Lecture Notes on Economics of Financial Risk Management

Xiaodong Zhu

March 20, 2011

1Incomplete draft for class uses only. Please do not circulate or cite without the author’s permission.

2Professor, Department of Economics, University of Toronto. Email: xzhu@chass.utoronto.ca
Chapter 1

Introduction

This course is about financial risk management. We will focus our discussion on why, when and where there is a need for risk management and how to measure and manage risk.

According to the Webster's New World Dictionary, risk is the *chance of injury, damage, or loss*. By this definition, humankind has always faced and had to deal with many kinds of risk such as war, famine, deadly diseases, bad weather, theft, robbery, etc.. Ideally, we would hope that we can eliminate the sources of risk so that we can live in a risk-free world. There has been some progress towards this goal. For example, most developed countries have almost certainly eliminated the possibility of famine. However, we are still living in a world that is far from risk-free. Despite enormous progress in many aspects of our daily life in the last few hundred years, risk is still everywhere in our life. A few examples:

- A firm may be taken over by another firm and its workers may be laid off.
- An earthquake or hurricane may hit a city.
- An electricity blackout may happen to a large region.
- A student may not be able to find a good job upon graduation.
- Your computer may be infected by a deadly virus or worm.
- Weather may be unexpectedly bad.
- Stock market may crash
• Tuition may increase
• US dollar may devalue dramatically against the Euro

When risk cannot be eliminated, the best way to deal with it is sharing it between lucky and unlucky ones. A major role of financial markets is to facilitate more efficient risk sharing among individuals, firms and governments. Financial innovations often occur to deal with new kinds of risk or to provide new ways to deal with old risk. Over the last three decades, financial innovations have been occurring at a frantic pace. Many new financial institutions, markets and products have been created. Markets for derivatives, in particular, exploded in the late 1980s and early 1990s. By 2001, the notional value of all derivatives outstanding amounted to $111 trillion. In comparison, the world GDP was only $31.3 trillion, and the capitalization of the stock markets of the industrialized countries was $21.2 trillion. The explosive growth in derivatives has made many people rich, and the market attracted talented individuals with very different backgrounds. Many Ph.D.s in economics, geophysics, chemical engineering, mathematics, computer science and other fields went to work on Wall Street. New professional programs in Mathematical Finance, Financial Engineering or Computational Finance have also been created in many universities.

With all the new derivatives that were invented or re–invented, one would expect that we should be much better at managing risk. Yet, today we still can not effectively manage many of the risk mentioned in the examples above. In the meantime, we heard all these news about spectacular losses associated with trading derivatives: Barings Bank, Orange County Government, Pocter and Gamble, Long Term Capital Management, etc. How does financial innovation help us to manage risk? Does financial innovation help us to manage risk? Could financial innovation create risk? What causes financial innovation? Should the society limit financial innovation? These are the central questions that we will try to address. To put things in perspective, it is useful to first take a brief look at the history of financial innovation.

1.1 A Brief History of Financial Innovation

Financial innovation is not a new phenomenon. Probably the earliest and one of the most important financial innovation is personal loan. A loan from one person to another can help both people to share risk and smooth their
income and expenditures over time. There is evidence that loans were used in many ancient civilizations. Banks that took deposits and lent money were established in Babylonia, Assyria and ancient Greece. They helped people to share risk and smooth their income and expenditures at a much larger scale.

Equities were issued by joint stock companies during the sixteenth century. Owners of equities could be those who work in the company or someone who simply provided financial capital. Issuing equities to “sleep partners” who did not actively involve in day–to–day operations allowed separation of ownership and control. This was very important because it allowed profit and risk sharing between those who knew how to produce and those who had money.

Bonds were first used by European governments in the sixteenth century. Collecting taxes was costly and time consuming, and it could be very disruptive to the economy if a government had to finance all of its war expenditures by levying a very high tax during the wartime. To smooth the tax burden over time, governments often financed their war expenditures by issuing bonds to the public. Today governments are still the biggest players in the bond market.

Initially, equities and bonds were mainly held by corporate insiders and wealthy investors. As the total amount of these securities became larger, more investors were needed to share the risk. This gave rise to secondary trading, first informally on the streets, latter in organized exchanges. The first exchange was the Amsterdam Stock Exchange opened in 1611, then the stock exchange in London in 1802 and New York Stock exchange in 1817 (known as New York Stock and Exchange Board then). Commodity futures exchanges emerged in Chicago in the 19th century. In the 1840s, Chicago became the commercial center for the Midwestern farm states. (The name of the basketball team Chicago Bulls originated from the fact that Chicago was a main distributor of beef in those days.) To make the trading of agriculture products more efficient, the Chicago Board of Trade (CBOT) was established in 1848. The establisher’s original purpose was to standardize the quantities and qualities of the grain being traded. However, contracts for the future delivery of grain started to trade soon after the establishment of the exchange. These allowed producers and processors to share the risk of fluctuations in prices.

The development of markets for financial derivatives is more recent. Financial futures and options on foreign exchange and interest rates were introduced and traded in the early 1970, and the futures on stock market indices
were introduced in the early 1980s. The financial innovation in this period is a response to the increasing volatility in the financial markets. There are three main reasons for the increasing volatility: (1) The oil crises in the early 1970s and early 1980s that directly and indirectly resulted in the significant drops in the stock prices and increases in commodity price volatility. (2) The increased inflation and interest rate risk. (3) The breakdown of Bretton Woods System and the end of fixed exchange rates for most industrialized countries.
Chapter 2

Arrow-Debreu Theory of Financial Markets

In discussing uncertainty and risk, two concepts are very important: states of nature and state contingent plan (contingent means depending on some outcome that is not yet known), which we define below:

- **states of nature**, which refer to different outcomes of some random event and

- **state contingent plan**, which specifies how much to be consumed (consumption plan), how much income to have (income plan), or what payoff to have (payoff plan) in each state of nature.

**Example 1.** Consider someone who has $100 and is considering whether to buy a Number 13 lottery ticket or not. The cost of the ticket is $5, and the winner receives $200. In this example, there are two states of nature: (1) number 13 is drawn and (2) number 13 is not drawn. If the individual does not buy the lottery ticket, then her plan is to have $100 in both states, which can also represented by a two-dimensional vector (100, 100). Here the first element of the vector represents the individual’s income in the state when number 13 is drawn and the second element the individual’s income in the other state. If the individual buys the ticket, then, her plan is

\[
\begin{align*}
295 & \quad \text{number 13 is drawn,} \\
95 & \quad \text{number 13 is not drawn.}
\end{align*}
\]
Alternatively, we can represent the plan as $(295, 95)$. We can also represent the payoffs of the lottery ticket as a payoff plan

$$\text{Lottery ticket} = \begin{cases} 200 & \text{number 13 is drawn,} \\ 0 & \text{number 13 is not drawn.} \end{cases}$$

or simply $(200, 0)$.

**Example 2.** Consider someone who has $35,000 initial wealth. There is a possibility that an accident will occur such that the individual will lose $10,000, maybe as a result of car accident or fire in the house. In this example, the two states of nature are "there is an accident" and "no accident". The individual’s wealth in this example is also a state-contingent plan:

$$\begin{cases} 25,000 & \text{there is an accident,} \\ 35,000 & \text{no accident,} \end{cases}$$

or $(25000, 35000)$. The individual can change her wealth plan by buying an insurance plan. Consider an insurance plan that pays $K$ dollars in the state of accident costs $0.01K$ dollars. If the individual buys the insurance plan, then her wealth after insurance would be

$$\begin{cases} 25,000 + K - 0.01K & \text{there is an accident,} \\ 35,000 - 0.01K & \text{no accident} \end{cases}$$

or $(25000 + 0.99K, 35000 - 0.01K)$. For example, if the individual pays $100 in both states of nature ($K = 10,000$), then she will get $10,000 from the insurance in the state when there is an accident and nothing when there is no accident. As a result, the individual’s wealth after insurance would be

$$\begin{cases} 34,900 & \text{there is an accident,} \\ 34,900 & \text{no accident} \end{cases}$$

or $(34900, 34900)$. We can graph all the feasible wealth plans the individual can have through insurance in Figure 1. Note that in the graph, we allow for the possibility that the amount of insurance $K$ to be negative. That is, the individual can have consumption plans under which she consumes less than $25,000 in the state of accident. In such cases, the individual is effectively short or selling insurance.

In both Example 1 and 2, we consider the cases when there are only two possible states, so any state contingent plan can be represented by a
two-dimensional vector. Similarly, for any positive integer \( N \), if there are \( N \) possible states, we can also represent a state contingent plan by an \( N \) dimensional vector. (See the appendix at the end of this chapter for a brief introduction to vectors and linear algebra.) So, a consumption plan can be represented by a vector \((c_1, c_2, ..., c_N)\), where \( c_i \) is how much to be consumed if the \( i \)th state is realized.

A financial security can also be represented by a vector \( a = (a_1, ..., a_N) \), where \( a_i \) is the security’s payoff in state \( i \) for \( i = 1, ..., N \). For example, \((1, ..., 1)\) represents a risk–free security that pays one unit of good in every state, and \((-1, -1, ..., 2, 2)\) a risky security that has negative payoffs in some states and positive payoffs in some other states.

**Definition** The financial market is **complete** if all financial securities can be traded. That is, for any arbitrary \( N \) numbers \( a_1, ..., a_N \), a security with payoff plan \( a = (a_1, ..., a_N) \) can be bought or sold in the market.

We denote the price of a security \( a \) by \( P(a) \). For any two securities \( a = (a_1, ..., a_N) \) and \( b = (b_1, ..., b_N) \), a third security can be formed by a portfolio of these two securities: \( \alpha a + \beta b \), where \( \alpha \) and \( \beta \) are any two real numbers. If \( \alpha \) (or \( \beta \)) is a negative number, then the portfolio includes short–selling (which means selling a security without actually owning it) rather than holding security \( a \) (or \( b \)). There is no arbitrage in the financial market if for any \( \alpha \) and \( \beta \), we have

\[
P(\alpha a + \beta b) = \alpha P(a) + \beta P(b) \tag{2.0.1}
\]

That is, if the payoff of a security can be replicated by a portfolio of some other securities, then the price of this security \((P(\alpha a + \beta b))\) should be the same as the cost of forming the portfolio \((\alpha P(a) + \beta P(b))\).

In the rest of this chapter, we will always assume that there is no-arbitrage unless stated otherwise.

Theoretically, it is useful to think of the following basic securities,

\[
\begin{align*}
b^{(1)} &= (1, 0, ..., 0) \\
b^{(2)} &= (0, 1, ..., 0) \\
&\vdots \\
b^{(N)} &= (0, ..., 0, 1)
\end{align*}
\]

Here, \( b^{(i)} \) is the security that pays one unit in state \( i \) and nothing in any other states.
Note that any security \( a = (a_1, ..., a_N) \) can be constructed as a portfolio of these basic securities:

\[
a = a_1 b(1) + ... + a_N b(N).
\]

Therefore, the market is complete if these basic securities and the portfolios of these basic securities can be traded in the market. Under the no-arbitrage condition, then, we have

\[
P(a) = a_1 P(b(1)) + ... + a_N P(b(N)). \tag{2.0.2}
\]

In the real world, securities that are traded are usually not basic securities. To determine whether the market is complete or not, then, all we need to know is whether the basic securities can be replicated by portfolios of securities that are traded in the market.

**Example 3.** There are only two states. There are two securities that are traded in the market: a risk-free bond \( B = (1, 1) \) and a risky stock \( A = (1, 2) \). Then, the basic securities can be formed as follows:

\[
\begin{align*}
b(1) &= 2B - A = (1, 0) \\
b(2) &= A - B = (0, 1)
\end{align*}
\]

that is, \( b(1) \) can be replicated by holding two units of \( B \) and short-selling one unit of \( A \), and \( b(2) \) can be replicated by holding one unit of \( A \) and short-selling one unit of \( B \). Thus, the market is complete in this example.

**Example 4.** There are three states. Again, there are two securities that are traded in the market: a risky stock \( A = (1, 2, 3) \) and a risk-free bond \( B = (1, 1, 1) \). In this example, the market is not complete since no portfolio of \( A \) and \( B \) can replicate the basic security \( b(1) \). To see this, let \( \alpha A + \beta B \) be a portfolio that replicates the payoff of \( b(1) \) in the first two states, that is

\[
\begin{align*}
\alpha + \beta &= 1, \\
2\alpha + \beta &= 0,
\end{align*}
\]

which implies that \( \alpha = -1 \) and \( \beta = 2 \). The payoff of this portfolio in the third state, then, is \( -1 \times 3 + 2 \times 1 = -1 \), which is different from zero payoff of \( b(1) \). So, there exists no combination of \( \alpha \) and \( \beta \) such that \( \alpha A + \beta B = b_1 \).
Remark: Example 4 is a special example of a more general property: For a market to be complete in an economy with $N$ states, there should be at least $N$ securities traded in the market that have linearly independent payoffs (i.e., the matrix that is formed by stacking the $N$ securities’ payoffs is invertible or nonsingular).

Example 5. The same environment as in example 4 except that derivative securities based on $A$ can be traded now. In this case, the basic security $b_{(3)}$ can be replicated by an option on the stock $A$. Let $C$ be the option on $A$ with strike price 2. That is, the holder of the option has the right, but not obligation, to buy one unit of $A$ with price 2 in any states. Apparently, the holder would buy the stock only in the state in which the payoff of $A$ is greater than 2, which is the third state. Thus, the payoff of the option is

$$C = \max\{A - 2, 0\} = (0, 0, 1) = b_{(3)}.$$  

With the new security $C$, then, the basic securities can be replicated as follows:

$$b_{(1)} = 2B - A + C,$$
$$b_{(2)} = A - B - 2C,$$
$$b_{(3)} = C.$$  

Thus, the originally incomplete market is completed with trading of options.

Appendix: Linear Algebra

2.0.1 Some Basic Definitions

Matrix—an array of numbers:

$$A \equiv [a_{ij}]_{m \times n} \equiv \begin{bmatrix}
    a_{11} & a_{12} & \ldots & a_{1n} \\
    a_{21} & a_{22} & \ldots & a_{2n} \\
    \vdots & \vdots & \ddots & \vdots \\
    a_{m1} & a_{m2} & \ldots & a_{mn}
\end{bmatrix}$$

Examples:
identity matrix

\[ I = \begin{bmatrix}
1 & 0 & \ldots & 0 \\
0 & 1 & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & 1
\end{bmatrix} \]

scalar matrix

\[ \begin{bmatrix}
c & 0 & \ldots & 0 \\
0 & c & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & c
\end{bmatrix} \]

2.0.2 Matrix Operations

Sum and Multiplication by a Scalar:

\[ A + B = [a_{ij} + b_{ij}] \]
\[ cA = [ca_{ij}] \]

some properties:

\[ A + B = B + A \]
\[ c(A + B) = cA + cB \]

column vector

\[ x = \begin{bmatrix}
x_1 \\
x_2 \\
\vdots \\
x_n
\end{bmatrix} \]

row vector

\[ y' = \begin{bmatrix}
y_1 \\
y_2 \\
\vdots \\
y_n
\end{bmatrix} \]
multiplication of vectors

\[ y^t x = \sum_{k=1}^{n} y_k x_k \]

multiplication of matrices

\[ A = [a_{ij}]_{m \times n}, \quad B = [b_{ij}]_{n \times p}, \quad C \equiv AB \equiv [c_{ij}]_{m \times p} \]

\[ A = \begin{bmatrix}
    a'_{(1)} \\
    a'_{(2)} \\
    \vdots \\
    a'_{(m)}
\end{bmatrix}, \quad B = \begin{bmatrix}
    b_1 & b_2 & \ldots & b_n
\end{bmatrix} \]

where

\[ a'_{(i)} = \begin{bmatrix}
    a_{i1} & a_{i2} & \ldots & a_{in}
\end{bmatrix} \]

\[ b_j = \begin{bmatrix}
    b_{1j} \\
    b_{2j} \\
    \vdots \\
    b_{nj}
\end{bmatrix} \]

\[ c_{ij} = a'_{(i)} b_{(j)} = \sum_{k=1}^{n} a_{ik} b_{kj}. \]

Some Properties:

(i) In general, \( AB \neq BA \).
(ii) \( IA = AI = A \).
(iii) \( Ax = x_1 a_1 + \ldots + x_n a_n \).

2.0.3 Inverse of a Matrix

Vectors \( a_1, \ldots, a_n \) are linearly independent if and only if none of them can be written as a linear combination of the others. Equivalently,

\[ x_1 a_1 + \ldots + x_n a_n = 0 \implies x_1 = \ldots = x_n = 0. \]
Proof: if one of the $x_i$'s is not zero, say $x_1 \neq 0$, then

$$a_1 = x_1^{-1}x_2a_2 + ... + x_1^{-1}x_na_n,$$

which is contradictory to the assumption.

