Models of the Short Rate

The short rate, \( r \), at time \( \mathcal{E} \) is the risk-free rate over an infinitesimally period of time \([\mathcal{E}, \mathcal{E} + \delta]\).
This is exactly the rate we used in our Black-Scholes PDE derivation (but way back then, all rates were constant).

We will look at models for the dynamics of the short rate (always in a risk-neutral world).

From our previous work, in a world that is forward risk-neutral with respect to the risk-free rate (i.e., the traditional risk neutral world), we have

\[
\frac{\delta}{\delta t} \hat{\mathcal{E}} = \hat{\mathcal{E}} \left( e^{-\hat{r}(\mathcal{E} - \mathcal{E})} \right)
\]

That is, the value of a derivative with payoff \( \hat{f} \) is a discounted expectation in the risk-neutral world, where

\[
\hat{r} = \frac{1}{\mathcal{E} - \mathcal{E}} \int_{\mathcal{E}}^{\mathcal{E} + \delta} r(t) \, dt
\]

is the average rate over \([\mathcal{E}, \mathcal{T}]\).

If we define the process for the short rate, we specify the entire zero coupon yield.

Letting \( P(\mathcal{E}, \mathcal{T}) \) be the price of a zero coupon bond at time \( \mathcal{E} \) maturing at \( \mathcal{T} \), we have:
\[ P(t, T) = \hat{P} \left( e^{-f(t-T)} \right) \]  

Price of \( P \) at \( T \) is \( \hat{P} \).

If we define \( R(t, T) \) as the continuously compounded interest rate at time \( t \) for term \( (T-t) \), we have

\[ P(t, T) = e^{-R(t, T)(T-t)} \]

and so

\[ R(t, T) = -\frac{1}{T-t} \ln (P(t, T)) \]

\[ = -\frac{1}{T-t} \ln \hat{P} \left( e^{-f(T-t)} \right) \]

Two broad categories may be defined for models of the short rate:

- Equilibrium models
- No-arbitrage models

In the first case, stylized features are incorporated into the model like mean reversion.

In the second, a parameterization of the observed term structure of interest rates is made.

Other differentiators obtain, but these largely capture the distinction. Alternatively, we may summarize these statements as

Equilibrium: model \( \rightarrow \) term structure
No-arbitrage: term structure \( \rightarrow \) model
Equilibrium Models

We will look at one-factor equilibrium models—that is, to process with one source of risk:

\[ dr = \mu(r)dt + \sigma(r)dZ \]

Both \( \mu(r) \) and \( \sigma(r) \) are functions of the rate, but not time.

(Important) Examples:

- Rendleman & Barter: \[ dr = \mu r dt + \sigma r dZ \]
- Vasicek: \[ dr = \alpha (b - r) dt + \sigma dZ \]
- Cox-Ingersoll-Ross: \[ dr = \alpha (b - r) dt + \sigma \sqrt{r} dZ \]

Both the Vasicek and CIR models incorporate mean reversion into the dynamics of \( r \); i.e., for large \( r \), the drift is negative, and for small \( r \), the drift is positive, both centering around \( b \).

Mean reversion in rates is a well-established stylized feature of rates... and a sharp contrast to stock dynamics.

As a result, there isn't much to say about the Rendleman & Barter model other than it came first.
Vasicek model

Sometimes referred to as the Ornstein-Uhlenbeck model.

In the model:

\[ dL_t = (a(b - L_t))dt + \sigma dZ_t \]

\( a, b, \sigma \) are each constants with:

\( a \): the drift rate
\( b \): the mean rate
\( \sigma \): the volatility.

When the short rate has these dynamics, a zero coupon bond price is:

\[ P(t, T) = A(t, T) e^{-B(t, T) r(t)} \]

Where:

\[ B(t, T) = \frac{1 - e^{-(a(T-t))}}{a} \]

\[ A(t, T) = \exp \left( \frac{B(t, T) - (T-t)}{a^2} \left( a^2 \int_{t}^{T} r(t') dt' - \frac{1}{2} \sigma^2 \right) - \frac{\sigma^2 B(t, T)^2}{4a} \right) \]

(We'll see this structure again for the CIR model.)

In the case \( a = 0 \), there is no drift and:

\[ B(t, T) = T - t \]

\[ A(t, T) = \exp \left( \frac{1}{2} \sigma^2 (T-t)^3 \right) \]
The distinguishing feature of the CIR model, then, is the stochastic term, where the standard deviation of the short rate becomes proportional to \( \sqrt{t} \); this is another stylized feature of rates.

