Sweden, Switzerland, the United Kingdom, and the United States. It met regularly in Basel, Switzerland, under the patronage of the Bank for International Settlements. The first major result of these meetings was a document titled "International Convergence of Capital Measurement and Capital Standards." This was referred to as "The 1988 BIS Accord" or just "The Accord." Later it became known as Basel I.

15.3 The 1988 BIS Accord

The 1988 BIS Accord was the first attempt to set international risk-based standards for capital adequacy. It has been subject to much criticism as being too simple and somewhat arbitrary. In fact, the Accord was a huge achievement. It was signed by all 12 members of the Basel Committee and paved the way for significant increases in the resources banks devote to measuring, understanding, and managing risks. The key innovation in the 1988 Accord was the Cooke ratio.

15.3.1 The Cooke Ratio

The Cooke ratio[2] considers credit risk exposures that are both on-balance-sheet and off-balance-sheet. It is based on what is known as the bank's total *risk-weighted assets* (also sometimes referred to as the *risk-weighted amount*). This is a measure of the bank's total credit exposure.

Credit risk exposures can be divided into three categories:

1. Those arising from on-balance sheet assets (excluding derivatives)
2. Those arising for off-balance sheet items (excluding derivatives)
3. Those arising from over-the-counter derivatives

Consider the first category. Each on-balance-sheet asset is assigned a risk weight reflecting its credit risk. A sample of the risk weights specified in the Accord is shown in Table 15.1. Cash and securities issued by governments of OECD countries (members of the Organisation for Economic Co-operation and Development) are considered to have virtually zero risk and have a risk weight of zero. Loans to corporations have a risk weight of 100%. Loans to banks and government agencies in OECD countries have a risk weight of 20%. Uninsured residential mortgages have a risk weight of 50%. The total of the risk-weighted assets for N on-balance-sheet items equals

$$\sum_{i=1}^{N} w_i L_i$$

where L_i is the principal amount of the ith item and w_i is its risk weight.

[2] The ratio is named after Peter Cooke, from the Bank of England.

Table 15.1 Risk Weights for On-Balance-Sheet Items

Risk Weight (%)	Asset Category
0	Cash, gold bullion, claims on OECD governments such as Treasury bonds or insured residential mortgages
20	Claims on OECD banks and OECD public-sector entities such as securities issued by U.S. government agencies or claims on municipalities
50	Uninsured residential mortgage loans
100	All other claims such as corporate bonds and less developed country debt, claims on non-OECD banks

Example 15.1

The assets of a bank consist of \$100 million of corporate loans, \$10 million of OECD government bonds, and \$50 million of residential mortgages. The total of the risk-weighted assets is

$$1.0 \times 100 + 0.0 \times 10 + 0.5 \times 50 = 125$$

or \$125 million.

Consider next the second category. This includes bankers' acceptances, guarantees, and loan commitments. A credit equivalent amount is calculated by applying a conversion factor to the principal amount of the instrument. Instruments that from a credit perspective are considered to be similar to loans, such as bankers' acceptances, have a conversion factor of 100%. Others, such as note issuance facilities (where a bank agrees that a company can issue short-term paper on pre-agreed terms in the future), have lower conversion factors.

Finally, consider the third category. For an over-the-counter derivative such as an interest rate swap or a forward contract, the credit equivalent amount is calculated as

$$\max(V, 0) + aL \tag{15.1}$$

where V is the current value of the derivative to the bank, a is an *add-on factor*, and L is the principal amount. The first term in equation (15.1) is the current exposure. If the counterparty defaults today and V is positive, the contract is an asset to the bank and the bank is liable to lose V. If the counterparty defaults today and V is negative, the contract is an asset to the counterparty and there will be neither a gain nor a loss to the bank. The bank's exposure is therefore $\max(V, 0)$. (More details on the way in which defaults are handled in the bilaterally cleared over-the-counter derivatives market are in Chapters 17 and 20.) The add-on amount, aL, is an allowance for the possibility of the exposure increasing in the future. Examples of the add-on factor, a, are shown in

Table 15.2 Add-On Factors as a Percent of Principal for Derivatives

Remaining Maturity (yr)	Interest Rate	Exchange Rate and Gold	Equity	Precious Metals Except Gold	Other Commodities
< 1	0.0	1.0	6.0	7.0	10.0
1 to 5	0.5	5.0	8.0	7.0	12.0
> 5	1.5	7.5	10.0	8.0	15.0

Table 15.2. Equation (15.1) is known as the *current exposure method* (CEM). The add-on factors were revised and extended in the years following 1988.

Example 15.2

A bank has entered into a \$100 million interest rate swap with a remaining life of four years. The current value of the swap is \$2.0 million. In this case, the add-on amount is 0.5% of the principal so that the credit equivalent amount is \$2.0 million plus \$0.5 million or \$2.5 million.

The credit equivalent amount arising from either the second or third category of exposures is multiplied by the risk weight for the counterparty in order to calculate the risk-weighted assets. The risk weights are similar to those in Table 15.1 except that the risk weight for a corporation is 0.5 rather than 1.0.

Example 15.3

Consider again the bank in Example 15.2. If the interest rate swap is with a corporation, the risk-weighted assets are 2.5×0.5 or \$1.25 million. If it is with an OECD bank, the risk-weighted assets are 2.5×0.2 or \$0.5 million.

Putting all this together, the total risk-weighted assets for a bank with N on-balance-sheet items and M off-balance-sheet items is

$$\sum_{i=1}^{N} w_i L_i + \sum_{j=1}^{M} w_j^* C_j \tag{15.2}$$

Here, L_i is the principal of the ith on-balance-sheet asset and w_i is the risk weight for the asset; C_j is the credit equivalent amount for the jth derivative or other off-balance-sheet item and w_j^* is the risk weight of the counterparty for this jth item.

15.3.2 Capital Requirement

The Accord required banks to keep capital equal to at least 8% of the risk-weighted assets. The capital had two components:

1. *Tier 1 Capital*. This consists of items such as equity and noncumulative perpetual preferred stock.[3] (Goodwill is subtracted from equity.[4])
2. *Tier 2 Capital*. This is sometimes referred to as Supplementary Capital. It includes instruments such as cumulative perpetual preferred stock,[5] certain types of 99-year debenture issues, and subordinated debt (i.e., debt subordinated to depositors) with an original life of more than five years.

Equity capital is the most important type of capital because it absorbs losses. If equity capital is greater than losses, a bank can continue as a going concern. If equity capital is less than losses, the bank is insolvent. In the latter case, Tier 2 capital becomes relevant. Because it is subordinate to depositors, it provides a cushion for depositors. If a bank is wound up after its Tier I capital has been used up, losses should be borne first by the Tier 2 capital and, only if that is insufficient, by depositors. (See Section 2.2.)

The Accord required at least 50% of the required capital (that is, 4% of the risk-weighted assets) to be in Tier 1. Furthermore, the Accord required 2% of risk-weighted assets to be common equity. (Note that the Basel Committee has updated its definition of instruments that are eligible for Tier 1 capital and its definition of common equity in Basel III.)

The bank supervisors in some countries require banks to hold more capital than the minimum specified by the Basel Committee and some banks themselves have a target for the capital they will hold that is higher than that specified by their bank supervisors.

15.4 The G-30 Policy Recommendations

In 1993, a working group consisting of end-users, dealers, academics, accountants, and lawyers involved in derivatives published a report that contained 20 risk management recommendations for dealers and end-users of derivatives and four recommendations for legislators, regulators, and supervisors. The report was based on a detailed survey of 80 dealers and 72 end-users worldwide. The survey involved both questionnaires and in-depth interviews. The report is not a regulatory document, but it has been influential in the development of risk management practices.

A brief summary of the important recommendations is as follows:

1. A company's policies on risk management should be clearly defined and approved by senior management, ideally at the board of directors level. Managers at all levels should enforce the policies.

---

[3] Noncumulative perpetual preferred stock is preferred stock lasting forever where there is a predetermined dividend rate. Unpaid dividends do not cumulate (that is, the dividends for one year are not carried forward to the next year).

[4] Goodwill is created when one company acquires another. It is the purchase price minus the book value of the assets. It represents the intangible assets acquired.

[5] In cumulative preferred stock, unpaid dividends cumulate. Any backlog of dividends must be paid before dividends are paid on the common stock.

2. Derivatives positions should be marked to market (i.e., revalued using a model that is consistent with market prices) at least once a day.
3. Derivatives dealers should measure market risk using a consistent measure such as value at risk. Limits to the market risks that are taken should be set.
4. Derivatives dealers should carry out stress tests to determine potential losses under extreme market conditions.
5. The risk management function should be set up so that it is independent of the trading operation.
6. Credit exposures arising from derivatives trading should be assessed based on the current replacement value of existing positions and potential future replacement costs.
7. Credit exposures to a counterparty should be aggregated in a way that reflects enforceable netting agreements. (We talk about netting in the next section.)
8. The individuals responsible for setting credit limits should be independent of those involved in trading.
9. Dealers and end-users should assess carefully both the costs and benefits of credit risk mitigation techniques such as collateralization and downgrade triggers. In particular, they should assess their own capacity and that of their counterparties to meet the cash flow requirement of downgrade triggers. (Downgrade triggers are discussed in Chapter 20.)
10. Only individuals with the appropriate skills and experience should be allowed to have responsibility for trading derivatives, supervising the trading, carrying out back office functions in relation to the trading, and so on.
11. There should be adequate systems in place for data capture, processing, settlement, and management reporting.
12. Dealers and end-users should account for the derivatives transactions used to manage risks so as to achieve a consistency of income recognition treatment between those instruments and the risks being managed.

15.5 Netting

Participants in the over-the-counter derivatives market have traditionally signed an International Swaps and Derivatives Association (ISDA) master agreement covering their derivatives trades. The word *netting* refers to a clause in the master agreement, which states that in the event of a default all transactions are considered as a single transaction. Effectively, this means that, if a company defaults on one transaction that is covered by the master agreement, it must default on all transactions covered by the master agreement.

Netting and ISDA master agreements will be discussed in Chapters 17 and 20. At this stage, we note that netting can have the effect of substantially reducing credit risk. Consider a bank that has three swap transactions outstanding with a particular counterparty. The transactions are worth +\$24 million, –\$17 million, and +\$8 million to the bank. Suppose that the counterparty experiences financial difficulties and defaults on its outstanding obligations. To the counterparty the three transactions have values of

−\$24 million, +\$17 million, and –\$8 million, respectively. Without netting, the counterparty would default on the first transaction, keep the second transaction, and default on the third transaction. Assuming no recovery, the loss to the bank would be \$32 (= 24 + 8) million. With netting, the counterparty is required to default on the second transaction as well. The loss to the bank is then \$15 (= 24 − 17 + 8) million.

More generally, suppose that a financial institution has a portfolio of N derivatives outstanding with a particular counterparty and that the current value of the ith derivative is V_i. Without netting, the financial institution's exposure in the event of a default today is

$$\sum_{i=1}^{N} \max(V_i, 0)$$

With netting, it is

$$\max\left(\sum_{i=1}^{N} V_i, 0\right)$$

Without netting, the exposure is the payoff from a portfolio of options. With netting, the exposure is the payoff from an option on a portfolio.

The 1988 Basel Accord did not take netting into account in setting capital requirements. From equation (15.1) the credit equivalent amount for a portfolio of derivatives with a counterparty under the Accord was

$$\sum_{i=1}^{N} [\max(V_i, 0) + a_i L_i]$$

where a_i is the add-on factor for the ith transaction and L_i is the principal for the ith transaction.

By 1995, netting had been successfully tested in the courts in many jurisdictions. As a result, the 1988 Accord was modified to allow banks to reduce their credit equivalent totals when enforceable bilateral netting agreements were in place. The first step was to calculate the net replacement ratio, NRR. This is the ratio of the current exposure with netting to the current exposure without netting:

$$\text{NRR} = \frac{\max(\sum_{i=1}^{N} V_i, 0)}{\sum_{i=1}^{N} \max(V_i, 0)}$$

The credit equivalent amount was modified to

$$\max\left(\sum_{i=1}^{N} V_i, 0\right) + (0.4 + 0.6 \times \text{NRR}) \sum_{i=1}^{N} a_i L_i$$

Table 15.3 Portfolio of Derivatives with a Particular Counterparty

Transaction	Principal, L_i	Current Value, V_i	Add-On Amount, a_iL_i
3-year interest rate swap	1,000	−60	5
6-year foreign exchange forward	1,000	70	75
9-month option on a stock	500	55	30

Example 15.4

Consider the example in Table 15.3, which shows a portfolio of three derivatives that a bank has with a particular counterparty. The third column shows the current mark-to-market values of the transactions and the fourth column shows the add-on amount calculated from Table 15.2. The current exposure with netting is $-60 + 70 + 55 = 65$. The current exposure without netting is $0 + 70 + 55 = 125$. The net replacement ratio is given by

$$\text{NRR} = \frac{65}{125} = 0.52$$

The total of the add-on amounts, $\sum a_iL_i$, is $5 + 75 + 30 = 110$. The credit equivalent amount when netting agreements are in place is $65 + (0.4 + 0.6 \times 0.52) \times 110 = 143.32$. Without netting, the credit equivalent amount is $125 + 110 = 235$. Suppose that the counterparty is an OECD bank so that the risk weight is 0.2. This means that the risk-weighted assets with netting is $0.2 \times 143.32 = 28.66$. Without netting, it is $0.2 \times 235 = 47$.

15.6 The 1996 Amendment

In 1995, the Basel Committee issued a consultative proposal to amend the 1988 Accord. This became known as the "1996 Amendment." It was implemented in 1998 and was then sometimes referred to as "BIS 98." The amendment involves keeping capital for the market risks associated with trading activities.

Marking to market is the practice of revaluing assets and liabilities daily using a model that is calibrated to current market prices. It is also known as *fair value accounting*. Banks are required to use fair value accounting for all assets and liabilities that are held for trading purposes. This includes most derivatives, marketable equity securities, foreign currencies, and commodities. These items constitute what is referred to as the bank's *trading book*.

Under the 1996 Amendment, the credit risk capital charge in the 1988 Accord continued to apply to all on-balance-sheet and off-balance-sheet items in the trading and banking book, except positions in the trading book that consisted of (a) debt and equity traded securities and (b) positions in commodities and foreign exchange. The Amendment introduced a capital charge for the market risk associated with all items in the trading book.[6]

The 1996 Amendment outlined a standardized approach for measuring the capital charge for market risk. The standardized approach assigned capital separately to each of debt securities, equity securities, foreign exchange risk, commodities risk, and options. No account was taken of correlations between different types of instruments. The more sophisticated banks with well-established risk management functions were allowed to use an "internal model-based approach" for setting market risk capital. This involved calculating a value-at-risk measure and converting it into a capital requirement using a formula specified in the 1996 Amendment. Most large banks preferred to use the internal model-based approach because it better reflected the benefits of diversification and led to lower capital requirements.

The value-at-risk measure used in the internal model-based approach was calculated with a 10-day time horizon and a 99% confidence level. It is the loss that has a 1% chance of being exceeded over a 10-day period. The capital requirement is

$$\max(\text{VaR}_{t-1}, m_c \times \text{VaR}_{\text{avg}}) + \text{SRC} \tag{15.3}$$

where m_c is a multiplicative factor, and SRC is a *specific risk charge*. The variable VaR_{t-1} is the previous day's value at risk and VaR_{avg} is the average value at risk over the past 60 days. The minimum value for m_c is 3. Higher values may be chosen by regulators for a particular bank if tests reveal inadequacies in the bank's value-at-risk model, as will be explained shortly.

The first term in equation (15.3) covers risks relating to movements in broad market variables such as interest rates, exchange rates, stock indices, and commodity prices. The second term, SRC, covers risks related to specific companies such as those concerned with movements in a company's stock price or changes in a company's credit spread.

Consider the first term and assume $m_c = 3$. In most circumstances, the most recently calculated VaR, VaR_{t-1}, is less than three times the average VaR over the past 60 days. This gives the capital requirement for movements in broad market variables as simply

$$3 \times \text{VaR}_{\text{avg}}$$

The most popular method for calculating VaR is historical simulation, described in Chapter 13. As explained in Chapter 12, banks almost invariably calculate a one-day 99% VaR in the first instance. When formulating the 1996 amendment, regulators explicitly stated that the 10-day 99% VaR can be calculated as $\sqrt{10}$ times the one-day 99% VaR. This means that, when the capital requirement for a bank is calculated as m_c times the average

[6] Certain nontrading book positions that are used to hedge positions in the trading book can be included in the calculation of the market risk capital charge.

10-day 99% VaR, it is to all intents and purposes $m_c \times \sqrt{10} = 3.16m_c$ times the average one-day 99% VaR. If $m_c = 3$, it is 9.48 times the average one-day VaR.

Consider next SRC. One security that gives rise to an SRC is a corporate bond. There are two components to the risk of this security: interest rate risk and credit risk of the corporation issuing the bond. The interest rate risk is captured by the first term in equation (15.3); the credit risk is captured by the SRC.[7] The 1996 Amendment proposed standardized methods for assessing the SRC, but allowed banks to use internal models once regulatory approval for the models had been obtained.

The internal model for SRC must involve calculating a 10-day 99% value at risk for specific risks. Regulators calculate capital by applying a multiplicative factor (similar to m_c) to the value at risk. This multiplicative factor must be at least 4 and the resultant capital must be at least 50% of the capital given by the standardized approach. A method for calculating SRC from credit spreads will be given in Section 21.5.

The total capital a bank was required to keep after the implementation of the 1996 Amendment was the sum of (a) credit risk capital equal to 8% of the risk-weighted assets (RWA) and (b) market risk capital as explained in this section. For convenience, an RWA for market risk capital was defined as 12.5 multiplied by the amount given in equation (15.3). This means that the total capital required for credit and market risk is given by

$$\text{Total Capital} = 0.08 \times (\text{credit risk RWA} + \text{market risk RWA}) \tag{15.4}$$

A bank had more flexibility in the type of capital it used for market risk. It could use Tier 1 or Tier 2 capital. It could also use what is termed Tier 3 capital. This consists of short-term subordinated debt with an original maturity of at least two years that is unsecured and fully paid up. (Tier 3 capital was eliminated under Basel III.)

15.6.1 Back-Testing

The BIS Amendment requires the one-day 99% VaR that a bank calculates to be back-tested over the previous 250 days. As described in Section 12.10, this involves using the bank's current procedure for estimating VaR for each of the most recent 250 days. If the actual loss that occurred on a day is greater than the VaR level calculated for the day, an "exception" is recorded. Calculations are carried out (a) including changes that were made to the portfolio on the day being considered and (b) assuming that no changes were made to the portfolio on the day being considered. (Regulators like to see both calculations.)

If the number of exceptions during the previous 250 days is less than 5, m_c is normally set equal to 3. If the number of exceptions is 5, 6, 7, 8, and 9, the value of the m_c is set equal

[7] As mentioned earlier, the 1988 credit risk capital charge did not apply to debt securities in the trading book under the 1996 Amendment.

to 3.4, 3.5, 3.65, 3.75, and 3.85, respectively. The bank supervisor has some discretion as to whether the higher multipliers are used. They will normally be applied when the reason for the exceptions is identified as a deficiency in the VaR model being used. If changes in the bank's positions during the day result in exceptions, the higher multiplier should be considered, but does not have to be used. When the only reason that is identified is bad luck, no guidance is provided for the supervisor. In circumstances where the number of exceptions is 10 or more, the Basel Amendment requires the multiplier to be set at 4. Problem 15.18 considers these guidelines in the context of the statistical tests we discussed in Section 12.10.

15.7 Basel II

The 1988 Basel Accord improved the way capital requirements were determined, but it does have significant weaknesses. Under the Accord, all loans by a bank to a corporation have a risk weight of 100% and require the same amount of capital. A loan to a corporation with a AAA credit rating is treated in the same way as one to a corporation with a B credit rating.[8] Also, in Basel I there was no model of default correlation.

In June 1999, the Basel Committee proposed new rules that have become known as Basel II. These were revised in January 2001 and April 2003. A number of Quantitative Impact Studies (QISs) were carried out prior to the implementation of the new rules to test them by calculating the amount of capital that would be required if the rules had been in place.[9] A final set of rules agreed to by all members of the Basel Committee was published in June 2004. This was updated in November 2005. Implementation of the rules began in 2007 after a further QIS.

The Basel II capital requirements applied to "internationally active" banks. In the United States, there are many small regional banks and the U.S. regulatory authorities decided that Basel II would not apply to them. (These banks are regulated under what is termed Basel IA, which is similar to Basel I.) In Europe, all banks, large or small, were regulated under Basel II. Furthermore, the European Union required the Basel II rules to be applied to securities companies as well as banks.

The Basel II is based on three "pillars":

1. Minimum Capital Requirements
2. Supervisory Review
3. Market Discipline

In Pillar 1, the minimum capital requirement for credit risk in the banking book is calculated in a new way that reflects the credit risk of counterparties. The capital

[8] Credit ratings are discussed in Section 1.7.

[9] One point to note about the QISs is that they do not take account of changes banks may choose to make to their portfolios to minimize their capital requirements once the new rules have been implemented.

requirement for market risk remains unchanged from the 1996 Amendment and there is a new capital charge for operational risk. The general requirement in Basel I that banks hold a total capital equal to 8% of risk-weighted assets (RWA) remains unchanged. When the capital requirement for a particular risk is calculated directly rather than in a way involving RWAs, it is multiplied by 12.5 to convert it into an RWA-equivalent. As a result it is always the case that

$$\begin{aligned}\text{Total Capital} = 0.08 \times (&\text{credit risk RWA} + \text{market risk RWA} \\ &+ \text{operational risk RWA})\end{aligned} \tag{15.5}$$

Pillar 2 is concerned with the supervisory review process. It covers both quantitative and qualitative aspects of the ways risk is managed within a bank. Supervisors are required to ensure that a bank has a process in place for ensuring that capital levels are maintained. Banks are expected to keep more than the minimum regulatory capital to allow for fluctuations in capital requirements and possible difficulties in raising capital at short notice. Regulators in different countries are allowed some discretion in how rules are applied (so that they can take account of local conditions), but overall consistency in the application of the rules is required. Pillar 2 places more emphasis on early intervention when problems arise. Supervisors are required to do far more than just ensuring that the minimum capital required under Basel II is held. Part of their role is to encourage banks to develop and use better risk management techniques and to evaluate those techniques. They should evaluate risks that are not covered by Pillar 1 (e.g., concentration risks) and enter into an active dialogue with banks when deficiencies are identified.

The third pillar, market discipline, requires banks to disclose more information about the way they allocate capital and the risks they take. The idea here is that banks will be subjected to added pressure to make sound risk management decisions if shareholders and potential shareholders have more information about those decisions.

15.8 Credit Risk Capital Under Basel II

For credit risk, Basel II specified three approaches:

1. The Standardized Approach
2. The Foundation Internal Ratings Based (IRB) Approach
3. The Advanced IRB Approach

However, the United States (which, as mentioned earlier, chose to apply Basel II only to large banks) decided that only the IRB approach can be used.

15.8.1 The Standardized Approach

The standardized approach is used by banks that are not sufficiently sophisticated (in the eyes of the regulators) to use the internal ratings approaches. The standardized approach

Table 15.4 Risk Weights as a Percent of Principal for Exposures to Countries, Banks, and Corporations Under Basel II's Standardized Approach

	AAA to AA–	A+ to A–	BBB+ to BBB–	BB+ to BB–	B+ to B–	Below B–	Unrated
Country*	0	20	50	100	100	150	100
Banks**	20	50	50	100	100	150	50
Corporations	20	50	100	100	150	150	100

*Includes exposures to the country's central bank.
**National supervisors have options as outlined in the text.

is similar to Basel I except for the calculation of risk weights.[10] Some of the new rules here are summarized in Table 15.4. Comparing Table 15.4 with Table 15.1, we see that the OECD status of a bank or a country is no longer considered important under Basel II. The risk weight for a country (sovereign) exposure ranges from 0% to 150% and the risk weight for an exposure to another bank or a corporation ranges from 20% to 150%. In Table 15.1, OECD banks were implicitly assumed to be lesser credit risks than corporations. An OECD bank attracted a risk weight of 20% while a corporation attracted a risk weight of 100%. Table 15.4 treats banks and corporations much more equitably. An interesting observation from Table 15.4 for a country, corporation, or bank that wants to borrow money is that it may be better to have no credit rating at all than a very poor credit rating. (Usually a company gets a credit rating when it issues a publicly traded bond.) Supervisors are allowed to apply lower risk weights (20% rather than 50%, 50% rather than 100%, and 100% rather than 150%) when exposures are to the country in which the bank is incorporated or to that country's central bank.

The risk weight for mortgages in the Basel II standardized approach was 35% and that for other retail loans was 75%. For claims on banks, the rules are somewhat complicated. Instead of using the risk weights in Table 15.4, national supervisors can choose to base capital requirements on the rating of the country in which the bank is incorporated. The risk weight assigned to the bank will be 20% if the country of incorporation has a rating between AAA and AA–, 50% if it is between A+ and A–, 100% if it is between BBB+ and B–, 150% if it is below B–, and 100% if it is unrated. Another complication is that, if national supervisors elect to use the rules in Table 15.4, they can choose to treat claims with a maturity less than three months more favorably so that the risk weights are 20% if the rating is between AAA+ and BBB–, 50% if it is between BB+ and B–, 150% if it is below B–, and 20% if it is unrated.

In December 2017, the Basel Committee created a revised more granular standardized approach for determining risk weights for credit exposures. It also limited the use of the Advanced IRB approaches in some situations. As explained in Section 16.4, the revised standardized approach will provide a floor for the determination of total capital requirements.

[10] Ratios calculated using the new weights are sometimes referred to as McDonough ratios after William McDonough, the head of the Basel Committee.

Example 15.5

Suppose that the assets of a bank consist of \$100 million of loans to corporations rated A, \$10 million of government bonds rated AAA, and \$50 million of residential mortgages. Under the Basel II standardized approach, the total of the risk-weighted assets is

$$0.5 \times 100 + 0.0 \times 10 + 0.35 \times 50 = 67.5$$

or \$67.5 million. This compares with \$125 million under Basel I. (See Example 15.1.)

15.8.2 Adjustments for Collateral

There are two ways banks can adjust risk weights for collateral. The first is termed the *simple approach* and is similar to an approach used in Basel I. The second is termed the *comprehensive approach*. Banks have a choice as to which approach is used in the banking book, but must use the comprehensive approach to calculate capital for counterparty credit risk in the trading book.

Under the simple approach, the risk weight of the counterparty is replaced by the risk weight of the collateral for the part of the exposure covered by the collateral. (The exposure is calculated after netting.) For any exposure not covered by the collateral, the risk weight of the counterparty is used. The minimum level for the risk weight applied to the collateral is 20%.[11] A requirement is that the collateral must be revalued at least every six months and must be pledged for at least the life of the exposure.

Under the comprehensive approach, banks adjust the size of their exposure upward to allow for possible increases in the exposure and adjust the value of the collateral downward to allow for possible decreases in the value of the collateral.[12] (The adjustments depend on the volatility of the exposure and the collateral.) A new exposure equal to the excess of the adjusted exposure over the adjusted value of the collateral is calculated and the counterparty's risk weight is applied to this exposure. The adjustments applied to the exposure and the collateral can be calculated using rules specified in Basel II or, with regulatory approval, using a bank's internal models. Where netting arrangements apply, exposures and collateral are separately netted and the adjustments made are weighted averages.

Example 15.6

Suppose that an \$80 million exposure to a particular counterparty is secured by collateral worth \$70 million. The collateral consists of bonds issued by an A-rated company. The

[11] An exception is when the collateral consists of cash or government securities with the currency of the collateral being the same as the currency of the exposure.

[12] An adjustment to the exposure is not likely to be necessary on a loan, but is likely to be necessary on an over-the-counter derivative. The adjustment is in addition to the add-on factor.

counterparty has a rating of B+. The risk weight for the counterparty is 150% and the risk weight for the collateral is 50%. The risk-weighted assets applicable to the exposure using the simple approach is

$$0.5 \times 70 + 1.50 \times 10 = 50$$

or \$50 million.

Consider next the comprehensive approach. Assume that the adjustment to exposure to allow for possible future increases in the exposure is +10% and the adjustment to the collateral to allow for possible future decreases in its value is –15%. The new exposure is

$$1.1 \times 80 - 0.85 \times 70 = 28.5$$

or \$28.5 million, and a risk weight of 150% is applied to this exposure to give risk-adjusted assets equal to \$42.75 million.

15.8.3 The IRB Approach

The model underlying the IRB approach is shown in Figure 15.1. Regulators base the capital requirement on the value at risk calculated using a one-year time horizon and a 99.9% confidence level. They recognize that expected losses are usually covered by the way a financial institution prices its products. (For example, the interest charged by a bank on a loan is designed to recover expected loan losses.) The capital required is therefore the value at risk minus the expected loss.

The value at risk is calculated using the one-factor Gaussian copula model of time to default that we discussed in Section 11.6. Assume that a bank has a very large number

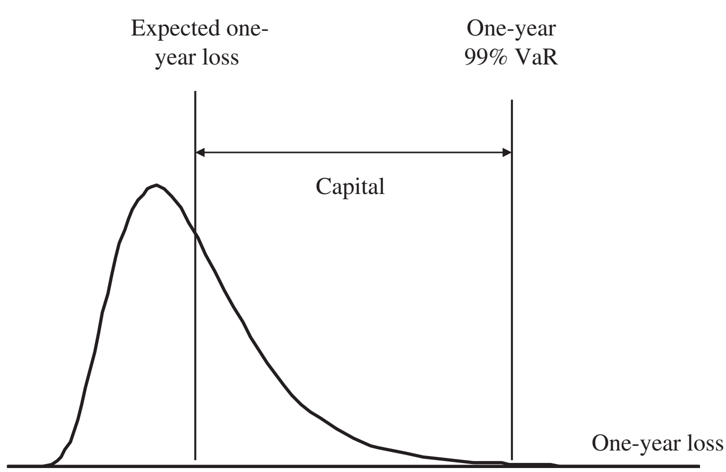


Figure 15.1 The Loss Probability Density Function and the Capital Required by a Financial Institution

of obligors and the ith obligor has a one-year probability of default equal to PD_i. The copula correlation between each pair of obligors is ρ.[13] As in Section 11.6, we define

$$\text{WCDR}_i = N\left[\frac{N^{-1}(\text{PD}_i) + \sqrt{\rho}N^{-1}(0.999)}{\sqrt{1-\rho}}\right] \tag{15.6}$$

where WCDR_i denotes the "worst-case default rate" defined so that the bank is 99.9% certain it will not be exceeded next year for the ith counterparty. Gordy's (2003) research shows that for a large portfolio of instruments (loans, loan commitments, derivatives, and so on) that have the same ρ, in a one-factor model the one-year 99.9% VaR is approximately[14]

$$\sum_i \text{EAD}_i \times \text{LGD}_i \times \text{WCDR}_i$$

where EAD_i is the exposure at default of the ith counterparty and LGD_i is the loss given default for the ith counterparty.

The variable EAD_i is the dollar amount that is expected to be owed by the ith counterparty at the time of default during the next year. The variable LGD_i is the proportion of EAD_i that is expected to be lost in the event of default. For example, if a bank expects to recover 30% of the amount owed in the event of default, $\text{LGD}_i = 0.7$.

The expected loss from defaults is

$$\sum_i \text{EAD}_i \times \text{LGD}_i \times \text{PD}_i$$

The capital required in Figure 15.1 is the excess of the 99.9% worst-case loss over the expected loss. It is therefore

$$\sum_i \text{EAD}_i \times \text{LGD}_i \times (\text{WCDR}_i - \text{PD}_i) \tag{15.7}$$

We now drop the subscripts and define for a counterparty:

PD: The probability that the counterparty will default within one year (expressed as a decimal)

EAD: The exposure at default (in dollars)

LGD: The loss given default or the proportion of the exposure that is lost if there is a default (expressed as a decimal)

Table 15.5 shows how WCDR depends on PD and ρ in the Gaussian copula model. When the correlation ρ is zero, WCDR = PD because in that case there is no default

[13] Note that the Basel Committee publications use R, not ρ, to denote the copula correlation.

[14] See M. B. Gordy, "A Risk-Factor Model Foundation for Ratings-Based Bank Capital Ratios," *Journal of Financial Intermediation* 12 (2003): 199–232.

Table 15.5 Dependence of One-Year 99.9% WCDR on PD and ρ

	PD = 0.1%	PD = 0.5%	PD = 1%	PD = 1.5%	PD = 2.0%
$\rho = 0.0$	0.1%	0.5%	1.0%	1.5%	2.0%
$\rho = 0.2$	2.8%	9.1%	14.6%	18.9%	22.6%
$\rho = 0.4$	7.1%	21.1%	31.6%	39.0%	44.9%
$\rho = 0.6$	13.5%	38.7%	54.2%	63.8%	70.5%
$\rho = 0.8$	23.3%	66.3%	83.6%	90.8%	94.4%

correlation and the percentage of loans defaulting can be expected to be the same in all years. As ρ increases, WCDR increases.

15.8.4 Exposure at Default

In the Foundation IRB approach, the calculation of EAD for derivatives is usually based on the CEM approach of Basel I.[15] In the Advanced IRB approach, banks can use their internal models to calculate EAD.

The first stage when internal models are used is to calculate expected exposure (EE) to each counterparty at a number of future times. (This generally involves Monte Carlo simulation and methods discussed in Chapter 20.) The Effective Expected Exposure (Effective EE) at a future time, t, is the maximum EE between time zero an time t. EAD is set equal to 1.4 times the Effective Expected Positive Exposure, which is the average of the Effective EEs over the first year.[16]

15.8.5 Corporate, Sovereign, and Bank Exposures

In the case of corporate, sovereign, and bank exposures, Basel II assumes a relationship between the correlation parameter, ρ, and the probability of default, PD, based on empirical research.[17] The formula is

$$\rho = 0.12\frac{1-\exp(-50\times \text{PD})}{1-\exp(-50)} + 0.24\left[1-\frac{1-\exp(-50\times \text{PD})}{1-\exp(-50)}\right]$$

Because $\exp(-50)$ is a very small number, this formula is to all intents and purposes

$$\rho = 0.12\left(1+e^{-50\times \text{PD}}\right) \tag{15.8}$$

As PD increases, ρ decreases. The reason usually given for this inverse relationship is as follows. As a company becomes less creditworthy, its PD increases and its probability of default becomes more idiosyncratic and less affected by overall market conditions.

[15] As will be discussed in Section 16.4, this was revised under a regulatory change known as SA-CCR.
[16] See Basel Committee on Banking Supervision, "The Application of Basel II to Trading Activities and the Treatment of Double Default Effects," July 2005.
[17] See J. Lopez, "The Empirical Relationship between Average Asset Correlation, Firm Probability of Default, and Asset Size," *Journal of Financial Intermediation* 13, no. 2 (2004): 265–283.

Table 15.6 Relationship between One-Year 99.9% WCDR and PD for Corporate, Sovereign, and Bank Exposures

PD	0.1%	0.5%	1%	1.5%	2.0%
WCDR	3.4%	9.8%	14.0%	16.9%	19.0%

Combining equation (15.8) with equation (15.6), we obtain the relationship between WCDR and PD in Table 15.6. WCDR is, as one would expect, an increasing function of PD. However, it does not increase as fast as it would if ρ were assumed to be independent of PD.

The formula for the capital required for the counterparty is

$$\text{EAD} \times \text{LGD} \times (\text{WCDR} - \text{PD}) \times \text{MA} \tag{15.9}$$

The meaning of the first three terms in this expression should be clear from our earlier discussion leading to equation (15.7). The variable MA is the maturity adjustment and is defined as

$$\text{MA} = \frac{1 + (\text{M} - 2.5) \times b}{1 - 1.5 \times b} \tag{15.10}$$

where

$$b = [0.11852 - 0.05478 \times \ln(\text{PD})]^2$$

and M is the maturity of the exposure. The maturity adjustment is designed to allow for the fact that, if an instrument lasts longer than one year, there is a one-year credit exposure arising from a possible decline in the creditworthiness of the counterparty as well as from a possible default by the counterparty. (Note that, when $\text{M} = 1$, MA is 1.0 and has no effect.) As mentioned in Section 15.7 (see Equation (15.5)), the risk-weighted assets (RWA) are calculated as 12.5 times the capital required

$$\text{RWA} = 12.5 \times \text{EAD} \times \text{LGD} \times (\text{WCDR} - \text{PD}) \times \text{MA}$$

so that the capital is 8% of RWA, 4% of which must be Tier 1.

Under the Foundation IRB approach, banks supply PD while LGD, EAD, and M are supervisory values set by the Basel Committee. PD is subject to a floor of 0.03% for bank and corporate exposures. LGD is set at 45% for senior claims and 75% for subordinated claims. When there is eligible collateral, in order to correspond to the comprehensive approach that we described earlier, LGD is reduced by the ratio of the adjusted value of the collateral to the adjusted value of the exposure, both calculated using the comprehensive approach. For derivatives, the EAD is calculated using the CEM "current exposure plus add-on" approach of Basel I and includes the impact of netting. M is set at 2.5 in most circumstances.

Under the Advanced IRB approach, banks supply their own estimates of the PD, LGD, EAD, and M for corporate, sovereign, and bank exposures. The PD can be reduced by credit mitigants such as credit triggers. (As in the case of the Foundation IRB approach, it is subject to a floor of 0.03% for bank and corporate exposures.) The two main factors influencing the LGD are the seniority of the debt and the collateral. In calculating EAD, banks can with regulatory approval use their own models. As mentioned, in the case of derivatives, the model is likely to involve Monte Carlo simulation to determine how expected exposure (after netting and collateral) will vary over the next year.

The capital given by equation (15.9) is intended to be sufficient to cover unexpected losses over a one-year period that we are 99.9% certain will not be exceeded. (As discussed earlier, the expected losses should be covered by a bank in the way it prices its products.) The WCDR is the default rate that (theoretically) happens once every thousand years. The Basel Committee reserved the right to apply a scaling factor (less than or greater than 1.0) to the result of the calculations in equation (15.9) if it finds that the aggregate capital requirements are too high or low. A typical scaling factor is 1.06.

Example 15.7

Suppose that the assets of a bank consist of \$100 million of loans to A-rated corporations. The PD for the corporations is estimated as 0.1% and the LGD is 60%. The average maturity is 2.5 years for the corporate loans. This means that

$$b = [0.11852 - 0.05478 \times \ln(0.001)]^2 = 0.247$$

so that

$$\text{MA} = \frac{1}{1 - 1.5 \times 0.247} = 1.59$$

From Table 15.6, the WCDR is 3.4%. Under the Basel II IRB approach, the risk-weighted assets for the corporate loans are

$$12.5 \times 100 \times 0.6 \times (0.034 - 0.001) \times 1.59 = 39.3$$

or \$39.3 million. This compares with \$100 million under Basel I and \$50 million under the standardized approach of Basel II. (See Examples 15.1 and 15.5 where a \$100 million corporate loan is part of the portfolio.)

15.8.6 Retail Exposures

The model underlying the calculation of capital for retail exposures is similar to that underlying the calculation of corporate, sovereign, and banking exposures. However, the Foundation IRB and Advanced IRB approaches are merged and all banks using the

Table 15.7 Relationship between One-Year 99.9% WCDR and PD for Retail Exposures

PD	0.1%	0.5%	1.0%	1.5%	2.0%
WCDR	2.1%	6.3%	9.1%	11.0%	12.3%

IRB approach provide their own estimates of PD, EAD, and LGD. There is no maturity adjustment, MA. The capital requirement is therefore

$$\text{EAD} \times \text{LGD} \times (\text{WCDR} - \text{PD})$$

and the risk-weighted assets are

$$12.5 \times \text{EAD} \times \text{LGD} \times (\text{WCDR} - \text{PD})$$

WCDR is calculated as in equation (15.6). For residential mortgages, ρ is set equal to 0.15 in this equation.[18] For qualifying revolving exposures, ρ is set equal to 0.04. For all other retail exposures, a relationship between ρ and PD is specified for the calculation of WCDR. This is

$$\rho = 0.03\frac{1-\exp(-35 \times \text{PD})}{1-\exp(-35)} + 0.16\left[1 - \frac{1-\exp(-35 \times \text{PD})}{1-\exp(-35)}\right]$$

Because $\exp(-35)$ is a very small number, this formula is to all intents and purposes

$$\rho = 0.03 + 0.13e^{-35\times\text{PD}} \tag{15.11}$$

Comparing equation (15.11) with equation (15.8), we see that correlations are assumed to be much lower for retail exposures than for corporate exposures. Table 15.7 is the table corresponding to Table 15.6 for retail exposures.

Example 15.8

Suppose that the assets of a bank consist of \$50 million of residential mortgages where the PD is 0.005 and the LGD is 20%. In this case, $\rho = 0.15$ and

$$\text{WCDR} = N\left[\frac{N^{-1}(0.005) + \sqrt{0.15}N^{-1}(0.999)}{\sqrt{1-0.15}}\right] = 0.067$$

The risk-weighted assets are

$$12.5 \times 50 \times 0.2 \times (0.067 - 0.005) = 7.8$$

[18]In the light of experience during the crisis, this is probably too low.

or \$7.8 million. This compares with \$25 million under Basel I and \$17.5 million under the Standardized Approach of Basel II. (See Examples 15.1 and 15.5 where \$50 million of residential mortgages is part of the portfolio.)

15.8.7 Guarantees and Credit Derivatives

The approach traditionally taken by the Basel Committee for handling guarantees and credit derivatives such as credit default swaps is the credit substitution approach. Suppose that a AA-rated company guarantees a loan to a BBB-rated company. For the purposes of calculating capital, the credit rating of the guarantor is substituted for the credit rating of the borrower so that capital is calculated as though the loan had been made to the AA-rated company. This overstates the credit risk because, for the lender to lose money, both the guarantor and the borrower must default (with the guarantor defaulting before the borrower).[19] The Basel Committee has addressed this issue. In July 2005, it published a document concerned with the treatment of double defaults under Basel II.[20] As an alternative to using the credit substitution approach, the capital requirement can be calculated as the capital that would be required without the guarantee multiplied by $0.15 + 160 \times PD_g$ where PD_g is the one-year probability of default of the guarantor.

15.9 Operational Risk Capital Under Basel II

In addition to changing the way banks calculate credit risk capital, Basel II requires banks to keep capital for operational risk. This is the risk of losses from situations where the bank's procedures fail to work as they are supposed to or where there is an adverse external event such as a fire in a key facility.

There are three approaches to calculating capital for operational risk in Basel II:

1. The Basic Indicator Approach
2. The Standardized Approach
3. The Advanced Measurement Approach

Which of these is used depends on the sophistication of the bank. The simplest approach is the Basic Indicator Approach. This sets the operational risk capital equal to the bank's average annual gross income over the last three years multiplied by 0.15.[21] The Standardized Approach is similar to the basic indicator approach except that a different

[19] Credit default swaps, which we discuss in Chapter 19, provide a type of insurance against default and are handled similarly to guarantees for regulatory purposes.

[20] See Bank for International Settlements, "The Application of Basel II to Trading Activities and the Treatment of Double Defaults," July 2005, available on www.bis.org.

[21] Gross income defined as net interest income plus non-interest income. Net interest income is the excess of income earned on loans over interest paid on deposits and other instruments that are used to fund the loans. Years where gross income is negative are not included in the calculations.

factor is applied to the gross income from different business lines. In the Advanced Measurement Approach, the bank uses its own internal models to calculate the operational risk loss that it is 99.9% certain will not be exceeded in one year. Similarly to the way credit risk capital is calculated in the IRB approach, operational risk capital is set equal to this loss minus the expected loss. One advantage of the advanced measurement approach is that it allows banks to recognize the risk-mitigating impact of insurance contracts subject to certain conditions. The calculation of operational risk is discussed further in Chapter 23.

As will be explained in Chapter 23, regulators do not feel the Advanced Measurement Approach has been successful and in 2016 proposed replacing it with a new, simpler approach.

15.10 Pillar 2: Supervisory Review

Pillar 2 of Basel II is concerned with the supervisory review process. Four key principles of supervisory review are specified:

1. Banks should have a process for assessing their overall capital adequacy in relation to their risk profile and a strategy for maintaining their capital levels.
2. Supervisors should review and evaluate banks' internal capital adequacy assessments and strategies, as well as their ability to monitor and ensure compliance with regulatory capital ratios. Supervisors should take appropriate supervisory action if they are not satisfied with the result of this process.
3. Supervisors should expect banks to operate above the minimum regulatory capital and should have the ability to require banks to hold capital in excess of this minimum.
4. Supervisors should seek to intervene at an early stage to prevent capital from falling below the minimum levels required to support the risk characteristics of a particular bank and should require rapid remedial action if capital is not maintained or restored.

The Basel Committee suggests that regulators pay particular attention to interest rate risk in the banking book, credit risk, and operational risk. Key issues in credit risk are stress tests used, default definitions used, credit risk concentration, and the risks associated with the use of collateral, guarantees, and credit derivatives.

The Basel Committee also stresses that there should be transparency and accountability in the procedures used by bank supervisors. This is particularly important when a supervisor exercises discretion in the procedures used or sets capital requirements above the minimum specified in Basel II.

15.11 Pillar 3: Market Discipline

Pillar 3 of Basel II is concerned with increasing the disclosure by a bank of its risk assessment procedures and capital adequacy. The extent to which regulators can force banks to increase their disclosure varies from jurisdiction to jurisdiction. However, banks are unlikely to ignore directives on this from their supervisors, given the potential of supervisors to make their life difficult. Also, in some instances, banks have to increase their disclosure in order to be allowed to use particular methodologies for calculating capital.

Regulatory disclosures are likely to be different in form from accounting disclosures and need not be made in annual reports. It is largely left to the bank to choose disclosures that are material and relevant. Among the items that banks should disclose are:

1. The entities in the banking group to which Basel II is applied and adjustments made for entities to which it is not applied
2. The terms and conditions of the main features of all capital instruments
3. A list of the instruments constituting Tier 1 capital and the amount of capital provided by each item
4. The total amount of Tier 2 capital
5. Capital requirements for credit, market, and operational risk
6. Other general information on the risks to which a bank is exposed and the assessment methods used by the bank for different categories of risk
7. The structure of the risk management function and how it operates

15.12 Solvency II

As discussed in Section 3.11, there are no international standards for the regulation of insurance companies. In the United States, insurance companies are regulated at the state level with some input from the Federal Insurance Office and the National Association of Insurance Commissioners. In Europe, the regulation of insurance companies is handled by the European Union. The long-standing regulatory framework in the European Union, known as Solvency I, was replaced by Solvency II in 2016. Whereas Solvency I calculates capital only for underwriting risks, Solvency II considers investment risks and operational risks as well.

Solvency II has many similarities to Basel II. There are three pillars. Pillar 1 is concerned with the calculation of capital requirements and the types of capital that are eligible. Pillar 2 is concerned with the supervisory review process. Pillar 3 is concerned with the disclosure of risk management information to the market. The three pillars are therefore analogous to the three pillars of Basel II.

Pillar 1 of Solvency II specifies a *minimum capital requirement* (MCR) and a *solvency capital requirement* (SCR). If its capital falls below the SCR level, an insurance company should, at minimum, deliver to the supervisor a plan to restore capital to above the SCR level. The supervisor might require the insurance company to take particular measures to correct the situation. The MCR is regarded as an absolute minimum level of capital. If capital drops below the MCR level, supervisors may assume control of the insurance company and prevent it from taking new business. It might force the insurance company into liquidation, transferring its policies to another company. The MCR is typically between 25% and 45% of the SCR.

There are two ways to calculate the SCR: the standardized approach and the internal models approach. The internal models approach involves a VaR calculation with a one-year time horizon and a 99.5% confidence limit. (The confidence level is therefore less than the 99.9% confidence level used in Pillar 1 of Basel II.) Longer time horizons with lower confidence levels are also allowed when the protection provided is considered equivalent. The SCR involves a capital charge for investment risk, underwriting risk, and operational risk. Investment risk is subdivided into market risk and credit risk. Underwriting risk is subdivided into risk arising from life insurance, non-life insurance (i.e., property and casualty), and health insurance.

Capital should be adequate to deal with large adverse events. Examples of the events considered in quantitative impact studies are:

1. A 32% decrease in global stock markets
2. A 20% decrease in real estate prices
3. A 20% change in foreign exchange rates
4. Specified catastrophic risk scenarios affecting property and casualty payouts
5. Health care costs increasing by a factor times the historical standard deviation of costs
6. A 10% increase in mortality rates
7. A 25% decrease in mortality rates
8. A 10% increase in expenses

The internal models are required to satisfy three tests. The first is a statistical quality test. This is a test of the soundness of the data and methodology used in calculating VaR. The second is a calibration test. This is a test of whether risks have been assessed in accordance with a common SCR target criterion. The third is a use test. This is a test of whether the model is genuinely relevant to and used by risk managers.

There are three types of capital in Solvency II. Tier 1 capital consists of equity capital, retained earnings, and other equivalent funding sources. Tier 2 capital consists of liabilities that are subordinated to policyholders and satisfy certain criteria concerning their availability in wind-down scenarios. Tier 3 capital consists of liabilities that are subordinated to policyholders and do not satisfy these criteria. Similarly to Basel II, the amount of capital that must be Tier 1, Tier 1 plus Tier 2, and Tier 1 plus Tier 2 plus Tier 3 is specified.

Summary

This chapter has provided an overview of capital requirements for banks throughout the world. The way in which regulators calculate the minimum capital a bank is required to hold has changed dramatically since the 1980s. Prior to 1988, regulators determined capital requirements by specifying minimum ratios for capital to assets or maximum ratios for assets to capital. In the late 1980s, both bank supervisors and the banks themselves agreed that changes were necessary. Derivatives trading was increasing fast. Also, banks were competing globally and it was considered important to create a level playing field by making regulations uniform throughout the world.

The 1988 Basel Accord assigned capital for credit risk for both on- and off-balance-sheet exposures. This involved calculating a risk-weighted asset for each item. The risk-weighted assets for an on-balance-sheet loan were calculated by multiplying the principal by a risk weight for the counterparty. In the case of derivatives such as swaps, banks were first required to calculate credit equivalent amounts. The risk-weighted assets were obtained by multiplying the credit equivalent amount by a risk weight for the counterparty. Banks were required to keep capital equal to 8% of the total risk-weighted assets. In 1995, the capital requirements for credit risk were modified to incorporate netting. As a result of an amendment introduced in 1996, the Accord was changed to include a capital charge for market risk. Sophisticated banks could base the capital charge on a value-at-risk calculation.

Basel II was proposed in 1999 and implemented by many banks in about 2007. It led to no immediate change to the capital requirement for market risk. Credit risk capital was calculated in a more sophisticated way than previously to reflect either (a) the credit ratings of obligors or (b) estimates made by the bank in conjunction with a default correlation parameter specified by regulators. In addition, there was a capital charge for operational risk.

Solvency II is a regulatory framework for insurance companies that was implemented by the European Union in 2016. It prescribes minimum capital levels for investment risk, underwriting risk, and operational risk. The structure of Solvency II is broadly similar to Basel II.

Further Reading

Bank for International Settlements. "Basel II: International Convergence of Capital Measurement and Capital Standards." June 2006, www.bis.org.

Gordy, M. B. "A Risk-Factor Model Foundation for Ratings-Based Bank Capital Ratios." *Journal of Financial Intermediation* 12 (2003): 199–232.