A square matrix $A$ is nonsingular or invertible if there exists a square matrix $B$ such that

$$AB = BA = I, \quad B \equiv A^{-1}.$$

(i) If $A$ is nonsingular, then its column vectors are linearly independent.

$$Ax = 0 \implies A^{-1}(Ax) = (A^{-1}A)x = x = 0.$$

(ii) If $A$ is nonsingular, then equation $Ax = y$ has a unique solution.

$$Ax = y \implies A^{-1}(Ax) = A^{-1}y \implies x = A^{-1}y.$$

(iii) If $A$ is nonsingular, then the determinant of $A$ is nonzero, $det(A) \neq 0$

### 2.0.4 Determinant of a Matrix

For a two dimensional matrix,

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

we have

$$det(A) = ad - cb$$

For a three dimensional matrix, if it has the following form:

$$A = \begin{bmatrix} a & b & e \\ c & d & f \\ 0 & 0 & g \end{bmatrix}$$

then, we have

$$det(A) = det\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) \times g$$
Chapter 3

Applications of Arrow-Debreu Theory: Pricing Options

Even though Arrow-Debreu Theory is an abstract economic theory, it has very practical applications in finance. In this section we illustrate the usefulness of this theory by showing how it can be used for pricing options and other derivative contracts.

3.1 Pricing Options in a Simple Two-State Economy

A financial market is open now (time zero) for trading securities that pays off at some time in the future. The length of the time period between now and the payoff date is $\Delta t$. In finance, we always use year as the unit of time. So $\Delta t = 1, 1/2, 1/4,$ and $1/12$ represent one year, half a year, a quarter and a month, respectively. There are two securities that are traded in the market: $A$ is a stock and $B$ is a bond. The price of the stock today is $S$ ($S > 0$) and the price of the bond today is 1. There are two possible states of nature at $\Delta t$, and the stock’s value is state-contingent. It will equal $SD$ in the first state and $SU$ in the second state, where $D$ and $U$ are two constants such that $0 < D < 1 < U$. So, the first state is the state when the stock price goes down (because $SD < S$), and the second state is the state when the stock price goes up (because $SU > S$). In contrast, the value of the bond at time $\Delta t$ will be $e^{R\Delta t}$ in both states of nature. Here $R$ is the annualized risk-free interest rate. Figure 2 is a graphical representation of the two assets.
We can represent the two assets using the Arrow-Debreu language: \( A = (SD, SU) \), \( B = (e^{R\Delta t}, e^{R\Delta t}) \), \( P(A) = S \) and \( P(B) = 1 \). It is clear that \( A \) and \( B \) are not proportional to each other. That is, they are linearly independent. Therefore, we know that the market is complete. In particular, the two basic securities, \( b_1 = (1, 0) \) and \( b_2 = (0, 1) \) can be replicated by portfolios of \( A \) and \( B \).

**Using Stock and Bond Prices to Find the Prices of the Basic Securities**

Let’s first see how we can form a portfolio of \( A \) and \( B \), \( \alpha_1 A + \beta_1 B \), such that its payoffs are the same as \( b_1 \). In other words, we are looking for \( \alpha_1 \) and \( \beta_1 \) such that

\[
\begin{align*}
\alpha_1 SD + \beta_1 e^{R\Delta t} &= 1, \\
\alpha_1 SU + \beta_1 e^{R\Delta t} &= 0.
\end{align*}
\]

(3.1.1)

(3.1.2)

Subtracting (3.1.2) from (3.1.1) yields the following:

\[
\alpha_1 S(D - U) = 1
\]

which implies that

\[
\alpha_1 = \frac{1}{S(D - U)} = -\frac{1}{S(U - D)}.
\]

Substituting the equation above into (3.1.2), we have

\[
-\frac{1}{S(U - D)} SU + \beta_1 e^{R\Delta t} = 0
\]

or

\[
\beta_1 e^{R\Delta t} = \frac{U}{U - D},
\]

which implies that

\[
\beta_1 = \frac{e^{-R\Delta t} U}{U - D}.
\]

Thus, we have

\[
b_{(1)} = -\frac{1}{S(U - D)} A + \frac{e^{-R\Delta t} U}{U - D} B.
\]

(3.1.3)
3.1. **PRICING OPTIONS IN A SIMPLE TWO-STATE ECONOMY**

Similarly, to find the portfolio of $A$ and $B$, $\alpha_2 A + \beta_2 B$, that replicates the payoffs of $b_{(2)}$, we solve the following equations:

\begin{align}
\alpha_2 S D + \beta_2 e^{R \Delta t} &= 0, \\
\alpha_2 S U + \beta_2 e^{R \Delta t} &= 1.
\end{align}

Subtracting (3.1.4) from (3.1.5) yields the following:

$$\alpha_2 S (U - D) = 1$$

which implies that

$$\alpha_2 = \frac{1}{S (U - D)}.$$ 

Substituting the equation above into (3.1.4), we have

$$\frac{1}{S (U - D)} SD + \beta_2 e^{R \Delta t} = 0$$

or

$$\beta_2 e^{R \Delta t} = - \frac{D}{U - D},$$

which implies that

$$\beta_2 = \frac{e^{-R \Delta t} D}{U - D}.$$ 

Thus, we have

$$b_{(2)} = \frac{1}{S (U - D)} \alpha_2 P(A) + \beta_2 e^{-R \Delta t} B.$$  

(3.1.6)

From equation (3.1.3) and (3.1.6) we can determine the prices of $b_{(1)}$ and $b_{(2)}$ using the non-arbitrage equation in Chapter 2, (2.0.1). Thus, we have

$$P(b_{(1)}) = \alpha_1 P(A) + \beta_1 P(B)$$

$$= - \frac{1}{S (U - D)} S + \frac{e^{-R \Delta t} U}{U - D}$$

or

$$P(b_{(1)}) = \frac{e^{-R \Delta t} U - 1}{U - D}.$$  

(3.1.7)

Similarly,

$$P(b_{(2)}) = \alpha_2 P(A) + \beta_2 P(B)$$

$$= \frac{1}{S (U - D)} S - \frac{e^{-R \Delta t} D}{U - D}.$$
or
\[ P(b(2)) = \frac{1 - e^{-R\Delta t}D}{U - D}. \quad (3.1.8) \]

No-Arbitrage Constraints on Basic Security Prices

Since both basic securities pay off a positive amount in one state and nothing in the other state, the prices of these two securities have to be positive. Otherwise, everyone would want to buy an infinite amount of these securities and the market would not be cleared. So, to be consistent with no-arbitrage, we need the following conditions to hold:
\[ P(b(1)) = \frac{e^{-R\Delta t}U - 1}{U - D} > 0 \]
\[ P(b(2)) = \frac{1 - e^{-R\Delta t}D}{U - D} > 0 \]

The first condition requires that \( e^{-R\Delta t}U > 1 \), or \( U > e^{R\Delta t} \), and the second condition requires that \( e^{-R\Delta t}D < 1 \), or \( D < e^{R\Delta t} \). That is, no-arbitrage requires the following to be true:
\[ D < e^{R\Delta t} < U. \quad (3.1.9) \]

Intuitively, the conditions are that the risk free return on bond has to be higher than the lowest possible return on stock (otherwise nobody would buy bond) and lower than the highest possible return on stock (otherwise nobody would buy stock).

Using the Basic Security Prices to Price All Securities

For any security in the market, \( a = (a_1, a_2) \), we can write it as a portfolio of the two basic securities:
\[ a = a_1b(1) + a_2b(2). \]

Again, using the non-arbitrage equation in Chapter 2, (2.0.1), we have
\[ P(a) = a_1P(b(1)) + a_2P(b(2)). \quad (3.1.10) \]

From (3.1.7) and (3.1.8), then, we have
\[ P(a) = \frac{e^{-R\Delta t}U - 1}{U - D}a_1 + \frac{1 - e^{-R\Delta t}D}{U - D}a_2. \]
Example 6. $S = 100$, $U = 1.1$, $D = 0.9$, $R = 0$, $\Delta t = 1$. That is, the payoffs of the stock are $(90, 110)$, and the payoffs of the bond are $(1, 1)$. Then,

$$P(b_{(1)}) = \frac{e^{-R\Delta t}U - 1}{U - D} = \frac{e^{-0.05 \times 1.1}1.1 - 1}{1.1 - 0.9} = \frac{0.1}{0.2} = \frac{1}{2},$$

$$P(b_{(2)}) = \frac{1 - e^{-R\Delta t}D}{U - D} = \frac{1 - e^{-0.05 \times 0.9}}{1.1 - 0.9} = \frac{0.1}{0.2} = \frac{1}{2}.$$

For a call option on stock (i.e., the option to buy stock) with strike price (or exercise price) 100, the option payoff is 0 when the stock price is 90 and 10 when the stock price is 110. So, $a = (0, 10)$ and the option price is given by

$$P(a) = \frac{1}{2} \times 0 + \frac{1}{2} \times 10 = 5.$$

For a put option on stock (i.e., the option to sell the stock) with strike price (or exercise price) 100, the option payoff is 10 when the stock price is 90 and 0 when the stock price is 110. So, $a = (10, 0)$ and the option price is given by

$$P(a) = \frac{1}{2} \times 10 + \frac{1}{2} \times 0 = 5.$$

Example 7. The same as in example 6 except that $R = 0.05$. Then,

$$P(b_{(1)}) = \frac{e^{-R\Delta t}U - 1}{U - D} = \frac{e^{-0.05 \times 1.1}1.1 - 1}{1.1 - 0.9} = 0.232,$$

$$P(b_{(2)}) = \frac{1 - e^{-R\Delta t}D}{U - D} = \frac{1 - e^{-0.05 \times 0.9}}{1.1 - 0.9} = 0.719.$$

So, the price of the call option on stock with strike price 100 is

$$P(a) = 0.232 \times 0 + 0.719 \times 10 = 7.19,$$

and the price of the put option on stock with strike price 100 is

$$P(a) = 0.232 \times 0 + 0.719 \times 0 = 2.32.$$

Note that as the interest rate increases from 0 to 0.05, the price of call option rises and the price of the put option declines.
Example 8. The same as in example 6, but now consider a security whose payoff is the stock price squared. So, \( a_1 = 90^2 = 8100 \), and \( a_2 = 110^2 = 12100 \)

\[
P(a) = \frac{1}{2} \times 8100 + \frac{1}{2} \times 12100 = 10100.
\]

The security that we consider in this example may look unusual. However, many unusual or exotic derivative contracts have in fact been traded in practice.

Example 9. The same as in example 6, but now consider a forward contract, i.e. a contract to buy the stock at a fixed price of 100. Note that the forward contract differs from the option contract in that the holder of the contract has to buy the stock. So, \( a = (-10, 10) \) and the price of the forward contract is

\[
P(a) = \frac{1}{2} \times (-10) + \frac{1}{2} \times 10 = 0.
\]

If the interest rate is 0.05, as in example 7, then,

\[
P(a) = 0.232 \times (-10) + 0.719 \times 10 = 4.87.
\]

Risk-Neutral or Risk-Adjusted Probabilities

From equation (3.1.9) we can see that to price securities, all we need to know are the state-contingent payoffs of the securities and the prices of the two basic securities, \( P(b_{(1)}) \) and \( P(b_{(2)}) \), which themselves are derived from the prices of the stock and bond traded in the market. This is somewhat surprising. For example, the basic security \( b_{(1)} \) pays one dollar in state one and nothing in the second state. One would think that the price of this security should depend on how likely it is that the state one will be realized as well as how much investors value having one dollar in state one. Yet, when we derive the price of the security, we did not use any such information. Only information on stock and bond prices are used. This is possible because under no-arbitrage condition, financial market is efficient and therefore the market prices of stocks and bonds have already contained all the information that is needed for pricing any securities.

There is, however, a probabilistic representation of the pricing of securi-
ties. Let
\[ q = \frac{P(b_{11})}{P(b_{11}) + P(b_{12})} = \frac{U - e^{R\Delta t}}{U - D}, \quad (3.1.11) \]
\[ 1 - q = \frac{P(b_{12})}{P(b_{11}) + P(b_{12})} = \frac{e^{R\Delta t} - D}{U - D}. \quad (3.1.12) \]

Under the no-arbitrage condition (3.1.9), we know that \( 0 < q < 1 \). So both \( q \) and \( 1 - q \) can be thought of as probabilities. From the construction, we have
\[ P(b_{11}) = \left[ P(b_{11}) + P(b_{12}) \right] q, \quad (3.1.13) \]
\[ P(b_{12}) = \left[ P(b_{11}) + P(b_{12}) \right] (1 - q). \quad (3.1.14) \]

From the equations for basic security prices that we derived earlier, (3.1.7) and (3.1.8), we have
\[ P(b_{11}) + P(b_{12}) = e^{-R\Delta t} U - 1 + 1 - e^{-R\Delta t} D = e^{-R\Delta t} U - e^{-R\Delta t} D = e^{-R\Delta t} (U - D). \]
So,
\[ P(b_{11}) + P(b_{12}) = e^{-R\Delta t} \]
Substituting it into (3.1.11) and (3.1.12) yields the following
\[ P(b_{11}) = e^{-R\Delta t} q, \]
\[ P(b_{12}) = e^{-R\Delta t} (1 - q). \]

Therefore, for any security \( a = (a_1, a_2) \), from equation (3.1.10), we have
\[ P(a) = e^{-R\Delta t} q a_1 + e^{-R\Delta t} (1 - q) a_2 = e^{-R\Delta t} [q a_1 + (1 - q) a_2]. \quad (3.1.15) \]

If we think of \( q \) and \( 1 - q \) as the probabilities of state one and two, respectively, then \( qa_1 + (1 - q)a_2 \) is simply the expected value of the security’s payoffs, and equation (3.1.15) says that the price of the security equals the expected payoff discounted by the risk-free interest rate. In other words, under these probabilities, we price a security as if we are risk neutral—all we care about is the expected payoff. For this reason, we call \( q \) and \( 1 - q \) the risk-neutral probabilities. This, however, does not mean that investors are risk-neutral. The risk-neutral probabilities are in general not the same as investors’ subjective probabilities. They are related to the subjective probabilities, but adjusted for the risk attitudes that investors have. We will go back to this point in more detail in the next chapter.
CHAPTER 3. APPLICATIONS OF ARROW-DEBREU THEORY: PRICING OPTIONS
Chapter 4

Individual and Social Gains from Risk Management

4.1 Individual’s Attitude Towards Risk

In the economy with risky outcomes, individuals’ consumption may be state dependent. Let \( c_i \) be an individual’s consumption in state \( i \), for \( i = 1, \ldots, N \) and let \( c = (c_1, \ldots, c_N) \) be the individual’s consumption plan. The individual’s preference over different consumption plans is represented by an expected utility function

\[
U(c) = E[u(c)] = \pi_1 u(c_1) + \ldots + \pi_N u(c_N),
\]

where \( \pi_i \) is the subjective probability that state \( i \) will be realized. In this chapter, we assume that all individuals share the same belief and have the same subjective probabilities of states. So, \( \pi_i \) can also be thought of as the “objective” probability that state \( i \) will be realized.

We call an individual risk-averse if for any two consumption plans, \( c^a = (c_1^a, \ldots, c_N^a) \) and \( c^b = (c_1^b, \ldots, c_N^b) \) that have the same expected consumption level, i.e.

\[
\pi_1 c_1^a + \ldots + \pi_N c_N^a = \pi_1 c_1^b + \ldots + \pi_N c_N^b,
\]

she prefers the one with lower variance. If an individual is indifferent between any two alternative consumption plans that have the same expected consumption level, then we call her risk-neutral.

An individual’s risk attitude is related to the curvature of her utility function \( u(.) \). If an individual’s utility function is linear, then she only cares
CHAPTER 4. INDIVIDUAL AND SOCIAL GAINS FROM RISK MANAGEMENT

about the expected consumption level and, therefore, is risk–neutral. If an individual’s utility function is concave, i.e. \( u''(x) < 0 \) for any \( x \), then the individual is risk-averse. By Jensen’s inequality\(^1\),

\[
U(c) = E[u(c)] \leq u(E[c]),
\]

where \( E[c] = \pi_1 c_1 + \ldots + \pi_N c_N \) is the expected consumption level, and the equality holds if and only if \( c_i = E[c] \) for all \( i = 1, \ldots, N \). Thus, a risk–averse individual strictly prefers to have a risk–free consumption plan (i.e., consumption does not vary across states) over any risky consumption plan (i.e., consumption varies across at least two different states). If her income is risky and if it is possible, she would trade her risky income for a risk–free consumption plan that allows her to consume her expected income in every states. In fact, she is willing to trade for a risk–free consumption plan that allows her to consume at a level that is below her expected income. In other words, she is willing to pay a positive risk–premium, measured by the reduction in her expected consumption, in order to eliminate the income risk.

