Interest Rate and Credit Models

2. Credit Spreads

Andrew Lesniewski

Baruch College
New York

Spring 2020

Outline

Credit risk

- One of the main themes of this course is *credit risk* and its modeling.
- Credit risk is the potential that a borrower or a counterparty on a transaction will fail to meet its obligations in accordance with agreed terms. This failure is referred to as a *default*.
- The exact definition of what constitutes an event of default (bankruptcy, debt restructuring, failure to pay, etc.) is governed by strict legal language, and is constantly ammended.
- Credit risk comes in several flavors:
  - (i) Consumer (eg. a consumer fails to make a payment due on a mortgage loan, car loan, credit card, or line of credit).
  - (ii) Corporate (eg. a company bond issuer is unable to make a payment on a coupon or principal payment when due).
  - (iii) Sovereign (eg. a government bond issuer is unable to make a payment on a coupon or principal payment when due).

Credit risk

- Like many other types of risk (FX risk, interest rate risk, prepaymant risk, ...), credit risk can be traded through a variety of credit sensitive instruments.
- These instruments fall into two broad categories:
  - (i) cash instruments (e.g. corporate coupon bonds, sovereign bonds, asset backed securities),
  - (ii) derivatives.

Credit sensitive instruments

- A basic credit sensitive instrument is a risky *coupon bond*.
  (i) Bond holder receives a coupon in excess of the risk free rate in return for assuming the credit risk of the issuing entity.
  (ii) The higher the default risk, the bigger the coupon.
- Consider first a riskless coupon bond with face value \$1 and annual coupon C.
  (i) A future cash flow at time T is discounted on the risk free interest rate curve with the discount factor $P(T)$.
  (ii) The price of the bond is the sum of the present values of all coupon payments and the final principal repayment:

$$\text{Price} = \sum_{j=1}^{n} P(T_j)C + P(T_n).$$

  (iii) The risk embedded in this bond is entirely interest rate risk (market risk).

Credit sensitive instruments

- This valuation formula has to be modified in the presence of credit risk:
  (i) The issuer may stop making coupon payments.
  (ii) The issuer may fail to repay the face value of the bond.
- In order to account for these events, we introduce the following concepts:
  (i) By $S(T)$ we denote the probability that the payment at time T will be made; we refer to it as the *survival probability*.
  (ii) The quantity $Q(T) = 1 - S(T)$ is the *default probability*; it is the probability that the default event will occur some time between now and T.
  (iii) By $R(T)$ we denote the fraction of the face value that the bond holder will receive upon default at time T; we refer to it as the *recovery value*.

Credit sensitive instruments

- The bond valuation formula for a credit sensitive bond has to be modified as follows:
  - (i) Each coupon payment at time T_j is multiplied by the survival probability.
  - (ii) The price of the bond is the sum of the present values of all coupon payments and the final principal repayment:

$$\text{Price} = \sum_{j=1}^{n} \big(S(T_j)P(T_j)C + \big(S(T_{j-1}) - S(T_j)\big)P(T_{j-1})R(T_{j-1})\big) + S(T_n)P(T_n).$$

- The risk has now two components: interest rate risk (market risk) and credit risk.
- Note that the natural discount factor for cash flows in a risky bond is he *risky discount factor*, $\mathcal{P}(T) = S(T)P(T)$:

$$\text{Price} = \sum_{j=1}^{n} \big(\mathcal{P}(T_j)C + \big(1 - S(T_j)/S(T_{j-1})\big)\mathcal{P}(T_{j-1})R(T_{j-1})\big) + \mathcal{P}(T_n).$$

Credit sensitive instruments

- The quantity $1 - S(T_j)/S(T_{j-1})$ is the *conditional* probability of default during the period $[T_{j-1}\,,\,T_j]$ provided that there was no default prior to T_{j-1}.
- The value of the risky discount factor takes into account the likelihood of default.
- If the discounting is done on a constant riskless rate, $P(T) = e^{-rT}$.
- Assume that $S(T) = e^{-\lambda T}$, with λ constant. We see then that $\mathcal{P}(T) = e^{-(r+\lambda)T}$, and so λ is the extra discounting to accommodate for the probability of default.
- λ is referred to as the *credit spread*.

Credit derivatives

- Credit derivatives are financial contracts that allow to transfer the *credit risk* of a *reference entity* from one counterparty to another.
- The counterparties on a credit derivative are a *protection seller* (the party that assumes the credit risk) and a *protection buyer*.
- Credit risk of the counterparties on credit derivatives (as well as on any other financial transaction) came to the forefront of interest after the spectacular failure of AIG Financial Products in 2008 followed up by the US government bailout.
- We will address the counterparty risk later in the course, for now assume no counterparty risk.
- A precise definition of a default event is largely a legal matter and it will not concern us here.

Credit derivatives

- Two main categories of credit derivatives are:
  (i) *Unfunded credit derivatives*. Under a contract of this type, the protection seller makes an initial payment that is used to settle any potential future credit events.
  (ii) *Funded credit derivatives*. These are bilateral contracts between two counterparties, each of which is responsible for making its contractual payments (i.e. payments of premiums and any cash or physical settlement amount) itself without recourse to other assets.
- Note that the protection buyer is exposed to the credit risk of the protection seller. This is an example of counterparty credit risk that we will discuss in detail later.

Credit derivatives

- Unfunded credit derivative products include the following products:
  (i) Single name credit default swap (CDS)
  (ii) Total return swap (TRS)
  (iii) CDS index products (CDX, iTraxx)
  (iv) Constant maturity credit default swap (CMCDS)
  (v) First to default credit default swap
  (vi) Portfolio credit default swap
  (vii) Credit default swap on asset backed securities (ABS CDS)
  (viii) Credit default swaption
  (ix) Recovery lock
  (x) Credit spread option
- Funded credit derivative products include the following products:
  (i) Credit-linked note (CLN)
  (ii) Synthetic collateralized debt obligation (SCDO)
  (iii) Constant proportion debt obligation (CPDO)
  (iv) Synthetic constant proportion portfolio insurance (Synthetic CPPI)

Credit modeling

- Default events are relatively infrequent.
- Historical default data are available for past events.
- Moody's Investor Services has an extensive database of corporate defaults, various companies collect data on consumer credit.
- The central questions of credit risk modeling are:
  (i) Probability of default
  (ii) Credit spread
  (iii) Recovery rate

Credit modeling

- There several different approaches to credit modeling:
  - (i) *Scoring*. Each entity is assigned a score. The scores can be produced on the basis of a mathematical model, purely judgmentally, or through a combination of both. Typically, credit scores are not associated with specific forecasts of default probabilities. Examples are Altman's Z-score (for corporations) and FICO scores (for consumers).
  - (ii) *Credit migration models*. Each entity is assigned a credit rating from "riskless" to "defaulted", and probabilities of transitions between these ratings are estimated. This approach is taken by credit rating agencies such Moody's, S&P, and Fitch.
  - (iii) *Structural models*. In this approach, the default probability is modeled through a Black-Scholes style stochastic model.
  - (iv) *Reduced form models*. This approach uses the conditional probability of default, called the *intensity* or *hazard rate*, as the starting point.