Lopez, J. A. "The Empirical Relationship between Average Asset Correlation, Firm Probability of Default, and Asset Size." *Journal of Financial Intermediation* 13, no. 2 (2004): 265–283.

Vasicek, O. "Probability of Loss on a Loan Portfolio." Working Paper, KMV, 1987. (Published in *Risk* in December 2002 under the title "Loan Portfolio Value.")

Practice Questions and Problems (Answers at End of Book)

15.1 "When a steel company goes bankrupt, other companies in the same industry benefit because they have one less competitor. But when a bank goes bankrupt other banks do not necessarily benefit." Explain this statement.

15.2 "The existence of deposit insurance makes it particularly important for there to be regulations on the amount of capital banks hold." Explain this statement.

15.3 An interest rate swap involves the exchange of a fixed rate of interest for a floating rate of interest with both being applied to the same principal. The principals are not exchanged. What is the nature of the credit risk for a bank when it enters into a five-year interest rate swap with a notional principal of \$100 million? Assume the swap is worth zero initially.

15.4 In a currency swap, interest on a principal in one currency is exchanged for interest on a principal in another currency. The principals in the two currencies are exchanged at the end of the life of the swap. Why is the credit risk on a currency swap greater than that on an interest rate swap?

15.5 A four-year interest rate swap currently has a negative value to a financial institution. Is the financial institution exposed to credit risk on the transaction? Explain your answer. How would the capital requirement be calculated under Basel I?

15.6 Estimate the capital required under Basel I for a bank that has the following transactions with a corporation. Assume no netting.

(a) A nine-year interest rate swap with a notional principal of \$250 million and a current market value of −\$2 million.

(b) A four-year interest rate swap with a notional principal of \$100 million and a current value of \$3.5 million.

(c) A six-month derivative on a commodity with a principal of \$50 million that is currently worth \$1 million.

15.7 What is the capital required in Problem 15.6 under Basel I assuming that the 1995 netting amendment applies?

15.8 All the derivatives transactions a bank has with a corporate client have a positive value to the bank. What is the value to the bank of netting provisions in its master agreement with the client?

15.9 Explain why the final stage in the Basel II calculations for credit risk (IRB), market risk, and operational risk is to multiply by 12.5.

15.10 What is the difference between the trading book and the banking book for a bank? A bank currently has a loan of \$10 million to a corporate client. At the end of the life of the loan, the client would like to sell debt securities to the bank instead of borrowing. How does this potentially affect the nature of the bank's regulatory capital calculations?

15.11 Under Basel I, banks do not like lending to highly creditworthy companies and prefer to help them issue debt securities. Why is this? Do you think this changed as a result of Basel II?

15.12 Banks sometimes use what is referred to as regulatory arbitrage to reduce their capital. What do you think this means?

15.13 Equation (15.9) gives the formula for the capital required under Basel II. It involves four terms being multiplied together. Explain each of these terms.

15.14 Explain the difference between the simple and the comprehensive approach for adjusting capital requirements for collateral.

15.15 Explain the difference between the standardized approach, the IRB approach, and the Advanced IRB approach for calculating credit risk capital under Basel II.

15.16 Explain the difference between the basic indicator approach, the standardized approach, and the advanced measurement approach for calculating operational risk capital under Basel II.

15.17 Suppose that the assets of a bank consist of \$200 million of retail loans (not mortgages). The PD is 1% and the LGD is 70%. What are the risk-weighted assets under the Basel II IRB approach? What are the Tier 1 and Tier 2 capital requirements?

15.18 Section 12.10 discusses how statistics can be used to accept or reject a VaR model. Section 15.6 discusses guidelines for bank supervisors in setting the VaR multiplier m_c. It explains that, if the number of exceptions in 250 trials is five or more, then m_c is increased. What is the chance of five or more exceptions if the VaR model is working well?

Further Questions

15.19 Why is there an add-on amount in Basel I for derivatives transactions? "Basel I could be improved if the add-on amount for a derivatives transaction depended on the value of the transaction." How would you argue this viewpoint?

15.20 Estimate the capital required under Basel I for a bank that has the following transactions with another bank. Assume no netting.

(a) A two-year forward contract on a foreign currency, currently worth \$2 million, to buy foreign currency worth \$50 million

(b) A long position in a six-month option on the S&P 500. The principal is \$20 million and the current value is \$4 million.

(c) A two-year swap involving oil. The principal is \$30 million and the current value of the swap is –\$5 million.

What difference does it make if the netting amendment applies?

15.21 A bank has the following transaction with a AA-rated corporation

(a) A two-year interest rate swap with a principal of \$100 million that is worth \$3 million

(b) A nine-month foreign exchange forward contract with a principal of \$150 million that is worth –\$5 million

(c) A long position in a six-month option on gold with a principal of \$50 million that is worth \$7 million

What is the capital requirement under Basel I if there is no netting? What difference does it make if the netting amendment applies? What is the capital required under Basel II when the standardized approach is used?

15.22 Suppose that the assets of a bank consist of \$500 million of loans to BBB-rated corporations. The PD for the corporations is estimated as 0.3%. The average maturity is three years and the LGD is 60%. What is the total risk-weighted assets for credit risk under the Basel II IRB approach? How much Tier 1 and Tier 2 capital is required? How does this compare with the capital required under the Basel II standardized approach and under Basel I?

Chapter 16

Basel II.5, Basel III, and Other Post-Crisis Changes

It was perhaps unfortunate for Basel II that its implementation date coincided, at least approximately, with the start of the worst crisis that financial markets had experienced since the 1930s. Some commentators have blamed Basel II for the crisis. They point out that it was a move toward self-regulation where banks, when calculating regulatory capital, had the freedom to use their own estimates of model inputs such as PD, LGD, and EAD. However, blaming Basel II may be unfair because, as explained in Chapter 6, the seeds of the crisis were sown well before Basel II was implemented.[1]

This chapter starts by discussing what has become known as Basel II.5. This is a collection of changes to the calculation of market risk capital that was put in place by the Basel Committee on Banking Supervision following the large losses experienced by banks during the crisis. The implementation date for Basel II.5 was December 31, 2011.

The chapter then moves on to consider Basel III, which was a major overhaul of bank regulations, published by the Basel Committee in December 2010. Basel III includes a series of rules concerned with increasing the amount of equity capital that banks have to keep for the risks being taken and tightening the definition of capital. An important new feature of Basel III is the specification of liquidity requirements that must be met by banks. Basel III is being phased in over several years. Full implementation is expected to be complete by 2019.

[1] Also, the United States was behind other countries in implementing Basel II. If Basel II had been fully implemented by the start of the crisis, capital levels at U.S. banks would probably have been lower.

The chapter also discusses other regulations that have been introduced since the 2008 crisis to complement the work of the Basel Committee. These include the Dodd–Frank Act, which was signed into law by President Barack Obama in the United States on July 21, 2010, and rules introduced in the European Union and the United Kingdom.

16.1 Basel II.5

During the credit crisis, it was recognized that some changes were necessary to the calculation of capital for market risk. These changes are referred to as Basel II.5 and, as already mentioned, the implementation date for them was December 31, 2011.[2] There are three changes involving:

1. The calculation of a stressed VaR;
2. A new incremental risk charge; and
3. A comprehensive risk measure for instruments dependent on credit correlation.

The measures have the effect of greatly increasing the market risk capital that large banks are required to hold.

16.1.1 Stressed VaR

The 1996 Amendment to Basel I, where capital was first required for market risk, allowed banks to base capital on a 10-day 99% VaR measure. Most banks use historical simulation to calculate VaR. This is described in Chapter 13. The assumption underlying historical simulation was that the changes in market variables during the following day would be a random sample from their daily changes (percentage or actual) observed during the previous one to four years. The 2003–2006 period was one where the volatility of most market variables was low. As a result, the market risk VaRs calculated during this period for regulatory capital purposes were also low. Furthermore, the VaRs continued to be too low for a period of time after the onset of the crisis, because much of the data used to calculate them continued to come from a low-volatility period.

This led the Basel Committee to introduce the stressed VaR measure. As explained in Section 13.1.3, stressed VaR is determined by basing calculations on how market variables moved during a 250-day (12-month) period of stressed market conditions, rather than on how they moved during the past one to four years. The historical simulation calculations to arrive at a stressed VaR measure assume that the changes in market variables during the next day are a random sample from their daily changes observed during the 250-day period of stressed market conditions.

Basel II.5 requires banks to calculate two VaRs. One is the usual VaR (based on the previous one to four years of market movements). The other is the stressed VaR

[2] See Basel Committee on Banking Supervision, "Revisions to the Basel II Market Risk Framework," February 2011.

(calculated from a stressed period of 250 days). The two VaR measures are combined to calculate a total capital charge. The formula for the total capital charge is

$$\max(\mathrm{VaR}_{t-1},\ m_c \times \mathrm{VaR}_{\mathrm{avg}}) + \max(\mathrm{sVaR}_{t-1},\ m_s \times \mathrm{sVaR}_{\mathrm{avg}})$$

where VaR_{t-1} and sVaR_{t-1} are the VaR and stressed VaR (with a 10-day time horizon and a 99% confidence level) calculated on the previous day. The variables $\mathrm{VaR}_{\mathrm{avg}}$ and $\mathrm{sVaR}_{\mathrm{avg}}$ are the average of VaR and stressed VaR (again with a 10-day time horizon and a 99% confidence level) calculated over the previous 60 days. The parameters m_s and m_c are multiplicative factors that are determined by bank supervisors and are at minimum equal to three. As explained in Section 15.6, the capital requirement prior to Basel II.5 was

$$\max(\mathrm{VaR}_{t-1}, m_c \times \mathrm{VaR}_{\mathrm{avg}})$$

Because stressed VaR is always at least as great as VaR, the formula shows that (assuming $m_c = m_s$) the impact of this new rule is to at least double the capital requirement. In practice, a trebling of the capital requirement as a result of the change was not unusual.

Originally it was considered that 2008 would constitute a good one-year period for the calculation of stressed VaR. Later it was realized that the one-year period chosen should reflect a bank's portfolio. A bank is now required to search for a one-year period that would be particularly stressful for its current portfolio. The stressed period used by one bank is therefore not necessarily the same as that used by another bank.

16.1.2 Incremental Risk Charge

In 2005, the Basel Committee became concerned that exposures in the trading book were attracting less capital than similar exposures in the banking book. Consider a bond. If held in the trading book, the capital would be calculated by applying a multiplier to the 10-day 99% VaR, as discussed in Section 15.6. If held in the banking book (and treated like a loan), capital for the bond would be calculated using VaR with a one-year time horizon and a 99.9% confidence level, as discussed in Section 15.8. The trading-book calculation usually gave rise to a much lower capital charge than the banking-book calculation. As a result, banks tended whenever possible to hold credit-dependent instruments in the trading book.[3]

Regulators proposed an incremental default risk charge (IDRC) in 2005 that would be calculated with a 99.9% confidence level and a one-year time horizon for instruments in the trading book that were sensitive to default risk. The effect of this would have been

[3] If a bank created ABSs from loans in the banking book, as described in Chapter 6, and then bought all the resultant tranches for its trading book, regulatory capital requirements would be lowered even though the bank's risks would be unchanged. This was one reason why banks wanted to securitize loans in the banking book.

that the capital requirement for these instruments equaled the maximum of the capital using trading book calculations and the capital using banking book calculations. In 2008, the Basel Committee recognized that most of the losses in the credit market turmoil of 2007 and 2008 were from changes in credit ratings, widening of credit spreads, and loss of liquidity, rather than solely as a result of defaults. It therefore amended its previous proposals to reflect this and the IDRC became the incremental risk charge (IRC).[4]

The IRC requires banks to calculate a one-year 99.9% VaR for losses from credit-sensitive products in the trading book taking both credit rating changes and defaults into account. Like the IDRC, the aim was to set capital equal to the maximum of that obtained using trading book calculations and that obtained using banking book calculations. Because the instruments subject to the IRC are in the trading book, it is assumed that a bank will have the opportunity to rebalance its portfolio during the course of the year so that default risk is mitigated. Banks are therefore required to estimate a liquidity horizon for each instrument subject to the IRC. The liquidity horizon represents the time required to sell the position or to hedge all material risks in a stressed market.

Suppose that the liquidity horizon for a bond with a credit rating of A is three months. For the purposes of the calculation of VaR over a one-year time horizon, the bank assumes that at the end of three months, if the bond's rating has changed or if it has defaulted, it is replaced by an A-rated bond similar to that held at the beginning of the period. The same thing happens at the end of six months and at the end of nine months. This is known as the *constant level of risk* assumption.

The impact of the constant level of risk assumption is that it is less likely that there will be a default. Instead, small losses are realized from ratings downgrades when rebalancing takes place. The assumption typically has the effect of reducing the one-year 99.9% VaR.[5] The minimum liquidity horizon for IRC is specified by the Basel Committee as three months.

The IRC therefore provides a measure of the default and credit migration risks of credit products over a one-year horizon at a 99.9% confidence level, taking into account the liquidity horizons of individual positions or sets of positions.

16.1.3 The Comprehensive Risk Measure

The comprehensive risk measure (CRM) is designed to take account of risks in what is known as the correlation book. This is the portfolio of instruments such as asset-backed securities (ABSs) and collateralized debt obligations (CDOs) that are sensitive to the correlation between the default risks of different assets. These instruments were discussed in Chapter 6. Suppose a bank has a AAA-rated tranche of an ABS. In the normal market environment, there is very little risk of losses from the tranche. However,

[4] Basel Committee on Banking Supervision, "Guidelines for Computing Capital for Incremental Risk in the Trading Book," July 2009.

[5] See C. Finger, "CreditMetrics and Constant Level of Risk," MSCI, 2010, for a discussion of the constant level of risk assumption.

Table 16.1 Standardized Capital Charge for Correlation-Dependent Instruments

External Credit Assessment	AAA to AA−	A+ to A−	BBB+ to BBB−	BB+ to BB−	Below BB− or Unrated
Securitizations	1.6%	4%	8%	28%	Deduction
Resecuritizations	3.2%	8%	18%	52%	Deduction

in stressed market environments when correlations increase, the tranche is vulnerable—as became apparent during the 2007–2009 crisis.

The CRM is a single capital charge replacing the incremental risk charge and the specific risk charge for instruments dependent on credit correlation. The Basel II.5 standardized approach for calculating the CRM is summarized in Table 16.1. Given the experience of the securitization market during the crisis (see Chapter 6), it is not surprising that capital charges are higher for resecuritizations (e.g., ABS CDOs) than for securitizations (e.g., ABSs). A deduction means than the principal amount is subtracted from capital, which is equivalent to a 100% capital charge.

Basel II.5 allows banks, with supervisory approval, to use their internal models to calculate the CRM. The models developed by banks have to be quite sophisticated to be approved by bank supervisors. For example, they must capture the cumulative impact of multiple defaults, credit spread risk, the volatility of implied correlations, the relationship between credit spreads and implied correlations, recovery rate volatility, the risk of hedge slippage, and potential hedge rebalancing costs. A routine and rigorous program of stress testing is also required. The capital charge calculated from an internal model is subject to a floor calculated from the capital given by the standardized approach.

16.2 Basel III

Following the 2007–2009 credit crisis, the Basel Committee realized that a major overhaul of Basel II was necessary. Basel II.5 increased capital requirements for market risk. The Basel Committee wanted to increase equity capital requirements further. In addition, it considered that the definition of equity capital needed to be tightened and that regulations were needed to address liquidity risk.

Basel III proposals were first published in December 2009. Following comments from banks, a quantitative impact study, and a number of international summits, the final version of the regulations was published in December 2010.[6] There are six parts to the regulations:

1. Capital Definition and Requirements
2. Capital Conservation Buffer

[6] See Basel Committee on Banking Supervision, "Basel III: A Global Regulatory Framework for More Resilient Banks and Banking Systems," June 2011, and Basel Committee on Banking Supervision, "Basel III: International Framework for Liquidity Risk Measurement Standards and Monitoring," December 2010.

3. Countercyclical Buffer
4. Leverage Ratio
5. Liquidity Risk
6. Counterparty Credit Risk

The regulations are being implemented gradually between 2013 and 2019.

16.2.1 Capital Definition and Requirements

Under Basel III, a bank's total capital consists of:

1. Tier 1 equity capital
2. Additional Tier 1 capital
3. Tier 2 capital

There is no Tier 3 capital.

Tier 1 equity capital (also referred to as common equity Tier 1, or CET1, capital) includes share capital and retained earnings but does not include goodwill or deferred tax assets. It must be adjusted downward to reflect defined benefit pension plan deficits but is not adjusted upward to reflect defined benefit plan surpluses. (See Section 3.12 for a discussion of defined benefit plans.) Changes in retained earnings arising from securitized transactions are not counted as part of capital for regulatory purposes. The same is true of changes in retained earnings arising from the bank's own credit risk. (The latter is referred to as DVA and will be discussed in Chapter 20.) There are rules relating to the inclusion of minority interests and capital issued by consolidated subsidiaries.

The additional Tier 1 (AT1) capital category consists of items, such as non-cumulative preferred stock, that were previously Tier 1 but are not common equity. Tier 2 capital includes debt that is subordinated to depositors with an original maturity of five years.

Common equity is referred to by the Basel Committee as "going-concern capital." When the bank is a going concern (i.e., has positive equity capital), common equity absorbs losses. Tier 2 capital is referred to as "gone-concern capital." When the bank is no longer a going concern (i.e., has negative capital), losses have to be absorbed by Tier 2 capital. Tier 2 capital ranks below depositors in a liquidation. While Tier 2 capital remains positive, depositors should in theory be repaid in full.

The capital requirements are as follows:

1. Tier 1 equity capital must be at least 4.5% of risk-weighted assets at all times.
2. Total Tier 1 capital (Tier 1 equity capital plus additional Tier 1 capital) must be at 6% of risk-weighted assets at all times.
3. Total capital (total Tier 1 plus Tier 2) must be at least 8% of risk-weighted assets at all times.

Basel I required Tier 1 equity capital to be at least 2% of risk-weighted assets and total Tier 1 capital to be at least 4% of risk-weighted assets. The Basel III rules are much more demanding because (a) these percentages have been increased and (b) the definition of what qualifies as equity capital for regulatory purposes has been tightened. However, the Tier 1 plus Tier 2 requirement is the same as under Basel I and Basel II.

The transitional arrangements were that Tier 1 equity capital and total Tier 1 capital must be 3.5% and 4.5%, respectively, by January 1, 2013. They must be 4% and 5.5%, respectively, by January 1, 2014. The new capital levels had to be in place by January 1, 2015. The new rules for the definition of what constitutes capital were phased in over a longer period stretching until January 1, 2018.

The Basel Committee also calls for more capital for "systemically important" banks, as will be discussed later in this chapter.

16.2.2 Capital Conservation Buffer

In addition to the capital requirements just mentioned, Basel III requires a capital conservation buffer in normal times consisting of a further amount of CET1 capital equal to 2.5% of risk-weighted assets. This provision is designed to ensure that banks build up capital during normal times so that it can be run down when losses are incurred during periods of financial difficulties. (The argument in favor of this is that it is much easier for banks to raise capital during normal times than during periods of stressed market conditions.) In circumstances where the capital conservation buffer has been wholly or partially used up, banks are required to constrain their dividends until the capital has been replenished. The dividend rules are summarized in Table 16.2. For example, if Tier 1 equity capital is 5.5% of risk-weighted assets, the minimum retained earnings is 80% so that the maximum dividends as a percent of retained earnings is 20%. The difference between the 4.5% basic equity capital requirement and the 7.0% equity capital requirement that includes the capital conservation buffer is in some ways analogous to the difference between the MCR and SCR in Solvency II (see Section 15.12).

The capital conservation buffer means that the Tier 1 equity capital that banks are required to keep in normal times (excluding other requirements that will be discussed later) is 7% of risk-weighted assets; total Tier 1 capital is required to be at least 8.5% of risk-weighted assets; Tier 1 plus Tier 2 capital is required to be at least 10.5% of risk-weighted assets. These numbers can decline to 4.5%, 6%, and 8% in stressed market conditions (because of losses), but banks are then under pressure to bring capital back up to the required levels. One of the consequences of the increased equity capital

Table 16.2 Dividend Restrictions Arising from the Capital Conservation Buffer

Tier 1 Equity Capital Ratio	Minimum Percent of Earnings Retained
4.000% to 5.125%	100%
5.125% to 5.750%	80%
5.750% to 6.375%	60%
6.375% to 7.000%	40%
>7%	0%

Table 16.3 Dividend Restrictions Arising from the Capital Conservation Buffer and 2.5% Countercyclical Buffer

Tier 1 Equity Capital Ratio	Minimum Percent of Earnings Retained
4.50% to 5.75%	100%
5.75% to 7.00%	80%
7.00% to 8.25%	60%
8.25% to 9.50%	40%
>9.50%	0%

requirement is that banks may find it difficult to achieve the returns on equity that they had during the 1990 to 2006 period. However, bank shareholders can console themselves that bank stock is less risky as a result of the extra capital.

The capital conservation buffer requirement is being phased in between January 1, 2016, and January 1, 2019.

16.2.3 Countercyclical Buffer

In addition to the capital conservation buffer, Basel III has specified a countercyclical buffer. This is similar to the capital conservation buffer, but the extent to which it is implemented in a particular country is left to the discretion of national authorities. The buffer is intended to provide protection for the cyclicality of bank earnings. The buffer can be set to between 0% and 2.5% of total risk-weighted assets and must be met with Tier 1 equity capital.

For jurisdictions where the countercyclical buffer is non-zero, Table 16.2 is modified. For example, when the countercyclical buffer is set at its maximum level of 2.5%, it is replaced by Table 16.3. Like the capital conservation buffer, the countercyclical buffer requirements will be phased in between January 1, 2016, and January 1, 2019.

16.2.4 Leverage Ratio

In addition to the capital requirements based on risk-weighted assets, Basel III specifies a minimum leverage ratio of 3%.[7] The leverage ratio is the ratio of a capital measure to an exposure measure. The capital measure is total Tier 1 capital. The exposure measure is the sum of (a) on-balance-sheet exposures, (b) derivatives exposures, (c) securities financing transaction exposures, and (d) off-balance-sheet items. No risk-weighting adjustments are made. The on-balance-sheet exposures include all assets on the balance sheet. Derivatives exposures are calculated as "replacement cost plus add-on" similar to the way they are calculated under Basel I (see Section 15.3). Securities financing transaction exposures include transactions such as repurchase agreements and security

[7] See Bank for International Settlements, "Basel III Leverage Ratio Framework and Disclosure Requirements," January 2014.

lending/borrowing when the transaction does not lead to balance sheet assets. Off-balance-sheet items include loan commitments, loan substitutes, acceptances, and letters of credit. The leverage ratio was introduced in January 2018 after a transition period.

Regulators in some countries have indicated that they regard Basel's leverage ratio of 3% as too low. In April 2014, regulators in the United States proposed rules requiring a 5% leverage ratio for eight large bank holding companies (which, as we will explain, are known as G-SIBs), and 6% for the FDIC-insured subsidiaries of these bank holding companies. In October 2014, the Bank of England's Financial Policy Committee increased the leverage ratio for UK banks to 4.05%, a level that could be increased to 4.95% in boom times to rein in excessive lending. In China the leverage ratio has been set at 4%.

Why did the Basel Committee introduce the leverage ratio? The reason is that regulators thought that banks had too much discretion in the way risk-weighted assets were calculated. They have far less discretion in the way "total exposure" is calculated. It should be emphasized that this does not mean that regulators are discarding capital requirements based on risk-weighted assets. They require banks to satisfy both (a) the ratios of capital to risk-weighted assets mentioned earlier in this chapter and (b) the ratio of capital to non-risk-weighted exposure leverage requirement explained here.

One of the two capital ratios is likely to be the critical ratio for a bank (i.e., the ratio it is closest to not complying with). If the risk-weighted assets ratio is the critical one, then arguably the leverage ratio can be justified because it provides useful additional information to regulators. But if the leverage ratio is the critical one (and there is some evidence that it has been for some U.S. banks), a bank might be encouraged to hold risky assets because they have the same effect on the leverage ratio as safe assets but provide a higher expected return. This could be an unintended adverse consequence for regulators.

16.2.5 Liquidity Risk

Prior to the crisis, the focus of the Basel regulations had been on ensuring that banks had sufficient capital for the risks they were taking. It turned out that many of the problems encountered by financial institutions during the crisis were not as a result of a shortage of capital. They were instead a result of liquidity risks taken by the banks.

Liquidity risks arise because there is a tendency for banks to finance long-term needs with short-term funding, such as commercial paper. Provided the bank is perceived by the market to be financially healthy, this is usually not a problem.[8] Suppose that a bank uses 90-day commercial paper to fund its activities. When one 90-day issue of commercial paper matures, the bank refinances with a new issue; when the new issue matures, it refinances with another issue; and so on. However, as soon as the bank experiences financial difficulties—or is thought to be experiencing financial difficulties—it is liable

[8] If the funds are being used to finance long-term fixed-rate loans and interest rates rise, net interest margins are squeezed. But this risk can be hedged with instruments such as interest rate swaps (see Example 5.1).

to become impossible for the bank to roll over its commercial paper. This type of problem led to the demise of Northern Rock in the United Kingdom and Lehman Brothers in the United States.

Basel III has introduced requirements involving two liquidity ratios that are designed to ensure that banks can survive liquidity pressures. The ratios are:

1. Liquidity Coverage Ratio (LCR); and
2. Net Stable Funding Ratio (NSFR)

The LCR focuses on a bank's ability to survive a 30-day period of liquidity disruptions. It is defined as:

$$\frac{\text{High-Quality Liquid Assets}}{\text{Net Cash Outflows in a 30-Day Period}}$$

The 30-day period considered in the calculation of this ratio is one of acute stress involving a downgrade of the bank's debt by three notches (e.g., from AA− to A−), a partial loss of deposits, a complete loss of wholesale funding, increased haircuts on secured funding (so that instruments posted as collateral are not valued as highly), and drawdowns on lines of credit. The Basel III regulations require the ratio to be greater than 100% so that the bank's liquid assets are sufficient to survive these pressures.

The NSFR focuses on liquidity management over a period of one year. It is defined as

$$\frac{\text{Amount of Stable Funding}}{\text{Required Amount of Stable Funding}}$$

The numerator is calculated by multiplying each category of funding (capital, wholesale deposits, retail deposits, etc.) by an available stable funding (ASF) factor, reflecting their stability. As shown in Table 16.4, the ASF for wholesale deposits is less than that for retail deposits, which in turn is smaller than that for Tier 1 or Tier 2 capital. The denominator

Table 16.4 ASF Factors for Net Stable Funding Ratio

ASF Factor	Category
100%	Tier 1 and Tier 2 capital Preferred stock and borrowing with a remaining maturity greater than one year
90%	"Stable" demand deposits and term deposits with a remaining maturity less than one year provided by retail or small business customers
80%	"Less Stable" demand deposits and term deposits with a remaining maturity less than one year provided by retail or small business customers
50%	Wholesale demand deposits and term deposits with a remaining maturity less than one year provided by non-financial corporates, sovereigns, central banks, multilateral development banks, and public-sector entities
0%	All other liability and equity categories

Table 16.5 RSF Factors for Net Stable Funding Ratio

RSF Factor	Category
0%	Cash
	Short-term instruments, securities, loans to financial entities if they have a residual maturity of less than one year
5%	Marketable securities with a residual maturity greater than one year if they are claims on sovereign governments or similar bodies with a 0% risk weight
20%	Corporate bonds with a rating of AA− or higher and a residual maturity greater than one year
	Claims on sovereign governments or similar bodies with a risk weight of 20%
50%	Gold, equity securities, bonds rated A+ to A−
65%	Residential mortgages
85%	Loans to retail and small business customers with a remaining maturity less than one year
100%	All other assets

is calculated from the items requiring funding. Each category of these is multiplied by a required stable funding (RSF) factor to reflect the permanence of the funding required. Some of the applicable factors are indicated in Table 16.5.

Basel III requires the NSFR to be greater than 100% so that the calculated amount of stable funding is greater than the calculated required amount of stable funding.

Example 16.1

A bank has the following balance sheet:

Cash	5	Retail Deposits (stable)	40
Treasury Bonds (> 1 yr)	5	Wholesale Deposits	48
Mortgages	20	Tier 2 Capital	4
Small Business Loans	60	Tier 1 Capital	8
Fixed Assets	10		
	100		100

The Amount of Stable Funding is

$$40 \times 0.9 + 48 \times 0.5 + 4 \times 1.0 + 8 \times 1.0 = 72$$

The Required Amount of Stable Funding is

$$5 \times 0 + 5 \times 0.05 + 20 \times 0.65 + 60 \times 0.85 + 10 \times 1.0 = 74.25$$

The NSFR is therefore

$$\frac{72}{74.25} = 0.970$$

or 97.0%. The bank does not therefore satisfy the NSFR requirement.

The new rules are tough and have the potential to dramatically change bank balance sheets. It was estimated in September 2014 that the rules will require U.S. banks to hold an extra $100 billion of liquid assets. The LCR requirement was scheduled by the Basel Committee to be implemented in stages between January 1, 2015, and January 1, 2019 (but regulators in the United States required full implementation by January 1, 2017). The NSFR requirement was scheduled to be implemented on January 1, 2018.

16.2.6 Counterparty Credit Risk

For each of its derivatives counterparties, a bank calculates a quantity known as the *credit value adjustment* (CVA). This is the expected loss because of the possibility of a default by the counterparty. The way in which the calculation of CVA is carried out is described in Chapter 20. Reported profit is reduced by the total of the CVAs for all counterparties.

As we will see in Chapter 20, the CVA for a counterparty can change because either (a) the risk factors underlying the value of the derivatives entered into with the counterparty change or (b) the credit spreads applicable to the counterparty's borrowing change. Basel III requires the CVA risk arising from changing credit spreads to be a component of market risk capital. As will be explained in Chapter 20, once CVA has been calculated, it is a relatively simple matter to calculate the delta and gamma with respect to a parallel shift in the term structure of the counterparty's credit spread. These can be used to calculate appropriate market risk measures.

The Basel Committee proposed changes to the calculation of capital for CVA risk in 2015.[9] A major objective was to allow banks that hedged the impact of both credit spread changes and risk factor changes on CVA to take both into account when calculating regulatory capital. Without this change, hedges for risk factor changes would increase capital charges.

16.2.7 G-SIBs, SIFIs, and D-SIBs

Regulators are particularly concerned that large, systemically important financial institutions keep sufficient capital to avoid a repeat of the government bailouts during the 2007 to 2009 credit crisis.

The term G-SIB stands for global systemically important bank, whereas the term SIFI (systemically important financial institution) is used to describe both banks and nonbanks that are considered to be systemically important. The popular view of SIFIs is that they are "too big to fail," and have been identified as the financial institutions that will have to be bailed out if they run into financial difficulties.

The systemic importance of a bank or other financial institution depends on the effect that its failure could have on the global financial system. This in turn depends on the

[9] See Basel Committee on Banking Supervision, "Review of Credit Valuation Adjustment Risk Framework," October 1, 2015.

nature of its activities and the contracts it has entered into with other financial institutions globally. The Basel Committee uses a scoring methodology to determine which banks are G-SIBs. Other approaches involving network theory have been attempted by some researchers.

In 2013, the Basel Committee published the final version of rules that call for more Tier 1 equity capital for G-SIBs.[10] G-SIBs are categorized according to whether the extra equity capital is 1%, 1.5%, 2%, 2.5%, or 3.5% of risk-weighted assets. The new rules took effect in January 2016.

A list of G-SIBs is published by the Financial Stability Board annually. In November 2017, the list included 30 banks. Of these, 17 were in the 1% category, 8 were in the 1.5% category, 4 were in the 2% category, and 1 (JPMorgan Chase) was in the 2.5% category. None were in the 3.5% category. G-SIBs are required to keep CET1 capital equal to a baseline 4.5% of risk-weighted assets plus a further 2.5% for the capital conservation buffer plus any extra amounts required by national supervisors. Suppose the total is X% of risk-weighted assets. From Section 16.2.1, all Tier 1 capital (CET1 plus AT1) must be at least (X + 1.5)% of risk-weighted assets, and Tier 1 plus Tier 2 must be at least (X + 3.5)% of risk-weighted assets.

In November 2014, the Financial Stability Board issued proposals concerning the total loss-absorbing capacity (TLAC) of G-SIBs. The proposals were developed in consultation with the Basel Committee on Banking Supervision and were in response to a call by G20 leaders at the 2013 St. Petersburg Summit. TLAC consists of instruments that can absorb losses and protect depositors in the event of a default. The instruments include equity, preferred shares, and subordinated debt. From January 1, 2019, the minimum TLAC for G-SIBs will be 16% of risk-weighted assets, increasing to 18% on January 1, 2022. Non-compliance with TLAC could impede a bank's ability to issue dividends.

Some countries are requiring greater capital than the minimums set by the Basel Committee and the Financial Stability Board or are implementing the rules earlier than required. An example of a country requiring that its banks hold high amounts of capital is Switzerland, which has two G-SIBs (UBS and Credit Suisse). The high capital requirements are understandable because the banks are large in relation to the Swiss economy, so that a failure of either of them would be catastrophic.

National regulators designate some banks that have not been classified as G-SIBs as domestic systemically important banks (D-SIBs). These banks may be subject to capital requirements higher than the minimum, extra disclosure requirements, or (see Chapter 22) stringent stress tests. In the United States in 2016, eight banks (Bank of America, Bank of New York Mellon, Citigroup, Goldman Sachs, JPMorgan Chase, Morgan Stanley, State Street, and Wells Fargo) were designated as G-SIBs, and all other banks with assets over $50 billion were designated as D-SIBs.

[10] See Basel Committee on Banking Supervision, "Global Systemically Important Banks: Updated Assessment Methodology and the Higher Loss Absorbency Requirement," July 2013.

16.3 Contingent Convertible Bonds

An interesting development in the capitalization of banks has been what are known as *contingent convertible bonds* (CoCos). Traditionally, convertible bonds have been bonds issued by a company where, in certain circumstances, the holder can choose to convert them into equity at a predetermined exchange ratio. Typically the bond holder chooses to convert when the company is doing well and the stock price is high. CoCos are different in that they automatically get converted into equity when certain conditions are satisfied. Typically, these conditions are satisfied when the company is experiencing financial difficulties.

CoCos are attractive to banks because in normal times the bonds are debt and allow the bank to report a relatively high return on equity. When the bank experiences financial difficulties and incurs losses, the bonds are converted into equity and the bank is able to continue to maintain an equity cushion and avoid insolvency. From the point of view of regulators, CoCos are potentially attractive because they avoid the need for a bailout. Indeed, the conversion of CoCos is sometimes referred to as a "bail-in." New equity for the financial institution is provided from within by private-sector bondholders rather than from outside by the public sector.

A key issue in the design of CoCos is the specification of the trigger that forces conversion and the way that the exchange ratio (number of shares received in exchange for one bond) is set. A popular trigger in the bonds issued so far is the ratio of Tier 1 equity capital to risk-weighted assets. Another possible trigger is the ratio of the market value of equity to book value of assets.

Lloyd's Banking Group, Rabobank Nederlands, and Credit Suisse were among the first banks to issue CoCos. Business Snapshot 16.1 provides a description of the bonds issued by Credit Suisse in 2011. These bonds get converted into equity if either Tier 1 equity capital falls below 7% of risk-weighted assets or the Swiss bank regulators determine that the bank requires public-sector support. It has been estimated that over $1 trillion of CoCos will be issued by banks during the decade beginning 2010 as they respond to the new Basel III regulatory requirements on capital adequacy.

CoCos (prior to conversion) qualify as additional Tier 1 capital if the trigger, defined in terms of the ratio of Tier 1 equity capital to risk-weighted assets, is set at 5.125% or higher. Otherwise they qualify as Tier 2 capital.

16.4 Use of Standardized Approaches and SA-CCR

In December 2017, the Basel Committee announced that, starting in 2022, it would require the implementation of a standardized approach for all capital calculations. A bank's total capital requirement will be the maximum of (a) that calculated as before using approved internal models and (b) a certain percentage of that given by the standardized approaches. The percentage will 50% in 2022, rising to 72.5% in 2027.

BUSINESS SNAPSHOT 16.1

Credit Suisse's CoCo Bond Issues

On February 14, 2011, Credit Suisse announced that it had agreed to exchange $6.2 billion of existing investments by two Middle Eastern investors, Qatar Holding LLC and the Olayan Group LLC, for CoCo bonds. The bonds automatically convert into equity if either of the following two conditions are satisfied:

1. The Tier 1 equity capital of Credit Suisse falls below 7% of risk-weighted assets.
2. The Swiss bank regulator determines that Credit Suisse requires public-sector support to prevent it from becoming insolvent.

Credit Suisse followed this announcement on February 17, 2011, with a public issue of $2 billion of CoCos. These securities have similar terms to ones held to the Middle Eastern investors and were rated BBB+ by Fitch. They mature in 2041 and can be called any time after August 2015. The coupon is 7.875%. Any concerns that the market had no appetite for CoCos were alleviated by this issue. It was 11 times oversubscribed.

Credit Suisse indicated that it plans to satisfy one-third of the non-equity capital requirement with bonds similar to those just described and two-thirds of the non-equity capital requirement with bonds where the conversion trigger is about 5% (rather than 7%) of risk-weighted assets.

The standardized approach for market risk capital applicable in 2022 will be explained in Chapter 18. As explained in Chapter 23, operational risk capital will be based entirely on a standardized approach. The standardized risk weights for credit risk are a revision of the Basel II weights in Section 15.8.1.

The CEM standardized approach for calculating exposure at default (EAD) for derivatives in Basel I has been improved with what is known as SA-CCR.[11] EAD is

$$1.4 \times (\mathrm{RC} + \mathrm{PFE})$$

where RC is the replacement cost and PFE is potential future exposure. For transactions where no collateral is posted, the RC is calculated as

$$\mathrm{RC} = \max(V, 0)$$

[11] See Basel Committee on Banking Supervision, "The Standardized Approach for Measuring Counterparty Credit Risk Exposures," March 2014.

For transactions that are subject to margin agreements this is adjusted to

$$RC = \max(V - C, D, 0)$$

where C is the net value of the collateral currently posted by the counterparty (with the collateral posted by the bank being considered negative). The variable D is the change in the V that could occur without any additional collateral being received (e.g., because of minimum transfer amounts in the collateral agreement). It reflects the terms of the collateral agreement. For example, if $V = C = 50$ and V could increase to 51 without any additional collateral being received, $D = 1$ and the RC = 1. In the case of non-cash collateral, a haircut is applied in the calculation of C.

The PFE is a relatively sophisticated way of calculating the add-on to the current exposure. Two factors taken into account are:

1. When the counterparty has posted excess collateral, risk is reduced and the reduction in risk increases as the amount of the excess collateral increases.
2. When no collateral is posted, credit exposure reduces as V becomes more negative. The PFE add-on amount when V is slightly negative should be greater than when V is highly negative.

16.5 Dodd–Frank Act

The Dodd–Frank Act in the United States was signed into law in July 2010. Its aim is to prevent future bailouts of financial institutions and protect the consumer. A summary of the main regulations is as follows:

1. Two new bodies, the Financial Stability Oversight Council (FSOC) and the Office of Financial Research (OFR), were created to monitor systemic risk and research the state of the economy. Their tasks are to identify risks to the financial stability of the United States, promote market discipline, and maintain investor confidence.
2. The orderly liquidation powers of the Federal Deposit Insurance Corporation (FDIC) were expanded. The Office of Thrift Supervision was eliminated.
3. The amount of deposits insured by the FDIC was increased permanently to $250,000. (Previously, the $250,000 limit had been regarded as temporary.)
4. Regulations were introduced requiring large hedge funds and similar financial intermediaries to register with the SEC and report on their activities.
5. A Federal Insurance Office was created to monitor all aspects of the insurance industry and work with state regulators.
6. Proprietary trading and other similar activities of deposit-taking institutions were curtailed. This is known as the "Volcker rule" because it was proposed by former Federal Reserve chairman Paul Volcker.

7. Some high-risk trading operations were required to be spun off into separately capitalized affiliates.
8. Standardized over-the-counter derivatives must be traded on electronic platforms known as swap execution facilities (SEFs). These are similar to exchanges and will, it is hoped, make prices in the OTC market more transparent. Standardized over-counter derivatives between financial institutions must be cleared by central clearing parties (CCPs). (See Chapter 17 for a further discussion of CCPs.) The Commodity Futures Trading Commission (CFTC) was given responsibility for monitoring the activities of CCPs and SEFs.
9. The Federal Reserve was required to set risk management standards for systemically important financial utilities engaged in activities such as payment, settlement, and clearing.
10. Protection for investors was increased and improvements were made to the regulation of securities.
11. Rating agencies were required to make the assumptions and methodologies behind their ratings more transparent and the potential legal liabilities of rating agencies were increased. An Office of Credit Ratings was created at the SEC to provide oversight of rating agencies.
12. The use of external credit ratings in the regulation of financial institutions was discontinued. (This provision of the Act brings Dodd–Frank into direct conflict with the Basel Committee, which, as we have seen in this chapter and the last one, does make some use of external credit ratings.)
13. A Consumer Financial Protection Bureau was created within the Federal Reserve to ensure that consumers get clear and accurate information when they shop for financial products, such as mortgages and credit cards.
14. Issuers of securitized products were required (with some exceptions) to keep 5% of each product created.
15. Federal bank regulators were required to issue regulations that discourage the use of compensation arrangements that might lead to excessive risk taking (e.g., compensation arrangements based on short-run performance). Shareholders were given a non-binding vote on executive compensation. A requirement that board compensation committees be made up of independent directors was introduced.
16. Mortgage lenders were required to make a reasonable good faith determination based on verified and documented information that the borrower has the ability to repay a loan. Failure to do this might lead to a foreclosure being disallowed.
17. Large financial firms were required to have board committees where at least one expert has risk management experience at a large complex firm.
18. The FDIC was allowed to take over a large financial institution when it failed and sell its assets, imposing losses on shareholders and creditors, with the costs of failures being paid for by the financial industry.
19. FSOC and OFR, which as mentioned earlier have the responsibility of monitoring systemic risk, are charged with identifying systemically important financial institutions (SIFIs).

20. The Federal Reserve Board and the FDIC require all SIFIs to prepare what is known as *living wills*, mapping out how they could raise funds in a crisis and how their operations would be wound up in the event of failure.

The Dodd–Frank Act did not define a future role for Fannie Mae and Freddie Mac, which were key players in the U.S. mortgage market. These agencies were taken over by the U.S. government in September 2008.

President Trump has indicated his intention to dismantle some parts of Dodd-Frank. The Volcker rule and the Consumer Financial Protection Bureau are the most likely targets.

16.6 Legislation in Other Countries

The large banks are truly global and when regulations vary throughout the world, they are liable to move all or part of their operations from one jurisdiction to another to obtain more favorable treatment. Although all countries are subject to the same Basel III rules, local regulators have some discretion in the application of rules, and legislation is not the same across all countries. In 2011, the Swiss bank UBS made headlines by suggesting that it might move its investment bank headquarters from Zurich to London, Singapore, or New York to avoid the higher capital requirements imposed by Swiss regulators. (It did not do this.)

The previous section outlined the rules introduced by legislators in the United States. Legislation in other countries has addressed some of the same issues. In the United Kingdom, an independent committee was chaired by Sir John Vickers to consider issues facing the banking industry, and new legislation in the form of the Financial Services (Banking Reform) Act 2013 was passed. In the European Union a committee headed by Erkki Liikanen was set up in November 2011 and published its report in October 2012.

Some of the rules and recommendations in the United Kingdom, European Union, and other countries are similar to those in the United States. For example, the Dodd–Frank Act requires originators of securitized products in the United States to keep 5% of all assets created. (See item 14 of the previous list.) A similar provision exists in the Capital Requirement Directive 2 (CRD2) of the European Union.[12] Most national legislators have agreed that standardized over-the-counter derivatives should be cleared through central clearing houses instead of being cleared bilaterally (see list item 8).

Sometimes the rules in different countries are similar but not exactly the same. The Dodd–Frank Act included the requirement that standardized derivatives be traded on swap execution facilities (see list item 8). The European Union has introduced similar trading platforms, known as organized trading facilities (OTFs). However, it is not as dogmatic about requiring their use for all trading in standardized products.

---

[12]However, Germany has increased the 5% to 10%.

The most controversial aspect of Dodd–Frank is probably the Volcker rule (see item 6), which prohibits banks from proprietary trading and restricts investment in hedge funds and private equity companies by banks and their affiliates. The rationale for the rule is reasonable: Banks should not be allowed to speculate with depositors' funds, because those funds are insured by the FDIC. The rule may be difficult to enforce, though, because for a bank with a large portfolio of positions, it can be difficult to tell whether a particular new trade is entered into for hedging or speculative purposes.[13] In spite of this, many U.S. banks have closed down their proprietary trading desks, and some employees working on those desks have left to join hedge funds. The Vickers committee in the United Kingdom recognized the importance of protecting depositors' funds but did not go so far as to prohibit proprietary trading. The retail operations of banks must (by 2019) be ring-fenced so that they are insulated from riskier activities such as trading and investment banking. The Liikanen committee in the European Union similarly requires a separation of core banking activities from proprietary trading and other high-risk activities.

Regulators in most countries consider living wills (see list item 20) to be important for SIFIs and are applying pressure on SIFIs to develop them. SIFIs tend to have developed complex organizational structures for tax and regulatory purposes. The living will requirement may result in this being simplified so that the different activities of a SIFI are in separately capitalized legal entities, not all of which need to be bailed out in the event of the SIFI's failure. Regulators have the option of forcing SIFIs to divest certain operations, or even break up entirely, if their living wills are deemed unsatisfactory. They hope that living wills avoid a replay of the panic and cross-border litigation that erupted when the Lehman Brothers investment bank collapsed in 2008. They also think that the multiyear litigation concerned with the termination of Lehman's derivatives portfolio could have been avoided if Lehman and its ISDA master agreements had been structured so that it was possible to keep the portfolio alive and to unwind it in an orderly way in the few days after bankruptcy.

Compensation is an important issue. Pre-crisis, the annual bonus was a large part of the compensation for many traders and other employees and led them to have a relatively short-term horizon in their decision making. If losses were incurred after the payment of a bonus, they did not have to return the bonus. Many banks have recognized the problem and voluntarily moved to bonuses that are deferred by being spread out over three to five years, rather than all being paid in one year. If a trader shows good results in one year and bad results in the next, some of the bonus applicable to the good year will be deferred and then "clawed back" during the bad year. The Dodd–Frank Act restrictions on pay in the financial sector are relatively mild. When financial institutions received funds during the crisis under the Troubled Asset Relief Program (TARP), compensation was restricted. But, as soon as the funds were paid back, banks had much more freedom in their compensation arrangements.[14]

---

[13]A joke circulating on Wall Street was that banks would have to employ psychologists or psychiatrists to determine what was going on in the mind of the trader at the time of a trade!

[14]Indeed, this was a major incentive to repay the TARP money as quickly as possible!

Some other countries have restricted compensation. Sometimes the restriction is temporary. For example, in 2009 the United Kingdom introduced a one-time "supertax" on bonuses in excess of £25,000. But in other cases it is permanent. For example, in April 2013, the European Union introduced CRD 4, which imposes a cap on bankers' bonuses. Under this directive the maximum ratio of bonus to fixed pay is 1.0 (with some flexibility to increase that ratio to 2.0 with shareholder approval).

Summary

The financial crisis that started in 2007 was the worst that many parts of the world had seen since the 1930s. Some financial institutions failed. Others had to be bailed out with taxpayers' money. Not surprisingly, both the Basel Committee and national governments decided that a major overhaul of the regulations affecting financial institutions was required.

The Basel II.5 regulations increased the capital banks were required to keep for market risk. They recognized that capital should reflect the volatilities and correlations experienced during stressed market conditions as well as during normal market conditions; they eliminated some of the ways banks could reduce regulatory capital by moving items from the banking book to the trading book; and they created a special capital requirement for derivatives dependent on credit correlation, which had been a particular problem during the crisis.

Basel III dramatically increased the amount of equity capital banks were required to keep. It also recognized that many of the problems of banks during the crisis were liquidity problems and imposed new liquidity requirements for financial institutions.

National governments have also introduced new rules for financial institutions. In the United States, the Dodd–Frank Act has many provisions designed to protect consumers and investors, avoid future bailouts, and monitor the functioning of the financial system more carefully. Similar, though not necessarily identical, regulations exist in other countries.

Exactly how Basel III and national legislation such as Dodd–Frank will be implemented is still somewhat uncertain—and this uncertainty is one of the major risks that banks face. How successful will the measures be once they have been implemented? We will not know this for some time. One problem facing regulators is what are referred to as *unintended consequences*. Basel I had the unintended consequence of discouraging loans to high-quality corporations because of the 100% risk weight that would be assigned. The 1996 Amendment and the development of the credit derivatives market that came after it encouraged banks to find ways of moving credit risks from the banking book to the trading book in order to reduce capital requirements. There will no doubt be unintended consequences of Basel III and the legislation that is being introduced throughout the world. Hopefully, the benefits of the new measures will outweigh any harm to the financial system arising from the unintended consequences.

Further Reading

Acharya, V. V., T. F. Cooley, M. P. Richardson, and I. Walter. *Regulating Wall Street: The Dodd–Frank Act and the New Architecture of Global Finance*. Hoboken, NJ: John Wiley & Sons, 2011.

Bank for International Settlements. "Basel III: Finalizing Post-Crisis Reforms," December 2017.

Basel Committee on Banking Supervision. "Basel III: A Global Regulatory Framework for More Resilient Banks and Banking Systems," June 2011.

Basel Committee on Banking Supervision. "Basel III: International Framework for Liquidity Risk Measurement Standards and Monitoring," December 2010.

Basel Committee on Banking Supervision. "Guidelines for Computing Capital for Incremental Risk in the Trading Book," July 2009.

Basel Committee on Banking Supervision. "Revisions to the Basel II Market Risk Framework," February 2011.

Finger, C. "CreditMetrics and Constant Level of Risk." MSCI, 2010.

Practice Questions and Problems (Answers at End of Book)

16.1 What are the three major components of Basel II.5?

16.2 What are the six major components of Basel III?

16.3 What is the difference between VaR as it has been traditionally measured and stressed VaR?

16.4 Explain how the incremental risk charge is calculated. Why was it introduced by the Basel Committee?

16.5 What is the difference between the capital required for a AAA-rated ABS with principal of \$100 million and a AAA-rated ABS CDO with a principal of \$100 million using the standardized approach?

16.6 By how much has the Tier 1 equity capital (including the capital conservation buffer) increased under Basel III compared with the Tier 1 equity capital requirement under Basel I and II?

16.7 Suppose that the Tier 1 equity ratio for a bank is 6%. What is the maximum dividend, as a percent of earnings, that can be paid if (a) there is no countercyclical buffer and (b) there is a 2.5% countercyclical buffer?

16.8 Explain how the leverage ratio differs from the usual capital ratios calculated by regulators.

16.9 Explain how the liquidity coverage ratio and the net stable funding ratio are defined.

16.10 How would the net stable funding ratio in Example 16.1 change if half the wholesale deposits were replaced by stable retail deposits?