**Insert Figure 3 Here**

Figure 3 illustrates the benefits of eliminating income risk for a risk–averse individual and the maximum risk–premium she is willing to pay. Consider an economy with two possible states, \( \{s_1, s_2\} \). The individual’s income in the two states are \( y_l \) and \( y_h \), respectively, where \( y_h > y_l \). So \( s_1 \) can be thought of as the bad state and \( s_2 \) the good state for the individual. The individual’s utility function is plotted in the figure as \( u(.) \). If the individual simply consumes her income, then she has a high level of consumption in the good state but a low level of consumption in the bad state. The expected utility of such a consumption plan is

\[
\pi_1 u(y_l) + \pi_2 u(y_h),
\]

---

\(^1\)Jensen’s Inequality refers to the following mathematical proposition: If a function, \( f(.) \), is concave (i.e., \( f''(x) < 0 \) for any \( x \)), then,

\[
\sum_{i=1}^{n} \alpha_i f(x_i) \leq f \left( \sum_{i=1}^{n} \alpha_i x_i \right)
\]

for any positive integer \( n \), and any \( \alpha_1 > 0, \ldots, \alpha_n > 0 \) such that \( \sum_{i=1}^{n} \alpha_i = 1 \). Furthermore, the equality holds if and only if \( x_1 = x_2 = \ldots = x_n \).
4.2. INDIVIDUAL DEMAND FOR RISK MANAGEMENT

which is represented graphically by the height of point A. If the individual trades her risky income for a risk–free consumption plan, \((x, x)\), then the expected utility of such a consumption plan is \(\pi_1 u(x) + \pi_2 u(x) = [\pi_1 + \pi_2]u(x) = u(x)\), which is graphically represented by the height of the utility function \(u(.)\) evaluated at \(x\). If \(x = E[y] = \pi_1 y_l + \pi_2 y_h\), then the expected utility she enjoys is represented by the height of point \(A'\), which is higher than point A. The lowest level \(x\) that the individual is willing to trade for is \(x^*\) at which the individual’s expected utility is the same as that from consuming her own risky income. The distance between \(E[y]\) and \(x^*\), \(E[y] - x^*\), is the maximum risk–premium or the reduction in the expected consumption level, that the individual is willing to pay for eliminating her income risk. Whether she will be able to trade at such a risk premium depends how risk–averse other individuals in the economy are and whether the individual’s income risk is diversifiable or not. In the next section, we will examine the conditions under which individuals will be able to trade financial securities to reduce the risk they face.

### 4.2 Individual Demand for Risk Management

Consider now an Arrow-Debreu economy that consists of \(M\) individuals, indexed by \(j = 1, \ldots, M\). Each individual's income is state-contingent. For any \(i = 1, \ldots, N\) and \(j = 1, \ldots, M\), let \(y_{j,i}\) be individual \(j\)'s income in state \(i\). So, individual \(j\)'s income over all possible states can be represented by a vector \(y_{(j)} = (y_{j,1}, \ldots, y_{j,N})\). The individual's problem is to choose a consumption plan, \(c_{(j)} = (c_{j,1}, \ldots, c_{j,N})\), which determines the amount of consumption in each possible state, in order to maximize his/her expected utility function:

\[
U(c_{(j)}) = E[c_{(j)}] = \pi_1 u_1(c_{j,1}) + \ldots + \pi_N u_N(c_{j,N}).
\]

For simplicity, we assume that individuals' utility function is of the following form:

\[
u_j(c) = \frac{1}{1 - \gamma_j} [c^{1-\gamma_j} - 1], \quad \gamma_j \geq 0.
\]

Remember that the relative risk aversion coefficient of a utility function is given by \(RRAC = -u''(c)c/u'(c)\) and it measures the degree of risk aversion of an individual with utility function \(u(.)\). For the utility function specified above, \(RRAC = \gamma_j\), which is a constant. So, the utility functions of this form are also called constant relative risk aversion utilities. The parameter
26 CHAPTER 4. INDIVIDUAL AND SOCIAL GAINS FROM RISK MANAGEMENT

\( \gamma_j \) represents the degree of individual \( j \)'s risk-aversion. The larger is \( \gamma_j \), the more risk-averse individual \( j \) is. If \( \gamma_j = 0 \), individual \( j \) is risk-neutral; if \( \gamma_j > 0 \), individual is risk averse; and if \( \gamma_j < 0 \), individual is risk loving.

We assume that the market is complete and individuals trade financial securities in order to achieve their optimal consumption plans. Individual \( j \)'s consumption plan can be accomplished if she can buy a portfolio that pays \( c_{j,i} \) in state \( i \) for all \( i = 1, ..., N \). From the no-arbitrage equation, we know that the cost of such a portfolio is

\[
c_{j,1}p_1 + ... + c_{j,N}p_N.
\]

The individual can finance its consumption plan by selling its income portfolio, which has value

\[
y_{j,1}p_1 + ... + y_{j,N}p_N.
\]

So, the individual’s problem is

\[
\max_{c_{j,1}, ..., c_{j,N}} \frac{1}{1 - \gamma_j} \sum_{i=1}^{N} \pi_i \left[ c_{j,i}^{1-\gamma_j} - 1 \right]
\]

subject to the budget constraint

\[
\sum_{i=1}^{N} c_{j,i}p_i \leq \sum_{i=1}^{N} y_{j,i}p_i.
\]

Define a Lagrangian \( L_j \) as follows:

\[
L_j = \frac{1}{1 - \gamma_j} \sum_{i=1}^{N} \pi_i \left[ c_{j,i}^{1-\gamma_j} - 1 \right] + \lambda_j \left[ \sum_{i=1}^{N} y_{j,i}p_i - \sum_{i=1}^{N} c_{j,i}p_i \right].
\]

Then, the individual’s optimal consumption plan is determined by the following first order conditions:

\[
\frac{\partial L_j}{\partial c_{j,i}} = \pi_i c_{j,i}^{-\gamma_j} - \lambda_j p_i = 0, \quad \text{for } i = 1, ..., N.
\] (4.2.1)

The individual’s optimal consumption plan is the solution to equation (4.2.1) and the following complementarity conditions:

\[
\lambda_j \geq 0, \quad \text{and} \quad \lambda_j \left[ \sum_{i=1}^{N} y_{j,i}p_i - \sum_{i=1}^{N} c_{j,i}p_i \right] = 0
\]
4.2. INDIVIDUAL DEMAND FOR RISK MANAGEMENT

From equation (4.2.1), we know that \( \lambda_j > 0 \). Thus, the above complementarity condition implies that

\[
\sum_{i=1}^{N} c_{j,i} p_i = \sum_{i=1}^{N} y_{j,i} p_i, \tag{4.2.2}
\]

or the individual’s budget constraint is binding.

From the first order condition (4.2.1), we can solve for the optimal consumption demand of a risk-averse individual \( j (\gamma_j > 0) \), \( c_{j,i} \),

\[
c_{j,i} = \lambda_j^{-1/\gamma_j} \left( \frac{\pi_i}{p_i} \right)^{1/\gamma_j} = w_j x_i^{1/\gamma_j}, \tag{4.2.3}
\]

where \( w_j = \lambda_j^{-1/\gamma_j} \) and \( x_i = \pi_i / p_i \). Equation (4.2.3) is a key equation for understanding the role of financial market in managing individuals’ income risk. Without fully solving the model, we can already derive from the equation some important implications of risk-sharing under a complete financial market.

Implied 1: A complete financial market allows individuals to diversify all idiosyncratic income risks.

Note that \( x_i \) is a variable that varies across states but does not depend on individual types. We can interpret it as a variable that represents the aggregate risk that is faced by every individual in the economy. Since \( w_j \) is a state-independent constant, equation (4.2.3) shows that each individual’s consumption exposure to risk is entirely through the aggregate risk variable \( x_i \). Therefore, any individual income risk that is uncorrelated with the aggregate risk \( x_i \) has no impact on consumption. In other words, with a complete financial markets, individuals can diversify away all idiosyncratic income risks.

Implication 2: A complete financial market allows individuals to share aggregate risk more efficiently.

Taking logarithm on both sides of equation (4.2.3), we have

\[
\ln c_{j,i} = \ln w_j + \gamma_j^{-1} \ln x_i. \tag{4.2.4}
\]

Let \( \sigma_j^2 \) and \( \sigma_x^2 \) be the variance of \( \ln c_j(s_i) \) and \( \ln x(s_i) \), respectively, then (4.2.4) implies that \( \sigma_j^2 = \gamma_j^{-2} \sigma_x^2 \) or

\[
\frac{\sigma_j}{\sigma_j'} = \frac{\gamma_{j'}}{\gamma_j}.
\]
That is, the volatility of an individual’s optimal consumption is inversely related to how risk–averse she is. Those who are highly risk–averse will end up having less risk in their consumption relative to those who are less risk–averse.

4.3 Equilibrium Prices of Arrow-Debreu Securities and Aggregate Risk

Since \( x_i = \pi_i / p_i \) and \( p_i \) is an endogenous variable, we need to solve for the equilibrium value of \( p_i \) to see how aggregate risk is determined in this economy. In this section we present two special cases when we can solve \( x_i \) and \( p_i \) explicitly.

Case 1: There exists some risk–neutral agents. Assume, for example, that individual 1 is risk–neutral, \( \gamma_1 = 0 \). Then, she is willing to take any income risk as long as her expected consumption level is not reduced. As a result, every risk–averse individual will completely eliminate her income risk by trading securities with the risk–neutral agent.\(^2\)

For \( j = 1 \), equation (4.2.1) becomes

\[
\pi_i - \lambda_1 p_i = 0
\]

or

\[
p_i = \lambda_1^{-1} \pi_i.
\]

(4.3.1)

This implies that \( x_i = \lambda_1 \) for any \( i = 1, ..., N \), and, from (4.2.3), for any individual \( j \) such that \( \gamma_j > 0 \),

\[
c_{j,i} = (\lambda_1 / \lambda_j)^{1/\gamma_j} \equiv \tau_j,
\]

which is a constant across all \( N \) different states. From (4.2.2), then, we have,

\[
\tau_j \sum_{i=1}^{N} p_i = \sum_{i=1}^{N} p_i y_{j,i}.
\]

\(^2\)We assume that everyone in this economy is a price-taker. This is equivalent to assuming that there are many individuals for each type and there are perfect competition among them. If there is only one individual who is risk-neutral, the individual would have monopoly power and therefore may charge a risk premium for taking income risk even if she is risk-neutral.
4.3. EQUILIBRIUM PRICES OF ARROW-DEBREU SECURITIES AND AGGREGATE RISK

Substituting (4.3.1) into the equation above yields:

\[ \mathbf{c}_j \sum_{i=1}^{N} \lambda_i^{-1} \pi_i = \sum_{i=1}^{N} \lambda_i^{-1} \pi_i y_{j,i}, \]

or

\[ \mathbf{c}_j = E[y_{(j)}] = \sum_{i=1}^{N} \pi_i y_{j,i}. \]

So, the risk-averse individual \( j \) will choose to fully insure herself from income fluctuations. The optimal consumption plan of individual \( j \) is

\[ (E[y_{(j)}], ..., E[y_{(j)}]). \]

In other words, when there are risk–neutral agents in the economy, then all risk–averse individuals can dump all of their income risk to the risk–neutral agents without paying any risk–premium.

Case 2: Individuals are equally risk–averse.

We assume that \( \gamma_j = \gamma \) for all \( j = 1, ..., M \). Note that \( p_i \) is the price of (one unit of) consumption in state \( i \). Like price of any commodity, \( p_i \) can be determined by the market clearing condition: Aggregate demand for consumption in state \( i \) equals the aggregate income in the same state. From (4.2.3), the aggregate demand for consumption in state \( i \) is

\[ C_i = \sum_{j=1}^{M} c_{j,i} = \sum_{j=1}^{M} \left( w_j x_i^{1/\gamma} \right) = \left( \sum_{j=1}^{M} w_j \right) x_i^{1/\gamma}. \]

Let \( Y_i = \sum_{j=1}^{M} y_{j,i} \) be the aggregate income in state \( i \). Then, the market clearing condition implies that

\[ C_i = \left( \sum_{j=1}^{M} w_j \right) x_i^{1/\gamma} = Y_i \]

or

\[ x_i = \left( \sum_{j=1}^{M} w_j \right)^{-\gamma} (Y_i)^{\gamma}. \] (4.3.2)
So, the aggregate risk variable $x_i$ is a time-invariant, increasing function of the aggregate income, $Y_i$. Substituting (4.3.2) into individual consumption demand equation (4.2.3) yields the following equation:

$$c_{j,i} = \frac{w_j}{\sum_{j=1}^{M} w_j} Y(s_i), \quad j = 1, ..., M. \quad (4.3.3)$$

That is, consumption of all individuals in this economy are proportional to aggregate income in the economy. Put it differently, when there is a complete financial market and all households have the same preferences, then, percentage deviation of any individual’s consumption from aggregate income or aggregate consumption should be zero in all states of the world.

### 4.4 Arrow-Debreu Prices, Stochastic Discount Factor, and Risk-Adjusted Probabilities

In Chapter 2, we showed that for any security $a = (a_1, ..., a_N)$, we have

$$P(a) = \sum_{i=1}^{N} a_i p_i.$$  

So, one can price any security in the market by using Arrow-Debreu prices, without any reference to probabilities of states. One natural question is: How are Arrow-Debreu prices related to the commonly agreed subjective probabilities (or objective probabilities) of the states of the world? Equation (4.3.2) provides an explicit answer to this question. Recall that $x_i = \pi_i/p_i$. From (4.3.2) we have

$$p_i = \pi_i \left( \sum_{j=1}^{M} w_j \right)^\gamma (Y_i)^{-\gamma} = \pi_i m_i, \quad (4.4.1)$$

where

$$m_i = \left( \sum_{j=1}^{M} w_j \right)^\gamma (Y_i)^{-\gamma}.$$  

So, the Arrow-Debreu price of state $i$, $p_i$, equals the probability of that state times an adjustment factor, $m_i$, which is also called the stochastic discount
4.5. EMPIRICAL EVIDENCE

You may think that the Arrow-Debreu theory that we used in this chapter is an over-simplified model with many unrealistic assumptions. However,
some empirical studies have shown that the predicted household consumption behavior are actually quite close to what we observe in real economies. For example, a paper by Robert Townsend (1994) shows that household consumption in poor India villages exhibit a pattern that are very close to the pattern of optimal consumption our theory predicted in equation (4.3.3).

You may also think that the Arrow-Debreu theory, with its assumption of complete financial market, may have its best application in an economy like US, because of the highly developed financial markets there. However, a paper by John Cochran (1991) shows that is not the case. While the data does show that household consumption is independent of many idiosyncratic income risks, the insurance is not complete. Certain types of idiosyncratic risks, such as long-term health problems and long spell of unemployment, are not insured. In other words, household’s consumption are still affected by these idiosyncratic shocks.

Finally, in a paper by Huw Lloyd-Ellis and Xiaodong Zhu (2001), it is shown that using financial market to smooth taxes can indeed yield welfare gains for the Canadian economy, as predicted by the theory in the last section. Furthermore, they estimate that the welfare gains can be substantial (around half a percent of GDP each year).
Chapter 5

Why Should Firms Manage Risk?

The discussions in last chapter should make it clear that there are many reasons why individuals may want to trade financial securities in order to manage their consumption risk. The justification for risk management by widely held public firms—those with a large and diverse group of shareholders, however, is not so straightforward. In fact, under certain conditions, one may even argue that risk management is irrelevant for these firms. In the next section, we examine in detail this argument, which was originally presented in the Nobel-prize-winning paper of Modigliani and Miller.

5.1 An Example: Risk Management by Company XYZ

Suppose that XYZ is a gold digging firm. It is known for sure that the firm will have 100,000 ounces of gold in a year. The price of the gold today is $400/ounce. So, if the gold price remains the same over the next year, the firm’s cash-flow (sales revenues in this case) would be $40 millions. However, the gold price is risky. Suppose that there is 50\% of chance that the gold price will be $440/ounce, and 50\% of chance that the gold price will be $360/ounce. So, the expected value of the firm is

\[
\bar{y} = 0.5 \times 440 \times 100,000 + 0.5 \times 360 \times 100,000 = 40,000,000.
\]

In generally, market value of the firm is not the same as the expected value of the firm because investors care about the risk of the firm and may
value the cash-flow in different states differently. Suppose that investors today value having $1 in the future when the gold price is $360 more than $1 when the gold price is $440. To be more specific, let’s call the state when the gold price is $440 the high state and the gold price is $360 the low state. Let’s also assume that the market value of having $1 in the high state is $3/8 and in the low state is $5/8. Then, the market value of the firm is

$$V(y) = \frac{3}{8} \times 440 \times 100,000 + \frac{5}{8} \times 360 \times 100,000 = 39,000,000.$$  

So, in this example, the market value of the firm is less than the expected value of the firm because the market investors dislike risk. We now study how the firm can use forward or option contract to manage its cash-flow risk and whether the risk management strategies help to increase the market value of the firm.

