Review GARCH & Credit Risk
Estimating Volatilities and Correlations

We discuss how to use historical data to extract estimates on the current and future levels of volatilities and correlations.

**Estimating Volatility:** Define $\sigma_n$ as the volatility of a market variable on day $n$, as estimated at the end of day $n - 1$. The square of the volatility, $\sigma_n^2$, on day $n$ is the **variance rate**. We described the standard approach to estimating $\sigma_n$ from historical data.

Suppose that the value of the market variable at the end of day $i$ is $S_i$. The variable $u_i$ is defined as the continuously compounded return during day $i$

$$u_i = \ln \frac{S_i}{S_{i-1}}$$

An unbiased estimate of the variance rate per day, $\sigma_n^2$ using the most recent $m$ observations on the $u_i$ is

$$\sigma_n^2 = \frac{1}{m - 1} \sum_{i=1}^{m} (u_{n-i} - \bar{u})^2$$

where $\bar{u}$ is the mean of the $u_i$'s:

$$\bar{u} = \frac{1}{m} \sum_{i=1}^{m} u_{m-i}$$
For the purposes of monitoring daily volatility, the formula is usually changed in a number of ways

1. $u_i$ is defined as the percentage change in the market variable between the end of the day $i - 1$ and the end of day $i$ so that

$$u_i = \frac{S_i - S_{i-1}}{S_{i-1}} \approx \ln(1 + \frac{S_i - S_{i-1}}{S_{i-1}}) \quad (1)$$

2. $\bar{u}$ is assumed to be zero

3. $m - 1$ is replaced by $m$

These three changes make very little difference to the estimates that are calculated, but they allow us to simplify the formula for the variance rate to

$$\sigma_n^2 = \frac{1}{m} \sum_{i=1}^{m} u_{n-i}^2 \quad (2)$$

where $u_i$ is given by (1).
The sigma given by (2) gives equal weight to

\[ u_{n-1}^2, u_{n-2}^2, \ldots, u_{n-m}^2 \]

Our objective is to estimate the current level of volatility \( \sigma_n \). Therefore, it makes sense to give more weight to recent data. One such model is

\[ \sigma_n^2 = \sum_{i=1}^{m} \alpha_i u_{n-i}^2 \]  

(3)

The variable \( \alpha_i \) is the amount of weight given to the observation \( i \) days ago. The \( \alpha \)'s are positive.

If we choose them so that \( \alpha_i < \alpha_j \) when \( i > j \), less weight is given to older observations. The weights must sum to unity, so we have

\[ \sum_{i=1}^{m} \alpha_i = 1 \]
Weighting Schemes, cont.

An extension of the idea, called $\text{ARCH}(m)$ or Autoregressive Conditional Heteroscedasticity in equation (3) is to assume that there is a long-run average variance rate and that this should be given some weight. This leads to the model that takes the form

$$\sigma_n^2 = \gamma V_L + \sum_{i=1}^{m} \alpha_i u_{n-i}^2$$

where $V_L$ is the long-run variance rate and $\gamma$ is the weight assigned to $V_L$. Because the weights must sum to unity, we have

$$\gamma + \sum_{i=1}^{m} \alpha_i = 1$$

Define $\omega = \gamma V_L$, the model equation becomes

$$\sigma_n^2 = \omega + \sum_{i=1}^{m} \alpha_i u_{n-i}^2$$
Exponentially Weighted Moving Average Model

The Exponentially Weighted Moving Average Model or EWMA is a particular case of (3) where the weights $\alpha_i$ decrease exponentially as we move back through time.

Specifically $\alpha_{i+1} = \lambda \alpha_i$, where $\lambda$ is a constant between 0 and 1.

The formula is

$$\sigma^2_n = \lambda \sigma^2_{n-1} + (1 - \lambda) u^2_{n-1}$$

(6)

The estimate $\sigma_n$ of the volatility of day $n$ is calculated from $\sigma_{n-1}$ and $u_{n-1}$.

We can see why this corresponds to exponentially decreasing weights. Continuing yields

$$\sigma^2_n = (1 - \lambda) \sum_{i=1}^{m} \lambda^{i-1} u^2_{n-i} + \lambda^m \sigma^2_{n-m}$$

(7)

For large $m$, the term $\lambda^m \sigma^2_{n-m}$ is sufficiently small to be ignored so that this is the same as a equation with

$$\alpha_i = (1 - \lambda) \lambda^{i-1}$$

The weights for the $u_i$ decline at rate $\lambda$ as we move back through time. Each weight is $\lambda$ times the previous weight.
The EWMA approach is attractive since only relatively little data needs to be stored.

At any given time, we need to store only the current estimate of the variance rate and the most recent observation on the value of the market variable.

When we get a new observation on the value of the market variable, we calculate a new daily percentage change and use equation

\[
\sigma_n^2 = \lambda \sigma_{n-1}^2 + (1 - \lambda) u_{n-1}^2
\]

to update our estimate of the variance rate.

The value of \( \lambda \) governs how responsive the estimate of the daily volatility is to the most recent daily percentage change.

A low value of \( \lambda \) leads to a great deal of weight being given to the \( u_{n-1}^2 \) when \( \sigma_n \) is calculated.

A high value of \( \lambda \) (close to 1.0) produces estimates of the daily volatility that respond relatively slowly to new information provided by the daily percentage change.
The GARCH(1,1) Model

A more generalized method to generate volatilities is the GARCH(1,1), or generalized autoregressive conditional heteroscedasticity.

- GARCH(1,1) differs from EWMA by including a long-run variance rate. \( \sigma_n^2 \) is calculated from a long-run average variance rate \( V_L \) as well as from \( \sigma_{n-1}^2 \) and \( u_{n-1} \):

\[
\sigma_n^2 = \gamma V_L + \alpha u_{n-1}^2 + \beta \sigma_{n-1}^2
\]

where \( \gamma \) is the weight assigned to \( V_L \), \( \alpha \) is the weight assigned to \( u_{n-1}^2 \), and \( \beta \) is the weight assigned to \( \sigma_{n-1}^2 \).