16.11 What is CVA? What new regulations concerning CVA were introduced in Basel III?

16.12 Explain how CoCo bonds work. Why are they attractive to (a) banks and (b) regulators?

Further Questions

16.13 Explain one way that the Dodd–Frank Act is in conflict with (a) the Basel international regulations and (b) the regulations introduced by other national governments.

16.14 A bank has the following balance sheet:

Cash	3	Retail Deposits (stable)	25
Treasury Bonds (>1 yr)	5	Retail Deposits (less stable)	15
Corporate Bonds Rated A	4	Wholesale Deposits	44
Mortgages	18	Preferred Stock (>1 yr)	4
Small Business Loans (<1 yr)	60	Tier 2 Capital	3
Fixed Assets	10	Tier 1 Capital	9
	100		100

(a) What is the Net Stable Funding Ratio?

(b) The bank decides to satisfy Basel III by raising more (stable) retail deposits and keeping the proceeds in Treasury bonds. What extra retail deposits need to be raised?

Chapter 17

Regulation of the OTC Derivatives Market

In Chapter 5, we distinguished between the over-the-counter (OTC) and exchange-traded markets. The exchange-traded market is a market where products developed by an exchange are bought and sold on a trading platform developed by the exchange. A market participant's trade must be cleared by a member of the exchange clearing house. The exchange clearing house requires margin (i.e., collateral) from its members, and the members require margin from the brokers whose trades they are clearing. The brokers in turn require margin from their clients.

The OTC market is a market where financial institutions, fund managers, and corporate treasurers deal directly with each other. An exchange is not involved. Before the 2007–2008 credit crisis, the OTC market was largely unregulated. Two market participants could enter into any trade they liked. They could agree to post collateral or not post collateral. They could agree to clear the trade directly with each other or use a third party. Also, they were under no obligation to disclose details of the trade to anyone else.

Since the crisis, the OTC market has been subject to a great deal of regulation. This chapter will explain the regulations and show that regulatory pressure is leading to the OTC market becoming more like the exchange-traded market.

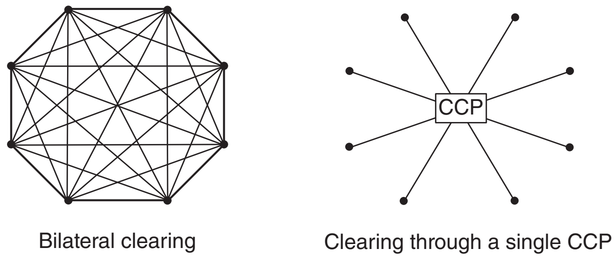


Figure 17.1 Bilateral and Central Clearing

17.1 Clearing in OTC Markets

We start by describing how transactions are cleared in the OTC market. There are two main approaches: central clearing and bilateral clearing. They are illustrated schematically in Figure 17.1 (which makes the simplifying assumption that there are only eight market participants and only one CCP). In bilateral clearing, market participants clear transactions with each other. In central clearing, a third party, known as a central counterparty (CCP), clears the transactions.

17.1.1 Margin

Before proceeding to describe bilateral and central clearing in more detail, we review the operation of margin accounts. *Margin* is the word now used to describe the collateral posted in OTC markets as well as exchange-traded markets.

Variation margin is the collateral posted to reflect the change in the value of a derivatives portfolio. Consider the situation where Party A is trading with Party B and the collateral agreement states that variation margin (with no threshold or minimum transfer amount) has to be posted by both sides.[1] This means that, if the value of outstanding transactions changes during a day so that they increase in value by $\$X$ to A (and therefore decrease in value by $\$X$ to B), B has to provide A with $\$X$ of acceptable collateral. The cumulative effect of variation margin is that, if outstanding derivatives have a value of $+\$V$ to A and $-\$V$ to B at a particular time, B should have posted a total of $\$V$ of collateral with A by that time.[2]

Variation margin provides some protection against a counterparty default. It would provide total protection in an ideal world where (a) the counterparty never owes any variation margin at the time of default and (b) all outstanding positions can be replaced

[1] A and B could be two derivatives dealers or a derivatives dealer and one of its clients. Also, one of A and B could be a CCP. A threshold is a minimum value of the portfolio to one side before it can demand margin, and the minimum transfer amount is the minimum change in value necessary for a margin to have to be posted.

[2] In this context, note that if A buys an option from B for \$10,000, it must pay \$10,000 to B, but B must then return the \$10,000 to A as variation margin.

at mid-market prices as soon as the counterparty defaults. In practice, defaulting counterparties often stop posting collateral several days before they default, and the non-defaulting counterparty is usually subject to a bid-offer spread as it replaces transactions.[3] To allow for adverse movements in the value of the portfolio during a period prior to default when no margin is being posted, market participants sometimes require *initial margin* in addition to variation margin. Note that, in this context, adverse market movements are increases in the value of the portfolio to the non-defaulting party, not decreases. This is because increases in the value during a period when variation margin is not being posted lead to increases in replacement costs.[4] Initial margin, which can change through time as the outstanding portfolio and relevant volatilities change, reflects the risk of a loss due to adverse market moves and the costs of replacing transactions.[5]

Most margin is cash, but the agreements in place may specify that securities can be posted instead of cash. The securities may be subject to a *haircut*. This means that the market value of the securities is reduced to determine their value for margin purposes. For example, a Treasury bond might be subject to a 10% haircut, indicating that, if its market value were $100, it would cover only $90 of a margin requirement.

Should cash margin earn interest? There is a difference between futures markets and OTC markets here. A futures exchange clearing house requires both initial margin and variation margin from members. Members earn interest on the initial margin. But they do not do so on variation margin because futures contracts are settled daily so that variation margin does not belong to the member posting it. In the case of OTC trades, interest is usually earned on all cash margin posted because trades are not settled daily.

17.1.2 Central Clearing

In central clearing, a central counterparty (CCP) handles the clearing. A CCP operates very much like an exchange clearing house. When two companies, A and B, agree to an over-the-counter derivatives transaction and decide to clear it centrally, they present it to a CCP. Assuming that the CCP accepts it, the CCP acts as an intermediary and enters into offsetting transactions with the two companies.

Suppose, for example, that the transaction is an interest rate swap where company A pays a fixed rate of 5% to company B on a principal of $100 million for five years and company B pays LIBOR to company A on the same principal for the same period of time. Two separate transactions are created. Company A has a transaction with the CCP

---

[3] As explained later, the non-defaulting counterparty is able to claim from the defaulting party the costs related to the bid-offer spread that it would incur in replacing the transaction.

[4] It may seem strange that a market participant would be worried about the value of its transactions increasing. But suppose a transaction with a defaulting counterparty is hedged with another transaction entered into with another counterparty (as is often the case). The transaction with the other party can be expected to lose value without any compensating gain on the defaulted transaction.

[5] As indicated earlier, the non-defaulting party is allowed to keep all margin posted by the defaulting party up to the amount that can be legitimately claimed.

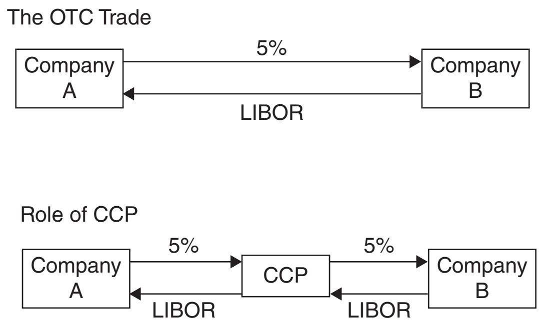


Figure 17.2 Role of CCP in OTC Markets

where it pays 5% and receives LIBOR on $100 million. Company B has a transaction with the CCP where it pays LIBOR and receives 5% on $100 million. The two companies no longer have credit exposure to each other. This is illustrated in Figure 17.2. If one or both parties to the transaction are not members of the CCP, they can clear the transaction through members.

Three large CCPs are

1. SwapClear (part of LCH Clearnet in London),
2. ClearPort (part of the CME Group in Chicago), and
3. ICE Clear Credit (part of the Intercontinental Exchange).

A CCP requires its members to provide initial margin and variation margin for the transactions being cleared. Typically, the initial margin is calculated so that there is a 99% probability that it will cover market moves over five days. This protects the CCP from losses as it tries to close out or replace the positions of a defaulting member.

Consider the swap in Figure 17.2. Suppose for simplicity that it is the only transaction each side has with the CCP. The CCP might require an initial margin of $0.5 million from each side. If, on the first day, interest rates fall so that the value of the swap to A goes down by $100,000, Party A would be required to pay a variation margin equal to this to the CCP, and the CCP would be required to pay the same amount to B. There could also be a change to the initial margin requirements determined by the CCP. If required margin is not paid by one of its members, the CCP closes out its transactions with that member. Cash and Treasury instruments are usually accepted as margin by CCPs. Typically the interest rate paid on cash balances is close to the overnight federal funds rate for U.S. dollars (and close to similar overnight rates for other currencies).

In practice, market participants are likely to have multiple transactions outstanding with the CCP at any given time. The initial margin required from a participant at any given time reflects the volatility of the value of its total position with the CCP. The role of a CCP in the OTC market is similar to the role of a clearing house in the exchange-traded market. The main difference is that transactions handled by the CCP are usually less standard than transactions in the exchange-traded market so that the calculation of margin requirements is more complicated.

The key advantage of clearing a transaction through a CCP is that OTC market participants do not need to worry about the creditworthiness of the counterparties they trade with. Credit risk is handled by the CCP using initial and variation margin.

A CCP requires its members to contribute to a default fund. (As mentioned, if one or both parties to a transaction are not members of the CCP, they can clear the transaction through members. They will then have to post margin with the members.) If a member fails to post margin when required, the member is in default and its positions are closed out. In closing out a member's positions, the CCP may incur a loss. A waterfall defines who bears the loss. The order in which the loss is funded is usually as follows:

1. The initial margin of the defaulting member
2. The default fund contribution of the member
3. The default fund contributions of other members
4. The equity of the CCP[6]

This is similar to the way losses in the event of a default are funded by an exchange clearing house.

17.1.3 Bilateral Clearing

In bilateral clearing, each pair of market participants enters into an agreement describing how all future transactions between them will be cleared. Typically this is an *ISDA master agreement*. (ISDA is short for International Swaps and Derivatives Association.) An annex to the agreement, known as the *credit support annex* (CSA), defines collateral arrangements. In particular, it defines what collateral (if any) has to be posted by each side, what assets are acceptable as collateral, what haircuts will be applied, and so on. The main body of the agreement defines what happens when one side defaults (e.g., by declaring bankruptcy, failing to make payments on the derivatives as they are due, or failing to post collateral when required). We will discuss this in more detail shortly.

17.1.4 Netting

We discussed netting in connection with the Basel I rules in Section 15.5. Netting is a feature of ISDA master agreements and a feature of the agreements between CCPs and their members. It states that all transactions between two parties are considered to be a single transaction when (a) collateral requirements are being calculated and (b) early terminations occur because of a default. As explained in Section 15.5, netting reduces credit risk because it means that the defaulting party cannot choose to default on transactions that are out-of-the-money while keeping transactions that are in-the-money.

[6] In some cases, the non-defaulting members are required to provide additional default fund contributions when there is a default, with a cap on the amount of these additional contributions. (This is true of both exchange clearing houses and CCPs.)

Netting can also save initial margin. Suppose Party A has two transactions with a CCP that are not perfectly correlated. The initial margin for the portfolio is likely to be less than that for the two transactions separately.

17.1.5 Events of Default

Derivatives transactions are treated differently from other transactions in the event that a market participant fails to meet its obligations. For example, in ISDA master agreements there is an early termination provision that takes precedence over bankruptcy rules. This states that, if there is an "event of default," the non-defaulting party has the right to terminate all transactions with the defaulting party after a short period of time has elapsed.[7] Events of default include declarations of bankruptcy, failure to make payments as they are due, and failure to post collateral when required.[8] Non-derivative contracts cannot always be terminated in this way. Another important difference between derivatives transactions and non-derivatives transactions is that in the case of derivatives transactions the non-defaulting party can take immediate possession of any collateral that has been posted by the defaulting party. It does not have to get a court order to allow it to do this.

If there is an event of default under an ISDA master agreement, the non-defaulting party calculates the mid-market value of outstanding transactions. It then adjusts this valuation in its favor by half the bid-offer spreads on the transactions for the purposes of calculating a settlement amount. This adjustment is compensation for the fact that it will have to trade with other dealers to replace the transactions and it will be subject to their bid-offer spreads when it does so. Suppose that one transaction has a mid-market value of \$20 million to the non-defaulting party and that the transaction is bid \$18 million, offer \$22 million. For the purposes of settlement, the transaction would be valued at \$22 million because this is what it would cost the non-defaulting party to replace the defaulting party's position in the transaction. If the non-defaulting party had the other side of the transaction so that its mid-market value were –\$20 million, it would be valued at –\$18 million for settlement purposes. In this case, the assumption is that a third party would be prepared to pay only \$18 million to take the defaulting party's position.

17.2 Post-Crisis Regulatory Changes

The OTC derivatives market was considered by many to have been partly responsible for the 2008 credit crisis. When the G20 leaders met in Pittsburgh in September 2009 in the aftermath of the 2008 crisis, they wanted to reduce systemic risk by regulating

---

[7] The non-defaulting party is not obliged to terminate transactions. Counterparties that are out-of-the-money sometimes consider that it is in their best interests not to terminate.

[8] Failure resolution mechanisms have been proposed where transactions are stayed (i.e., not terminated) for a period of time even if there is a bankruptcy filing, provided margin/collateral continues to be posted. These would allow the derivatives portfolios of bankrupt market participants to be unwound in an orderly way.

the OTC market. The statement issued by the leaders after the meeting included the following paragraph:

> *All standardized OTC derivative contracts should be traded on exchanges or electronic trading platforms, where appropriate, and cleared through central counterparties by end-2012 at the latest. OTC derivative contracts should be reported to trade repositories. Non-centrally cleared contracts should be subject to higher capital requirements. We ask the FSB and its relevant members to assess regularly implementation and whether it is sufficient to improve transparency in the derivatives markets, mitigate systemic risk, and protect against market abuse.*

The results of this were three major changes affecting OTC derivatives:

1. A requirement that all standardized OTC derivatives be cleared through CCPs. Standardized derivatives include plain vanilla interest rate swaps (which account for the majority of OTC derivatives traded) and default swaps on credit indices. The purpose of this requirement is to reduce systemic risk (see Business Snapshot 15.1). It leads to derivatives dealers having less credit exposure to each other so that their interconnectedness is less likely to lead to a collapse of the financial system.
2. A requirement that standardized OTC derivatives be traded on electronic platforms. This is to improve transparency. The thinking is that, if there is an electronic platform for matching buyers and sellers, the prices at which products trade should be readily available to all market participants.[9] The platforms are called *swap execution facilities* (SEFs) in the United States and *organized trading facilities* (OTFs) in Europe. In practice, standardized products once they have been traded on these platforms are passed automatically to a CCP.
3. A requirement that all trades in the OTC market be reported to a central trade repository. This requirement provides regulators with important information on the risks being taken by participants in the OTC market. It is partly a response to the AIG fiasco where regulators were not aware of the huge risks being taken by a subsidiary of AIG until the insurance company asked to be bailed out.

The first two of these requirements apply only to transactions between two financial institutions (or between a financial institution and a non-financial company that is considered to be systemically important because of the volume of its OTC derivatives trading). Derivatives dealers can therefore continue to trade with many of their non-financial corporate clients in the same way that they did pre-crisis.

About 25% of OTC transactions were cleared through CCPs pre-crisis and the remaining 75% were cleared bilaterally. As a result of the new rules, these percentages have flipped so that approximately 75% of OTC transactions are now cleared through CCPs while 25% are cleared bilaterally.

[9] An issue here is that the type of electronic platform that is appropriate for swaps may not be the same as the one that is used by exchanges. Swaps are traded intermittently with large notional principals. Futures and options on an exchange trade continually and the size of trades is usually much smaller.

17.2.1 Uncleared Trades

Following another G20 meeting in 2011, the rules have been tightened for non-standard OTC derivatives. These are the derivatives that are not covered by the rules just mentioned. They are cleared bilaterally rather than centrally and are referred to as *uncleared trades*. Regulations, which are being implemented between 2016 and 2020, require uncleared trades between two financial institutions (or between a financial institution and a non-financial company that is considered to be systemically important) to be subject to rules on the margin that has to be posted. Previously, one of the attractions of bilateral clearing was that market participants were free to negotiate any credit support annex to their ISDA master agreements.

The rules state that both initial margin and variation margin must be posted for uncleared trades by both sides. Variation margin was fairly common in the OTC market pre-crisis (particularly in trades between derivatives dealers), but initial margin was rare. When entering into a transaction with a much less creditworthy counterparty, a derivatives dealer might insist on the counterparty posting initial margin. But the posting of initial margin by both sides was almost unheard of in the bilaterally cleared market.

Variation margin is usually transmitted directly from one counterparty to the other. Initial margin when posted by both sides cannot be handled in this way. If, for example, A transmitted \$1 million of initial margin to B and B transmitted \$1 million of initial margin to A, the initial margin would not serve the desired purpose because the transfers would cancel each other. For this reason the regulations require initial margin to be transmitted to a third party, where it is held in trust.

17.2.2 Determination of Initial Margin: SIMM

For the new rules on uncleared transactions to work, the two sides to an ISDA master agreement must agree on the variation margin and initial margin. The variation margin requires agreement on the valuation of outstanding transactions, and procedures have been established for resolving any disagreements on this. The calculation of initial margin is more complicated than valuing the transactions and there is more scope for different models to give different results. As a result there have been attempts to develop an industry standard.

Initial margin is specified in the regulations for portfolios of uncleared transactions between two parties as the gain in value over 10 days that we are 99% certain will not be exceeded in stressed market conditions. Note that initial margin is the mirror image of VaR. When we are calculating VaR, we are determining extreme percentiles of the loss distribution, but when we are calculating initial margin we are determining extreme percentiles of the gain distribution. This is because exposure increases as the uncollateralized value of a portfolio increases.

The Basel Committee proposed a grid approach for calculating initial margin, which specified initial margin as a percentage of notional principal for different types of transactions. This was unpopular because it did not incorporate netting. If a market participant entered into a certain transaction on Day 1 and an almost offsetting transaction on Day 5, both with the same counterparty, the initial margin on Day 5 would be almost double that on Day 1—even though the net exposure to the counterparty would be close to zero. ISDA proposed what is known as the Standard Initial Margin Model (SIMM) as a way of overcoming this. This model has now been approved by regulators.

SIMM uses the model-building approach described in Chapter 14.[10] Delta and vega risks are handled using the weighted sensitivities and risk weights as in equation (14.8) so that

$$\text{IM (Delta and Vega)} = \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n} \delta_i \delta_j \rho_{ij} W_i W_j}$$

where the W_i is the risk weight for risk factor i (specified by the regulators), δ_i is the sensitivity of the position held to risk factor i (determined by the bank), and ρ_{ij} is the correlation between risk factors i and j (specified by the regulators). Because a 10-day time horizon with 99% confidence is used, a possible formula for W_i is

$$W_i = \sqrt{10} \times N^{-1}(0.99)\sigma_i \tag{17.1}$$

where σ_i is the daily volatility (or standard deviation, in the case of interest rates, credit spreads, and volatilities) of the ith risk factor in stressed market conditions.

To calculate the incremental effect on initial margin of gamma risk, SIMM first considers the situation where all deltas are zero and there is no cross gamma. From equations (14.11) and (14.12), the mean and standard deviation of the change in the value of the portfolio over one day are:

$$E(\Delta P) = \frac{1}{2}\sum_i \gamma_i \sigma_i^2$$

$$SD(\Delta P) = \sqrt{\frac{1}{2}\sum_{i,j} \rho_{ij}^2 \gamma_i \gamma_j \sigma_i^2 \sigma_j^2}$$

where γ_i is the gamma with respect to the ith risk factor.

Estimates of the mean and standard deviation of portfolio change over 10 days are obtained by replacing σ_i with $\sqrt{10}\sigma_i$. Defining

$$C_i = \frac{1}{2}\gamma_i(\sqrt{10}\sigma_i)^2$$

the mean, m, and standard deviation, s, of the 10-day change are therefore given by

$$m = \sum_i C_i$$

$$s^2 = 2\sum_{i,j} \rho_{ij}^2 C_i C_j$$

[10] See "ISDA SIMM: From Principles to Model Specification," ISDA, March 3, 2016.

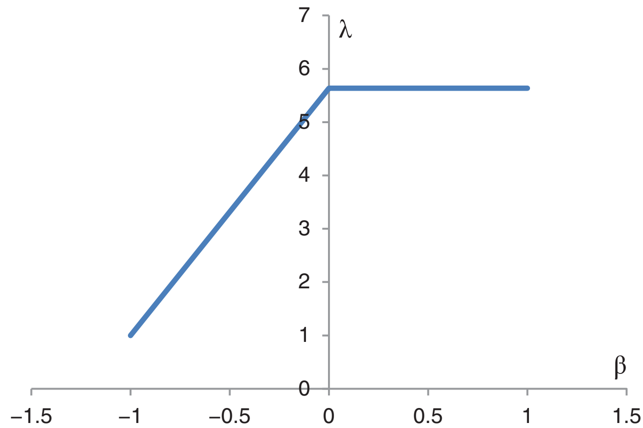


Figure 17.3 Relation between λ and β

SIMM then sets

$$\text{IM (Gamma)} = m + \lambda s/\sqrt{2}$$

The parameter λ in this equation is (see Problem 17.14) defined in terms of

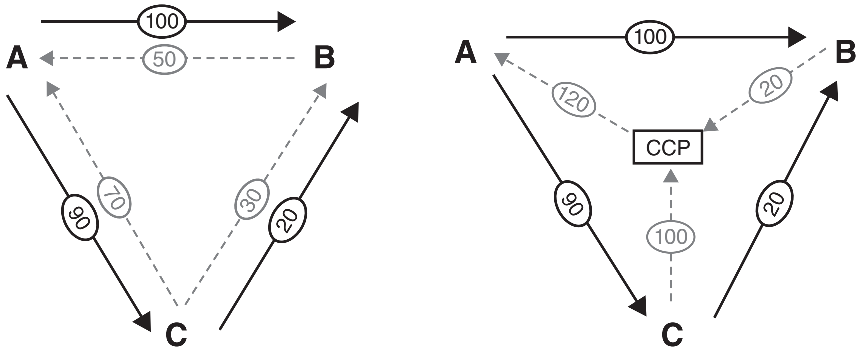
$$\beta = \frac{\sum_i C_i}{\sum_i |C_i|}$$

as indicated in Figure 17.3. This relationship produces results that have the right properties and correspond closely with tests carried out using Monte Carlo simulation.

There are a number of other details in SIMM. To simplify matters, gamma is calculated from vega using the relationship between the two that holds for European options. Risk factors are divided into buckets, and some risk factors involve term structures with vertices and are handled as we have described in Chapter 14. There are rules specified for calculating the correlations ρ_{ij} both within buckets and between buckets.

17.3 Impact of the Changes

The new regulations have led to a world where more collateral is required for OTC derivatives transactions. Pre-crisis, most OTC transactions were cleared bilaterally and an initial margin was usually not required. Under the new regulations, most transactions will be cleared through CCPs where both initial and variation margin will be required from both sides. Furthermore, transactions that are cleared bilaterally between financial institutions will require even more collateral than they would if they could be cleared through CCPs.



Dealer	Exposure after bilateral netting
A	0
B	100
C	20
Average	40

Dealer	Exposure after netting including CCP	Exposure after netting excluding CCP
A	120	0
B	120	120
C	90	90
Average	110	70

Figure 17.4 Example Where There Are Three Market Participants, One CCP, and Two Product Types
One product type (represented by dotted lines) can be cleared; the other (represented by solid lines) cannot.

As discussed by Duffie and Zhu, there is one potential partial offset to the huge increase in collateral requirements mandated by the new rules.[11] Under central clearing there is the potential for more netting. In Figure 17.1, under bilateral clearing a market participant has many different netting sets, one for each of the other market participants. Under central clearing, there is only one netting set. Bank A can, for example, net its transactions where Bank B is the counterparty with its transactions where Bank C is the counterparty, provided that all go through the same CCP.

Figure 17.1, however, is a simplification. It suggests that the choice is between a 100% bilateral world and a world where all transactions are cleared through a single CCP. The reality is that (a) there will be a number of CCPs and it is quite likely that they will not cooperate with each other to reduce initial margin requirements, and (b) some transactions will continue to be cleared bilaterally; so banks will face a situation that is a mixture of the two worlds depicted in Figure 17.1.

It is even possible that the new rules requiring the use of CCPs could reduce rather than increase netting in some cases. This is illustrated by Figure 17.4, which shows the situation where there are three market participants and one CCP. The exposures represented by the dotted lines are standard transactions that can be cleared centrally. Those represented by the solid lines are nonstandard transactions that cannot be cleared

[11]See D. Duffie and H. Zhu, "Does a Central Counterparty Reduce Counterparty Risk?" *Review of Asset Pricing Studies* 1 (2011): 74–95.

centrally. For example, in B's dealings with A, the nonstandard transactions are worth 100 to B and −100 to A; the standard transactions are worth +50 to A and −50 to B.

Without central clearing, the average exposure before collateral of the three parties is +40. With central clearing, the average exposure is 110 when the exposure to the CCP is included and 70 when it is not. Central clearing is likely to increase the collateral market participants have to post in this simple situation. This happens because without the central clearing rules standard transactions can be netted with nonstandard transactions, but with the central clearing rules this is no longer possible.

Most experts think that there will be an increase in netting, but the overall effect of the changes will be an increase in margin requirements. Pre-crisis, relatively few OTC derivatives attracted initial margin. Post-crisis, the vast majority of OTC derivatives will require initial margin. A related consideration is that, as more transactions are cleared through CCPs, more of the funds of a financial institution will be tied up in default fund contributions.

17.3.1 Liquidity

Most of the collateral required under the new regulations will have to be in the form of cash or government securities. An increasingly important consideration for all derivatives market participants is therefore liquidity. Not only will the collateral posted at any given time be a drain on liquidity, but banks will have to keep a sufficient quantity of liquid assets on hand to ensure that they are able to meet any margin calls. (Margin calls from a CCP have to be met almost immediately.) As we saw in Chapter 16, Basel III has recognized the importance of liquidity by proposing two new liquidity ratios that banks must adhere to. Capital has in the past been the key metric in determining the profitability of different business units and different projects at a bank. In the future, a two-dimensional metric involving capital and liquidity is likely to be used. Often there will be a trade-off between capital and liquidity in that a project will look attractive from a capital perspective and unattractive from a liquidity perspective, or vice versa.

17.3.2 Rehypothecation

Liquidity pressures are likely to increase because of another post-crisis change. What is known as "rehypothecation" was common in some jurisdictions (particularly the United Kingdom) pre-crisis. (See Business Snapshot 17.1.) It involved a dealer using collateral posted with it by one counterparty to satisfy a collateral demand by another counterparty. It is estimated that pre-crisis about \$4 trillion of collateral was required in derivatives markets, but that because of rehypothecation only \$1 trillion of new collateral was posted.[12] In other words, each item of collateral was used on average four times. Rehypothecation will be restricted under new rules developed by the Basel Committee and the

[12] See M. Singh and J. Aitken, "The (Sizable) Role of Rehypothecation in the Shadow Banking System," Working Paper, International Monetary Fund, 2010.

BUSINESS SNAPSHOT 17.1

Rehypothecation

A practice in the management of collateral known as rehypothecation can cause problems. If Party A posts collateral with Party B and rehypothecation is permitted, Party B can use the same collateral to satisfy a demand for collateral from Party C; Party C can then the use the collateral to satisfy a demand for collateral from Party D; and so on. In 2007, it was estimated that U.S. banks had more than \$4 trillion of collateral, but that this was created by using \$1 trillion of original collateral in conjunction with rehypothecation. Rehypothecation was particularly common in the United Kingdom, where title to collateral is transferred.

After Lehman declared bankruptcy in September 2008, clients (particularly European hedge fund clients) found it difficult to get a return of the collateral they had posted with Lehman because it had been rehypothecated. As a result of this experience, many market participants are more cautious than they used to be, and clauses in CSAs banning or limiting rehypothecation are now common.

International Organization of Securities Commissions (IOSCO). These rules allow initial margin to be rehypothecated once, but only if certain conditions are satisfied. Variation margin can be rehypothecated. But increasingly dealers themselves impose restrictions on rehypothecation because they do not want to be disadvantaged in the same way that some of Lehman's counterparties were (see Business Snapshot 17.1).

17.3.3 The Convergence of OTC and Exchange-Traded Markets

The developments we have been discussing are blurring the distinction between OTC derivatives and exchange-traded derivatives. Many OTC transactions are now traded on platforms similar to exchanges and cleared through organizations similar to exchange clearing houses. As time goes by, more OTC transactions are likely to be classified as "standard" so that the percentage of OTC transactions handled similarly to exchange-traded transactions will increase. What is more, even those OTC transactions between financial institutions that are cleared bilaterally may begin to look more like exchange-traded transactions. This is because margin has to be posted with a third party, and we can expect organizations (somewhat similar to exchange clearing houses) to be set up to facilitate this.

It is also the case that exchanges are increasingly trying to offer less standard products to institutional investors in an attempt to take business away from the OTC market. As a result, while OTC markets are moving in the direction of becoming more like exchange-traded markets, exchange-traded markets are moving in the opposite direction and becoming more like OTC markets. Many CCPs and exchanges have a common ownership and will find areas for cooperation on margin requirements and business practices. Whether a transaction is being cleared through an exchange or a CCP may not be important in the future because it will be handled in the same way by the same organization.

17.4 CCPs and Bankruptcy

The key objective of regulators is to reduce systemic risk. Some commentators have criticized the new derivatives regulations as replacing too-big-to-fail banks by too-big-to-fail CCPs. It certainly would be a disaster for the financial system if a major CCP such as LCH Clearnet's SwapClear and CME's ClearPort were to fail. In theory, as described in Hull (2012), it is possible to design the contract between CCPs and their members so that it is virtually impossible for a CCP to fail.[13] In practice, it is considered important that a CCP has "skin in the game." It is then motivated to take good decisions with respect to key issues such as whether a new member should be admitted, how initial margins should be set, and so on.

The main reason why it makes sense to replace too-big-to-fail banks by too-big-to-fail CCPs is that CCPs are much simpler organizations than banks. They are therefore much simpler to regulate than banks. In essence, regulators need ensure only that the CCP follows good practices in (a) choosing members, (b) valuing transactions, and (c) determining initial margins and default fund contributions. In the case of banks, a myriad of different, much more complex activities must be monitored. It is of course important for regulators to ensure that CCPs are not allowed to become more complex organizations by expanding outside their core activity of intermediating derivatives transactions.

Summary

Prior to the 2007–2008 credit crisis, the over-the-counter (OTC) derivatives market was largely unregulated. Two market participants could agree to any transaction they liked and then reach any agreement they liked on how the transaction would be cleared. They were also free to choose any arrangements they liked for the posting of collateral. This is no longer the case. The OTC derivatives market is now subject to a great deal of regulation throughout the world. The extent to which the OTC derivatives market should be blamed for the crisis is debatable, but post-crisis regulatory changes are having more effect on this market than on almost any other sector of the economy.

Most standard OTC derivatives between two financial institutions must be cleared through central counterparties. These are very similar to exchanges. They require initial margin and variation margin to be posted by both sides. Nonstandard transactions

[13]See J. Hull, "CCPs, Their Risks, and How They Can Be Reduced," *Journal of Derivatives* 20, no. 1 (Fall 2012): 26–29.

between financial institutions will continue to be cleared bilaterally, but are subject to regulation on the collateral that must be posted. Specifically, transactions between financial institutions are subject to initial margin (segregated) and variation margin (transferred from one side to the other when the value of outstanding transactions changes).

What will the derivatives world look like in 15 or 20 years? Present trends indicate that there will be a convergence between OTC and exchange-traded markets, and the distinction between the two will become blurred. But it should be acknowledged that there is no certainty that this trend will continue. The OTC market as it existed before the crisis was very profitable for a few large banks. It is possible that they will chip away at the regulations so that they are able eventually to find a way of creating a new OTC market somewhat similar to the one that existed before the crisis. A battle is likely to take place pitting the determination of regulators against the ingenuity of banks.

Further Reading

Basel Committee on Banking Supervision and IOSCO. "Margin Requirements for Non-Centrally Cleared Derivatives," September 2013.

Duffie, D., and H. Zhu. "Does a Central Counterparty Reduce Counterparty Risk?" *Review of Asset Pricing Studies* 1 (2011): 74–95.

Hull, J. "CCPs, Their Risks, and How They Can Be Reduced." *Journal of Derivatives* 20, no. 1 (Fall 2012): 26–29.

Hull, J. "The Changing Landscape for Derivatives." *Journal of Financial Engineering* 1, no. 2 (2014).

Hull, J. "OTC Derivatives and Central Clearing: Can All Transactions Be Cleared?" *Financial Stability Review* 14 (July 2010): 71–89.

Singh, M., and J. Aitken. "The (Sizable) Role of Rehypothecation in the Shadow Banking System." Working Paper, International Monetary Fund, 2010.

Practice Questions and Problems (Answers at End of Book)

17.1 What is the difference between using an ISDA master agreement and using a CCP for clearing OTC transactions?

17.2 Explain the new regulations introduced since the 2007–2008 crisis that (a) require certain transactions to be cleared using CCPs and (b) require extra collateral for some transactions cleared bilaterally.

17.3 Why might the regulations introduced since the 2007–2008 crisis create liquidity problems for some financial institutions?

17.4 What is meant by a haircut in a collateral agreement?

17.5 Explain the meaning of "event of default" and an "early termination" in an ISDA master agreement.

17.6 In Figure 17.4 where the CCP is used, suppose that half of the transactions between A and B that are represented by the solid line are moved to the CCP. What effect does this have on (a) the average exposure of the three parties including their exposures to the CCP and (b) the average exposure of the three parties excluding their exposures to the CCP?

17.7 What credit risks is a company taking when it becomes a member of a CCP and clears transactions through the CCP?

17.8 "Netting affects the collateral that has to be posted and the settlement in the event of an early termination." Explain.

17.9 What is rehypothecation?

17.10 Why are CCPs easier to regulate than banks?

17.11 What is claimed by the non-defaulting party in an early termination under an ISDA master agreement?

Further Questions

17.12 "Some aspects of the new regulations can be expected to increase the amount of collateral posted for derivatives, and some can be expected to reduce it." Explain this statement.

17.13 In Figure 17.4 where the CCP is used, suppose that an extra transaction between A and C that is worth 140 to A is cleared bilaterally. What effect does this have on the tables in Figure 17.4?

17.14 Consider a delta-neutral position in a single asset with a gamma (measured with respect to percentage changes in the asset) of γ ($\gamma > 0$). Suppose that the 10-day return on the asset is normally distributed with a mean of zero and a standard deviation σ.

(a) What is an expression for the change in the portfolio value, ΔP, over 10 days as a function of γ, σ, and a random sample from a standard normal distribution.

(b) The square of a standard normal has a chi-squared distribution with one degree of freedom. Write the expression for ΔP in (a) as a function of a random sample from such a chi-squared distribution and C, where $C = \gamma\sigma^2/2$.

(c) The 99 percentile of a chi-squared distribution with one degree of freedom is 6.63. (See CHISQ.INV in Excel.) Use this to show that the 99 percentile of ΔP is $6.63C$.

(d) The mean and variance of a chi-squared distribution with one degree of freedom are 1 and 2. Show that this result is consistent with Figure 17.3 and the formula for IM (Gamma) in Section 17.2.2.

(e) Show that when $\gamma < 0$ the 99 percentile of ΔP is less than zero so that Figure 17.3 then gives IM (Gamma) $= 0$.

Chapter 18

Fundamental Review of the Trading Book

In May 2012, the Basel Committee on Banking Supervision issued a consultative document proposing major revisions to the way regulatory capital for market risk is calculated. This is referred to as the "Fundamental Review of the Trading Book" (FRTB).[1] The Basel Committee then followed its usual process of requesting comments from banks, revising the proposals, and carrying out Quantitative Impact Studies (QISs).[2] The final version of the rules was published by the Basel Committee in January 2016.[3] This require banks to implement the new rules in 2019, but in December 2017 the implemented year was revised to 2022.

FRTB's approach to determining capital for market risk is much more complex than the approaches previously used by regulators. The purpose of this chapter is to outline its main features.

---

[1] See Bank for International Settlements, "Consultative Document: Fundamental Review of the Trading Book," May 2012.

[2] QISs are calculations carried out by banks to estimate the impact of proposed regulatory changes on capital requirements.

[3] See Bank for International Settlements, "Minimum Capital Requirements for Market Risk," January 2016.

18.1 Background

The Basel I calculations of market risk capital were based on a value at risk (VaR) calculated for a 10-day horizon with a 99% confidence level. The VaR was "current" in the sense that calculations made on a particular day were based on the behavior of market variables during an immediately preceding period of time (typically, one to four years). Basel II.5 required banks to calculate a "stressed VaR" measure in addition to the current measure. As explained in Sections 13.1 and 16.1, this is VaR where calculations are based on the behavior of market variables during a 250-day period of stressed market conditions. To determine the stressed period, banks were required to go back through time searching for a 250-day period where the observed movements in market variables would lead to significant financial stress for the current portfolio.

FRTB changes the measure used for determining market risk capital. Instead of VaR with a 99% confidence level, it uses expected shortfall (ES) with a 97.5% confidence level. The measure is actually stressed ES with a 97.5% confidence. This means that, as in the case of stressed VaR, calculations are based on the way market variables have been observed to move during stressed market conditions.

For normal distributions, VaR with a 99% confidence and ES with a 97.5% confidence are almost exactly the same. Suppose losses have a normal distribution with a mean μ and standard deviation σ. The 99% VaR is $\mu + 2.326\sigma$ while the 97.5% expected shortfall is $\mu + 2.338\sigma$.[4] (See Problem 18.2.) For non-normal distributions, they are not equivalent. When the loss distribution has a heavier tail than a normal distribution, the 97.5% ES can be considerably greater than the 99% VaR.

Under FRTB, the 10-day time horizon used in Basel I and Basel II.5 is changed to reflect the liquidity of the market variable being considered. FRTB considers changes to market variables that would take place (in stressed market conditions) over periods of time reflecting their liquidity. The changes are referred to as *shocks*. The market variables are referred to as *risk factors*. The periods of time considered are referred to as *liquidity horizons*. Five different liquidity horizons are specified: 10 days, 20 days, 40 days, 60 days, and 120 days. The allocation of risk factors to these liquidity horizons is indicated in Table 18.1.

FRTB specifies both a standardized approach and an internal models approach for calculating market risk capital. Even when banks have been approved to use the internal models approach, they are required by regulators to calculate required capital under both approaches. This is consistent with the Basel Committee's plans to use standardized approaches to provide a floor for capital requirements. As discussed in Section 16.4, in December 2017, the Basel Committee announced a move to a situation where total required capital is at least 72.5% of that given by standardized approaches. It will achieve

[4] From equation (12.2), the ES for a normal distribution with mean μ and standard deviation σ is $\mu + \sigma \exp(-Y^2/2)/[\sqrt{2\pi}(1 - X)]$ where X is the confidence level and Y is the point on a normal distribution that has a probability of $1 - X$ of being exceeded. This can also be written $\mu + \sigma^2 f(\text{VaR})/(1 - X)$ where f is the probability density function for the loss.

Table 18.1 Allocation of Risk Factors to Liquidity Horizons

Risk Factor	Horizon (days)
Interest rate (dependent on currency)	10–60
Interest rate volatility	60
Credit spread: sovereign, investment grade	20
Credit spread: sovereign, non-investment grade	40
Credit spread: corporate, investment grade	40
Credit spread: corporate, non-investment grade	60
Credit spread: other	120
Credit spread volatility	120
Equity price: large cap	10
Equity price: small cap	20
Equity price: large cap volatility	20
Equity price: small cap volatility	60
Equity: other	60
Foreign exchange rate (dependent on currency)	10–40
Foreign exchange volatility	40
Energy price	20
Precious metal price	20
Other commodities price	60
Energy price volatility	60
Precious metal volatility	60
Other commodities price volatility	120
Commodity (other)	120

this by 2027 with a five-year phase-in period. These changes are a culmination of a trend by the Basel Committee since the 2008 crisis to place less reliance on internal models and to use standardized models to provide a floor for capital requirements.

A difference between FRTB and previous market risk regulatory requirements is that most calculations are carried out at the trading desk level. Furthermore, permission to use the internal models approach is granted on a desk-by-desk basis. Therefore it is possible that, at a particular point in time, a bank's foreign currency trading desk has permission to use the internal models approach while the equity trading desk does not.

In earlier chapters, we saw how the ways in which capital is calculated for the trading book and the banking book are quite different. This potentially gives rise to regulatory arbitrage where banks choose to allocate instruments to either the trading book or the banking book so as to minimize capital. In Basel II.5, the incremental risk charge made this less attractive. FRTB counteracts regulatory arbitrage by defining more clearly than previously the differences between the two books.

18.2 Standardized Approach

Under the standardized approach, the capital requirement is the sum of three components: a risk charge calculated using a risk sensitivity approach, a default risk charge, and a residual risk add-on.

Consider the first component. Seven risk classes (corresponding to trading desks) are defined (general interest rate risk, foreign exchange risk, commodity risk, equity risk, and three categories of credit spread risk). Within each risk class, a delta risk charge, vega risk charge, and curvature risk charge are calculated.

The delta risk charge for a risk class is calculated using the risk weights and weighted sensitivity approach described in Section 14.6:

$$\text{Risk Charge} = \sum_i \sum_j \rho_{ij}\delta_i\delta_j W_i W_j \tag{18.1}$$

where the summations are taken over all risk factors in the risk class. The risk weights, W_i, and the correlations between risk factors, ρ_{ij}, are determined by the Basel Committee.[5] The weighted sensitivities (or deltas), δ_i, are determined by the bank. As explained in Chapter 14, in the case of risk factors such as equity prices, exchange rates, or commodity prices, the deltas measure the sensitivity of the portfolio to percentage changes. For example, if a 1% increase in a commodity price would increase the value of a portfolio by \$3,000, the delta would be 3,000/0.01 = 300,000. In the case of risk factors such as interest rates and credit spreads, the deltas are defined in terms of absolute changes. For example, if the effect of an interest rate increasing by one basis point (0.0001) is to reduce the value of a portfolio by \$200, the delta with respect to that interest rate would be $-200/0.0001 = -2{,}000{,}000$.

Consider how the risk weights, W_i, might be set by regulators. Suppose first that all risk factors are equity prices, exchange rates, or commodity prices, so the deltas are sensitivities to percentage changes. If W_i were set equal to the daily volatility of risk factor i for all i, the risk charge in equation (18.1) would equal the standard deviation of change in the value of the portfolio per day. This corresponds to the Markowitz result discussed in Chapters 1 and 14. If W_i were set equal to the daily volatility of risk factor i in stressed market conditions (the stressed daily volatility) for all i, equation (18.1) would give the standard deviation of the daily change of the portfolio in stressed market conditions. In practice, the W_i are set equal to multiples of the stressed daily volatility to reflect the liquidity horizon and the confidence level that regulators wish to consider. Suppose that the stressed daily volatility of risk factor i is estimated as 2% and that the risk factor has a 20-day liquidity horizon. The risk weight might be set as $0.02 \times \sqrt{20} \times 2.338 = 0.209$. (Note that the 2.338 multiplier reflects the amount by which a standard deviation has to be multiplied to get ES with a 97.5% confidence when a normal distribution is assumed.)

Now suppose that the risk factors are interest rates and credit spreads so that deltas are sensitivities with respect to actual changes measured in basis points. The W_i for risk factor i is set equal to a multiple of the stressed daily standard deviation for all i. If the multiple were 1, the formula would give the standard deviation of the value of the portfolio in

[5] Banks are required to test the effect of multiplying the correlations specified by the Basel Committee by 1.25, 1.00, and 0.75 and then set the capital charge equal to the greatest result obtained.

one day. In practice the multiple is determined as just described to reflect the liquidity horizon and confidence level.

Vega risk is handled similarly to delta risk.[6] A vega risk charge is calculated for each risk class using equation (18.1). The risk factors (counted by the i and j) are now volatilities. The summation is taken over all volatilities in the risk class. The parameter δ_i is actually a vega. It is the sensitivity of the value of the portfolio to small changes in volatility i.[7] The parameter ρ_{ij} is the correlation between changes in volatility i and volatility j, and W_i is the risk weight for volatility i. The latter is determined similarly to the delta risk weights to reflect the volatility of the volatility i, its liquidity horizon, and the confidence level.

There are assumed to be no diversification benefits between risk factors in different risk classes and between the vega risks and delta risks within a risk class. The end product of the calculations we have described so far is therefore the sum of the delta risk charges across the seven risk classes plus the sum of the vega risk charges across the seven risk classes.

18.2.1 Term Structures

In the case of risk factors such as interest rates, volatilities, and credit spreads, there is usually a term structure defined by a number of points. For example, an interest rate term structure is typically defined by 10 points. These are the zero-coupon interest rates for maturities of 3 months, 6 months, 1 year, 2 years, 3 years, 5 years, 10 years, 15 years, 20 years, and 30 years. As discussed in Chapter 14, each vertex of the term structure is a separate risk factor for the purposes of using equation (18.1). The delta of a portfolio with respect to a one basis point move in one of the vertices on the term structure is calculated by increasing the position of the vertex by one basis point while making no change to the other vertices. The Basel Committee defines risk weights for each vertex of the term structure and correlations between the vertices of the same term structure.

A simplification is used when correlations between points on different term structures are defined. The correlations between point A on term structure 1 and point B on term structure 2 are assumed to be the same for all A and B. This was one of the alternatives considered in Chapter 14; see equation (14.6).

18.2.2 Curvature Risk Charge

The curvature risk charge is a capital charge for a bank's gamma risk exposure under the standardized approach. Consider the exposure of a portfolio to the ith risk factor. Banks are required to test the effect of increasing and decreasing the risk factor by its risk weight, W_i. If the portfolio is linearly dependent on the risk factor, the impact of

[6] As discussed in Chapter 14, this works well because most of the value of a derivative is in many cases approximately linearly dependent on volatility.

[7] Banks can choose whether it is percentage or actual changes in volatility that are considered.

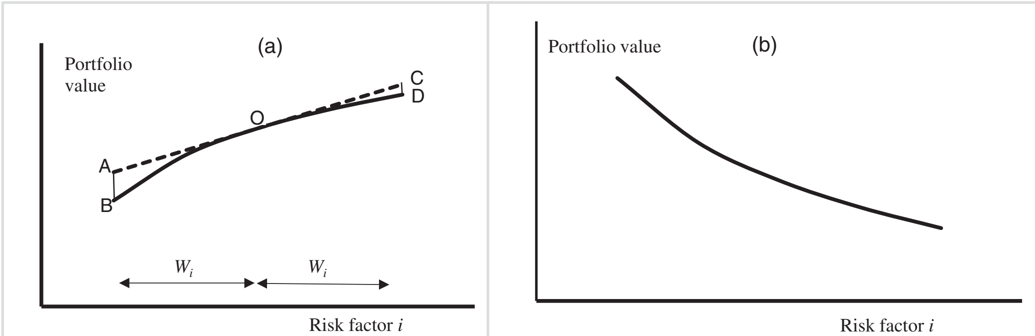


Figure 18.1 Calculation of Curvature Risk Charge for a Risk Factor
In Figure 18.1a, the curvature risk charge is AB; in Figure 18.1b, it is zero.

an increase of W_i in the risk factor is $W_i\delta_i$. Similarly, the impact of a decrease of W_i in the risk factor is $-\delta_i W_i$. To evaluate the impact of curvature net of the delta effect, the standardized approach therefore calculates

1. $W_i\delta_i$ minus the impact of a increase of W_i in the risk factor, and
2. $-W_i\delta_i$ minus the impact of a decrease in the risk factor of W_i.

The curvature risk charge for the risk factor is the greater of these two. If the impact of curvature net of delta is negative, it is counted as zero. The calculation is illustrated in Figure 18.1. In Figure 18.1a, the portfolio value is currently given by point O. If there were no curvature, an increase of W_i in the risk factor would lead to the portfolio value at point C, whereas a decrease of W_i in the risk factor would lead to the portfolio value at point A. Because of curvature, an increase of W_i leads to the portfolio value at point D, and a decrease of W_i leads to the portfolio value at point B. Since AB > CD, the risk charge is AB. In Figure 18.1b, the risk charge is zero because curvature actually increases the value of the position (relative to what delta would suggest) for both increases and decreases in the risk factor. (Figure 18.1a could correspond to a short position in an option; Figure 18.1b could correspond to a long position in an option.)

When there are several risk factors, each is handled similarly to Figure 18.1. When there is a term structure (e.g., for interest rates, credit spreads, and volatilities), all points are shifted by the same amount for the purpose of calculating the effect of curvature. The shift is the largest W_i for the points on the term structure. In the case of an interest rate term structure, the W_i corresponding to the three-month vertex is often the largest W_i, so this would define an upward and downward parallel shift in the term structure. The delta effect is removed for each point on the term structure by using the δ_i for that point.

The curvature risk charges for different risk factors are combined to determine a total curvature risk charge. When diversification benefits are allowed, aggregation formulas broadly similar to those used for deltas are used with correlations specified by the Basel Committee.

18.2.3 Default Risk Charge

Risks associated with counterparty credit spread changes are handled separately from risks associated with counterparty defaults in FRTB. In the standardized approach, credit spread risks are handled using the delta/vega/curvature approach described earlier. Default risks, sometimes referred to as *jump-to-default* (JTD) risks, are handled by a separate default risk charge. This is calculated by multiplying each exposure by a loss given default (LGD) and a default risk weight. Both the LGD and the risk weight are specified by the Basel Committee. For example, the LGD for senior debt is specified as 75% and the default risk for a counterparty rated A is 3%. Equity positions are subject to a default risk charge with an LGD = 100%. Rules for offsetting exposures are specified.

18.2.4 Residual Risk Add-On

The residual risk add-on considers risks that cannot be handled by the delta/vega/curvature approach described earlier. It includes exotic options when they cannot be considered as linear combinations of plain vanilla options. The add-on is calculated by multiplying the notional amount of the transaction by a risk weight that is specified by the Basel Committee. In the case of exotic options the risk weight is 1%.

18.2.5 A Simplified Approach

In this section, we have described the standardized approach that the Basel Committee requires all large banks to use. It is worth noting that in June 2017 the Basel Committee published a consultative document outlining a simplified standardized approach that it proposes for smaller banks.[8] The full approach is simplified in a number of ways. For example, vega and gamma risk do not have to be considered. This should make FRTB more attractive to jurisdictions such as the United States that have many small banks that tend to enter into only relatively simple transactions.

18.3 Internal Models Approach

The internal models approach requires banks to estimate stressed ES with a 97.5% confidence. FRTB does not prescribe a particular method for doing this. Typically the historical simulation approach described in Chapter 13 is likely to be used. Risk factors are allocated to liquidity horizons as indicated in Table 18.1. Define:

Category 1 Risk Factors: Risk factors with a time horizon of 10 days
Category 2 Risk Factors: Risk factors with a time horizon of 20 days
Category 3 Risk Factors: Risk factors with a time horizon of 40 days

[8] See Basel Committee on Banking Supervision, "Simplified Alternative to the Standardized Approach to Market Risk Capital Requirements," June 2017.

