cases} \widehat{\varphi}_\alpha(x) - \varepsilon & \text{if } \widehat{\varphi}_\alpha(x) < \infty, \\ 1/\varepsilon & \text{if } \widehat{\varphi}_\alpha(x) = \infty. \end{cases}$$

Taking $\lambda(t) = I_{\{t=T\}}\psi_\varepsilon(S(T))$, we conclude from (5.8) that

$$v(0,S(0);\alpha) \geq -\varepsilon + \mathbb{E}\big[I_{\{\widehat{\varphi}_\alpha(S(T))<\infty\}}e^{-rT}\widehat{\varphi}_\alpha(S(T))\big] + \frac{e^{-rT}}{\varepsilon}\mathbb{P}\{\widehat{\varphi}_\alpha(S(T)) = \infty\}.$$

Letting $\varepsilon \downarrow 0$, we obtain $v(0,S(0);\alpha) \geq \mathbb{E}\big[e^{-rT}\widehat{\varphi}_\alpha(S(T))\big]$. Equation (2.7) is proved.

6 Examples

In this section, we give examples of options whose upper hedging prices can be computed using either Theorem 3.1 or 5.3. In both these theorems, the path-dependent payoff function g is assumed to be lower semicontinuous. Some option contracts are written with upper-semicontinuous payoffs. However, one can usually trivially modify an upper-semicontinuous payoff so as to obtain a lower-semicontinuous payoff, and then the theorems apply. Our first example highlights the danger of applying the theorems naively to upper-semicontinuous payoff functions.

Example 6.1 (Cactus option) Consider an option whose payoff at expiration date T is 1 if and only if $S(T) = K$, where K is a fixed positive number. Otherwise, the payoff is zero. The payoff can be written as $\varphi(S(T))$, where $\varphi(x) \triangleq I_{\{x=K\}}$ is upper semicontinuous rather than lower semicontinuous. If we ignore this fact and attempt to use Theorem 2.2 (which via Remark 5.4 follows from Theorem 5.3) to compute the upper hedging price, we would first determine

$$\widehat{\varphi}_\alpha(x) \triangleq \sup_{\lambda\geq 0} e^{-\alpha\lambda}\varphi(xe^{-\lambda}) = \left(\frac{K}{x}\right)^\alpha I_{\{x\geq K\}}, \quad x \geq 0,$$

and then compute $\mathbb{E}\big[e^{-rT}\widehat{\varphi}_\alpha(S(T))\big]$. This last quantity is strictly positive. However, the option is clearly worth zero, since there is zero probability that $S(T) = K$. To correctly compute the upper hedging price, one should replace the given φ by its lower-semicontinuous envelope $\varphi_* \equiv 0$.

Example 6.2 (Digital put) The payoff for the digital put is

$$g(S(\cdot)) = I_{\{S(T)<K\}}$$

where K is positive. This can be written as $g(S(\cdot)) = \varphi(S(T))$, where $\varphi(x) \triangleq I_{\{x<K\}}$. According to Theorem 2.2, we should first determine the face-lift

$$\widehat{\varphi}_\alpha(x) = \begin{cases} 1 & \text{if } 0 \leq x \leq K, \\ \left(\frac{K}{x}\right)^\alpha & \text{if } x \geq K, \end{cases}$$

and then the upper hedging price can be computed to be

$$v(0,S(0);\alpha) = e^{-rT}\,\mathbf{E}[\widehat{\varphi}_\alpha(S(T))]$$
$$= e^{-rT}N(-d) + e^{(1+\alpha)(\frac{1}{2}\alpha\sigma^2 - r)T}\left(\frac{K}{S(0)}\right)^\alpha N(d - \alpha\sigma\sqrt{T}),$$

where $d \triangleq \frac{1}{\sigma\sqrt{T}}\left[\log\frac{S(0)}{K} + \left(r - \frac{1}{2}\sigma^2\right)T\right]$.

Example 6.3 (Discrete barrier option) The in-the-money knock-out call described in Section 1 was discussed in considerable detail in Section 4. Here we modify the payoff by assuming the option can only knock out at discrete check times $0 < t_1 < t_2 < \cdots < t_I \le T$, i.e.,

$$g(S(\cdot)) = \left(S(T) - K\right)^+ \prod_{i=1}^{I} I_{\{S(t_i)<B\}}.$$

The payoff function g is of the form treated in Proposition 5.2, and thus

$$g_*(S,\lambda) = \left(S(T)e^{-\lambda(T)} - K\right)^+ \prod_{i=1}^{I} I_{\{S(t_i)e^{-\lambda(t_i)}<B\}}.$$

The supremum in (5.7) is approached by processes λ which are constant between the check times, and jump at the check times "just enough" to prevent knock-out. More precisely, let $\{B_n\}_{n=1}^\infty$ be a sequence converging up to B. For each n, define

$$\lambda_n(t) \triangleq \max_{\{i;\,t_i\le t\}}\left(\log S(t_i) - \log B_n\right)^+, \quad 0 \le t \le T.$$

Then $S(t_i)e^{-\lambda_n(t_i)} \le B_n$ for each $i \in \{1,\ldots,I\}$, and λ_n is the smallest process in $\mathcal{R}$ which forces these inequalities. Starting with the maximization problem (5.7) in Theorem 5.3 and applying the arguments which led to (4.5) and (4.6), we obtain for the upper hedging price

$$v(0,S(0);\alpha) = \lim_{n\to\infty}\mathbf{E}\left[e^{-rT-\alpha\lambda_n(T)}\left(S(T)e^{-\lambda_n(T)} - K\right)^+\right]$$
$$= \mathbf{E}\left[e^{-rT-\alpha\lambda^*(T)}\left(S(T)e^{-\lambda^*(T)} - K\right)^+\right],$$

where

$$\lambda^*(t) \triangleq \max_{\{i;\,t_i\le t\}}\left(\log S(t_i) - \log B\right)^+, \quad 0 \le t \le T.$$

This may be rewritten as

$$v(0,S(0);\alpha) = e^{-rT}\,\mathbf{E}\left[\left(1 \wedge \min_{1\le i\le I}\frac{B}{S(t_i)}\right)^\alpha\left(S(T)\left(1 \wedge \min_{1\le i\le I}\frac{B}{S(t_i)}\right) - K\right)^+\right].$$

The computation has been reduced to a finite-dimensional Gaussian integration. If the barrier depends on time, we need only to replace the ratios $B/S(t_i)$ by $B(t_i)/S(t_i)$ in the last formula.