As in Vasicek, in the CIR model, we have

\[
P(t, T) = A(t, T) \exp \left[ -B(t, T) (T - t) \right]
\]

where now

\[
B(t, T) = \frac{2 \left( e^{\frac{1}{2} (T-t)} - 1 \right)}{(\delta + a) \left( e^{\frac{1}{2}(T-t)} - 1 \right) + 2 \gamma}
\]

and

\[
A(t, T) = \left( \frac{2 \delta e^{\frac{1}{2}(a+\gamma)(T-t)}}{(\delta + a) \left( e^{\frac{1}{2}(T-t)} - 1 \right) + 2 \gamma} \right)^{\frac{2ab}{\delta + 2}}
\]

and

\[
\gamma = \sqrt{a^2 + 2\delta^2}
\]
Properties common to both models

In both cases we have

$$\frac{\delta P}{\delta r} = -B(t,T)A(t,T)e^{-B(t,T)r(t)}$$

$$= -B(t,T)P(t,T)$$

Further, for

$$R(t,T) = -\frac{1}{T-t} \ln P(t,T)$$

we have

$$R(t,T) = -\frac{1}{T-t} \ln \left[ A(t,T)e^{-B(t,T)r(t)} \right]$$

$$= -\frac{1}{T-t} \left( \ln A(t,T) - B(t,T)r(t) \right)$$

(*)

Giving an explicit formula for the term structure of rates.

NB: $R(t,T)$ is linearly dependent on $r(t)$.

Further, the shape of the term structure is dependent on $A$ and $B$. 
We can measure the sensitivity of a bond to changes in the short rate as

\[ -\vec{D} \cdot \vec{Q} = \frac{\partial Q}{\partial r} \]

for \( Q \), the price of the bond.

**Example**

In the case that \( Q = P(t, T) \) and the short rate follows either the Vasicek or CIR model, we have

\[ \frac{\partial P}{\partial r} = -B \cdot P \]

and hence

\[ \vec{D} = B(t, T) \]

**Example**

Consider a three-year zero-coupon bond.

In this case \( \vec{D} = 3 \), meaning that an 10 bps parallel shift in the yield curve leads to a -3.00% (term structure) change in the bond price; viz., if \( r_0 = 10 \) bps, the bond price decreases by 3.01 = 0.3%.

In Vasicek's model, if \( \alpha = 0.2 \), we have

\[ \vec{D} = B(0.3) : \frac{(1 - e^{-0.2 \cdot 3})}{0.2} = 2.75 \]

so that a 10 bps parallel shift in the term structure leads to a -0.28% change in bond price.
The above can be aggregated to the portfolio level (of course this is true as it's just a partial derivative):

$$\frac{\sigma^2}{\alpha} \frac{\delta^2}{\delta t^2} = -\frac{1}{\alpha} \frac{\delta}{\delta t} \left( \sum_{i=1}^{n} p_i \cdot Q_i \right)$$

$$= -\frac{1}{\alpha} \sum_{i=1}^{n} p_i \cdot \frac{\delta Q_i}{\delta t}$$

$$= -\frac{1}{\alpha} \sum_{i=1}^{n} p_i \cdot (-Q_i \cdot \delta_i)$$

$$= \frac{1}{\alpha} \sum_{i=1}^{n} p_i \cdot Q_i \cdot \delta_i$$

That is, $\hat{D}$ is a weighted average of $\hat{E_i}$. What is $\hat{C}$?

**Question:** Can you calculate a convexity measurement, $\hat{C}$, similar to the original convexity calculation presented by Hull? Specifically, if we have a zero coupon bond priced under the Vasicek model, what is $\hat{C}$?

Write a Taylor expansion for $\Delta^2$ in $\Delta r$ and $(\Delta r)^2$. 


SDE for \( P(t,T) \) \text{ [in the traditional risk neutral world]} 

In the traditional risk neutral world we know that

\[
\frac{\partial P(t,T)}{\partial t} = r(t)P(t,T)dt + \frac{1}{2} \sigma^2(t) \frac{\partial^2 P(t,T)}{\partial T^2} dt
\]

For some \( f \) ... which we can find by Ito (as usual)

\text{ Vasicek Case }\]