Category 4 Risk Factors: Risk factors with a time horizon of 60 days
Category 5 Risk Factors: Risk factors with a time horizon of 120 days

As we shall see, all calculations are based on considering 10-day changes in the risk factors. In Basel I and Basel II.5, banks are allowed to deduce the impact of 10-day changes from the impact of one-day changes using a $\sqrt{10}$ multiplier. In FRTB, banks are required to consider changes over periods of 10 days that occurred during a stressed period in the past. Econometricians naturally prefer that non-overlapping periods be used when VaR or ES is being estimated using historical simulation, because they want observations on the losses to be independent. However, this is not feasible when 10-day changes are considered, because it would require a very long historical period. FRTB requires banks to base their estimates on overlapping 10-day periods. The first simulation trial assumes that the percentage changes in all risk factors over the next 10 days will be the same as their changes between Day 0 and Day 10 of the stressed period; the second simulation trial assumes that the percentage changes in all risk factors over the next 10 days will be the same as their changes between Day 1 and Day 11 of the stressed period; and so on.

Banks are first required to calculate ES when 10-day changes are made to all risk factors. (We will denote this by ES_1.) They are then required to calculate ES when 10-day changes are made to all risk factors in categories 2 and above with risk factors in category 1 being kept constant. (We will denote this by ES_2.) They are then required to calculate ES when 10-day changes are made to all risk factors in categories 3, 4, and 5 with risk factors in categories 1 and 2 being kept constant. (We will denote this by ES_3.) They are then required to calculate ES when 10-day changes are made to all risk factors in categories 4 and 5 with risk factors in categories 1, 2, and 3 being kept constant. (We will denote this by ES_4.) Finally, they are required to calculate ES_5, which is the effect of making 10-day changes only to category 5 risk factors.

The liquidity-adjusted ES is calculated as

$$\sqrt{\mathrm{ES}_1{}^2 + \sum_{j=2}^{5}\left(\mathrm{ES}_j\sqrt{\frac{\mathrm{LH}_j - \mathrm{LH}_{j-1}}{10}}\right)^2} \tag{18.2}$$

where LH_j is the liquidity horizon for category j. To understand equation (18.2), suppose first that all risk factors are in category 1 or 2 so that only ES_1 and ES_2 are calculated. It is assumed that the behavior of all risk factors during a 10-day period is independent of the behavior of category 2 risk factors during a further 10-day period. An extension of the square root rule discussed in Chapter 13 then leads to the liquidity-adjusted ES being

$$\sqrt{\mathrm{ES}_1^2 + \mathrm{ES}_2^2}$$

Now suppose that there are also category 3 risk factors. The expression $\sqrt{\mathrm{ES}_1^2 + \mathrm{ES}_2^2}$ would be correct if the category 3 risk factors had a 20-day instead of a 40-day liquidity horizon. We assume that the behavior of the category 3 risk factors over an additional 20 days is independent of the behavior of all the risk factors over the periods already considered. We also assume that the ES for the category 3 risk factors over 20 days is $\sqrt{2}$ times their ES over 10 days. This leads to a liquidity-adjusted ES of:

$$\sqrt{\mathrm{ES}_1^2 + \mathrm{ES}_2^2 + 2\mathrm{ES}_3^2}$$

Continuing in this way, we obtain equation (18.2). This is referred to as the *cascade approach* to calculating ES (and can be used for VaR as well).

Calculations are carried out for each desk. If there are six desks, this means the internal models approach, as we have described it so far, requires $5 \times 6 = 30$ ES calculations. As mentioned, the use of overlapping time periods is less than ideal because changes in successive historical simulation trials are not independent. This does not bias the results, but it reduces the effective sample size, making results more noisy than they would otherwise be.

FRTB represents a movement away from basing calculations on one-day changes. Presumably the Basel Committee has decided that, in spite of the lack of independence of observations, a measure calculated from 10-day changes provides more relevant information than a measure calculated from one-day changes. This could be the case if changes on successive days are not independent, but changes in successive 10-day periods can reasonably be assumed to be independent.

The calculation of a stressed measure (VaR or ES) requires banks to search for the period in the past when market variable changes would be worst for their current portfolio. (The search must go back as far as 2007.) When Basel II.5 was implemented, a problem was encountered in that banks found that historical data were not available for some of their current risk factors. It was therefore not possible to know how these risk factors would have behaved during the 250-day periods in the past that were candidates for the reference stressed period. FRTB handles this by allowing the search for stressed periods to involve a subset of risk factors, provided that at least 75% of the current risk factors are used. The expected shortfalls that are calculated are scaled up by the ratio of ES for the most recent 12 months using all risk factors to ES for the most recent 12 months using the subset of risk factors. (This potentially doubles the number of ES calculations from 30 to 60.)

Banks are required to calculate ES for the whole portfolio as well for each of six trading desks. The ES for a trading desk is referred to as a *partial expected shortfall*. It is determined by shocking the risk factors belonging to the trading desk while keeping all other risk factors fixed. The sum of the partial expected shortfalls is always greater than the ES for the whole portfolio. What we will refer to as the weighted expected shortfall (WES) is a weighted average of (a) the ES for the whole portfolio and (b) the sum of the partial expected shortfalls. Specifically:

$$\mathrm{WES} = \lambda \times \mathrm{EST} + (1 - \lambda) \times \sum_j \mathrm{ESP}_j$$

where EST is the expected shortfall calculated for the total portfolio and ESP_j is the jth partial expected shortfall. The parameter λ is set by the Basel Committee to be 0.5.

Some risk factors are categorized as *non-modelable*. Specifically, if there are less than 24 observations on a risk factor in a year or more than one month between successive observations, the risk factor is classified as non-modelable. Such risk factors are handled by special rules involving stress tests.

The total capital requirement for day t is

$$\max(\text{WES}_{t-1} + \text{NMC}_{t-1}, m_c \times \text{WES}_{\text{avg}} + \text{NMC}_{\text{avg}})$$

where WES_{t-1} is the WES for day $t-1$, NMC_{t-1} is the capital charge calculated for non-modelable risk factors on day $t-1$, WES_{avg} is the average WES for the previous 60 days, and NMC_{avg} is the average capital charge calculated for the non-modelable risk factors over the previous 60 days. The parameter m_c is at minimum 1.5.

18.3.1 Back-Testing

FRTB does not back-test the stressed ES measures that are used to calculate capital under the internal models approach for two reasons. First, it is more difficult to back-test ES than VaR. Second, it is not possible to back-test a stressed measure at all. The stressed data upon which a stressed measure is based are extreme data that statistically speaking are not expected to be observed with the same frequency in the future as they were during the stressed period.

FRTB back-tests a bank's models by asking each trading desk to back-test a VaR measure calculated over a one-day horizon and the most recent 12 months of data. Both 99% and 97.5% confidence levels are to be used. If there are more than 12 exceptions for the 99% VaR or more than 30 exceptions for the 97.5% VaR, the trading desk is required to calculate capital using the standardized approach until neither of these two conditions continues to exist.

Banks may be asked by regulators to carry out other back-tests. Some of these could involve calculating the p-value of the profit or loss on each day. This is the probability of observing a profit that is less than the actual profit or a loss that is greater than the actual loss. If the model is working perfectly, the p-values obtained should be uniformly distributed.

18.3.2 Profit and Loss Attribution

Another test used by the regulators is known as *profit and loss attribution*. Banks are required to compare the actual profit or loss in a day with that predicted by their models. Two measures must be calculated. The measures are:

$$\frac{\text{Mean of } U}{\text{Standard Deviation of } V}$$

$$\frac{\text{Variance of } U}{\text{Variance of } V}$$

where U denotes the difference between the actual and model profit/loss in a day and V denotes the actual profit/loss in a day.[9] Regulators expect the first measure to be between –10% and +10% and the second measure to be less than 20%. When there are four or more situations in a 12-month period where the ratios are outside these ranges, the desk must use the standardized approach for determining capital.

18.3.3 Credit Risk

As mentioned, FRTB distinguishes two types of credit risk exposure to a company:

1. *Credit spread risk* is the risk that the company's credit spread will change, causing the mark-to-market value of the instrument to change.
2. *Jump-to-default risk* is the risk that there will be a default by the company.

Under the internal models approach, the credit spread risk is handled in a similar way to other market risks. Table 18.1 shows that the liquidity horizon for credit spread varies from 20 to 120 days and the liquidity horizon for a credit spread volatility is 120 days. The jump-to-default risk is handled in the same way as default risks in the banking book. In the internal models approach, the capital charge is based on a VaR calculation with a one-year time horizon and a 99.9% confidence level.

18.3.4 Securitizations

The comprehensive risk measure (CRM) charge was introduced in Basel II.5 to cover the risks in products created by securitizations such as asset-backed securities and collateralized debt obligations (see Section 16.1). The CRM rules allow a bank (with regulatory approval) to use its own models. The Basel Committee has concluded that this is unsatisfactory because there is too much variation in the capital charges calculated by different banks for the same portfolio. It has therefore decided that under FRTB the standardized approach must be used for securitizations.

18.4 Trading Book vs. Banking Book

The FRTB addresses whether instruments should be put in the trading book or the banking book. Roughly speaking, the trading book consists of instruments that the bank intends to trade. The banking book consists of instruments that are expected to be held to maturity. Instruments in the banking book are subject to credit risk capital whereas those in the trading book are subject to market risk capital. The two sorts of capital are calculated in quite different ways. This has in the past given rise to regulatory arbitrage. For example, as we have mentioned in earlier chapters, banks have often chosen to hold

[9] The "actual" profit/loss should be the profit and loss that would occur if there had been no trading in a day. This is sometimes referred to as the *hypothetical profit and loss*.

credit-dependent instruments in the trading book because they are then subject to less regulatory capital than they would be if they had been placed in the banking book.

The FRTB attempts to make the distinction between the trading book and the banking book clearer and less subjective. To be in the trading book, it will no longer be sufficient for a bank to have an "intent to trade." It must be able to trade and manage the underlying risks on a trading desk. The day-to-day changes in value should affect equity and pose risks to solvency. The FRTB provides rules for determining for different types of instruments whether they should be placed in the trading book or the banking book.

An important point is that instruments are assigned to the banking book or the trading book when they are initiated and there are strict rules preventing them from being subsequently moved between the two books. Transfers from one book to another can happen only in extraordinary circumstances. (Examples given of extraordinary circumstances are the closing of trading desks and a change in accounting standards with regard to the recognition of fair value.) Any capital benefit as a result of moving items between the books will be disallowed.

Summary

FRTB is a major change to the way capital is calculated for market risk. After 20 years of using VaR with a 10-day time horizon and 99% confidence to determine market risk capital, regulators are switching to using ES with a 97.5% confidence level and varying time horizons. The time horizons, which can be as high as 120 days, are designed to incorporate liquidity considerations into the capital calculations. The change that is considered to a risk factor when capital is calculated reflects movements in the risk factor over a period of time equal to the liquidity horizon in stressed market conditions.

The Basel Committee has specified a standardized approach and an internal models approach. Even when they have been approved by their supervisors to use the internal models approach, banks must also implement the standardized approach. Regulatory capital under the standardized approach is based on formulas involving the delta, vega, and gamma exposures of the trading book. Regulatory capital under the internal models approach is based on the calculation of stressed expected shortfall. Calculations are carried out separately for each trading desk.

Further Reading

Bank for International Settlements. "Minimum Capital Requirements for Market Risk," January 2016.

Practice Questions and Problems (Answers at End of Book)

18.1 Outline the differences between the way market risk capital is calculated in (a) Basel I, (b) Basel II.5, and (c) the FRTB.

18.2 Use footnote 4 to verify that when losses have a normal distribution with mean μ and standard deviation σ the 97.5% expected shortfall is $\mu + 2.338\sigma$.

18.3 Explain why the use of overlapping time periods proposed by the FRTB does not give independent observations on the changes in variables.

18.4 What are the advantages of expected shortfall over value at risk?

18.5 What is the difference between the trading book and the banking book? Why are regulators concerned about specifying whether an instrument should be one or the other in the FRTB?

18.6 How are credit trades handled under the FRTB?

Further Question

18.7 Suppose that an investor owns the \$10 million portfolio in Table 13.1 on September 30, 2014. The values of the DJIA, FTSE 100, CAC 40, and Nikkei 225 on that day were 17,042.90, 6,622.7, 4,416.24, and 16,173.52. The exchange rates on that day were 1.6211 USD per GBP, 0.7917 EUR per USD, and 109.64 JPY per USD. Suppose that the 250 days ending September 9, 2008, constitute the stressed period for the portfolio. Calculate the 97.5% expected shortfall using the overlapping periods method in conjunction with historical simulation and the cascade approach. Relevant data on the indices is on the author's website (see "Worksheets for Value at Risk Example"). For the purposes of this problem, assume that DJIA and FTSE have a 10-day liquidity horizon, CAC 40 has a 40-day liquidity horizon, and Nikkei 225 has a 20-day liquidity horizon. For each day during the stressed period, consider the change in a variable over a 10-day period ending on the day.

Part Four

CREDIT RISK

Chapter 19

Estimating Default Probabilities

As explained in Chapter 15, regulators have for a long time required banks to keep capital for credit risk. Under Basel II banks can, with approval from bank supervisors, use their own estimates of default probabilities to determine the amount of capital they are required to keep. This has led banks to search for better ways of estimating these probabilities.

In this chapter, we discuss a number of different approaches to estimating default probabilities and explain the key difference between risk-neutral and real-world estimates. The material we cover will be used in Chapter 20 when we examine how the price of a derivative in the bilaterally cleared over-the-counter market can be adjusted for counterparty credit risk, and in Chapter 21 when we discuss the calculation of credit value at risk.

19.1 Credit Ratings

As explained in Section 1.7, rating agencies such as Moody's, S&P, and Fitch provide ratings describing the creditworthiness of corporate bonds.[1] Because a credit rating is

[1] In theory, a credit rating is an attribute of a bond issue, not a company. However, in many cases all bonds issued by a company have the same rating. A rating is therefore often referred to as an attribute of a company.

designed to provide information about credit quality, one might expect frequent changes in credit ratings as positive and negative information reaches the market. In fact, ratings change relatively infrequently. One of the objectives of rating agencies when they assign ratings is stability. For example, they want to avoid ratings reversals, where a company's bonds are downgraded and then upgraded a few weeks later. Ratings therefore change only when there is reason to believe that a long-term change in the company's creditworthiness has taken place. The reason for this is that bond traders are major users of ratings. Often they are subject to rules governing what the credit ratings of the bonds they hold must be. (For example, many bond funds are allowed to hold only investment-grade bonds.) If these ratings changed frequently, they might have to do a large amount of trading (and incur high transaction costs) just to satisfy the rules.

A related point is that rating agencies try to "rate through the cycle." Suppose that the economy exhibits a downturn and this has the effect of increasing the probability of a company defaulting in the next six months, but makes very little difference to the company's probability of defaulting over the next three to five years. A rating agency would not usually change the company's credit rating in these circumstances.

Companies such as Moody's KMV and Kamakura provide estimates of a company's probability of default that are based on its equity price and other variables. These estimates do not have stability as one of their objectives and tend to respond more quickly to market information than credit ratings. The types of models that are used to produce the estimates will be discussed in Section 19.8.

19.1.1 Internal Credit Ratings

Most banks have procedures for rating the creditworthiness of their corporate and retail clients. This is a necessity. The ratings published by rating agencies are usually available only for companies that have issued publicly traded debt. As a result, many small and medium-sized companies do not have credit ratings from rating agencies. As explained in Chapter 15, the internal-ratings-based (IRB) approach in Basel II allows banks to use their internal ratings in determining the probability of default, PD.

Internal-ratings-based approaches for estimating PD typically involve profitability ratios such as return on assets and balance sheet ratios such as current assets divided by current liabilities (the current ratio) and debt to equity. Banks recognize that it is cash rather than profits that is necessary to repay a loan. They typically take the financial information provided by a company and convert it to a cash flow statement. This allows them to estimate how easy it will be for a company to service its debt.

19.1.2 Altman's Z-Score

Edward Altman has pioneered the use of accounting ratios to predict default. In 1968, he developed what has become known as the Z-score.[2] Using a statistical technique

[2] See E. I. Altman, "Financial Ratios, Discriminant Analysis, and the Prediction of Corporate Bankruptcy," *Journal of Finance* 23, no. 4 (September 1968): 589–609.

known as discriminant analysis, he attempted to predict defaults from five accounting ratios:

X_1: Working capital/Total assets
X_2: Retained earnings/Total assets
X_3: Earnings before interest and taxes/Total assets
X_4: Market value of equity/Book value of total liabilities
X_5: Sales/Total assets

For publicly traded manufacturing companies, the original Z-score was

$$Z = 1.2X_1 + 1.4X_2 + 3.3X_3 + 0.6X_4 + 0.999X_5 \tag{19.1}$$

If the Z-score was greater than 3.0, the company was considered unlikely to default. If it was between 2.7 and 3.0, there was reason to be "on alert." If it was between 1.8 and 2.7, there was a good chance of default. If it was less than 1.8, the probability of a financial embarrassment was considered to be very high. Equation (19.1) was estimated from a sample of 66 publicly traded manufacturing companies. Of these, 33 failed within one year and 33 did not fail within one year. The model proved to be fairly accurate when tested out of sample (i.e., on a set of firms different from that used to estimate equation 19.1). Both Type I errors (companies that were predicted not to go bankrupt but did do so) and Type II errors (companies that were predicted to go bankrupt, but did not do so) were small.[3] Variations on the model have been developed for manufacturing companies that are not publicly traded and for non-manufacturing companies.

Example 19.1

Consider a company for which working capital is 170,000, total assets are 670,000, earnings before interest and taxes is 60,000, sales are 2,200,000, the market value of equity is 380,000, total liabilities is 240,000, and retained earnings is 300,000. In this case, $X_1 = 0.254$, $X_2 = 0.448$, $X_3 = 0.0896$, $X_4 = 1.583$, and $X_5 = 3.284$. The Z-score is

$$1.2 \times 0.254 + 1.4 \times 0.448 + 3.3 \times 0.0896 + 0.6 \times 1.583 + 0.999 \times 3.284 = 5.46$$

The Z-score indicates that the company is not in danger of defaulting in the near future.

The Z-score methodology has been revised and extended since Altman's original research more than 45 years ago and can now be used to produce probabilities of default for applications such as Basel II.

[3] Type I errors are liable to be more costly to the lending department of a commercial bank than Type II errors.

19.2 Historical Default Probabilities

Table 19.1 is typical of the data that are produced by rating agencies. It shows the default experience through time of companies that started with a certain credit rating. For example, Table 19.1 shows that a bond with an initial Moody's credit rating of Baa has a 0.177% chance of defaulting by the end of the first year, a 0.461% chance of defaulting by the end of the second year, and so on. The probability of a bond defaulting during a particular year can be calculated from the table. For example, the probability that a bond initially rated Baa will default during the second year of its life is $0.461 - 0.177 = 0.284\%$.

Table 19.1 shows that, for investment-grade bonds, the probability of default in a year tends to be an increasing function of time. (For example, the probabilities of a bond initially rated Aa defaulting during years one, two, three, four, and five are 0.021%, 0.039%, 0.050%, 0.082%, and 0.106%, respectively.) This is because the bond issuer is initially considered to be creditworthy and the more time that elapses, the greater the possibility that its financial health will decline. For bonds with a poor credit rating, the probability of default is often a decreasing function of time. (For example, the probabilities that a bond initially in the lowest (Caa-C) category will default during years one, two, three, four, and five are 10.624%, 8.046%, 6.773%, 5.531%, and 4.569%, respectively.) The reason here is that, for a bond with a poor credit rating, the next year or two may be critical. If the issuer survives this period, its financial health is likely to have improved.

19.2.1 Hazard Rates

From Table 19.1, we can calculate the probability of a Caa-C bond defaulting during the third year as $25.443 - 18.670 = 6.773\%$. We will refer to this as the *unconditional default probability*. It is the probability of default during the third year as seen at time zero. The probability that the Caa-rated bond will survive until the end of year two is $100 - 18.670 = 81.330\%$. The probability that it will default during the third year conditional on no earlier default is therefore $0.06773/0.81330$ or 8.33%.

The 8.33% we have just calculated is a conditional default probability for a one-year time period. When we consider a conditional default probability for a short time

Table 19.1 Average Cumulative Issuer-Weighted Default Rates (%), 1970–2016

Time (years)	1	2	3	4	5	7	10	15	20
Aaa	0.000	0.011	0.011	0.031	0.085	0.195	0.386	0.705	0.824
Aa	0.021	0.060	0.110	0.192	0.298	0.525	0.778	1.336	2.151
A	0.055	0.165	0.345	0.536	0.766	1.297	2.224	3.876	5.793
Baa	0.177	0.461	0.804	1.216	1.628	2.472	3.925	7.006	10.236
Ba	0.945	2.583	4.492	6.518	8.392	11.667	16.283	23.576	29.733
B	3.573	8.436	13.377	17.828	21.908	28.857	36.177	43.658	48.644
Caa-C	10.624	18.670	25.443	30.974	35.543	42.132	50.258	53.377	53.930

SOURCE: Moody's.

period of length Δt, we get a measure known as the *hazard rate* or *default intensity*. The hazard rate, $\lambda(t)$, at time t is defined so that $\lambda(t)\Delta t$ is the probability of default between time t and $t + \Delta t$ conditional on no default between time zero and time t. If $V(t)$ is the cumulative probability of the company surviving to time t (i.e., no default by time t), the unconditional default probability between times t and $t + \Delta t$ is $[V(t) - V(t + \Delta t)]$. The probability of default between times t and $t + \Delta t$ conditional on no earlier default is $[V(t) - V(t + \Delta t)]/V(t)$. Hence

$$\frac{V(t) - V(t + \Delta t)}{V(t)} = \lambda(t)\Delta t$$

or

$$\frac{V(t + \Delta t) - V(t)}{\Delta t} = -\lambda(t)\, V(t)$$

Taking limits

$$\frac{dV(t)}{dt} = -\lambda(t)\, V(t)$$

from which

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

Defining $Q(t)$ as the probability of default by time t, so that $Q(t) = 1 - V(t)$ gives

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

or

$$Q(t) = 1 - e^{-\bar{\lambda}(t)t} \tag{19.2}$$

where $\bar{\lambda}(t)$ is the average hazard rate between time zero and time t.

Example 19.2

Suppose that the hazard rate is a constant 1.5% per year. The probability of a default by the end of the first year is $1 - e^{-0.015\times1} = 0.0149$. The probability of a default by the end of the second year is $1 - e^{-0.015\times2} = 0.0296$. The probabilities of a default by the end of the third, fourth, and fifth years are similarly 0.0440, 0.0582, and 0.0723. The unconditional probability of a default during the fourth year is $0.0582 - 0.0440 = 0.0142$. The probability of default in the fourth year, conditional on no earlier default, is $0.0142/(1 - 0.0440) = 0.0149$.

Table 19.2 Recovery Rates on Corporate Bonds as a Percent of Face Value, 1983 to 2016, Issuer Weighted

Class	**Average Recovery Rate (%)**
1st lien bond	52.8
2nd lien bond	44.6
Senior unsecured bond	37.2
Senior subordinated bond	31.1
Subordinated bond	31.9
Junior subordinated bond	23.2

SOURCE: Moody's.

19.3 Recovery Rates

When a company goes bankrupt, those who are owed money by the company file claims against the company.[4] Sometimes there is a reorganization in which these creditors agree to a partial payment of their claims. In other cases, the assets are sold by the liquidator and the proceeds are used to meet the claims as far as possible. Some claims typically have priorities over other claims and are met more fully.

The *recovery rate* for a bond is normally defined as the price at which it trades about one month after default as a percent of its face value. As we saw in Chapter 15, the Basel II formulas are expressed in terms of the loss given default (LGD). The percentage recovery rate is 100 minus the percentage loss given default.

Table 19.2 provides historical data on average recovery rates for different categories of bonds in the United States. This varies from 52.8% for those who have a first lien on assets to 23.2% for those who rank after other lenders.

19.3.1 The Dependence of Recovery Rates on Default Rates

In Chapter 6, we saw that one of the lessons from the credit crisis of 2007–2008 is that the average recovery rate on mortgages is negatively related to the mortgage default rate. As the mortgage default rate increases, foreclosures lead to more houses being offered for sale and a decline in house prices. This in turn results in a decline in recovery rates.

The average recovery rate on corporate bonds exhibits a similar negative dependence on default rates.[5] In a year when the number of bonds defaulting is low, economic conditions are usually good and the average recovery rate on those bonds that do default might be as high as 60%; in a year when the default rate on corporate bonds is high, economic conditions are usually poor and the average recovery rate on the defaulting bonds might

[4] In the United States, the claim made by a bondholder is the bond's face value plus accrued interest.

[5] See E. I. Altman, B. Brady, A. Resti, and A. Sironi, "The Link between Default and Recovery Rates: Theory, Empirical Evidence, and Implications," *Journal of Business* (November 2005): 2203–2228. The correlation is also discussed in publications by Moody's Investors Service. It finds that the average recovery rate in a year and the non-investment-grade default rate are highly correlated.

be as low as 30%. The result of the negative dependence is that a bad year for defaults is doubly bad for a lender because it is usually accompanied by a low recovery rate.

19.4 Credit Default Swaps

A credit default swap (CDS) is a derivative that allows market partipants to trade credit risks. As indicated in Business Snapshot 19.1, the market for this product grew rapidly until 2007 and then declined. The simplest type of CDS is an instrument that provides insurance against the risk of a default by a particular company. The company is known as the *reference entity* and a default by the company is known as a *credit event*. The buyer of the insurance obtains the right to sell bonds issued by the company for their face value when a credit event occurs and the seller of the insurance agrees to buy the bonds for their face value when a credit event occurs.[6] The total face value of the bonds that can be sold is known as the credit default swap's *notional principal*.

The buyer of a CDS makes periodic payments to the seller until the end of the life of the CDS or until a credit event occurs. These payments are usually made in arrears every quarter.

An example will help to illustrate how a typical deal is structured. Suppose that two parties enter into a five-year credit default swap on December 20, 2018. Assume that the notional principal is $100 million and the buyer agrees to pay 90 basis points per year (quarterly in arrears) for protection against default by the reference entity.

The CDS is shown in Figure 19.1. If the reference entity does not default (that is, there is no credit event), the buyer receives no payoff and pays approximately $225,000 ($= 0.25 \times 0.0090 \times 100{,}000{,}000$) on March 20, June 20, September 20, and December 20 of each of the years 2019, 2020, 2021, 2022, and 2023.[7] If there is a credit event, a substantial payoff is likely. Suppose that the buyer notifies the seller of a credit event on May 20, 2021 (five months into the third year). If the contract specifies physical settlement, the buyer of protection has the right to sell to the seller of protection bonds issued by the reference entity with a face value of $100 million for $100 million. If, as is now usual, there is a cash settlement, a two-stage auction process is used to determine the mid-market value of the cheapest deliverable bond several days after the credit event. Suppose the auction indicates that the cheapest deliverable bond is worth $35 per $100 of face value. The cash payoff would be $65 million.

The regular payments from the buyer of protection to the seller of protection cease when there is a credit event. However, because these payments are made in arrears, a final accrual payment by the buyer is usually required. In our example, where there is a default on May 20, 2021, the buyer would be required to pay to the seller the amount of

[6] The face value (or par value) of a bond is the principal amount that the issuer will repay at maturity if it does not default.

[7] The payments are not exactly $225,000 because of the impact of day count conventions.

BUSINESS SNAPSHOT 19.1

The CDS Market

In 1998 and 1999, the International Swaps and Derivatives Association developed a standard contract for trading credit default swaps in the over-the-counter market. After that, the market grew very fast. The Bank for International Settlements (BIS) started producing statistics for the size of the credit derivatives market in December 2004. At that time, the total notional principal underlying outstanding contracts was estimated to be about $6 trillion. It peaked at $58 trillion in December 2007. But it had fallen to about $10 trillion by December 2016. Banks and other financial institutions are both buyers and sellers of protection. Banks tend to be net buyers of protection, and insurance companies tend to be net sellers of protection. Banks sometimes use credit default swaps to hedge an exposure to a borrower. It is then the case that the financial institution bearing the credit risk of a loan is different from the financial institution that did the original credit checks.

During the credit turmoil that started in August 2007, regulators became concerned that CDSs were a source of systemic risk. (See Business Snapshot 15.1 for a discussion of systemic risk.) No doubt their concerns arose in part because of the losses experienced by the insurance company AIG. This was a big seller of protection on the AAA-rated tranches created from mortgages (see Chapter 6). The protection proved very costly to AIG and a failure of AIG would have led to big losses elsewhere in the financial system. AIG was bailed out by the United States government in September 2008.

CDSs have come under criticism during the European sovereign debt crisis. Some legislators feel that speculative activity in credit default swap markets has exacerbated the debt problems of countries such as Greece; naked CDS positions on sovereign debt (where credit protection on a country is bought without an underlying exposure) were banned in Europe in 2013.

During 2007 and 2008, trading ceased in many types of credit derivatives, but plain vanilla CDSs, which provide protection against a single company or country defaulting, continued to trade actively (albeit with dramatically increased spreads). This is because their structure is easy to understand compared with, say, ABS CDOs (see Chapter 6). However, as pointed out in Business Snapshot 19.2, flaws in this structure are emerging. This may account for their declining popularity.

There was a huge number of CDS contracts outstanding with Lehman Brothers as the reference entity when Lehman Brothers declared bankruptcy in September 2008. The recovery rate (determined by an auction process) was only about eight cents on the dollar, so that the payout to the buyers of protection was equal to about 92% of the notional principal. There were predictions that some sellers of protection would be unable to pay and that further bankruptcies would occur, but on the settlement day (October 21, 2008) everything went smoothly.

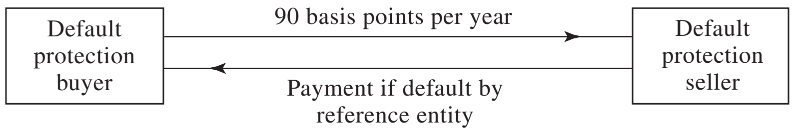


Figure 19.1 Credit Default Swap

the annual payment accrued between March 20, 2021, and May 20, 2021 (approximately $150,000), but no further payments would be required.

The total amount paid per year, as a percent of the notional principal, to buy protection is known as the *CDS spread*. (In our example, the CDS spread is 90 basis points.) Several large banks are market makers in the credit default swap market. For a five-year credit default swap on a company, a market maker might quote: bid 250 basis points, offer 260 basis points. This means that the market maker is prepared to buy protection by paying 250 basis points per year (i.e., 2.5% of the principal per year) and to sell protection for 260 basis points per year (i.e., 2.6% of the principal per year).

Many different companies and countries are reference entities for the CDS contracts traded. Contracts with maturities of five years are most popular, but other maturities such as 1, 2, 3, 7, and 10 years are also traded. Usually, contracts mature on the following standard dates: March 20, June 20, September 20, and December 20. The effect of this is that the actual time to maturity of a contract when it is initiated is close to, but not necessarily the same as, the number of years to maturity that is specified. Suppose you call a dealer on November 15, 2018, to buy five-year protection on a reference entity. The contract would probably last until December 20, 2023. Your first payment would be due on December 20, 2018, and would equal an amount covering the November 15, 2018, to December 20, 2018, period.[8] After that, payments would be made quarterly.

A key aspect of a CDS contract is the definition of a credit event (i.e., a default). Usually a credit event is defined as a failure to make a payment as it becomes due, a restructuring of debt, or a bankruptcy. Restructuring is sometimes excluded in North American contracts, particularly in situations where the yield on the company's debt is high. Potential flaws in the CDS market are discussed in Business Snapshot 19.2. They may account for the market's declining popularity.

19.4.1 The Cheapest-to-Deliver Bond

Usually a CDS specifies that a number of different bonds can be delivered in the event of a default. The bonds typically have the same seniority, but they may not sell for the same

[8] If the time to the first date is less than one month, the first payment is typically on the second payment date, otherwise it is on the first payment date.

BUSINESS SNAPSHOT 19.2

Is a CDS a Flawed Contract?

There is one important difference between credit default swaps and most other over-the-counter derivatives. The other over-the-counter derivatives depend on interest rates, exchange rates, equity indices, commodity prices, and so on. There is no reason to assume that any one market participant has better information than other market participants about these variables.

Credit default swap spreads depend on the probability that a particular company will default during a particular period of time. Arguably, some market participants have more information to estimate this probability than others. A financial institution that works closely with a particular company by providing advice, making loans, and handling new issues of securities is likely to have more information about the creditworthiness of the company than another financial institution that has no dealings with the company. Economists refer to this as an *asymmetric information* problem.

A blatant abuse of the CDS market was reported in the *Financial Times* on January 11, 2018. A U.S. housebuilder, Hovnanian, obtained attractive funding from the hedge fund GSO Capital Partners with the agreement that it would skip an interest payment, thereby manufacturing a technical default. GSO had bought protection on Hovnanian and obtained a big payoff. Assuming that this is a legal arrangement, it may lead to CDSs, as currently structured, disappearing altogether.

percentage of face value immediately after a default.[9] This gives the holder of a CDS what is known as a *cheapest-to-deliver bond* option. As already mentioned, an auction process is usually used to determine the value of the cheapest-to-deliver bond and, therefore, the payoff to the buyer of protection.

The determination of CDS spreads and the valuation of CDS transactions is discussed in Appendix K.

19.4.2 Credit Indices

Participants in credit markets have developed indices to track credit default swap spreads. In 2004, there were agreements between different producers of indices that

[9] There are a number of reasons for this. The claim that is made in the event of a default is typically equal to the bond's face value plus accrued interest. Bonds with high accrued interest at the time of default therefore tend to have higher prices immediately after default. Also, the market may judge that in the event of a reorganization of the company some bondholders will fare better than others.

led to some consolidation. Two important standard portfolios used by index providers are:

1. CDX NA IG, a portfolio of 125 investment-grade companies in North America
2. iTraxx Europe, a portfolio of 125 investment-grade companies in Europe

These portfolios are updated on March 20 and September 20 each year. Companies that are no longer investment grade are dropped from the portfolios and new investment-grade companies are added.[10]

Suppose that the five-year CDX NA IG index is quoted by a market maker as bid 165 basis points, offer 166 basis points. The quotes mean that a trader can buy CDS protection on all 125 companies in the index for 166 basis points per company. Suppose an investor wants \$800,000 of protection on each company. The total cost is $0.0166 \times 800{,}000 \times 125$ or \$1,660,000 per year. (The investor could similarly sell \$800,000 of protection on each of the 125 companies for a total of \$1,650,000 per annum.) When a company defaults, the investor receives the usual CDS payoff and the annual payment is reduced by $1{,}660{,}000/125 = \$13{,}280$. There is an active market in buying and selling CDS index protection for maturities of 3, 5, 7, and 10 years. The maturities for these types of contracts on the index are usually December 20 and June 20. (This means that a "five-year" contract lasts between $4\frac{3}{4}$ and $5\frac{1}{4}$ years.) Roughly speaking, the index is the average of the CDS spreads on the companies in the underlying portfolio.[11]

19.4.3 The Use of Fixed Coupons

To facilitate trading, the precise way in which CDS and CDS index transactions work is a little more complicated than has been described up to now. In practice, CDS and CDS indices trade like bonds. For each of the standard transactions that trade, a fixed coupon and a recovery rate are specified. The periodic protection payments are equal to the coupon. The CDS spread is quoted in the market. When there is a trade, the present value of the difference between payments equal to the spread and payments equal to the coupon for the remaining life of the transaction is estimated using a specified procedure. If the quoted spread is less than the coupon, the seller of protection pays this present value

[10] On September 20, 2014, the Series 22 iTraxx Europe portfolio and the Series 23 CDX NA IG portfolio were defined. The series numbers indicate that by the end of September 2014 the iTraxx Europe portfolio had been updated 21 times and the CDX NA IG portfolio had been updated 22 times.

[11] More precisely, the index is slightly lower than the average of the credit default swap spreads for the companies in the portfolio. To understand the reason for this, consider a portfolio consisting of two companies, one with a spread of 1,000 basis points and the other with a spread of 10 basis points. To buy protection on the companies would cost slightly less than 505 basis points per company. This is because the 1,000 basis points is not expected to be paid for as long as the 10 basis points and should therefore carry less weight. Another complication for CDX NA IG, but not iTraxx Europe, is that the definition of default applicable to the index includes restructuring whereas the definition for CDS contracts on the underlying companies may not.

to the buyer of protection at the time of the trade. If the quoted spread is greater than the coupon, the buyer of protection pays this present value to the seller of protection at the time of the trade. After that, the buyer of protection then pays the coupon times the principal applicable to those companies in the index that have not yet defaulted to the seller.

19.5 Credit Spreads

The credit spread is the extra rate of interest per annum required by investors for bearing a particular credit risk. CDS spreads, which were explained in the previous section, provide one measure of the credit spread. Another is the bond yield spread. This is the amount by which the yield on a corporate bond exceeds the yield on a similar risk-free bond. We now show that the two should be approximately equal.

19.5.1 CDS Spreads and Bond Yields

A CDS can be used to hedge a position in a corporate bond. Suppose that an investor buys a five-year corporate bond yielding 7% per year for its face value and at the same time enters into a five-year CDS to buy protection against the issuer of the bond defaulting. Suppose that the CDS spread is 200 basis points or 2% per annum. The effect of the CDS is to convert the corporate bond to a risk-free bond (at least approximately). If the bond issuer does not default, the investor earns 5% per year (when the CDS spread is netted against the corporate bond yield). If the bond issuer does default, the investor earns 5% up to the time of the default. Under the terms of the CDS, the investor is then able to exchange the bond for its face value. This face value can be invested at the risk-free rate for the remainder of the five years.

This argument shows that the n-year CDS spread should be approximately equal to the excess of the par yield on an n-year corporate bond over the par yield on an n-year risk-free bond.[12] If it is markedly less than this, an investor can earn more than the risk-free rate by buying the corporate bond and buying protection. If it is markedly greater than this, an investor can borrow at less than the risk-free rate by shorting the corporate bond and selling CDS protection. As will be discussed later, these are not perfect arbitrages, and the extent to which they can be used is in practice influenced by liquidity constraints and other considerations.

19.5.2 The Risk-Free Rate

CDSs provide a direct estimate of the credit spread. To calculate a credit spread from a bond yield, it is necessary to make an assumption about the risk-free rate. When bond yield spreads are quoted by bond traders, the risk-free rate that is used is usually the yield

[12]The par yield on an n-year bond is the coupon rate per year that causes the bond to sell for its par value (i.e., its face value).

on a Treasury bond of similar maturity. For example, a bond trader might quote the yield on a particular corporate bond as being a spread of 250 basis points over Treasuries.

A number of researchers have compared bond yields to CDS spreads to imply a risk-free rate. This involves matching the maturities of CDSs and bonds and implying a risk-free rate from the arbitrage arguments given above. For example, if the five-year bond yield is 4.7% and the five-year CDS spread is 80 basis points, the implied five-year risk-free rate is 3.9%.

As discussed in Section 9.2, traders have traditionally used LIBOR/swap rates as proxies for risk-free rates when valuing derivatives. The research indicates that this practice has carried over to the credit market. Implied risk-free rates are much closer to the LIBOR/swap rates than to the Treasury rates. One estimate puts implied risk-free rates at about 10 basis points less than the LIBOR/swap rate.[13] This estimate is plausible. As explained in Section 9.2, the credit risk in a swap rate is the credit risk from making a series of short-term loans to AA-rated counterparties and 10 basis points is a reasonable credit spread for a short-term AA-rated instrument.

19.5.3 Asset Swaps

Asset swaps provide a convenient reference point for traders in credit markets because they give direct estimates of the excess of bond yields over LIBOR/swap rates.

To explain how asset swaps work, consider the situation where an asset swap spread for a particular bond is quoted as 150 basis points. There are three possible situations:

1. The bond sells for its par value of 100. The swap then involves one side (Company A) paying the coupon on the bond and the other side (Company B) paying LIBOR plus 150 basis points.[14]
2. The bond sells below its par value, say, for 95. The swap is then structured so that Company A pays \$5 per \$100 of notional principal at the outset. After that, Company A pays the bond's coupons and Company B pays LIBOR plus 150 basis points.
3. The underlying bond sells above par, say, for 108. The swap is then structured so that Company B makes a payment of \$8 per \$100 of principal at the outset. After that, Company A pays the bond's coupons and Company B pays LIBOR plus 150 basis points.

The effect of structuring asset swaps in this way is that the present value of the asset swap spread (150 basis points in our example) is the amount by which the price of the corporate bond is exceeded by the price of a similar risk-free bond where the risk-free rate is assumed to be given by the LIBOR/swap curve (see Problem 19.16).

---

[13]See J. Hull, M. Predescu, and A. White, "The Relationship between Credit Default Swap Spreads, Bond Yields, and Credit Rating Announcements," *Journal of Banking and Finance* 28 (November 2004): 2789–2811.

[14]Note that it is the promised coupons that are exchanged. The exchanges take place regardless of whether the bond defaults.

19.5.4 CDS–Bond Basis

The CDS–bond basis is the excess of the CDS spread over the bond yield spread for a company.

$$\text{CDS–Bond Basis} = \text{CDS Spread} - \text{Bond Yield Spread}$$

The bond yield spread is calculated relative to the LIBOR/swap benchmark. Usually it is assumed to be the asset swap spread.

The arbitrage argument given above, relating CDS spreads and bond yields, suggests that the CDS–bond basis should be close to zero. In fact, there are a number of reasons why it deviates from zero. For example:

1. The bond may sell for a price that is significantly different from par. (Bond prices above par tend to give rise to a negative basis; bond prices below par tend to give rise to a positive basis.)
2. There is counterparty default risk in a CDS. (This pushes the basis in a negative direction.)
3. There is a cheapest-to-deliver bond option in a CDS. (This pushes the basis in a positive direction.)
4. The payoff in a CDS does not include accrued interest on the bond that is delivered. (This pushes the basis in a negative direction.)
5. The restructuring clause in a CDS contract may lead to a payoff when there is no default. (This pushes the basis in a positive direction.)
6. LIBOR is greater than the risk-free rate being assumed by the market. (This pushes the basis in a positive direction.)

Prior to the market turmoil starting in 2007, the basis tended to be positive. For example, De Witt estimates that the average CDS bond basis in 2004 and 2005 was 16 basis points.[15] During the credit crisis, the basis was at times very negative but, as explained by Bai and Collin-Dufresne (2011), it was difficult for financial institutions to arbitrage between bonds and CDSs because of a shortage of liquidity and other considerations.[16] Since the crisis, the magnitude of the CDS-bond basis (sometimes positive and sometimes negative) has become much smaller.

19.6 Estimating Default Probabilities from Credit Spreads

We now discuss how default probabilities can be estimated from credit spreads.

[15] See J. De Witt, "Exploring the CDS–Bond Basis," Working Paper no. 104, National Bank of Belgium, 2006.

[16] See J. Bai and P. Collin-Dufresne, "The CDS–Bond Basis," Working Paper, SSRN 2024531, 2011.

19.6.1 Approximate Calculation

Suppose that a five-year credit spread (CDS spread, bond yield spread, or asset swap spread) for a company is 240 basis points and that the expected recovery rate in the event of a default is 40%. The holder of a corporate bond issued by the company must be expecting to lose 240 basis points (or 2.4% per year) from defaults. Roughly speaking, the credit spread can be considered to be an average loss rate. Given the recovery rate of 40%, this leads to an estimate of the average probability of a default per year over the five-year period, conditional on no earlier default, of $0.024/(1 - 0.4)$ or 4%. In general:

$$\bar{\lambda} = \frac{s(T)}{1 - R} \tag{19.3}$$

where $s(T)$ is the credit spread (which should in theory be expressed with continuous compounding) for a maturity of T, R is the recovery rate, and $\bar{\lambda}$ is the average hazard rate between time zero and time T.

If credit spreads are known for a number of different maturities, the term structure of the hazard rate can be bootstrapped (at least approximately), as the following example illustrates.

Example 19.3

Suppose that the CDS spreads for 3-, 5-, and 10-year instruments are 50, 60, and 100 basis points, respectively, and the expected recovery rate is 60%. The average hazard rate over three years is approximately $0.005/(1 - 0.6) = 0.0125$. The average hazard rate over five years is approximately $0.006/(1 - 0.6) = 0.015$. The average hazard rate over 10 years is approximately $0.01/(1 - 0.6) = 0.025$. From this we can estimate that the average hazard rate between year 3 and year 5 is $(5 \times 0.015 - 3 \times 0.0125)/2 = 0.01875$. The average hazard rate between year 5 and year 10 is $(10 \times 0.025 - 5 \times 0.015)/5 = 0.035$.

19.6.2 A More Exact Calculation

The calculation we have just given works well for CDS spreads. It also works well for bond yield spreads and asset swap spreads when the underlying bond is selling for close to its par value. We now consider a more exact calculation for situations when the underlying bond's price is not close to par.

Suppose that a five-year corporate bond with a principal of 100 provides a coupon of 6% per annum (paid semiannually) and that the yield on the bond is 7% per annum (with continuous compounding). The yield on a similar risk-free bond is 5% (again with continuous compounding). The yields imply that the price of the corporate bond is 95.34 and the price of the risk-free bond is 104.09. The expected loss from default over the five-year life of the bond is therefore 104.09 – 95.34, or \$8.75. For simplicity, we

Table 19.3 Calculation of Loss from Default on a Bond in Terms of the Default Probabilities per Year, Q

Time (yrs)	Def. Prob.	Recovery Amount ($)	Default-Free Value ($)	Loss ($)	Discount Factor	PV of Expected Loss ($)
0.5	Q	40	106.73	66.73	0.9753	65.08Q
1.5	Q	40	105.97	65.97	0.9277	61.20Q
2.5	Q	40	105.17	65.17	0.8825	57.52Q
3.5	Q	40	104.34	64.34	0.8395	54.01Q
4.5	Q	40	103.46	63.46	0.7985	50.67Q
Total						288.48Q

Notional Principal = $100.

suppose that the unconditional probability of default per year is the same each year and equal to Q. Furthermore, we assume defaults can happen only at times 0.5, 1.5, 2.5, 3.5, and 4.5 years (immediately before coupon payment dates). Risk-free rates are assumed to be 5% (with continuous compounding) for all maturities and the recovery rate (i.e., recovery as a percent of principal) is assumed to be 40%. (The analysis can be extended so that defaults happen more frequently.)

Table 19.3 calculates the expected loss from defaults in terms of Q. To illustrate the calculations, consider the 3.5-year row in Table 19.3. The expected value of the default-free bond at time 3.5 years (calculated using the forward risk-free interest rates) is

$$3 + 3e^{-0.05\times 0.5} + 3e^{-0.05\times 1.0} + 103e^{-0.05\times 1.5} = 104.34$$

Given the definition of recovery rates in Section 19.3, the amount recovered if there is a default is 40 so that the loss given default is 104.34 − 40 or $64.34. The present value of this loss is 54.01 and the expected loss is therefore 54.01Q.

Table 19.3 shows that the total expected loss is 288.48Q. Setting this equal to the 8.75 expected loss calculated earlier, we obtain a value for Q of 8.75/288.48, or 3.03%. The calculations we have given assume that the default probability is the same in each year and that defaults take place at just one time during the year. We can extend the calculations to assume that defaults take place more frequently. Also, instead of assuming a constant unconditional probability of default, we can assume a constant hazard rate or assume a particular pattern for the variation of the default probability with time. With several bonds, we can estimate several parameters describing the term structure of default probabilities. Suppose, for example, that we have bonds maturing in 3, 5, 7, and 10 years and we assume a step function for the default probability. We could use the first bond to estimate the default probability per year for the first three years, the second bond to estimate the default probability per year for years 4 and 5, the third bond to estimate the default probability per year for years 6 and 7, and the fourth bond to estimate the default probability per year for years 8, 9, and 10 (see Problems 19.15 and 19.24). This approach is analogous to the bootstrap procedure for estimating the term structure of interest rates in Appendix B.

Table 19.4 Cumulative Default Probabilities Compared with Credit Spreads

Rating	Cumulative 7-Year Default Probabilities (%), 1970–2016	7-Year Credit Spread (bp) 1996–2007
Aaa	0.195	35.74
Aa	0.525	43.67
A	1.297	68.68
Baa	2.472	127.53
Ba	11.667	280.28
B	28.857	481.04
Caa	42.132	1,103.70

19.7 Comparison of Default Probability Estimates

We now compare the default probability estimates calculated from historical data with those calculated from credit spreads. Table 19.4 shows:

- The seven-year average cumulative probability of default for companies with different credit ratings. This is taken from data published by Moody's. It is the seven-year column of Table 19.1.
- The average credit spread for bonds with different credit ratings between December 1996 and June 2007. This is taken from data on bond yields published by Merrill Lynch and an assumption that the risk-free rate is 10 basis points less than the seven-year swap rate. (See discussion of risk-free rate proxy in Section 19.5.2.) The bonds had an average maturity of about seven years.

The table considers data on bond yields only up to the start of the credit crisis. During the credit crisis, credit spreads soared. If that unusual period had been included, the results we present, which show that hazard rates calculated from credit spreads are higher than those calculated from historical default probabilities, would become much more extreme.

Table 19.5 shows estimates of the average seven-year hazard rates.[17] The historical seven-year hazard rates are calculated so that they are consistent with the cumulative default rates in Table 19.4. Suppose that Q is the seven-year cumulative default probability. From equation (19.2)

$$Q = 1 - e^{-7\bar{\lambda}}$$

where $\bar{\lambda}$ is the average hazard rate so that

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

[17]The analysis here is an update to that in J. Hull, M. Predescu, and A. White, "Bond Prices, Default Probabilities, and Risk Premiums," *Journal of Credit Risk* 1, no. 2 (Spring 2005): 53–60.

Table 19.5 Average Seven-Year Hazard Rates

Rating	Historical Hazard Rate (%)	Hazard Rate from Credit Spread (%)	Ratio	Difference
Aaa	0.028	0.596	21.4	0.568
Aa	0.075	0.728	9.7	0.653
A	0.186	1.145	6.1	0.959
Baa	0.358	2.126	5.9	1.768
Ba	1.772	4.671	2.6	2.899
B	4.864	8.017	1.6	3.153
Caa	7.814	18.395	2.4	10.581

% per annum.

Consider, for example, an A-rated company. The cumulative seven-year default rate is 1.217% or 0.01217 so that the average hazard rate is

$$-\frac{1}{7}\ln(1-0.01297) = 0.00186$$

or 0.186%. The hazard rates estimated from the credit spreads in Table 19.5 are based on equation (19.3). The recovery rate is assumed to be 40%. Consider again A-rated bonds. The average seven-year credit spread in Table 19.4 is 68.68 basis points or 0.006868. The average seven-year hazard rate is therefore

$$0.006868/(1-0.4) = 0.01145$$

Table 19.5 shows that the hazard rates calculated from credit spreads before the credit crisis are higher than those calculated from a long period of historical data. The ratio of the hazard rate backed out of bond prices to the hazard rate calculated from historical data is high for investment-grade bonds and tends to decline as the credit quality declines. By contrast, the difference between the two hazard rates tends to increase as credit quality declines.[18]

Table 19.6 provides another way of looking at these results. It shows the excess return over the risk-free rate (still assumed to be the seven-year swap rate minus 10 basis points) earned by investors in bonds with different credit ratings. Consider again an A-rated bond. The average spread over Treasuries is 111 basis points. Of this, 42 basis points are accounted for by the average spread between seven-year Treasuries and our proxy for the risk-free rate. A spread of 11 basis points is necessary to cover expected defaults. (This equals the historical hazard rate in Table 19.5 multiplied by one minus the assumed

[18]Other studies have identified a similar difference between the two types of default probability estimates. See, for example, J. S. Fons, "The Default Premium and Corporate Bond Experience," *Journal of Finance* 42, no. 1 (March 1987): 81–97, and E. I. Altman, "Measuring Corporate Bond Mortality and Performance," *Journal of Finance* 44, no. 4 (September 1989): 909–922.