Example 6.4 (Vanilla put) We compute the upper hedging price of the vanilla put as a prelude to Examples 6.5 and 6.6. The payoff of the vanilla put is $g(S(\cdot)) = \varphi(S(T))$, where $\varphi(x) = (K - x)^+$ and K is a positive constant. According to Proposition 5.2 and Theorem 5.3, the upper hedging price is

$$v(0, S(0); \alpha) = \sup_{\lambda \in \mathcal{R}} \mathbb{E}\big[e^{-rT-\alpha\lambda(T)}\big(K - S(T)e^{-\lambda(T)}\big)^+\big], \qquad (6.1)$$

where we take $I = 1$ and $t_1 = T$ in the definition of $\mathcal{R}$, meaning that the processes are continuous except for a possible jump at time T. Theorem 2.2 applies, and asserts that the upper hedging price is

$$v(0, S(0); \alpha) = e^{-rT}\,\mathbb{E}\big[\widehat{\varphi}_\alpha(S(T); K)\big],$$

where the face-lift is given by

$$\widehat{\varphi}_\alpha(x; K) \triangleq \sup_{\lambda \geq 0} e^{-\alpha\lambda}\big(K - xe^{-\lambda}\big)^+ = \begin{cases} K - x & \text{if } 0 \leq x \leq \frac{\alpha K}{1+\alpha}, \\ \frac{K}{1+\alpha}\left(\frac{\alpha K}{(1+\alpha)x}\right)^\alpha & \text{if } x \geq \frac{\alpha K}{1+\alpha}. \end{cases} \qquad (6.2)$$

On the other hand, in the case $\alpha > 0$, maximizing the integrand in (6.1) for every value of $S(T)$ shows that a process $\lambda \in \mathcal{R}$ is a maximizer if

$$\lambda(T) = \left(\log S(T) - \log \frac{\alpha K}{1+\alpha}\right)^+.$$

Example 6.5 (Lookback put) We consider the lookback put payoff

$$g(S(\cdot)) = M(S) - S(T) \quad \text{with maximum} \quad M(S) \triangleq \sup_{0 \leq t \leq T} S(t).$$

According to Proposition 5.2 and Theorem 5.3, the upper hedging price is

$$v(0, S(0); \alpha) = \sup_{\lambda \in \mathcal{R}} \mathbb{E}\big[e^{-rT-\alpha\lambda(T)}\big(M(Se^{-\lambda}) - S(T)e^{-\lambda(T)}\big)\big].$$

We maximize $M(Se^{-\lambda})$ over λ by choosing λ to be identically zero on $[0, T)$, and this results in $M(Se^{-\lambda}) = M(S)$. The upper hedging price is obtained by then choosing $\lambda(T) \geq 0$ so as to maximize $\mathbb{E}\big[e^{-rT-\alpha\lambda(T)}\big(M(S) - S(T)e^{-\lambda(T)}\big)\big]$. This is the maximization problem of (6.1) with $M(S)$ replacing the strike price K. While Theorem 2.2 does not apply, direct calculation as in Example 6.4 for $\alpha > 0$ shows that a maximizing process in $\mathcal{R}$ is given by

$$\lambda^*(t) = \left(\log S(T) - \log \frac{\alpha M(S)}{1+\alpha}\right)^+ I_{\{t=T\}}, \quad 0 \leq t \leq T.$$

Example 6.6 (Asian put) We next consider the Asian payoff

$$g(S(\cdot)) = \big(A(S) - S(T)\big)^+$$

with arithmetic average $A(S) \triangleq \frac{1}{T}\int_0^T S(t)\,dt$. The upper hedging price is

$$v(0, S(0); \alpha) = \sup_{\lambda \in \mathcal{R}} \mathbb{E}\big[e^{-rT-\alpha\lambda(T)}\big(A(Se^{-\lambda}) - S(T)e^{-\lambda(T)}\big)^+\big].$$

Once again, for $\alpha > 0$, a maximizing λ is identically zero on $[0, T)$, and for this process $A(Se^{-\lambda}) = A(S)$. The upper hedging price is obtained by choosing $\lambda(T) \geq 0$ so as to maximize

$$\mathbb{E}\big[e^{-rT-\alpha\lambda(T)}\big(A(S) - S(T)e^{-\lambda(T)}\big)^+\big].$$

This is the maximization problem of (6.1) with $A(S)$ replacing the strike price K. A maximizing process in $\mathcal{R}$ is

$$\lambda^*(t) = \left(\log S(T) - \log \frac{\alpha A(S)}{1+\alpha}\right)^+ I_{\{t=T\}}, \quad t \in [0, T].$$

Example 6.7 (Book of two barrier options) Upper hedging prices for individual exotic options are often too high to permit sales except in thinly traded over-the-counter markets. However, the theory developed in this paper can be applied to books of derivatives, and because the upper hedging methodology exploits natural hedges within the book, the upper hedging price of the book can be considerably less than the sum of the upper hedging prices of the individual assets in the book. The difficulty with pricing books, of course, is in solving the resulting stochastic control problem of Theorem 3.1 or 5.3.

In this example, we determine the upper hedging price of a book of two in-the-money knock-out calls of the type discussed in Sections 1 and 4. These calls have a common maturity T, a common strike price K, and the barriers L and U are related by $0 < K < L < U$. The "low barrier" call has payoff

$$g^L(S(\cdot)) = \big(S(T) - K\big)^+ I_{\{\max_{0\leq t\leq T} S(t) < L\}},$$

and the "high barrier" call has payoff

$$g^U(S(\cdot)) = \big(S(T) - K\big)^+ I_{\{\max_{0\leq t\leq T} S(t) < U\}}.$$

Let $v^L(t, x; \alpha)$ and $v^U(t, x; \alpha)$ be the functions given by (4.7), (4.9) with B replaced by L and U, respectively. These functions provide the upper hedging prices of the calls, except at the respective barriers. They further satisfy their respective versions of (4.12), (4.13) and (4.14).

Using Theorem 3.1 to determine the upper hedging price of a book consisting of one call with barrier U and $\kappa \geq 0$ calls with barrier L, we must compute

$$v(0, S(0); \alpha) = \sup_{\lambda \in \mathcal{C}} \mathbb{E}\big[e^{-rT-\alpha\lambda(T)}\big(\kappa g^L(Se^{-\lambda}) + g^U(Se^{-\lambda})\big)\big]. \tag{6.3}$$

We shall instead compute

$$\begin{aligned} &v^*(0, S(0); \alpha) \\ &= \sup_{\lambda \in \mathcal{C}} \mathbb{E}\big[e^{-rT-\alpha\lambda(T)}(S_\lambda(T) - K)^+\big(\kappa I_{\{M_\lambda(T)\leq L\}} + I_{\{M_\lambda(T)\leq U\}}\big)\big], \end{aligned} \tag{6.4}$$

where $S_\lambda(t) = S(t)e^{-\lambda(t)}$, $M_\lambda(t) = \max_{0\leq u\leq t} S_\lambda(u)$, and subsequently argue that $v(0, S(0); \alpha)$ and $v^*(0, S(0); \alpha)$ agree except when $S(0) = L$ or $S(0) = U$, in which cases we have $v(0, L; \alpha) = \lim_{x\downarrow L} v^*(0, x; \alpha)$ and $v(0, U; \alpha) = 0$.