**Managing risk with forward contract**

If the firm enters into a short position in a forward contract so that it will sell 100,000 ounces of gold to the other party of the contract at a fixed price $k$ per ounce. Then the firm’s cash-flow will simply be $k \times 100,000$ dollars in both states. So, the market value of the firm is

$$V(y) = \frac{3}{8} \times k \times 100,000 + \frac{5}{8} \times k \times 100,000 = k \times 100,000.$$  

To see if the market value of the firm has changed with hedging using the forward contract, we need to know what is the delivery price $k$. Since a forward contract cost zero to enter for both parties, to be consistent with no arbitrage the delivery price must be such that the market value of the contract is zero. In the high state, the forward contract value to the firm is $(k - 440) \times 100,000$, and in the low state, the value is $(k - 360) \times 100,000$. So, the market value of the forward contract to the firm is

$$\left(\frac{3}{8}\right) \times (k - 440) \times 100,000 + \left(\frac{5}{8}\right) \times (k - 360) \times 100,000$$

$$= k \times 100,000 - 390 \times 100,000.$$  

For the value of the forward contract to be zero, the delivery price has to be $390 per ounce. In that case, the market value of the firm is 39,000,000 dollars, which is the same as the value of the firm without hedging.

**Managing risk with option contract**
Now consider the case when the firm hedges the gold price risk with an option contract rather than forward contract. For example, the firm could eliminate the downside risk by buying a put option, which gives the firm the right to sell 100,000 ounces of gold at the price of $390. This would insure that the firm’s sales revenue will be no less than $390*100,000. So, in the high state, the firm will not exercise the put option, and sell the gold in the market at $440 per ounce, receiving $440*100,000 in revenues. In the low state, the firm will exercise the put option and sell it at $390 per ounce, receiving $390*100,000. So, the market value of the firm’s future cash-flow is

\[(3/8) \times 440 \times 100,000 + (5/8) \times 390 \times 100,000 = 40,875,000.\]

However, the firm has to pay for the put option it bought. The economic value of the option is 0 in the high state and \((390 - 360) \times 100,000 = 3,000,000\) in the low state. So, the market value of the put option is

\[(3/8) \times 0 + (5/8) \times 3,000,000 = 1,875,000.\]

The market value of the firm equals the market value of its cash-flow, $40,875,000, minus the cost of hedging, $1,875,000, which equals $39,000,000. Again, the value of the firm is unaffected by hedging.

The examples about the firm XYZ show that when derivatives contracts are valued in the market based on no-arbitrage, then hedging does not change the market value of the firm. This is a result that has been shown more generally by Franco Modigliani and Merton Miller (MM theorem), which states that a firm’s financial policy is irrelevant in determining the market value of the firm. If shareholders can manage risk themselves, then, there is no need for risk management for the firm.

5.2 Modigliani–Miller Theorem

In a very influential paper, Modigliani and Miller (1958) show that, under some fairly general (but strong) conditions, financial policies (including hedging) can not change the market value of a firm. We used the example of XYZ firm to illustrate this result in the last chapter. Now we use the complete market framework in the previous section to prove this very strong and powerful result.

Consider a firm whose income in state \(i\) is \(y_i\), for \(i = 1, ..., N\). From Chapter 2, we know that the market value of the firm (i.e. the market value
of the income plan \((y_1, \ldots, y_N)\) is

\[
V = P(y) = \sum_{i=1}^{N} y_i p_i.
\]

If the firm buys a financial security \(a = (a_1, \ldots, a_N)\) to change the distribution of its income across states, then, the new income profile is given by \(\tilde{y} = (y_1 + a_1, \ldots, y_N + a_N)\), and the firm’s market value is now

\[
\tilde{V} = P(\tilde{y}) - P(a)
\]

where \(P(a)\) is the market value of the security \(a\). But

\[
P(\tilde{y}) = \sum_{i=1}^{N} [y_i + a_i] p_i,
\]

and

\[
P(a) = \sum_{i=1}^{N} a_i p_i.
\]

So,

\[
\tilde{V} = \sum_{i=1}^{N} [y_i + a_i] p_i - \sum_{i=1}^{N} a_i p_i = \sum_{i=1}^{N} y_i p_i = V.
\]

That is, the value of the firm before and after the firm purchased the financial security are the same—the firm’s financial policy does not affect the firm’s market value. This irrelevance result is the now celebrated Modigliani–Miller Theorem.

Note that in the proof of the theorem above, we have not put any restrictions on the income \(\tilde{y}\) implied by the firm’s financial policy. In reality, however, firms may have limited liability and, therefore, the firm’s income should always be non-negative: when the implied income is negative, the firm will simply go bankrupt and pays zero to those who have claims on the firm. In the following we show that even if there is the possibility of default, Modigliani–Miller Theorem still holds.

Assume that the firm chooses a combination of equity and bond to finance its investment. Let the firm issue a bond which promises to pay the bond holder a fixed amount \(R\). Let \(V_D\) be the funds that the firms raised through the bond issue. Then, the total value to the firm is the value of equity income
plus the funds raised through the bond issue. The payoffs to the equity and bond holders are state contingent. In a good state when the investment return is higher than the promised payment to the bond holder, the firm will pay $R$ to the bond holder and the residual income will be paid to the equity holder as dividend. In a bad state when the investment return is below the promised payment $R$, however, the firm will go bankrupt, the bond holder will seize all of the firm’s income (which is less than $R$), and the equity holder receives no income from the firm. Thus, the state–contingent payoffs for the equity holder and the bond holder are as follows:

$$E_i = \begin{cases} y_i - R & \text{if } y_i \geq R; \\ 0 & \text{if } y_i < R. \end{cases}$$

and

$$D_i = \begin{cases} R & \text{if } y_i \geq R; \\ y_i & \text{if } y_i < R. \end{cases}$$

Thus, the market value of the equity income and debt are

$$V_E = \sum_{i=1}^{N} E_i p_i = \sum_{\{i: y_i < R\}} 0 \cdot p_i + \sum_{\{i: y_i \geq R\}} (y_i - R)p_i,$$

and

$$V_D = \sum_{i=1}^{N} D_i p_i = \sum_{\{i: y_i < R\}} y_i p_i + \sum_{\{i: y_i \geq R\}} Rp_i.$$

Summing the above two equation we have,

$$V_E + V_D = \sum_{i=1}^{N} y_i p_i = V_y.$$

That is, the composition of equity and debt in the financing of the firm’s investment does not affect the total value of the investment. The shareholders’ value of the firm is always the value of the firm’s income.

To understand the demand for risk management by firms, then, we need to understand why MM Theorem may not hold. There are several important assumptions behind the MM Theorem:

- shareholders can manage risk themselves
• there exist a complete, frictionless financial market;
• investment opportunities is independent of the financial decisions;
• no bankruptcy cost;
• no taxes.

Note that the first assumption is not needed for the MM Theorem per se. But it is an important reason why shareholders care only about the market value, not the risk of a firm. We now consider in turn each of these assumptions and discuss why and when these assumptions may not hold and, therefore, there are needs for firms to manage risk.

5.3 Shareholders may not be able to manage risk effectively themselves.

Risk management usually involves transaction costs. For small shareholders, the transaction costs may be too high relative to the benefits of risk management. Therefore, they may not want to manage risk themselves. Nevertheless, they would benefit from risk management since they are risk-averse. By having the firm manage the risk, the transaction costs that would have incurred independently by many small shareholders are now incurred by the firm only. Because of the economy of scale, the average transaction cost to each individual shareholder is much smaller in this case and may justify the use of risk management by the firm. A crucial condition for this argument to be valid, however, is that shareholders are similar to each other so that there are risk management strategies that can benefit most of the shareholders.

Sometimes large shareholders may find it difficult to manage themselves effectively the firm’s risk. This is especially true for inside shareholders such as Bill Gates, who is a large shareholder of Microsoft. Like any other individuals, Bill Gates may want to hedge his risk by hedging against negative shocks to Microsoft that is beyond his control. However, if he actually tries to implement a hedge by trading derivatives on Microsoft’s stocks, say put options, market may perceive it as a bad signal about the Microsoft business and, therefore, its stock value may fall due to the fact that Bill Gates is hedging against potential fall in the Microsoft stocks.
5.4 Controlling Underinvestment with Hedging

Consider the following example: From previous investment, a firm’s income is $1000 with probability 1/2 and $200 with probability 1/2. The firm has $500 debt outstanding. After the income of the firm is realized, and before the debt is due, the firm is presented with a risk-free investment project: investing $600 yields $800 in return.

If the firm’s income is $1000, then, by taking the project, the firm’s total income would be $1,200. After paying back the debt, the equity holders would have a residual value of $700. If the project is not taken, however, the total income would be $1000, and the equity holders would have a residual value of $500. So, the equity holder will invest in the project.

If the firm’s income is $200, then, to invest in the project, the firm has to come up with additional $400. After the investment, the total income would be $800, of which $500 will go to the bond holders and the equity holders would have a residual value of $300. Taking into account the $400 additional investment, the equity holder incur a net loss of $100. If the investment project is not taken, then the firm would go into default, the bond holders receive $200 and the equity holders receive no income. In this case, the equity holders do not have incentive to invest in the project even if it is clearly profitable project. The reason that the equity holders do not want to invest in the project is that all the additional returns are captured by the bond holders when the firm is in financial distress.

If the firm has hedged so that the income in both state is $600 (assuming that there is no hedging cost), then the firm will always take the investment project. So, hedging in this case ensures that the profitable project is taken by the firm and, therefore, increases the value of the firm.

5.5 Reducing Tax Liabilities with Hedging

Governments often give tax credits (i.e., investment tax credits for research and development expenditures) to firms. When a firm’s income is too low, they will not be able to take advantage of such credits. Risk management can help firms to fully take advantage of tax credits by stabilizing income. To see why this is the case, consider the following simple example.

Let $y$ be a firm’s income, which is subject to a flat tax rate, $\tau$, $0 < \tau < 1$. Suppose that the government allows the firm to write off its first $d$ dollars of
income as tax credits. So, if the firm’s income is greater than \( d \), then its tax liability is \( \tau(y - d) \); and if the firm’s income is less or equal to \( d \), then it does not have to pay any taxes. That is, the tax liability of the firm with income level \( y \) is

\[
T = \tau \max(y - d, 0).
\]

Figure 4 plots the tax liability as a function of the firm’s income, \( y \). Suppose that with equal probabilities (0.5) the firm’s income may be \( y_l \) or \( y_h \), and that \( y_l < d < y_h \). Then, the firm’s expected tax liability is

\[
E[T] = \frac{1}{2} \tau(y_h - d).
\]

If the firm has hedged so that its income always equals (assuming no hedging cost for simplicity)

\[
E[y] = \frac{1}{2} y_h + \frac{1}{2} y_l,
\]

then the tax liability is certain and equals

\[
T_{hedging} = \tau \max(E[y] - d, 0)
\]

Because \( y_l < d \), we have

\[
E[y] - d = \frac{1}{2}(y_h - d) + \frac{1}{2}(y_l - d) < \frac{1}{2}(y_h - d).
\]

So,

\[
T_{hedging} < \tau \max(\frac{1}{2}(y_h - d), 0) = \frac{1}{2} \tau(y_h - d).
\]

That is, the tax liability under hedging is less than the expected tax liability without hedging. (See Figure 4.)

Insert Figure 4 Here
5.6 Debt and Hedging Policies

The reason that a firm may experience financial distress often has to do with problems of repaying the debt the firm had incurred. Why would a firm issue debt then? According to the MM theorem, debt financing does not increase the value of the firm, and therefore, there is not reason for the firm to have debt financing, especially when debt may increase the probability of financial distress.

One reason firms issue debt is to reduce its tax liabilities, because interest payments on debt can be deducted from firms’ cash flows. Therefore, in choosing its debt level, a firm is faced with a trade-off between tax reduction and increases in the probability of financial distress. We now consider a simple model that can be used to study this trade-off. Consider the model in Section 4.2. We now introduce a flat-rate tax on firm’s cash-flow. The tax liabilities for the firm is:

\[ T_i = \begin{cases} 
  t(y_i - R) & \text{if } y_i \geq R; \\
  0 & \text{if } y_i < R.
\end{cases} \]

So, the equity holders’ dividend is given by

\[ E_i = \begin{cases} 
  (1 - t)(y_i - R) & \text{if } y_i \geq R; \\
  0 & \text{if } y_i < R.
\end{cases} \]

In addition, we also introduce a cost of default. That is, the value that the debt holder receives in the case of default is \( y_i - cy_i \), where \( cy_i \) is the cost associated with default. So, the debt holder’s payoff is given by

\[ D_i = \begin{cases} 
  R & \text{if } y_i \geq R; \\
  (1 - c)y_i & \text{if } y_i < R.
\end{cases} \]

Thus, the market value of the equity and debt are, respectively,

\[ V_E = \sum_{i=1}^{N} E_i p_i = \sum_{\{i: y_i < R\}} 0 \cdot p_i + (1 - t) \sum_{\{i: y_i \geq R\}} (y_i - R)p_i, \]

and

\[ V_D = \sum_{i=1}^{N} D_i p_i = (1 - c) \sum_{\{i: y_i < R\}} y_i p_i + \sum_{\{i: y_i \geq R\}} Rp_i. \]
Summing the above two equations we have

$$V(R) = V_E + V_D = \sum_{i=1}^{N} y_i p_i - c \left[ \sum_{\{i: y_i < R\}} y_i p_i \right] - t \left[ \sum_{\{i: y_i \geq R\}} (y_i - R) p_i \right].$$

Note that the value of the firm is now a function of the amount of liabilities outstanding, $R$. That is, with taxes and default cost, leverage has an effect on the value of the firm.

Consider the following example: $N = 2$. $y_1 = e - h$, and $y_2 = e + h$. Furthermore, assume that $p_1 = p_2 = 1/2$. Now we examine the firm’s optimal choice of debt level, $R$.

If $R \leq e - h$, then, $y_i \geq R$ for both $i = 1$ and 2. Thus,

$$V(R) = \sum_{i=1}^{2} y_i p_i - t \left[ \sum_{i=1}^{2} (y_i - R) p_i \right] = (1 - t)e + tR.$$

In this case, the optimal debt level should be the maximum possible, $R = e - h$, which implies that

$$V^1 = e - th.$$

Next, consider the case when $e - h < R \leq e + h$. In this case, $y_1 < R$ and $y_2 \geq R$. So, we have

$$V(R) = \sum_{i=1}^{2} y_i p_i - c y_1 p_1 - t(y_2 - R) p_2 = e - \frac{1}{2} c(e - h) - \frac{1}{2} t(e + h - R).$$

Again, it is optimal for the firm to choose the maximum possible debt in this case, $R = e + h$, which implies that

$$V^2 = e - \frac{1}{2} c(e - h) = (1 - \frac{1}{2} c)e + \frac{1}{2} ch.$$

Finally, if the firm chooses the debt level that is above $e + h$, then, the firm will for sure go bankrupt and we have

$$V^3 = (1 - c) \sum_{i=1}^{2} y_i p_i = (1 - c)e.$$
5.6. DEBT AND HEDGING POLICIES

Apparently, the firm will not choose a debt level that is above $e + h$, because the value of the firm in that case, $V^3$, is always lower than the value of the firm when the debt level is at $e + h$, $V^2$. Whether the firm will choose the debt level to be $e - h$, or $e + h$, however, depends on the values of $c$ and $t$.

\[ V^2 - V^1 = th - \frac{1}{2}c(e - h). \]

So, the higher the tax rate is, the more likely that the firm will choose a high debt level to reduce its tax liabilities. On the other hand, the higher the cost of default, $c$, the more likely that the firm will choose a relatively low level of debt to reduce the expected default cost. The optimal debt level, then, is chosen based on trading off these two factors.