- The weights must sum as

\[
\gamma + \alpha + \beta = 1
\]

- EWMA is a particular case of GARCH(1,1) with \( \gamma = 0 \), \( \alpha = 1 - \lambda \), and \( \beta = \lambda \).

- The \((1, 1)\) in GARCH(1,1) implies that \( \sigma_n^2 \) is based on the most recent observation of \( u^2 \) and the most recent estimate of the variance rate.
Setting $\omega = \gamma V_L$, the GARCH(1,1) model can also be written

$$\sigma_n^2 = \omega + \alpha u_{n-1}^2 + \beta \sigma_{n-1}^2$$

This is the form of the model that is usually used for the purposes of estimating the parameters.

Once $\omega$, $\alpha$, and $\beta$ have been estimated, we can calculate $\gamma$ as $1 - \alpha - \beta$. The long-term variance $V_L$ can then be calculated as $\omega / \gamma$.

For a stable GARCH(1,1) process we require $\alpha + \beta < 1$. Otherwise the weight applied to the long-term variance is negative.
GARCH(1,1), cont.

The Weights:
Substituting for $\sigma_{n-1}^2$ in the GARCH(1,1) model, we obtain

$$\sigma_n^2 = \omega + \alpha u_{n-1}^2 + \beta (\omega + \alpha u_{n-2}^2 + \beta \sigma_{n-2}^2)$$

Substituting for $\sigma_{n-2}^2$ we get

$$\sigma_n^2 = \omega + \beta \omega + \beta^2 \omega + \alpha u_{n-1}^2 + \alpha \beta u_{n-2}^2 + \alpha \beta^2 u_{n-3}^2 + \beta^3 \sigma_{n-3}^2$$

Continuing on we see that the weight applied to $u_{n-i}^2$ is $\alpha \beta^{i-1}$.

The weights decline exponentially at rate $\beta$.

The parameter $\beta$ can be interpreted as a decay rate. It is similar to $\lambda$ in the EWMA model.

It defines the relative importance of the observations on the $u$’s in determining the current variance rate.

Mean Reversion

- The GARCH(1,1) model recognizes that over time the variance tends to get pulled back to a long-run average level of $V_L$. 
Choosing Between Models

In practice, variance rates tend to be mean reverting. The GARCH(1,1) model incorporates mean reversion, whereas EWMA model does not. The GARCH(1,1) model is therefore theoretically more appealing than the EWMA model.

In circumstances where the best-fit value of $\omega$ turns out to be negative, the GARCH(1,1) model is not stable and it makes sense to switch to EWMA model.
Now appropriate to discuss how the parameters in the models we have been considering are estimated from historical data. We use the **maximum likelihood method**. We choose parameters that maximize the chance of the data occurring.

Consider the problem of estimating a variance of a variable $X$ from $m$ observations on $X$ when the underlying distribution is normal with mean zero.

We assume that the observations are $u_1, u_2, \ldots, u_m$ and that the mean of the underlying distribution is zero. Denote the variance by $v$. The likelihood of $u_i$ being observed is the probability density function for $X$ when $X = u_i$. This is

$$\frac{1}{\sqrt{2\pi v}} \exp \left[ -\frac{u_i^2}{2v} \right]$$

The likelihood of $m$ observations occurring in the order in which they are observed is

$$\prod_{i=1}^{m} \left[ \frac{1}{\sqrt{2\pi v}} \exp \left[ -\frac{u_i^2}{2v} \right] \right]$$

(8)
Using the maximum likelihood method, the best estimate of $v$ is the value that maximizes this expression.

Maximizing an expression is equivalent to maximizing the logarithm of the expression. Taking logarithms of the expression in (8).

Ignoring constant multiplicative factors, it can be seen that we wish to maximize

$$
\sum_{i=1}^{m} \left[-\ln(v) - \frac{u_i^2}{v}\right]
$$

or

$$
-m \ln(v) - \sum_{i=1}^{m} \frac{u_i^2}{v}
$$

Differentiating this expression with respect to $v$ and setting the result equation to zero, we see that the maximum likelihood estimator of $v$ is

$$
\frac{1}{m} \sum_{i=1}^{m} u_i^2
$$
Estimating GARCH(1,1) Parameters

We now consider how the maximum likelihood method can be used to estimate the parameters when GARCH(1,1) or some other volatility updating scheme is used.

Define \( v_i = \sigma_i^2 \) as the variance estimated for day \( i \). We assume that the probability distribution of \( u_i \) conditional on the variance is normal. A similar analysis to the one just given shows the best parameters are the ones that maximize

\[
\prod_{i=1}^{m} \left[ \frac{1}{\sqrt{2\pi v_i}} \exp \left( -\frac{u_i^2}{2v_i} \right) \right]
\]

Taking logarithms we see that this is equivalent to maximizing

\[
\sum_{i=1}^{m} \left[ -\ln(v_i) - \frac{u_i^2}{2v_i} \right]
\]

This is the same as the expression above except that \( v \) is replaced by \( v_i \) is replaced by \( v_i \). We search iteratively to find the parameters in the model that maximize the expression in (9)