Credit migration matrix

- Credit rating agencies periodically review the credit quality of each issuer, and may change their rating.
- Such a change will likely impact the perception of credit worthiness of the name.
- Historical observations of such changes are summarized in the form of a *credit migration matrix*.

Credit migration matrix

- For example, here is a one-year credit migration matrix from Moody's data:

$\Pi =$

Grade	Aaa	Aa	A	Baa	Ba	B	D
Aaa	91.027%	6.998%	1.003%	0.650%	0.238%	0.059%	0.025%
Aa	7.003%	85.823%	5.997%	0.704%	0.266%	0.147%	0.060%
A	2.000%	10.865%	80.251%	6.159%	0.397%	0.238%	0.090%
Baa	0.299%	0.999%	3.798%	90.624%	3.680%	0.400%	0.200%
Ba	0.151%	0.902%	3.701%	7.002%	72.855%	12.889%	2.500%
B	0.007%	0.047%	0.217%	0.405%	8.898%	78.849%	11.576%
D	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	100.000%

- The probabilities in Moody's transition matrix are historical ("P-measure") probabilities.
- A matrix with risk neutral ("Q-measure") probabilities would typically lead to higher default probabilities.

Credit as a Markov chain

- We consider a *finite state* space $E = \{E_1, \ldots, E_K\}$, and a sequence of random variables X_t, with $X_t \in E$. Such a sequence is called a *discrete-time stochastic process*.
- A *Markov chain* is a discrete-time stochastic process such that

$$\mathsf{P}(X_{t+1} = e | X_t = e_t, X_{t-1} = e_{t-1}, \ldots, X_0 = e_0) = \mathsf{P}(X_{t+1} = e | X_t = t_n).$$

- The property above is called the *Markov property*.
- A Markov chain is called *homogeneous*, if the conditional probability $\mathsf{P}(X_{t+1} = e | X_t = e_t)$ is independent of the time t.
- We call the matrix

$$\Pi_{ij} = \mathsf{P}(X_{t+1} = E_j | X_t = E_i)$$

the *transition matrix*. For a homogeneous process, Π is independent of t.

Credit as a Markov chain

- The probability distribution of the initial state is denoted by $p_0(i)$ and is given by $p_0(i) = \mathrm{P}(X_0 = E_i)$.
- Using Bayes' rule and the Markov property we find that the t-step transition probability

$$\Pi(t)_{ij} = \mathrm{P}(X_{s+t} = E_j | X_s = E_i)$$

  is given by

$$\Pi(t) = \Pi^t.$$

- The probability distribution p_t at time t is given by

$$p_t^{\mathrm{T}} = p_0^{\mathrm{T}} \Pi^t.$$

Credit as a Markov chain

- The credit migration matrix is a starting point to a Markov chain model of credit ratings.
- Assume that there are K "live" (non-default) states, and a $(K+1)$-st default state D. The transition probabilities between the states are given by the entries of the credit migration matrix.
- Notice that D is an absorbing state: the probability of transitioning out of it is zero.
- By the Markov property, the t-year transition matrix is given by:

$$\Pi_t = \Pi^t,$$

  for integer t (of course, $\Pi^1 = \Pi$).
- The transition matrix does not have to be built out of historical data (it can be produced e.g. by a model).

Credit as a Markov chain

- For convenience, one may choose to work with a *continous-time Markov chains* X_t. We than have:

$$
\mathsf{P}(X_{t_{n+1}} = e | X_{t_n} = e_n, X_{t_{n-1}} = e_{n-1}, \ldots, X_{t_0} = e_0) = \mathsf{P}(X_{t_{n+1}} = e | X_{t_n} = e_n), .
$$

  for all choices of observation times $t_0 < \ldots < t_{n-1} < t_n < t_{n+1}$.
- In the continuous time framework, it is convenient to work with a so called *generator matrix* Λ. For a homogeneous Markov chain, it is defined by

$$
\begin{aligned}
\Pi_t &= e^{\Lambda t} \\
&\triangleq \sum_{k=0}^{\infty} \frac{(t\Lambda)^k}{k!} .
\end{aligned}
$$

- Notice that $\Pi_1 = e^{\Lambda} = \Pi$.
- Calibrating the generator matrix to historical data (or to a model) allows us to calculate transition probabilities for all time horizons.

Credit as a Markov chain

- For example, the generator matrix used to produce the credit migration matrix above is given by:

$\Lambda =$

Grade	Aaa	Aa	A	Baa	Ba	B	D
Aaa	-0.0971753	0.0788	0.0087	0.0065	0.0026	0.0004	0.000175
Aa	0.0788	-0.16071	0.072	0.0051	0.0029	0.00143	0.000482
A	0.0182	0.13035	-0.22666	0.0718	0.0031	0.0025	0.000705
Baa	0.0025	0.0083	0.0433	-0.10179	0.0452	0.001	0.001491
Ba	0.0009	0.0079	0.0463	0.0846	-0.32923	0.1711	0.018426
B	0	0	0	0	0.1182	-0.24746	0.129255
D	0	0	0	0	0	0	0

- Notice that the transition matrix satisfies the following ODE (forward Kolmogorov equation):

$$\frac{d\Pi_t}{dt} = \Lambda\Pi_t,$$

with $\Pi_0 = I$.