Now let’s see how hedging affect the firm’s value and the firm’s choice of debt level. Since $p_1 = p_2 = 1/2$, we know that the market value of $(e - h, e + h)$ is the same as the market value of $(e, e)$. Therefore, the firm can choose to have its after hedge cash-flow to be $e$ in both states. In this case, if the firm choose the debt level at $R$ to be the same as $e$. Then, the firm will not go default and there will be no taxes to be paid. The total value of the firm is in this case,

\[ V = e. \]

Apparently, with hedging, the firm has eliminated the trade-off between default and tax reduction.
Chapter 6
Bond Analytics

6.1 Zero-Coupon Bonds and Coupon Bonds

There are two kinds of bonds. One are the bonds that pay the principal plus interest all at once at the maturity. Such bonds are called zero-coupon bonds. The total amount that will be paid a at the maturity is called the face value of the bond. For example, consider a one-year zero-coupon bond whose current value is $96. Suppose the interest on the bond is $4. Then the bond will pay $96 + $4 = $100 at the end of one year, and the face value of the bond is $100. The effective continuous compounding interest rate on this bond $R$ is determined by the following equation:

$$96 = e^{-R}100$$

or $R = \ln(100/96) = 4.08\%$. Here, $R$ is the interest rate that can be used to discount the face value of the bond to get the current market price of the bond.

Now, for any time $t$, let $Z(t)$ denote the zero coupon bond with a face value of $1$ paid at time $t$. Let $R(t)$ be the corresponding interest rate for the zero-coupon bond. Then, we have $P(Z(t)) = e^{-R(t)t}$.

The second type of bonds are those that pay coupons in regular time intervals as well as the principal at the maturity. For example, a two year US treasury bond with an annual coupon rate of 8\% and principal of $100 will have four coupon payments of $4$ (half of the annual coupon) at six month, one year, one and half year and two year plus the principal payment of $100 at two year. Such a coupon bond can be thought of a combination
of four zero coupon bonds: a six month zero coupon bond with face value of $4, a one year zero coupon bond with face value of $4, a one and half year zero coupon bond with face value of $4, and a 2 year zero coupon bond with face value of $104 (principal plus coupon). That is the two year US treasury bond is a portfolio of four zero-coupon bonds:

\[4Z(0.5) + 4Z(1) + 4Z(1.5) + 104Z(2),\]

and the value of this bond is

\[
P(4Z(0.5) + 4Z(1) + 4Z(1.5) + 104Z(2))
= 4P(Z(0.5)) + 4P(Z(1)) + 4P(Z(1.5)) + 104P(Z(2))
= e^{-R(0.5)0.5}4 + e^{-R(1)}4 + e^{-R(1.5)1.5}4 + e^{-R(2)2}104.
\]

Now consider a portfolio of two coupon bonds. One is a one year coupon bond with a coupon rate of 6% and principal of $100. The other is a two year coupon bond like the one above. We can write the one year coupon bond as

\[3Z(0.5) + 103Z(1).\]

So, the portfolio of the two bonds can be written as

\[
3Z(0.5) + 3Z(1) + 4Z(0.5) + 4Z(1) + 4Z(1.5) + 104Z(2)
= 7Z(0.5) + 107Z(1) + 4Z(1.5) + 104Z(2).
\]

Thus, a portfolio of coupon bonds can also be written as a portfolio of zero coupon bonds. The value of the portfolio is

\[
7Z(0.5) + 107Z(1) + 4Z(1.5) + 104Z(2)
= e^{-R(0.5)0.57} + e^{-R(1)107} + e^{-R(1.5)1.5}4 + e^{-R(2)2}104.
\]

The examples above show that a portfolio of bonds always consists of a finite number of payments at a finite number of dates. So, we can always express the portfolio as a portfolio of zero coupon bonds:

\[c_1Z(t_1) + c_2Z(t_2) + ... + c_nZ(t_n).\]  

(6.1.1)  

In the first example above, we have \(n = 1\), \(t_1 = 1\), and \(c_1 = 100\). In the example of the two year treasury bond, we have \(n = 4\), \(t_1 = 0.5\), \(t_2 = 1\), \(t_3 = 1.5\), and \(t_4 = 2\); and \(c_1 = c_2 = c_3 = 4\) and \(c_4 = 104\). Finally, in the last
example, we also have \( n = 4 \), \( n = 4 \), \( t_1 = 0.5 \), \( t_2 = 1 \), \( t_3 = 1.5 \), and \( t_4 = 2 \). However, we have \( c_1 = 7 \), \( c_2 = 107 \), \( c_3 = 4 \), and \( c_4 = 104 \).

The value of the bond portfolio in (6.1.1) can be written as:

\[
B = P(c_1 Z(t_1) + c_2 Z(t_2) + ... + c_n Z(t_n)) = e^{-R(t_1)t_1}c_1 + e^{-R(t_2)t_2}c_2 + ... + e^{-R(t_n)t_n}c_n.
\]

(6.1.2)

### 6.2 Bond Yield and Duration

Equation (6.1.2) shows that the value of a bond portfolio is affected by \( n \) potentially different interest rates, \( R(t_1),...,R(t_n) \). The reason is that the portfolio has cash-flows that will occur at different dates and therefore each of these cash-flow will be discounted by an appropriate interest rate. It would be convenient, however, if we can use one interest rate to discount all the cash-flows in the portfolio for determining the value of the portfolio. Such an interest rate is call the yield of the bond portfolio and is defined mathematically as the value of \( y \) that solves the following equation:

\[
B = e^{-R(t_1)t_1}c_1 + e^{-R(t_2)t_2}c_2 + ... + e^{-R(t_n)t_n}c_n = e^{-yt_1}c_1 + e^{-yt_2}c_2 + ... + e^{-yt_n}c_n.
\]

(6.2.1)

For a zero-coupon bond \( Z(t) \), we have

\[
B = P(Z(t)) = e^{-R(t)t}
\]

and apparently the yield of the bond is simply \( R(t) \). For a coupon bond like that in (6.1.1), however, the yield of the bond does not correspond to a particular interest rate. Instead, it is some kind of average of all the relevant interest rates: \( R(t_1) \) through \( R(t_n) \). It is clear from (6.2.1) that the value of the bond is inversely related to the yield of the bond: As yield increases, the bond value declines and vice versa. One way to measure the impact of the yield on the value of the bond is calculate the derivative of the bond price with respect to the yield. From (6.2.1), then, we have

\[
\frac{dB}{dy} = - \left( e^{-yt_1}c_1 t_1 + e^{-yt_2}c_2 t_2 + ... + e^{-yt_n}c_n t_n \right).
\]

This equation can be rewritten as follows:

\[
dB = - \left( e^{-yt_1}c_1 t_1 + e^{-yt_2}c_2 t_2 + ... + e^{-yt_n}c_n t_n \right) dy
\]
or
\[ \frac{dB}{B} = - \frac{e^{-yt_1}c_1 t_1 + e^{-yt_2}c_2 t_2 + \ldots + e^{-yt_n}c_n t_n}{B} dy \]

If we define
\[ D = \frac{e^{-yt_1}c_1 t_1 + e^{-yt_2}c_2 t_2 + \ldots + e^{-yt_n}c_n t_n}{B}, \tag{6.2.2} \]
then, we have
\[ \frac{dB}{B} = -D dy. \tag{6.2.3} \]

The left hand of equation (6.2.3) is the return of the bond: change in the value of the bond as a percentage of the current value of the bond. The equation shows that the bond return is inversely related to the change in bond yield. Furthermore, the magnitude of the negative impact of the yield on the bond is measured by the quantity \( D \), which is called the duration of the bond.

Note that the duration of the bond is a weighted average of the cash-flow payment time:
\[ D = \left( \frac{e^{-yt_1}c_1}{B} \right) t_1 + \left( \frac{e^{-yt_2}c_2}{B} \right) t_2 + \ldots + \left( \frac{e^{-yt_n}c_n}{B} \right) t_n. \]

If the bond is a zero-coupon bond, then the duration is simply the maturity of the bond. For coupon bonds, however, the duration is generally less than the maturity of the bond because coupons are paid before the maturity date. In general, however, the duration increases with the maturity for both zero-coupon and coupon bonds.

### 6.3 Value at Risk (VaR): An Example

From equation (6.2.3) we can see that the return of a bond is negatively related to the change in bond yield. If the yield goes down, then the bond return increases. Furthermore, the magnitude of the return increase is measured by the duration. So, if an investor thinks that the bond yield will go down and she wants to maximize the return of her investment in a bond portfolio, then she should choose a portfolio with a high duration. This can be accomplished by investing in long-term bonds. There is a downside of such an investment strategy, however. If the bond yield goes up rather than goes down, then the bond return will be negative and the higher the duration is,
the large the decline in the return of the bond. So, the strategy of investing in long duration bond portfolio has the potential of generating high positive returns as well as large negative returns. In other words, investing in a bond portfolio with a long duration is highly risky.

How do we quantify the risk involved in the investment in bonds? One apparent way is to calculate the variance of the bond returns. From equation (6.2.3), we have

\[
Var \left( \frac{dB}{B} \right) = D^2 Var(dy).
\]

So, the higher the variance of yields changes and the higher the duration of a bond portfolio, the larger the variance of the bond returns.

In recent years, another measure of risk has been widely used. The measure is called Value at Risk, or VaR. Put it simply, VaR is an upper bound on the potential loss such that the probability that a loss will exceed this upper bound is very small. For a bond portfolio, we know that

\[
dB = -DBdy.
\]

Therefore, over a reasonably short time interval, the change in the portfolio’s value is approximately

\[
\Delta B = -DB\Delta y \quad (6.3.1)
\]

The potential loss of the bond portfolio is measured by \(-\Delta B\). A 99% VaR of the bond portfolio can be defined as follows:

\[
Pr (\text{potential loss from the portfolio} \leq \text{VaR}_{99\%}) = 99%.
\]

From (6.3.1), this equivalent to

\[
Pr (-\Delta B = DB\Delta y \leq \text{VaR}_{99\%}) = 99%,
\]

or

\[
Pr \left( \Delta y \leq \frac{\text{VaR}_{99\%}}{DB} \right) = 99\%. \quad (6.3.2)
\]

So, to find out the VaR of a bond portfolio, we need to know the distribution of yield change, \(\Delta y\). In practice, it is usually assumed that \(\Delta y\) has a normal distribution with mean zero and variance \(\sigma^2 \Delta t\), where \(\sigma^2\) is the annual variance of yield change and \(\Delta t\) is the length of the time interval during which we are calculating the potential loss. Under this assumption, if we define

\[
x = \frac{\Delta y}{\sigma \sqrt{\Delta t}}
\]
then $x$ is a standardized normal variable with mean zero and variance one. The equation (6.3.2) can now be rewritten as the following:

$$\Pr \left( x = \frac{\Delta y}{\sigma \sqrt{\Delta t}} \leq \frac{\text{VaR}_{99\%}}{DB\sigma \sqrt{\Delta t}} \right) = 99\%.$$  

From the normal distribution table, we know that

$$\Pr \left( x = \frac{\Delta y}{\sigma \sqrt{\Delta t}} \leq 2.33 \right) = 99\%.$$  

Thus, we have,

$$\frac{\text{VaR}_{99\%}}{DB\sigma \sqrt{\Delta t}} = 2.33$$

or

$$\text{VaR}_{99\%} = 2.33DB\sigma \sqrt{\Delta t}.$$  

(6.3.3)
Chapter 7

Pricing Forwards and Swaps

7.1 Forwards

Throughout this chapter, we will repeatedly use the following property of no-arbitrage:

\[ P_0(\alpha x_T + \beta y_T) = \alpha P_0(x_T) + \beta P_0(y_T). \]

Here, \( P_0(w_T) \) is the time zero market price of a security whose payoff at time \( T \) is \( w_T \), which could be a random variable if the security is risky.

Let \( S_t \) be the price of an asset at time \( t \). Consider a forward contract on the asset which requires the buyer of the contract to buy the asset at time \( T \) with a fixed price \( K \). The payoff of this forward contract for the buyer (who takes a long position in the contract) is then \( S_T - K \) at time \( T \). Note that having \( K \) dollar at \( T \) is equivalent to having \( K \) units of the zero-coupon bond that pays one dollar at \( T \). Let \( Z_t(T) \) be the market price of this zero coupon bond at time \( t \) \((t \leq T)\). Then, the payoff of the forward contract can be written as \( S_T - KZ_T(T) \) (note that \( Z_T(T) = 1 \)) and therefore the market value of the contract at time zero is

\[ P_0(S_T - KZ_T(T)) = P_0(S_T) - KP_0(Z_T(T)). \]

Let \( r \) be the risk-free, continuously compounding interest rate. Then,

\[ P_0(Z_T(T)) = Z_0(T) = e^{-rT}. \]

So the time zero value of the forward contract is

\[ P_0(S_T) - e^{-rT}K. \]
The forward price of the asset at time zero, $F_0$, is the value of $K$ such that the value of the forward contract is zero. That is, $F_0$ is the solution to the following equation

$$P_0(S_T - F_0 Z_T(T)) = P_0(S_T) - e^{-rT} F_0 = 0,$$

which implies that

$$F_0 = e^{rT} P_0(S_T).$$

So, to determine the value of a forward contract and to determine the forward price of an asset, we need to figure out how to determine $P_0(S_T)$. This is done again by no-arbitrage argument.

1. $S_T$ is the price of a financial asset that pays no income.
   In this case, $S_T$ is simply the payoff from holding the asset until time $T$. Since the cost of buying the asset at time zero is $S_0$, no-arbitrage implies that
   $$P_0(S_T) = S_0$$
   and therefore
   $$P_0(S_T - K Z_T(T)) = S_0 - e^{-rT} K, \text{ and } F_0 = e^{rT} S_0.$$

2. $S_T$ is the price of a financial asset that pays a fixed income $I$ at time $T$.
   In this case, the payoff of holding the asset until time $T$ is $S_T + I Z_T(T)$ instead of $S_T$, so we have the following no-arbitrage condition:
   $$P_0(S_T + I Z_T(T)) = S_0.$$
   But
   $$P_0(S_T + I) = P_0(S_T) + e^{-rT} I.$$
   So, we have
   $$P_0(S_T) = S_0 - e^{-rT} I, \text{ and } F_0 = e^{rT}(S_0 - e^{-rT} I) = e^{rT} S_0 - I.$$

3. $S_T$ is the price of a financial asset that pays a fixed income $I$ at time $T_1 < T$. 

7.1. FORWARDS

Having \( I \) at \( T_1 \) is the same as having \( IZ_{T_1}(T_1) \) at \( T_1 \). So, from (2), we have

\[
S_0 = P_0(S_T + IZ_{T_1}(T_1))
\]
\[
= P_0(S_T) + IP_0(Z_{T_1}(T_1))
\]
\[
= P_0(S_T) + e^{-rT_1}I.
\]

\[
-e^{-rT_1}e^{r(T-T_1)}I = S_0 - e^{-rT_1}I, \text{ and}
\]
\[
F_0 = e^{rT}(S_0 - e^{-rT_1}I) = e^{rT}S_0 - e^{r(T-T_1)}I.
\]

Thus,

\[
P_0(S_T) = S_0 - e^{-rT_1}I
\]

and

\[
F_0 = e^{rT}(S_0 - e^{-rT_1}I) = e^{rT}S_0 - e^{r(T-T_1)}I.
\]

(4) \( S_T \) is the price of a stock that pays a fixed dividend yield \( d \). That is, the payoff of owning the asset from 0 to \( T \) is \( S_T e^{dT} \).

No-arbitrage condition implies that

\[
S_0 = P_0(S_T e^{dT}) = e^{dT}P_0(S_T).
\]

Thus,

\[
P_0(S_T) = e^{-dT}S_0, \text{ and } F_0 = e^{rT}(e^{-dT}S_0) = e^{(r-d)T}S_0.
\]

(5) \( S_T \) is the price of a foreign currency (the exchange rate).

In this case, owning one unit of foreign currency at time zero will give you \( e^{r_fT} \) units of foreign currency at time \( T \), where \( r_f \) is the risk-free interest rate in the foreign country. In domestic dollar units, the payoff is \( S_T e^{r_fT} \).

This is equivalent to an asset that pays a fixed continuous dividend yield \( r_f \).

From (4), then, we have

\[
P_0(S_T) = e^{-r_fT}S_0, \text{ and } F_0 = e^{(r-r_f)T}S_0.
\]

Note that \( F_0 \) is the forward price of a foreign currency, or the forward exchange rate.

(6) \( S_T \) is the price of a commodity.
Holding commodities may be costly. Let $U$ be the (time zero) present value of the storage cost, then, no-arbitrage condition says that

$$U + S_0 = P_0(S_T).$$

Thus,

$$F_0 = e^{rT}(S_0 + U).$$

If the holder of the commodity have to pay storage cost continuously and the cost is proportional to the value of the commodity, then, the present value of holding the commodity to time $T$ is $S_0e^{uT}$, where $u$ is the rate of storage cost. In this case, the no-arbitrage condition is

$$S_0e^{uT} = P_0(S_T),$$

which implies that

$$F_0 = e^{(r+u)T}S_0.$$

Note that effect of the storage cost $u$ on the forward price is like having a negative dividend $-u$.