Table 19.6 Expected Excess Return on Bonds

	Bond Yield Spread over Treasuries (bp)	Spread of Risk-Free Rate over Treasuries (bp)	Spread for Historical Defaults (bp)	Expected Excess Return (bp)
Aaa	78	42	2	34
Aa	86	42	5	39
A	111	42	11	58
Baa	169	42	21	106
Ba	322	42	106	174
B	523	42	292	189
Caa	1,146	42	469	635

recovery rate of 0.4.) This leaves an expected excess return (after expected defaults have been taken into account) of 58 basis points.[19]

Tables 19.5 and 19.6 show that a large percentage difference between default probability estimates translates into a relatively small expected excess return on the bond. For Aaa-rated bonds, the ratio of the two hazard rates is 21.4, but the expected excess return is only 34 basis points. The expected excess return tends to increase as credit quality declines.

The excess return in Table 19.6 does not remain constant through time. Credit spreads, and therefore excess returns, were high in 2001, 2002, and the first half of 2003. Credit spreads were very high during the crisis and, if excess returns had been calculated for the crisis period, they would have been high as well.

19.7.1 Real-World vs. Risk-Neutral Probabilities

The risk-neutral valuation argument is explained in Section 7.2. It shows that we can value cash flows on the assumption that all investors are risk neutral (that is, on the assumption that they do not require a premium for bearing risks). When we do this, we get the right answer in the real world as well as in the risk-neutral world.

The theoretical underpinnings of risk-neutral default probabilities are also explained in Section 7.2. These default probabilities, sometimes also called implied default probabilities, are the ones obtained when credit spreads are used for estimation. By contrast, the default probabilities estimated from historical data are real-world default probabilities (sometimes also called physical default probabilities). The expected excess return in Table 19.6 arises directly from the difference between real-world and risk-neutral default probabilities. If there was no expected excess return, the real-world and risk-neutral default probabilities would be the same, and vice versa.

[19]To avoid complications, the calculations do not make any adjustments for the impact of compounding frequency issues on spreads, returns, and hazard rates.

Why do we see such big differences between real-world and risk-neutral default probabilities? As we have just argued, this is the same as asking why corporate bond traders earn more than the risk-free rate on average.

One reason for the results is that corporate bonds are relatively illiquid and the returns on bonds are higher than they would otherwise be to compensate for this. But this is a small part of what is going on. In normal markets, it explains perhaps 25 basis points of the excess return in Table 19.6. Another possible reason for the results is that the subjective default probabilities of bond traders are much higher than those given in Table 19.1. Bond traders may be allowing for depression scenarios much worse than anything seen in the period covered by their data. However, it is difficult to see how this can explain a large part of the excess return that is observed.[20]

By far the most important reason for the results in Tables 19.5 and 19.6 is that bonds do not default independently of each other. (To put this another way, default correlation is a feature of financial markets.) Evidence for this is that default rates vary markedly from year to year. Moody's statistics (see Table 11.6) show that between 1970 and 2016, the default rate per year for all rated companies ranged from a low of 0.088% in 1979 to a high of 4.996% in 2009. This year-to-year variation in default rates gives rise to systematic risk (i.e., risk that cannot be diversified away). Bond traders earn an excess expected return for bearing this risk. In this respect, bond traders are no different from equity traders. The average extra return earned by equity traders for bearing systematic risk is often assumed to be 5% or 6% per year. From Table 19.6 we see that the excess return earned by bond traders is much less than this for high-quality bonds. However, as the bond's credit quality decreases, it becomes more like equity and the excess return earned tends to increase.

What causes default correlation and the resultant systematic risk? One explanation is the economy. Good macroeconomic conditions decrease the probability of default for all companies; bad macroeconomic conditions increase the probability of default for all companies. (In Vasicek's model, which was discussed in Chapter 11, the factor F can be regarded as representing the overall health of the economy.) Another explanation is what is known as "contagion." This is discussed in Business Snapshot 19.3.

In addition to systematic risk, which we have just talked about, there is nonsystematic (or idiosyncratic) risk associated with each bond. If we were talking about stocks, we would argue that investors can diversify the nonsystematic risk by choosing a portfolio of, say, 30 stocks. They should not therefore demand a risk premium for bearing nonsystematic risk. For bonds the arguments are not so clear cut. Bond returns are highly skewed, with limited upside. (For example, on an individual bond there might be a 99.75% chance of a 7% return in a year, and a 0.25% chance of a –60% return in the year, the first outcome corresponding to no default and the second to default.) The nonsystematic component

[20] In addition to producing Table 19.1, which is based on the 1970 to 2016 period, Moody's produces a similar table based on the 1920 to 2016 period. When this table is used, historical hazard rates for investment-grade bonds in Table 19.5 rise somewhat. However, the non-investment-grade historical hazard rates decline.

BUSINESS SNAPSHOT 19.3

Contagion

Credit contagion is the process whereby a problem in one sector of the world economy leads to problems in other, unrelated sectors. When Russia defaulted on its debt in 1998, there was a flight to quality and credit spreads on all bonds increased. During the credit crisis that started in 2007, there was a similar flight to quality and again credit spreads increased. The accompanying recession led to a record number of companies defaulting in 2009. In 2011, problems experienced by Greece caused investors to be reluctant to buy the debt of other countries such as Spain, Ireland, Portugal, and Italy. As a result, credit spreads on the debt issued by these countries increased sharply.

The reasons for credit contagion have been debated by researchers. Systemic risk (see Business Snapshot 15.1) is one potential source of contagion. It may be the case that investors become more risk averse when they lose money in one sector. It may be the case that problems in one sector lead investors to become more pessimistic about other, unrelated sectors. Whatever the reason, entities in unrelated sectors are liable to find it more difficult to fund their activities and, as a result, may become more likely to default.

of this risk is difficult to "diversify away."[21] It requires tens of thousands of different bonds to be held. In practice, many bond portfolios are far from fully diversified. As a result, bond traders may earn an extra return for bearing nonsystematic risk as well as for bearing the systematic risk mentioned in the previous paragraph.

19.7.2 Which Estimates Should Be Used?

At this stage it is natural to ask whether we should use real-world or risk-neutral default probabilities in the analysis of credit risk. This question is discussed in a more general context in Chapter 7. The answer depends on the purpose of the analysis. When valuing credit derivatives or estimating the impact of default risk on the pricing of instruments, we should use risk-neutral default probabilities. This is because the analysis calculates the present value of expected future cash flows and almost invariably (implicitly or explicitly) involves using risk-neutral valuation. When carrying out scenario analyses to calculate potential future losses from defaults, we should use real-world default probabilities. For example, the probability of default used to calculate regulatory capital is a real-world default probability.

[21] See J. D. Amato and E. M. Remolona, "The Credit Spread Puzzle," *BIS Quarterly Review* (December 2003): 51–63.

19.8 Using Equity Prices to Estimate Default Probabilities

When we use a table such as Table 19.1 to estimate a company's real-world probability of default, we are relying on the company's credit rating. Unfortunately, credit ratings are revised relatively infrequently. This has led some analysts to argue that equity prices can provide more up-to-date information for estimating default probabilities.

In 1974, Merton proposed a model where a company's equity is an option on the assets of the company.[22] Suppose, for simplicity, that a firm has one zero-coupon bond outstanding and that the bond matures at time T. Define

V_0: Value of company's assets today.
V_T: Value of company's assets at time T.
E_0: Value of company's equity today.
E_T: Value of company's equity at time T.
D: Amount of debt interest and principal due to be repaid at time T.
σ_V: Volatility of assets (assumed constant).
σ_E: Instantaneous volatility of equity.

If $V_T < D$, it is (at least in theory) rational for the company to default on the debt at time T. The value of the equity is then zero. If $V_T > D$, the company should make the debt repayment at time T and the value of the equity at this time is $V_T - D$. Merton's model, therefore, gives the value of the firm's equity at time T as

$$E_T = \max(V_T - D, 0)$$

This shows that the equity of a company is a call option on the value of the assets of the company with a strike price equal to the repayment required on the debt. The Black–Scholes–Merton formula (see Appendix E at the end of this book) gives the value of the equity today as

$$E_0 = V_0 N(d_1) - De^{-rT} N(d_2) \tag{19.4}$$

where

$$d_1 = \frac{\ln(V_0/D) + \left(r + \sigma_V^2/2\right)T}{\sigma_V\sqrt{T}}$$

$$d_2 = d_1 - \sigma_V\sqrt{T}$$

and N is the cumulative normal distribution function.

[22]See R. Merton, "On the Pricing of Corporate Debt: The Risk Structure of Interest Rates," *Journal of Finance* 29 (1974): 449–470.

Under Merton's model, the company defaults when the option is not exercised. The probability of this can be shown to be $N(-d_2)$. To calculate this, we require V_0 and σ_V. Neither of these is directly observable. However, if the company is publicly traded, we can observe E_0. This means that equation (19.4) provides one condition that must be satisfied by V_0 and σ_V. We can also estimate σ_E. From a result in stochastic calculus known as Ito's lemma,

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

Here $\partial E/\partial V$ is the delta of the equity. From Appendix E it is $N(d_1)$, so that

$$\sigma_E E_0 = N(d_1)\sigma_V V_0 \tag{19.5}$$

This provides another equation that must be satisfied by V_0 and σ_V. Equations (19.4) and (19.5) provide a pair of simultaneous equations that can be solved for V_0 and σ_V.[23]

Example 19.4

The value of a company's equity is \$3 million and the volatility of the equity is 80%. The debt that will have to be paid in one year is \$10 million. The risk-free rate is 5% per annum. In this case, $E_0 = 3$, $\sigma_E = 0.80$, $r = 0.05$, $T = 1$, and $D = 10$. Solving equations (19.4) and (19.5) yields $V_0 = 12.40$ and $\sigma_V = 0.2123$. The parameter, d_2, is 1.1408, so that the probability of default is $N(-d_2) = 0.127$ or 12.7%. The market value of the debt is $V_0 - E_0$ or 9.40. The present value of the promised payment on the debt is $10e^{-0.05\times 1} = 9.51$. The expected loss on the debt is therefore $(9.51 - 9.40)/9.51$ or about 1.2% of its no-default value.

19.8.1 Extensions of the Basic Model

The basic Merton's model we have just presented has been extended in a number of ways. For example, one version of the model assumes that a default occurs whenever the value of the assets falls below a barrier level. Another allows payments on debt instruments to be required at more than one time. Many analysts have found the implied volatility of equity issued by a company to be a good predictor of the probability of default. (The higher the implied volatility, the higher the probability of default.) Hull et al. (2004) show that this is consistent with Merton's model.[24] They provide a way of implementing Merton's model using two equity implied volatilities and show that the resulting model provides results comparable to those provided by the usual implementation of the model.

[23]To solve two non-linear equations of the form $F(x, y) = 0$ and $G(x, y) = 0$, we can use the Solver routine in Excel to find the values of x and y that minimize $[F(x, y)]^2 + [G(x, y)]^2$.

[24]See J. Hull, I. Nelken, and A. White, "Merton's Model, Credit Risk, and Volatility Skews," *Journal of Credit Risk* 1, no. 1 (2004): 1–27.

19.8.2 Performance of the Model

How well do the default probabilities produced by Merton's model and its extensions correspond to actual default experience? The answer is that Merton's model and its extensions produce a good ranking of default probabilities (risk-neutral or real-world). This means that a monotonic transformation can be estimated to convert the probability of default output from Merton's model into a good estimate of either the real-world or risk-neutral default probability. Moody's KMV and Kamakura provide a service that transforms a default probability produced by Merton's model into a real-world default probability. CreditGrades uses Merton's model to estimate credit spreads, which are closely linked to risk-neutral default probabilities. The default probability, $N(-d_2)$, is in theory a risk-neutral default probability because it is calculated from an option pricing model. It may seem strange for Moody's KMV and Kamakura to use it to estimate a real-world default probability. Given the nature of the calibration process we have just described, the underlying assumption is that the rankings of risk-neutral default probabilities, real-world default probabilities, and default probabilities produced by Merton's model are all the same.

19.8.3 Real-World vs. Risk-Neutral Default Probabilities

Merton's model provides a way of understanding why default probabilities are higher in the risk-neutral world than in the real world. In a risk-neutral world, the expected growth rate of the value of the company's assets is the risk-free rate. In the real world, the growth rate of the company's assets is usually higher than this (reflecting a risk premium demanded by the market). The probability of the value of the assets dropping below the face value of the debt at a future time is therefore higher in the risk-neutral world than in the real world.

19.8.4 Distance to Default

The term *distance to default* has been coined to describe the output from Merton's model. This is the number of standard deviations the asset price must change for default to be triggered T years in the future. One definition of distance to default is d_2 or

$$\frac{\ln V_0 - \ln D + \left(r - \sigma_V^2/2\right)T}{\sigma_V\sqrt{T}}$$

As the distance to default declines, the company becomes more likely to default. In Example 19.4, the one-year distance to default is 1.14 standard deviations.

Summary

The estimation of default probabilities and recovery rates is an important activity for risk managers. If a company has issued publicly traded debt, credit ratings provide one source of information. Rating agencies such as Moody's provide extensive statistics on default rates for companies that have been awarded particular credit ratings. The recovery rate is the value of a bond shortly after default as a percentage of its face value. Rating agencies provide statistics on recovery rates for different types of bonds.

There are a number of sources of information about credit spreads. The credit default swap (CDS) market is one such source. A CDS is an instrument where one company buys from another company protection against a third company or country (the reference entity) defaulting on its obligations. The CDS spread is the amount paid per year for protection as a percentage of the notional principal. Two other sources of a company's credit spreads are the excess of yields on bonds issued by the company over the risk-free rate and asset swap spreads. Risk-neutral default probabilities can be calculated from credit spreads and an assumption about recovery rates.

The default probabilities that are based on historical data, such as those produced by rating agencies, are termed real-world or physical default probabilities. Risk-neutral default probabilities are higher than real-world default probabilities. Risk-neutral default probabilities should be used for valuation. Real-world default probabilities should be used for scenario analysis. Either type of probability can be calculated using a model developed by Robert Merton in 1974, provided that the model is calibrated appropriately.

Further Reading

Altman, E. I. "Measuring Corporate Bond Mortality and Performance." *Journal of Finance* 44 (1989): 902–922.

Duffie, D., and K. Singleton. "Modeling Term Structures of Defaultable Bonds." *Review of Financial Studies* 12 (1999): 687–720.

Fons, J. S. "The Default Premium and Corporate Bond Experience." *Journal of Finance* 42, no. 1 (March 1987): 81–97.

Hull, J., M. Predescu, and A. White. "Bond Prices, Default Probabilities, and Risk Premiums." *Journal of Credit Risk* 1, no. 2 (Spring 2005): 53–60.

Hull, J., M. Predescu, and A. White. "Relationship between Credit Default Swap Spreads, Bond Yields, and Credit Rating Announcements." *Journal of Banking and Finance* 28 (November 2004): 2789–2811.

Kealhofer, S. "Quantifying Credit Risk I: Default Prediction." *Financial Analysts Journal* 59, no. 1 (2003): 30–44.

Kealhofer, S. "Quantifying Credit Risk II: Debt Valuation." *Financial Analysts Journal* 59, no. 3 (2003): 78–92.

Litterman, R., and T. Iben. "Corporate Bond Valuation and the Term Structure of Credit Spreads." *Journal of Portfolio Management* (Spring 1991): 52–64.

Merton, R. C. "On the Pricing of Corporate Debt: The Risk Structure of Interest Rates." *Journal of Finance* 29 (1974): 449–470.

Rodriguez, R. J. "Default Risk, Yield Spreads, and Time to Maturity." *Journal of Financial and Quantitative Analysis* 23 (1988): 111–117.

Practice Questions and Problems (Answers at End of Book)

19.1 How many different ratings does Moody's use for investment-grade companies? What are they?

19.2 How many different ratings does S&P use for investment-grade companies? What are they?

19.3 Calculate the average hazard rate for a B-rated company during the first year from the data in Table 19.1.

19.4 Calculate the average hazard rate for a Ba-rated company during the third year from the data in Table 19.1.

19.5 A credit default swap requires a semiannual payment at the rate of 60 basis points per year. The principal is $300 million and the credit default swap is settled in cash. A default occurs after four years and two months, and the calculation agent estimates that the price of the cheapest deliverable bond is 40% of its face value shortly after the default. List the cash flows and their timing for the seller of the credit default swap.

19.6 Explain the two ways a credit default swap can be settled.

19.7 Explain the difference between risk-neutral and real-world default probabilities.

19.8 What is the formula relating the payoff on a CDS to the notional principal and the recovery rate?

19.9 The spread between the yield on a three-year corporate bond and the yield on a similar risk-free bond is 50 basis points. The recovery rate is 30%. Estimate the average hazard rate per year over the three-year period.

19.10 The spread between the yield on a five-year bond issued by a company and the yield on a similar risk-free bond is 80 basis points. Assuming a recovery rate of 40%, estimate the average hazard rate per year over the five-year period. If the spread is 70 basis points for a three-year bond, what do your results indicate about the average hazard rate in years 4 and 5?

19.11 Should researchers use real-world or risk-neutral default probabilities for (a) calculating credit value at risk and (b) adjusting the price of a derivative for default?

19.12 How are recovery rates usually defined?

19.13 Verify (a) that the numbers in the second column of Table 19.5 are consistent with the numbers in the second column of Table 19.4 and (b) that the numbers in the fourth column of Table 19.6 are consistent with the numbers in Table 19.5 and a recovery rate of 40%.

19.14 A four-year corporate bond provides a coupon of 4% per year payable semiannually and has a yield of 5% expressed with continuous compounding. The risk-free yield curve is flat at 3% with continuous compounding. Assume that defaults can take place at the end of each year (immediately before a coupon or principal payment) and the recovery rate is 30%. Estimate the risk-neutral default probability on the assumption that it is the same each year using the approach in Table 19.3.

19.15 A company has issued three- and five-year bonds, each of which has a coupon of 4% per annum payable annually. The yields on the bonds (expressed with continuous compounding) are 4.5% and 4.75%, respectively. Risk-free interest rates are 3.5% with continuous compounding for all maturities. The recovery rate is 40%. Defaults can take place halfway through each year. The unconditional risk-neutral default rates per year are Q_1 for years 1 to 3 and Q_2 for years 4 and 5. Estimate Q_1 and Q_2.

19.16 Suppose that in an asset swap, B is the market price of the bond per dollar of principal, B^* is the default-free value of the bond per dollar of principal, and V is the present value of the asset swap spread per dollar of principal. Show that $V = B^* - B$.

19.17 Show that, under Merton's model in Section 19.8, the credit spread on a T-year zero-coupon bond is $-\ln[N(d_2) + N(-d_1)/L]/T$ where $L = De^{-rT}/V_0$.

19.18 The value of a company's equity is \$2 million and the volatility of its equity is 50%. The debt that will have to be repaid in one year is \$5 million. The risk-free interest rate is 4% per annum. Use Merton's model to estimate the probability of default. (*Hint*: The Solver function in Excel can be used for this question.)

19.19 A five-year credit default swap entered into on June 20, 2013, requires quarterly payments at the rate of 400 basis points per year. The principal is \$100 million. A default occurs after four years and two months. The auction process finds the price of the cheapest deliverable bond to be 30% of its face value. List the cash flows and their timing for the seller of the credit default swap.

19.20 "The position of a buyer of a credit default swap is similar to the position of someone who is long a risk-free bond and short a corporate bond." Explain this statement.

19.21 Why is there a potential asymmetric information problem in credit default swaps?

19.22 Suppose that the LIBOR/swap curve is flat at 6% with continuous compounding and a five-year bond with a coupon of 5% (paid semiannually) sells for 90.00. How much would the bond be worth if it were a risk-free bond? What is the present value of the expected loss from defaults? How would an asset swap on the bond be structured? What is the asset swap spread that would be calculated in this situation?

Further Questions

19.23 Suppose that a three-year corporate bond provides a coupon of 7% per year payable semiannually and has a yield of 5% (expressed with semiannual compounding). The yield for all maturities on risk-free bonds is 4% per annum (expressed with semiannual compounding). Assume that defaults can take place every six months (immediately before a coupon payment) and the recovery rate is 45%. Estimate the default probabilities assuming (a) the unconditional default probabilities are the same on each possible default date and (b) the default probabilities conditional on no earlier default are the same on each possible default date.

19.24 A company has issued one- and two-year bonds providing 8% coupons, payable annually. The yields on the bonds (expressed with continuous compounding) are 6.0% and 6.6%, respectively. Risk-free rates are 4.5% for all maturities. The recovery rate is 35%. Defaults can take place halfway through each year. Estimate the risk-neutral default rate each year.

19.25 The value of a company's equity is \$4 million and the volatility of its equity is 60%. The debt that will have to be repaid in two years is \$15 million. The risk-free interest rate is 6% per annum. Use Merton's model to estimate the expected loss from default, the probability of default, and the recovery rate (as a percentage of the no-default value) in the event of default. (*Hint:* The Solver function in Excel can be used for this question as indicated in footnote 23.)

Chapter 20

CVA and DVA

As explained in Chapter 17, there are two ways in which over-the-counter derivatives transactions are cleared: through central clearing parties (CCPs) and through bilateral agreements. As a result of regulations introduced since the credit crisis, transactions between financial institutions in standardized derivatives have to be cleared through CCPs. Transactions in nonstandard derivatives between financial institutions are cleared bilaterally, but regulations being introduced will require a great deal of collateral to be posted so that there will be virtually no credit risk. This chapter focuses on bilaterally cleared over-the-counter derivatives transactions that entail credit risk. This includes transactions with non-systemically important end users and transactions with other counterparties that were entered into before the new rules came into effect.

Assessing the credit risk for derivatives transactions is much more complicated than assessing the credit risk for loans because the future exposure (i.e., the amount that could be lost in the event of a default) is not known. If a bank makes a $10 million uncollateralized five-year loan to a client with repayment of principal at the end, the bank knows that its exposure is approximately $10 million at all times during the five-year period. If the bank instead enters into a five-year uncollateralized interest rate swap with the client, the future exposure is much less certain. This is because the future value of the swap depends on movements in interest rates. If the value of the interest rate swap to the bank becomes positive, the exposure is equal to the value of the swap (because this is what the bank could lose in the event of a counterparty default). If the value becomes negative, the exposure is zero (because in that case the bank would not lose anything in the case of a counterparty default).

The chapter explains the credit value adjustment (CVA), which is the expected loss to a derivatives dealer from a default by the counterparty, and the debit (or debt) value adjustment (DVA), which is the expected gain to the dealer (loss to the counterparty) from a default by the dealer.

20.1 Credit Exposure on Derivatives

We start by quickly reviewing the nature of a dealer's exposure to a counterparty created by derivatives transactions. Consider first the simple case where there is only one derivatives transaction outstanding between a derivatives dealer and a counterparty. We assume that the transaction is cleared bilaterally and no collateral is posted by either the dealer or the counterparty.

Three possible situations can be distinguished:

1. The derivative is always a liability to the dealer.
2. The derivative is always an asset to the dealer.
3. The derivative can become either an asset or a liability to the dealer.

An example of a derivative in the first category is a short position in an option; an example in the second category is a long position in an option; an example in the third category is a forward contract or an interest rate swap transaction.

In the first case, the dealer has no credit exposure to the counterparty. If the counterparty defaults (e.g., because it is bankrupt), the transaction is terminated.[1] As explained in Section 17.1.5, the settlement will require the dealer to pay the negative value of the derivative (adjusted to reflect the bid–offer spread costs that the dealer will incur in replacing the transaction) to the estate of the defaulting counterparty. There should therefore be neither a gain nor a loss to the dealer.

In the second case, the dealer always has credit exposure to the counterparty. If the counterparty defaults, the dealer is liable to experience a loss. The dealer is an unsecured creditor for the value of the derivative (adjusted to reflect the bid–offer spread costs that the dealer will incur in replacing the transaction). The amount of the loss depends on the value of the derivative at the time of the bankruptcy.

The third case is more complicated. The dealer may or may not have credit exposure to the counterparty in the future. If the counterparty defaults when the value of the derivative is positive to the dealer, the dealer is an unsecured creditor, as in the second

[1] Note that, because the derivative is an asset to the counterparty, it must be the case that, if the counterparty goes bankrupt, it has run into financial difficulties for reasons unrelated to the derivative.

case. If the counterparty defaults when the value is negative to the dealer, no loss or gain to the dealer results, as in the first case.

When calculating their exposures, dealers typically do not include bid–offer spread adjustments mentioned at the end of Section 17.1.5. In the simple example we have been considering, the dealer's net exposure is therefore

$$\max(V, 0)$$

where the variable, V, is the market value of the derivative at the time of the default.

When there are many derivatives transactions outstanding between the dealer and the counterparty, the bilateral agreement between the dealer and its counterparties means that they are netted (see Section 15.5). The dealer's exposure at any given time, assuming no collateral is posted, is therefore $\max(V, 0)$, with V equal to the net mark-to-market value of all outstanding transactions.

In calculating its credit risk, the dealer is interested in both its credit exposure today and what that exposure might be in the future. Regulators have recognized this. As explained in Section 15.3, Basel I used the CEM approach where the loan-equivalent exposure is the current exposure plus an add-on amount. A version of this was used to calculate exposure at default (EAD) in Basel II and was revised in SA-CCR (see Section 16.4). The procedure for calculating EAD in the Advanced IRB approach of Basel II involves calculating what is termed the Effective Expected Positive Exposure (see Section 15.8.4).

In this chapter we focus on the calculation of CVA and DVA, which are estimates of expected losses from possible defaults.

20.2 CVA

A dealer calculates what is known as a *credit value adjustment* or CVA for each counterparty with which it has bilaterally cleared OTC derivatives.[2] This is an estimate of its expected loss from a default by the counterparty. CVA reduces the value of the derivatives on the balance sheet, and an increase (decrease) in the total CVA during a period leads to a decrease (or increase) in profits for the period.

Suppose that T is the life of the longest derivative outstanding with a counterparty. To calculate CVA, the period of time between time zero (today) and time T is divided into a number of intervals (e.g., zero to one month, one to three months, etc.).

---

[2] For further reading on CVA and the way it is calculated, see J. Hull and A. White, "CVA and Wrong Way Risk," *Financial Analysts Journal* 65, no. 5 (September/October 2012): 58–69; E. Canabarro and D. Duffie, "Measuring and Marking Counterparty Risk," Chapter 9 in *Asset/Liability Management for Financial Institutions*, ed. L. Tilman (New York: Institutional Investor Books, 2003); E. Picault, "Calculating and Hedging Exposure, CVA, and Economic Capital," in *Counterparty Credit Risk Modeling*, ed. M. Pykhtin (London: Risk Books, 2005); and J. Gregory, *Counterparty Credit Risk: The New Challenge for Financial Markets*, 2nd ed. (Hoboken, NJ: John Wiley & Sons, 2012).

Suppose that the ith time interval runs from t_{i-1} to t_i ($t_0 = 0$) and there are n intervals. Define

q_i: the risk-neutral probability of a loss from a default by the counterparty during the ith time interval
v_i: the present value of the expected net exposure of the dealer to the counterparty (after collateral) at the midpoint of the ith time interval, conditional on a default
R: the estimated recovery rate of the dealer in the event of a default by the counterparty when it is an unsecured creditor

Assuming that the net exposure is independent of the probability of default, the present value of the expected loss from a default during the ith interval can be estimated as

$$(1 - R)q_i v_i$$

and the total expected loss is

$$\text{CVA} = \sum_{i=1}^{n} (1 - R)q_i v_i \tag{20.1}$$

The q_i are risk-neutral default probabilities. (This is because the calculation of CVA involves the valuation of potential future cash flows and, as explained in Chapter 7, it is correct to use risk-neutral, rather than real-world, default probabilities for valuation.) They are estimated from the counterparty's credit spreads, as discussed in Section 19.6. Suppose that s_i is an estimate of the counterparty's credit spread for a maturity of t_i. From equation (19.3), an estimate of the average hazard rate between time zero and time t_i is

$$\bar{\lambda}_i = \frac{s_i}{1 - R} \tag{20.2}$$

and the probability of no default between times zero and t_i is

$$e^{-\bar{\lambda}_i t_i}$$

so that

$$q_i = e^{-\bar{\lambda}_{i-1} t_{i-1}} - e^{-\bar{\lambda}_i t_i}$$

or, substituting from equation (20.2)

$$q_i = \exp\left(-\frac{s_{i-1} t_{i-1}}{1 - R}\right) - \exp\left(-\frac{s_i t_i}{1 - R}\right) \tag{20.3}$$

The v_i are usually calculated using Monte Carlo simulation. The market variables determining the future value of the transactions that the dealer has with the counterparty are simulated in a risk-neutral world between time zero and time T. On each simulation trial the exposure of the dealer to the counterparty at the midpoint of each interval is calculated. These exposures are averaged and v_i is set equal to the present value of the average exposure at the midpoint of the ith interval. Dealers may have transactions with thousands of counterparties so that the calculation of the v_i for all of them can be computationally very intensive.

20.2.1 Collateral and Cure Periods

Collateral agreements must be incorporated into the calculation of the v_i. A collateral agreement between two sides is in the credit support annex (CSA) of their ISDA master agreement. This gives the details on how the collateral posted by each side is calculated and specifies haircuts on noncash collateral. Suppose that C is defined as the collateral posted by the counterparty at the time of a default. (If C is negative, $-C$ is the collateral posted by the dealer with the counterparty at the time of the default.) The exposure, E, of the dealer to the counterparty is in all situations

$$E = \max(V - C,\ 0) \tag{20.4}$$

where V is the market value of outstanding transactions to the dealer at the time of the default.

Equation (20.4) reflects the fact that collateral posted by the counterparty can be used to reduce (and possibly eliminate) an exposure from a positive V. If the dealer has posted collateral with the counterparty ($C < 0$) it will not be returned and, to the extent that it is greater than the value $-V$ of the transactions to the counterparty, it is an exposure. As mentioned, in CVA and DVA calculations exposure is usually defined in terms of the mid-market value of transactions. (In Section 17.1.5, we explained that for the purposes of the claim made in the event of default, mid-market values are adjusted by half the bid-offer spread.)

It is usually assumed that a period of time elapses between the time when a counterparty stops posting collateral and the close-out of transactions. This period of time is known as the *cure period* (or *margin period of risk*). It is typically 10 or 20 days. The effect of the cure period is that the collateral at the time of a default does not reflect the value of the portfolio at the time of the default. It reflects the value 10 or 20 days earlier.

Suppose that the midpoint of the ith interval is t_i^* so that $t_i^* = (t_{i-1} + t_i)/2$ and c is the cure period. The Monte Carlo simulation to calculate the v_i must be structured so that the value of the derivatives portfolio with the counterparty is calculated at times $t_i^* - c$ as well as at time t_i^* ($i = 1, 2, \dots, n$). On each simulation trial the value at time $t_i^* - c$ is used in conjunction with the CSA to calculate the collateral available at time t_i^*. The net exposure at time t_i^* is then calculated using equation (20.4).[3] The calculations are illustrated in Example 20.1.

[3] This is a simplification. It assumes that the non-defaulting party does not post any collateral or return any collateral during the cure period.

Example 20.1

There is a two-way zero threshold collateral agreement between a bank and its counterparty. This means that the total collateral posted by one side with the other side on a day is specified as $\max(V, 0)$, where V is the value of the outstanding transactions to the other side. The cure period is 20 days. This means that the collateral posted by one side at the time of a default can be assumed to equal $\max(V, 0)$ where V is the value to the other side 20 days earlier. Suppose that time τ is the midpoint of one of the intervals used in the bank's CVA calculation.

1. On a particular simulation trial, the value of outstanding transactions to the bank at time τ is 50 and their value 20 days earlier is 45. The calculation assumes that the bank has collateral worth 45 in the event of a default at time τ. The bank's exposure is therefore 5, the uncollateralized value it has in the derivatives transactions.
2. On a particular simulation trial, the value of outstanding transactions to the bank at time τ is 50 and their value 20 days earlier is 55. In this case, it is assumed that the bank will have adequate collateral and its exposure is zero.
3. On a particular simulation trial, the value of outstanding transactions to the bank at time τ is -50 and the value 20 days earlier is -45. In this case, the bank is assumed to have posted less than 50 of collateral with the counterparty in the event of a default at time τ and its exposure is zero.
4. On a particular simulation trial, the value of outstanding transactions to the bank at time τ is -50 and the value 20 days earlier is -55. In this case, it is assumed that 55 of the bank's collateral is held by the counterparty 20 days before time τ and, in the event of a default at time τ, none of it is returned. The bank's exposure is therefore 5, the excess collateral it has posted.

20.2.2 Peak Exposure

In addition to calculating CVA, dealers usually calculate the *peak exposure* at the midpoint of each interval. This is a high percentile of the exposures given by the Monte Carlo simulation trials. Suppose for example that the percentile chosen is 97.5% and there are 10,000 Monte Carlo simulation trials. The peak exposure at time t_i^* is the 250th highest exposure recorded at that time. The *maximum peak exposure* is the maximum of the peak exposures across all the t_i^*.

There is a theoretical issue here (which is often ignored in practice). To calculate CVA, we simulate the behavior of market variables in the risk-neutral world and discount at the risk-free rate. (As explained in Chapter 7, it is correct to do this when valuation is the objective.) When we calculate peak exposure, we are carrying out a scenario analysis. We are in effect asking, "How bad can our exposure to the counterparty get in the future?" For this purpose, as explained in Chapter 7, we should in theory simulate the behavior of market variables in the real world, not the risk-neutral world.[4]

20.2.3 Downgrade Triggers

Netting and collateral agreements are important ways in which credit risk is reduced in bilaterally cleared derivatives transactions. Sometimes the credit support annex to an ISDA master agreement also includes a downgrade trigger. This is a clause stating that if the credit rating of one side falls below a certain level, the other side has the right to ask for more collateral or to terminate transactions.

AIG provides an example of the operation of a downgrade trigger. Many of AIG's transactions stated that AIG did not have to post collateral provided its credit rating remained above AA. However, once it was downgraded below AA, collateral was required. AIG was downgraded below AA by all three rating agencies on September 15, 2008. This led to collateral calls, which AIG was unable to meet. Its bankruptcy was avoided by a massive government bailout.

Other companies that have run into problems because of downgrade triggers include Enron, Xerox, and Ambac. Downgrade triggers do not provide protection to a company's counterparties from a big jump in a company's credit rating (for example, from AA to default). Also, downgrade triggers work well for a company's counterparties only when the company is making relatively little use of them. If a company has many downgrade triggers in its contracts, an AIG-type situation can arise where a downgrade below a specified credit rating can lead to huge cash demands on the company. If these cash demands cannot be met, an immediate bankruptcy follows.

20.3 The Impact of a New Transaction

When a new derivatives transaction is being negotiated by a dealer with a counterparty, its incremental effect on CVA may influence the terms offered. If the value of the new transaction is positively correlated with other transactions entered into by the dealer with the counterparty, the incremental effect on CVA is likely to be positive. If this correlation is negative, the new transaction can have the effect of reducing CVA.

Suppose, for example, that a dealer has one transaction outstanding with a counterparty: a five-year forward foreign currency transaction where the counterparty is buying the currency. If the counterparty then expresses interest in entering into a three-year forward foreign currency transaction, the competitiveness of the dealer's quote might depend on which side of the transaction the counterparty wants to take. If the counterparty wants to buy the currency, the new transaction would have the effect of increasing

[4] When the peak exposure is being calculated for times that are not far in the future, the difference between simulating in the real world and in the risk-neutral world is usually quite small.

CVA (making the transaction less attractive to the dealer). On the other hand, if the counterparty wants to sell the currency, it would be netted with the existing transaction and have the effect of reducing CVA (making the transaction more attractive to the dealer). These CVA/netting considerations suggest that when transactions are cleared bilaterally and are not fully collateralized, a company will sometimes get the most favorable quotes from a dealer with whom it already has transactions outstanding rather than from a dealer with whom it has done no previous business.

Calculating CVA is computationally intensive. Often dealers have hundreds or thousands of transactions outstanding with a counterparty.[5] Calculating the incremental effect of a new transaction on CVA by recomputing CVA is not usually feasible. Luckily there is a computationally efficient approach for calculating incremental CVA.

When the CVA calculations in Section 20.2 are carried out, the paths followed by market variables on each simulation trial and the value of the portfolio on each simulation trial are stored.[6] When a potential new transaction is being considered, its value at future times is calculated for the values of the market variables that were obtained on the simulation trials. This gives the incremental impact of the transaction on the future portfolio value for each of the last-used Monte Carlo simulation trials at all future times. The incremental impact on the exposure at each time for each Monte Carlo trial can then be calculated. From this, the incremental effect on the average exposure at each time can be calculated and equation (20.1) can then be used to calculate the incremental effect on CVA.

To illustrate this, suppose that a portfolio with a counterparty depends only on the price of gold and that the price of gold at time 2.5 years on the 545th simulation trial when CVA is calculated is $1,572, with the value of the portfolio to the dealer being $2.4 million. Assuming no collateral is posted, this is also the value of the exposure. If 2.5 years is the midpoint of the 20th time step, this means that v_{20} is the present value of $2.4 million received at time 2.5 years. We suppose that this is $2.3 million.

Suppose that, shortly after CVA is calculated, a potential new transaction with the counterparty, dependent on the price of gold, is contemplated. This transaction is valued at all times for the paths followed by the price of gold on all simulation trials. Suppose that, on the 545th simulation trial, it is calculated that the value of the new transaction will be –$4.2 million at time 2.5 years (when the price of gold is $1,572). This means that the portfolio value reduces from $2.3 million to –$1.9 million at time 2.5 years on the 545th simulation trial as a result of the new transaction. This reduces the exposure to zero so that the new v_{20} is zero. The new transaction therefore has the effect of reducing v_{20} by $2.3 million. Similar calculations are carried out for all simulation trials and all times.

---

[5] Lehman Brothers, for example, had about 1.5 million derivatives transactions with about 8,000 counterparties at the time of its failure.

[6] To be precise, the values of all market variables at the t_i^* and $t_i^* - c$ for $1 \leq i \leq n$ and the MTM portfolio values at these times are stored until the next time CVA is calculated (at which time they can be deleted).

Define Δv_i as the average change in v_i across all the simulation trials. The incremental effect on CVA of the new transaction is estimated as

$$\sum_{i=1}^{n}(1-R)q_i\Delta v_i$$

20.4 CVA Risk

A dealer has one CVA for each counterparty. These CVAs can themselves be regarded as derivatives. They are particularly complex derivatives. Indeed, the CVA for a counterparty is more complex than any of the transactions between the dealer and the counterparty because it is contingent on the net value of all the transactions between the dealer and the counterparty.

When CVA increases (decreases), the income reported by a derivatives dealer decreases (increases). For this reason, many dealers consider it prudent to try to hedge CVAs in the same way that they hedge other derivatives. This means that they must calculate the Greek letters (delta, gamma, vega, etc.) discussed in Chapter 8.

The variables affecting the v_i are market variables such as interest rates, exchange rates, commodity prices, and so on. Calculating Greek letters for these is liable to be computationally quite time consuming. For example, to calculate the delta of CVA with respect to an exchange rate, it is necessary to make a small change to the exchange rate and recompute CVA. A technique known as adjoint differentiation can be used to ease the computational burden.[7]

The variables affecting the q_i are the credit spreads of the counterparty for different maturities. From equation (20.3)

$$q_i = \exp\left(-\frac{s_{i-1}t_{i-1}}{1-R}\right) - \exp\left(-\frac{s_i t_i}{1-R}\right)$$

From equation (20.1)

$$\text{CVA} = \sum_{i=1}^{n}(1-R)q_i v_i$$

[7] See, for example, M. Giles and P. Glasserman, "Smoking Adjoints: Fast Monte Carlo Greeks," *Risk* 19, no. 1 (2006): 88–92; and M. Henrard, "Adjoint Algorithmic Differentiation: Calibration and the Implicit Function Theorem," *Journal of Computational Finance* 17, no. 4 (2014): 37–47.

Using a delta/gamma approximation, the change in CVA resulting from a small parallel shift, Δs, in the term structure of credit spreads (with all the market variables determining the v_i being assumed to remain fixed) is therefore

$$\Delta(\text{CVA}) = \sum_{i=1}^{n}\left[t_i \exp\left(-\frac{s_i t_i}{1-R}\right) - t_{i-1}\exp\left(-\frac{s_{i-1}t_{i-1}}{1-R}\right)\right] v_i \Delta s + \frac{1}{2(1-R)}\sum_{i=1}^{n}\left[t_{i-1}^2 \exp\left(-\frac{s_{i-1}t_{i-1}}{1-R}\right) - t_i^2\exp\left(-\frac{s_i t_i}{1-R}\right)\right] v_i (\Delta s)^2 \tag{20.5}$$

This can be calculated without difficulty once the v_i are known.

Basel III requires dealers to use this equation to incorporate the risks arising from changes in credit spreads into market risk capital calculations. Originally, risks arising from changes in the market variables affecting the v_i were not included in market risk capital calculations. This had the effect of discouraging dealers from hedging the risks associated with the v_i because the impact of hedging would be to increase rather than reduce the market risk assessed for capital purposes. This issue was addressed by the Basel Committee in 2015.[8]

20.5 Wrong-Way Risk

Up to now we have assumed that the probability of default is independent of the exposure. A situation where there is a positive dependence between the two, so that the probability of default by the counterparty tends to be high (low) when the dealer's exposure to the counterparty is high (low), is referred to as wrong-way risk. A situation where there is negative dependence, so that the probability of default by the counterparty tends to be high (low) when the dealer's exposure to the counterparty is low (high), is referred to as right-way risk.

A subjective evaluation of the amount of wrong-way or right-way risk in transactions with a counterparty requires a good knowledge of the counterparty's business, in particular the nature of the risks facing the business. It also requires knowledge of the transactions the counterparty has entered into with other dealers. The latter is difficult to know with any precision.

One situation in which wrong-way risk tends to occur is when a counterparty is using a credit default swap to sell protection to the dealer. (AIG is an example here.) When a dealer buys protection from the counterparty and the credit spread of the reference entity increases, the value of the protection to the dealer becomes positive. However, because the credit spreads of different companies tend to be correlated, it is likely that

[8] See Basel Committee on Banking Supervision, "Review of Credit Valuation Adjustment Risk Framework," October 1, 2015.

the credit spread of the counterparty has increased so that the counterparty's calculated probability of default has also increased. Similarly, right-way risk tends to occur when a counterparty is buying credit protection from the dealer.

A situation in which a company is speculating by entering into many similar trades with one or more dealers is likely to lead to wrong-way risk for these dealers. This is because the company's financial position and therefore its probability of default is likely to be affected adversely if the trades move against the company.

If a company enters into transactions with a dealer to partially hedge an existing exposure, there should in theory be right-way risk. This is because, when the transactions move against the counterparty, it will be benefiting from the unhedged portion of its exposure so that its probability of default should be relatively low.[9]

A simple way of dealing with wrong-way risk is to use what is termed the "alpha" multiplier to increase v_i in the version of the model where the v_i and q_i are assumed to be independent. The effect of this is to increase CVA by the alpha multiplier. Basel rules set alpha equal to 1.4, but allow banks to use their own models, with a floor for alpha of 1.2. This means that, at minimum, the CVA has to be 20% higher than that given by the model where the v_i and q_i are assumed to be independent. If a bank does not have its own model for wrong-way risk, it has to be 40% higher. Estimates of alpha reported by banks range from 1.07 to 1.10.

Some models have been developed to capture the dependence of the probability of default on the exposure. For example, Hull and White (2012) propose a simple model where the hazard rate at time t is a function of variables observable at that time.[10] Their parameter describing the extent of the dependence can be either estimated subjectively or estimated by relating past credit spreads for the counterparty to what the value of the current portfolio would have been in the past. Implementing the model involves relatively minor modifications to the calculations outlined in Section 20.2.

20.6 DVA

Debit value adjustment (DVA) is the mirror image of CVA.[11] Whereas CVA is the expected cost to the dealer from a possible default by the counterparty, DVA is the expected cost to the counterparty because the dealer might default. It is the counterparty's CVA. If DVA is a cost to the counterparty, it must be a benefit to the dealer. (This is because derivatives are zero sum games: the gain to one side always equals the loss to the other side.) The benefit arises from the fact that the dealer when it defaults avoids payments that would otherwise be required on outstanding derivatives. Accounting

[9] An exception could be when the counterparty is liable to run into liquidity problems. Although the assets being hedged have increased in value, the counterparty might be unable to post collateral when required. An example here is Ashanti Goldfields (see Business Snapshot 24.2).

[10] See J. Hull and A. White, "CVA and Wrong Way Risk," *Financial Analysts Journal* 68, no. 5 (September/October 2012): 58–69.

[11] DVA is sometimes also referred to as "debt value adjustment."

standards recognize both CVA and DVA. The book value of the derivatives outstanding with a counterparty to be calculated is

$$f_{\mathrm{nd}} - \mathrm{CVA} + \mathrm{DVA}$$

where f_{nd} is the value of the derivatives assuming neither side will default.

DVA can be calculated at the same time as CVA. Equation (20.1) gives DVA rather than CVA if R is the recovery rate of the counterparty in the event of a default by the dealer, v_i is the present value of the counterparty's exposure to the dealer at the midpoint of the ith interval, and q_i is the probability of a default by the dealer during the ith interval. The counterparty's net exposure to the dealer takes account of collateral posted by the dealer with the counterparty in the same way that the dealer's net exposure to the counterparty takes account of the collateral posted by the counterparty with the dealer.

One surprising effect of DVA is that when the credit spread of a derivatives dealer increases, DVA increases. This in turn leads to an increase in the reported value of the derivatives on the books of the dealer and a corresponding increase in its profits. Some banks reported several billion dollars of profits from this source in the third quarter of 2011. Regulators are uncomfortable with this and have excluded DVA gains and losses from the definition of common equity in the determination of regulatory capital.

20.7 Some Simple Examples

To illustrate the ideas presented in this chapter, we now present some simple examples.

20.7.1 Single Transaction with Positive Value

Suppose first that a dealer has a single derivatives transaction with a counterparty that is bound to have a positive value to the dealer and a negative value to the counterparty at all future times. (An example would be the situation where the dealer has bought an option from the counterparty.) We suppose that no collateral has to be posted by the counterparty and, for ease of exposition, assume that payoffs on the derivatives transaction occur only on its expiration date.

The exposure of the dealer to the counterparty at a future time is the value of the transaction at that time. The present value of the expected exposure at time t_i, which we have denoted by v_i, is therefore the present value of the expected value of the transaction at time t_i. Because we are assuming no payoffs before maturity, the present value of the expected value of the transaction at time t_i is always equal to its value today.

Equation (20.1) therefore becomes

$$\mathrm{CVA} = (1 - R)f_0 \sum_{i=1}^{n} q_i$$

where f_0 is the value of the derivative today, assuming no defaults. If f_0^* is the value after defaults are taken into account:

$$f_0^* = f_0 - \mathrm{CVA}$$

or

$$f_0^* = f_0 \left[1 - (1 - R) \sum_{i=1}^{n} q_i \right] \tag{20.6}$$

This means that the effect of defaults is to reduce the value of the derivative by a proportional amount equal to the cumulative risk-neutral probability of default during the life of the derivative times one minus the recovery rate.

Now consider an unsecured zero-coupon bond issued by the counterparty that promises \$1,000 at time T. Define B_0 as the value of the bond assuming no possibility of default and B_0^* as the actual value of the bond. Assuming that the bond ranks equally with the derivative in the event of a default, it will have the same recovery rate. Similarly to equation (20.6)

$$B_0^* = B_0 \left[1 - (1 - R) \sum_{i=1}^{n} q_i \right] \tag{20.7}$$

From equations (20.6) and (20.7)

$$\frac{f_0^*}{f_0} = \frac{B_0^*}{B_0} \tag{20.8}$$

If y is the yield on a risk-free zero-coupon bond maturing at time T and y^* is the yield on a zero-coupon bond issued by the counterparty that matures at time T, then $B_0 = e^{-yT}$ and $B_0^* = e^{-y^*T}$ so that equation (20.8) gives

$$f_0^* = f_0 e^{-(y^* - y)T} \tag{20.9}$$

This shows that the derivative can be valued by increasing the discount rate that is applied to the expected payoff by $y^* - y$.

Example 20.2

Consider a two-year uncollateralized over-the-counter option sold by a company with a value, assuming no possibility of default, of \$3. Suppose that two-year zero-coupon bonds issued by the company have a yield that is 1.5% greater than a similar risk-free zero-coupon bond. The value of the option is

$$3e^{-0.015 \times 2} = 2.91$$

or \$2.91.

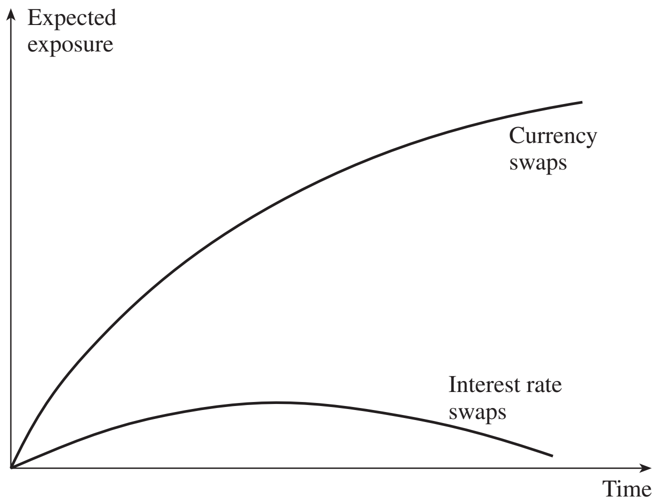


Figure 20.1 Expected Exposure on a Matched Pair of Offsetting Interest Rate Swaps and a Matched Pair of Offsetting Currency Swaps

20.7.2 Interest Rate Swaps vs. Currency Swaps

Consider next the situation where the dealer has entered into a pair of swaps that offset each other with two different counterparties. Figure 20.1 compares the total expected future exposure on the two transactions when they are (a) currency swaps and (b) interest rate swaps. The expected exposure on the interest rate swaps starts at zero, increases, and then decreases. By contrast, expected exposure on the currency swaps increases steadily with the passage of time. The main reason for the difference is that principals are exchanged at the end of the life of a currency swap and there is uncertainty about the exchange rate at that time. By contrast, toward the end of the life of the interest rate swap, very little is still to be exchanged. The impact of default risk for a dealer in currency swaps is therefore much greater than for a dealer in interest rate swaps. The q_i that the dealer calculates for counterparties in equation (20.1) are the same regardless of the transaction, but the v_i are on average greater for currency swaps.