Credit as a Markov chain

- It is possible (and, sometimes, necessary) to extend the Markov chain model to time dependent generators Λ_t.
- The starting point is the ODE:

$$\frac{d\Pi_t}{dt} = \Lambda_t \Pi_t .$$

- In general (when the matrices Λ_t do not commute for different values of t), it is generally impossible to solve it in closed form.
- Numerical solution, satisfying $\Pi_0 = I$, is given by:

$$\begin{aligned} \Pi_t &= \mathrm{T} \exp \int_0^t \Lambda_t ds \\ &\triangleq \lim_{n \to \infty} \prod_{j=1}^{n} \left(I + \Lambda\big(\tfrac{(j-1)t}{n}\big) \tfrac{t}{n} \right). \end{aligned}$$

Girsanov's theorem

- Girsanov's theorem plays a key conceptual role in arbitrage free pricing theory, a fact that will be explained below.
- Girsanov's theorem is a culmination of efforts by a number of mathematicians studying the effect of "change of variables" in the measure P on the properties of martingales under that measure.
- We consider a Brownian motion $W(t)$, and the associated probability space $(\Omega, \mathscr{F}, \mathrm{P})$, where Ω is the sample space, $\mathscr{F} = (\mathscr{F}_t)_{t\geq 0}$, is the filtered information set, and P is the probability measure.
- By E (or E^P, when we want to be precise) we denote the expected value with respect to the measure P.

Girsanov's theorem

- We say that a measure Q on Ω is *absolutely continuous* with respect to P if there exists a positive function D (called the *Radon-Nikodym derivative*) such that

$$Q(A) = \int_A D(\omega)\, dP(\omega)\,, \tag{1}$$

for $A \subset \Omega$.

- Equivalently,

$$\frac{dQ}{dP}(\omega) = D(\omega)\,. \tag{2}$$

Girsanov's theorem

- In other words, the "volume element" $d\mathrm{Q}$ is always proportional to the "volume element" $d\mathrm{P}$, with the proportionality factor being a positive function throughout the probability space.
- In the context of a Brownian motion, we also require that the Radon-Nikodym derivative respect the filtration by time, i.e. the identity above holds if we condition on the information up to time t:

$$\frac{d\mathrm{Q}}{d\mathrm{P}}(\omega, t) = D(\omega, t) . \tag{3}$$

- Two probability measures Q and P are called *equivalent*, if Q is absolutely continuous with respect to P and P is absolutely continuous with respect to Q.

Girsanov's theorem

- Consider now a diffusion process:

$$dX(t) = \Delta(t, X(t))dt + C(t, X(t))dW(t)\,. \tag{4}$$

- A natural question arises: can we transform a diffusion process into a diffusion process with a different drift,

$$dX(t) = \widetilde{\Delta}(t, X(t))dt + C(t, X(t))d\widetilde{W}(t)\,. \tag{5}$$

by a change to an equivalent probability measure Q?

Girsanov's theorem

- In particular, can we make the new process a *martingale*?
- Recall that a process $X(t)$ is a martingale if $\mathsf{E}^{\mathsf{Q}}[|X(t)|] < \infty$, for all t, and

$$X(s) = \mathsf{E}^{\mathsf{Q}}[X(t)|\mathscr{F}_s], \tag{6}$$

  where $\mathsf{E}^{\mathsf{Q}}[\,\cdot\,|\mathscr{F}_s]$ denotes the conditional expected value.
- In other words, given all information up to time s, the expected value of future values of a martingale is $X(s)$.
- If the process $X(t)$ is a *martingale*, the diffusion above is driftless, i.e. $\widetilde{\Delta}(t, X(t)) = 0$.
- An affirmative answer to this question is provided by Girsanov's theorem.

Girsanov's theorem

- One might proceed heuristically as follows. Write

$$
\begin{aligned}
dX(t) &= \widetilde{\Delta}(t)\,dt + C(t)\left(\frac{\Delta(t)-\widetilde{\Delta}(t)}{C(t)}\,dt + dW(t)\right) \\
&= \widetilde{\Delta}(t)\,dt + C(t)\,d\widetilde{W}(t)\,,
\end{aligned}
\tag{7}
$$

where

$$
\begin{aligned}
\widetilde{W}(t) &= W(t) + \int_0^t \frac{\Delta(s)-\widetilde{\Delta}(s)}{C(s)}\,ds \\
&\equiv W(t) - \int_0^t \theta(s)\,ds.
\end{aligned}
\tag{8}
$$

- This looks like a new Brownian motion!

Girsanov's theorem

- Girsanov's theorem asserts that, under some technical assumptions on the drift and diffusion coefficients, $\widetilde{W}(t)$ is indeed a Brownian motion provided that the probability measure is modified appropriately.
- More precisely, define the stochastic process:

$$D(t) = \exp\left(\int_0^t \theta(s)\, dW(s) - \frac{1}{2}\int_0^t \theta(s)^2\, ds\right). \tag{9}$$

- Note that we have changed our notation: as always when dealing with stochastic processes, we have suppressed the argument ω in D, and made the dependence on t explicit.
- We now define the equivalent measure Q with

$$\frac{dQ}{dP}(t) = D(t). \tag{10}$$

Girsanov's theorem

- *Girsanov's theorem*. Assume that the following technical condition (*Novikov's condition*) holds:

$$\mathrm{E}^{\mathrm{P}}\left[\exp\left(\frac{1}{2}\int_0^t \theta(s)^2\, ds\right)\right] < \infty. \tag{11}$$

Then:

(i) The process $D(t)$ is a martingale under P. Furthermore, it satisfies the following stochastic differential equation:

$$dD(t) = \theta(t)\, D(t)\, dW(t)\,. \tag{12}$$

(ii) $\widetilde{W}(t)$ is a Wiener process under Q.

Girsanov's theorem

- We have stated Girsanov's theorem for a one-dimensional Brownian motion. This assumption is not essential and, using a bit of linear algebra, one can easily formulate a version of Girsanov's theorem for an arbitrary multidimensional Brownian motion.
- Consider a standard d-dimensional Brownian motion $W(t) = \big(W_1(t), \ldots, W_d(t)\big)$ (i.e. the components of $W(t)$ are uncorrelated), and let $X(t) = \big(X_1(t), \ldots, X_d(t)\big)$ be a d-dimensional diffusion process:

$$dX(t) = \Delta(t, X(t))dt + C(t, X(t))dW(t)\,. \tag{13}$$

- Define the stochastic process:

$$D(t) = \exp\Big(\int_0^t \theta(s)^{\mathrm{T}}\, dW(s) - \frac{1}{2}\int_0^t \theta(s)^{\mathrm{T}}\,\theta(s)\,ds\Big). \tag{14}$$

where

$$\theta(t) = C(t, X(t))^{-1}\big(\widetilde{\Delta}(t, X(t)) - \Delta(t, X(t))\big). \tag{15}$$

Girsanov's theorem

- *Girsanov's theorem (multidimensional version)*. Assume that the following technical condition (*Novikov's condition*) holds:

$$\mathrm{E}^{\mathrm{P}}\left[\exp\left(\frac{1}{2}\int_0^t \theta(s)^{\mathrm{T}}\,\theta(s)\,ds\right)\right] < \infty. \tag{16}$$

Then:

(i) The process $D(t)$ is a martingale under P. Furthermore, it satisfies the following stochastic differential equation:

$$dD(t) = \theta(t)^{\mathrm{T}}\,D(t)\,dW(t)\,. \tag{17}$$

(ii) $\widetilde{W}(t)$ is a Wiener process under Q.