Sometimes people hold commodities for reasons other than hedging or speculation. For example, a firm may hold inventories of a commodity to meet unexpected future demand. In this case, holding commodity may actually generate some benefits to the holder. As a result, the actual economic cost of holding a commodity is less than $S_0e^{uT}$. In this case, a firm is willing to hold the commodity even if $P_0(S_T) < S_0e^{uT}$. This implies that

$$F_0 = e^{rT}P_0(S_T) < e^{(r+u)T}S_0.$$

That is, since the party who delivers the commodity at $T$ enjoys some benefits from holding the commodity, she is willing to charge a forward price that is lower than implied by pure financial arbitrage. We measure the benefits by the so-called convenience yield $y$, which is the number such that

$$F_0 = e^{(r+u)T}S_0e^{-yT} = e^{(r+u-y)T}S_0.$$

When $F_0 < e^{(r+u)T}S_0$, the convenience yield $y > 0$.

**Problem 1**: Let $Z_t(t_2)$ be the time $t$ price of a zero-coupon bond that pays one dollar at time $t_2$. Consider a forward contract that allows one to buy the zero-coupon bond at $t_1$ ($t_1 < t_2$). Derive the expression for forward price the
zero-coupon bond, $P(t_1, t_2)$. Suppose that $R(0, t)$ is the zero-coupon interest rate for the period from 0 to $t$, (i.e., $Z_0(t) = e^{-R(0,t)t}$), find the forward interest rate $f(t_1, t_2)$ such that

$$P(t_1, t_2) = e^{-f(t_1,t_2)(t_2-t_1)}.$$ 

**Problem 2:** Consider a forward contract on a 5 year coupon bond that settles in one year. Suppose that the bond pays a coupon of $C$ dollar semi-annually and that the zero-coupon interest rate is constant, $r$. Let $B_0$ be the price of the 5 year coupon bond today. Express the forward price as a function of $r$, $C$ and $B_0$.

### 7.2 Swaps

Consider a standard interest rate swap. Let $A$ and $B$ denote the party that pays floating interest rates and the party that pays a fixed interest rate, respectively. Let $T$ be the maturity of the swap, and assume that the frequency of the payment is $m$ times per year. So the total number of payment dates are $n = mT$, and the time interval between two consecutive payments are $\Delta = 1/m$. For example, for a 5 year swap with quarterly payment frequency we have $T = 5$, $m = 4$, $n = 20$ and $\Delta = 0.25$.

Let $t_i = i/m$ be the $i$Th payment date of the swap. Then, we have $t_{i-1} = t_i - \Delta$. (In the example above, $t_1 = 0.25$, $t_2 = 0.5$, ..., $t_{19} = 4.75$ and $t_{20} = 5$. Let $X$ be the notional amount of the swap. For any $t$, let $R_t(\Delta)$ be the annualized LIBOR (London Interbank Offered Rate) of maturity $\Delta$ that is determined in the market at time $t$. By definition, this interest rate is the rate such that the if you borrow 1 dollar in the interbank market at time $t$, you are supposed to repay $1 + \frac{1}{m}R_t(\Delta)$ dollar at $t + \Delta$. By no arbitrage, we have

$$1 = P_t(1 + \frac{1}{m}R_t(\Delta) \text{ at } t + \Delta)$$

Thus, the time $t$ value of a zero-coupon bond with maturity $\Delta$ is

$$Z_t(t+\Delta) = P_t(\$1 \text{ at } t+\Delta) = \frac{1}{(1 + \frac{1}{m}R_t(\Delta))}.$$ 

In the standard interest rate swap, the payments of party $A$ and $B$ on date $t_i$ are $\frac{1}{m}R_{t_{i-1}}(\Delta)X$ and $\frac{1}{mRX}$, respectively. The value of party $B$'s payments
is easy to determine. It is simply a portfolio of \( n \) zero-coupon bonds with maturity \( t_i \) \( i = 1, \ldots, n \) and size \( \frac{1}{m} R X \). So,

\[
V_B = \left[ \sum_{i=1}^{n} Z_0(t_i) \right] \frac{1}{m} R X.
\]

The determination of the value of party A’s payments, however, is more complicated because, except for the first payment, all the future \( n-1 \) payments are uncertain at time 0. However, for the \( i \)th payment, \( \frac{1}{m} R t_{i-1} X \) we know that

\[
P_{t_{i-1}} \left( \frac{1}{m} R t_{i-1} (\Delta) X \right) = X P_{t_{i-1}} \left( \frac{1}{m} R t_{i-1} (\Delta) \right)
= X \left[ P_{t_{i-1}} \left( \frac{1}{m} R t_{i-1} (\Delta) + 1 - 1 \right) \right]
= X \left[ P_{t_{i-1}} \left( \frac{1}{m} R t_{i-1} (\Delta) + 1 \right) - Z_{t_{i-1}} (t_{i-1} + \Delta) \right]
= X \left[ \left( \frac{1}{m} R t_{i-1} (\Delta) + 1 \right) P_{t_{i-1}} (\$1 \text{ at } t_i) - Z_{t_{i-1}} (t_{i-1} + \Delta) \right]
= X \left[ \left( \frac{1}{m} R t_{i-1} (\Delta) + 1 \right) \frac{1}{(1 + \frac{1}{m} R t_{i-1} (\Delta))} - Z_{t_{i-1}} (t_{i-1} + \Delta) \right]
= X \left[ 1 - Z_{t_{i-1}} (t_{i-1} + \Delta) \right] .
\]

Thus, the time zero value of the \( i \)th payment of party A is

\[
P_0 \left( P_{t_{i-1}} \left( \frac{1}{m} R t_{i-1} (\Delta) X \right) \right)
= P_0 \left( X \left[ 1 - Z_{t_{i-1}} (t_{i-1} + \Delta) \right] \right)
= X P_0 (\$1 - Z_{t_{i-1}} (t_{i-1} + \Delta) \text{ at } t_{i-1})
= X \left[ Z_0(t_{i-1}) - Z_0(t_i) \right] .
\]

Note that \( t_{i-1} + \Delta = t_i \). Therefore, the total value of the floating side payments is

\[
V_A = \sum_{i=1}^{n} X \left[ Z_0(t_{i-1}) - Z_0(t_i) \right]
= X \left[ Z_0(t_0) - Z_0(t_n) \right] .
\]
Note that $t_0 = 0$ and $t_n = T$. So, we have

$$V_A = X \left[1 - Z_0(T)\right].$$

So, for the party that receives floating rates, the value of the swap is

$$P_{\text{float}} = V_A - V_B = X \left[1 - Z_0(T)\right] - \left[\sum_{i=1}^{n} Z_0(t_i)\right] \frac{1}{m} RX$$

or

$$P_{\text{float}} = X - \left\{ \left[\sum_{i=1}^{n} Z_0(t_i)\right] \frac{1}{m} RX + Z_0(T)X \right\}$$

Note that the second term (inside the big bracket) is the value of a coupon bond with principal $X$ and annualized coupon rate $\overline{R}$.

When the two parties entered into the swap, the value of the payments on both sides must be the same. So, at the time when the swap is entered, the fixed interest rate $\overline{R}$ must be such that

$$V_A = X \left[1 - Z_0(T)\right] = V_B = \left[\sum_{i=1}^{n} Z_0(t_i)\right] \frac{1}{m} RX,$$

which implies that

$$\overline{R} = m \frac{1 - Z_0(T)}{\sum_{i=1}^{n} Z_0(t_i)}.$$

This fixed rate that makes the swap has zero market value is called the swap rate.

**Problem 3.** Assume that the (continuous compounding) zero coupon interest rates are given by the following table:

1. Calculate the swap rate for 1 year, 2 year and 5 year swaps, respectively. (Assuming semi-annual payment. You can use interpolation to find missing zero-coupon interest rates.)

2. Consider the five year swap. Suppose that after two months, the zero-coupon yield curve remains the same as in the table above. Calculate the value of the swap that was entered two months ago for the party who receives floating interest rates.
CHAPTER 7. PRICING FORWARDS AND SWAPS

<table>
<thead>
<tr>
<th>Maturity</th>
<th>annualized zero coupon rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 month</td>
<td>0.5%</td>
</tr>
<tr>
<td>2 month</td>
<td>0.75%</td>
</tr>
<tr>
<td>3 month</td>
<td>1%</td>
</tr>
<tr>
<td>6 month</td>
<td>1.25%</td>
</tr>
<tr>
<td>9 month</td>
<td>1.5%</td>
</tr>
<tr>
<td>1 year</td>
<td>1.75%</td>
</tr>
<tr>
<td>2 year</td>
<td>2.5%</td>
</tr>
<tr>
<td>3 year</td>
<td>3.5%</td>
</tr>
<tr>
<td>4 year</td>
<td>4.5%</td>
</tr>
<tr>
<td>5 year</td>
<td>5.5%</td>
</tr>
</tbody>
</table>

7.3 Forward Rates and Swaps

If two parties negotiate on a future loan transaction, what interest rate should they agree on so that there would be no arbitrage? Consider the following case: Party A plans to borrow $100 from party B a year from now and will pay the borrowed amount plus interest back to B 3 month later. The interest rate used in the transaction is called the forward rate. Let $F$ be the simple compounding annualized interest rate, then the transaction involves the following cash flows:

- In 1 year, B pays A 100 dollars.
- In 1.25 year, A pays B $100(1 + 0.25F)$ dollars.

For there to be no-arbitrage, the current value of those two payments have to be the same. So, the forward rate $f$ has to satisfy the following condition:

$$P_0(100 \text{ in 1 year}) = P_0(100(1 + 0.25F) \text{ in 1.25 year})$$

Using the definition of zero-coupon bonds, we have

$$P_0(100 \text{ in 1 year}) = 100Z_0(1)$$

and

$$P_0(100(1 + 0.25F) \text{ in 1.25 year}) = 100(1 + 0.25F)Z_0(1.25)$$

Thus, we have

$$100Z_0(1) = 100(1 + 0.25F)Z_0(1.25)$$
which implies

\[ F = \frac{Z_0(1) - Z_0(1.25)}{0.25Z_0(1.25)} \]

More generally, using the argument similar to the one above we can show that the annualized simple compounding forward rate for a loan in time \( t \) that repays back in time \( t' \) is

\[ F(t, t') = \frac{Z_0(t) - Z_0(t')}{(t' - t)Z_0(t')} \]

From the previous section, we know that the value of the \( i \)th floating payment is

\[ X\left[Z_0(t_{i-1}) - Z_0(t_i)\right] \]

From the forward rate formula above we can rewrite the value of the \( i \)th floating payment of a swap as

\[ Z_0(t_i) \Delta F(t_{i-1}, t_i) X \]

That is, the value of receiving the floating rate is the same as receiving the corresponding forward rate. Therefore, the swap rate is the fixed rate such that

\[ \sum_{i=1}^{n} Z_0(t_i) \Delta \left[F(t_{i-1}, t_i) - \overline{R}\right] X = 0 \]

which implies another way of calculating the swap rate:

\[ \overline{R} = \frac{\sum_{i=1}^{n} Z_0(t_i) F(t_{i-1}, t_i)}{\sum_{i=1}^{n} Z_0(t_i)} \]

In other words, the swap rate is a weighted average of the forward rates.

**Problem 4.** Using the table in Problem 3 to calculate the following forward rates: \( F(1, 2), F(2, 3), F(3, 4), F(4, 5) \).

### 7.4 From Simple Compounding to Continuous Compounding

The forward rate that we studied in the previous section is a simple compounding rate. To find the continuous compounding forward rate, \( f(t, t') \),
note the definition is that
\[ e^{(t' - t)f(t, t')} = 1 + (t' - t)F(t, t') \]

Substituting the formula for simple compounding forward rate into the equation above yields the following:
\[ e^{(t' - t)f(t, t')} = 1 + (t' - t) \frac{Z_0(t) - Z_0(t')}{Z_0(t') - Z_0(t)} = \frac{Z_0(t)}{Z_0(t')} \]

Let \( R(t) \) be the continuous compounding zero-coupon rate of maturity \( t \), that is, \( Z_0(t) = e^{-tR(t)} \), then we have
\[ e^{(t' - t)f(t, t')} = \frac{e^{-tR(t)}}{e^{-t'R(t')}} = e^{t'R(t') - tR(t)} \]

which implies the following formula for the continuous compounding forward rate
\[ f(t, t') = \frac{t'R(t') - tR(t)}{t' - t} \].
Chapter 8

Ito Calculus and Black-Scholes Formula

8.1 Ito Process

Financial variables such as stock prices or interest rates are a function of time and some random components which make the future changes of these variables unpredictable. Ito process is a natural way to capture this property of financial variables. We say \( x_t \) follows an Ito process if the change in \( x_t \) over time is given by the following equation:

\[
dx_t = a(x_t, t)dt + b(x_t, t)dz_t
\]

Here, \( a(x_t, t) \) and \( b(x_t, t) \) are two deterministic functions of \( x_t \) and time \( t \), and \( dz_t \) is a normally distributed random variable with mean zero and variance \( dt \). We can calculate the expected value and variance of \( dx_t \) as follows:

\[
E[dx_t] = E[a(x_t, t)dt + b(x_t, t)dz_t] = a(x_t, t)dt + b(x_t, t)E[dz_t] = a(x_t, t)dt
\]

\[
Var[a(x_t, t)dt + b(x_t, t)dz_t] = b^2(x_t, t)Var[dz_t] = b^2(x_t, t)dt
\]

So, the first and second term of (8.1.1) represent the expected and unexpected changes of \( dx_t \) respectively. For (8.1.1) to represent an Ito process, it is also required that for any \( t \) and \( t' \), \( t \neq t' \), \( dz_t \) and \( dz_{t'} \) are independent. That is, the changes of the random variable \( z_t \) over any two disjoint time intervals are independent of each other.
Example 1. Let $S_t$ be the price of a stock at time $t$. It is usually assumed that the change in stock price is governed by the following process:

$$dS_t = \mu S_t dt + \sigma S_t dz_t$$

In this case, $a(S_t, t) = \mu S_t$ and $b(S_t, t) = \sigma S_t$. Under this process, the stock return over the time interval $dt$ is

$$\frac{dS_t}{S_t} = \mu dt + \sigma dz_t$$

That is, the expected stock return is $\mu dt$ and the unexpected stock return is normally distributed with mean zero and variance $\sigma^2 dt$. Parameter $\mu$ is called the annualized expected return and $\sigma$ is called the annual volatility of the stock.

Example 2. Let $r_t$ be the short rate (the interest rate on a loan that will be repaid an instant later) at time $t$. The following is one common way of describing the change of the short rate over time:

$$dr_t = \beta (\bar{r} - r_t) dt + \sigma dz_t$$

Here $\beta$ and $\bar{r}$ are positive constant. So, the expected change in the interest rate is $\beta (\bar{r} - r_t) dt$, which is positive if $r_t$ is lower than $\bar{r}$ and negative if $r_t$ is greater than $\bar{r}$. That is, the interest rate tends to increase when its value is low and to decline if its value is high. For this reason, we call the process a mean-reverting process. The parameter $\beta$ determines the speed of mean reversion and $\bar{r}$ is the long-run mean that the interest rate is reverting to.

8.2 Ito’s Lemma

The price of a derivative security is generally a function of both the time and the price of the underlying security. For example, the price of an option on a stock is a function of both the stock price and the time to expiration date. If the price of the underlying security follows an Ito process, what about the price of the derivative security? Ito’s Lemma shows that it also follows an Ito process.