<table>
<thead>
<tr>
<th>Date</th>
<th>Day $i$</th>
<th>$S_i$</th>
<th>$w_i$</th>
<th>$v_i = \sigma_i^2$</th>
<th>$-ln(v_i) - \frac{u_i^2}{v_i}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>06-Jan-88</td>
<td>1</td>
<td>0.007728</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>07-Jan-88</td>
<td>2</td>
<td>0.007779</td>
<td>0.006599</td>
<td></td>
<td></td>
</tr>
<tr>
<td>08-Jan-88</td>
<td>3</td>
<td>0.007746</td>
<td>-0.004242</td>
<td>0.00004355</td>
<td>9.683</td>
</tr>
<tr>
<td>09-Jan-88</td>
<td>4</td>
<td>0.007816</td>
<td>0.009037</td>
<td>0.00004198</td>
<td>8.1329</td>
</tr>
<tr>
<td>11-Jan-88</td>
<td>5</td>
<td>0.007837</td>
<td>0.002687</td>
<td>0.00004455</td>
<td>9.8568</td>
</tr>
<tr>
<td>12-Jan-88</td>
<td>6</td>
<td>0.007924</td>
<td>0.011101</td>
<td>0.000042240</td>
<td>7.1529</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13-Aug-97</td>
<td>2421</td>
<td>0.008643</td>
<td>0.003374</td>
<td>0.00007626</td>
<td>9.3321</td>
</tr>
<tr>
<td>14-Aug-97</td>
<td>2422</td>
<td>0.008493</td>
<td>-0.0137309</td>
<td>0.00007092</td>
<td>5.3294</td>
</tr>
<tr>
<td>15-Aug-97</td>
<td>2423</td>
<td>0.008495</td>
<td>0.000144</td>
<td>0.00008417</td>
<td>9.3824</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

22,063.5763

In the example the optimal values of the parameters are

$$\omega = 0.00000176 \quad \alpha = 0.0626 \quad \beta = 0.8976$$

and the maximum value of the function in the equation is 22,063.5763.
The numbers were calculated on the final iteration of the search for the optimal $\omega, \alpha, \beta$.

The long term variance rate $V_L$ in the example is

$$\frac{\omega}{1 - \alpha - \beta} = \frac{0.00000176}{0.0398} = 0.00004422$$

The long-term volatility then is $\sqrt{0.00004422} = 0.665\%$. 
Returning to the table we see the GARCH(1,1) volatility for Japanese yen changed over the 10-year period covered by the data.

![Figure 19.1] Daily volatility of the yen:USD exchange rate, 1988–1997.

Most of the time, the volatility was between 0.4% and 0.8% per day, but volatilities over 1% were experienced during some periods.
Variance Sampling

An alternative and more robust approach to estimating the parameters in GARCH(1,1) is known as variance sampling. This involves setting the long-run average variance rate $V_L$ equal to the sample variance calculated from the data.

- The value of $\omega$ the equals $V_L(1 - \alpha - \beta)$ and only two parameters have to be estimated. For our data the sample variance is 0.00004341, which gives a daily volatility of 0.659%.
- Setting $V_L$ equal to the sample variance, the values of $\alpha$ and $\beta$ that maximize the objective function in equation

$$
\sum_{i=1}^{m} \left[ -\ln(v_i) - \frac{u_i^2}{v_i} \right]
$$

are 0.0607 and 0.8990, respectively. The value of the objective function is 22,063.537, only marginally below the value of 22,063.5763 obtained using the previous method.
- When the EWMA model is used, the estimation procedure is relatively simple. We set $\omega = 0$, $\alpha = 1 - \lambda$, and $\beta = \lambda$ and only one parameter has to be estimated.
In the data in our table, the value of $\lambda$ that maximizes the objective function in

$$\sum_{i=1}^{m} \left[ -\ln(v_i) - \frac{u_i^2}{v_i} \right]$$

is 0.9686 and the value of the objective function is 21,995.8377.

Both GARCH(1,1) and EWMA method can be implemented by using the Solver routine in Excel to search for the values of the parameters that maximize the likelihood function. The routine works well provided that we structure our spreadsheet so that the parameters we are searching for have roughly equal values.
Using GARCH(1,1) to forecast future volatility

The variance rate estimated at the end of the day \( n - 1 \) for the day \( n \), using GARCH(1,1) is

\[
\sigma_n^2 = (1 - \alpha - \beta) V_L + \alpha u_{n-1}^2 + \beta \sigma_{n-1}^2.
\]

On day \( n + t \) in the future, we get

\[
\sigma_{n+t}^2 - V_L = \alpha \left( u_{n+t-1}^2 - V_L \right) + \beta \left( \sigma_{n+t-1}^2 - V_L \right)
\]

The expected value of \( u_{n+t-1}^2 \) is \( \sigma_{n+t-1}^2 \); hence

\[
E[\sigma_{n+t}^2 - V_L] = (\alpha + \beta) E[\sigma_{n+t-1}^2 - V_L]
\]

where \( E \) denotes the expected value. Using this equation repeatedly we find

\[
E[\sigma_{n+t}^2 - V_L] = (\alpha + \beta)^t \left( \sigma_n^2 - V_L \right)
\]

or

\[
E[\sigma_{n+t}^2] = V_L + (\alpha + \beta)^t \left( \sigma_n^2 - V_L \right)
\] (10)
Formula (10) forecasts the volatility on day $n + t$ using the information available at the end of day $n - 1$. In the EWMA model, $\alpha + \beta = 1$ and equation (10) shows that the expected future variance rate equals the current variance rate.

When $\alpha + \beta < 1$, the final term in the equation becomes progressively smaller as $t$ increases.

The variance rate exhibits mean reversion with a reversion level of $V_L$ and a reversion rate of $1 - \alpha - \beta$. Our forecast of the future variance rate tends towards $V_L$ as we look further and further ahead. This analysis emphasizes the point that we must have $\alpha + \beta < 1$ for a stable GARCH(1,1) process.

![Figure 19.2](image)

**Figure 19.2** Expected path for the variance rate when (a) current variance rate is above long-term variance rate and (b) current variance rate is below long-term variance rate.
In the yen-dollar exchange rate example considered earlier $\alpha + \beta = 0.962$, and $V_L = 0.00004422$. Suppose that or estimate of the current variance rate per day is 0.00006 (corresponding to a volatility of 0.7% per day). In 10 days the expected variance rate is

$$0.00004422 + 0.9602^{10} (0.00006 - 0.00004422) = 0.00005473$$

The expected volatility per day is 0.74%, still well above the long-term volatility of 0.665% per day.