Structural credit models

- Structural models of credit risk historically emerged first, they originated with Merton's paper [3] applying the Black-Scholes framework to the problem of default.
- The structural approach allows for relating the economic context, such as balance sheet information, to understand the definition of the event of default.
- We begin with the discussion of Merton's model.
- It is a single period (extremely simplified) stochastic model of a firm with time horizon T.

Merton's firm value model

- The capital structure of the firm consists only of
  (i) debt whose value is B, and
  (ii) equity whose value is S.
- The debt is a zero coupon bond with face value F which matures at T.
- The equity consists of non-dividend paying shares.
- The total value $V(t)$ of the firm at time t is thus

$$V(t) = B(t) + S(t).$$

- At time T, the firm is solvent if $V(T) \geq F$, and it is in default, if $V(T) < F$.

Merton's firm value model

- The debt holders' payoff at T is

$$\begin{aligned} B(T) &= \min(F, V(T)) \\ &= V(T) - (V(T) - F)^{+}, \end{aligned}$$

  where we have used the notation $x^{+} = \max(x, 0)$.
- The equity holders' payoff at T is

$$\begin{aligned} S(T) &= (V(T) - F)^{+} \\ &= V(T) - B(T). \end{aligned}$$

- We assume that the value of the firm follows a lognormal diffusion:

$$dV(t) = V(t)(\mu dt + \sigma_V dW(t)),$$

  where μ is the firm's rate of return, and σ_V is the volatility of the value process.

Merton's firm value model

- We assume that the pricing is done in the risk neutral measure Q associated with a riskless rate r.
- From the Black-Scholes theory we get

$$
\begin{aligned}
S(t,T) &= V(t)\,N(d_1) - Fe^{-r(T-t)}N(d_2),\\
B(t,T) &= Fe^{-r(T-t)}N(d_2) + V(t)\,N(-d_1).
\end{aligned}
$$

- Here $N(x)$ is the cumulative normal distribution function, and

$$
\begin{aligned}
d_1 &= \frac{\log\frac{V(t)}{F} + (r + \frac{1}{2}\sigma_V^2)(T-t)}{\sigma_V\sqrt{T-t}},\\
d_2 &= d_1 - \sigma_V\sqrt{T-t}.
\end{aligned}
$$

Merton's firm value model

- We can now calculate the credit metrics defined earlier within Merton's model:
  (i) Survival probability:

$$\begin{aligned} S(T) &= \mathsf{Q}(V(T) \geq F) \\ &= N(d_2). \end{aligned}$$

  (ii) Recovery rate:

$$R(t) = \frac{V(t)\,N(-d_1)}{Fe^{-r(T-t)}N(-d_2)}\,.$$

  (iii) Implied volatility for the stock process:

$$\begin{aligned} \sigma_S &= \frac{\partial S(t)}{\partial V(t)}\frac{V(t)}{S(t)}\,\sigma_V \\ &= N(d_1)\,\frac{V(t)}{S(t)}\,\sigma_V. \end{aligned}$$

Merton's firm value model

- Merton's model leads to an explicit expression for the credit spread.
- Recall that the credit spread s is defined by

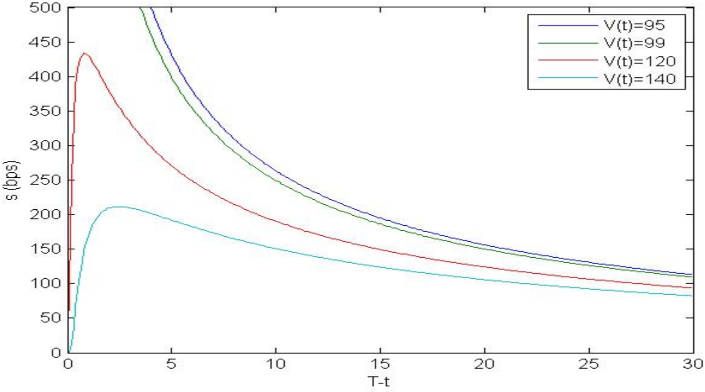
$$B(t) = Fe^{-(r+s)(T-t)}.$$

- Hence,

$$\begin{aligned} s &= -\frac{1}{T-t}\log\frac{B(t,T)}{F} - r \\ &= -\frac{1}{T-t}\log\Big(N(d_2) + e^{r(T-t)}\,\frac{V(t)}{F}\,N(-d_1)\Big). \end{aligned}$$

Merton's firm value model

- The graph below shows the credit spreads for different values $V(t)$ implied by Merton's model.



Black and Cox model

- One of the problems with Merton's original model is that the default can occur only at one time, namely the time horizon T.
- A natural extension, due to Black and Cox [2] assumes that there is a *safety covenant* in place, and the firm is obliged to reimburse its debt holders, if $V(t)$ falls below a predefined safety level.
- Such a safety level is given by a continuous barrier $H(t)$ and the firm defaults as soon as $V(t) \leq H(t)$.
- From the point of view of option pricing, the price of the bond is related to the price of a barrier option.
- To describe this, let τ be the default time of the firm. Then

$$\tau = \inf\{t > 0 : \; V(t) \leq H(t)\}.$$

Black and Cox model

- The choice of the barrier $H(t)$ should be financially meaningful.
- Observe that if $H(t) > F$, for all t, then the debt holders would always be protected from losses.
- Therefore, we require that at the very least $H(T) \leq F$.
- A natural choice is

$$H(t) = H_0 e^{at},$$

where H_0 is selected so that $H_0 < Fe^{-aT}$.