Ito’s Lemma. Let $x_t$ be an Ito process as defined in (8.1.1) and $y_t$ is a differentiable function of $x_t$ and $t$, $y_t = G(x_t, t)$. Then,

$$dy_t = \frac{\partial G}{\partial x_t} dx_t + \frac{\partial G}{\partial t} dt + \frac{1}{2} \frac{\partial^2 G}{\partial x_t^2} \sigma^2 (x_t, t) dt$$
or equivalently,

\[ dy_t = \left[ \frac{\partial G}{\partial x_t} a(x_t, t) + \frac{\partial G}{\partial t} + \frac{1}{2} \frac{\partial^2 G}{\partial x_t^2} b^2(x_t, t) \right] dt + \frac{\partial G}{\partial x_t} b(x_t, t) dz_t \]

**Example 3.** Let \( S_t \) be the price of a stock that follows the process in Example 1. Let \( F_t \) be the time \( t \) forward price of the stock that settles at time \( T \). From Chapter 7 we know that \( F_t = e^{r(T-t)} S_t \). So, in this case, we have

\[ \frac{\partial F_t}{\partial S_t} = e^{r(T-t)}, \quad \frac{\partial F_t}{\partial t} = -re^{r(T-t)} S_t, \quad \frac{\partial^2 F_t}{\partial S_t^2} = 0 \]

Thus, from Ito’s Lemma, we have

\[ dF_t = \left[ e^{r(T-t)} \mu S_t - re^{r(T-t)} \mu S_t \right] dt + e^{r(T-t)} \sigma S_t dz_t \]

or

\[ dF_t = (\mu - r) F_t dt + \sigma F_t dz_t \]

**Example 4.** \( y_t = \ln S_t \). In this case,

\[ \frac{\partial y_t}{\partial S_t} = S_t^{-1}, \quad \frac{\partial y_t}{\partial t} = 0, \quad \frac{\partial^2 y_t}{\partial S_t^2} = -S_t^{-2} \]

From Ito’s Lemma, we have

\[ d\ln S_t = \left[ S_t^{-1} \mu S_t - \frac{1}{2} S_t^{-2} \sigma^2 S_t^2 \right] dt + S_t^{-1} \sigma S_t dz_t \]

or

\[ d\ln S_t = \left( \mu - \frac{1}{2} \sigma^2 \right) dt + \sigma dz_t \]

**Problem 1.** Suppose that \( E_t \) is the Canadian dollar price of one US dollar and that

\[ dE_t = \mu E_t dt + \sigma E_t dz_t \]

What process does the US dollar price of one Canadian dollar follow?
8.3 Risk-neutral Pricing

From example 4 in the last section, we know that the infinitesimal change in the log of stock price is normally distributed with mean \((\mu - \frac{1}{2}\sigma^2) dt\) and variance \(\sigma^2 dt\). Integrating the infinitesimal changes to get the change from 0 to \(T\), we have

\[
\ln S_T - \ln S_0 = \int_0^T \left( \mu - \frac{1}{2}\sigma^2 \right) dt + \int_0^T \sigma d\tau = \left( \mu - \frac{1}{2}\sigma^2 \right) T + \sigma (Z_T - Z_0)
\]

Since \(Z_T - Z_0\) is normally distributed with mean zero and variance \(T\). Let \(u = (Z_T - Z_0)/\sqrt{T}\), then \(u\) is normally distributed with mean zero and variance one. From the equation above, we have,

\[
\ln S_T = \ln S_0 + \left( \mu - \frac{1}{2}\sigma^2 \right) T + \sigma \sqrt{T} u
\]

or

\[
S_T = S_0 e^{\left( \mu - \frac{1}{2}\sigma^2 \right) T + \sigma \sqrt{T} u} \tag{8.3.1}
\]

That is, the future stock price \(S_T\) has a log-normal distribution.

Consider a derivative contract that has a payoff function \(h(S_T)\) at time \(T\). From Chapter 3 and 4 we know that the value of the derivative contract is the expected payoff discounted by the risk-free interest rate. Here, the expectation is taken under the risk-neutral probability distribution. That is,

\[
P_0 \left( h(S_T) \right) = e^{-rT} \hat{E} \left[ h(S_T) \right]
\]

where \(r\) is the risk-free interest rate and \(\hat{E}\) means expectation under the risk-neutral distribution. What is the risk-neutral distribution of the stock price \(S_T\)? It can be shown that it is still a log-normal distribution like in equation (8.3.1) but with possibly a different expected return parameter. For a stock that does not pay dividend, instead of \(\mu\), the expected return under the risk-neutral distribution is the risk-free interest rate \(r\). So, under the risk-neutral distribution, we have

\[
S_T = S_0 e^{(r - \frac{1}{2}\sigma^2) T + \sigma \sqrt{T} u} \tag{8.3.2}
\]

Thus, the time zero price of an European call option on the stock with strike price \(K\) is

\[
c_0 = e^{-rT} \hat{E} \left[ \max (S_T - K, 0) \right]
\]
where the expectation is taken under the assumption that $S_T$ follows the distribution given in equation (8.3.2). The result is the Black-Scholes formula:

**Black-Scholes Formula:**

$$c_0 = S_0 N(d_1) - e^{-rT} K N(d_2)$$

Here, $N(.)$ is the cumulative standard normal distribution function and

$$d_1 = \frac{\ln (S_0/K) + (r + \frac{1}{2} \sigma^2)T}{\sigma \sqrt{T}}$$

$$d_2 = \frac{\ln (S_0/K) + (r - \frac{1}{2} \sigma^2)T}{\sigma \sqrt{T}}$$

**Delta:** From the Black-Scholes formula it can easily be shown that

$$\frac{\partial c_0}{\partial S_0} = N(d_1)$$

This partial derivative is called the delta of the call option. It measures the sensitivity of the option price to the change in the price of the underlying stock. We will show below that calculating delta is important for hedging option against stock market risk.

**Put-Call Parity and Pricing of Put Options:**

Since the payoffs of the call and put options are $c_T = \max(S_T - K, 0)$ and $p_T = \max(K - S_T, 0)$, respectively. We have

$$c_T - p_T = S_T - K$$

By no-arbitrage principle,

$$c_0 - p_0 = P_0(S_T - K) = S_0 - e^{-rT} K$$

This is called the put-call parity. From this parity, we can derive the price and delta of the put option:

$$p_0 = c_0 - S_0 + e^{-rT} K$$

$$\frac{\partial p_0}{\partial S_0} = N(d_1) - 1$$
Problem 2. Show that the probability of exercising the call option \( \text{Prob} [S_T > K] \) is \( N(d_2) \).

Problem 3. Suppose that the stock price is currently $90 and the risk-free interest rate is 2%. Calculate the value of an option that pays $1 if the stock price exceeds $100 one year from now.

8.4 Merton’s No-Arbitrage Argument

Instead of using the risk-neutral pricing method above, the Black-Scholes formula was derived using an no-arbitrage argument suggested by Robert Merton. Let \( f_t = F(S_t, t) \) be the price of a derivative security. From Ito’s Lemma we have

\[
dF = \left[ \frac{\partial F}{\partial S_t} \mu S_t + \frac{\partial F}{\partial t} + \frac{1}{2} \frac{\partial^2 F}{\partial S_t^2} \sigma^2 S_t^2 \right] dt + \frac{\partial F}{\partial S_t} \mu S_t \, dz_t
\]

That is, the unexpected change in the price of the derivative security is also a linear function of \( dz_t \), the same shock that moves the stock price change unexpectedly. This suggests that one can form a risk-neutral portfolio by shorting an appropriate amount of the underlying stock. Let \( \delta = \frac{\partial f}{\partial S_t} \) and \( A_t = f_t - \delta S_t \). From the formula above we have

\[
dA_t = df_t - \delta dS_t = \left[ \frac{\partial F}{\partial t} + \frac{1}{2} \frac{\partial^2 F}{\partial S_t^2} \sigma^2 S_t^2 \right] dt
\]

That is, the change in the value of the portfolio is fully predictable or risk-free over the infinitesimal time interval \( dt \). By no-arbitrage principle, the return of the portfolio has to be the same as the risk-free interest rate. So, we have

\[
dA_t = rA_t \, dt
\]

or

\[
\frac{\partial F}{\partial t} + \frac{1}{2} \frac{\partial^2 F}{\partial S_t^2} \sigma^2 S_t^2 = rF - r \frac{\partial F}{\partial S_t} S_t
\]

Rewriting the equation as follows:

\[
\frac{\partial F}{\partial t} + \frac{\partial F}{\partial S_t} rS_t + \frac{1}{2} \frac{\partial^2 F}{\partial S_t^2} \sigma^2 S_t^2 = rF
\]

This is a second-order partial differential equation for the unknown function \( f \). It has exactly the same form as the standard heat equation in physics.
Whether this equation has an analytical solution depends on the terminal condition (or payoff at the terminal date) for the function $F$. For the European call option, the terminal condition is $F(S_T, T) = \max(S_T - K, 0)$. It happens that this terminal condition does yield an analytical solution to the equation and the result is the Black-Scholes formula.

An important insight of the argument presented here is the importance of delta, the amount of underlying security that one needs to short in order to hedge the derivative security.
Chapter 9
Value at Risk

In this chapter we discuss the general method of calculation value-at-risk (VaR) that can be applied to bonds, stocks and any derivative securities.

Let $P_t$ be the market value of a portfolio, which may include fixed income, equity, derivatives and other securities. The general rule of thumb is that the

$$\text{VaR}_{99\%} = 2.37 \sqrt{\text{Variance}(dP_t)} \quad (9.0.1)$$

For each problem at hand, the question is how to figure out the variance of the change in market value of the portfolio.

**Example 1.** If $P_t$ is a fixed income portfolio, then

$$dP_t = -DP_t dy_t$$

and

$$\text{Variance}(dP_t) = \text{Variance}(-DP_t dy_t) = D^2 P_t^2 \sigma_y^2 dt$$

where $D$ is the duration of the portfolio, $y_t$ the yield, and $\sigma_y$ the annual volatility of the yield change.

**Example 2.** If $P_t$ is a stock, then,

$$dP_t = \mu P_t dt + \sigma P_t dz_t$$

and

$$\text{Variance}(dP_t) = \text{Variance}(\sigma P_t dz_t) = \sigma^2 P_t^2 dt.$$
Example 3. If $P_t$ is a derivative contract on stock. Then $P_t = F(S_t, t)$ for some function $F$. By Ito’s Lemma, we have

$$dP_t = \left[ \frac{\partial F(S_t, t)}{\partial S_t} \mu S_t + \frac{\partial F(S_t, t)}{\partial t} + \frac{1}{2} \frac{\partial^2 F(S_t, t)}{\partial S_t^2} \sigma^2 S_t^2 \right] dt + \frac{\partial F(S_t, t)}{\partial S_t} \sigma S_t dz_t$$

and

$$\text{Variance}(dP_t) = \text{Variance} \left( \frac{\partial F(S_t, t)}{\partial S_t} \sigma S_t dz_t \right) = \sigma^2 S_t^2 \left( \frac{\partial F(S_t, t)}{\partial S_t} \right)^2 dt.$$ 

Example 4. If $P_t$ is the market value of a swap contract that pays fixed and receives floating interest rates with notional principal being 100, then,

$$P_t = 100 - B_t$$

where $B_t$ is the market value of a bond that pays fixed interest rate and 100 is the value of the floating payments plus principal. Thus,

$$dP_t = -dB_t$$

and the variance of $dP_t$ is the same as the variance of $dB_t$, which we already discussed in example 1.
Chapter 10

Credit Risk

Credit risk is about the risk of default. It applies to loans, bonds and swaps. We usually assume that the treasury bonds issued by the US federal government has little credit risk and use their prices and yields as benchmarks. For all other borrowers, the interests on their loans or yields on the bonds they issue are higher than the yields on the treasury bonds of corresponding maturities. For example, on March 15, 2011, the three month interest rate on a T-bill (US treasury bill) is 0.102%, but the three-month LIBOR, which is the interest rate charged on three month loans among banks, is 0.309%. So the borrowing cost of a bank is 20 basis points higher than that of the US government. This difference, called the interest rate spread, is demanded by the lenders to compensate the risk of default that they bear. Similarly, the coupon rate on a five-year treasury bond issued on that day is 2.125%, while the coupon rate on a five-year bond issued by J.P.Morgan (Rated A+) is 3.45%. So the yield spread between the A+ rated firm and the treasury is 1.315%, which is again a compensation for the risk of default borne by the bond investors.

The simplest corporate bonds are non-callable bonds. That is, the firm that issues the bond does not have the option to buy back the bonds from the investors at a pre-specified price. There are two types of non-callable corporate debt: fixed coupon bonds and floating notes. The first type are bonds that pays a fixed coupon rate that is usually quoted as the treasury plus a spread and the second type pays a floating rate that equals LIBOR plus a spread. For example, an eight-year fixed coupon bond issued by an AAA rated firm in 1983 paid treasury plus 60 basis points (0.6%) coupon, and a floating notes of the same maturity issued that year paid LIBOR plus
25 basis points.

In the example above, the yield spread on the fixed-coupon bond is higher than the yield spread on the floating note. But this is not always the case. Sometimes the spread on a floating note may be higher than the spread on a fixed-coupon bond with the same maturity. In general, if the credit risk is independent of the interest rate risk, then the spread on floating notes will be higher (lower) if the forward rate increase (decrease) over time. When forward rate increases over time, investor expects more payments in the future, but the risk of default increases with time. So the investors of floating notes have more exposure to credit risk and therefore demand a higher spread.

In summary, when investors face the risk of default on bonds issued by a firm, they demand a spread over the treasury rate as a compensation for the risk. The higher the risk of default, the higher the spread would be. So, the yield spread between a corporate bond and a treasury bond of the same maturity contains information on investor’s view of the credit risk they face. In the following we discuss how these information can be used to price a credit derivative instrument called credit default swap.

## 10.1 Credit Default Swaps

Credit Default Swaps (CDS) is a credit derivative that is used to hedge or speculate on credit risk (the risk of default). There are CDS on a particular firm or institution (single name CDS) and CDS on a grouped of firms or institutions. Here we will discuss CDS on a particular firm or institution only. A CDS is a transaction between two parties. One party buys protection against the default risk by paying selling party a constant premium at regular time intervals for a pre-specified length of period. During the period, if default occurs, the seller of the CDS will compensate the buyer the full amount of his loss due to default. So a CDS is a derivative contract based on a risky bond or risky loan.

Let’s first define some notations:

- $P(t_1, t_2) = \text{forward price of risky zero coupon bond issued at } t_1 \text{ that promises to pay } $1 \text{ at } t_2$. (Think of this as a bond issued by a firm, a
government or a financial institution that has potential risk of default. Like in the case of risk-free bonds, one can use the bootstrap methods to calculate the zero-coupon interest rates of corporate bond by using the information on the yield of risky coupon bonds.)

- \( P^*(t_1, t_2) \) = forward price of a risk-free zero coupon bond issued at \( t_1 \) that will pay $1 at \( t_2 \). (Think of this as a bond issued by the US Treasury.)
- \( f(t_1, t_2) \) = risky forward rate (i.e. \( P(t_1, t_2) = e^{-f(t_1,t_2)(t_2-t_1)} \).)
- \( f^*(t_1, t_2) \) = risk-free forward rate (i.e. \( P^*(t_1, t_2) = e^{-f^*(t_1,t_2)(t_2-t_1)} \).)
- \( s(t_1, t_2) = f(t_1, t_2) - f^*(t_1, t_2) \) = forward spread between risky and risk-free forward rates.
- \( h(t_1, t_2) \) = expected proportional loss from the risky bond between \( t_1 \) and \( t_2 \).
- \( Q(t_1, t_2) \) = survival probability (or probability of no default) between \( t_1 \) and \( t_2 \).
- \( \delta \) = recovery rate.

Calculating \( h(t_1, t_2) \) from forward spread: The expected proportional loss should be embedded in the market price of the risky bond. So, we have

\[
P(t_1, t_2) = P^*(t_1, t_2) (1 - h(t_1, t_2))
\]

(Note that here the expectation is taken under the risk-neutral probability distribution.) Thus,

\[
h(t_1, t_2) = 1 - \frac{P(t_1, t_2)}{P^*(t_1, t_2)} = 1 - e^{-s(t_1,t_2)(t_2-t_1)}.
\]

Finding survival probability: Probability of default is \( 1 - Q(t_1, t_2) \). If the issuer does not default, full promised value will be received. If the issuer defaults, then only proportion \( \delta \) of the promised value will be recovered. So, the expected value of the risky bond is

\[
P(t_1, t_2) = Q(t_1, t_2)P^*(t_1, t_2) + [1 - Q(t_1, t_2)] \delta P^*(t_1, t_2)
\]
So,

\[ P^*(t_1, t_2) (1 - h(t_1, t_2)) = Q(t_1, t_2)P^*(t_1, t_2) + [1 - Q(t_1, t_2)] \delta P^*(t_1, t_2) \]

which implies

\[ 1 - h(t_1, t_2) = \delta + Q(t_1, t_2)(1 - \delta) \]

or

\[ Q(t_1, t_2) = 1 - \frac{h(t_1, t_2)}{1 - \delta} \]

So, if one can estimate the recovery rate \( \delta \), then survival probability can be calculated from the equation above.

Credit Default Swap with maturity \( T \):

For a buyer of credit protection, the party pays a constant amount \( c \), which is called spread, at time \( t_i \) if there has been no default up until \( t_i \), for each of the \( n \) prespecified time: \( t_1, ..., t_n (= T) \).