We first construct a function $w(t,x,y)$ which we shall show is almost the upper hedging price of the book at time $t \in [0,T]$ if at that time the stock price is $S(t) = x > 0$ and the maximum stock price to date is $M(t) \triangleq \max_{0 \le u \le t} S(u) = y \ge x$. We begin by setting

$$w(t,x,y) \triangleq v^U(t,x;\alpha) I_{\{y \le U\}}, \quad 0 \le t \le T,\, y > L,\, 0 < x \le y. \tag{6.5}$$

We next define $\bar{\kappa} \triangleq \kappa + 1$ and

$$t^* \triangleq T \wedge \min\{t \ge 0;\, \bar{\kappa} v^L(s,L;\alpha) \ge v^U(s,L;\alpha)\ \forall s \in [t,T]\}.$$

We set

$$w(t,x,y) \triangleq \bar{\kappa} v^L(t,x;\alpha), \quad t^* < t \le T,\, 0 < x \le y \le L. \tag{6.6}$$

Finally, for $0 \le t \le t^*$ and $0 \le x \le y \le L$, we define $w(t,x,y)$ to be the solution to the Black-Scholes partial differential equation

$$w_t(t,x) + rxw_x(t,x) + \frac{1}{2}\sigma^2 x^2 w_{xx}(t,x) = rw(t,x), \tag{6.7}$$

subject to the boundary conditions

$$w(t,0) = 0, \qquad 0 \le t < t^*, \tag{6.8}$$

$$w(t,L) = v^U(t,L;\alpha), \qquad 0 \le t < t^*, \tag{6.9}$$

$$w(t^*,x) = \bar{\kappa} v^L(t^*,x;\alpha), \quad 0 < x \le L. \tag{6.10}$$

In other words, in the region $0 \le t \le t^*$, $0 < x \le y \le L$, the function $w(t,x,y)$ is the Black-Scholes price of a derivative security which knocks out at L, paying rebate $v^U(s,L;\alpha)$ if the knock-out occurs at time $s < t^*$, and otherwise expires at time t^*, paying $\bar{\kappa} v^L(t^*,S(t^*);\alpha)$ upon expiration. Because $\alpha w(t,x) + xw_x(t,x)$ also satisfies the Black-Scholes equation and $\alpha w(t,x) + xw_x(t,x) \ge 0$ on the boundary $(\{t^*\} \times [0,L)) \cup ([0,t^*) \times \{L\})$ (see (4.14), satisfied by both v^L and v^U), we have from the maximum principle that

$$\alpha w(t,x) + xw_x(t,x) \ge 0, \quad 0 \le t \le t^*,\, 0 \le x \le L. \tag{6.11}$$

We show that $v^*(0,S(0);\alpha)$ given by (6.4) is actually $w(0,S(0),S(0))$. If $S(0) > U$, then both $v^*(0,S(0);\alpha)$ and $w(0,S(0),S(0))$ are zero. If $L < S(0) \le U$, then the computation of $v^*(t,S(0);\alpha)$ reduces to the single-option problem solved in Corollary 4.1 with $B = U$, and hence $v^*(0,S(0);\alpha) = v^U(0,S(0);\alpha) = w(t,S(0),S(0))$ by (6.5).

In remains to prove the equality when $0 < S(0) \le L$. This requires the proof of an inequality in each direction. For the first inequality, we let $\lambda \in \mathcal{C}$ be given and define $\Theta_L \triangleq \inf\{t \ge 0;\, S_\lambda(t) > L\}$, $\Theta_U \triangleq \inf\{t \ge 0;\, S_\lambda(t) > U\}$. Itô's formula implies

$$
\begin{aligned}
&w(0,S(0),S(0))\\
&= w(0,S(0)) \qquad\qquad (6.12)\\
&= \mathbb{E}\big[e^{-r(t^*\wedge\Theta_L)-\alpha\lambda(t^*\wedge\Theta_L)}w(t^*\wedge\Theta_L, S_\lambda(t^*\wedge\Theta_L))\big]\\
&\quad+\mathbb{E}\bigg[\int_0^{t^*\wedge\Theta_L} e^{-rt-\alpha\lambda(t)}\big(\alpha w(t,S_\lambda(t))+S_\lambda(t)w(t,S_\lambda(t))\big)\,d\lambda(t)\bigg]\\
&\geq \mathbb{E}\big[e^{-r\Theta_L-\alpha\lambda(\Theta_L)}v^U(\Theta_L,L;\alpha)I_{\{\Theta_L<t^*\}}\\
&\qquad+\tilde{\kappa}e^{-rt^*-\alpha\lambda(t^*)}v^L(t^*,S_\lambda(t^*);\alpha)I_{\{\Theta_L\geq t^*\}}\big].
\end{aligned}
$$

We continue with the case $\Theta_L < t^*$, again using Itô's formula, this time to obtain

$$
\begin{aligned}
&\mathbb{E}\big[e^{-r\Theta_L-\alpha\lambda(\Theta_L)}v^U(\Theta_L,L;\alpha)I_{\{\Theta_L<t^*\}}\big]\\
&= \mathbb{E}\big[e^{-r(T\wedge\Theta_U)-\alpha\lambda(T\wedge\Theta_U)}v^U(T\wedge\Theta_U,S_\lambda(T\wedge\Theta_U);\alpha)I_{\{\Theta_L<t^*\}}\big]\\
&\quad+\mathbb{E}\bigg[\int_{\Theta_L}^{T\wedge\Theta_U} e^{-rt-\alpha\lambda(t)}\big(\alpha v^U(t,S_\lambda(t);\alpha) \qquad\qquad (6.13)\\
&\qquad\qquad+S_\lambda(t)v^U_x(t,S_\lambda(t);\alpha)\big)\,d\lambda(t)I_{\{\Theta_L<t^*\}}\bigg]\\
&\geq \mathbb{E}\big[e^{-r\Theta_U-\alpha\lambda(\Theta_U)}v^U(\Theta_U,U;\alpha)I_{\{\Theta_L<t^*,\Theta_U<T\}}\\
&\qquad\qquad+e^{-rT-\alpha\lambda(T)}v^U(T,S_\lambda(T);\alpha)I_{\{\Theta_L<t^*,\Theta_U\geq T\}}\big]\\
&\geq \mathbb{E}\big[e^{-rT-\alpha\lambda(T)}(S_\lambda(T)-K)^+I_{\{\Theta_L<t^*,M_\lambda(T)\leq U\}}\big].
\end{aligned}
$$

We also continue (6.12) in the case $\Theta_L \geq t^*$. In this case, we have

$$
\begin{aligned}
&\tilde{\kappa}\,\mathbb{E}\big[e^{-rt^*-\alpha\lambda(t^*)}v^L(t^*,S_\lambda(t^*);\alpha)I_{\{\Theta_L\geq t^*\}}\big]\\
&= \tilde{\kappa}\,\mathbb{E}\big[e^{-r(T\wedge\Theta_L)-\alpha\lambda(T\wedge\Theta_L)}v^L(T\wedge\Theta_L,S_\lambda(T\wedge\Theta_L);\alpha)I_{\{\Theta_L\geq t^*\}}\big]\\
&\quad+\tilde{\kappa}\,\mathbb{E}\bigg[\int_{t^*}^{T\wedge\Theta_L} e^{-rt-\alpha\lambda(t)}\big(\alpha v^L(t,S_\lambda(t);\alpha) \qquad\qquad (6.14)\\
&\qquad\qquad+S_\lambda(t)v^L_x(t,S_\lambda(t);\alpha)\big)\,d\lambda(t)I_{\{\Theta_L\geq t^*\}}\bigg]\\
&\geq \tilde{\kappa}\,\mathbb{E}\big[e^{-r\Theta_L-\alpha\lambda(\Theta_L)}v^L(\Theta_L,L;\alpha)I_{\{t^*\leq\Theta_L<T\}}\big]\\
&\quad+\tilde{\kappa}\,\mathbb{E}\big[e^{-rT-\alpha\lambda(T)}v^L(T,S_\lambda(T);\alpha)I_{\{\Theta_L\geq T\}}\big]\\
&\geq \mathbb{E}\big[e^{-r\Theta_L-\alpha\lambda(\Theta_L)}v^U(\Theta_L,L;\alpha)I_{\{t^*\leq\Theta_L<T\}}\big]\\
&\quad+\tilde{\kappa}\,\mathbb{E}\big[e^{-rT-\alpha\lambda(T)}(S_\lambda(T)-K)^+I_{\{M_\lambda(T)\leq L\}}\big],
\end{aligned}
$$