Black and Cox model

- As before, we assume that $V(t)$ follows the lognormal process

$$dV(t) = V(t)\left(\mu dt + \sigma_V dW(t)\right),$$

and so, under the risk neutral measure Q,

$$V(t) = V_0 \exp\left(\sigma_V W(t) + \left(r - \frac{\sigma_V^2}{2}\right)t\right).$$

- As a consequence,

$$\begin{aligned}\left\{V(t) \leq H(t)\right\} &= \left\{\log V(t) \leq \log H(t)\right\} \\ &= \left\{\sigma_V W(t) + (r - \sigma_V^2/2 - a)t \leq \log \frac{H_0}{V_0}\right\}.\end{aligned}$$

Black and Cox model

- Consequently,

$$\begin{aligned}\mathrm{Q}(\tau \leq t) &= \mathrm{Q}\Big(\min_{s\leq t} \frac{V(s)}{H(s)} \leq 1\Big) \\ &= \mathrm{Q}\Big(\min_{s\leq t} \left(W(s) + bs\right) \leq \frac{1}{\sigma_V} \log \frac{H_0}{V_0}\Big),\end{aligned}$$

where $b = (r - \sigma_V^2/2 - a)/\sigma_V$.

- From the mathematical point of view, this is the classic problem of *the first passage time for the Brownian motion with drift*.
- We will digress now to briefly present the key ideas of the solution to this problem.

Black and Cox model

- Consider the following stochastic processes:

$$M(t) = \max_{s \leq t} (W(s) + bs),$$
$$m(t) = \min_{s \leq t} (W(s) + bs).$$

- $M(t)$ is referred to as the *running maximum*, while $m(t)$ is referred to as the *running minimum* of the Brownian motion with a drift.
- We are concerned with determining the probability distributions of $M(t)$ and $m(t)$.
- To this end, we define the following function:

$$\Phi(\alpha, \beta, b, t) = N\big(\tfrac{\alpha - bt}{\sqrt{t}}\big) - e^{2b\beta} N\big(\tfrac{\alpha - 2\beta - bt}{\sqrt{t}}\big),$$

for $\alpha < \beta$.

Black and Cox model

- The probability distributions of $M(t)$ and $m(t)$ can be expressed in terms of the function Φ as follows.
- For any α and $\beta \geq 0$,
  - (i) $\mathrm{P}\big(M(t) \leq \beta,\ W(t) + bt \leq \alpha\big) = \Phi(\min(\alpha, \beta), \beta, b, t)$,
  - (ii) $\mathrm{P}\big(m(t) \geq -\beta,\ W(t) + bt \geq \alpha\big) = \Phi(\min(-\alpha, \beta), \beta, -b, t)$,
  - (iii) $\mathrm{P}\big(M(t) \leq \beta\big) = \Phi(\beta, \beta, b, t)$,
  - (iv) $\mathrm{P}\big(m(t) \geq -\beta\big) = \Phi(\beta, \beta, -b, t)$.
- Note first that (i) implies the other three formulas by changing the sign of $W(t)$ and noting that

$$\min_{s \leq t}\big(-W(s) + bs\big) = -\max_{s \leq t}\big(W(s) - bs\big).$$

- Also, (iii) follows from (i) by letting $\alpha \to \infty$, and (iv) follows from (ii) by letting $\alpha \to -\infty$.
- We thus have to prove (i) only.

Black and Cox model

- The proof of (i) proceeds in two steps.
- Consider first $b = 0$, and $\alpha \leq \beta$. We have:

$$\begin{aligned} \mathrm{P}\big(M(t) \leq \beta, W(t) \leq \alpha\big) &= \mathrm{P}\big(W(t) \leq \alpha\big) - \mathrm{P}\big(M(t) \geq \beta, W(t) \leq \alpha\big) \\ &= \mathrm{P}\big(W(t) \leq \alpha\big) - \mathrm{P}\big(M(t) \geq \beta, W(t) \geq 2\beta - \alpha\big), \end{aligned}$$

  where we have used the reflection principle for Brownian motion.
- Therefore,

$$\begin{aligned} \mathrm{P}\big(M(t) \leq \beta, W(t) \leq \alpha\big) &= \mathrm{P}\big(W(t) \leq \alpha\big) - \mathrm{P}\big(W(t) \geq 2\beta - \alpha\big) \\ &= N(\alpha/\sqrt{t}) - N((\alpha - 2\beta)/\sqrt{t}) \\ &= \Phi(\min(\alpha, \beta), \beta, b, t). \end{aligned}$$

Black and Cox model

- The case of $b \neq 0$ is now a consequence of Girsanov's theorem.
- Let P denote the original measure, and let Q denote the martingale measure obtained after the drift has been removed. The Radon-Nikodym derivative is

$$\begin{aligned} D(t) &= \frac{d\mathrm{Q}}{d\mathrm{P}}(t) \\ &= e^{-bW(t) - \frac{1}{2}b^2 t}. \end{aligned}$$

- Therefore,

$$\begin{aligned} &\mathrm{E}^{\mathrm{P}}\left[1_{M(t) \leq \beta} 1_{W(t)+bt \leq \alpha}\right] \\ &\qquad = \mathrm{E}^{\mathrm{Q}}\left[\frac{d\mathrm{P}}{d\mathrm{Q}}(t) 1_{M(t) \leq \beta} 1_{W(t)+bt \leq \alpha}\right] \\ &\qquad = \mathrm{E}^{\mathrm{Q}}\left[e^{b(\widetilde{W}(t)-bt)+\frac{1}{2}b^2 t} 1_{M(t) \leq \beta} 1_{\widetilde{W}(t) \leq \alpha}\right]. \end{aligned}$$

Black and Cox model

- Therefore,

$$\begin{aligned}
&\mathrm{E}^{\mathrm{P}}\left[1_{M(t)\leq\beta}1_{W(t)+bt\leq\alpha}\right] \\
&\qquad = \int_{-\infty}^{\alpha} e^{b(y+bt)-\frac{1}{2}b^2t}\int_{y^+}^{\beta}\frac{\partial^2}{\partial x\partial y}\Phi(y,x,0,t)dx\,dy \\
&\qquad = \frac{1}{\sqrt{t}}\int_{-\infty}^{\alpha} e^{b(y+bt)-\frac{1}{2}b^2t}\big(n(-|y|/\sqrt{t})-n((y-2\beta)/\sqrt{t})\big)dx\,dy,
\end{aligned}$$

  where $n(x)=(2\pi)^{-1/2}\exp(-x^2/2)$.

- Carrying out the integration we obtain formula (i).