However, the expected variance rate in 100 days is

$$0.00004422 + 0.9602^{100} (0.00006 - 0.00004422) = 0.00004449$$

and the expected volatility per day is 0.667%, very close to the long-term volatility.
Volatility Term Structures

Suppose it is day $n$. Define

$$V(t) = E[\sigma_{n+t}^2]$$

and

$$a = \ln \frac{1}{\alpha + \beta}$$

then (10) becomes

$$V(t) = V_L + e^{-at} [V(0) - V_L]$$

here $V(t)$ is an estimate of the instantaneous variance rate in $t$ days. The average variance rate per day between today and time $T$ is given by

$$\frac{1}{T} \int_0^T V(t) dt = V_L + \frac{1 - e^{-aT}}{aT} [V(0) - V_L]$$

- The longer the life of the option, the closer this is to $V_L$.
- Define $\sigma(T)$ as the volatility per annum that should be used to price a $T$-day option under GARCH(1,1). Assuming 252 days per year, $\sigma(T)^2$ is 252 times the average variance rate per day or

$$\sigma(T)^2 = 252 \left[ V_L + \frac{1 - e^{-aT}}{aT} [V(0) - V_L] \right] \tag{11}$$
The market prices of different options on the same asset are often used to calculate a volatility term structure. This is the relationship between the implied volatilities of the options and their maturities.

Equation (11) can be used to estimate a volatility term structure based on the GARCH(1,1) model. The estimated volatility term structure is not usually the same as the actual volatility term structure. However, it is often used to predict the way that the actual volatility term structure will respond to volatility changes.

When the current volatility is above the longer-term volatility, GARCH(1,1) model estimates a downward-sloping volatility term structure. When the current volatility, it estimates an upward-sloping volatility term structure. In the case of the yen-dollar exchange rate $a = \ln(1/0.9602) = 0.0406$ and $V_L = 0.00004422$. Suppose that the current variance rate per day, $V(0)$ is estimated as 0.00006 per day. It follows from equation (11) that

$$\sigma(T)^2 = 252 \left[ 0.00004422 + \frac{1 - e^{-0.0406T}}{0.0406T} [0.00006 - 0.00004422] \right]$$

where $T$ is measured in days. we find

<table>
<thead>
<tr>
<th>Option life (days)</th>
<th>10</th>
<th>20</th>
<th>50</th>
<th>100</th>
<th>500</th>
</tr>
</thead>
<tbody>
<tr>
<td>Option volatility(% per year)</td>
<td>12.00</td>
<td>11.59</td>
<td>11.33</td>
<td>11.00</td>
<td>10.65</td>
</tr>
</tbody>
</table>

for different values of $T$. 
Impact of Volatility Changes

Equation (11) can be written as

\[ \sigma(T)^2 = 252 \left[ V_L + \frac{1 - e^{-aT}}{aT} \left[ \frac{\sigma(0)^2}{252} - V_L \right] \right] \]

When \( \sigma(0) \) changes by \( \Delta \sigma(0) \), \( \sigma(T) \) changes by

\[ \frac{1 - e^{-aT}}{aT} \frac{\sigma(0)}{\sigma(T)} \Delta \sigma(0) \]  

(12)

We get a formula for the impact

<table>
<thead>
<tr>
<th>Option life (days)</th>
<th>10</th>
<th>20</th>
<th>50</th>
<th>100</th>
<th>500</th>
</tr>
</thead>
<tbody>
<tr>
<td>Increase in volatility (% per year)</td>
<td>0.84</td>
<td>0.61</td>
<td>0.46</td>
<td>0.27</td>
<td>0.06</td>
</tr>
</tbody>
</table>
Correlations

Correlations are important, as seen from last week, for computing VaR. We show how correlation estimates can be updated in a similar way as volatility estimates.

The correlation between two variables $X$ and $Y$ can be defined by

$$\frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y}$$

where $\sigma_X$ is the standard deviation of $X$, $\sigma_Y$ is the standard deviation of $Y$, and $\text{cov}(X, Y)$ is the covariance between $X$ and $Y$. The covariance between $X$ and $Y$ is defined as

$$E[(X - \mu_X)(Y - \mu_Y)]$$

where $\mu_X$ and $\mu_Y$ are the means of $X$ and $Y$. Easier to develop methods for the covariances as opposed to the correlations.
Define \(x_i\) and \(y_i\) as the percentage changes in the values of \(X\) and \(Y\) between the end of day \(i - 1\) and the end of day \(i\):

\[
x_i = \frac{X_i - X_{i-1}}{X_{i-1}} \quad y_i = \frac{Y_i - Y_{i-1}}{Y_{i-1}}
\]

where \(X_i\) and \(Y_i\) are the values of \(X\) and \(Y\) at the end day \(i\).

We also define the following:

- \(\sigma_{x,n}\) - daily volatility of variable \(X\), estimated for day \(n\)
- \(\sigma_{y,n}\) - daily volatility of variable \(Y\), estimated for day \(n\)
- \(\text{cov}_n\) - daily covariance between daily changes in \(X\) and \(Y\), estimated for day \(n\)

Then we estimate the correlation between \(X\) and \(Y\) on day \(n\) as

\[
\frac{\text{cov}_n}{\sigma_{x,n}\sigma_{y,n}}
\]

We use an equal-weighting scheme and assuming that the means of \(x_i\) and \(y_i\) are zero, then we can estimate the variance rates of \(X\) and \(Y\) from the most recent \(m\) observations as

\[
\sigma_{x,n} = \frac{1}{m} \sum_{i=1}^{m} x_{n-i}^2 \quad \sigma_{y,n} = \frac{1}{m} \sum_{i=1}^{m} y_{n-i}^2
\]

A similar estimate for the covariance between \(X\) and \(Y\) is

\[
\text{cov}_n = \frac{1}{m} \sum_{i=1}^{m} x_{n-i} y_{n-i}
\]

(13)
One alternative for updating covariances is an EWMA model similar we find

\[ \text{cov}_n = \lambda \text{cov}_{n-1} + (1 - \lambda) x_{n-1}y_{n-1} \]

A similar analysis to that presented for the EWMA volatility model shows that the weights given observations on the \( x_i y_i \) decline as we move back through time. The lower the value of \( \lambda \) the greater the weight that is given to recent observations.
Once all the variances and covariances have been computed, a variance-covariance matrix can be constructed. When \( i \neq j \), the \((i, j)\) element of the matrix shows the covariance between variable \( i \) and variable \( j \). When \( i = j \) it shows the variance of variable \( i \).