where the definition of t^* is used to obtain the last inequality. Finally, Itô's formula implies

$$
\begin{aligned}
&\mathbb{E}\big[e^{-r\Theta_L-\alpha\lambda(\Theta_L)}v^U(\Theta_L,L;\alpha)I_{\{t^*\le\Theta_L<T\}}\big]\\
&=\mathbb{E}\big[e^{-r(T\wedge\Theta_U)-\alpha\lambda(T\wedge\Theta_U)}v^U(T\wedge\Theta_U,S_\lambda(T\wedge\Theta_U);\alpha)I_{\{t^*\le\Theta_L<T\}}\big]\\
&\quad+\mathbb{E}\Bigg[\int_{\Theta_L}^{T\wedge\Theta_U}e^{-rt-\alpha\lambda(t)}\big(\alpha v^U(t,S_\lambda(t);\alpha) \qquad (6.15)\\
&\qquad\qquad+S_\lambda(t)v^U_x(t,S_\lambda(t);\alpha)\big)\,d\lambda(t)I_{\{t^*\le\Theta_L<T\}}\Bigg]\\
&\ge\mathbb{E}\big[e^{-rT-\alpha\lambda(T)}v^U(T,S_\lambda(T);\alpha)I_{\{t^*\le\Theta_L<T,\,\Theta_U\ge T\}}\big]\\
&\ge\mathbb{E}\big[e^{-rT-\alpha\lambda(T)}\big(S_\lambda(T)-K\big)^+I_{\{t^*\le\Theta_L<T,\,M_\lambda(T)\le U\}}\big].
\end{aligned}
$$

Combining (6.12)–(6.15), we see that

$$
\begin{aligned}
&w(0,S(0),S(0)) \qquad (6.16)\\
&\ge\mathbb{E}\big[e^{-rT-\alpha\lambda(T)}\big(S_\lambda(T)-K\big)^+\big(I_{\{\Theta_L<T,\,M_\lambda(T)\le U\}}+\tilde{\kappa}I_{\{M_\lambda(T)\le L\}}\big)\big]\\
&=\mathbb{E}\big[e^{-rT-\alpha\lambda(T)}\big(S_\lambda(T)-K\big)^+\big(\kappa I_{\{M_\lambda(T)\le L\}}+I_{\{M_\lambda(T)\le U\}}\big)\big].
\end{aligned}
$$

Recalling (6.4) and using the fact that $\lambda\in\mathcal{C}$ is arbitrary, we conclude

$$w(0,S(0),S(0))\ge v^*(0,S(0);\alpha). \qquad (6.17)$$

For the reverse inequality in the case $0<S(0)\le L$, we define

$$\lambda^*(t)=\begin{cases}\max\limits_{0\le u\le t}\big(\log S(u)-\log U\big)^+ & \text{if } 0\le t\le t^*,\\ \max\limits_{0\le u\le t}\big(\log S(u)-\log U\big)^+I_{\{M(t^*)>L\}} & \\ \quad+\max\limits_{t^*\le u\le t}\big(\log S(u)-\log L\big)^+I_{\{M(t^*)\le L\}} & \text{if } t^*\le t\le T.\end{cases} \qquad (6.18)$$

Then S_{λ^*} never exceeds U, and if by time t^*, S_{λ^*} has not exceeded L, then it never exceeds L. The process λ^* is the minimal process which guarantees these properties; in particular, λ^* grows only when S_{λ^*} is at either U or L. Replacing λ in (6.12) by λ^*, we have equality because $\lambda^*\equiv 0$ on $[0,t^*\wedge\Theta_L]$. Replacing λ by λ^* in (6.13), we again have equality because on the set $\{\Theta_L<t^*\}$, the process λ^* grows only when $S_{\lambda^*}=U$, and

$$\alpha v^U(t,U;\alpha)+Uv^U_x(t,U;\alpha)=0.$$

Furthermore, $\{\Theta_L<t^*,\Theta_U<T\}=\varnothing$, since $M_{\lambda^*}(T)\le U$. With λ replaced by λ^*, (6.14) becomes an equality because on the set $t^*\le t\le T\wedge\Theta_L$, the process λ^* grows only when $S_{\lambda^*}=L$ and

$$\alpha v^L(t,L;\alpha)+Lv^L_x(t,L;\alpha)=0.$$

Furthermore, $\{t^*\le\Theta_L<T\}=\varnothing$ because $M_{\lambda^*}(T)\le L$ on $\Theta_L(T)\ge t^*$. For this same reason, all terms in (6.15) are zero when λ is replaced by λ^*. These

observations lead to equality in (6.16) when λ is replaced by λ^*, and according to (6.4).

$$w(0, S(0), S(0)) = \mathbf{E}\big[e^{-rT-\alpha\lambda^*(T)}(S_{\lambda^*}(T) - K)^+\big(\kappa I_{\{M_{\lambda^*}(T)\le L\}} + I_{\{M_{\lambda^*}(T)\le U\}}\big)\big] \\ \le v^*(0, S(0); \alpha). \tag{6.19}$$

Due to (6.17), equality has to hold here.

To establish the relationship between $v^*(0, S(0); \alpha)$ and the upper hedging price $v(0, S(0); \alpha)$ of (6.3), we start with the case $S(0) < U$ and choose two sequences of barriers $\{L_n\}_{n=1}^{\infty}$ and $\{U_n\}_{n=1}^{\infty}$ with $L_n \uparrow L$ and $U_n \uparrow U$ satisfying $L_n < L \le U_n$ and $S(0) \le U_n < U$ for all $n \in \mathbb{N}$. Let λ_n be given by (6.18) with L and U replaced by L_n and U_n, respectively. Then $\lambda_n \in \mathcal{C}$ and $\lambda_n \downarrow \lambda^*$ pointwise, so $S_{\lambda_n} \uparrow S_{\lambda^*}$ and $M_{\lambda_n} \uparrow M_{\lambda^*}$ pointwise. In addition, $\{M_{\lambda_n}(T) \le U_n\} = \Omega = \{M_{\lambda^*}(T) \le U\}$ and for $S(0) \in (0, U) \setminus \{L\}$, we also have

$$\lim_{n\to\infty} I_{\{M_{\lambda_n}(T)\le L_n\}} = \lim_{n\to\infty} I_{\{M(t^*)\le L_n\}} = I_{\{M(t^*)\le L\}} = I_{\{M_{\lambda^*}(T)\le L\}}, \quad \text{a.s.}$$

It follows from (6.3), $L_n < L$, $U_n < U$, Fatou's lemma and (6.19) with equality that, for $S(0) \in (0, U) \setminus \{L\}$,

$$\begin{aligned} &v(0, S(0); \alpha) \\ &\ge \liminf_{n\to\infty} \mathbf{E}\big[e^{-rT-\lambda_n(T)}\big(S_{\lambda_n}(T) - K\big)^+\big(\kappa I_{\{M_{\lambda_n}(T)<L\}} + I_{\{M_{\lambda_n}(T)<U\}}\big)\big] \\ &\ge \liminf_{n\to\infty} \mathbf{E}\big[e^{-rT-\lambda_n(T)}\big(S_{\lambda_n}(T) - K\big)^+\big(\kappa I_{\{M_{\lambda_n}(T)\le L_n\}} + I_{\{M_{\lambda_n}(T)\le U_n\}}\big)\big] \\ &\ge \mathbf{E}\big[e^{-rT-\lambda^*(T)}\big(S_{\lambda^*}(T) - K\big)^+\big(\kappa I_{\{M_{\lambda^*}(T)\le L\}} + I_{\{M_{\lambda^*}(T)\le U\}}\big)\big] \\ &= v^*(0, S(0); \alpha). \end{aligned}$$