Black and Cox model

- Let us now go back to the Black and Cox model. We have

$$
\begin{aligned}
\mathsf{Q}\Big(\min_{s\leq t}\ (W(s)+bs)\leq d\Big) &= 1-\Phi(-d,-d,-b,t)\\
&= 1-N\big(\tfrac{-d+bt}{\sqrt{t}}\big)+e^{2bd}N\big(\tfrac{d+bt}{\sqrt{t}}\big),
\end{aligned}
$$

  where $d=\frac{1}{\sigma_V}\log\frac{H_0}{V_0}$.
- Using the fact that $1-N(-x)=N(x)$, we see that the risk neutral probability of default is given by

$$
\mathsf{Q}(\tau<t)=N\big(\tfrac{d-bt}{\sqrt{t}}\big)+e^{2bd}\,N\big(\tfrac{d+bt}{\sqrt{t}}\big).
$$

Black and Cox model

- The equity holders' payoff at maturity T is given by

$$(V(T) - F)^{+} 1_{\tau \geq T} = (V_0 e^{aT} e^{\sigma_V (W(T)+bT)} - F)^{+} 1_{m(T) \geq d}.$$

- This is the payoff of a down and out call option.
- Its expected value gives the equity value of the company.
- As a consequence, the equity value in the Black and Cox model is smaller than the equity value in Merton's model.

Black and Cox model

- In the event of default, the debt holders' payoff for is $V(\tau) = H(\tau)$.
- The recovery value can be computed by integrating the discounted $H(s)$ through the time of default τ.
- The recovery value is thus

$$B^{\mathrm{rec}}(t, T) = -\int_t^T e^{-r(s-t)} H(s) \frac{\partial}{\partial s} \Phi(-d(t), -d(t), -b, s-t) ds,$$

where $d(t) = \frac{1}{\sigma_V} \log \frac{H(t)}{V(t)}$.

Black and Cox model

- This integral can be calculated explicitly.
- The value of the bond at time t prior to default is the sum of the payment at maturity (B^{mat}) and the recovery value (B^{rec}).
- The payment at maturity term can be written as the risk neutral price of a difference of two barrier call options:

$$\begin{aligned} B^{\text{mat}}(t,T) &= \mathrm{E}^{\mathrm{Q}}\left[e^{-r(T-t)}\left(V(T)-(V(T)-F)^{+}\right)1_{\tau>T}\,|\,\mathscr{F}_t\right] \\ &= e^{-r(T-t)}F\Phi(-d(t),-d(t),-b,T-t). \end{aligned}$$

Reduced form models

- In a structural model of credit, a default event is an *endogenous* feature of the model and it results as a consequence of the (stochastic) value of the entity falling below a certain threshold.
- In order to take all factors properly into account, such a description requires a detailed knowledge of the capital structure of the entity, and lead to complicated computational problems.
- In contrast, in the *reduced form* approach, a default event is an *exogenous* occurrence, which is governed exclusively by the assumed probability distribution of a random time τ, namely the *time to default*.
- While structural models offer a better intuitive insight into the nature of the credit of the entity, reduce form models are generally more transparent and numerically tractable.
- We begin by reviewing the basics of the theory of counting processes.

Reduced form models

- Credit events are examples of financial processes with discrete outcomes (unlike, say, stock prices which are usually treated as continuous processes).
- The basic probabilistic set up to model event risk is based on the concept of a random time τ.
- The stochastic process $1_{\tau \leq t}$ defines a filtration $(\mathscr{G}_t)_{t \geq 0}$, which is referred to as the *information set*.
- Of key importance to financial event modeling are *counting processes*.

Counting processes

- An stochastic process $N(t)$ is called a counting process if:
  (i) $N(t)$ is defined on a probability space$(\Omega, (\mathscr{G}_t)_{t\geq 0}, \mathrm{P})$.
  (ii) $N(t)$ is integer valued, non-negative, with $N(0) = 0$.
  (iii) Sample paths of $N(t)$ are right continuous, piecewise constant, may have jumps of size 1.
- Jump times $\tau_1 < \tau_2 < \ldots$ are stopping times at which the process $N(t)$ increases by 1.
- The *intensity* $\lambda(t)$ of $N(t)$ is defined by

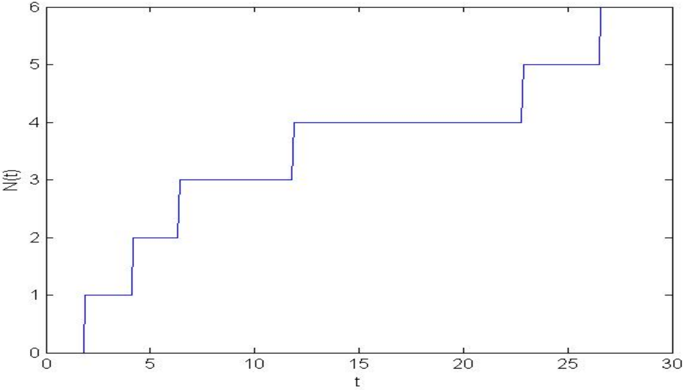
$$\lambda(t)\,dt = \mathrm{P}(N(t+dt) - N(t) = 1 \,|\, \mathscr{G}_{t\_})$$

- Alternatively,

$$\lambda(t)\,dt = \mathrm{E}[N(t+dt) - N(t) \,|\, \mathscr{G}_{t\_}\,.]$$

Counting processes

The graph below shows a sample path of a counting process.



Counting processes

- The simplest counting process is given by $N(t) \in \{0, 1\}$ and constant intensity λ. In this case,

$$
\begin{aligned}
\mathsf{P}(N(t) = 0) &= \mathsf{P}(\tau \geq t) \\
&= e^{-\lambda t}.
\end{aligned}
$$

  This probability distribution is called the exponential distribution.
- Indeed,

$$
\begin{aligned}
\mathsf{P}(\tau \geq t) &= \lim_{n \to \infty} \prod_{j=0}^{n} \Big(1 - \mathsf{P}\Big(\frac{j}{n}\, t \leq \tau \leq \frac{j+1}{n}\, t\Big)\Big) \\
&= \lim_{n \to \infty} \prod_{j=0}^{n} \Big(1 - \lambda \frac{t}{n}\Big) \\
&= \Big(1 - \lambda \frac{t}{n}\Big)^{n} \\
&= e^{-\lambda t}.
\end{aligned}
$$

