Chapter 16

Heath-Jarrow-Morton

16 The HJM framework

Up to this point we have studied interest models where the short rate r is the only explanatory variable. The main advantages with such models are as follows.

1. Specifying r as the solution of an SDE allows us to use Markov process theory, so we may work within a PDE framework.
2. In particular it is often possible to obtain analytical formulas for bond prices and derivatives.

The main drawbacks of short rate models are as follows.

1. From an economic point of view it seems unreasonable to assume that the entire money market is governed by only one explanatory variable.
2. It is hard to obtain a realistic volatility structure for the forward rates without introducing a very complicated short rate model.
3. As the short rate model becomes more realistic, the inversion of the yield curve described above becomes increasingly more difficult.

These, and other considerations, have led various authors to propose models that use more than one state variable. One obvious idea would, for example, be to present an a priori model for the short rate as well as for some long rate, and one could of course also model one or several intermediary interest rates. The method proposed by Heath-Jarrow-Morton is at the far end of this spectrum - they choose the entire forward rate curve as their (infinite dimensional) state variable.

We will now specify the HJM framework, and we start by specifying everything under a given objective measure P.

Assumption: We assume that, for every fixed $T > 0$, the forward rate $f(t, T)$ has a stochastic differential which under the objective measure P is given by

$$\begin{cases} df(t,T) = \alpha(t,T)dt + \sigma(t,T)dW^P(t,T), \\ f(0,T) = f^*(0,T) \end{cases}$$

where W^P is a (d-dimensional) P-Wiener process whereas $\alpha(t, T)$ and $\sigma(t, T)$ is adapted processes.

Note that conceptually this is one stochastic differential in the t-variable for each fixed choice of T. The index T thus only serves as a "mark" or "parameter" in order to indicate which maturity we are looking at. Also note that we use the observed forward rated curve $\{f^*(0, T); T \geq 0\}$ as the initial condition. This will automatically give us a perfect fit between observed and theoretical bond prices at $t = 0$, thus relieving us of the task of inverting the yield curve.

It is important to observe that the HJM approach is not a proposal for a specific model, like for example, the Vasicek model. It is instead a framework to be used for analysing interest rates models. Every short rate models can be equivalently formulated in forward rate terms, and for every forward rate model, the arbitrage free price of a contingent T-claim X will still be given by the pricing formula

$$\Pi[0, X] = E^Q\left[\exp\left\{-\int_0^T r(s)ds\right\} \cdot X\right]$$

where the spot rate is as usual given by $r(t) = f(t, t)$.

Suppose now that we have specified α, σ and $\{f^*(0, T\}; T \geq 0\}$. Then we have specified the entire forward rate structure and thus, by the relation

$$p(t,T) = \exp\left\{-\int_t^T f(t,s)ds\right\}$$

we have in fact specified the entire term structure $\{p(t, T); T > 0, 0 \leq t \leq T\}$. Since we have d sources of randomness (one for every Wiener process), and an infinite number of traded assets (one bond for each maturity T), we run a clear risk of having introduced arbitrage possibilities into the bond market. The first question we pose is thus very natural: How must the processes α and σ be related in order that the induced system of bond prices admits no arbitrage possibilities? The answer is given by the HJM drift condition that relates α to σ.

We start, as usually with the assumptions that there exists a local risk-free security with the price process B given by

$$\begin{cases} dB(t) = r(t)B(t)dt \\ B(0) = 1 \end{cases}$$

where the spot rate is given by $r(t) = f(t, t)$.

We also assume that exist an equivalent probability measure $Q \sim P$ such as each Z^T-process is a Q-martingale on [0, T], where the discounted bond prices Z^T is defined as

$$Z^T(t) = \frac{p(t,T)}{B(t)}$$

We also know that the dynamic of the forward rates imply the following dynamic for the bond prices:

$$dp(t,T) = p(t,T)\{r(t) + b(t,T)\}dt + p(t,T)a(t,T)dW(t)$$

where

$$df(t,T) = \alpha(t,T)dt + \sigma(t,T)dW(t)$$

and $a(t, T)$ and $b(t, T)$ are given by:

$$\begin{cases} a(t,T) = -\int_t^T \sigma(t,u)du \\ b(t,T) = -\int_t^T \alpha(t,u)du + \frac{1}{2}a(t,T)^2 \end{cases}$$