The reverse inequality is obvious. Since the case $S(0) > U$ is trivial, we have established

$$v(0, S(0); \alpha) = v^*(0, S(0); \alpha), \quad \forall\, S(0) \in (0, \infty) \setminus \{L, U\}.$$

It is clear that $v(0, U; \alpha) = 0$, since both options are knocked-out at the initial time. Finally, if $S(0) = L$, then the "low barrier" option is knocked out at the initial time, and by Corollary 4.1,

$$v(0, L; \alpha) = v^U(0, L; \alpha) = \lim_{x\downarrow L} v^U(0, x; \alpha) = \lim_{x\downarrow L} v^*(0, x; \alpha).$$

The construction of the upper hedging price for the book of two barrier options is complete. $\diamondsuit$

7 Proof of Theorem 3.1

We denote by

$$\mathcal{L} \triangleq \big\{\lambda; \lambda \text{ is a nondecreasing, } \{\mathcal{F}(t); 0 \le t \le T\}\text{-adapted process,} \\ \text{Lipschitz in } t \text{ uniformly in } \omega, \text{ with } \lambda(0) = 0\big\},$$

the class of processes over which the supremum in (2.4) is taken. In a first step, we show that the supremum in (2.4) can be reduced to the set $\mathcal{L}_{\mathrm{pl}}$ of piecewise linear, $\{\mathcal{F}^W(t); 0 \le t \le T\}$-adapted processes $\lambda \in \mathcal{L}$, meaning that for every $\lambda \in \mathcal{L}_{\mathrm{pl}}$ there exist $m \in \mathbb{N}$, a partition $0 = t_0 < t_1 < \cdots < t_m = T$, and bounded, $\mathcal{F}^W(t_i)$-measurable $a_i\colon \Omega \to [0,\infty)$ such that

$$\lambda(t,\omega) = \sum_{i=0}^{m-1} a_i(\omega)\big((t_{i+1} \wedge t) - t_i\big)^+, \quad t \in [0,T],\ \omega \in \Omega. \tag{7.1}$$

Lemma 7.1 *Let g be a nonnegative measurable function defined on $C_+[0,T]$. We have*

$$\sup_{\lambda\in\mathcal{L}} \mathbb{E}_\lambda\big[e^{-rT-\alpha\lambda(T)}g(S)\big] = \sup_{\lambda\in\mathcal{L}_{\mathrm{pl}}} \mathbb{E}_\lambda\big[e^{-rT-\alpha\lambda(T)}g(S)\big].$$

Proof: Since $\mathcal{L}_{\mathrm{pl}} \subset \mathcal{L}$, there is just one inequality to prove. Consider $\lambda \in \mathcal{L}$. Then there is a bounded, adapted process $\lambda'\colon [0,T] \times \Omega \to [0,\infty)$ such that

$$\lambda(t,\omega) = \int_0^t \lambda'(s,\omega)\,ds, \quad t \in [0,T].$$

As in the construction of the stochastic integral (see, e.g., [12], Chap. 3, Lemma 2.4), one can prove the existence of a sequence $\{\lambda_n'\}_{n\in\mathbb{N}}$ of processes $\lambda_n'\colon [0,T] \times \Omega \to [0,\infty)$ of the form

$$\lambda_n'(t,\omega) = \sum_{i=0}^{m_n-1} a_{i,n}(\omega)1_{(t_{i,n},t_{i+1,n}]}(t), \quad t \in [0,T],\ \omega \in \Omega,$$

where $m_n \in \mathbb{N}$, $0 = t_{0,n} < t_{1,n} < \cdots < t_{m_n,n} = T$ and every $a_{i,n}\colon \Omega \to [0,\infty)$ is $\mathcal{F}(t_{i,n})$-measurable and bounded by the Lipschitz constant of λ, such that

$$\lim_{n\to\infty} \mathbb{E}\left[\int_0^T |\lambda_n'(t) - \lambda'(t)|^2\,dt\right] = 0.$$

By changing $a_{i,n}$ on a set of $\mathbb{P}$-measure zero if necessary, we may assume that every $a_{i,n}$ is $\mathcal{F}^W(t_{i,n})$-measurable. By the definition of the stochastic integral,

$$\lim_{n\to\infty} \int_0^T \lambda_n'(t)\,dW(t) = \int_0^T \lambda'(t)\,dW(t) \tag{7.2}$$

in $L^2(\Omega, \mathcal{F}(T), \mathbb{P})$. By passing to a subsequence if necessary, we may assume that

$$\lim_{n\to\infty} \int_0^T |\lambda_n'(t) - \lambda'(t)|^2\,dt = 0 \quad \mathbb{P}\text{-almost surely}$$

and that the convergence in (7.2) is also $\mathbb{P}$-almost sure. Define

$$\lambda_n(t,\omega) = \int_0^t \lambda_n'(s,\omega)\,ds, \quad t \in [0,T],\ \omega \in \Omega.$$

Then $\lambda_n(T) \to \lambda(T)$ almost surely. Let Z_λ denote the density given by (2.5) and let Z_{λ_n} denote the corresponding density for λ_n. By Fatou's lemma,

$$\begin{aligned}\mathbb{E}_\lambda[e^{-\alpha\lambda(T)}f] &= \mathbb{E}[e^{-\alpha\lambda(T)}fZ_\lambda]\\ &\le \liminf_{n\to\infty} \mathbb{E}[e^{-\alpha\lambda_n(T)}fZ_{\lambda_n}] = \liminf_{n\to\infty} \mathbb{E}_{\lambda_n}[e^{-\alpha\lambda_n(T)}f],\end{aligned}$$

where $f \triangleq e^{-rT}g(S)$. ◇

Lemma 7.2 *Let g be a nonnegative lower-semicontinuous function defined on $C_+[0,T]$. We have*

$$\sup_{\lambda\in\mathcal{L}_{\mathrm{pl}}} \mathbb{E}\big[e^{-rT-\alpha\lambda(T)}g(Se^{-\lambda})\big] = \sup_{\lambda\in\mathcal{C}} \mathbb{E}\big[e^{-rT-\alpha\lambda(T)}g(Se^{-\lambda})\big].$$

PROOF: Since $\mathcal{L}_{\mathrm{pl}} \subset \mathcal{C}$, there is just one inequality to prove. Consider $\lambda \in \mathcal{C}$. Given $n \in \mathbb{N}$, define $a_{0,n} = 0$, $t_{0,n} = 0$, $t_{i,n} = iT2^{-n}$ and

$$a_{i,n}(\omega) = \min\Big\{2^n, \frac{\lambda(t_{i,n},\omega) - \lambda(t_{i-1,n},\omega)}{T2^{-n}}\Big\} \tag{7.3}$$

for all $i \in \{1,2,\dots,2^n\}$ and $\omega \in \Omega$. By changing $a_{i,n}$ on a set of $\mathbb{P}$-measure zero if necessary, we may assume that $a_{i,n}$ is $\mathcal{F}^W(t_{i,n})$-measurable for every $i \in \{1,\dots,2^n\}$. Then

$$\lambda_n(t,\omega) = \sum_{i=0}^{2^n-1} a_{i,n}(\omega)\big((t_{i+1,n}\wedge t) - t_{i,n}\big)^+, \quad t \in [0,T],\ \omega\in\Omega,$$

is in $\mathcal{L}_{\mathrm{pl}}$.