20.7.3 Single Forward Transaction

For another example, we suppose that a dealer has a single forward transaction with a counterparty giving the dealer the right to buy an asset from the counterparty at time T for a price K. No collateral has to be posted. Suppose that the forward price of the asset today is F_0 (which is known) and at time t ($t \leq T$) it is F_t (which is unknown). As explained in Appendix C, the value of the transaction at time t is

$$(F_t - K)e^{-r(T-t)}$$

where r is the risk-free interest rate (assumed constant).

The exposure at time t is

$$\max[(F_t - K)e^{-r(T-t)}, 0] = e^{-r(T-t)} \max[(F_t - K), 0] \tag{20.10}$$

The present value of the exposure is therefore $e^{-r(T-t)}$ times the value of a derivative that pays off $\max[(F_t - K), 0]$ at time t. This is an option on a forward price. From Appendix E, the value of the derivative is

$$e^{-rt}[F_0N(d_1) - KN(d_2)]$$

where

$$d_1 = \frac{\ln(F_0/K) + \sigma^2 t/2}{\sigma\sqrt{t}}$$

and

$$d_2 = \frac{\ln(F_0/K) - \sigma^2 t/2}{\sigma\sqrt{t}}$$

and σ is the volatility of the forward price of the asset. It follows from equation (20.10) that the present value of the exposure at time t is

$$e^{-rT}[F_0N(d_1) - KN(d_2)]$$

The exposure, v_i, at a possible default time, t_i, is therefore

$$v_i = e^{-rT}[F_0N(d_{1,i}) - KN(d_{2,i})] \tag{20.11}$$

where

$$d_{1,i} = \frac{\ln(F_0/K) + \sigma^2 t_i/2}{\sigma\sqrt{t_i}}$$

and

$$d_{2,i} = \frac{\ln(F_0/K) - \sigma^2 t_i/2}{\sigma\sqrt{t_i}}$$

Example 20.3

Suppose that a bank has entered into a forward contract to buy 1 million ounces of gold from a mining company in two years for \$1,500 per ounce. The current forward price for the contract is \$1,600 per ounce. The probability of the company defaulting during the first year is 2% and the probability of the company defaulting during the second year is 3%. Defaults are assumed to happen at the midpoints of the years. The risk-free rate is 5% per annum. The financial institution anticipates a 30% recovery in the event of a default. The volatility of the forward price of gold when the forward contract expires in two years is 20%.

In this case, from equation (20.11)

$$v_1 = e^{-0.05\times 2}[1{,}600N(d_{1,1}) - 1{,}500N(d_{2,1})]$$

where

$$d_{1,1} = \frac{\ln(1{,}600/1{,}500) + 0.2^2 \times 0.5/2}{0.2\sqrt{0.5}} = 0.5271$$

$$d_{2,1} = \frac{\ln(1{,}600/1{,}500) - 0.2^2 \times 0.5/2}{0.2\sqrt{0.5}} = 0.3856$$

so that $v_1 = 132.38$. Also, from equation (20.11)

$$v_2 = e^{-0.05\times 2}[1{,}600N(d_{1,2}) - 1{,}500N(d_{2,2})]$$

where

$$d_{1,2} = \frac{\ln(1{,}600/1{,}500) + 0.2^2 \times 1.5/2}{0.2\sqrt{1.5}} = 0.3860$$

$$d_{2,2} = \frac{\ln(1{,}600/1{,}500) - 0.2^2 \times 1.5/2}{0.2\sqrt{1.5}} = 0.1410$$

so that $v_2 = 186.65$.

Other variables are: $q_1 = 0.02$, $q_2 = 0.03$, and $R = 0.3$ so that

$$\text{CVA} = (1 - 0.3) \times (0.02 \times 132.38 + 0.03 \times 186.65) = 5.77$$

The no-default value of the forward contract is $(1{,}600 - 1{,}500)e^{-2\times 0.05} = 90.48$ or \$90.48. The value after the possibility of defaults has been taken into account is therefore

$$90.48 - 5.77 = 84.71$$

or \$84.71.

This example can be extended to the situation where the mining company can default more frequently or to calculate DVA (see Problems 20.13 and 20.14).

Summary

A derivatives dealer's credit value adjustment (CVA) for its bilaterally cleared derivatives with a counterparty is the present value of the expected future loss from the possibility of the counterparty defaulting. A derivatives dealer typically has systems in place for calculating the credit value adjustments (CVAs) for all the counterparties with which it has bilaterally cleared derivatives. The systems must simulate the market variables underlying outstanding transactions with each counterparty so that the expected net exposure at future times conditional on a default can be estimated. The calculations take account of netting and collateral agreements. The simplest assumption is that the probability of default is independent of the exposure. Procedures can be developed to incorporate wrong-way risk (where there is positive dependence between the two) and right-way risk (where there is negative dependence between the two).

CVAs are complex derivatives, and many dealers hedge them in the same way that they hedge other derivatives. There are two types of risks associated with CVAs: the risk that the counterparty's credit spread will change and the risk that there will be movements in the values of the underlying market variables. Once a CVA has been calculated, it is relatively easy to quantify the first risk, and the original Basel III regulations require dealers to keep market risk capital for this risk. Calculating the sensitivity of CVA to movements in the underlying market variables is more difficult, but regulatory capital has been proposed for this risk as well.

A dealer's debit value adjustment (DVA) for its bilaterally cleared derivatives with a counterparty is the present value of the expected future loss to the counterparty from the possibility of the dealer defaulting. This is a gain to the dealer. It can be calculated at the same time as CVA.

Further Reading

Basel Committee on Banking Supervision. "Basel III: A Global Regulatory Framework for More Resilient Banks and Banking Systems," www.bis.org/publ/bcbs189dec2010.pdf, December 2010.

Canabarro, E., and D. Duffie. "Measuring and Marking Counterparty Risk." Chapter 9 in *Asset/Liability Management for Financial Institutions*, edited by L. Tilman. New York: Institutional Investor Books, 2003.

Gregory, J. *Counterparty Credit Risk: The New Challenge for Financial Markets*. 2nd ed. Hoboken, NJ: John Wiley & Sons, 2012.

Hull, J., and A. White. "The Impact of Default Risk on the Prices of Options and Other Derivative Securities." *Journal of Banking and Finance* 19 (1995): 299–322.

Pengelley, M. "CVA Melee." *Risk* 24, no. 2 (2011): 37–39.

Picault, E. "Calculating and Hedging Exposure, CVA, and Economic Capital for Counterparty Credit Risk." In *Counterparty Credit Risk Modeling*, edited by M. Pykhtin. London: Risk Books, 2005.

Sokol, A. "A Practical Guide to Monte Carlo CVA." Chapter 14 in *Lessons from the Crisis*, edited by A. Berd. London: Risk Books, 2010.

Practice Questions and Problems (Answers at End of Book)

20.1 Explain why a new transaction by a bank with a counterparty can have the effect of increasing or reducing the bank's credit exposure to the counterparty.

20.2 A company offers to post its own equity as collateral. How would you respond?

20.3 Suppose that a financial institution has two derivatives transactions outstanding with different counterparties, X and Y. Which of the following is/are true?
(a) The total expected exposure in one year on the two transactions is the sum of the expected exposure on the transaction with X and the expected exposure on the transaction with Y.
(b) The total present value of the cost of defaults is the sum of the present value of the cost of defaults on the transaction with X plus the present value of the cost of defaults on the transaction with Y.
(c) The 95th percentile for the total exposure in one year on both transactions is the sum of the 95th percentile for the exposure in one year on the transaction with X and the 95th percentile for the exposure in one year on the transaction with Y.

Explain your answers.

20.4 "In the absence of collateral and other transactions between the parties, a long forward contract subject to credit risk is a combination of a short position in a no-default put and a long position in a call subject to credit risk." Explain this statement.

20.5 Suppose that the spread between the yield on a three-year riskless zero-coupon bond and a three-year zero-coupon bond issued by a corporation is 120 basis points. By how much do standard option pricing models such as Black–Scholes–Merton overstate the value of a three-year option sold by the corporation? Assume there is only this one transaction between the corporation and its counterparty and no collateral is posted.

20.6 Can the existence of default triggers increase default risk? Explain your answer.

20.7 Give two examples of when (a) wrong-way risk and (b) right-way risk can be expected to be observed.

20.8 Explain the term "cure period."

20.9 "Netting means that CVA cannot be calculated on a transaction-by-transaction basis." Explain this statement.

20.10 "DVA can improve the bottom line when a bank is experiencing financial difficulties." Explain why this is so.

20.11 What part of CVA risk is considered a component of market risk in Basel III?

20.12 A CSA between a dealer and one of its counterparties states that collateral has to be posted by both sides with zero thresholds. If the cure period is assumed to be 15 days, under what circumstances will the dealer's CVA model lead to losses?

Further Questions

20.13 Extend Example 20.3 to calculate CVA when default can happen in the middle of each month. Assume that the default probability during the first year is 0.001667 per month and the default probability during the second year is 0.0025 per month.

20.14 Calculate DVA for the bank in Example 20.3. Assume that the bank can default in the middle of each month and that the default probability is 0.001 per month for the two years. Assume that the recovery rate for the counterparty when the bank defaults is 40%.

20.15 Consider a European call option on a non-dividend-paying stock where the stock price is \$52, the strike price is \$50, the risk-free rate is 5%, the volatility is 30%, and the time to maturity is one year. Answer the following questions assuming no recovery in the event of default, that the probability of default is independent of the option valuation, no collateral is posted, and no other transactions between the parties are outstanding.

(a) What is the value of the option, assuming no possibility of a default?

(b) What is the value of the option to the buyer if there is a 2% chance that the option seller will default at maturity?

(c) Suppose that, instead of paying the option price up front, the option buyer agrees to pay the option price (with accumulated interest) at the end of the option's life. By how much does this reduce the cost of defaults to the option buyer in the case where there is a 2% chance of the option seller defaulting?

(d) If in case (c) the option buyer has a 1% chance of defaulting at the end of the life of the option, what is the default risk to the option seller? Discuss the two-sided nature of default risk in the case and the value of the option to each side.

20.16 Suppose that the spread between the yield on a three-year riskless zero-coupon bond and a three-year zero-coupon bond issued by a bank is 210 basis points. The Black–Scholes–Merton price of an option is \$4.10. How much should you be prepared to pay for it if you buy it from a bank?

Chapter 21

Credit Value at Risk

The value at risk (VaR) measure is central to the determination of regulatory capital for credit risk and to much of the credit risk management carried out internally by both financial and nonfinancial corporations. This chapter covers alternative approaches for calculating credit risk VaR.

Credit risk VaR is defined similarly to market risk VaR. It is the credit risk loss over a certain time period that will not be exceeded with a certain confidence level. Some credit risk VaR models consider only losses from defaults; others consider losses from downgrades or credit spread changes as well as from defaults.

Banks calculate credit risk VaR to determine both regulatory capital and economic capital. The regulatory capital requirements for credit risk are discussed in Chapters 15 and 16. Economic capital, which is discussed in Chapter 26, is a financial institution's own estimate of the capital it requires for the risks it is taking and is used to calculate return on capital measures for its business units. Sometimes the VaR model that a bank chooses to use to determine credit risk economic capital is different from the one it is required to use for the determination of regulatory capital.

Value at risk is a scenario analysis measure. As explained in Chapter 7, the probability of default estimates used to develop alternative horizon-date scenarios should therefore be real-world estimates. Risk-neutral estimates should then be used to value the underlying portfolio on the VaR horizon date. It will be recalled from Chapter 19 that real-world default probabilities are lower than risk-neutral default probabilities.

The time horizon for credit risk VaR is often longer than that for market risk VaR. Market risk VaR is often calculated with a one-day time horizon. Credit risk VaR, for instruments that are not held for trading, is usually calculated with a one-year time horizon. Historical simulation (see Chapter 13) is the main tool used to calculate market risk VaR, but a more elaborate model is usually necessary to calculate credit risk VaR.

A key aspect of any credit risk VaR model is credit correlation. Defaults for different companies do not happen independently of each other. During an economic downturn, most companies are adversely affected and become more likely to default. When the economy is faring well, they are favorably affected and less likely to default. This relationship between default rates and economic factors is a major reason for credit correlation. If credit correlation increases (as it tends to do in stressed economic conditions), the risk for a financial institution with a portfolio of credit exposures increases.

21.1 Ratings Transition Matrices

The methods used by financial institutions for calculating credit VaR often involve ratings transition matrices. These are matrices showing the probability of a company migrating from one rating category to another during a certain period of time. They are based on historical data. The rating categories can be either those used internally by the financial institution or those produced by rating agencies such as Moody's, S&P, and Fitch. Table 21.1 shows the one-year transition matrix produced by Moody's, based on following the performance of all the companies that Moody's rated between 1970 and 2016. For example, it shows that a company that starts with a rating of A has a 90.90% probability of still being rated A at the end of a year. There is a 2.64% chance that it will be upgraded to Aa by the end of the year, a 5.67% chance that it will be downgraded to Baa, and so on. The probability that it will default during the year is only 0.06%, or 6 chances in 10,000.

If we assume that the rating change in one period is independent of that in another period, Table 21.1 can be used to calculate a transition matrix for periods other than one

Table 21.1 One-Year Ratings Transition Matrix, 1970–2016, with Probabilities Expressed as Percentages and Adjustments for Transition to the WR (without rating) Category

Initial Rating	Rating at Year-End								
	Aaa	**Aa**	**A**	**Baa**	**Ba**	**B**	**Caa**	**Ca-C**	**Default**
Aaa	90.94	8.36	0.59	0.08	0.02	0.00	0.00	0.00	0.00
Aa	0.87	89.68	8.84	0.45	0.07	0.04	0.02	0.00	0.02
A	0.06	2.64	90.90	5.67	0.51	0.12	0.04	0.01	0.06
Baa	0.04	0.16	4.44	90.16	4.09	0.75	0.17	0.02	0.18
Ba	0.01	0.05	0.47	6.66	83.03	7.90	0.78	0.12	0.99
B	0.01	0.03	0.16	0.51	5.32	82.18	7.39	0.61	3.79
Caa	0.00	0.01	0.03	0.11	0.46	7.82	78.52	3.30	9.75
Ca-C	0.00	0.00	0.07	0.00	0.80	3.19	11.41	51.28	33.24
Default	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	100.00

SOURCE: Moody's.

Table 21.2 Five-Year Ratings Transition Matrix Calculated from Table 21.1 with Probabilities Expressed as Percentages

Initial	Rating at Year-End								
Rating	**Aaa**	**Aa**	**A**	**Baa**	**Ba**	**B**	**Caa**	**Ca–C**	**Default**
Aaa	62.75	28.14	7.58	1.16	0.20	0.07	0.02	0.00	0.03
Aa	2.96	60.27	29.98	5.27	0.82	0.33	0.13	0.01	0.20
A	0.40	9.00	65.74	19.73	3.19	1.08	0.32	0.04	0.55
Baa	0.18	1.45	15.45	63.40	12.25	4.33	1.18	0.13	1.68
Ba	0.06	0.39	3.57	19.72	43.71	19.78	5.10	0.55	7.16
B	0.04	0.16	0.94	3.83	13.22	43.21	16.56	1.65	20.39
Caa	0.01	0.06	0.26	0.94	3.46	17.61	34.25	3.53	39.88
Ca–C	0.00	0.03	0.21	0.51	2.09	7.30	12.50	4.53	72.82
Default	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	100.00

year. For example, a transition matrix for two years can be calculated by multiplying the matrix by itself. The five-year transition matrix, which is shown in Table 21.2, is calculated as the fifth power of the matrix in Table 21.1. Table 21.2 shows, not surprisingly, that the probability of a company keeping the same credit rating over five years is much lower than it is over one year, and default probabilities over five years are much higher than over one year.

The credit rating change over a period less than a year is not quite so easy to calculate. For example, estimating a transition matrix for six months involves taking the square root of the matrix in Table 21.1; estimating the transition matrix for three months involves taking the fourth root of the matrix; and so on. The calculation methodology is explained in Appendix J at the end of the book and software for performing the calculations is on the author's website: www-2.rotman.utoronto.ca/∼hull/riskman.

Table 21.3 shows the rating transition calculated for a period of one month from the data in Table 21.1. As might be expected, the probability that a company's credit rating will stay the same over a period of one month is very high.

Table 21.3 One-Month Ratings Transition Matrix Calculated from Table 21.1 with Probabilities Expressed as Percentages

Initial	Rating at Year-End								
Rating	**Aaa**	**Aa**	**A**	**Baa**	**Ba**	**B**	**Caa**	**Ca–C**	**Default**
Aaa	99.21	0.77	0.02	0.01	0.00	0.00	0.00	0.00	0.00
Aa	0.08	99.08	0.81	0.02	0.00	0.00	0.00	0.00	0.00
A	0.00	0.24	99.19	0.52	0.04	0.01	0.00	0.00	0.00
Baa	0.00	0.01	0.40	99.12	0.39	0.05	0.01	0.00	0.01
Ba	0.00	0.00	0.03	0.63	98.42	0.78	0.04	0.01	0.07
B	0.00	0.00	0.01	0.03	0.53	98.32	0.75	0.06	0.30
Caa	0.00	0.00	0.00	0.01	0.02	0.79	97.94	0.41	0.82
Ca–C	0.00	0.00	0.01	0.00	0.08	0.32	1.42	94.55	3.62
Default	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	100.00

The assumption that the credit rating change in one period is independent of that in another period is not exactly true. If a company has been downgraded recently, it is more likely to be downgraded again in the next short period of time.[1] (This phenomenon is sometimes referred to as ratings momentum.) However, the independence assumption is not too unreasonable for many purposes.[2]

21.2 Vasicek's Model

The Basel II internal-ratings-based (IRB) capital requirements for credit risk in the banking book are based on Vasicek's Gaussian copula model (see Sections 11.6 and 15.8). This is a way of calculating high percentiles of the distribution of the default rate for a portfolio of loans. As in Chapter 11, we define WCDR(T, X) as the Xth percentile of the default rate distribution during a period of length T. (It will be recalled that WCDR denotes worst case default rate.) Vasicek's model relates WCDR(T, X) to the probability of default in time T, PD, and a parameter, ρ, describing credit correlation. The formula, which is proved in Section 11.6, is

$$\text{WCDR}(T, X) = N\left(\frac{N^{-1}(\text{PD}) + \sqrt{\rho}N^{-1}(X)}{\sqrt{1-\rho}}\right) \tag{21.1}$$

For an individual loan, if EAD is the exposure at default and LGD is the loss given default, the Xth percentile of the loss distribution is

$$\text{WCDR}(T, X) \times \text{EAD} \times \text{LGD}$$

A result by Gordy (2003) enables us to extend this in a one-factor world.[3] If we have a large portfolio of n loans where each loan is a small part of the total portfolio, the Xth percentile of the loss distribution is approximately

$$\sum_{i=1}^{n} \text{WCDR}_i(T, X) \times \text{EAD}_i \times \text{LGD}_i \tag{21.2}$$

where $\text{WCDR}_i(T, X)$, EAD_i, and LGD_i are the values of WCDR, EAD, and LGD for the ith loan in the portfolio.

---

[1] For a discussion of this, see E. Altman and D. Kao, "The Implications of Corporate Bond Rating Drift," *Financial Analysts Journal* (May–June 1992): 64–75, and D. Lando and T. Skodeberg, "Analyzing Rating Transitions and Rating Drift with Continuous Observations," *Journal of Banking and Finance* 26 (2002): 423–444.

[2] When the five-year transition matrix in Table 21.2 is compared with the actual five-year transition matrix published by Moody's, there are some differences. For example, companies in the top three rating categories are more likely to default or suffer big downgrades. This is consistent with ratings momentum.

[3] See M. Gordy, "A Risk-Factor Model Foundation for Ratings-Based Capital Rules," *Journal of Financial Intermediation* 12, no. 3 (July 2003): 199–233.

As explained in Chapter 15, regulatory capital for the banking book is set equal to the expression in equation (21.2) with T equal to one year and X equal to 99.9%. Sometimes the expression under the summation sign in equation (21.2) is multiplied by a maturity adjustment factor MA_i, to allow for the fact that, if it lasts longer than one year, the ith loan might deteriorate without defaulting. In the foundation IRB approach, banks estimate PDs, but EADs, LGDs, and MAs are set by the Basel II rules. In the advanced IRB approach, banks estimate PDs, EADs, LGDs, and MAs. But in all cases the parameter ρ is set by Basel II rules.

When Vasicek's model is used to determine economic capital, banks are free to make their own estimates of ρ. Structural models of the type discussed in Section 19.8 can be used to show that ρ for two companies should be roughly equal to the correlation between the returns on the assets of the companies.[4] As an approximation, this is the same as the correlation between the returns on their equities. One way of determining ρ for a portfolio of exposures to companies is therefore to calculate the average correlation between the returns on equities for the companies. If the companies are not publicly traded, the average correlation between representative publicly traded companies can be used instead. Another approach is the maximum likelihood approach described in Section 11.6.2.

As pointed out in Chapter 11, Vasicek's model has the disadvantage that it incorporates very little tail correlation. Replacing the Gaussian copula model with another copula model (see Section 11.6.3 for some alternatives) can remedy this.

21.3 Credit Risk Plus

In 1997, Credit Suisse Financial Products proposed a methodology for calculating VaR that it termed Credit Risk Plus.[5] It involves analytic approximations that are well established in the insurance industry.

Suppose that a financial institution has n loans of a certain type and the probability of default for each loan during a year is q so that the expected number of defaults is qn. Assuming that default events are independent, the probability of m defaults is given by the binomial distribution as

$$\frac{n!}{m!(n-m)!}q^m(1-q)^{n-m}$$

If q is small and n large, this can be approximated by the Poisson distribution:

$$\text{Prob}(m \text{ defaults}) = \frac{e^{-qn}(qn)^m}{m!} \tag{21.3}$$

[4] See for example, J. Hull, M. Predescu, and A. White, "The Valuation of Correlation Dependent Derivatives Using a Structural Model," *Journal of Credit Risk* 6, no. 3 (Fall 2010): 99–132.

[5] See Credit Suisse Financial Products, "Credit Risk Management Framework" (October 1997).

This equation is approximately true even if the probability of default is not the same for each loan, provided that all the probabilities of default are small and q equals the average probability of default over the next year for the loans in the portfolio.

In practice, we are uncertain about what the default rate, q, will be during the next year. As shown in Table 11.6, default rates vary greatly from year to year. A convenient assumption is that qn, the expected number of defaults, has a gamma distribution with mean μ and standard deviation σ.

The Poisson distribution in equation (21.3) then becomes a negative binomial distribution

$$\text{Prob}(m \text{ defaults}) = p^m(1-p)^\alpha \frac{\Gamma(m+\alpha)}{\Gamma(m+1)\Gamma(\alpha)} \tag{21.4}$$

where $\alpha = \mu^2/\sigma^2$, $p = \sigma^2/(\mu + \sigma^2)$, and $\Gamma(x)$ is the gamma function.[6]

Table 21.4 shows the probability distribution for the number of defaults when $\mu = 4$ for various values of σ. The first column ($\sigma = 0$) is calculated from equation (21.3). The remaining columns are calculated from equation (21.4). It can be seen that as σ tends to zero, the negative binomial distribution in equation (21.4) tends to the same probability distribution for the number of defaults as the Poisson distribution. As σ increases, the probability of an extreme outcome involving a large number of defaults increases. Suppose that the loss from a default is constant. As indicated in the final row of the table, when $\sigma = 0$ so that there is no uncertainty about the default rate, the 99.9% VaR is calculated as the loss when there are 11 defaults. When $\sigma = 10$ so that there is a great deal of uncertainty about the default rate, the 99.9% VaR is considerably larger. It is the loss when there are 98 defaults.

A more flexible way of building uncertainty into the default rate is by using Monte Carlo simulation. Typically there are several different categories of exposures with different but correlated default rates. One way of proceeding is as follows:

1. Sample an overall default rate. This could be a random sample from data such as that in Table 11.6.
2. Develop a model relating the default rate for each category of obligors to the overall default rate. A simple model is obtained by regressing default rates for each category of obligors against the overall default rate.
3. Sample a number of defaults for each category of obligors.
4. Sample a loss given default for each default in each category.
5. Calculate the total loss from the defaults.
6. Repeat steps 1 to 5 many times to build up a probability distribution for the total loss.
7. Calculate the required VaR from the total loss probability distribution.

[6] In Excel, the GAMMALN function returns the natural logarithm of the gamma function, $\Gamma(x)$. When x is an integer $\Gamma(x) = (x-1)!$

Table 21.4 Probability Distribution for Number of Defaults When the Expected Number of Defaults Is 4.0.

Number of Defaults	Standard Deviation of Default Rate						
	0	**0.1**	**0.5**	**1.0**	**2.0**	**5.0**	**10.0**
0	0.0183	0.0184	0.0207	0.0281	0.0625	0.2814	0.5938
1	0.0733	0.0734	0.0777	0.0901	0.1250	0.1553	0.0913
2	0.1465	0.1466	0.1486	0.1531	0.1563	0.1098	0.0509
3	0.1954	0.1952	0.1923	0.1837	0.1563	0.0833	0.0353
4	0.1954	0.1951	0.1895	0.1746	0.1367	0.0653	0.0268
5	0.1563	0.1561	0.1516	0.1396	0.1094	0.0523	0.0214
6	0.1042	0.1041	0.1026	0.0978	0.0820	0.0423	0.0177
7	0.0595	0.0596	0.0603	0.0614	0.0586	0.0346	0.0150
8	0.0298	0.0298	0.0315	0.0353	0.0403	0.0285	0.0129
9	0.0132	0.0133	0.0148	0.0188	0.0269	0.0236	0.0113
10	0.0053	0.0053	0.0064	0.0094	0.0175	0.0196	0.0099
11	0.0019	0.0019	0.0025	0.0045	0.0111	0.0163	0.0088
12	0.0006	0.0007	0.0009	0.0020	0.0069	0.0137	0.0079
13	0.0002	0.0002	0.0003	0.0009	0.0043	0.0115	0.0071
14	0.0001	0.0001	0.0001	0.0004	0.0026	0.0096	0.0064
15	0.0000	0.0000	0.0000	0.0001	0.0016	0.0081	0.0058
16	0.0000	0.0000	0.0000	0.0001	0.0009	0.0068	0.0053
17	0.0000	0.0000	0.0000	0.0000	0.0005	0.0058	0.0048
18	0.0000	0.0000	0.0000	0.0000	0.0003	0.0049	0.0044
19	0.0000	0.0000	0.0000	0.0000	0.0002	0.0041	0.0041
20	0.0000	0.0000	0.0000	0.0000	0.0001	0.0035	0.0038
99.9 percentile of number of defaults	11	11	12	13	17	39	98

There are a number of ways in which the Monte Carlo simulation can be refined. An analysis of the data in Table 11.6 shows that the default rate in one year is not independent of the default rate in the previous year. This indicates that randomly sampling a default rate from a table such as Table 11.6 to determine next year's default rate may not be the best approach. It may be preferable to develop a model that relates the default rate in one year to the default rate in the previous year or other economic factors observable in the previous year.

Note that the uncertainty about the default rate plays a key role in the analysis. Without this uncertainty, there is no default correlation and, as indicated in Table 21.4, there is very little chance of a large number of defaults. As the uncertainty about the default rate increases, default correlation increases and a large number of defaults becomes more likely. The default correlation arises because all companies have the same default rate, which can be high or low. Without default correlation, the loss probability distribution is fairly symmetrical. With default correlation, it is positively skewed, as indicated in Figure 21.1.

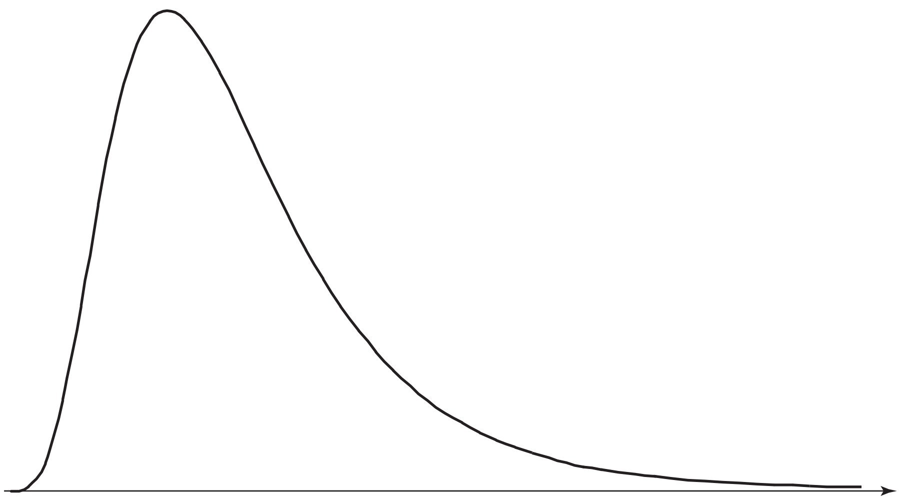


Figure 21.1 General Shape of Probability Distribution of Default Losses

21.4 Creditmetrics

Vasicek's model and Credit Risk Plus estimate the probability distribution of losses arising from defaults. The impact of downgrades is not considered.[7] We now consider the CreditMetrics model, which is designed so that it can take account of downgrades as well as defaults. This model was proposed by JPMorgan in 1997. It is based on a rating transition matrix such as the one shown in Table 21.1. The ratings can be the internal ratings used by the bank (derived from its own historical data) or those produced by rating agencies.

Calculating a one-year credit VaR for a portfolio of transactions with many counterparties involves carrying out a Monte Carlo simulation. On each simulation trial, the credit ratings of all counterparties at the end of one year are determined. (The procedure for doing this will be explained shortly.) The credit loss for each counterparty is then calculated. If the end-of-year credit rating for a counterparty is not default, the calculation of the credit loss is given by valuing all the transactions with the counterparty at the one-year point. If the end-of-year credit rating is default, the credit loss is the exposure at default times one minus the recovery rate.[8]

The calculations require the term structure of credit spreads for each rating category. A simple assumption is that it is the same as that observed in the market today. An alternative assumption is that there is a credit spread index that has a certain probability distribution and that all credit spreads are linearly dependent on the index.

[7] However, the maturity adjustment factor does allow for downgrades in the regulatory implementation of Vasicek's model.

[8] The recovery rate is typically sampled from an empirical distribution.

When transactions with a counterparty involve derivatives, an extension of the CVA calculations described in Chapter 20 is necessary. Recall that, for each counterparty, a CVA calculation involves dividing the time between today and the end of the life of the longest transaction with the counterparty into n intervals. We defined q_i as the probability of default during the ith time interval, v_i is the present value of the expected net exposure at the midpoint of the interval after collateral has been taken into account, and R is the recovery rate. The expected cost of defaults from one year onward, as seen at time zero, is

$$\sum_{i=j}^{n}(1-R)q_i v_i$$

where the jth interval starts at time one year.

On each CreditMetrics simulation trial where there is no default during the first year, the term structure of credit spreads at the one-year point gives the probability of default in each interval from year one onward using equation (20.3). Suppose that q_i^* is the probability of default for the ith interval ($i \geq j$) on a particular simulation trial. The credit loss for that trial is

$$\sum_{i=j}^{n}(1-R)(q_i^* - q_i)v_i \tag{21.5}$$

Note that, if the counterparty's credit rating improves during the year, this credit loss is likely to be negative.

If there is a default during the first year on a particular simulation trial, the timing of the default is sampled and the exposure at default determined. This is multiplied by one minus the recovery rate to determine the loss.

The results of the CreditMetrics Monte Carlo simulation determine a probability distribution for total credit losses arising from defaults and downgrades across all counterparties. The required value at risk can be calculated from this distribution.

21.4.1 The Correlation Model

In sampling to determine credit losses, the credit rating changes for different counterparties are not assumed to be independent. A Gaussian copula model is used to construct a joint probability distribution of rating changes. (See Section 11.5 for a discussion of copula models.) The copula correlation between the rating transitions for two companies is typically set equal to the correlation between their equity returns. (The capital asset pricing model discussed in Chapter 1 provides a one-factor model for equity returns. As described in Section 11.4.2, this can be used to determine an internally consistent correlation matrix.)

Suppose that we are simulating the rating change of an A-rated and a B-rated company over a one-year period using the transition matrix in Table 21.1. Suppose that the correlation between the equity returns of the two companies is 0.2. On each simulation trial we would sample two variables x_A and x_B from standard normal distributions so that their correlation is 0.2. The variable x_A determines the new rating of the A-rated company and variable x_B determines the new rating of the B-rated company. From Table 21.1, the probability of an A-rated company moving to Aaa, Aa, A,… is 0.0006, 0.0264, 0.9090, and so on. Because

$$N^{-1}(0.0006) = -3.2389$$

$$N^{-1}(0.0006 + 0.0264) = -1.9268$$

$$N^{-1}(0.0006 + 0.0256 + 0.9090) = 1.5220$$

the A-rated company gets upgraded to Aaa if $x_A < -3.2389$, it becomes Aa-rated if $-3.2389 < x_A < -1.9268$, it stays A-rated if $-1.9268 < x_A < 1.5220$, and so on. Table 21.1 also shows that the probability of a B-rated company moving to Aaa, Aa, A,… is 0.0001, 0.0003, 0.0016, and so on. Because

$$N^{-1}(0.0001) = -3.7190$$

$$N^{-1}(0.0001 + 0.0003) = -3.3528$$

$$N^{-1}(0.0001 + 0.0003 + 0.0016) = -2.8782$$

the B-rated becomes Aaa-rated if $x_B < -3.7190$, it becomes Aa-rated if $-3.7190 < x_B < -3.3528$, it becomes A-rated if $-3.3528 < x_B < -2.8782$, and so on. The A-rated company defaults if $x_A > N^{-1}(0.9994)$, that is, when $x_A > 3.2389$. The B-rated company defaults when $x_B > N^{-1}(0.9621)$, that is, when $x_B > 1.7756$. This example is illustrated in Figure 21.2.

It is interesting to note that, if both the CreditMetrics and the Credit Risk Plus models were based on the same set of assumptions, they should in theory predict the same probability distribution for losses over the long term. It is the timing of losses that is different. Suppose, for example, that you hold a certain loan in your portfolio. In year 1, the borrower gets downgraded from A to BBB; in year 2, it gets downgraded from BBB to B; in year 3 it defaults. You could assume that there are no losses in years 1 and 2 and calculate loss in year 3 (the Credit Risk Plus approach). Alternatively, you can calculate separate revaluation losses in years 1, 2, and 3 (the CreditMetrics approach). The losses under the second approach should in theory add up to the losses under the first approach.

21.5 Credit Spread Risk

The values of most credit-sensitive products in the trading book depend critically on credit spreads. Calculating VaR or ES for portfolios including these products therefore involves examining potential credit spread changes. One possibility is to use historical simulation to calculate a one-day 99% VaR. This one-day 99% VaR can then be scaled up by $\sqrt{10}$ to obtain a 10-day 99% VaR. It is first necessary to collect credit spread changes for each company to which the bank is exposed over the last 500 days (or some

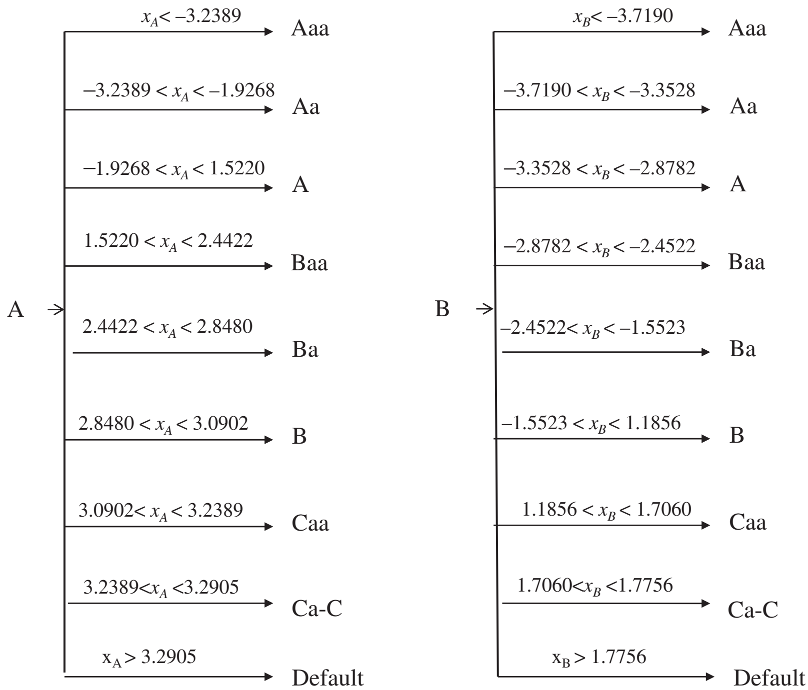


Figure 21.2 The CreditMetrics Correlation Model: Transition of A-Rated and B-Rated Companies to a New Rating after One Year (x_A and x_B are sampled from standard normal distributions with the correlation equal to the correlation between the equity returns of A and B)

other period). The first historical simulation scenario assumes that the percentage credit spread changes for all companies are the same as they were on the first day; the second scenario assumes that they are the same as they were on the second day; and so on. However, there are problems with this approach. If the company is alive today, it did not default in the past and the calculations carried out therefore assume no probability of default in the future.[9] Another problem is that credit spreads are not updated daily for all companies and so the historical data may not be of high quality.

An alternative approach is to use a version of the CreditMetrics approach. A 10-day rating transitions matrix is calculated as indicated in Section 21.1. The ratings transition matrix defines the probability of a company moving from one rating category to another during a 10-day period. The historical data on credit rating changes define a probability distribution for the credit spread changes associated with each rating category at the end of 10 days. A Monte Carlo simulation is carried out. On each trial, the transition matrix is sampled from to determine whether each company stays in the same rating category, changes to a different rating category, or defaults. The probability distribution of credit spread changes is also sampled from to determine the credit spread at the end of the

[9] One way of dealing with this is to assume that default takes place when the credit spread exceeds a certain level. Another is to separately consider "jump-to-default" risk, which is the approach taken by regulators.

10 days associated with each rating category. This enables the value of the portfolio at the end of 10 days to be calculated on each trial and a VaR to be obtained.

Credit correlation can be introduced in two ways. First, a Gaussian copula model can be used to correlate the credit rating changes of different companies, as explained in Section 21.4.1. Second, the credit spread changes for different rating categories can be assumed to be perfectly (or nearly perfectly) correlated so that when the spreads for A-rated instruments move up, those for instruments with other ratings do as well.

Example 21.1

Consider a company that owns a single two-year zero-coupon bond with a principal of \$1,000. Suppose that the risk-free rate is 3% and the current credit spread is 200 basis points so that the bond yield is 5%. (Rates are expressed with annual compounding.) The current price of the bond is $1{,}000/1.05^2 = \$907.03$. Suppose that the bond's current rating is Ba and, approximating the numbers in Table 21.3, that during the next month there is a 0.6% chance that it will increase to Baa, a 98.5% chance that it will stay the same, a 0.8% chance that it will decrease to B, and a 0.1% chance that it will default. If it defaults, the bond will be worth \$400. For each possible rating category there are three equally likely credit spreads. In basis points these are 80, 100, and 120 for Baa; 160, 200, and 240 for Ba; and 400, 450, and 500 for B.

The worst outcome is clearly a default. This has a probability of 0.1% and will lead to a loss of $907.03 - 400 = \$507.03$. The next worst outcome involves a transition to a rating of B when the credit spread for that rating category is 500 basis points. This has a probability of $0.8/3 = 0.267\%$. Because the bond will have about 1.917 years to maturity, its price will be $1{,}000/1.08^{1.917} = \$862.85$, for a loss of $907.03 - 862.85 = \$44.17$. A complete list of outcomes with their probabilities is shown in Table 21.5.

Table 21.5 Outcomes for Example 21.1

Rating	Spread (bp)	Probability	Bond Value (\$)	Loss (\$)
Default		0.100%	400.00	507.03
B	500	0.267%	862.85	44.17
B	450	0.267%	870.56	36.47
B	400	0.267%	878.38	28.65
Ba	240	32.833%	904.11	2.92
Ba	200	32.833%	910.72	−3.70
Ba	160	32.833%	917.41	−10.38
Baa	120	0.200%	924.17	−17.14
Baa	100	0.200%	927.58	−20.55
Baa	80	0.200%	931.01	−23.98

The table shows that, if the confidence level is higher than 99.9%, the VaR is \$507.03; if it is between 99.9% and 99.633%, the VaR is \$44.17; and so on. When the confidence level is 99%, the VaR is a modest \$2.92.

21.5.1 Constant Level of Risk Assumptions

When calculating VaR or ES, banks sometimes make a constant level of risk assumption. This means that they assume a portfolio is rebalanced periodically to bring the risk back to where it was initially.

Suppose, for example, that a company invests in Baa bonds and is considering two alternative trading strategies:

1. Hold bonds for one year and then sell (the "buy and hold" strategy).
2. Rebalance at the end of each month so that if bonds are no longer Baa they are sold and replaced with Baa bonds (the "constant level of risk" strategy).

The first strategy leads to outcomes given by the Baa row of Table 21.1, while, in the case of the second strategy, the outcomes given by the Baa row of Table 21.3 are repeated 12 times.

Consider first defaults. From Table 21.1, the probability of a default under the buy-and-hold strategy is 0.0018. From Table 21.3, the probability of a default each month is 0.0001215.[10] The probability of one default when the constant level of risk strategy is used is therefore

$$12 \times 0.0001215 \times 0.9998785^{11} = 0.001456$$

This is 19% less than that under the buy-and-hold strategy. The probability of two or more defaults is negligible.

Consider next the probability of a downgrade to C. Under the buy-and-hold strategy, the probability of this is 0.0017 from Table 21.1. When the constant level of risk strategy is used, the probability of one downgrade to C is

$$12 \times 0.0001277 \times 0.9998723^{11} = 0.001530$$

which is a 10% reduction over the corresponding number for the buy-and-hold strategy. The probability of more than one downgrade to C is negligible.

Finally, consider the probability of a downgrade to Ba. Under the buy-and-hold strategy, the probability of this is 0.0409 from Table 21.1. When the constant level of risk strategy is used, the probability of one downgrade to Ba is

$$12 \times 0.0038735 \times 0.9961265^{11} = 0.045505$$

which is a 10% increase over the buy-and-hold strategy. The probability of more than one downgrade in this case is about 0.001.

[10] Note that these results are based on numbers with more decimal places than those in Table 21.3. See results on the author's website.

The calculations illustrate that the buy-and-hold strategy will lead to greater losses from defaults and big downgrades than the constant level of risk strategy. But the probabilities of losses from a small rating change are less. VaR and ES are generally smaller when the constant level of risk strategy is used.

Summary

Credit VaR is defined similarly to the way VaR is defined for market risk. It is the credit loss that will not be exceeded over some time horizon with a specified confidence level. Basel regulations require credit VaR to be calculated with a one-year time horizon and a 99.9% confidence level. For economic capital purposes a different confidence level may be used.

We have discussed three methods for calculating credit VaR: Vasicek's model, Credit Risk Plus, and CreditMetrics. Vasicek's model is based on a one-factor Gaussian copula model of the time to default for the companies in a portfolio. It is used in the determination of regulatory capital. Credit Risk Plus uses procedures that are similar to those used in the insurance industry to calculate default losses from assumptions about the probabilities of default of individual companies. CreditMetrics is different from the other two methods in that it considers losses from both defaults and downgrades. It uses a Gaussian copula model in conjunction with a ratings transition matrix.

To determine VaR or ES for instruments dependent on credit spreads, one approach is to collect historical data on the credit spread changes of companies and use a historical simulation methodology similar to that for market risk. Another is to model the rating transitions of companies and movements in the average credit spreads associated with different rating categories.

Further Reading

Credit Suisse Financial Products. "Credit Risk Management Framework" (October 1997).

Finger, C. C. "Creditmetrics and Constant Level of Risk." Working Paper, RiskMetrics Group-MSCI, September 2010.

Gordy, M. "A Risk-Factor Model Foundation for Ratings-Based Capital Rules." *Journal of Financial Intermediation* 12, no. 3 (July 2003): 199–233.

JP Morgan. "CreditMetrics–Technical Document" (April 1997).

Vasicek, O. "Probability of Loss on a Loan Portfolio." Working Paper, KMV, 1987. (Published in *Risk* in December 2002 under the title "Loan Portfolio Value.")

Practice Questions and Problems (Answers at End of Book)

21.1 Explain the difference between Vasicek's model, the Credit Risk Plus model, and CreditMetrics as far as the following are concerned: (a) when a credit loss is recognized and (b) the way in which default correlation is modeled.

21.2 Explain what is meant by the constant level of risk assumption.

21.3 Use the transition matrix in Table 21.1 to calculate a transition matrix for two years. What is the probability of a company rated Aaa staying Aaa during the two years? What is the probability of it moving to Aa?

21.4 Use the transition matrix in Table 21.1 and software on the author's website to calculate a transition matrix for six months. What is the probability of a company rated Aaa staying Aaa during the six months? What is the probability of it moving to Aa?

21.5 How can historical simulation be used to calculate a one-day 99% VaR for the credit risk of bonds in the trading book? What are the disadvantages?

21.6 A bank has 100 one-year loans, each with a 1% probability of default. What is the probability of six or more defaults, assuming independence?

21.7 Repeat Problem 21.6 on the assumption that the probability of default applicable to all loans is uncertain. It is equally likely to be 0.5% or 1.5%.

21.8 What is the autocorrelation for the default rates in Table 11.6, which can also be found on a spreadsheet on the author's website? What are the implications of this for a Credit Risk Plus model?

Further Questions

21.9 Explain carefully the distinction between real-world and risk-neutral default probabilities. Which is higher? A bank enters into a credit derivative where it agrees to pay \$100 at the end of one year if a certain company's credit rating falls from A to Baa or lower during the year. The one-year risk-free rate is 5%. Using Table 21.1, estimate a value for the derivative. What assumptions are you making? Do they tend to overstate or understate the value of the derivative?

21.10 Suppose that a bank has a total of \$10 million of small exposures of a certain type. The one-year probability of default is 1% and the recovery rate averages 40%. Estimate the 99.5% one-year credit VaR using Vasicek's model if the copula correlation parameter is 0.2.

21.11 Use the transition matrix in Table 21.1 and software on the author's website to calculate the transition matrix over 1.25 years.

Part Five

OTHER TOPICS

Chapter 22

Scenario Analysis and Stress Testing

Stress testing involves evaluating the impact of extreme, but plausible, scenarios that are not considered by value at risk (VaR) or expected shortfall (ES) models. If there is one lesson to be learned from the market turmoil that started in the summer of 2007, it is that more emphasis should be placed on stress testing and less emphasis should be placed on the mechanistic application of VaR/ES models. VaR/ES models are useful, but they are inevitably backward looking. Risk management is concerned with what might happen in the future.

This chapter considers the different approaches that can be used to generate scenarios for stress testing and how the results should be used. It explains that the financial crisis of 2007–2008 has caused regulators to require banks to conduct more stress testing and that increasingly regulators are defining the stress tests themselves in an attempt to ensure that banks have sufficient capital to withstand adverse scenarios.

22.1 Generating the Scenarios

The most popular approach for calculating market risk VaR or ES is the historical simulation approach covered in Chapter 13. This approach assumes that data from the past are a good guide to what will happen over the next 1 to 10 days. But if an event has not occurred during the period covered by the data, it will not affect the results when the usual historical simulation method is used.

We have discussed a number of ways VaR/ES calculations can be modified so that they reflect more than the simple assumption that the future movements in market variables will be a random sample from the recent past. In particular:

1. Volatility scaling (see Section 13.3) can lead to more extreme outcomes being considered when the market is highly volatile.
2. Extreme value theory (see Section 13.5) provides a way of extending the tails of the loss distribution obtained from historical data.
3. Calculating stressed VaR or stressed ES (see Section 13.1) considers the impact of a particularly difficult period of 250 days that has occurred in the past.

But the nature of a VaR/ES calculation is that it is backward looking. Events that could happen, but are quite different from those that occurred during the period covered by the data, are not taken into account. Stress testing is an attempt to overcome this weakness of the VaR/ES measure.

Stress testing involves estimating how the portfolio of a financial institution would perform under scenarios involving extreme (but plausible) market moves. The exchange rate example in Table 10.1 illustrates the occurrence of extreme market moves. It shows that five-standard-deviation moves happen about every five years on average, whereas they would happen hardly ever (about once every 7,000 years) if the distributions were normal.

A key issue in stress testing is the way in which scenarios are chosen. We now consider alternative procedures.

22.1.1 Stressing Individual Variables

One approach is to use scenarios where there is a large move in one variable and other variables are unchanged. Examples of scenarios of this type that are sometimes considered are:

1. A 100-basis-point parallel shift (up or down) in a yield curve.
2. Increasing or decreasing all the implied volatilities used for an asset by 50% of current values.
3. Increasing or decreasing an equity index by 10%.
4. Increasing or decreasing the exchange rate for a major currency by 6%.
5. Increasing or decreasing the exchange rate for a minor currency by 20%.

The impact of small changes in a variable is measured by its delta, as explained in Chapter 8. The impact of larger changes can be measured by a combination of delta and gamma. Here we are considering changes that are so large that it is likely to be unreliable to estimate the change in the value of a portfolio using Greek letters.

22.1.2 Scenarios Involving Several Variables

Usually, when one market variable shows a big change, others do as well. This has led financial institutions to develop scenarios where several variables change at the same time. A common practice is to use extreme movements in market variables that have occurred in the past. For example, to test the impact of an extreme movement in U.S. equity prices, a company might set the percentage changes in all market variables equal to those on October 19, 1987 (when the S&P 500 moved by 22.3 standard deviations). If this is considered to be too extreme, the company might choose January 8, 1988 (when the S&P 500 moved by 6.8 standard deviations). Other dates when there were big movements in equity prices are September 11, 2001, when terrorists attacked the World Trade Center in New York, and September 15, 2008, when Lehman Brothers declared bankruptcy. To test the effect of extreme movements in UK interest rates, the company might set the percentage changes in all market variables equal to those on April 10, 1992 (when 10-year bond yields moved by 8.7 standard deviations).

Another approach is to magnify what has happened in the past to generate extreme scenarios. For example, we might take a period of time when there were moderately adverse market movements and create a scenario where all variables move by three or five times as much as they did then. The problem with this approach is that correlations increase in stressed market conditions, and increasing the movements in all market variables by a particular multiple does not increase correlation.

Some historical scenarios are one-day shocks to market variables. Others, particularly those involving credit and liquidity, involve shocks that take place over several days, several weeks, or even several months. It is important to include volatilities in the market variables that are considered. Typically, extreme movements in market variables such as interest rates and exchange rates are accompanied by large increases in the volatilities of these variables and large increases in the volatilities of a wide range of other variables. Some scenarios are likely to involve big movements in commodity prices, such as the plunge in oil prices in the second half of 2014. Others may involve a situation where there is a flight to quality combined with a shortage of liquidity and an increase in credit spreads. This was what happened in August 1998 when Russia defaulted on its debt and in July and August 2007 when investors lost confidence in the products created from the securitization of subprime mortgages (see Chapter 6).

22.1.3 Scenarios Generated by Management

History never repeats itself exactly. This may be partly because traders are aware of past financial crises and try to avoid making the same mistakes as their predecessors. The U.S. mortgage market led to the credit crisis that started in 2007. It is unlikely that future credit crises will be a result of mortgage-lending criteria being relaxed—but it is likely that there will be credit crises in the future.