The dynamic of Z^T is then given by:

$$dZ(t,T) = b(t,T)Z(t,T)dt + a(t,T)Z(t,T)dW(t)$$

Therefore we have to ask the question, if there exists a Girsanov transformation that for all T at the same time removes the drift.

$$dQ = L(T)dP \qquad \text{on } \mathcal{F}_T$$

where

$$\begin{cases} dL(t) = g(t)L(t)dW(t) \\ L(0) = 1 \end{cases}$$

for some process $g(t)$. From Girsanov theorem we get

$$dW(t) = g(t)dt + dV(t)$$

where *V* is a *Q*-Wiener process. We then have

$$\begin{aligned} dZ(t,T) &= \frac{\partial Z(t,T)}{\partial p(t,T)}dp(t,T) + \frac{\partial Z(t,T)}{\partial B(t)}dB(t) = \frac{1}{B(t)}dp(t,T) - \frac{p(t,T)}{B^2(t)}dB(t) \\ &= Z(t,T)\{r(t)+b(t,T)\}dt + Z(t,T)a(t,T)dW(t) - Z(t,T)r(t)dt \\ &= b(t,T)Z(t,T)dt + a(t,T)Z(t,T)dW(t) \\ &= \{b(t,T) + g(t)a(t,T)\}Z(t,T)dt + a(t,T)Z(t,T)dV(t) \end{aligned}$$

We must have

$$g(t,T) = -\frac{b(t,T)}{a(t,T)}$$

This Girsanov kernel, $g(t, T)$ holds for a given T. Therefore a martingale measure Q^T will be generated such as Z^T becomes martingale. Remark! This depends on our choice of T, so there is no guarantee that Q^S for $S \neq T$ is Q^T-martingale. If there exist a Girsanov transformation that make all Z^T-processes martingale at the same time, then $g(t, T)$ must be independent of the choice of T.

Theorem 16.1. *The following statements are equivalent*

- *There exists a measure Q^T that makes all Z^T processes martingales.*
- *For all T and S we have*

$$\frac{b(t,T)}{a(t,T)} = \frac{b(t,S)}{a(t,S)}$$

for all $t \leq min(T, S)$.

- *The Girsanov kernel $g(t, T)$ is independent of T.*
- *For each S and T we have*

$$\alpha(t,T) = -\sigma(t,T)\left\{g(t,S) - \int_t^T \sigma(t,s)ds\right\}$$

Proof: We only prove the last statement since the others are obvious. We have

$$g(t,S) = -\frac{b(t,T)}{a(t,T)}$$

which in detail gives

$$g(t,S)a(t,T) = -b(t,T)$$

$$g(t,S)\int_t^T \sigma(t,u)du = -\int_t^T \alpha(t,u)du + \frac{1}{2}a(t,T)^2$$

$$g(t,S)\int_t^T \sigma(t,s)ds = -\int_t^T \alpha(t,s)ds + \frac{1}{2}\left\{\int_t^T \sigma(t,s)ds\right\}^2$$

so

$$\int_t^T \alpha(t,s)ds = -g(t,S)\int_t^T \sigma(t,s)ds + \frac{1}{2}\left\{\int_t^T \sigma(t,s)ds\right\}^2$$

Take the derivative with respect to T and we are finished.

Theorem 16.2. *If one of the statements above holds, then the market is free of arbitrage.*

It is natural to call the function $g(t, T)$, the market price of risk. We remember that on a market free of arbitrage, the market price of risk is the same for all securities.

Suppose one of the statements above hold. Then we can define a unique measure Q that makes all discounted bond prices martingales. The question we ask us is, how does the forward process look like under this measure? The answer is surprisingly simple.

Theorem 16.3. *Let the forward dynamic under P be given by:*

$$df(t,T) = \alpha(t,T)dt + \sigma(t,T)dW(t)$$

Then, if any of the above statements holds, the forward rates under Q are given by

$$df(t,T) = \alpha^*(t,T)dt + \sigma(t,T)dV(t)$$

where

$$\alpha^*(t,T) = \sigma(t,T)\int_t^T \sigma(t,s)ds$$

Proof: After Girsanov transformation, the Q-dynamic is given by

$$df(t,T) = \{\alpha(t,T) + g(t)\sigma(t,T)\}dt + \sigma(t,T)dV(t)$$

Using

$$\alpha(t,T) = -\sigma(t,T)\left\{g(t,S) - \int_t^T \sigma(t,s)ds\right\}$$

$\Rightarrow$

$$df(t,T) = \left\{-\sigma(t,T)\left\{g(t) - \int_t^T \sigma(t,s)ds\right\} + g(t)\sigma(t,T)\right\}dt + \sigma(t,T)dV(t)$$

$$= \left\{\sigma(t,T)\int_t^T \sigma(t,s)ds\right\}dt + \sigma(t,T)dV(t) = \alpha^*(t,T)dt + \sigma(t,T)dV(t)$$