To prove the $\mathbb{P}$-almost sure uniform convergence of $\{\lambda_n\}_{n\in\mathbb{N}}$ to λ, consider an $\omega \in \Omega$ so that (7.3) holds for all $n \in \mathbb{N}$ and $i \in \{1,\dots,2^n\}$. Since the mapping $t \mapsto \lambda(t,\omega)$ is uniformly continuous and nondecreasing on $[0,T]$, the modulus of continuity

$$m(t,\omega) = \sup_{s\in[0,T-t]} (\lambda(s+t,\omega) - \lambda(s,\omega)), \quad t\in[0,T],$$

is a nondecreasing continuous function with $m(0,\omega) = 0$. In particular, there exists $k \in \mathbb{N}$ satisfying $m(T2^{-k},\omega) < T$. For every $n \ge k$ and $i \in \{1,\dots,2^n\}$, the quotient in (7.3) is less than 2^n; hence $\lambda_n(t_{i+1,n},\omega) = \lambda(t_{i,n},\omega)$ for all $i \in \{0,\dots,2^n-1\}$. If $t \in [t_{i,n}, t_{i+1,n}]$ with $i \in \{1,\dots,2^n-1\}$, then $\lambda(t_{i-1,n},\omega) = \lambda_n(t_{i,n},\omega) \le \lambda_n(t,\omega) \le \lambda_n(t_{i+1,n},\omega) = \lambda(t_{i,n},\omega)$, because λ_n is nondecreasing. Note that $\lambda_n(t,\omega) = 0$ for $t \in [0,t_{1,n}]$. Because λ is also nondecreasing, we have

$$\sup_{t\in[0,T]} |\lambda(t,\omega) - \lambda_n(t,\omega)| \le m(T2^{-n},\omega) \downarrow 0 \quad \text{as } n\to\infty.$$

Using the lower semicontinuity of g and Fatou's lemma, we compute

$$\begin{aligned}\mathbb{E}\big[e^{-rT-\alpha\lambda(T)}g(Se^{-\lambda})\big] &\le \mathbb{E}\Big[e^{-rT-\alpha\lambda(T)} \liminf_{n\to\infty} g(Se^{-\lambda_n})\Big]\\ &\le \liminf_{n\to\infty} \mathbb{E}\big[e^{-rT-\alpha\lambda_n(T)}g(Se^{-\lambda_n})\big],\end{aligned}$$

which implies the desired inequality. ◇

PROOF OF THEOREM 3.1: In light of Theorem 2.1 and Lemmas 7.1, 7.2, it suffices to show

$$\sup_{\lambda\in\mathcal{L}_{\mathrm{pl}}} \mathbb{E}_\lambda\big[e^{-rT-\alpha\lambda(T)}g(S)\big] = \sup_{\bar\lambda\in\mathcal{L}_{\mathrm{pl}}} \mathbb{E}\big[e^{-rT-\alpha\bar\lambda(T)}g(Se^{-\bar\lambda})\big].$$

In other words, for each $\lambda \in \mathcal{L}_{\mathrm{pl}}$, we will construct a $\bar\lambda \in \mathcal{L}_{\mathrm{pl}}$ such that

$$\mathbb{E}_\lambda\big[e^{-rT-\alpha\lambda(T)}g(S)\big] = \mathbb{E}\big[e^{-rT-\alpha\bar\lambda(T)}g(Se^{-\bar\lambda})\big], \tag{7.4}$$

and conversely, for each $\bar\lambda \in \mathcal{L}_{\mathrm{pl}}$, we will construct $\lambda \in \mathcal{L}_{\mathrm{pl}}$ satisfying (7.4).

For each $\lambda \in \mathcal{L}_{\mathrm{pl}}$, we define $\varphi_\lambda : \Omega \to \Omega$ by

$$\varphi_\lambda(\omega)(t) \triangleq \omega(t) + \frac{1}{\sigma}\lambda(t,\omega), \quad t \in [0,T], \ \omega \in \Omega. \tag{7.5}$$

Note that φ_λ is $\mathcal{F}^W(t)/\mathcal{F}^W(t)$-measurable for every $t \in [0,T]$. We show that φ_λ is bijective. To see that φ_λ is injective, we suppose that $\varphi_\lambda(\omega_1) = \varphi_\lambda(\omega_2)$. The process λ has a representation of the form (7.1), and in terms of this representation we define

$$I \triangleq \max\{i;\ \omega_1(t) = \omega_2(t)\ \forall t \in [0,t_i]\}.$$

Note that $0 \le I \le m$. If $I < m$, then $a_i(\omega_1) = a_i(\omega_2)$ for all $i \le I$, which implies that $\lambda(t,\omega_1) = \lambda(t,\omega_2)$ for all $t \le t_{I+1}$. Therefore, for $0 \le t \le t_{I+1}$, we have

$$\omega_1(t) = \varphi_\lambda(\omega_1)(t) - \frac{1}{\sigma}\lambda(t,\omega_1) = \varphi_\lambda(\omega_2)(t) - \frac{1}{\sigma}\lambda(t,\omega_2) = \omega_2(t),$$

and the definition of I is contradicted. It follows that $I = m$ and $\omega_1 = \omega_2$. To see that φ_λ is surjective, we let $\bar\omega \in \Omega$ be given. We set $\omega(0) = 0$ and define inductively, for $i \in \{0,1,\dots,m-1\}$,

$$\omega(t) \triangleq \bar\omega(t) - \frac{1}{\sigma}\sum_{j=0}^{i} a_j(\omega)\big((t_{j+1}\wedge t) - t_j\big)^+, \quad t_i < t \le t_{i+1}.$$

Then $\bar\omega = \varphi_\lambda(\omega)$. This construction also shows that φ_λ^{-1} is $\mathcal{F}^W(t)/\mathcal{F}^W(t)$-measurable for every $t \in [0,T]$.