Not all variance-covariance matrices are internally consistent. The condition for an \( N \times N \) variance-covariance matrix \( \Omega \) to be internally consistent is

\[
w^T \Omega w \geq 0
\]

for all \( N \times 1 \) vectors \( w \), where \( w^T \) is the transpose of \( w \). Such matrices are positive semi-definite.

To ensure that a positive-semidefinite matrix is produced, variances and covariances should be calculated consistently. For example, if variances are calculated by giving equal weight to the last \( m \) data items, the same should be done for covariances.
• In practice, volatility is a stochastic variable. Unlike asset price, it is not directly observable.
• We define $u_i$ to be the percentage change in a market variable between the end of day $i-1$ and the end of day $i$. The variance rate of the market variable is calculated as a weighted average of the $u_i^2$'s. The key feature of the schemes that have been discussed here is that they do not give equal weight to the observations on the $u_i$'s.
• The more recent an observation, the greater the weight assigned to it. In EWMA and GARCH(1,1) models, the weights assigned to observations decrease exponentially as the observations become older.
• The GARCH(1,1) model differs from the EWMA model in that some weight is also assigned to the long-run average variance rate. Both the EWMA and GARCH(1,1) models have structures that enable forecasts of the future level of variance rate to be produced relatively easily.
• Maximum likelihood methods are usually used to estimate parameters from historical data in GARCH(1,1) and similar models. These methods involve using an iterative procedure to determine the parameter values that maximize the chance or likelihood that the historical data will occur.
• Once its parameters have been determined, a mood can be judged by how well it removes autocorrelation from the $u_i$'s.
• For every model that is developed to track variances there is a corresponding model that can be developed to track covariances. The procedures can be used to update the complete variance-covariance matrix used in VaR calculations.
Credit Risk

Credit Risk arises from the probability that borrowers and counterparts in derivatives transactions may default.

We attempt to quantify the risk associated to credit risk.

Credit Ratings

- Moody’s and S&P provide ratings that describe the creditworthiness of corporate bonds.

<table>
<thead>
<tr>
<th>Moody’s Rating</th>
<th>S&amp;P Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aaa</td>
<td>AAA</td>
</tr>
<tr>
<td>Aa</td>
<td>AA</td>
</tr>
<tr>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>Baa</td>
<td>BBB</td>
</tr>
<tr>
<td>Ba</td>
<td>BB</td>
</tr>
<tr>
<td>B</td>
<td>B</td>
</tr>
<tr>
<td>Caa</td>
<td>CCC</td>
</tr>
</tbody>
</table>

Bonds with Aaa rating are considered to have little to no chance of default.

- Moody’s subdivides categories such as Aa to Aa1, Aa2, Aa3, etc.
- S&P subdivides categories such as AA to AA+, AA, AA−, etc.
- Only Aaa or AAA are not subdivided.
Historical Default Probabilities

We can consider the historical default rates of the certain class of corporate bonds:

For example:

- A bond with an initial credit rating of $A$ has a 0.23% chance of defaulting by the end of the third year.
- A bond with an initial credit rating of $C'aa$ has a 69.83% chance of defaulting by the end of the seventh year.

<table>
<thead>
<tr>
<th>Term (years)</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>7</th>
<th>10</th>
<th>15</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aaa</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.04</td>
<td>0.12</td>
<td>0.29</td>
<td>0.62</td>
<td>1.21</td>
<td>1.55</td>
</tr>
<tr>
<td>Aa</td>
<td>0.02</td>
<td>0.03</td>
<td>0.06</td>
<td>0.15</td>
<td>0.24</td>
<td>0.43</td>
<td>0.68</td>
<td>1.51</td>
<td>2.70</td>
</tr>
<tr>
<td>A</td>
<td>0.02</td>
<td>0.09</td>
<td>0.23</td>
<td>0.38</td>
<td>0.54</td>
<td>0.91</td>
<td>1.59</td>
<td>2.94</td>
<td>5.24</td>
</tr>
<tr>
<td>Baa</td>
<td>0.20</td>
<td>0.57</td>
<td>1.03</td>
<td>1.62</td>
<td>2.16</td>
<td>3.24</td>
<td>5.10</td>
<td>9.12</td>
<td>12.59</td>
</tr>
<tr>
<td>Ba</td>
<td>1.26</td>
<td>3.48</td>
<td>6.00</td>
<td>8.59</td>
<td>11.17</td>
<td>15.44</td>
<td>21.01</td>
<td>30.88</td>
<td>38.56</td>
</tr>
<tr>
<td>B</td>
<td>6.21</td>
<td>13.76</td>
<td>20.65</td>
<td>26.66</td>
<td>31.99</td>
<td>40.79</td>
<td>50.02</td>
<td>59.21</td>
<td>60.73</td>
</tr>
<tr>
<td>Caa</td>
<td>23.65</td>
<td>37.20</td>
<td>48.02</td>
<td>55.56</td>
<td>60.83</td>
<td>69.36</td>
<td>77.91</td>
<td>80.23</td>
<td>80.23</td>
</tr>
</tbody>
</table>

(Source: Moody’s)
We can compute the probability of default for a particular year from the table.