In many ways, the scenarios that are most useful in stress testing are those generated by senior management or by the economics group within a financial institution. Senior management and the economics group are in a good position to use their understanding of markets, world politics, the economic environment, and current global uncertainties to develop plausible scenarios that would lead to large losses. Sometimes, the scenarios produced are based on things that have happened in the past, but are adjusted to include key features of the current financial and economic environment.

One way of developing the scenarios is for a committee of senior management to meet periodically and "brainstorm" answers to the simple question: "What can go wrong?" Clemens and Winkler (1999) have done studies to investigate the optimal composition of a committee of this type.[1] Their conclusions are that (a) the committee should have three to five members, (b) the backgrounds of the committee members should be different, and (c) there should be a healthy dialogue between members of the committee. It is important that members of the committee are able step back from their day-to-day responsibilities to consider the big picture.

It is not always the case that senior management's thinking has to be highly innovative in order for it to come up with relevant scenarios. In 2005 and 2006, many commentators realized that the U.S. housing market was experiencing a bubble and that sooner or later the bubble would burst. It is easy to be wise after the event, but one reasonable scenario for the stress-testing committee to propose during that period would have been a 20% or 30% decline in house prices in all parts of the country.

It is important that senior management and the board of directors understand and recognize the importance of stress testing. They have the responsibility for taking strategic decisions based on the stress-testing results. One advantage of involving senior management in the development of the scenarios to be used in stress testing is that it should naturally lead to a "buy in" by them to the idea that stress testing is important. The results generated from scenarios that are created by individuals who have middle management positions are unlikely to be taken as seriously.

22.1.4 Core vs. Peripheral Variables

When individual variables are stressed or scenarios are generated by management, the scenarios are likely to be incomplete in that movements of only a few (core) market variables are specified. One approach is to set changes in all other (peripheral) variables to zero, but this is likely to be unsatisfactory. Another approach is to regress the peripheral variables on the core variables that are being stressed to obtain forecasts for them conditional on the changes being made to the core variables. These forecasts (as point forecasts or probability distributions) can be incorporated into the stress test.

This is known as *conditional stress testing* and is discussed by Kupiec (1999).[2] Kim and Finger (2000) carry this idea further by using what they call a "broken arrow" stress test. In this, the correlation between the core variables and the peripheral variables

[1] See R. Clemens and R. Winkler, "Combining Probability Distributions from Experts in Risk Analysis," *Risk Analysis* 19, no. 2 (April 1999): 187–203.

[2] P. Kupiec, "Stress Testing in a Value at Risk Framework," *Journal of Derivatives* 6 (1999): 7–24.

is based on what happens in stressed market conditions rather than what happens on average.[3]

22.1.5 Making Scenarios Complete

Scenarios should be carefully examined in an attempt to make sure that all the adverse consequences are considered. The scenarios should include not only the immediate effect on the financial institution's portfolio of a shock to market variables, but also any "knock-on" effect resulting from many different financial institutions being affected by the shock in the same way and responding in the same way. Many people have said that they recognized the real estate bubble in the United States would burst in 2007, but did not realize how bad the consequences would be. They did not anticipate that many financial institutions would experience losses at the same time with the result that there would be a flight to quality with severe liquidity problems and a huge increase in credit spreads.

Another example of knock-on effects is provided by the failure of Long-Term Capital Management (LTCM) in 1998 (see Business Snapshot 22.1). LTCM tended to have long positions in illiquid securities and short positions in liquid securities. Its failure was caused by a flight to quality following Russia's default on its debt. Many investors were only interested in buying liquid securities. Spreads between liquid and illiquid securities increased. LTCM had done stress tests looking at the impact of flights to quality similar to those that had occurred pre-1998. What it did not allow for was the knock-on effect. Many hedge funds were following similar trading strategies to those of LTCM in 1998. When the flight to quality started, they were all forced to unwind their positions at the same time. Unwinding meant selling illiquid securities and buying liquid securities, reinforcing the flight to quality and making it more severe than previous flights to quality.

Scenarios should ideally be dynamic so that the response to the scenario of the financial institution doing the stress test, as well as the responses of other financial institutions, are considered. Consider, for example, the situation where a financial institution has sold call options dependent on an underlying asset and maintains delta neutrality. A shock where there is a large increase in the asset price will lead to an immediate loss on the option position. To maintain delta neutrality (see Section 8.1), large amounts of the asset will have to be bought at a relatively high price. The cost of subsequent delta hedging is liable to depend on the volatility of the asset price. Often the volatility will be high, with big increases and decreases in the asset price before it settles down. The subsequent delta hedging is then particularly expensive.

22.1.6 Reverse Stress Testing

Reverse stress testing involves the use of computational procedures to search for scenarios that lead to a failure of the financial institution. It has become an important tool in risk management, and many bank supervisory bodies now require banks to have reverse stress testing programs in place.

---

[3] See J. Kim and C. C. Finger, "A Stress Test to Incorporate Correlation Breakdown," *Journal of Risk* 2, no. 3 (Spring 2000): 5–19.

BUSINESS SNAPSHOT 22.1

Long-Term Capital Management's Big Loss

Long-Term Capital Management (LTCM), a hedge fund formed in the mid-1990s, always collateralized its transactions. The hedge fund's investment strategy was known as convergence arbitrage. A very simple example of what it might do is the following. It would find two bonds, X and Y, issued by the same company promising the same payoffs, with X being less liquid (i.e., less actively traded) than Y. The market always places a value on liquidity. As a result, the price of X would be less than the price of Y. LTCM would buy X, short Y, and wait, expecting the prices of the two bonds to converge at some future time.

When interest rates increased, the company expected both bonds to move down in price by about the same amount so that the collateral it paid on bond X would be about the same as the collateral it received on bond Y. Similarly, when interest rates decreased, LTCM expected both bonds to move up in price by about the same amount so that the collateral it received on bond X would be about the same as the collateral it paid on bond Y. It therefore expected that there would be no significant outflow of funds as a result of its collateralization agreements.

In August 1998, Russia defaulted on its debt and this led to what is termed a "flight to quality" in capital markets. One result was that investors valued liquid instruments more highly than usual and the spreads between the prices of the liquid and illiquid instruments in LTCM's portfolio increased dramatically. The prices of the bonds LTCM had bought went down and the prices of those it had shorted increased. It was required to post collateral on both. The company was highly leveraged and found it difficult to make the payments required under the collateralization agreements. The result was that positions had to be closed out and there was a total loss of about $4 billion. If the company had been less highly leveraged it would probably have been able to survive the flight to quality and could have waited for the prices of the liquid and illiquid bonds to become closer.

Example 22.1

As a simple example of reverse stress testing, suppose a financial institution has positions in four European call options on an asset. The asset price is $50, the risk-free rate is 3%, the volatility is 20%, and there is no income on the asset. The positions, strike prices, and lives of the options are as indicated in the table below. The current value of the position (in millions of dollars) is −25.90. A numerical procedure can be used to search for one-day changes in the asset price and the volatility that will lead to the greatest losses. Some bounds should be put on the changes that are considered. We assume that the asset price will not decrease below \$40 or increase above \$60. It is assumed that the volatility will not fall below 10% or rise above 30%.

Position (000s)	**Strike Price**	**Life (years)**	**Position Value (\$ millions)**
+250	50	1.0	1,176.67
−125	60	1.5	−293.56
−75	40	0.8	−843.72
−50	55	0.5	−65.30
Total			−25.90

Using the formulas for valuing European options (see Appendix E) in conjunction with Excel's Solver, the worst loss is found to be when the volatility decreases to 10% and the asset price falls to \$45.99. The loss is \$289.38 million. Reverse stress testing therefore shows that the financial institution is most exposed to a reduction of about 8% in the asset price combined with a sharp decline in volatility.

This might seem to be an unreasonable scenario. It is unlikely that volatility would go down sharply when the asset price declines by 8%. Solver could be run again with the lower bound to volatility being 20% instead of 10%. This gives a worst-case loss occurring when the volatility stays at 20% and the asset price falls to \$42.86. The loss is then \$87.19 million.

Searching over all the market variables to which a financial institution is exposed in the way indicated in Example 22.1 is in practice usually not computationally feasible. One approach is to identify 5 to 10 key market variables and assume that changes in other variables are dependent on changes in these variables.

Another way of simplifying the search process is to impose some structure on the problem. A principal components analysis (see Section 9.7) can be carried out on the changes in market variables (ideally using data from stressed market conditions) and then a search can be conducted to determine the changes in the principal components that generate large losses. This reduces the dimension of the space over which the search is conducted and should lead to fewer implausible scenarios.

An alternative approach is for the risk management group to impose a structure on scenarios. For example, a number of adverse scenarios that have occurred in the past can be specified. An analyst can then search to find what multiplier must be applied to the changes observed in all market variables under the scenarios in order for a particular loss level to be reached. For example, it might be concluded that a financial institution could survive a repeat of market movements during the credit crisis of 2008, but that if movements were 50% greater there would be serious problems.

Reverse stress testing can be used as a tool to facilitate brainstorming by the stress-testing committee. Prior to a meeting of the stress-testing committee, analysts can use reverse stress testing to come up with several scenarios that would be disastrous to the financial institution. These scenarios, along with others they generate themselves, are then considered by the stress-testing committee. They use their judgment to eliminate some of the analysts' scenarios as implausible and modify others so that they become plausible and are retained for serious evaluation.

22.2 Regulation

The Basel Committee requires market risk calculations that are based on a bank's internal VaR models to be accompanied by "rigorous and comprehensive" stress testing. Similarly, banks using the IRB approach in Basel II (advanced or foundation) to determine credit risk capital are required to conduct stress tests to determine the robustness of their assumptions.

In May 2009, the Basel Committee issued the final version of its recommendations on stress-testing practices and how stress testing should be supervised by regulators.[4] The recommendations emphasize the importance of stress testing in determining how much capital is necessary to absorb losses should large shocks occur. They make the point that stress testing is particularly important after long periods of benign conditions because such conditions tend to lead to complacency.

The recommendations highlight the importance of top management and board involvement in stress testing. In particular, top management and board members should be involved in setting stress-testing objectives, defining scenarios, discussing the results of stress tests, assessing potential actions, and decision making. It makes the point that the banks that fared well in the financial crisis that started in mid-2007 were the ones where senior management as a whole took an active interest in the development and operation of stress testing, with the results of stress testing serving as an input into strategic decision making. Stress testing should be conducted across all areas of the bank. It should not be the case that each area conducts its own stress test.

The Basel recommendations make the point that many of the scenarios chosen pre-2007 were based on historical data and much less severe than what actually happened. Specific recommendations for banks are:

1. Stress testing should form an integral part of the overall governance and risk management culture of the bank. Stress testing should be actionable, with the results from stress-testing analyses impacting decision making at the appropriate management level, including strategic business decisions of the board and senior management. Board and senior management involvement in the stress-testing program is essential for its effective operation.
2. A bank should operate a stress-testing program that promotes risk identification and control, provides a complementary risk perspective to other risk management tools, improves capital and liquidity management, and enhances internal and external communication.

[4] See "Principles for Sound Stress-Testing Practices and Supervision," Basel Committee on Banking Supervision, May 2009.

3. Stress-testing programs should take account of views from across the organization and should cover a range of perspectives and techniques.
4. A bank should have written policies and procedures governing the stress-testing program. The operation of the program should be appropriately documented.
5. A bank should have a suitably robust infrastructure in place, which is sufficiently flexible to accommodate different and possibly changing stress tests at an appropriate level of granularity.
6. A bank should regularly maintain and update its stress-testing framework. The effectiveness of the stress-testing program, as well as the robustness of major individual components, should be assessed regularly and independently.
7. Stress tests should cover a range of risks and business areas, including at the firmwide level. A bank should be able to integrate effectively across the range of its stress-testing activities to deliver a complete picture of firmwide risk.
8. Stress-testing programs should cover a range of scenarios, including forward-looking scenarios, and aim to take into account system-wide interactions and feedback effects.
9. Stress tests should feature a range of severities, including events capable of generating most damage whether through size of loss or through loss of reputation. A stress-testing program should also determine what scenarios could challenge the viability of the bank (reverse stress tests) and thereby uncover hidden risks and interactions among risks.
10. As part of an overall stress-testing program, a bank should aim to take account of simultaneous pressures in funding and asset markets, and the impact of a reduction in market liquidity on exposure valuation.
11. The effectiveness of risk mitigation techniques should be systematically challenged.
12. The stress-testing program should explicitly cover complex and bespoke products such as securitized exposures. Stress tests for securitized assets should consider the underlying assets, their exposure to systematic market factors, relevant contractual arrangements and embedded triggers, and the impact of leverage, particularly as it relates to the subordination level in the issue structure.
13. The stress-testing program should cover pipeline and warehousing risks.[5] A bank should include such exposures in its stress tests regardless of their probability of being securitized.
14. A bank should enhance its stress-testing methodologies to capture the effect of reputational risk. The bank should integrate risks arising from off-balance-sheet vehicles and other related entities in its stress-testing program.
15. A bank should enhance its stress-testing approaches for highly leveraged counterparties in considering its vulnerability to specific asset categories or market movements and in assessing potential wrong-way risk related to risk-mitigating techniques.

[5] "Pipeline and warehousing" risks refer to risks associated with assets that are awaiting securitization, but might not be securitized if market conditions change. These risks led to losses during the onset of the crisis.

The recommendations for bank supervisors are:

16. Supervisors should make regular and comprehensive assessments of a bank's stress-testing programs.

17. Supervisors should require management to take corrective action if material deficiencies in the stress-testing program are identified or if the results of stress tests are not adequately taken into consideration in the decision-making process.

18. Supervisors should assess and, if necessary, challenge the scope and severity of firm-wide scenarios. Supervisors may ask banks to perform sensitivity analysis with respect to specific portfolios or parameters, use specific scenarios, or evaluate scenarios under which their viability is threatened (reverse stress-testing scenarios).

19. Under the Pillar 2 (supervisory review process) of the Basel II framework, supervisors should examine a bank's stress-testing results as part of a supervisory review of both the bank's internal capital assessment and its liquidity risk management. In particular, supervisors should consider the results of forward-looking stress testing for assessing the adequacy of capital and liquidity.

20. Supervisors should consider implementing stress-test exercises based on common scenarios.

21. Supervisors should engage in a constructive dialogue with other public authorities and the industry to identify systemic vulnerabilities. Supervisors should also ensure that they have the capacity and the skills to assess banks' stress-testing programs.

22.2.1 Scenarios Chosen by Regulators

Bank regulators require banks to consider extreme scenarios and then make sure they have enough capital for those scenarios. There is an obvious problem here. Banks want to keep their regulatory capital as low as possible. They therefore have no incentive to consider extreme scenarios that would lead to a bank supervisor telling them that their capital requirements need to be increased. There is therefore a natural tendency for the scenarios they consider to be "watered down" and fairly benign.

One approach to overcoming this problem is for regulators themselves to provide the scenarios (see Recommendations 18 and 20). Regulators in many jurisdictions (including the United Kingdom, the European Union, and the United States) now regularly carry out their own stress tests to determine whether the banks they are supervising have enough capital to survive extreme adverse events. If banks fail the stress test, they are required to raise more capital and may be subject to dividend restrictions.

In the United States, the Federal Reserve has since 2009 carried out an annual stress test for all banks with consolidated assets of more than $50 billion (i.e., G-SIBs and D-SIBs). Since 2011 this has been called the Comprehensive Capital Analysis and Review (CCAR). Banks are required to submit a capital plan (including planned dividends). The scenarios created by the Federal Reserve are recession scenarios similar to those in 1973–1975, 1981–1982, and 2007–2009. The scenarios include projections for about 25 variables, including macroeconomic variables such as gross domestic product (GDP) growth, unemployment rate, stock market indices, and house price indices.

Banks in the United States with consolidated assets over $10 billion are subject to the Dodd–Frank Act Stress Test (DFAST). The scenarios considered in DFAST are similar to those in CCAR. However, banks do not have to submit a capital plan, as the management of capital is assumed to be based on a standard set of assumptions. Both CCAR and DFAST involve a huge amount of work for banks with thousands of pages of justification for the models used being produced.

By choosing scenarios themselves, regulators are able to focus the attention of banks on recession scenarios that are of concern to regulators. If regulators see many banks taking positions with similar risks, they can require all banks to consider a particular set of scenarios that give rise to adverse results for the positions. The downside of regulators generating scenarios themselves is that part of the reason for the increased focus by supervisors on stress testing is that they want to encourage financial institutions to spend more time generating and worrying about potential adverse scenarios. If supervisors do the work in generating the scenarios, this may not happen and banks will focus only on the scenarios generated by regulators. In an ideal regulatory environment, both management-generated scenarios and supervisor-generated scenarios are evaluated.

There is a danger that financial institutions will find actions they can take that allow them to pass the stress tests without making them any safer. An extreme example of this is in Business Snapshot 22.2. When regulators defined key scenarios for life insurance companies and pension funds in Denmark, some of those companies responded by hedging against the particular scenarios used by regulators—and only against those scenarios.[6] This is not what regulators intended. Each scenario used in stress testing should be viewed as representative of a range of things that might happen. Financial institutions should ensure that their capital will be in good shape not just for the specified scenarios, but also for other similar or related scenarios. Regulators are now well aware of this type of problem and have procedures aimed at ensuring that banks do not game the system in this way.

22.3 What to Do with the Results

The biggest problem in stress testing is using the results effectively. All too often, the results of stress testing are ignored by senior management. A typical response is, "Yes, there are always one or two scenarios that will sink us. We cannot protect ourselves against everything that might happen." One way of trying to avoid this sort of response is to involve senior management in the development of scenarios, as outlined earlier. A better response on the part of senior management would be, "Are the risks associated with these scenarios acceptable? If not, let's investigate what trades we can put on or other actions we can take to make these types of risks more acceptable."

[6] The information in Business Snapshot 22.2 is from P. L. Jorgensen, "Traffic Light Options," *Journal of Banking and Finance* 31, no. 12 (December 2007): 3698–3719.

BUSINESS SNAPSHOT 22.2

Traffic Light Options

In June 2001, the Danish Financial Supervisory Authority (DFSA) introduced a "traffic light" solvency stress-testing system. This requires life insurance companies and pension funds to submit semiannual reports indicating the impact on them of certain predefined shocks. The "red light scenario" involves a 70-basis-point decrease in interest rates, a 12% decline in stock prices, and an 8% decline in real estate prices. If capital falls below a specified critical level in this scenario, the company is categorized with "red light status" and is subject to more frequent monitoring with monthly reports being required. The "yellow light scenario" involves a 100-basis-point decrease in interest rates, a 30% decline in stock prices, and a 12% decline in real estate prices. If capital falls below the critical level in this scenario, the company is categorized with "yellow light status" and has to submit quarterly reports. When the company's capital stays above the critical levels for the red and yellow light scenarios, the company has a "green light status" and is subject to normal semiannual reporting.

Derivatives dealers have developed products for helping life insurance and pension funds keep a green light status. These are known as traffic light options. They pay off in the traffic light scenarios so as to give a boost to the financial institution's performance when these scenarios are considered. Traffic light options pay off only when the movements close to those specified by regulators in the red light or yellow light scenario occur. The options are therefore not too expensive.

The problem for both senior management and the risk management group is that they have two separate reports on their desks concerning what could go wrong. One report comes from VaR or ES models, the other from internal stress testing. Which one should they base their decision making on?

22.3.1 Integrating Stress Testing and VaR Calculations

Berkowitz (2000) suggests that stress testing will be taken more seriously if its results are integrated into the calculation of VaR.[7] This can be done by assigning a probability to each stress scenarios that is considered. Suppose that a financial institution has considered n_s stress scenarios and the total probability assigned to the stress scenarios is p. Assume

[7] See J. Berkowitz, "A Coherent Framework for Stress Testing," *Journal of Risk* 2, no. 2 (Winter 1999/2000): 5–15.

further that there are n_v VaR scenarios generated using historical simulation in the usual way. An analyst can assume that there are a total of $n_s + n_v$ scenarios. The n_s stress scenarios have probability p and the n_v historical scenarios have probability $1 - p$.

Unfortunately human beings are not good at estimating a subjective probability for the occurrence of a rare event. To make the task feasible for the stress-testing committee, one approach is to ask the stress-testing committee to allocate each scenario to categories with preassigned probabilities. The categories might be:

1. Probability = 0.05%. Extremely unlikely. One chance in 2,000.
2. Probability = 0.2%. Very unlikely, but the scenario should be given about the same weight as the 500 scenarios used in the historical simulation analysis.
3. Probability = 0.5%. Unlikely, but the scenario should be given more weight than the 500 scenarios used in the historical simulation analysis.

Example 22.2

Suppose that, in the example in Section 13.1, five stress scenarios are considered. They lead to losses ($000s) of 235, 300, 450, 750, and 850. The probabilities assigned to the scenarios are 0.5%, 0.2%, 0.2%, 0.05%, and 0.05%, respectively. The total probability of the stress scenarios is, therefore, 1%. This means that the probability assigned to the scenarios generated by historical simulation is 99%. Assuming that equal weighting is used, each historical simulation scenario is assigned a probability of $0.99/500 = 0.00198$. Table 13.4 is therefore replaced by Table 22.1. The probabilities assigned to scenarios are

Table 22.1 Losses Ranked from Highest to Lowest

Scenario	Loss ($000s)	Probability	Cumulative Probability
s5	850.000	0.00050	0.00050
s4	750.000	0.00050	0.00100
v494	477.841	0.00198	0.00298
s3	450.000	0.00200	0.00498
v339	345.435	0.00198	0.00696
s2	300.000	0.00200	0.00896
v349	282.204	0.00198	0.01094
v329	277.041	0.00198	0.01292
v487	253.385	0.00198	0.01490
s1	235.000	0.00500	0.01990
v227	217.974	0.00198	0.02188
v131	205.256	0.00198	0.02386
v238	201.389	0.00198	0.02584
…	…	…	…
…	…	…	…
…	…	…	…

For Example 22.2, s1, s2, … are the stress scenarios; v1, v2, … are the VaR historical simulation scenarios.

accumulated from the worst scenario to the best.[8] The VaR level when the confidence level is 99% is the first loss for which the cumulative probability is greater than 0.01. In our example this is $282,204.

Rebonato (2010) suggests a more elaborate approach to assessing probabilities of scenarios involving a careful application of a well-known result in statistics, Bayes' theorem, and what are known as Bayesian networks.[9] The probability of a scenario consisting of two events is equal to the probability of the first event happening times the probability of the second event happening conditional on the first event having happened. Similarly the probability of a scenario consisting of three events is the probability of the first event happening times the probability of the second event happening conditional that the first event has happened times the probability of the third event happening conditional that the first two events have happened. Rebonato's approach provides a way of evaluating the conditional probabilities.

22.3.2 Subjective vs. Objective Probabilities

There are two types of probability estimates: objective and subjective. An *objective probability* for an event is a probability calculated by observing the frequency with which the event happens in repeated trials. As an idealized example of an objective probability, consider an urn containing red balls and black balls in an unknown proportion. We want to know the probability of a ball drawn at random from the urn being red. We could draw a ball at random, observe its color, replace it in the urn, draw another ball at random, observe its color, replace it in the urn, and so on, until 100 balls have been drawn. If 30 of the balls that have been drawn are red and 70 are black, our estimate for the probability of drawing a red ball is 0.3. Unfortunately, most objective probabilities calculated in real life are usually less reliable than the probability in this example, because the probability of the event happening does not remain constant for the observations that are available and the observations may not be independent.

A *subjective probability* is a probability derived from an individual's personal judgment about the chance of a particular event occurring. The probability is not based on historical data. It is a degree of belief. Different people are liable to have different subjective probabilities for the same event.

The probabilities in historical simulation are objective whereas the probabilities assigned to the scenarios in stress testing are subjective. Many analysts are uncomfortable with subjective probabilities because they are not based on data. Also, it is unfortunately the case that political considerations may play a part in a financial institution's decision

[8] This is the same procedure that we used when weights were assigned to historical simulation scenarios. (See Table 13.5 in Section 13.3.)

[9] See Riccardo Rebonato, *Coherent Stress Testing: A Bayesian Approach to Financial Stress* (Chichester, UK: John Wiley & Sons, 2010).

to focus on historical data. If you use historical data and things go wrong, you can blame the data. If you use subjective judgments that have been provided by a group of people, those people are liable to be blamed.

However, if it is based only on objective probabilities, risk management is inevitably backward looking and fails to capitalize on the judgment and expertise of senior managers. It is the responsibility of those managers to steer the financial institution so that catastrophic risks are avoided.

Summary

Stress testing is an important part of the risk management process. It leads to a financial institution considering the impact of extreme scenarios that are ignored by a traditional VaR or ES analysis, but that do happen from time to time. Once plausible scenarios have been evaluated, the financial institution can take steps to lessen the impact of the particularly bad ones. One advantage of a comprehensive stress-testing program is that a financial institution obtains a better understanding of the nature of the risks in its portfolio.

Scenarios can be generated in a number of different ways. One approach is to consider extreme movements in just one market variable while keeping others fixed. Another is to use the movements in all market variables that occurred during periods in the past when the market experienced extreme shocks. The best approach is to ask a committee of senior management and economists to use their judgment and experience to generate the plausible extreme scenarios. Sometimes financial institutions carry out reverse stress testing where algorithms are used to search for scenarios that would lead to large losses. Scenarios should be as complete as possible and include the impact of knock-on effects as well as the initial shock to market variables. The market turmoil starting in summer 2007 shows that, in some cases, the knock-on effect can be significant and include a flight to quality, an increase in credit spreads, and a shortage of liquidity.

Regulators require financial institutions to keep sufficient capital for stress scenarios. Regulators in the United States, European Union, and United Kingdom regularly develop stress scenarios to be evaluated by all the large financial institutions they are responsible for. This may lead to some of the financial institutions having to raise more capital and has the potential to uncover systemic risk problems.

If subjective probabilities are assigned to the extreme scenarios that are considered, stress testing can be integrated with a VaR analysis. This is an interesting idea, but was not one of the approaches outlined in the Basel Committee consultative document published in May 2009.

Further Reading

Alexander, C., and E. A. Sheedy. "Developing a Stress-Testing Framework Based on Market Risk Models." *Journal of Banking and Finance* 32, no. 10 (2008): 2220–2236.

Aragonés, J. R., C. Blanco, and K. Dowd. "Incorporating Stress Tests into Market Risk Modeling." *Derivatives Quarterly* 7 (Spring 2001): 44–49.

Aragonés, J. R., C. Blanco, and K. Dowd. "Stress Tests, Market Risk Measures, and Extremes: Bringing Stress Tests to the Forefront of Market Risk Management." In *Stress Testing for Financial Institutions: Applications, Regulations, and Techniques*, edited by D. Rösch and H. Scheule. London: Risk Books, 2008.

Basel Committee on Banking Supervision. "Principles for Sound Stress-Testing Practices and Supervision," May 2009.

Berkowitz, J. "A Coherent Framework for Stress Testing." *Journal of Risk* 2, no. 2 (Winter 1999/2000): 5–15.

Bogle, J. C. "Black Monday and Black Swans." *Financial Analysts Journal* 64, no. 2 (March/April 2008): 30–40.

Clemens, R., and R. Winkler. "Combining Probability Distributions from Experts in Risk Analysis." *Risk Analysis* 19, no. 2 (April 1999): 187–203.

Duffie, D. "Systemic Risk Exposures: A 10-by-10-by-10 Approach." Working Paper, Stanford University, 2011.

Glasserman, P., C. Kang, and W. Kang. "Stress Scenario Selection by Empirical Likelihood." *Quantitative Finance* 15 (2015): 25–41.

Hassani, B. *Scenario Analysis in Risk Management*. Cham, Switzerland: Springer, 2016.

Hua, P., and P. Wilmott. "Crash Courses." *Risk* 10, no. 6 (June 1997): 64–67.

Kim, J., and C. C. Finger. "A Stress Test to Incorporate Correlation Breakdown." *Journal of Risk* 2, no. 3 (Spring 2000): 5–19.

Kupiec, P. "Stress Testing in a Value at Risk Framework." *Journal of Derivatives* 6 (1999): 7–24.

Rebonato, R. *Coherent Stress Testing: A Bayesian Approach to the Analysis of Financial Stress*. Chichester, UK: John Wiley & Sons, 2010.

Taleb, N. N. *The Black Swan: The Impact of the Highly Improbable*. New York: Random House, 2007.

Practice Questions and Problems (Answers at End of Book)

22.1 Explain three different ways that scenarios can be generated for stress testing.

22.2 What is reverse stress testing? How is it used?

22.3 Why might the regulatory environment lead to a financial institution underestimating the severity of the scenarios it considers?

22.4 What are traffic light options? What are their drawbacks?

22.5 Why is it important for senior management to be involved in stress testing?

22.6 What are the advantages and disadvantages of bank regulators choosing some of the scenarios that are considered for stress testing?

22.7 Explain the difference between subjective and objective probabilities.

22.8 Suppose that, in the example in Section 13.1, seven stress scenarios are considered. They lead to losses ($000s) of 240, 280, 340, 500, 700, 850, and 1,050. The subjective probabilities assigned to the scenarios are 0.5%, 0.5%, 0.2%, 0.2%, 0.05%, 0.05%, and 0.05%, respectively. What is the new one-day 99% VaR that would be calculated using the procedure discussed in Section 22.3?

22.9 Suppose that the positions in the four options in Example 22.1 are changed to 200, −70, −120, and −60. Use Solver to calculate the worst-case scenario for a daily change. Consider asset prices between \$40 and \$60 and volatilities between 10% and 30%.

Further Questions

22.10 What difference does it make to the worst-case scenario in Example 22.1 if (a) the options are American rather than European and (b) the options are barrier options that are knocked out if the asset price reaches \$65? Use Solver to search over asset prices between \$40 and \$60 and volatilities between 18% and 30%.

22.11 What difference does it make to the VaR calculated in Example 22.2 if the exponentially weighted moving average model is used to assign weights to scenarios as described in Section 13.3?

Chapter 23

Operational Risk

In 1999, the Basel Committee announced plans to assign capital for operational risk in the new Basel II regulations. This met with some opposition from banks. The chairman and CEO of one major international bank described it as "the dopiest thing I have ever seen." However, bank supervisors did not back down, because they knew that many of the big losses of banks over the previous 10 years had been operational risk losses—not credit risk or market risk losses.

Operational risk has also led to some big losses since the 1999 announcement. A big rogue trader loss at Société Générale in 2008 is discussed in Business Snapshot 5.5, and there was a similar loss at UBS in 2011. The London Whale loss at JPMorgan Chase in 2012 is discussed in Business Snapshot 25.1. In 2014, it was announced that the French bank BNP Paribas would pay $9 billion (roughly one year's profit) to the U.S. government for violating economic sanctions by moving dollar-denominated transactions through the American banking system on behalf of Sudanese, Iranian, and Cuban parties. The bank was also banned from conducting certain U.S. transactions for a year.

An increasingly important type of operational risk for banks is cyber risk. Banks have sophisticated systems in place to protect themselves from cyber attacks, but the attacks themselves are becoming more sophisticated. Also, banks are making increasing use of computer systems and the Internet, providing more avenues for cyber fraud. Customers and employees have to be continually educated so that the bank's data remain secure. Cyber attacks on banks are attractive to criminals because, to quote the bank robber Willie Sutton, "that's where the money is." They are also attractive to terrorists because of their potential to damage a nation's economic security and way of life.

Some regulators now regard operational risk as the most important risk facing banks. To quote Thomas J. Curry, head of the Office of the Comptroller of the Currency (OCC) in the United States, in 2012: "Given the complexity of today's banking markets and the sophistication of the technology that underpins it, it is no surprise that the OCC deems operational risk to be high and increasing. Indeed, it is currently at the top of the list of safety and soundness issues for the institutions we supervise." He goes on to argue that operational risk is more important than credit risk.[1] Most banks have always had some framework in place for managing operational risk. However, the prospect of new capital requirements led them to greatly increase the resources they devote to measuring and monitoring operational risk.

It is much more difficult to quantify operational risk than credit or market risk. Operational risk is also more difficult to manage. Financial institutions make a conscious decision to take a certain amount of credit and market risk, and there are many traded instruments that can be used to reduce these risks. Operational risk, by contrast, is a necessary part of doing business. An important part of operational risk management is identifying the types of risk that are being taken and should be insured against. There is always a danger that a huge loss will be incurred from taking an operational risk that ex ante was not even recognized as a risk.

It might be thought that a loss such as that at Société Générale (SocGen, see Business Snapshot 5.5) was a result of market risk because it was movements in market variables that led to it. However, it should be classified as an operational risk loss because it involved fraud. (Jérôme Kerviel created fictitious trades to hide the big bets he was making.) Suppose there was no fraud. If it was part of the bank's policy to let traders take huge risks, then the loss would be classified as a market risk loss. But, if this was not part of the bank's policy and there was a breakdown in its controls, it would be classified as an operational risk loss. The SocGen example illustrates that operational risk losses are often contingent on market movements. If the market had moved in Kerviel's favor, there would have been no loss. The fraud and the breakdown in SocGen's control systems might then never have come to light.

There are some parallels between the operational risk losses of banks and the losses of insurance companies. Insurance companies face a small probability of a large loss from a hurricane, earthquake, or other natural disaster. Similarly, banks face a small probability of a large operational risk loss. But there is one important difference. When insurance companies lose a large amount of money because of a natural disaster, all companies in the industry tend to be affected and often premiums rise the next year to cover losses. Operational risk losses tend to affect only one bank. Because it operates in a competitive

[1] Speech to the Exchequer Club, May 16, 2012.

environment, the bank does not have the luxury of increasing prices for the services it offers during the following year.

23.1 Defining Operational Risk

There are many different ways in which operational risk can be defined. It is tempting to consider operational risk as a residual risk and define it as any risk faced by a financial institution that is not market risk or credit risk. To produce an estimate of operational risk, we could then look at the financial institution's financial statements and remove from the income statement (a) the impact of credit losses and (b) the profits or losses from market risk exposure. The variation in the resulting income would then be attributed to operational risk.

Most people agree that this definition of operational risk is too broad. It includes the risks associated with entering new markets, developing new products, economic factors, and so on. Another possible definition is that operational risk, as its name implies, is the risk arising from operations. This includes the risk of mistakes in processing transactions, making payments, and so on. This definition of risk is too narrow. It does not include major risks such as the risk of a rogue trader such as Jérôme Kerviel.

We can distinguish between internal risks and external risks. Internal risks are those over which the company has control. The company chooses whom it employs, what computer systems it develops, what controls are in place, and so on. Some people define operational risks as all internal risks. Operational risk then includes more than just the risk arising from operations. It includes risks arising from inadequate controls such as the rogue trader risk and the risks of other sorts of employee fraud.

Bank regulators favor including more than just internal risks in their definition of operational risk. They include the impact of external events such as natural disasters (for example, a fire or an earthquake that affects the bank's operations), political and regulatory risk (for example, being prevented from operating in a foreign country by that country's government), security breaches, and so on. All of this is reflected in the following definition of operational risk produced by the Basel Committee on Banking Supervision in 2001:

> *The risk of loss resulting from inadequate or failed internal processes, people, and systems or from external events.*

Note that this definition includes legal risk but does not include reputation risk and the risk resulting from strategic decisions.

Operational risks result in increases in the bank's costs or decreases in its revenue. Some operational risks interact with credit and market risk. For example, when mistakes are made in a loan's documentation, it is usually the case that losses result if and only if the counterparty defaults. When a trader exceeds limits and misreports positions, losses result if and only if the market moves against the trader.

23.2 Categorization of Operational Risks

The Basel Committee on Banking Supervision has identified seven categories of operational risk.[2] These are:

1. *Internal fraud:* Acts of a type intended to defraud, misappropriate property, or circumvent regulations, the law, or company policy (excluding those concerned with diversity or discrimination) involving at least one internal party. Examples include intentional misreporting of positions, employee theft, insider trading on an employee's own account, and rogue trading.
2. *External fraud:* Acts by a third party of a type intended to defraud, misappropriate property, or circumvent the law. Examples include robbery, forgery, check kiting, and damage from computer hacking.
3. *Employment practices and workplace safety:* Acts inconsistent with employment, health, or safety laws or agreements, or that result in payment of personal injury claims or claims relating to diversity or discrimination issues. Examples include worker compensation claims, violation of employee health and safety rules, organized labor activities, discrimination claims, and general liability (for example, a customer slipping and falling at a branch office).
4. *Clients, products, and business practices:* Unintentional or negligent failure to meet a professional obligation to clients and the use of inappropriate products or business practices. Examples are fiduciary breaches, misuse of confidential customer information, improper trading activities on the bank's account, money laundering, and the sale of unauthorized products.
5. *Damage to physical assets:* Loss or damage to physical assets from natural disasters or other events. Examples include terrorism, vandalism, earthquakes, fires, and floods.
6. *Business disruption and system failures:* Disruption of business or system failures. Examples include hardware and software failures, telecommunication problems, and utility outages.
7. *Execution, delivery, and process management:* Failed transaction processing or process management, and disputes with trade counterparties and vendors. Examples include data entry errors, collateral management failures, incomplete legal documentation, unapproved access given to clients' accounts, non-client counterparty misperformance, and vendor disputes.

Eight different business lines within a bank can be distinguished:

1. Corporate finance
2. Trading and sales
3. Retail banking
4. Commercial banking

[2] See Basel Committee on Banking Supervision, "Sound Practices for the Management and Supervision of Operational Risk," Bank for International Settlements, February 2003.

5. Payment and settlement
6. Agency services
7. Asset management
8. Retail brokerage

Combining the seven categories of risk with the eight business lines gives a total of $7 \times 8 = 56$ potential sources of operational risk for a bank.

23.3 Regulatory Capital Under Basel II

We now discuss how regulatory capital is determined under Basel II. As we will see in Section 23.4, the Basel Committee has indicated its intention to totally change its procedures for determining operational risk regulatory capital. In particular, it stated that the Advanced Measurement Approach (AMA), which is a complex procedure for determining the one-year 99% VaR for operational risk losses, is to be abandoned. However, many banks have devoted huge resources to implementing AMA, and it is unlikely to disappear altogether. Indeed, regulators in some countries have indicated that their banks are welcome to continue using AMA. We will therefore review AMA in this section.

Banks have three alternatives for determining operational risk regulatory capital under Basel II. The simplest approach is the *basic indicator approach*. Under this approach, operational risk capital is set equal to 15% of annual gross income over the previous three years. Gross income is defined as net interest income plus noninterest income.[3] A slightly more complicated approach is the *standardized approach*. In this, a bank's activities are divided into the eight business lines mentioned earlier: corporate finance, trading and sales, retail banking, commercial banking, payment and settlement, agency services, asset management, and retail brokerage. The average gross income over the last three years for each business line is multiplied by a "beta factor" for that business line and the result summed to determine the total capital. The beta factors are shown in Table 23.1. The third alternative is the *advanced measurement approach* (AMA). In this, the operational risk

Table 23.1 Beta Factors in the Standardized Approach

Business Line	**Beta Factor**
Corporate finance	18%
Trading and sales	18%
Retail banking	12%
Commercial banking	15%
Payment and settlement	18%
Agency services	15%
Asset management	12%
Retail brokerage	12%

[3] Net interest income is the excess of income earned on loans over interest paid on deposits and other instruments that are used to fund the loans (see Section 2.2).

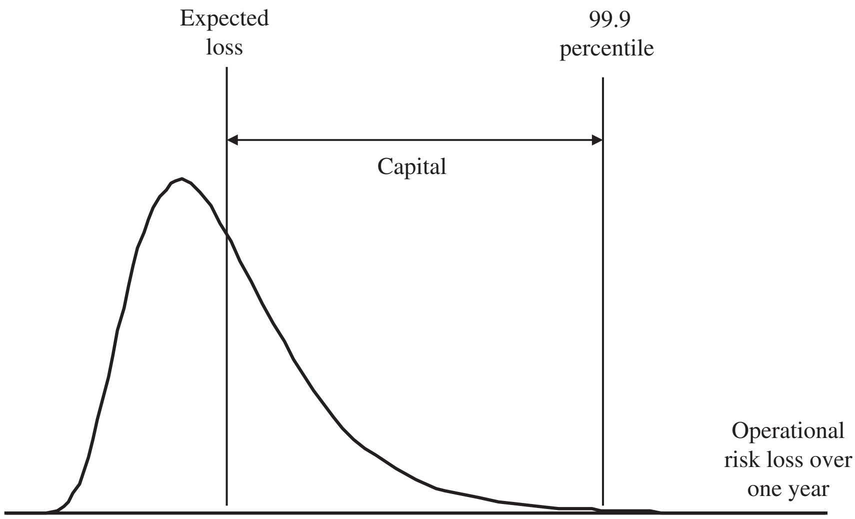


Figure 23.1 Calculation of Capital for Operational Risk

regulatory capital requirement is calculated by the bank internally using qualitative and quantitative criteria. Similarly to credit capital, it is based on a one-year 99.9% VaR.

The objective of banks using the AMA approach for operational risk is analogous to their objectives when they attempt to quantify credit risk. They would like to produce a probability distribution of losses such as that shown in Figure 23.1. Assuming that they can convince regulators that their expected operational risk cost is incorporated into their pricing of products, capital is assigned to cover unexpected costs. The unexpected loss is the difference between the one-year 99.9% VaR and the expected one-year loss.

As mentioned, there are $7 \times 8 = 56$ combinations of risk types and business lines. Banks must estimate one-year 99.9% VaRs for each combination and then aggregate them, using the approach described in Section 12.9, to determine a single one-year 99.9% operational risk VaR measure.

23.3.1 Loss Severity and Loss Frequency

There are two distributions that are important in estimating potential operational risk losses for a risk type/business line combination. One is the *loss frequency distribution* and the other is the *loss severity distribution*. The loss frequency distribution is the distribution of the number of losses observed during one year. The loss severity distribution is the distribution of the size of a loss, given that a loss occurs. It is usually assumed that loss severity and loss frequency are independent.

For loss frequency, the natural probability distribution to use is a Poisson distribution. This distribution assumes that losses happen randomly through time so that in any short period of time Δt there is a probability $\lambda \Delta t$ of a loss occurring. The probability of n losses in T years is

$$e^{-\lambda T} \frac{(\lambda T)^n}{n!}$$

The parameter λ can be estimated as the average number of losses per year. For example, if during a 10-year period there were a total of 12 losses, λ would be estimated as 1.2 per year. A Poisson distribution has the property that the mean frequency of losses equals the variance of the frequency of losses.[4]

For the loss-severity probability distribution, a lognormal probability distribution is often used. The parameters of this probability distribution are the mean and standard deviation of the logarithm of the loss.

The loss-frequency distribution can be combined with the loss severity distribution for each risk type/business line combination to determine a loss distribution. Monte Carlo simulation can be used for this purpose.[5] As mentioned earlier, the usual assumption is that loss severity is independent of loss frequency. On each simulation trial, we proceed as follows:

1. We sample from the frequency distribution to determine the number of loss events ($= n$) in one year.
2. We sample n times from the loss severity distribution to determine the loss experienced for each loss event ($L_1, L_2, \dots, L_n$).
3. We determine the total loss experienced ($= L_1 + L_2 + \cdots + L_n$).

When many simulation trials are used, we obtain a total loss distribution for losses of the type being considered. The 99.9 percentile of the distribution can then be determined.

Figure 23.2 illustrates the procedure. In this example, the expected loss frequency is 3 per year and the loss severity is drawn from a lognormal distribution. The logarithm of each loss ($ millions) is assumed to have a mean of zero and a standard deviation of 0.4. The Excel worksheet used to produce Figure 23.2 is on the author's website: www-2.rotman.utoronto.ca/hull/riskman.

The Basel Committee requires the implementation of AMA to involve four elements: internal data, external data, scenario analysis, and business environment and internal control factors.[6] We will now consider each of these in turn.

23.3.2 Internal Data

Unfortunately, there is usually relatively little historical data available within a bank to estimate loss severity and loss frequency distributions for particular types of losses. Many

---

[4] If the mean frequency is greater than the variance of the frequency, a binomial distribution may be more appropriate. If the mean frequency is less than the variance, a negative binomial distribution (mixed Poisson distribution) may be more appropriate.

[5] Combining the loss severity and loss frequency distribution is a very common problem in insurance. Apart from Monte Carlo simulation, two approaches that are used are Panjer's algorithm and fast Fourier transforms. See H. H. Panjer, "Recursive Evaluation of a Family of Compound Distributions," *ASTIN Bulletin* 12 (1981): 22–29.

[6] See Basel Committee on Banking Supervision, "Operational Risk: Supervisory Guidelines for the Advanced Measurement Approach," June 2011.

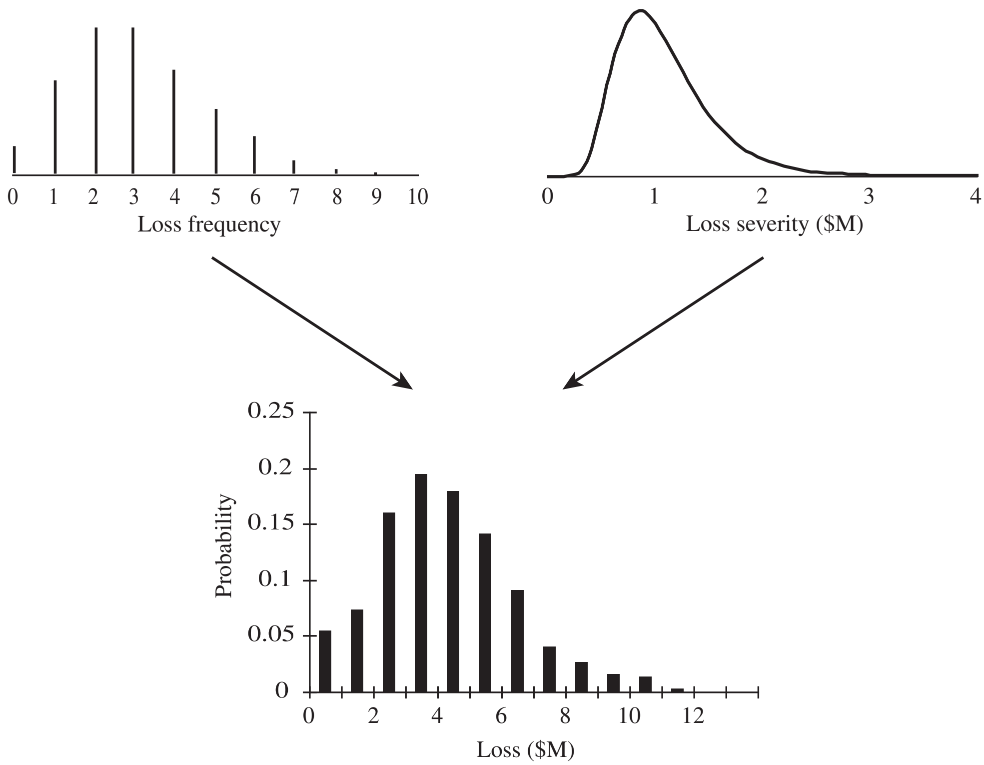


Figure 23.2 Calculation of Loss Distribution from Loss Frequency and Loss Severity

banks have not in the past kept records of operational risk losses. They are doing so now, but it may be some time before a reasonable amount of historical data is available. It is interesting to compare operational risk losses with credit risk losses in this respect. Traditionally, banks have done a much better job at documenting their credit risk losses than their operational risk losses. Also, in the case of credit risks, a bank can rely on a wealth of information published by credit rating agencies to assess probabilities of default and expected losses given default. Similar data on operational risk have not been collected in a systematic way.

There are two types of operational risk losses: high-frequency low-severity losses (HFLSLs) and low-frequency high-severity losses (LFHSLs). An example of the first is credit card fraud losses. An example of the second is rogue trader losses. A bank should focus its attention on LFHSLs. These are what create the tail of the loss distribution. A particular percentile of the total loss distribution can be estimated as the corresponding percentile of the total LFHSL distribution plus the average of the total HFLSL. Another reason for focusing on LFHSLs is that HFLSLs are often taken into account in the pricing of products.

By definition, LFHSLs occur infrequently. Even if good records have been kept, internal data are liable to be inadequate, and must be supplemented with external data and scenario analysis. As we will describe, external data can be used for the loss severity distribution. The loss frequency distribution must be specific to the bank and based on internal data and scenario analysis estimates.

23.3.3 External Data

There are two sources of external data. The first is data consortia, which are companies that facilitate the sharing of data between banks. (The insurance industry has had mechanisms for sharing loss data for many years and banks are now doing this as well.) The second is data vendors, who are in the business of collecting publicly available data in a systematic way. External data increases the amount of data available to a bank for estimating potential losses. It also has the advantage that it can lead to a consideration of types of losses that have never been incurred by the bank, but which have been incurred by other banks.

Both internal and external historical data must be adjusted for inflation. In addition, a scale adjustment should be made to external data. If a bank with a revenue of \$10 billion reports a loss of \$8 million, how should the loss be scaled for a bank with a revenue of \$5 billion? A natural assumption is that a similar loss for a bank with a revenue of \$5 billion would be \$4 million. But this estimate is probably too small. For example, research by Shih et al. (2000) suggests that the effect of firm size on the size of a loss experienced is non-linear.[7] Their estimate is

$$\text{Estimated Loss for Bank A} = \text{Observed Loss for Bank B} \times \left(\frac{\text{Bank A Revenue}}{\text{Bank B Revenue}}\right)^{\alpha}$$

where $\alpha = 0.23$. This means that in our example the bank with a revenue of \$5 billion would experience a loss of $8 \times 0.5^{0.23} = \$6.82$ million. After the appropriate scale adjustment, data obtained through sharing arrangements with other banks can be merged with the bank's own data to obtain a larger sample for determining the loss severity distribution.

The loss data available from data vendors are data taken from public sources such as newspapers and trade journals. Data from vendors cannot be used in the same way as internal data or data obtained through sharing arrangements because they are subject to biases. For example, only large losses are publicly reported, and the larger the loss, the more likely it is to be reported.

Data from vendors are most useful for determining *relative* loss severity. Suppose that a bank has good information on the mean and standard deviation of its loss severity distribution for internal fraud in corporate finance, but not for external fraud in corporate finance or for internal fraud in trading and sales. Suppose that the bank estimates the mean and standard deviation of its loss severity distribution for internal fraud in corporate finance as \$50 million and \$30 million. Suppose further that external data indicate that, for external fraud in corporate finance, the mean severity is twice that for internal fraud in corporate finance and the standard deviation of the severity is 1.5 times as great. In the

---

[7] See J. Shih, A. Samad-Khan, and P. Medapa, "Is the Size of an Operational Loss Related to Firm Size?" *Operational Risk Magazine* 2, no. 1 (January 2000). Whether Shih et al.'s results apply to legal risks is debatable. The size of a settlement in a large lawsuit against a bank can be governed by how much the bank can afford.

absence of a better alternative, the bank might assume that its own severity for external fraud in corporate finance has a mean of $2 \times 50 = \$100$ million and a standard deviation of severity equal to $1.5 \times 30 = \$45$ million. Similarly, if the external data indicate that the mean severity for internal fraud in trading and sales is 2.5 times that for internal fraud in corporate finance and the standard deviation is twice as great, the bank might assume that its own severity for internal fraud in trading and sales has a mean of $2.5 \times 50 = \$125$ million and a standard deviation of $2 \times 30 = \$60$ million.