Now let $\lambda \in \mathcal{L}_{\mathrm{pl}}$ be given. We define $\bar\lambda \in \mathcal{L}_{\mathrm{pl}}$ by

$$\bar\lambda(\cdot,\bar\omega) \triangleq \lambda(\cdot,\varphi_\lambda^{-1}(\bar\omega)), \quad \forall \bar\omega \in \Omega, \tag{7.6}$$

and verify that (7.4) holds. According to Girsanov's theorem, if we impose on ω the measure $\mathbb{P}_\lambda$ given by (2.5), then $\bar\omega = \varphi_\lambda(\omega)$ is distributed according to

Wiener measure $\mathbb{P}$. Therefore,

$$
\begin{aligned}
&\mathbb{E}_\lambda\left[e^{-rT-\alpha\lambda(T)}g(S)\right] \\
&= \int_\Omega e^{-rT-\alpha\lambda(T,\omega)} g\big(S(0)\exp(\sigma W(\cdot,\omega)+\mu)\big)\,\mathbb{P}_\lambda(d\omega) \\
&= \int_\Omega e^{-rT-\alpha\bar\lambda(T,\bar\omega)} g\big(S(0)\exp(\sigma W(\cdot,\omega)+\lambda(\cdot,\omega)+\mu-\bar\lambda(\cdot,\bar\omega))\big)\,\mathbb{P}_\lambda(d\omega) \\
&= \int_\Omega e^{-rT-\alpha\bar\lambda(T,\bar\omega)} g\big(S(0)\exp(\sigma\bar\omega+\mu-\bar\lambda(\cdot,\bar\omega))\big)\,\mathbb{P}_\lambda(d\omega) \\
&= \int_\Omega e^{-rT-\alpha\bar\lambda(T,\bar\omega)} g\big(S(0)\exp(\sigma\bar\omega+\mu-\bar\lambda(\cdot,\bar\omega))\big)\,\mathbb{P}(d\bar\omega) \\
&= \int_\Omega e^{-rT-\alpha\bar\lambda(T,\bar\omega)} g\big(S(0)\exp(\sigma W(\cdot,\bar\omega)+\mu-\bar\lambda(\cdot,\bar\omega))\big)\,\mathbb{P}(d\bar\omega) \\
&= \mathbb{E}\big[e^{-rT-\alpha\bar\lambda(T)}g(Se^{-\bar\lambda})\big],
\end{aligned}
$$

which is (7.4).

For the converse, let $\bar\lambda \in \mathcal{L}_{\text{pl}}$ be given. The function $\psi_{\bar\lambda}\colon \Omega \to \Omega$ defined by

$$\psi_{\bar\lambda}(\bar\omega)(t) \triangleq \bar\omega(t) - \frac{1}{\sigma}\bar\lambda(t,\bar\omega), \quad t \in [0,T],\ \bar\omega \in \Omega$$

is bijective, and both it and its inverse are $\mathcal{F}^W(t)/\mathcal{F}^W(t)$-measurable for every $t \in [0,T]$ (it is formally merely $\varphi_{-\bar\lambda}$). Therefore, we can define $\lambda \in \mathcal{L}_{\text{pl}}$ by

$$\lambda(\cdot,\omega) \triangleq \bar\lambda(\cdot,\psi_{\bar\lambda}^{-1}(\omega)), \quad \forall \omega \in \Omega. \tag{7.7}$$

According to the definitions and with the notation $\omega = \psi_{\bar\lambda}(\bar\omega)$, we have

$$\varphi_\lambda(\omega) = \omega + \frac{1}{\sigma}\lambda(\cdot,\omega) = \psi_{\bar\lambda}(\bar\omega) + \frac{1}{\sigma}\bar\lambda(\cdot,\bar\omega) = \bar\omega.$$

In other words, φ_λ and $\psi_{\bar\lambda}$ are inverse functions and the relationship (7.7) between λ and $\bar\lambda$ coincides with the relationship (7.6). It follows that (7.4) again holds, and the theorem is proved. $\diamond$

8 Proof of Theorem 5.3

We use the notation introduced in Section 5.

Lemma 8.1 *Given $g\colon C_+[0,T] \to [0,\infty)$, the map $g_*\colon C_+[0,T] \times R[0,T] \to [0,\infty)$ defined by (5.3) is lower semicontinuous in the second argument with respect to the topology of pointwise convergence on $R[0,T]$. If g is lower semicontinuous, then*

$$g_*(y,\lambda) = g(ye^{-\lambda}) \quad \forall (y,\lambda) \in C_+[0,T] \times (R[0,T] \cap C_+[0,T]). \tag{8.1}$$

PROOF: To prove lower semicontinuity in the second argument, let $y \in C_+[0,T]$ and $\lambda \in R[0,T]$ be given, and let $\lambda_n \to \lambda$ pointwise, where each λ_n is in $R[0,T]$. Let $\varepsilon > 0$ be given. According to the definition of g_*, for each n we may choose $\eta_n \in R[0,T] \cap C_+[0,T]$ such that $g(ye^{-\eta_n}) \leq \varepsilon + g_*(y, \lambda_n)$ and $d(\eta_n, \lambda_n) < \frac{1}{n}$. But then $\eta_n \to \lambda$ pointwise, which implies

$$g_*(y,\lambda) \leq \liminf_{n\to\infty} g(ye^{-\eta_n}) \leq \varepsilon + \liminf_{n\to\infty} g_*(y, \lambda_n).$$

Since $\varepsilon > 0$ is arbitrary, we have lower semicontinuity of $g_*(y,\cdot)$ at λ.

Assume now that g is lower semicontinuous and $(y,\lambda) \in C_+[0,T] \times (R[0,T] \cap C_+[0,T])$. Let $\lambda_n \to \lambda$ pointwise. Remark 5.1 shows that $\lambda_n \to \lambda$ uniformly. We have

$$g(ye^{-\lambda}) \leq \liminf_{n\to\infty} g(ye^{-\lambda_n}),$$

and minimizing over sequences $\{\lambda_n\}_{n=1}^{\infty}$ we obtain $g(ye^{-\lambda}) \leq g_*(y,\lambda)$. Using the constant sequence $\{\lambda\}_{n=1}^{\infty}$, we obtain the reverse inequality. ◇

PROOF OF THEOREM 5.3: Since $\mathcal{C} \subset \mathcal{R}$, the inequality

$$v(0, S(0); \alpha) \leq \sup_{\lambda \in \mathcal{R}} \mathbb{E}\big[e^{-rT-\alpha\lambda(T)} g_*(S,\lambda)\big] \tag{8.2}$$

follows immediately from Theorem 3.1 and Lemma 8.1. For the reverse inequality, we show that

$$\sup_{\lambda \in \mathcal{C}} \mathbb{E}\big[e^{-rT-\alpha\lambda(T)} g(Se^{-\lambda})\big] \geq \sup_{\lambda \in \mathcal{R}} \mathbb{E}\big[e^{-rT-\alpha\lambda(T)} g_*(S,\lambda)\big]. \tag{8.3}$$

To do this, we choose a process $\lambda \in \mathcal{R}$, and approximate it by a sequence $\{\lambda_n\}_{n=1}^{\infty}$ of processes in $\mathcal{C}$.

Let $\lambda \in \mathcal{R}$ be given, and assume for the moment that $\lambda(T) \leq C$. We need to approximate λ by processes which are continuous. According to the definition of $\mathcal{R}$, there are finitely many pre-specified times $0 < t_1 < \cdots < t_I \leq T$ at which λ can be discontinuous. Denote the jumps of λ by

$$\alpha_i \triangleq \lambda(t_i) - \lambda(t_i-), \quad i = 1, \ldots, I,$$

and denote the continuous part of λ by

$$\lambda^c(t) \triangleq \lambda(t) - \sum_{\{i; t_i \leq t\}} \alpha_i, \quad 0 \leq t \leq T.$$

Set $M_i(t) \triangleq \mathbb{E}[\alpha_i \,|\, \mathcal{F}(t)]$ for $0 \leq t \leq T$. Each M_i is a bounded, nonnegative martingale, relative to the Brownian filtration $\{\mathcal{F}(t); 0 \leq t \leq T\}$, and must therefore have a continuous modification ([12], Theorem 3.13 of Chapter 1 and Problem 4.16 of Chapter 3). Without loss of generality, we assume therefore that each M_i is continuous.