- The probability of a *Ba* bond defaulting in the third year is
  \[ 6.00 - 3.48 = 2.52\% \]

- The probability of a *Caa* bond defaulting between the fifth and seventh year is
  \[ 69.36 - 60.83 = 8.53\% \]

The default probabilities are an increasing function of time, since over time a company with strong credit may deteriorate. Companies with poor credit may default in a short period of time.

We define two quantities:

- **Unconditional default probability**: This is the probability of default during a particular year as seen from the initial year.

  **Example**: The unconditional default probability of a *Caa* bond defaulting in the third year is
  \[ 48.02 - 37.20 = 10.82\% \]
• **Default intensities** or **Hazard rates**: This is the probability of default during a particular year as seen from the initial year conditioned on no default occurring earlier.

**Example**: The probability that a *Caa* bond will last past year two is

\[ 100 - 37.20 = 62.80\% \]

Therefore, the default intensity is

\[ \frac{0.1082}{0.6280} = 17.23\% \]

If we instead compute the default intensity \( \lambda(t) \) at time \( t \) over a shorter length of time \( \Delta t \). Then \( \lambda(t) \Delta t \) is the probability of default between time \( t \) and time \( t + \Delta t \) conditional on no earlier default.
Default Intensities

If $V(t)$ is the cumulative probability of the company surviving to time $t$, then

$$V(t + \Delta t) - V(t) = -\lambda(t)V(t)\Delta t$$

Taking limits we get

$$\frac{dV(t)}{dt} = e^{-\int_0^t \lambda(\tau)d\tau}$$

Define $Q(t)$ as the probability of default by time $t$. It follows that

$$Q(t) = 1 - e^{-\int_0^t \lambda(\tau)d\tau}$$

or

$$Q(t) = 1 - e^{-\overline{\lambda(t)}t} \quad (14)$$

where $\overline{\lambda(t)}$ is the average default intensity between time 0 and time $t$. 
Recovery Rates

- When a company goes bankrupt, those that are owed money by the company file claims against the assets of the company.
- Either the company reorganizes and creditors agree to partial payments or the assets are liquidated and used to meet outstanding claims.
- Recovery rates for a bond is normally defined as the bond’s market value immediately after a default, as a percent of the face value.
- Senior secured debt holders receive 51 cents to the dollar owed on average; whereas junior subordinated debt holders receive 25 cents to the dollar owed on average.

<table>
<thead>
<tr>
<th>Class</th>
<th>Average recovery rate %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Senior secured</td>
<td>51.6</td>
</tr>
<tr>
<td>Senior unsecured</td>
<td>36.1</td>
</tr>
<tr>
<td>Senior subordinated</td>
<td>32.5</td>
</tr>
<tr>
<td>Subordinated</td>
<td>31.1</td>
</tr>
<tr>
<td>Junior subordinated</td>
<td>24.5</td>
</tr>
</tbody>
</table>

- Recovery rates are significantly negatively correlated with default rates.
It found that the following relationship provides a good fit to the data:

\[
\text{Average recovery rate} = 50.3 - 6.3 \times \text{Average default rate}
\]

where both the average recovery rate and the average default rate are measured as percentages.

**Estimating Default Probabilities from Bond Prices**

- Probabilities of default can be estimate from the prices of bonds it has issued.
- Assume that the reason a corporate bond sells for less than a risk-free bond is the possibility of default.
- Consider an approximate calculation. Suppose that a bond yields 200 basis points more than a similar risk-free bond and that the expected recovery rate in the event of default is 40%.
- The holder of a corporate bond must be expecting to lose 200 basis points (2% per year) from defaults. Given the recovery rate of 40%, this leads to an estimate of the probability of a default per year conditional on no earlier default of
  \[
  \frac{0.02}{1-0.4} = 2.22\%.
  \]
- In general
  \[
  h = \frac{s}{1 - R}
  \]
where \( h \) is the default intensity per year, \( s \) is the spread of the corporate bond yield over the risk-free rate, and \( R \) is the expected recovery rate.
A more exact calculation, suppose that the corporate bond we have been considering lasts for 5 years, provides a coupon 6% per annum (paid semiannually) and that the yield on the corporate bond is 7% per annum (with continuous compounding).

- The yield on a similar risk-free bond is 5% (with continuous compounding).
- The yields imply that the price of the corporate bond is 95.34 and the price of the risk-free bond is 104.09.
- The expected loss from default over the 5-year life of the bond is therefore 104.09-95.34 = 8.75.
- Suppose that the probability of default per year (assumed in this simple example to be the same each year) is $Q$.

Consider for example the loss of default at 3.5 years. The expected value of the risk-free bond at time 3.5 years (using forward interest rates) is

$$3 + 3e^{-0.05\times0.5} + 3e^{-0.05\times1.0} + 3e^{-0.05\times1.5} = 104.34$$

Given the default of recovery rates in the previous section, the amount recovered if there is a default is 40, so that the loss given default is 104.34-40 or $64.34.$
The present value of this loss is 54.01. The expected loss is therefore $54.01Q$.

### Table 20.3 Calculation of loss from default on a bond in terms of the default probabilities per year, $Q$. Notional principal = $100.

<table>
<thead>
<tr>
<th>Time (years)</th>
<th>Default probability</th>
<th>Recovery amount ($)</th>
<th>Risk-free value ($)</th>
<th>Loss given default ($)</th>
<th>Discount factor</th>
<th>$PV$ of expected loss ($)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>$Q$</td>
<td>40</td>
<td>106.73</td>
<td>66.73</td>
<td>0.9753</td>
<td>65.08 $Q$</td>
</tr>
<tr>
<td>1.5</td>
<td>$Q$</td>
<td>40</td>
<td>105.97</td>
<td>65.97</td>
<td>0.9277</td>
<td>61.20 $Q$</td>
</tr>
<tr>
<td>2.5</td>
<td>$Q$</td>
<td>40</td>
<td>105.17</td>
<td>65.17</td>
<td>0.8825</td>
<td>57.52 $Q$</td>
</tr>
<tr>
<td>3.5</td>
<td>$Q$</td>
<td>40</td>
<td>104.34</td>
<td>64.34</td>
<td>0.8395</td>
<td>54.01 $Q$</td>
</tr>
<tr>
<td>4.5</td>
<td>$Q$</td>
<td>40</td>
<td>103.46</td>
<td>63.46</td>
<td>0.7985</td>
<td>50.67 $Q$</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>288.48 $Q$</strong></td>
</tr>
</tbody>
</table>

The total expected loss is $288.48Q$. Setting this equal to 8.75, we obtain a value of $Q$ equal to 3.03%. The calculations we have give assume that the default probability is the same in each year and that defaults take place at just one time during the year. We can extend the calculations to assume that defaults take place at just one time during the year.