If default occurs between \( t_{i-1} \) and \( t_i \), the buyer receives a payment that equals the losses from default at \( t_i \).

The present expected value of out payments is

\[ V_A = c \sum_{i=1}^{n} P^*(0, t_i)Q(0, t_i) \]

and the value of potential payments received is

\[ V_B = h(0, T)P^*(0, T) \]

Thus, the no-arbitrage value of the spread of CDS is the value such that \( V_A = V_B \), or

\[ c = \frac{h(0, T)P^*(0, T)}{\sum_{i=1}^{n} P^*(0, t_i)Q(0, t_i)} \]

Example. Suppose that zero-rates of treasury bond are 5% for all maturities. Now suppose that the zero-rates for risky bond are 7% for all maturities.

1. Calculate the expected proportional loss for any period \( (t_1, t_2) \).

\* The forward spread is 2% for any \( t_1 \) and \( t_2 \). So, \( h(t_1, t_2) = 1 - e^{-0.02(t_2-t_1)} \)
2. If the recovery rate is 0.5, what is the survival probability between $(t_1, t_2)$?

   $Q(t_1, t_2) = 1 - \frac{h(t_1, t_2)}{\delta} = 1 - 2 \left( 1 - e^{-0.02(t_2-t_1)} \right) = 2e^{-0.02(t_2-t_1)} - 1$

3. What is the market spread for a CDS? (Assuming annual payment frequency.)

   \[ c = \frac{\left[1 - e^{-0.02T}\right] e^{-0.05T}}{\sum_{i=1}^{n} e^{-0.05t_i} \left[2e^{-0.02t_i} - 1\right]} \]

   - See attached excel spreadsheet for calculating the CDS spread for different maturities.

When there is an active market of CDS, the formula for the CDS spread can also be used along with yield spread to estimate the recovery rate. To see this, we substitute the formula for the survival probability into the CDS spread formula to arrive at the following equation:

\[ c = \frac{h(0, T) P^*(0, T)}{\sum_{i=1}^{n} P^*(0, t_i) \left[1 - \frac{h(0, t_i)}{\delta}\right]} \]

Since the information on yield spreads will allow us to calculate $h(0, t_i)$ for all $i$, the only unknown in the right hand side of this equation is the recovery rate $\delta$. If we know what the value of CDS spread $c$ is in the market, then we can use this equation to find the implied value of $\delta$ that will match the predicted $c$ with the observed $c$.

### 10.2 Using Stock Market Information to Assess Credit Risk

In the previous section we discussed how to use the information on yield spreads to assess a company’s credit risk. However, for some companies there may not be an active market for the bonds they issued. As a result, we may not have up to date information about the market value of its debt, which is needed to determine yield spreads. In these cases, Robert Merton
suggested a method that utilizes a company’s stock price to assess its credit risk. The method is based on the idea that a firm’s value is the sum of equity and debt and the former is effectively an option on the underlying assets of the firm.

Let $A_t$ be the value of the firm’s underlying asset at time $t$ and $F$ be the value of outstanding debt that has to be paid back at some specific time $T$. Then, the equity and debt values at time $T$ are

$$E_T = \max\{A_T - F, 0\} \quad D_T = \min\{F, A_T\}$$

What we want to know is the survival probability evaluated at time $t = 0$ (the probability of default is just one minus the survival probability):

$$Q = \text{Prob}_0 [A_T \geq F]$$

We assume that the underlying asset follows the following Ito process under the risk-neutral distribution:

$$dA_t = \tilde{r}A_t dt + \sigma_A A_t dz_t$$

Here $\tilde{r} = r - \lambda$, $r$ is the risk-free interest rate and $\lambda$ is the cash payout rate (e.g., the dividend/asset ratio). From Problem 2 of Chapter 8 we know that

$$Q = N(d_2)$$

where $N(.)$ is the standard normal cumulative distribution function and

$$d_2 = \frac{\ln(A_0/F) + (\tilde{r} - \frac{1}{2}\sigma_A^2)T}{\sigma_A \sqrt{T}}$$

However, we may not have good information about the current value of the assets, $A_0$, and its volatility, $\sigma_A$. Fortunately, the Black-Scholes formula for options allows us to infer those values from the current price and volatility of the firm’s equity.

From the Black-Scholes formula, we have

$$E_0 = A_0 e^{-\lambda T} N(d_1) - e^{-rT} F N(d_2)$$

(10.2.1)

where

$$d_1 = d_2 + \sigma_A \sqrt{T}$$

\footnote{For simplicity we assume that there is no tax nor default cost.}
Note that the formula here is slightly different from the formula in Chapter 8 because we allow for the asset to have cash payouts. Using Ito’s Lemma, we also have

$$Var[dE_0] = \left(\frac{\partial E}{\partial A}\right)^2 \sigma^2 A^2$$

and we also know that

$$\frac{\partial E}{\partial A} = e^{-\lambda T} N(d_1)$$

So, the volatility of the equity price is

$$\sigma_E = \sqrt{Var[dE_0]} \frac{\partial E}{A E_0} \frac{A_0}{\partial A} \sigma_A \quad (10.2.2)$$

Since we can observe equity price $E_0$ and estimate its volatility $\sigma_E$, equation (10.2.1) and (10.2.2) can be used to jointly solve for the values of $A_0$ and $\sigma_A$.

**Problem 1.** There are two companies that have the same current value of assets, the same debt outstanding and the same cash payout rate. If one company has higher asset volatility than the other one, what do you think would be the differences in equity price and default probability of the two companies?

**Problem 2.** Use Ito’s Lemma to derive equation (10.2.2).
Chapter 11

Using Swaps to Hedge Interest Rate Risk

In this chapter we discuss two examples when interest rate swap can help a firm to hedge its interest rate risk.

Hedging Interest Rate Risk by Mortgage Banks

Mortgage banks take deposits from households and then lend the money out as mortgages to home buyers. Because the deposits are usually short term, the interest rates paid on deposits have to move with the market. Otherwise, depositors will take their money elsewhere. Mortgages, on the other hand, are long-term loans and many of them have a fixed interest rate. Suppose that a mortgage bank received from depositors 100 million dollars and lent them out as 5 year mortgages at a fixed rate equals the 5-year treasury rate plus a spread. Suppose that the bank has some reserve money in its safe so that it can meet the demand for occasional cash withdraws from the depositors. Also suppose that there will always be new depositors replacing old depositor so that the total value of the pool of deposits stay at 100 million dollars. Finally, suppose that the average term of the deposits is 3-month and the average interest rate the bank pays to the depositors is the 3-month LIBOR.

On March 17, for example, the 5-year treasury yield is 1.909% and the 3-month LIBOR is 0.3%. If the bank started this business today, they probably could charge a fixed rate of 3.5% on the five year mortgages. So, if
the LIBOR stays at its current level, the bank will make a profit of \((3.5-0.3)\% \times 100\text{ million} = 3.2\text{ million dollars} \) each year. In reality, however, LIBOR may go down or go up in the next five years. If the LIBOR goes above 3.5% within the next five year, however, the bank will start to lose money. To hedge this risk, the bank may enter into a fix for floating swap. On the swap market, the fixed rate on a 5 year swap for a AAA rated bank is 2.35%. This bank is rated slightly lower than AAA and therefore it has to pay a higher fixed rate, say 2.85%, on the 5-year swap. As a result of entering the swap, the firm will now make a profit of \((3.5-2.85)\% \times 100\text{ million} = 650,000\text{ each year} \) for sure, independent of how the LIBOR will fluctuate in the future—if the mortgage borrowers will not repay their loans early.

If the LIBOR continues to be low, some borrowers may want to switch to a mortgage with an adjustable rate rather than paying the fixed rate. Let’s say that 20 million worth of the mortgage is paid off in one year. Then, after one year, the bank will receive \(3.5\% \times 80\text{ million} = 2.8\text{ million dollars} \) each year. But the bank still have to honor the swap and pay \(2.85\% \times 100\text{ million} = 2.85\text{ million} \) each year. In that case, the bank would lose 50,000 dollars each year. To hedge this pre-payment risk, the bank may choose to include a clause in the swap contract that gives it the option to reduce the notional amount of the swap by 20 million dollars at any time. This is equivalent to have a standard swap with a notional amount of 80 million dollars and another swap with a notional amount of 20 million dollars plus an option to cancel the swap. The later is called a cancelable swap. The option to cancel the swap is called a swaption. Of course, the bank has to pay a premium for the option and it will only have the option if the cost of it is less than the potential loss of 50,000 dollars per year.

### Using Swaps to Reduce Financing Cost

Consider the following case. There are two companies, A and B. Both companies can raise funds from the corporate debt market in two alternative ways: issuing a fixed coupon bond that pays a coupon rate equals the treasury rate plus a spread \(S^i_{\text{fix}}\) or a floating note that pays a coupon rate equals the LIBOR plus a spread \(S^i_{\text{float}}\). Here \(i = A\) or \(B\). Suppose that company A has a better credit rating than company B so that \(S^A_{\text{fix}} > S^B_{\text{fix}}\) and \(S^A_{\text{float}} > S^B_{\text{float}}\). Furthermore, suppose that company B prefers to have a fixed interest payment on its debt and company A prefers to have a floating payment on its
debt. Company B has two alternatives:

1. Issue a fixed coupon bond.

2. Issue a floating note, but then enter into an interest rate swap with company A so that it will receive LIBOR and pay a fixed rate that equals treasury rate plus $\overline{S}$.

Let $y$ be the treasury rate. The financing cost for B is $y + S_{\text{fix}}^B$ and $S_{\text{float}}^B + y + \overline{S}$ in the second case. It will choose the second financing method if $y + S_{\text{fix}}^B > S_{\text{float}}^B + y + \overline{S}$ or equivalently

$$\overline{S} < S_{\text{fix}}^B - S_{\text{float}}^B \quad (11.0.1)$$

Company A also has two alternatives:

1. Issue a floating note.

2. Issue a fixed coupon bond, but then enter into an interest rate swap with company B so that it will receive treasury rate plus $\overline{S}$ pay LIBOR.

The financing cost for B is $\text{LIBOR} + S_{\text{float}}^A$ and $y + S_{\text{fix}}^A + \text{LIBOR} - (y + \overline{S})$ in the second case. It will choose the second financing method if $S_{\text{fix}}^A - \overline{S} < S_{\text{float}}^A$ or equivalently

$$\overline{S} > S_{\text{fix}}^A - S_{\text{float}}^A \quad (11.0.2)$$

For both companies to be willing to enter into the interest rate swap, then, condition (11.0.1) and (11.0.2) both have to hold. This is only possible if

$$S_{\text{fix}}^A - S_{\text{float}}^A < S_{\text{fix}}^B - S_{\text{float}}^B$$

In this case, it is usually said that Company B has a comparative advantage in issuing floating notes and company A has a comparative advantage in issuing fixed coupon bonds. Any $\overline{S}$ between $S_{\text{fix}}^A - S_{\text{float}}^A$ and $S_{\text{fix}}^B - S_{\text{float}}^B$ will make the two companies to be willing to do the swap. The gain from the swap for A is $S_{\text{float}}^A - (S_{\text{fix}}^A - \overline{S})$ and the gain from the swap for B is $(S_{\text{fix}}^B - S_{\text{float}}^B) - \overline{S}$. So the total gain is $(S_{\text{fix}}^B - S_{\text{float}}^B) - (S_{\text{fix}}^A - S_{\text{float}}^A)$.

**Example.** In 1983, B.F. Goodrich was a US tire company that had a BBB rating and Rabobank-Nederland was an AAA rated financial institution in Holland. The market conditions was such that for a $50 million debt, $S_{\text{fix}}^A = \ldots$
0.6%, $S^A_{\text{float}} = 0.25\%$, $S^B_{\text{fix}} = 2.4\%$ and $S^B_{\text{float}} = 0.5\%$. (The yield curve was inverted–short term rate was higher than long-term rate–at that time so that the spread on floating notes are lower than on fixed coupon bonds.) In this case, the total gain is $(2.4\%-0.5\%)-(0.6\%-0.25\%)=1.55\%$. That is, by doing the swap, the two companies would save $1.55\%\times50\text{ million} =775,000\text{ dollars}$ each year. Such a large “arbitrage” opportunity exists partly because the two companies are in two different bond markets (US and Europe) and back then the market integration was not as high as today and the interest rate swaps, was just starting to be used by banks to arbitrage the differences between the two markets. Today, it is quite rare that such a large total gain could be realized by using interest rate swaps.
Chapter 12

Using Options to Hedge Uncertain Price Exposures

If a firm knows exactly its exposure to some market price risk, it can use either forwards or options to hedge the Market risk. In this chapter we discuss how options can be used to hedge price risk when the firm’s exposure is uncertain.

12.1 Hedging in the Cases of Known Exposures

We first start with an example when the exposure is known. A gold mining company that knows it will have 10,000 ounces of gold to sell in a fixed time in the future, $T$. So the firm’s revenue at $T$ will be $10,000S_T$, where $S_T$ is the gold price at $T$. To hedge the gold price risk, it can simply take a short position of a forward contract on 10,000 ounces of gold that settles at $T$. Let $F$ be the forward price, the firm’s revenue after hedging will be $10,000S_T + 10,000(F - S_T) = 10,000F$. This will completely eliminate the firm’s exposure to gold price risk. Alternatively, if the firm does not want to give up the upside gains in the case when gold price increases but still wants to cover the down side risk of declining gold price, it can buy an European put option for selling 10,000 ounces of gold at $T$ at a pre-specified price $K$. So the after hedging revenue will be $10,000S_T + 10,000max\{K - S_T, 0\} = 10,000max(S_T, K)$. (Of course, the firm has to pay a premium upfront to have the option.)
12.2 Hedging in the Cases of Unknown Exposures

In reality, however, the firm’s exposure to price risk may be uncertain. For example, the firm may know how much gold it will be able to produce, but does not know exactly when they will be available to the market. In that case, the firm can still choose to use a short position forward contract to hedge the risk as long as it chooses a settlement date $T$ that is sufficiently far in the future so that it can be sure that the gold will be available by then. Alternatively, the firm can use an American style put option so that it can choose any time before $T$ to sell the gold at a prespecified price $K$. In this case the firm can sell the gold at the time when it’s available.

Sometimes the firm may not even sure how much gold it will have. Suppose that the firm only knows that the amount of gold it will have at $T$ is between 9,000 to 11,000 ounces. In that case, the firm’s exposure is somewhere between $9,000S_T$ and $11,000S_T$. In this case, the firm can buy a put option to sell 11,000 ounces of gold at a fixed price $K$. If the gold price is above $K$, then the firm will let the option expire and sell whatever amount of gold it has at the market price. If the gold price is below $K$, the firm will sell 11,000 ounces of gold at price $K$ for sure: even if the firm production results in less than 11,000 ounces of gold, the firm can buy the rest of the gold from the market and then sell to the other party of the option contract at price $K$. This is profitable since the market price is below $K$. So, in this case, the firm’s revenue will be

\[
\begin{align*}
Q S_T, & \quad \text{if } S_T \geq K \\
Q K + (11,000 - Q)(K - S_T), & \quad \text{if } S_T < K
\end{align*}
\]

With the put option, the firm ensures that its revenue will be greater or equal to $QK$ for any $Q$ between 9,000 and 11,000.

The option used above to cover the downside risk could be quite expensive. To reduce the hedging cost, the firm may also short a forward contract on 9,000 ounces of gold and buy a put option on 2,000 ounces of gold. In this case, the cost of hedging will only be $2/9$ of the original hedging cost. The firm’s after-hedging revenue will be

\[
\begin{align*}
9,000K + (Q - 9,000)S_T, & \quad \text{if } S_T \geq K \\
9,000K + (Q - 9,000)K + (11,000 - Q)(K - S_T), & \quad \text{if } S_T < K
\end{align*}
\]
Note that $9,000K + (Q - 9,000)K + (11,000 - Q)(K - S_T) = QK + (11,000 - Q)(K - S_T)$. In this case, the firm also ensures that its revenue is greater or equal to $QK$.

**Problem.** Suppose a firm can produce a car in either Canada or US. The cost in Canada is $20,000 Canadian dollar and the cost in US is $19,000 US dollar. Currently the Canada-US exchange rate is 1. The firm will start to produce in six month. Whether the firm will produce in the US or Canadian depends on the exchange rate. Let $E_t$ be the Canadian dollar price of one US dollar. Then, the firm will produce in US if and only if $19,000E_T < 20,000$ ($T = 0.5$). So, the firm’s cost is a function of the exchange rate as well:

$$C_T = \min \{20,000, 19,000E_T\}$$

If the firm would like to hedge the exchange rate, what derivative contract should it use?