Choose $N \in \mathbb{N}$ so that $t_1 \geq 1/N$. For all $n \geq N$ and $t \in [0,T]$, define

$$\lambda_n(t) \triangleq \lambda^c(t) + \sum_{i=1}^{I} \left[1 \wedge n\left(t - t_i + \frac{1}{n}\right)^+\right] \max_{t_i - \frac{1}{n} \leq s \leq t \wedge t_i} M_i(s).$$

Then λ_n is continuous, adapted, nondecreasing, and satisfies $\lambda_n(0) = 0$. If, in addition, n satisfies $t_{i+1} - t_i \geq 1/n$ for all $i \in \{1, \ldots, I-1\}$, then

$$\lambda_n(t_i) = \lambda^c(t_i) + \sum_{j=1}^{i} \max_{t_j - \frac{1}{n} \leq s \leq t_j} M_j(s)$$

for all $i \in \{1, \ldots, I\}$; in particular,

$$\lim_{n\to\infty} \lambda_n(t_i) = \lambda^c(t_i) + \sum_{j=1}^{i} \alpha_j = \lambda(t_i).$$

For $t \in [0, t_1)$ and sufficiently large n, we have $t \leq t_1 - 1/n$ and $\lambda_n(t) = \lambda^c(t) = \lambda(t)$. For $t \in (t_i, t_{i+1})$, we have $t \leq t_{i+1} - 1/n$ for sufficiently large n, and then

$$\lambda_n(t) = \lambda^c(t) + \sum_{j=1}^{i} \max_{t_j - \frac{1}{n} \leq s \leq t_j} M_j(s) \to \lambda^c(t) + \sum_{j=1}^{i} \alpha_j = \lambda(t).$$

In other words, $\lambda_n \in C$ and $\lambda_n \to \lambda$ pointwise almost surely.

We relax the condition that $\lambda(T) \leq C$. Assuming only that $\lambda(T) < \infty$ almost surely, we define for each $m \in \mathbb{N}$ the process $\lambda_m = m \wedge \lambda$. We have just proved that for each m we may construct a sequence $\{\lambda_{m,n}\}_{n=N}^{\infty}$ in C such that $\lambda_{m,n} \to \lambda_m$ pointwise almost surely as $n \to \infty$. Let d be the metric of pointwise convergence on $R[0,T]$ defined in Section 5. We may choose a subsequence $\{\lambda_{m_k}\}_{k=1}^{\infty}$ such that $\mathbb{P}\{d(\lambda_{m_k}, \lambda) \geq 1/k\} \leq 2^{-k}$ for all $k \in \mathbb{N}$, and then choose n_k such that $\mathbb{P}\{d(\lambda_{m_k,n_k}, \lambda_{m_k}) \geq 1/k\} \leq 2^{-k}$. Then

$$\mathbb{P}\left\{d(\lambda_{m_k,n_k}, \lambda) \geq \frac{2}{k}\right\} \leq \mathbb{P}\left\{d(\lambda_{m_k,n_k}, \lambda_{m_k}) \geq \frac{1}{k} \text{ or } d(\lambda_{m_k}, \lambda) \geq \frac{1}{k}\right\} \leq 2^{-k+1}.$$

The Borel-Cantelli Lemma implies $\mathbb{P}\{d(\lambda_{m_k,n_k}, \lambda) \geq 2/k \text{ infinitely often}\} = 0$, and hence $\lambda_{m_k,n_k} \to \lambda$ pointwise almost surely as $k \to \infty$.

In either case, whether $\lambda \in \mathcal{R}$ is bounded or not, there is a sequence $\{\lambda_n\}_{n=N}^{\infty}$ of processes in C such that $\lambda_n \to \lambda$ pointwise almost surely. For this sequence, we have from Fatou's Lemma and the lower semicontinuity of $g_*(S, \cdot)$ that

$$\begin{aligned}
\sup_{\eta \in C} \mathbb{E}\big[e^{-rT-\alpha\eta(T)} g(Se^{-\eta})\big] &\geq \liminf_{n\to\infty} \mathbb{E}\big[e^{-rT-\alpha\lambda_n(T)} g(Se^{-\lambda_n})\big] \\
&\geq \mathbb{E}\Big[\liminf_{n\to\infty} e^{-rT-\alpha\lambda_n(T)} g(Se^{-\lambda_n})\Big] \\
&\geq \mathbb{E}\big[e^{-rT-\alpha\lambda(T)} g_*(S, \lambda)\big].
\end{aligned}$$

Taking the supremum of the right-hand side over $\lambda \in \mathcal{R}$, we obtain (8.3). ◇

References

[1] Borodin, A.N. and Salminen, P., *Handbook of Brownian Motion – Facts and Formulae*, Birkhäuser, Basel, 1996.

[2] BOUCHARD, B. and TOUZI, N. (1999) Explicit solution of the multivariate super-replication problem undei transaction costs, preprint, Université Paris Dauphine.

[3] BERTSEKAS, D. P. and SHREVE, S. E., *Stochastic Optimal Control: The Discrete Time Case*, Academic Press, Orlando, FL, 1978. Reissued by Athena Scientific, Belmont, MA, 1996.

[4] BROADIE, M., CVITANIĆ, J. and SONER, H. M. (1998) On the cost of super-replication with transaction costs, *Rev. Financial Studies* **11**, 59–79.

[5] CVITANIĆ, J. and KARATZAS, I. (1993) Hedging contingent claims with constrained portfolios. *Annals Appl. Probability* **3**, 652–681.

[6] CVITANIĆ, J., PHAM, H. and TOUZI, N. (1999) Super-replication in stochastic volatility models under portfolio constraints, *J. Applied Probab.* **36**.

[7] CVITANIĆ, J., PHAM, H. and TOUZI, N. (1999) A closed-form solution for the problem of super-replication under transaction costs, *Finance and Stochastics* **3**, 35–54.

[8] EL KAROUI, N. and QUENEZ, M. C. (1995) Dynamic programming and the pricing of contingent claims in an incomplete market, *SIAM J. Control Optim.* **33**, 29–66.

[9] FREEDMAN, D. (1974) Optimal reward operator in special classes of dynamic programming problems, *Ann. Probability* **2**, 942–949.

[10] KARATZAS, K. and KOU, S. (1996) On the pricing of contingent claims under constraints, *Ann. Appl. Probab.* **6**, 321–369.

[11] KARATZAS, K. and KOU, S. (1998) Hedging American contingent claims with constrained portfolios, *Finance and Stochastics* **2**, 215–258.

[12] KARATZAS, I. and SHREVE, S., *Brownian Motion and Stochastic Calculus*, Springer-Verlag, New York, 1991.

[13] KARATZAS, I. and SHREVE, S., *Methods of Mathematical Finance*, Springer-Verlag, New York, 1998.

[14] SONER, H. M. and TOUZI, N. (1999) Super-replication under gamma constraint, preprint, Dept. Operations Research and Financial Engineering, Princeton University.

[15] TOUZI, N. (1999) Direct characterization of the value of super-replication under stochastic volatility and portfolio constraints, preprint, Université Paris Dauphine.

[16] VARADARAJAN, V. S. (1958) Weak convergence of measures on separable metric spaces, *Sankhya* **19**, 15–22.

[17] ZHANG, P., *Exotic Options: A Guide to Second Generation Options*, World Scientific, Singapore, 1997.