- The calculations assume that the default probability is the same each year and that defaults take place at just one time during the year. We can extend the calculations to assume that defaults can take place more frequently.
- Instead of assuming a constant unconditional probability of default we can assume a constant default intensity or assume a particular pattern for the variation of default probabilities with time.
With several bonds we can estimate several parameters describing the term structure of default probabilities.

Suppose we have bonds maturing in 3, 5, 7, and 10 years.

● We could use the first bond to estimate the default probability per year for the first 3 years.
● We could use the second bond to estimate default probability for years 4 and 5.
● Etc.
● The approach is analogous to the bootstrap procedure for calculating the zero-coupon yield curve.
A key issue when bond prices are used to estimate default probabilities is the meaning of the terms "risk-free rate" and "risk-free bond".

From equation (15), the spread $s$ is the excess of the corporate bond yield over the yield on a similar risk-free bond.

In our previous table, the risk-free value of the bond must be calculated using the risk-free rate.

Benchmark risk-free rate this is usually used in quoting corporate bond yields is the yield on similar Treasury bonds.

Traders usually use LIBOR/swap rates as proxies for risk-free rates instead of the Treasury rates when valuing derivatives. Also use these rates as risk-free rates when calculating default probabilities.

When determining default prob. from bond prices, the spread $s$ in (15) is the spread of the bond yield over the LIBOR/swap rate.

The risk-free discount rates used in the table before are LIBOR/swap zero rates.

Credit default swaps (to be discussed next week) can be used to imply the risk-free rate assumed by traders. The rate used appears to be approximately equal to the LIBOR/swap rate minus 10 basis points on average.
In practice traders often use **asset swap** spreads as a way of extracting default probabilities from bond prices. This is because asset swap spreads provide a direct estimate of the spread of bond yields over the LIBOR/swap curve.

Consider how an **asset swap** works. Consider the situation where an asset swap spread for a particular bond is quoted as 150 basis points. There are three possible situations:

1. The bond sells for its par value of 100. The swap then involves one side (company A) paying the coupon on the bond and the other side (company B) paying LIBOR plus 150 basis points.
2. The bond sells below its par value, say, for 95. The swap is then structured so that in addition to the coupons company A pays $5 per $100 of notional principal at the outset.
3. The underlying bond sells above par, say, for 108. Company B would then make a payment of $8 per $100 of principal at the outset.

The effect of this is that the present value of the asset swap spread is the amount by which the price of the corporate bond is exceeded by the price of a similar risk-free bond where the risk-free rate is assumed to be given by the LIBOR/swap curve.
Consider again our example where the LIBOR/swap zero curve is flat at 5%.

- Suppose that instead of knowing the bond’s price we know that the risk-free bond exceeds the value of the corporate bond is the present value of 150 basis points per year for 5 years.
- Assuming semiannual payments, this is $6.55 per $100 of principal.
- Total loss in this case would be set equal to $6.55. This means that the default probability per year $Q$ would be $\frac{6.55}{288.48} = 2.27\%$. 
Comparison of Default Probability Estimates

The default probabilities estimate from historical data are much less than those derived from bond prices.

<table>
<thead>
<tr>
<th>Rating</th>
<th>Historical default intensity</th>
<th>Default intensity from bonds</th>
<th>Ratio</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aaa</td>
<td>0.04</td>
<td>0.67</td>
<td>16.8</td>
<td>0.63</td>
</tr>
<tr>
<td>Aa</td>
<td>0.06</td>
<td>0.78</td>
<td>13.0</td>
<td>0.72</td>
</tr>
<tr>
<td>A</td>
<td>0.13</td>
<td>1.28</td>
<td>9.8</td>
<td>1.15</td>
</tr>
<tr>
<td>Baa</td>
<td>0.47</td>
<td>2.38</td>
<td>5.1</td>
<td>1.91</td>
</tr>
<tr>
<td>Ba</td>
<td>2.40</td>
<td>5.07</td>
<td>2.1</td>
<td>2.67</td>
</tr>
<tr>
<td>B</td>
<td>7.49</td>
<td>9.02</td>
<td>1.2</td>
<td>1.53</td>
</tr>
<tr>
<td>Caa</td>
<td>16.90</td>
<td>21.30</td>
<td>1.3</td>
<td>4.40</td>
</tr>
</tbody>
</table>

For companies that start with a particular rating, the average annual default intensity over 7 years calculated from (a) historical data and (b) bond prices.
The calculation of default intensities, using historical data are based on (14) and our table of cumulative defaults on bond types. We have for example:

\[ \bar{\lambda}(7) = -\frac{1}{7} \ln [1 - Q(7)] \]

where \( \bar{\lambda}(t) \) is the average default intensity (or hazard rate) by time \( t \) and \( Q(t) \) is the cumulative probability of default by time \( t \).

- The values of \( Q(7) \) are taken directly from our table.
- Consider, for example, an A-rated company. The value of \( Q(7) \) is 0.0091. The average 7-year default intensity is 0.13% since

\[ \bar{\lambda}(7) = -\frac{1}{7} \ln(0.9909) = 0.0013 \]

- The calculations on bond prices are based on (15) and bond yields from Merrill Lynch.
- The recovery rate is assumed to be 40% and the risk-free interest rate is assume to be the 7-year swap rate minus 10 basis points (the 10 basis point subtraction is from empirical data)
- For example, for A-rated bonds, the average Merrill Lynch yield was 6.274%. The average swap rate was 5.605%, so that the average risk-free rate was 5.505%. This gives the average 7-year default probability as

\[ \frac{0.06274 - 0.05505}{1 - 0.4} = 0.0128 \]

or 1.28%. 