By Ito we have

\[
\frac{\partial P}{\partial t} = \frac{\partial P}{\partial t} + \text{ first order term } + \frac{1}{2} \frac{\partial^2 P}{\partial t^2} \left( \sigma(t) \right)^2
\]

\[
= \frac{\partial P}{\partial t} + \left( a(t) \mathbf{b} \cdot (b - \mathbf{r}) \right) dt + \sigma^2(t) \frac{\partial^2 P}{\partial t^2} dt
\]

\[
= \left[ \frac{\partial P}{\partial t} + \left( a(t) \mathbf{b} \cdot (b - \mathbf{r}) \right) + \frac{1}{2} \frac{\partial^2 P}{\partial t^2} \sigma^2(t) \right] dt + \frac{\partial P}{\partial t} \frac{\partial^2 P}{\partial T^2} dt
\]

\text{ In the traditional risk neutral world, the drift is simply } r(t) P, \text{ giving }

\[
\frac{\partial P}{\partial t} = r(t)P(t,T)dt + \frac{\partial P}{\partial t} \frac{\partial^2 P}{\partial T^2} dt
\]

\text{ and since under Vasicek we have }

\[
\frac{\partial P}{\partial t} = -B \cdot P, \text{ we have }
\]

\[
\frac{\partial P(t,T)}{\partial t} = r(t)P(t,T)dt + -B \cdot P(t,T) \frac{\partial^2 P}{\partial T^2} dt
\]
One more comment: Vasicek's model can give negative rates. (This used to seem crazy).

**Fitting the Models - Vasicek**

Begin by aggregating short-term rates — on a fixed time interval; e.g., weekly or daily. Call this data

\[ \{ r_i \}_{i=1}^N \]

**Regression**

The discrete version of Vasicek implies

\[ \Delta r_i = a \cdot (b - r_i) \Delta t + \sigma \cdot \epsilon \sqrt{\Delta t} \]

\[ \Delta r_i = ab \Delta t - a \Delta t \cdot r_i + \sigma \sqrt{\Delta t} \cdot \epsilon \]

We may therefore regress \( \Delta r_i \) on \( r_i \) as

\[ \Delta r_i = \alpha + \beta r_i + \tilde{\epsilon} \]

The regression coefficients \( \hat{\alpha} \) and \( \hat{\beta} \) satisfy

\[ -a \Delta t = \hat{\beta} \quad \Rightarrow \quad \hat{\alpha} = -\frac{\hat{\beta}}{\Delta t} \]

\[ a \cdot b \cdot \Delta t = \hat{\alpha} \quad \Rightarrow \quad \hat{\beta} = \frac{a \Delta t}{\hat{\alpha}} \]

\[ sd(\tilde{\epsilon}) = \sigma \sqrt{\Delta t} \quad \Rightarrow \quad \sigma = \frac{sd(\tilde{\epsilon})}{\sqrt{\Delta t}} \]
Maximum Likelihood

Looking again at the discrete version of Vasicek, we have

\[ \Delta r_i = a \left( b - r_i \right) \Delta t + \sqrt{\Delta t} \epsilon \]

so that conditioned on having observed \( r_i \) and fixed \( a, b, \) and \( \sigma \), \( \Delta r_i \) is normally distributed as

\[ \Delta r_i \sim N \left( a \left( b - r_i \right) \Delta t, \sigma^2 \Delta t \right) \]

The likelihood function for \( \Delta r_i \) is therefore

\[ f \left( \Delta r_i \mid r_i; a, b, \sigma \right) = \frac{1}{\sqrt{2\pi \sigma^2 \Delta t}} \exp \left[ -\frac{1}{2} \frac{\left( \Delta r_i - a \left( b - r_i \right) \Delta t \right)^2}{\sigma^2 \Delta t} \right] \]

The likelihood function is given by

\[ l_0 \left( a, b, \sigma \right) = \prod_i f \left( \Delta r_i \mid r_i; a, b, \sigma \right) \]

with log-likelihood

\[ l \left( a, b, \sigma \right) = \sum_i - \frac{1}{2} \ln \left( 2\pi \sigma^2 \Delta t \right) - \frac{1}{2} \left[ \frac{\left( \Delta r_i - a \left( b - r_i \right) \Delta t \right)^2}{\sigma^2 \Delta t} \right] \]

\[ \propto \sum_i \left( -\ln (\sigma^2 \Delta t) - \frac{\left( \Delta r_i - a \left( b - r_i \right) \Delta t \right)^2}{\sigma^2 \Delta t} \right) \]
NB: These estimates are for the real world. To have the dynamics in risk-neutral world, we have to use the change of numeraire results, obtaining
\[ \Delta r = (a(b-r) - \lambda \sigma) \Delta t + \sigma \Delta \epsilon \]
with $\lambda$ the market price of interest risk.