The result is a little bit surprising, since the Q-dynamic is completely determined by the diffusion function $\sigma(t, T)$. Therefore, if the process $\sigma(t, T)$ is deterministic, the forward rates are independent of the market price of risk. This is also true in a more complex situation where the forward rate solves a system of stochastic differential equations.

If we remember what we found when we went from df^T to dp^T and if we use the super index * in the drift and diffusion under Q we have:

$$df(t,T) = \alpha^*(t,T)dt + \sigma^*(t,T)dV(t)$$
$$dp(t,T) = p(t,T)\left\{r(t) + b^*(t,T)\right\}dt + p(t,T)a^*(t,T)dW(t)$$

$$\begin{cases} a^*(t,T) = -\int_t^T \sigma^*(t,u)du \\ b^*(t,T) = -\int_t^T \alpha^*(t,u)du + \frac{1}{2}a^*(t,T)^2 \end{cases}$$

We know that a Girsanov transformation do not change the drift, so we must have $\sigma^*(t, T) = \sigma(t, T)$. Then we know that under Q the yield is given by the short rate. Therefore we must have $b^*(t, T) = 0$.

16.1 The HJM program

To use the HJM framework we will use the following program:

1. Fix a filtrated probability space (Ω, $\mathcal{F}$, P, $\underline{\mathcal{F}}$) and a Wiener process V. The filtration is the natural, generated by V.
2. Specify the choice of volatility structure for the forward rates by for each $T > 0$, explicit give the process $\sigma(t, T)$.
3. Define the drift of the forward rates by

$$\alpha(t,T) = \sigma(t,T)\int_t^T \sigma(t,s)ds$$

4. Observe on the market, the initial forward structure $\{f^*(0, T); T > 0\}$.
5. Integrate the forward rates with the equations

$$f(t,T) = f^*(0,T) + \int_0^t \alpha(u,T)du + \int_0^t \sigma(u,T)dV(u)$$

6. Calculate the bond prices as

$$p(t,T) = \exp\left\{-\int_t^T f(t,u)du\right\}$$

7. Calculate derivative prices based on $p(t, T)$.

16.1.1 Ho-Lee model

To see how this work, we use the simplest we can think of, a constant volatility $\sigma(t, T) = \sigma > 0$. If we use the HJM equation

$$df(t,T) = \alpha^*(t,T)dt + \sigma(t,T)dV(t)$$

where

$$\alpha^*(t,T) = \sigma(t,T)\int_t^T \sigma(t,s)ds$$

We then get

$$df(t,T) = \left(\sigma\int_t^T \sigma ds\right)dt + \sigma dV(t) = \sigma^2(T-t)dt + \sigma dV(t)$$

We see that the drift is then given by: $\alpha(t, T) = \sigma^2(T - t)$. We then get

$$f(t,T) = f^*(0,T) + \int_0^t \sigma^2(T-u)du + \int_0^t \sigma dV(u) = f^*(0,T) + \sigma^2 t(T - \frac{t}{2}) + \sigma V(t)$$

We remember that a Wiener process at time $t = 0$ is zero: $V(0) = 0$. We recognize these rates as the one we got from the Ho-Lee model. Remark how easy we get them in the HJM framework. We also get the bond prices

$$p(t,T) = \exp\left\{ -\int_t^T f^*(0,u)du - \frac{\sigma^2 Tt}{2}(T-t) - \sigma(T-t)V(t) \right\}$$

i.e.,

$$p(t,T) = \frac{p^*(0,T)}{p^*(0,t)} \exp\left\{ -\frac{\sigma^2 Tt}{2}(T-t) - \sigma(T-t)V(t) \right\}$$