Remarks

From our chart we see

- The ratio of default probability backed out of bond prices to the default calculated from historical data tends to decline as the credit quality declines with the ratio very high for investment grade companies.
- The difference between the two default probabilities tends to increase as credit quality declines.

Real-World vs. Risk-Neutral Probabilities

Default probabilities implied from bond yields are risk-neutral probabilities of default. To explain why, consider the calculations of default probabilities of $Q$ table.

- These calculations assume that expected default losses can be discounted at the risk-free rate
- The risk-neutral valuation principle shows that this is a valid procedure providing the expected losses are calculated in a risk-neutral world. This means that the default probability $Q$ must be a risk-neutral probability.

On the other hand default probabilities implied from historical data are real-world default probabilities or **physical probabilities**. The expected excess return between real-world and risk-neutral default probabilities.
Real-World vs. Risk-Neutral Probabilities

Why do we see such big differences between real-world and risk-neutral default probabilities?

1. Corporate bonds are relatively illiquid and bond traders demand an extra return to compensate for this.
2. The subjective default probabilities of bond traders may be much higher than those given in our first table. Bond traders may be allowing for depression scenarios much worse than anything seen during the period from 1970 to 2003.
3. Bonds do not default independently of each other. This is the most important reason for the results in our last table. There are periods of time when default rates are very low and periods of time when they are very high. This gives rise to systematic risk (i.e. risk that cannot be diversified away) and bond traders should require an expected excess return for bearing the risk. The variation in default rates from year to year may be due to overall economic conditions or it may be because a default by one company has a ripple effect resulting in defaults by other companies. (The latter is referred to by researchers as **credit contagion**).
4. Bond returns are highly skewed with limited upside. As a result it is much more difficult to diversify risks in a bond portfolio than in an equity portfolio. A very large number of different bonds must be held. In practice many bond portfolios are far from fully diversified. As a result bond traders may require an extra return for bearing unsystematic risk in addition to the systematic risk mentioned above.
In order to get the default probabilities we rely on the company’s credit rating. Unfortunately, credit ratings are revised relatively infrequently. This leads analysts to argue that equity prices can provide more up-to-date information for estimating default probabilities.

Merton proposed a model where a company’s equity is an option on the assets of the company. Suppose that a firm has one zero-coupon bond outstanding and that the bond matures at time $T$. Define

$V_0$ Value of company’s assets today
$V_T$ Value of company’s assets at time $T$
$E_0$ Value of company’s equity today
$E_T$ Value of company’s equity at time $T$
$D$ Amount of debt interest and principal due to be repaid at time $T$
$\sigma_V$ Volatility of assets (assumed constant)
$\sigma_E$ Instantaneous volatility of equity
Consider two cases:

- If $V_T < D$, it is (at least in theory) rational for the company to default on the debt at time $T$. The value of the equity is then zero.
- If $V_T > D$, the company should make the debt repayment at time $T$ and the value the equity at this time is $V_T - D$.

Merton’s model, therefore, gives the value of the firm’s equity at time $T$ as

$$E_T = \max\{V_T - D, 0\}$$

This shows that the equity is a call option on the value of the assets with a strike price equal to the repayment required on the debt.

Black-Scholes formula gives the value of the equity today as

$$E_0 = V_0 N(d_1) - De^{-rT}N(d_2)$$

(16)

where

$$d_1 = \frac{\ln \frac{V_0}{D} + (r + \frac{\sigma^2}{2})T}{\sigma V \sqrt{T}} \quad d_2 = d_1 - \sigma V \sqrt{T}$$

The value of the debt today is $V_0 - E_0$. 
• The risk-neutral probability that the company will default on the debt is

\[ N(-d_2) \]

• To calculate this, we require \( V_0 \) and \( \sigma_V \). Neither of these are directly observable. However if the company is publicly traded, we can observe \( E_0 \).

• This means (16) provides one condition that must be satisfied by \( V_0 \) and \( \sigma_V \).

• We can also estimate \( \sigma_E \). From Itô's lemma

\[ \sigma_E E_0 = \frac{\partial E}{\partial V} \sigma_V V_0 \]

or

\[ \sigma_E E_0 = N(d_1) \sigma_V V_0 \] (17)

This provides another equation that must be satisfied by \( V_0 \) and \( \sigma_V \).

This provides two nonlinear equations (16) and (17) to get \( V_0 \) and \( \sigma_V \).
Example

The value of a company’s equity is $3 million and the volatility of the equity is 80%.

• The debt that will have to be paid in 1 year is $10 million. The risk-free rate is 5% per annum. Thus
  \[ E_0 = 3 \quad \sigma_E = 0.8 \quad r = 0.05 \quad T = 1 \quad D = 10 \]

• Solving (16) and (17) yields
  \[ V_0 = 12.40 \quad \sigma_V = 0.2123 \]

The parameter \( d_2 = 1.1408 \), so that the probability of default is \( N(-d_2) = 0.127 \) or 12.7%.

• The market value of the debt is \( V_0 - E_0 \) or 9.40. The present value of the promised payment on the debt is
  \[ 10e^{-0.05 \times 1} = 9.51 \]

• The expected loss on the debt is therefore, \( \frac{9.51 - 9.40}{9.51} = 1.2\% \) of its no-default value.

• Comparing this with the probability of default gives the expected recovery in the event of a default as \( \frac{12.7 - 1.2}{12.7} = 91\% \).
Midterm: March 12th
Homework: Due Mar. 5, 5PM.
Graded:

- Problem 19.15, 19.16, 19.18