Poisson process

- Another important example of a counting process is a *Poisson process*.
- A Poisson process is defined by the two requirements:
  (i) The probability that given $N(t) = n$, $N(t + dt) = n + 1$ is λdt,
  (ii) The probability that given $N(t) = n$, $N(t + dt) > n + 1$ is 0.
- Here, $\lambda > 0$ is the intensity of the Poisson process.
- Notice that the Poisson process has independent increments and is Markovian.
- We shall show that

$$\mathrm{P}(N(t) = n) = \frac{(\lambda t)^n}{n!}\, e^{-\lambda t}.$$

Poisson process

- Indeed, using $N(0) = 0$,

$$
\begin{aligned}
\mathrm{P}(N(t) = n) &= \lim_{N\to\infty} \binom{N}{n} \Big(\lambda\,\frac{t}{N}\Big)^n \Big(1 - \lambda\,\frac{t}{N}\Big)^{N-n} \\
&= \lim_{N\to\infty} \frac{N^N e^{-N}\sqrt{2\pi N}}{n!(N-n)^{N-n} e^{-N+n}\sqrt{2\pi(N-n)}} \Big(\lambda\,\frac{t}{N}\Big)^n \Big(1 - \lambda\,\frac{t}{N}\Big)^{N-n} \\
&= \frac{(\lambda t)^n}{n!} \lim_{N\to\infty} \frac{N^{N-n+1/2}}{(N-n)^{N-n+1/2} e^n} \Big(1 - \lambda\,\frac{t}{N}\Big)^{N-n} \\
&= \frac{(\lambda t)^n}{n!} \lim_{N\to\infty} \frac{1}{(1-\frac{n}{N})^{N-n+1/2} e^n} \Big(1 - \lambda\,\frac{t}{N}\Big)^{N-n} \\
&= \frac{(\lambda t)^n}{n!} \lim_{N\to\infty} \frac{1}{(1-\frac{n}{N})^{N} e^n} \Big(1 - \lambda\,\frac{t}{N}\Big)^{N} \\
&= \frac{(\lambda t)^n}{n!}\, e^{-\lambda t},
\end{aligned}
$$

where we have used Stirling's approximation $n! \approx (n/e)^n \sqrt{2\pi n}$.

Poisson process

- Furthermore, we have
$$E[N(t)] = \lambda t,$$
and
$$\mathrm{Var}[N(t)] = \lambda t.$$
- An important property of the Poisson process is that it is memoryless, meaning that
$$P(N(t+h) = n+m \,|\, N(t) = n) = P(N(h) = m).$$

Poisson process

- Let us consider the time of the first jump (or first arrival) τ_1 in a Poisson process.
- This is simply our first example of a counting process, and so its survival probability is given by the exponential distribution:

$$\mathrm{P}(\tau_1 > t) = e^{-\lambda t},$$

with the first moment

$$\mathrm{E}[\tau_1] = \frac{1}{\lambda}\,.$$

- The density of τ_1 can be found by differentiation of the cumulative distribution,

$$\mathrm{P}(\tau_1 \in [t, t + dt]) = e^{-\lambda t}\lambda dt.$$

- The distributions of all interarrival times $\tau_2 - \tau_1, \tau_3 - \tau_2, \ldots$, are all identical to the distribution of τ_1 (due to the memoryless property).

Inhomogeneous Poisson process

- Another important example of a counting process is an *inhomogeneous Poisson process*.
- It is a generalization of the Poisson procces, in which the intensity is a deterministic function of time $\lambda(t)$.
- The calculations involving inhomogeneous Poisson processes follow the lines of the homogeneous Poisson processes. It is easy to see that all the results stated above for a homogeneous Poisson process hold also for an inhomogeneous Poisson process with λ replaced by the average $\frac{1}{t}\int_0^t \lambda(s)\,ds$.
- In particular, the probability distribution is given by

$$\mathrm{P}(N(t)=n)=\frac{\left(\int_0^t \lambda(s)\,ds\right)^n}{n!}\,e^{-\int_0^t \lambda(s)ds}.$$

Inhomogeneous Poisson process

- Furthermore,

$$\mathrm{E}[N(t)] = \int_0^t \lambda(s)\, ds,$$
$$\mathrm{Var}[N(t)] = \int_0^t \lambda(s)\, ds$$

- For the properties of the first arrival time we have:

$$\mathrm{P}(\tau_1 > t) = e^{-\int_0^t \lambda(s) ds},$$
$$\mathrm{P}(\tau_1 \in [t, t+dt]) = e^{-\int_0^t \lambda(s) ds} \lambda(t)\, dt.$$

- The integral $A(t) = \int_0^t \lambda(s)\, ds$ is called the *compensator* of the process $N(t)$.

Cox processes

- A general class of counting processes are *Cox processes*.
- A Cox process $N(t)$ is associated with an information set of the form $\mathscr{G}_t \vee \mathscr{F}_t$, $t > 0$, where $\mathscr{G}_t$ is the information set pertaining to jump times, and where $\mathscr{F}_t$ is the "background" information set.
- One can think of $\mathscr{F}_t$ as containing the background economic information.
- A point process $N(t)$ is a Cox process if, conditioned on $\mathscr{F}_t$, $N(t)$ is an inhomogeneous Poisson process with a deterministic intensity $\lambda(s)$, $0 \leq s \leq t$.

Cox processes

- Intuitively, a Cox process can be thought of as an inhomogeneous Poisson process whose intensity is stochastic itself.
- We can think about it in the following way: We draw an intensity path first, and then - conditional on this path - we use the properties the inhomogeneous Poisson process.
- The compensator of a Cox process is the integrated intensity process $A(t) = \int_0^t \lambda(s)\, ds$.
- In particular,

$$\mathrm{P}\big(N(t) = n \,|\, \{\lambda(s)\}_{0\leq s\leq t}\big) = \frac{A(t)^n}{n!}\, e^{-A(t)},$$

and, taking expectation over all paths $\{\lambda(s)\}$, we get the absolute probability distribution:

$$\mathrm{P}(N(t) = n) = \frac{1}{n!}\,\mathrm{E}\big[A(t)^n\, e^{-A(t)}\big].$$

Cox processes

- As another example, let us find $\mathrm{P}(\tau_1 > t)$. We have

$$\mathrm{P}\big(\tau_1 > t \,|\, \{\lambda(s)\}_{0\leq s\leq t}\big) = e^{-A(t)},$$

and by taking expectation over all paths $\{\lambda(s)\}$, we get the absolute probability:

$$\mathrm{P}(\tau_1 > t) = \mathrm{E}\big[e^{-A(t)}\big].$$

- Finally, the absolute expectation and variance of a Cox process are given by

$$\begin{aligned}\mathrm{E}[N(t)] &= \mathrm{E}[A(t)],\\ \mathrm{Var}[N(t)] &= \mathrm{E}[A(t)],\end{aligned}$$

respectively.