16.1.2 Hull-White model

If we use a Gaussian forward rates with volatility given by

$$\sigma(t,T) = \sigma e^{-a(T-t)}$$

We get

$$df(t,T) = \alpha^*(t,T)dt + \sigma(t,T)dV(t)$$

where

$$\alpha^*(t,T) = \sigma(t,T)\int_t^T \sigma(t,s)ds$$

Therefore

$$df(t,T) = \left(\sigma e^{-a(T-t)} \int_t^T \sigma e^{-a(T-s)} ds \right) dt + \sigma e^{-a(T-t)} dV(t)$$

$$= \frac{\sigma^2}{a}\left(e^{-a(T-t)} - e^{-2a(T-t)} \right) dt + \sigma e^{-a(T-t)} dV(t)$$

Integrating with respect to t, we obtain

$$f(t,T) = f^*(0,T) + \frac{\sigma^2}{2a^2}\left(1 - e^{-aT}\right)^2 - \frac{\sigma^2}{2a^2}\left(1 - e^{-a(T-t)}\right)^2 + \sigma \int_0^t e^{-a(T-s)} dV(s)$$

Introducing the notation

$$X(t) = \int_0^t e^{-a(t-s)} dV(s)$$

And using the fact that

$$\int_0^t e^{-a(T-s)} dV(s) = e^{-a(T-t)} X(t)$$

We obtain the following formulas for $f(t, T)$ and $r(t)$

$$f(t,T) = f^*(0,T) + \sigma e^{-a(T-t)} X(t) + \frac{\sigma^2}{2a^2}\left(1-e^{-aT}\right)^2 - \frac{\sigma^2}{2a^2}\left(1-e^{-a(T-t)}\right)^2$$

and

$$r(t) = f^*(0,t) + \sigma X(t) + \frac{\sigma^2}{2a^2}\left(1-e^{-at}\right)^2$$

We recognize these rates as the Hull-White (extended Vasicek) model. The formula shows that the forward rates and the instantaneous short rate are linear functions of the same Gaussian process $X(t)$, so we observe a perfect correlation of the forward rates.

The asymptotical behaviour of the short rate is given by

$$r(t) = f(0,t) + \sigma X(t) + \frac{\sigma^2}{2a^2}$$

which is a Gaussian random variable with mean

$$\mu_\infty(t) = f(0,t) + \frac{\sigma^2}{2a^2}$$

and variance

$$\sigma_\infty^2 = \sigma^2 E\left[\left(\int_0^t e^{-a(t-s)} dV(s)\right)^2\right] = \sigma^2 \int_0^t e^{-2a(t-s)} ds = \frac{\sigma^2}{2a}\left(1-e^{-2at}\right)$$

We see that the short rate fluctuation have a non-trivial asymptotic probability distribution. This fact is known as **mean-reversion** of the spot rate and a is called the **rate of mean reversion**.

16.1.3 The general situation

In a mode general situation we can let the volatility depends on the forward rates and then solve a system of stochastic differential equations under Q. In more detail

1. Specify σ as function of three variables: t, T and $f(t, T)$.
2. Solve

$$\begin{cases} df(t,T) = \alpha(t,T)dt + \sigma(t,T,f(t,T))dV(t) \\ f(0,T) = f^*(0,T) \end{cases}$$

where

$$\alpha(t,T) = \sigma(t,T,f(t,T))\int_t^T \sigma(t,s,f(t,s))ds$$

The question to ask at this point is under what condition on σ we can solve the equations above. The situation is complex since this is infinite number of coupled equations where $\alpha(t, T)$ at time t, not only depends on the actual forward rates, $f(t, T)$ but also all forward rates $f(t, s)$ with $t \leq s \leq T$. But more difficult is the problem with σ. If we do not specify σ well enough, α which is quadratic in σ can explode and give infinite forward rates. This gives bond prices of zero and possible arbitrage situations. If σ is Lipschitz continuous in $f(t, s)$, positive and uniformed limited, then there exist a solution to the system for all initial forward rates.

16.2 A change of perspective

The main result of the HJM approach consists in providing the extension of the Black and Scholes (1973) reasoning to the fixed income sector using forward rates. This can be done as there exists a one to one correspondence between instantaneous forward rates and bond prices. Bonds are traded assets, so we can apply the procedure of replacing the drift coefficient with the short rate under a risk neutral probability measure. Passing from spot rates to forward rates, thus, allows us to incorporate directly arbitrage restrictions without specifying in advance the market price of risk.