Zero coupons and survival probabilities

- Note that there is a close analogy between the intensity $\lambda(t)$ in a Cox model and the short rate $r(t)$ in an interest rate model.
- Let $S(t, T)$ denote the time t survival probability $S(t, T) = \mathrm{P}(\tau_1 > T)$, and let $P(t, T)$ be the time t discount bond maturing at T. Then,

$$S(0, T) = \mathrm{E}\left[e^{-\int_0^T \lambda(s)ds}\right],$$
$$P(0, T) = \mathrm{E}\left[e^{-\int_0^T r(s)ds}\right].$$

- More generally, at time $t \geq 0$,

$$S(t, T) = 1_{\tau_1 > t}\mathrm{E}\left[e^{-\int_t^T \lambda(s)ds}\right],$$
$$P(t, T) = \mathrm{E}\left[e^{-\int_t^T r(s)ds}\right].$$

Note the presence of the process $1_{\tau_1 > t}$ in the first formula: the time t survival probability for T is zero, if the event has occurred prior to t.

Risky zero coupon

- Consider now a risky zero coupon bond with a zero recovery rate.
- Assuming that the credit intensity follows a Cox process, its price is given by:

$$
\begin{aligned}
\mathcal{P}(0,T) &= \mathsf{E}\big[e^{-\int_0^T r(s)ds}1_{\tau_1>T}\big]\\
&= \mathsf{E}\big[\mathsf{E}\big[e^{-\int_0^T r(s)ds}1_{\tau_1>T} \mid \{\lambda(s)\}_{0\le s\le T}\big]\big]\\
&= \mathsf{E}\big[\mathsf{E}\big[e^{-\int_0^T r(s)ds}e^{-\int_0^T \lambda(s)ds} \mid \{\lambda(s)\}_{0\le s\le T}\big]\big]\\
&= \mathsf{E}\big[e^{-\int_0^T (r(s)+\lambda(s))ds}\big].
\end{aligned}
$$

- Notice that the presence of event risk (i.e. credit) amounts to increasing the discounting rate by the intensity of default. We have already noted this earlier in these notes.

Risky zero coupon

- In particular, if the default probability is *independent* of the rates probability, then,

$$\begin{aligned}\mathcal{P}(0, T) &= \mathrm{E}[e^{-\int_0^T r(s)ds}]\mathrm{E}[1_{\tau_1 > T}] \\ &= P(0, T)S(0, T).\end{aligned}$$

  We have thus recovered the formula for the risky coupon bond discussed earlier as a special case of our general framework.

- For any time $t \geq 0$, we have the following expression for the risky zero coupon bond:

$$\mathcal{P}(t, T) = 1_{\tau_1 > t}\mathrm{E}\big[e^{-\int_t^T (r(s)+\lambda(s))ds}\big].$$

  Note the time t value of the bond is zero, if the default happens prior to t.

Gaussian structural model

- As an explicit example, we consider the model in which both the rate and intensity are Gaussian processes:

$$dr(t) = \mu(t)\,dt + \sigma_r dW(t),$$
$$d\lambda(t) = \theta(t)\,dt + \sigma_\lambda dZ(t),$$

with correlated Brownian motions:

$$dW(t)\,dZ(t) = \rho dt.$$

- Our goal is to compute $\mathcal{P}(0, T)$.

Gaussian structural model

- Note that the pathwise solution to this model is explicitly given by:

$$
\begin{aligned}
r(t) &= r_0 + \int_0^t \mu(s)\, ds + \sigma_r W(t)\,, \\
\lambda(t) &= \lambda_0 + \int_0^t \theta(s)\, ds + \sigma_\lambda Z(t)\,.
\end{aligned}
$$

- These are Gaussian variables with means

$$
\begin{aligned}
\mathrm{E}[r(t)] &= r_0 + \int_0^t \mu(s)\, ds, \\
\mathrm{E}[\lambda(t)] &= \lambda_0 + \int_0^t \theta(s)\, ds,
\end{aligned}
$$

and covariance matrix (for $t \leq u$)

$$
\mathrm{Cov}(r(t), \lambda(u)) = \begin{pmatrix} \sigma_r^2 t & \rho\sigma_r\sigma_\lambda t \\ \rho\sigma_r\sigma_\lambda t & \sigma_\lambda^2 u \end{pmatrix}.
$$

Gaussian structural model

- Hence, the mean and variance of the Gaussian variable $X \triangleq \int_0^T (r(t) + \lambda(t))dt$ are given by

$$\mathrm{E}[X] = (r_0 + \lambda_0)T + \int_0^T (T - t)(\mu(t) + \theta(t))dt,$$

and

$$\begin{aligned}\mathrm{Var}[X] &= 2(\sigma_r^2 + 2\rho\sigma_r\sigma_\lambda + \sigma_\lambda^2) \iint_{0\le t\le u\le T} t\,dt\,du \\ &= (\sigma_r^2 + 2\rho\sigma_r\sigma_\lambda + \sigma_\lambda^2)\,\frac{T^3}{3},\end{aligned}$$

respectively.

Gaussian structural model

- Using the fact that

$$\mathrm{E}[\exp(-X)] = \exp\Big(-\mathrm{E}[X] + \frac{1}{2}\mathrm{Var}[X]\Big),$$

  we find

$$\mathcal{P}(0, T) = P(0, T)S(0, T)e^{\rho\sigma_r\sigma_\lambda T^3/3}.$$

- The Gaussian model implies thus that positive correlation between rates and credit implies a higher zero coupon bond price.
- Later in the course we will discuss other examples of Cox processes that are used in credit modeling.

References

- Bielecki, T. R., and Rutkowski, M.: *Credit Risk: Modeling, Valuation and Hedging*, Springer (2002).
- Black, F., and Cox, J. C.: Valuing corporate securities: Some effects of bond indenture provisions, *J. Finance*, **31**, 351 - 367 (1976).
- Merton, R.: On the pricing of corporate debt: the risk structure of interest rates, *J. Finance*, **29**, 449 - 470 (1974).
- O'Kane, D.: *Modeling Single-Name and Multi-Name Credit Derivatives*, Wiley (2008).