The noticeable fact is that the drift, determined by arbitrage arguments, depends only on the volatility parameters, and this resembles the Black and Scholes (1973) results. In this sense it can be said that the HJM can be considered the true extension of their methodology to the fixed income sector. Up to now it might seem that HJM comes in at no cost, but this is not the case. Switching to forward rates and relying only on volatility calibration, has two main drawbacks: firstly, under the risk free measure, the forward rates are biased estimators of the future spot rates; secondly there may be cases in which the spot rate does not follow a Markov process. This is a somewhat unpleasant feature of the HJM approach because of the heavy computational difficulties arising in non-Markovian contexts.

Jeffrey in a 1995 article derived general conditions on the volatility structures in HJM under which markovness is still retained. To be more specific, he provides necessary and sufficient conditions such that one can determine which volatility structures are allowable in a Markovian spot interest rate context and the set of allowable initial term structures corresponding to a given volatility structure. We will not discuss this here, because this is out of the scope at this point.

Mari (2003) has proposed a perturbative extension of a model for Bond Prices within the affine class. This model is set up with the property of consistency with arbitrary initial term structures. Affine structures posse's very interesting properties: first of all they are mathematically very tractable; as a consequence they allow for risk analysis and estimation via closed form solutions of PDE or via solutions of ODE of the first order; moreover they can be estimated using maximum likelihood techniques. In the model it is assumed that the discount factor is a smooth function of the spot rate and of maturity *T*.

Under the risk neutral measure, the stochastic dynamics of the term structure is given by

$$\begin{cases} \frac{dp(t,T,r(t))}{p} = r(t)dt + \sigma_p(t,T,r(t))dW^*(t) \\ p(0,T,r(0)) = p^*(0,T) \end{cases}$$

Mari (2003) proved that the model can be fitted consistently with arbitrary initial term structures and the implied spot rate follows a Markov process if and only of the following condition holds

$$\sigma_p(t,T,r(t)) = \sqrt{h(t)+k(t)r(t)}B(t,T)$$

with

$$B(t,T) = 2\frac{C'(t)-A(t)}{k(t)}\left[\frac{1}{C(t)} - \frac{1}{\int_t^T A(u)du + C(t)}\right]$$

and

$$A(t) = \frac{1}{2}\left\{C'(t) - \sqrt{C'^2(t) - 2k(t)C^2(t)}\right\}$$

where $h(t)$, $k(t)$ and $C(t)$ are functions that can be arbitrarily chosen which must satisfy the condition $C'(t)^2 \geq 2k(t)C^2(t)$. Under this condition Mari shows that the solution of the term structure is:

$$p(t,T,r(t)) = \frac{p^*(0,T)}{p^*(0,t)}\exp\left\{\begin{matrix} f^*(0,t)B(t,T) - \int_0^t H(u)B(u,T)du + \\ \frac{1}{2}\int_0^t \sigma^2(u,T,f^*(0,t))B^2(u,T)du \end{matrix}\right\} e^{-r(t)B(t,T)}$$

where $H(t)$ is the solution of the Volterra integral equation of the first kind:

$$\int_0^t H(u)B(u,T)du = G(t)$$

with

$$g(t) = -\frac{1}{2}\int_0^t \sigma^2(u,T,f^*(0,t))B^2(u,T)du$$

and f^* as usual is the initial forward rate curve. As a Corollary, Mari proves that the dynamics of the spot rate is described by

$$dr(t) = \{a(t) - b(t)r(t)\}\,dt + \sqrt{h(t) + k(t)r(t)}dW(t)$$

where

$$\begin{cases} a(t) = \dfrac{\partial f^*(0,t)}{\partial t} + b(t)f^*(0,t) - H(t) \\ b(t) = \dfrac{\partial B(t,T)}{\partial T^2}\big|_{T=t} \end{cases}$$

The innovation of the paper consists in the explicit determination of the function $a(t)$ which is the term accounting for the initial term structure, and in bringing to the forefront the Volterra equation, as a device to overcame the obstacle met by Hull and White (1992). In the following, Mari goes on considering some applications.

Gaussian Models, the CIR volatility structure and the generalized CIR volatility structure, in which $h(t) \neq 0$. The problem is that, in general, Volterra equation does not admit closed form solution, except for simple case, as the Vasicek one. In fact for the extended Vasicek model, this method gives the same solution of the Hull and White (1992). As for the generalized CIR model, a perturbative solution of the Volterra equation is proposed.