23.3.4 Scenario Analysis

Scenario analysis has become a key tool for the assessment of operational risk under the AMA. The aim of scenario analysis is to generate scenarios covering the full range of possible LFHSLs. Some of these scenarios might come from the bank's own experience, some might be based on the experience of other banks, some might come from the work of consultants, and some might be generated by the risk management group in conjunction with senior management or business unit managers. The Basel Committee estimates that at many banks the number of scenarios considered that give rise to a loss of greater than 10 million euros is approximately 20 times larger than the number of internal losses of this amount.

An operational risk committee consisting of members of the risk management group and senior management should be asked to estimate loss severity and loss frequency parameters for the scenarios. As explained earlier, a lognormal distribution is often used for loss severity and a Poisson distribution is often used for loss frequency. Data from other banks may be useful for estimating the loss severity parameters. The loss frequency parameters should reflect the controls in place at the bank and the type of business it is doing. They should reflect the views of the members of the operational risk committee. Similarly to Section 22.3, a number of categories of loss frequency can be defined, such as:

1. Scenario happens once every 1,000 years on average ($\lambda = 0.001$)
2. Scenario happens once every 100 years on average ($\lambda = 0.01$)
3. Scenario happens once every 50 years on average ($\lambda = 0.02$)
4. Scenario happens once every 10 years on average ($\lambda = 0.1$)
5. Scenario happens once every 5 years on average ($\lambda = 0.2$)

The committee can be asked to assign each scenario that is developed to one of the categories.

One difference between this scenario analysis and the one in Chapter 22 is that there is no model for determining losses and, if data are not available, the parameters of the loss severity distribution have to be estimated by the committee. One approach is to ask the committee to estimate an average loss and a high loss that the committee is 99% certain will not be exceeded. A lognormal distribution can then be fitted to the estimates.

Fortunately, the operational risk environment does not usually change as fast as the market and credit risk environment so that the amount of work involved in developing scenarios and keeping them up to date should not be as onerous. Nevertheless, the approach we have described does require a great deal of senior management time. The relevant scenarios for one bank are often similar to those for other banks and, to lessen the burden on the operational risk committee, there is the potential for standard scenarios to be developed by consultants or by bank industry associations. However, the loss frequency estimates should always be specific to a bank and reflect the controls in place in the bank and the type of business it is currently doing.

As in the case of market and credit risk stress testing, the advantage of generating scenarios using managerial judgment is that they include losses that the financial institution has never experienced, but could incur. The scenario analysis approach leads to management thinking actively and creatively about potential adverse events. This can have a number of benefits. In some cases, strategies for responding to an event so as to minimize its severity are likely to be developed. In other cases, proposals may be made for reducing the probability of the event occurring at all.

Whether scenario analysis or internal/external data approaches are used, distributions for particular loss types have to be combined to produce a total operational risk loss distribution. The correlations assumed for the losses from different operational risk categories can make a big difference to the one-year 99.9% VaR that is calculated, and therefore to the AMA capital. Chapter 26 builds on material in Section 12.9 to explain how correlations can be used to aggregate economic capital requirements across market risk, credit risk, and operational risk. The same approach can be used to aggregate different operational risk capital requirements. It is often argued that operational risk losses are largely uncorrelated with each other and there is some empirical support for this view. If the zero-correlation assumption is made, Monte Carlo simulation can be used in a straightforward way to sample from the distribution of losses for each scenario to obtain a total distribution of risk losses.

23.3.5 Business Environment and Internal Control Factors

Business environment and internal control factors (BEICFs) should be taken into account when loss severity and loss frequency are estimated. These include the complexity of the business line, the technology used, the pace of change, the level of supervision, staff turnover rates, and so on. For example, factors influencing the estimates made for the rogue trader scenario might be the level of supervision of traders, the level of trade surveillance, and the strengths or weaknesses of the systems used by the middle and back office.

23.4 The Standardized Measurement Approach

One thing the Basel Committee is always looking for is consistency across banks in the application of its regulations. If different banks would calculate quite different capital

Table 23.2 Calculation of BI Component from BI

BI (billions of euros)	BI Component (millions of euros)
0 to 1	110 × BI
1 to 3	110 + 150 × (BI–1)
3 to 10	410 + 190 × (BI–3)
10 to 30	1,740 + 230 × (BI–10)
>30	6,340 + 290 × (BI–30)

charges from the same set of data, the Basel Committee feels that the underlying regulation is not working well. This was the situation with AMA. The Basel Committee hoped that the flexibility in AMA would ultimately lead to the emergence of best practice. In fact, the inherent complexity of AMA and the wide range of internal modeling practices led to a high variability in capital charges.

In March 2016, the Basel Committee issued a consultative document indicating that it was going to replace all previous approaches for determining operational risk capital with a new approach: the Standardized Measurement Approach (SMA).[8] Banks that invested heavily in implementing AMA were understandably annoyed and, as indicated earlier, it may well be the case that some national supervisors will allow them to continue using AMA.

SMA is simpler than AMA. It first defines a quantity known as the *Business Indicator* (BI). This is a measure of bank size, similar to the Gross Income (GI) measure used in the Basel II basic indicator approach. Key differences are:

1. Whereas a negative income (from, for example, trading) would reduce GI, it does not do so for BI.
2. BI includes some items that produce operational risk but are omitted from GI.
3. Operating expenses reduce GI, but because they can lead to operational risk they increase BI.

The *BI Component* for a bank is calculated from its BI using the piecewise linear relationship indicated in Table 23.2. The dollar increase in BI Component for a $1 increase in BI increases as the size of the bank increases. For small banks with BIs in the range 0 to 1 billion euros, it is 0.11. It rises to 0.15 for banks with BIs in the 1 billion to 3 billion euro range, to 0.19 for banks with BIs in the 3 billion to 10 billion euro range, to 0.23 for banks with BIs in the 10 billion to 30 billion euro range, and to 0.29 for banks with BIs above 30 billion euros.

[8] See Basel Committee on Banking Supervision, "Consultative Document: Standardized Approach for Operational Risk," March 2016.

Banks with BIs greater than one billion euros are required to calculate a *Loss Component* as follows:

$$\text{Loss Component} = 7X + 7Y + 5Z$$

where X is the average total annual loss from operational risk, Y is the average total annual loss including only losses greater than 10 million euros, and Z is the average total annual loss including only losses greater than 100 million euros. The averages should be over 10 years of good-quality loss data.[9] The Loss Component and BI Component are calculated so that, according to Basel Committee estimates, Loss Component equals BI Component for an average bank.

The *Internal Loss Multiplier* is calculated as

$$\text{Internal Loss Multiplier} = \ln\left(e - 1 + \frac{\text{Loss Component}}{\text{BI Component}}\right)$$

where e is the exponential constant (2.718). The calculations are designed so that the estimated Internal Loss Multiplier for an average bank is 1. Banks with a loss experience above (below) the industry average will have an Internal Loss Multiplier greater than (less than) 1. The lowest possible Internal Loss Multiplier is $\ln(1.718)$ or 0.514.

Finally, the SMA capital requirement can be calculated. For small banks whose BI is less than 1 billion euros, the SMA capital is set equal to the BI Component (so that these small banks are all regarded as "average"). For other banks it is calculated (in millions of euros) as:

$$\text{SMA Capital} = 110 + (\text{BI Component} - 110) * \text{Internal Loss Multiplier}$$

23.5 Preventing Operational Risk Losses

Up to now we have focused on the measurement of operational risk for the purposes of determining capital. A key part of an operational risk manager's job is to be proactive and try to prevent losses before they occur. We now consider some of the ways this can be done.

One approach is to monitor what is happening at other banks and try and learn from their mistakes. When a \$700 million rogue trader loss happened at a Baltimore subsidiary of Allied Irish Bank in 2002, risk managers throughout the world studied the situation carefully and asked: "Could this happen to us?" Business Snapshot 23.1 describes a situation concerning a British local authority in the late 1980s. It immediately led to all

[9] During a transition period, banks that do not have 10 years of good-quality loss data may use a minimum of five years of data.

BUSINESS SNAPSHOT 23.1

The Hammersmith and Fulham Story

Between 1987 and 1989, the London Borough of Hammersmith and Fulham in Great Britain entered into about 600 interest rate swaps and related transactions with a total notional principal of about six billion pounds. The transactions appear to have been entered into for speculative rather than hedging purposes. The two employees of Hammersmith and Fulham who were responsible for the trades had only a sketchy understanding of the risks they were taking and how the products they were trading worked.

By 1989, because of movements in sterling interest rates, Hammersmith and Fulham had lost several hundred million pounds on the swaps. To the banks on the other side of the transactions, the swaps were worth several hundred million pounds. The banks were concerned about credit risk. They had entered into offsetting swaps to hedge their interest rate risks. If Hammersmith and Fulham defaulted, they would still have to honor their obligations on the offsetting swaps and would take a huge loss.

What actually happened was not a default. Hammersmith and Fulham's auditor asked to have the transactions declared void because Hammersmith and Fulham did not have the authority to enter into the transactions. The British courts agreed. The case was appealed and went all the way to the House of Lords, Britain's highest court. The final decision was that Hammersmith and Fulham did not have the authority to enter into the swaps, but that they ought to have the authority to do so in the future for risk management purposes. Needless to say, banks were furious that their contracts were overturned in this way by the courts.

banks instituting procedures for checking that counterparties had the authority to enter into derivatives transactions.

23.5.1 Causal Relationships

Operational risk managers should try to establish causal relations between decisions taken and operational risk losses. Does increasing the average educational qualifications of employees reduce losses arising from mistakes in the way transactions are processed? Will a new computer system reduce the probabilities of losses from system failures? Are operational risk losses correlated with the employee turnover rate? Can the risk of a rogue trader be reduced by the way responsibilities are divided between different individuals and by the way traders are motivated?

One approach to establishing causal relationships is statistical. If we look at 12 different locations where a bank operates and find a high negative correlation between the education of back-office employees and the cost of mistakes in processing transactions, it might well make sense to do a cost-benefit analysis of changing the educational requirements for a back-office job in some of the locations. In some cases, a detailed analysis of the cause of losses may provide insights. For example, if 40% of computer failures can be attributed to the fact that the current hardware is several years old and less reliable than newer versions, a cost-benefit analysis of upgrading is likely to be useful.

23.5.2 RCSA and KRIs

Risk control and self-assessment (RCSA) is an important way in which banks try to achieve a better understanding of their operational risk exposures. It involves asking the managers of business units to identify their operational risks. Sometimes questionnaires and scorecards designed by senior management or consultants are used.

A by-product of any program to measure and understand operational risk is likely to be the development of key risk indicators (KRIs).[10] Risk indicators are key tools in the management of operational risk. The most important indicators are prospective. They provide an early warning system to track the level of operational risk in the organization. Examples of key risk indicators that could be appropriate in particular situations are

1. Staff turnover
2. Number of failed transactions
3. Number of positions filled by temps
4. Ratio of supervisors to staff
5. Number of open positions
6. Percentage of staff who did not take 10 days consecutive leave in the last 12 months

The hope is that key risk indicators can identify potential problems and allow remedial action to be taken before losses are incurred. It is important for a bank to quantify operational risks, but it is even more important to take actions that control and manage those risks.

23.5.3 E-Mails and Phone Calls

One way in which operational risk costs can be mitigated is by educating employees to be very careful about what they write in e-mails and, when they work on the trading floor, what they say in phone calls. Lawsuits or regulatory investigations contending that a financial institution has behaved inappropriately or illegally are a major source of operational risk. One of the first things that happens when a case is filed is "discovery," where the financial institution is required to provide all relevant internal communications. These

[10]These are sometimes referred to as Business Environment and Internal Control Factors (BEICFs).

have often proved to be very embarrassing and have made it difficult for financial institutions to defend themselves. (Fabrice Tourre's e-mail in Business Snapshot 27.3 is an example of an embarrassing e-mail.) Before sending an e-mail or making a recorded phone call, an employee should consider the question "How could it hurt my employer if this became public knowledge?"

23.6 Allocation of Operational Risk Capital

Operational risk capital should be allocated to business units in a way that encourages them to improve their operational risk management. Methods for doing this are discussed in Sections 12.8 and 26.6. If a business unit can show that it has taken steps to reduce the frequency or severity of a particular risk, it should be allocated less capital. This will have the effect of improving the business unit's return on capital (and possibly lead to the business unit manager receiving an increased bonus).

Note that it is not always optimal for a manager to reduce a particular operational risk. Sometimes the costs of reducing the risk outweigh the benefits of reduced capital so that return on allocated capital decreases. A business unit should be encouraged to make appropriate calculations and determine the amount of operational risk that maximizes return on capital.

The overall result of operational risk assessment and operational risk capital allocation should be that business units become more sensitive to the need for managing operational risk. Hopefully, operational risk management will be seen to be an important part of every manager's job. A key ingredient for the success of any operational risk program is the support of senior management. The Basel Committee on Banking Supervision is very much aware of this. It recommends that a bank's board of directors be involved in the approval of a risk management program and that it review the program on a regular basis.

23.7 Use of Power Law

In Section 10.4, we introduced the power law. This states that, for a wide range of variables

$$\text{Prob}(v > x) = Kx^{-\alpha}$$

where v is the value of the variable, x is a relatively large value of v, and K and α are constants. We covered the theoretical underpinnings of the power law and maximum likelihood estimation procedures when we looked at extreme value theory in Section 13.5.

De Fontnouvelle et al. (2003), using data on losses from vendors, found that the power law holds well for the large losses experienced by banks.[11] This makes the calculation

[11]See P. de Fontnouvelle, V. DeJesus-Rueff, J. Jordan, and E. Rosengren, "Capital and Risk: New Evidence on Implications of Large Operational Risk Losses," *Journal of Money, Credit and Banking* 38, no. 7 (October 2006): 1819–1846.

of VaR with high degrees of confidence, such as 99.9%, easier. Loss data (internal or external) and scenario analysis are employed to estimate the power law parameters using the maximum likelihood approach in Chapter 13. The 99.9 percentile of the loss distribution can then be estimated using equation (13.9).

When loss distributions are aggregated, the distribution with the heaviest tails tends to dominate. This means that the loss with the lowest α defines the extreme tails of the total loss distribution.[12] Therefore, if all we are interested in is calculating the extreme tail of the total operational risk loss distribution, it may only be necessary to consider one or two business line/risk type combinations.

23.8 Insurance

An important decision for operational risk managers is the extent to which operational risks should be insured against. Insurance policies are available on many different kinds of risk ranging from fire losses to rogue trader losses. In Section 3.7 we discussed the moral hazard and adverse selection risks faced by insurance companies. We now review these again in the context of operational risk.

23.8.1 Moral Hazard

One of the risks facing a company that insures a bank against operational risk losses is moral hazard. This is the risk that the existence of the insurance contract will cause the bank to behave differently than it otherwise would. This changed behavior increases the risks to the insurance company. Consider, for example, a bank that insures itself against robberies. As a result of the insurance policy, it may be tempted to be lax in its implementation of security measures—making a robbery more likely than it would otherwise have been.

Insurance companies have traditionally dealt with moral hazard in a number of ways. Typically there is a *deductible* in any insurance policy. This means that the bank is responsible for bearing the first part of any loss. Sometimes there is a *coinsurance provision* in a policy. In this case, the insurance company pays a predetermined percentage (less than 100%) of losses in excess of the deductible. In addition, there is nearly always a *policy limit*. This is a limit on the total liability of the insurer. Consider again a bank that has insured itself against robberies. The existence of deductibles, coinsurance provisions, and policy limits is likely to provide an incentive for a bank not to relax security measures in its branches. The moral hazard problem in rogue trader insurance is discussed in Business Snapshot 23.2.

[12]The parameter ξ in extreme value theory (see Chapter 13) equals $1/\alpha$, so it is the loss distribution with the largest ξ that defines the extreme tails.

BUSINESS SNAPSHOT 23.2

Rogue Trader Insurance

A rogue trader insurance policy presents particularly tricky moral hazard problems. An unscrupulous bank could enter into an insurance contract to protect itself against losses from rogue trader risk and then choose to be lax in its implementation of trading limits. If a trader exceeds the trading limit and makes a large profit, the bank is better off than it would be otherwise. If a large loss results, a claim can be made under the rogue trader insurance policy. Deductibles, coinsurance provisions, and policy limits may mean that the amount recovered is less than the loss incurred by the trader. However, potential net losses to the bank are likely to be far less than potential profits, making the lax trading limits strategy a good bet for the bank.

Given this problem, it is perhaps surprising that some insurance companies do offer rogue trader insurance policies. These companies tend to specify carefully how trading limits are implemented within the bank. They may require that the existence of the insurance policy not be revealed to anyone on the trading floor. They are also likely to want to retain the right to investigate the circumstances underlying any loss.

From the bank's point of view, the lax trading limits strategy we have outlined may be very short-sighted. The bank might well find that the costs of all types of insurance rise significantly as a result of a rogue trader claim. Also, a large rogue trader loss (even if insured) would cause its reputation to suffer.

23.8.2 Adverse Selection

The other major problem facing insurance companies is adverse selection. This is where an insurance company cannot distinguish between good and bad risks. It offers the same price to everyone and inadvertently attracts more of the bad risks. For example, banks without good internal controls are more likely to enter into rogue trader insurance contracts; banks without good internal controls are more likely to buy insurance policies to protect themselves against external fraud.

To overcome the adverse selection problem, an insurance company must try to understand the controls that exist within banks and the losses that have been experienced. As a result of its initial assessment of risks, it may not charge the same premium for the same contract to all banks. As time goes by, it gains more information about the bank's operational risk losses and may increase or reduce the premium charged. This is much the same as the approach adopted by insurance companies when they sell automobile insurance to a driver. At the outset, the insurance company obtains as much information on the driver as possible. As time goes by, it collects more information on the driver's risk (number of accidents, number of speeding tickets, etc.) and modifies the premium charged accordingly.

23.9 Sarbanes–Oxley

Largely as a result of the Enron bankruptcy, the Sarbanes–Oxley Act was passed in the United States in 2002. This provides another dimension to operational risk management for both financial and nonfinancial institutions in the United States. The act requires boards of directors to become much more involved with day-to-day operations. They must monitor internal controls to ensure risks are being assessed and handled well.

The act specifies rules concerning the composition of the board of directors of public companies and lists the responsibilities of the board. It gives the SEC the power to censure the board or give it additional responsibilities. A company's auditors are not allowed to carry out any significant non-auditing services for the company.[13] Audit partners must be rotated. The audit committee of the board must be made aware of alternative accounting treatments. The CEO and CFO must prepare a statement to accompany the audit report to the effect that the financial statements are accurate. The CEO and CFO are required to return bonuses in the event that financial statements are restated. Other rules concern insider trading, disclosure, personal loans to executives, reporting of transactions by directors, and the monitoring of internal controls by directors.

Summary

In 1999, bank supervisors indicated their intention to charge capital for operational risk as part of Basel II. This has led banks to carefully consider how they should measure and manage operational risk. Bank supervisors have identified seven different types of operational risks and eight different business lines. The most sophisticated banks attempted to quantify risks for each of the 56 risk type/business line combinations.

Operational risk losses of a particular type can in theory be treated in much the same way as actuaries treat losses from insurance policies. A frequency of loss distribution and a severity of loss distribution can be estimated and then combined to form a total loss distribution. The advanced measurement approach (AMA) in Basel II required banks to use internal data, external data, scenario analysis, and business environment and risk control factors to estimate these distributions.

In 2016, the Basel Committee announced that AMA was not working as well as hoped and indicated that it was planning to require banks to use a new approach: the Standardized Measurement Approach (SMA). As part of SMA, the Basel Committee developed the Business Indicator, a measure of the size of a bank that is relevant to its

[13]Enron's auditor, Arthur Andersen, provided a wide range of services in addition to auditing. The company did not survive the litigation that followed the downfall of Enron.

potential operational risk exposure. This is used in conjunction with the bank's average loss experience over the previous 10 years to determine operational risk regulatory capital.

Part of the job of operational risk managers is to try to increase awareness among employees of operational risk so that losses are prevented. They should also try to understand what determines operational risk losses and develop key risk indicators to track the level of operational risk in different parts of the organization.

Once operational risk capital has been estimated, it is important to develop procedures for allocating it to business units. This should be done in a way that encourages business units to reduce operational risk when this can be done without incurring excessive costs.

The power law introduced in Chapter 10 seems to apply to operational risk losses. This makes it possible to use extreme value theory (see Chapter 13) to estimate the tails of a loss distribution from empirical data. When several loss distributions are aggregated, it is the loss distribution with the heaviest tail that dominates. In principle, this makes the calculation of VaR for total operational risk easier.

Many operational risks can be insured against. However, most policies include deductibles, coinsurance provisions, and policy limits. As a result, a bank is always left bearing part of any risk itself. Moreover, the way insurance premiums change as time passes is likely to depend on the claims made and other indicators that the insurance company has of how well operational risks are being managed.

The whole process of measuring, managing, and allocating operational risk is still in its infancy. As time goes by and data are accumulated, improvements to our procedures for managing operational risk are likely to emerge. One of the key problems is that there are two sorts of operational risk: high-frequency low-severity risks and low-frequency high-severity risks. The former are relatively easy to quantify, but the latter pose the biggest risks.

Bank supervisors seem to have succeeded in their objective of making banks more sensitive to the importance of operational risk. In many ways, the key benefit of an operational risk management program lies in making managers more aware of the importance of operational risk and perhaps leading to them thinking about it differently.

Further Reading

Bank for International Settlements. "Consultative Document: Standardized Measurement Approach for Operational Risk," March 2016.

Bank for International Settlements. "Operational Risk: Supervisory Guidelines for the Advanced Measurement Approach," June 2011.

Bank for International Settlements. "Sound Practices for the Management and Supervision of Operational Risk," February 2003.

Baud, N., A. Frachot, and T. Roncalli. "Internal Data, External Data and Consortium Data for Operational Risk Management: How to Pool Data Properly." Working Paper, Groupe de Recherche Operationelle, Credit Lyonnais, 2002.

Brunel, V. "Operational Risk Modelled Analytically." *Risk* 27, no. 7 (July 2014): 55–59.
Chorafas, D. N. *Operational Risk Control with Basel II: Basic Principles and Capital Requirements*. Oxford, UK: Elsevier, 2003.
Davis, E., ed. *The Advanced Measurement Approach to Operational Risk*. London: Risk Books, 2006.
De Fontnouvelle, P., V. DeJesus-Rueff, J. Jordan, and E. Rosengren. "Capital and Risk: New Evidence on Implications of Large Operational Risk Losses." *Journal of Money, Credit, and Banking* 38, no. 7 (October 2006): 1819–1846.
Dutta, K., and D. Babbel. "Scenario Analysis in the Measurement of Operational Risk Capital: A Change of Measure Approach." *Journal of Risk and Insurance* 81, no. 2 (2014): 303–334.
Girling, P. X. *Operational Risk Management: A Complete Guide to a Successful Operational Risk Framework*. Hoboken, NJ: John Wiley & Sons, 2013.
Lambrigger, D. D., P. V. Shevchenko, and M. V. Wüthrich. "The Quantification of Operational Risk Using Internal Data, Relevant External Data, and Expert Opinion." *Journal of Operational Risk* 2, no. 3 (Fall 2007): 3–28.
McCormack, P., A. Sheen, and P. Umande. "Managing Operational Risk: Moving Towards the Advanced Measurement Approach." *Journal of Risk Management in Financial Institutions* 7, no. 3 (Summer 2014): 239–256.

Practice Questions and Problems (Answers at End of Book)

23.1 What risks are included by regulators in their definition of operational risks? What risks are not included?

23.2 Suppose that external data show that a loss of $100 million occurred at a bank with annual revenues of $1 billion. Your bank has annual revenues of $3 billion. What is the implication of the external data for losses that could occur at your bank? (Use Shih's result.)

23.3 Suppose that there is a 90% probability that operational risk losses of a certain type will not exceed $20 million. The power law parameter, α, is 0.8. What is the probability of losses exceeding (a) $40 million, (b) $80 million, and (c) $200 million?

23.4 Discuss how moral hazard and adverse selection are handled in car insurance.

23.5 Give two ways Sarbanes–Oxley affects the CEOs of public companies.

23.6 When is a trading loss classified as a market risk and when is it classified as an operational risk?

23.7 Discuss whether there is (a) moral hazard and (b) adverse selection in life insurance contracts.

23.8 What are external loss data? How are they obtained? How are they used in determining operational risk loss distributions for a bank?

23.9 What distributions are commonly used for loss frequency and loss severity?

23.10 Give two examples of key risk indicators that might be monitored by a central operational risk management group within a bank.

23.11 The worksheet used to produce Figure 23.2 is on the author's website. What is the mean and standard deviation of the loss distribution? Modify the inputs to the simulation to test the effect of changing the loss frequency from three to four.

23.12 What are the key inputs to determining regulatory capital under the Standardized Measurement Approach?

23.13 Consider two banks, each with a BI component of 300 million euros. Bank A has suffered 10 operational risk losses of 20 million euros in the past 10 years. Bank B has suffered one 200 million euro loss in the past 10 years. What is the operational risk regulatory capital for each bank under SMA?

Further Questions

23.14 Suppose that there is a 1% probability that operational risk losses of a certain type exceed $10 million. Use the power law to estimate the 99.97% worst-case operational risk loss when the α parameter equals (a) 0.25, (b) 0.5, (c) 0.9, and (d) 1.0.

23.15 Consider the following two events: (a) A bank loses $1 billion from an unexpected lawsuit relating to its transactions with a counterparty and (b) an insurance company loses $1 billion because of an unexpected hurricane in Texas. Suppose that you have the same investment in shares issued by both the bank and the insurance company. Which loss are you more concerned about? Why?

23.16 The worksheet used to produce Figure 23.2 is on the author's website. How does the loss distribution change when the loss severity has a beta distribution with upper bound of five, lower bound of zero, and the other parameters both one?

23.17 A bank has a BI of of 5.5 billion euros. It has had eight operational risk losses in the past 10 years. The amounts of the losses in millions of euros are: 3, 7, 15, 65, 85, 150, 250, and 300. What is the bank's operational risk regulatory capital under SMA?

Chapter 24

Liquidity Risk

The credit crisis that started in the middle of 2007 has emphasized the importance of liquidity risk for both financial institutions and their regulators. Some financial institutions that relied on wholesale deposits for their funding experienced problems as investors lost confidence in financial institutions. Moreover, financial institutions found that some instruments for which there had previously been a liquid market could only be sold at fire-sale prices during the crisis.

It is important to distinguish solvency from liquidity. Solvency refers to a company having more assets than liabilities, so that the value of its equity is positive. Liquidity refers to the ability of a company to make cash payments as they become due. Financial institutions that are solvent can—and sometimes do—fail because of liquidity problems. Consider a bank whose assets are mostly illiquid mortgages. Suppose the assets are financed 90% with deposits and 10% with equity. The bank is comfortably solvent. But it could fail if there is a run on deposits with 25% of depositors suddenly deciding to withdraw their funds. In this chapter we will examine how Northern Rock, a British bank specializing in mortgage lending, failed largely because of liquidity problems of this type.

It is clearly important for financial institutions to manage liquidity carefully. Liquidity needs are uncertain. A financial institution should assess a worst-case liquidity scenario and make sure that it can survive that scenario by either converting assets into cash or raising cash in some other way. The new Basel III requirements are designed to ensure that banks do this.

Liquidity is also an important consideration in trading. A liquid position in an asset is one that can be unwound at short notice. As the market for an asset becomes less liquid, traders are more likely to take losses because they face bigger bid–offer spreads.

For an option or other derivative, it is important for there to be a liquid market for the underlying asset so that the trader has no difficulty in doing the daily trades necessary to maintain delta neutrality (see Chapter 8).

This chapter discusses different aspects of liquidity risk. It considers liquidity trading risk and liquidity funding risk. It also looks at what are termed "liquidity black holes." These are situations where a shock to financial markets causes liquidity to almost completely dry up.

24.1 Liquidity Trading Risk

If a financial institution owns 100, 1,000, 10,000, or even 100,000 shares in IBM, liquidity risk is not a concern. Several million IBM shares trade on the New York Stock Exchange every day. The quoted price of the shares is very close to the price that the financial institution would be able to sell the shares for. However, not all assets are as readily convertible into cash. For example, a $100 million investment in the bonds of a non-investment-grade U.S. company might be quite difficult to sell at close to the market price in one day. Shares and debt of companies in emerging markets are likely to be even less easy to sell.

The price at which a particular asset can be sold depends on

1. The mid-market price of the asset, or an estimate of its value
2. How much of the asset is to be sold
3. How quickly it is to be sold
4. The economic environment

When there is a market maker who quotes a bid and offer price for a financial asset, the financial institution can sell the asset at the bid and buy at the offer. However, it is usually stated (or understood) that a particular quote is good for trades up to a certain size. Above that size, the market maker is likely to increase the bid–offer spread. This is because the market maker knows that, as the size of a trade increases, the difficulty of hedging the exposure created by the trade also increases.

When there is no market maker for a financial instrument, there is still an implicit bid–offer spread. If a financial institution approaches another financial institution (or an interdealer broker) to do a trade, the price depends on which side of the trade it wants to take. The bid–offer spread for an asset can vary from 0.05% of the asset's mid-market price to as much as 5%, or even 10%, of its mid-market price.

The general nature of the relationship between bid quotes, offer quotes, and trade size is indicated in Figure 24.1. The bid price tends to decrease and the offer price tends to increase with the size of a trade. For an instrument where there is a market maker, the bids and offers are the same up to the market maker's size limit and then start to diverge.

Figure 24.1 describes the market for large deals between sophisticated financial institutions. It is interesting to note that bid–offer spreads in the retail market sometimes show the opposite pattern to that in Figure 24.1. Consider, for example, an individual who

Figure 24.1 Bid and Offer Prices as a Function of Quantity Transacted

approaches a branch of a bank wanting to do a foreign exchange transaction or invest money for 90 days. As the size of the transaction increases, the individual is likely to get a better quote.

The price that can be realized for an asset often depends on how quickly it is to be liquidated and the economic environment. Suppose you want to sell your house. Sometimes the real estate market is referred to as a "seller's market." Almost as soon as you put the house on the market, you can expect to get several offers and the house will be sold within a week. In other markets, it may take six months or more to sell the house. In the latter case, if you need to sell the house immediately, you will have to reduce the asking price well below the estimated market value.

Financial assets are similar to real assets as far as this is concerned. Sometimes liquidity is tight (e.g., after the Russian default of 1998 and after the subprime crisis of 2007–2008). Liquidating even a relatively small position can then be time-consuming and is sometimes impossible. On other occasions, there is plenty of liquidity in the market and relatively large positions can be unwound without difficulty.

Liquidating a large position can be affected by what is termed *predatory trading*. This occurs when a market participant, say Company X, has a large position and other market participants guess that it will have to be unwound in the near future. The other market participants attempt to profit by doing similar trades to those they expect from Company X. For example, if it is expected that Company X will have to sell a large position in a particular stock, they short the stock in anticipation of a price decline. This makes it more difficult than it would otherwise be for Company X to exit from its position at competitive prices. To avoid predatory trading, financial institutions emphasize to employees the importance of keeping their positions and their future trading plans confidential. Predatory trading was an issue for the trader known as the London Whale in 2012 (see Business Snapshot 25.1) and for Metallgesellschaft (see Business Snapshot 24.3). In the case of Long-Term Capital Management (see Business Snapshot 22.1), positions were unwound slowly over a period of time under the supervision of the Federal Reserve to avoid predatory trading.

Another problem in the market for financial assets is that, when one financial institution finds that it needs to unwind a position, it is often the case that many other financial institutions with similar positions need to do the same thing. The liquidity normally present in the market then evaporates. This is the "liquidity black hole" phenomenon that will be discussed later in this chapter.

24.1.1 The Importance of Transparency

One thing that the market has learned from the credit crisis of 2007 is that transparency is important for liquidity. If the nature of an asset is uncertain, it is not likely to trade in a liquid market for very long.

As explained in Chapter 6, it had become common practice in the years prior to 2007 to form portfolios of subprime mortgages and other assets and to create financial instruments by securitizing, re-securitizing, and re-re-securitizing the credit risk. Many of the financial instruments were even more complicated than indicated in Chapter 6. Sometimes the ABS CDOs that were created included non-mortgage assets and even tranches from other ABS CDOs. After August 2007, market participants realized that they knew very little about the risks in the instruments they had traded. Moreover, it was very difficult for them to find out very much about this. Belatedly, they realized they had been using credit ratings as a substitute for an understanding of the instruments.

After August 2007, the instruments created from subprime mortgages became illiquid. Financial institutions had no idea how to mark to market investments that they had been scrambling to buy a few months earlier. They realized that they had purchased highly complicated credit derivatives and that they did not have the tools to value them. They lacked both the necessary models and solid information about the assets in the portfolios underlying the derivatives.

Other well-defined credit derivatives, such as credit default swaps, continued to trade actively during the credit crisis. The lesson from all this is that the market can sometimes get carried away trading complex products that are not transparent, but, when it comes to its senses, liquidity for the products soon disappears. When the products do trade again, prices are likely to be low and bid–offer spreads are likely to be high. As mentioned in Chapter 6, in July 2008 Merrill Lynch agreed to sell \$30.6 billion of ABS CDO tranches (previously rated AAA) to Lone Star Funds for 22 cents on the dollar.

24.1.2 Measuring Market Liquidity

One measure of the market liquidity of an asset is its bid–offer spread. This can be measured either as a dollar amount or as a proportion of the asset price. The dollar bid–offer spread is

$$p = \text{Offer price} - \text{Bid price}$$

The proportional bid–offer spread for an asset is defined as

$$s = \frac{\text{Offer price} - \text{Bid price}}{\text{Mid-market price}}$$

where the mid-market price is halfway between the bid and the offer price. Sometimes it is convenient to work with the dollar bid–offer spread, p, and sometimes with the proportional bid–offer spread, s.

The mid-market price can be regarded as the fair price. In liquidating a position in an asset, a financial institution incurs a cost equal to $s\alpha/2$ where α is the dollar (mid-market) value of the position. This reflects the fact that trades are not done at the mid-market price. A buy trade is done at the offer price and a sell trade is done at the bid price.

One measure of the liquidity of a book is how much it would cost to liquidate the book in normal market conditions within a certain time. Suppose that s_i is an estimate of the proportional bid–offer spread in normal market conditions for the ith financial instrument held by a financial institution and α_i is the dollar value of the position in the instrument. Then

$$\text{Cost of liquidation (normal market)} = \sum_{i=1}^{n} \frac{s_i \alpha_i}{2} \tag{24.1}$$

where n is the number of positions. Note that although diversification reduces market risk, it does not necessarily reduce liquidity trading risk. However, as explained earlier, s_i increases with the size of position i. Holding many small positions rather than a few large positions therefore tends to entail less liquidity risk. Setting limits to the size of any one position can therefore be one way of reducing liquidity trading risk.

Example 24.1

Suppose that a financial institution has bought 10 million shares of one company and 50 million ounces of a commodity. The shares are bid \$89.5, offer \$90.5. The commodity is bid \$15, offer \$15.1. The mid-market value of the position in the shares is $90 \times 10 =$ \$900 million. The mid-market value of the position in the commodity is $15.05 \times 50 =$ \$752.50 million. The proportional bid–offer spread for the shares is 1/90 or 0.01111. The proportional bid–offer spread for the commodity is 0.1/15.05 or 0.006645. The cost of liquidation in a normal market is

$$900 \times 0.01111/2 + 752.5 \times 0.006645/2 = 7.5$$

or \$7.5 million.

Another measure of liquidity is the cost of liquidation in stressed market conditions within a certain time period. Define μ_i and σ_i as the mean and standard deviation of the proportional bid–offer spread for the ith financial instrument held. Then:

$$\text{Cost of liquidation (stressed market)} = \sum_{i=1}^{n} \frac{(\mu_i + \lambda\sigma_i)\alpha_i}{2} \tag{24.2}$$

The parameter λ gives the required confidence level for the spread. If, for example, we are interested in considering "worst case" spreads that are exceeded only 1% of the time, and if it is assumed that spreads are normally distributed, then $\lambda = 2.326$.

Example 24.2

Suppose that in Example 24.1 the mean and standard deviation for the bid–offer spread for the shares are \$1.0 and \$2.0, respectively. Suppose further that the mean and standard deviation for the bid–offer spread for the commodity are both \$0.1. The mean and standard deviation for the proportional bid–offer spread for the shares are 0.01111 and 0.02222, respectively. The mean and standard deviation for the proportional bid–offer spread for the commodity are both 0.006645. Assuming the spreads are normally distributed, the cost of liquidation that we are 99% confident will not be exceeded is

$$\begin{aligned} &0.5 \times 900 \times (0.01111 + 2.326 \times 0.02222) \\ &\quad + 0.5 \times 752.5 \times (0.006645 + 2.326 \times 0.006645) = 36.58 \end{aligned}$$

or \$36.58 million. This is almost five times the cost of liquidation in normal market conditions.

In practice, bid–offer spreads are not normally distributed and it may be appropriate to use a value of λ that reflects their empirical distribution. For example, if it is found that the 99 percentile point of the distribution is 3.6 standard deviations above the mean for a particular category of financial instruments, λ can be set equal to 3.6 for those instruments.

Equation (24.2) assumes that spreads in all instruments are perfectly correlated. This may seem overly conservative, but it is not. When liquidity is tight and bid–offer spreads widen, they tend to do so for all instruments. It makes sense for a financial institution to monitor changes in the liquidity of its book by calculating the measures in equations (24.1) and (24.2) on a regular basis. As we have seen, the bid–offer spread depends on how quickly a position is to be liquidated. The measures in equations (24.1) and (24.2) are therefore likely to be decreasing functions of the time period assumed for the liquidation.

24.1.3 *Liquidity-Adjusted VaR*

Market value at risk, which we discussed in Chapters 12 to 14, is designed to calculate an estimate of a "worst case" change in the mark-to-market valuation of the trading book. The measures in equations (24.1) and (24.2) are designed to calculate the cost of liquidating a book if market prices do not change. Although VaR and liquidity risk measures deal with different types of risks, some researchers have suggested combining them into a *liquidity-adjusted VaR* measure. One definition of liquidity-adjusted VaR is regular VaR plus the cost of unwinding positions in a normal market. From equation (24.1) this gives

$$\text{Liquidity-Adjusted VaR} = \text{VaR} + \sum_{i=1}^{n} \frac{s_i \alpha_i}{2} \tag{24.3}$$

Alternatively it can be defined as regular VaR plus the cost of unwinding positions in a stressed market. From equation (24.2) this gives[1]

$$\text{Liquidity-Adjusted VaR} = \text{VaR} + \sum_{i=1}^{n} \frac{(\mu_i + \lambda\sigma_i)\alpha_i}{2}$$

24.1.4 *Unwinding a Position Optimally*

A trader wishing to unwind a large position in a financial instrument has to decide on the best trading strategy. If the position is unwound quickly, the trader will face large bid–offer spreads, but the potential loss from the mid-market price moving against the trader is small. If the trader chooses to take several days to unwind the position, the bid–offer spread the trader faces each day will be lower, but the potential loss from the mid-market price moving against the trader is larger.

This type of problem is discussed by Almgren and Chriss (2001).[2] Suppose that the size of a position is V units and that a trader has to decide how to liquidate it over an n-day period. In this case, it is convenient to define the bid–offer spread in dollars rather than as a proportion. Define the dollar bid–offer spread when the trader trades q units in one day as $p(q)$ dollars. Define q_i as the units traded on day i and x_i as the size of the trader's position at the end of day i ($1 \le i \le n$). It follows that $x_i = x_{i-1} - q_i$ for $1 \le i \le n$ where x_0 is defined as the initial position size, V.

Each trade costs half the bid–offer spread. The total of the costs related to the bid–offer spread is therefore

$$\sum_{i=1}^{n} q_i \frac{p(q_i)}{2}$$

[1] This was suggested in A. Bangia, F. Diebold, T. Schuermann, and J. Stroughair, "Liquidity on the Outside," *Risk* 12 (June): 68–73.

[2] See R. Almgren and N. Chriss, "Optimal Execution of Portfolio Transactions," *Journal of Risk* 3 (Winter 2001): 5–39.

Suppose that the mid-market price changes are normally distributed with a standard deviation of σ per day and trading takes place at the beginning of a day. The variance of the change in the value of the trader's position on day i is $\sigma^2 x_i^2$. Assuming that price changes on successive days are independent, the variance of the change in the value of the position applicable to the unwind is therefore

$$\sum_{i=1}^{n} \sigma^2 x_i^2$$

A trader might reasonably wish to minimize VaR after trading costs have been considered. This corresponds to minimizing something similar to the liquidity-adjusted VaR measure in equation (24.3). The trader's objective is to choose the q_i so that

$$\lambda \sqrt{\sum_{i=1}^{n} \sigma^2 x_i^2} + \sum_{i=1}^{n} q_i \frac{p(q_i)}{2}$$

is minimized subject to

$$\sum_{i=1}^{n} q_i = V$$

with the x_i being calculated from V and the q_i, as indicated above. The parameter λ measures the confidence level in the VaR estimate. For example, when the confidence level is 99%, and daily price changes are assumed to be normally distributed, $\lambda = 2.326$. Once the $p(q)$ function has been estimated, Excel's Solver can be used for the optimization.

Example 24.3

A trader wishes to unwind a position of 100 million units in an asset over five days. Suppose that the bid–offer spread p (measured in dollars) as a function of the daily trading volume is

$$p(q) = a + be^{cq}$$

where $a = 0.1$, $b = 0.05$, and $c = 0.03$ and the amount traded, q, is measured in millions of units.

The standard deviation of the price change per day is \$0.1. A spreadsheet for calculating the optimal strategy can be downloaded from the author's website: www2.rotman.utoronto.ca/~hull/riskman. When the confidence level is 95%, the amounts that should be traded on days 1, 2, 3, 4, and 5 are 48.9, 30.0, 14.1, 5.1, and 1.9 million units, respectively. As the VaR confidence level is reduced, the amounts traded per day show less variability. For example, when the confidence level is 90%, they are 45.0, 29.1, 15.6, 7.0, and 3.3 million units, respectively. When the confidence level is 75% they are 36.1, 26.2, 17.7, 11.6, and 8.4 million units, respectively. In the limit when the confidence level is set equal to 50%, so that the trader is interested only in expected costs, not in the standard deviation of costs, 20 million units should be traded each day.

As this example illustrates, when a position is to be closed out over n days, more than $1/n$ of the position should be traded on the first day. This is because the longer any part of the position is held, the greater the risk of adverse market moves.

24.1.5 Other Measures of Market Liquidity

So far we have focused on bid–offer spread as a measure of market liquidity. Many other measures have been proposed. The volume of trading per day (i.e., the number of times the asset trades in a day) is an important measure. When an asset is highly illiquid, the volume of trading in a day is often zero. The price impact of a trade of a certain size is another measure. A measure somewhat similar to this, but more easily calculated, was proposed by Amihud (2002).[3] It is the average of

$$\frac{\text{Absolute value of daily return}}{\text{Daily dollar volume}}$$

over all days in the period considered. This measure is widely used by researchers. Amihud shows that an asset's expected return increases as its liquidity decreases. In other words, investors do get compensated for illiquidity.

24.2 Liquidity Funding Risk

We now move on to consider liquidity funding risk. This is the financial institution's ability to meet its cash needs as they arise. As mentioned at the outset of this chapter, liquidity is not the same as solvency. Financial institutions that are solvent (i.e., have positive equity) can, and sometimes do, fail because of liquidity problems. Northern Rock, a British mortgage lender, is a case in point (see Business Snapshot 24.1).

Liquidity funding problems at a financial institution can be caused by:

1. Liquidity stresses in the economy (e.g., a flight to quality such as that seen during the 2007 to 2009 crisis). Investors are then reluctant to provide funding in situations where there is any credit risk at all.

[3] See Y. Amihud, "Illiquidity and Stock Returns: Cross-Section and Time-Series Effects," *Journal of Financial Markets* 5 (2002): 31–56.

2. Overly aggressive funding decisions. There is a tendency for all financial institutions to use short-term instruments to fund long-term needs, creating a liquidity mismatch. Financial institutions need to ask themselves: "How much of a mismatch is too much?"
3. A poor financial performance, leading to a lack of confidence. This can result in a loss of deposits and difficulties in rolling over funding.

Often, when a company experiences severe liquidity problems, all three of these have occurred at the same time. The key to managing liquidity risk is predicting cash needs and ensuring that they can be met in adverse scenarios. Some cash needs are predictable. For example, if a bank has issued a bond, it knows when coupons will have to be paid. Others, such as those associated with withdrawals of deposits by retail customers and drawdowns by corporations on lines of credit that the bank has granted, are less predictable. As the financial instruments entered into by financial institutions have become more complex, cash needs have become more difficult to predict. For example, downgrade triggers (see Section 20.2.3), guarantees provided by a financial institution, and possible defaults by counterparties in derivatives transactions can have an unexpected impact on cash resources.

24.2.1 Sources of Liquidity

The main sources of liquidity for a financial institution are:

1. Holdings of cash and Treasury securities
2. The ability to liquidate trading book positions
3. The ability to borrow money at short notice
4. The ability to offer favorable terms to attract retail and wholesale deposits at short notice
5. The ability to securitize assets (such as loans) at short notice
6. Borrowings from the central bank

We now consider each of these in turn.

Cash and Treasury Securities Cash and Treasury securities are excellent sources of liquidity. Cash is of course always available to meet liquidity needs and Treasury securities issued by countries such as the United States and the United Kingdom can generally be converted into cash at short notice without any problem. However, cash and Treasury securities are relatively expensive sources of liquidity. There is a trade-off between the liquidity of an asset and the return it provides. In order to be profitable, a financial institution needs to invest in assets such as loans to corporations that provide a higher rate of return than Treasury instruments. There is therefore a limit to the cash and Treasury securities that can reasonably be held.

BUSINESS SNAPSHOT 24.1

Northern Rock

Northern Rock, a British bank, was founded in 1997 when the Northern Rock Building Society floated shares on the London Stock Exchange. In 2007, it was one of the top five mortgage lenders in the United Kingdom. It had 76 branches and offered deposit accounts, savings accounts, loans, and house/contents insurance. The bank grew rapidly between 1997 and 2007. Some of its mortgages were securitized through a subsidiary, Granite, that was based in the Channel Islands.

Northern Rock relied on selling short-term debt instruments for much of its funding. Following the subprime crisis of August 2007, the bank found it very difficult to replace maturing instruments because institutional investors became very nervous about lending to banks that were heavily involved in the mortgage business. The bank's assets were sufficient to cover its liabilities so it was not insolvent. To quote from the Financial Services Authority (FSA) in September 2007: "The FSA judges that Northern Rock is solvent, exceeds its regulatory capital requirement, and has a good quality loan book." But Northern Rock's inability to fund itself was a serious problem. It approached the Bank of England for funding on September 12, 2007, and borrowed about £3 billion from the Tripartite Authority (Bank of England, the Financial Services Authority, and HM Treasury) in the following few days.

On September 13, 2007, the BBC business editor Robert Peston broke the news that the bank had requested emergency support from the Bank of England. On Friday, September 14, there was a run on the bank. Thousands of people lined up for hours to withdraw their funds. This was the first run on a British bank in 150 years. Some customers held their funds in an "Internet-only" account, which they were unable to access due to the volume of customers trying to log on. On Monday, September 17, worried savers continued to withdraw their funds. An estimated £2 billion was withdrawn between September 12 and September 17, 2007.

Depositor insurance in the UK guaranteed 100% of the first £2,000 and 90% of the next £33,000. Late on September 17, 2007, the British Chancellor of the Exchequer, Alistair Darling, announced that the British government and the Bank of England would guarantee all deposits held at Northern Rock. As a result of this announcement and subsequent advertisements in major UK newspapers, the lines outside Northern Rock's branches gradually disappeared. Northern Rock's shares, which had fallen from £12 earlier in the year to £2.67, rose 16% on Mr. Darling's announcement.

During the months following September 12, 2007, Northern Rock's emergency borrowing requirement increased. The Bank of England insisted on a penalty rate of interest to discourage other banks from taking excessive risks. Northern Rock raised some funds by selling assets, but by February 2008 the emergency borrowing reached £25 billion. The bank was then nationalized and the management of the bank was changed. It was split into Northern Rock plc and Northern Rock (Asset Management), with the company's bad debt being put in Northern Rock (Asset Management). In November 2011, Northern Rock plc was bought from the British government for £747 million by the Virgin Group, which is headed by the colorful entrepreneur Sir Richard Branson.

The Northern Rock story illustrates just how quickly liquidity problems can lead to a bank spiraling downward. If the bank had been managed a little more conservatively and had paid more attention to ensuring that it had access to funding, it might have survived.

Liquidating Trading Book Positions Liquidity funding risk is related to liquidity trading risk, considered in Section 24.1, because one way a financial institution can meet its funding requirements is by liquidating part of its trading book. It is therefore important for a financial institution to quantify the liquidity of its trading book so that it knows how easy it would be to use the book to raise cash. The financial institution wants to make sure that it will be able to survive stressed market conditions where there is a general shortage of liquidity. The financial institution's analysis should therefore be based on stressed market conditions, not normal market conditions. This is the reason why the measures discussed in the previous section concerned with bid–offer spreads in stressed markets are important.

Ability to Borrow When markets are unstressed, a creditworthy bank usually has no problem in borrowing money, but in stressed market conditions there is a heightened aversion to risk. This leads to higher interest rates, shorter maturities for loans, and in some cases a refusal to provide funds at all. Financial institutions should monitor the assets that can be pledged as collateral for loans at short notice. A financial institution can (at a cost) mitigate its funding risks somewhat by arranging lines of credit. For example, Countrywide, an originator of mortgages in the United States, had a syndicated loan facility of \$11.5 billion, which it was able to use during the credit crisis of 2007. (This helped keep the company afloat, but it still experienced significant problems and was taken over by Bank of America in January 2008.) As Business Snapshot 24.1 shows, Northern Rock, a similar British mortgage lender, did not fare so well.

Wholesale and Retail Deposits Wholesale deposits are a more volatile source of funding than retail deposits and can disappear quickly in stressed market conditions. Even retail deposits are not as stable as they used to be because it is very easy to compare interest rates offered by different financial institutions and make transfers via the Internet. Unfortunately, liquidity problems tend to be market-wide rather than something that affects one or two financial institutions. When one financial institution wants to increase its retail or wholesale deposit base for liquidity reasons by offering more attractive rates of interest, others usually want to do the same thing and the increased funding is likely to be difficult to achieve.

Securitization As mentioned in Chapter 2, banks have found the "originate-to-distribute" model attractive. Rather than keep illiquid assets such as loans on their balance sheet, they have securitized them. The structures developed for doing this were discussed in Chapter 6. Prior to August 2007, securitization was an important source of liquidity for banks. However, this source of liquidity dried up almost overnight in August 2007 as investors decided that the securitized products were too risky. "Originate-to-distribute" was no longer a viable strategy, and banks had to fund their loans. Not surprisingly, banks became a lot less willing to lend.

Securitization led to other liquidity problems in August 2007. Banks had entered into liquidity backstop arrangements on the asset-backed commercial paper (ABCP) that was used to fund debt instruments, such as mortgages, prior to their securitization. When buyers could not be found, they had to buy the instruments themselves. In some cases, in order to avoid their reputations being adversely impacted, they had to provide financial support to conduits and other off-balance-sheet vehicles that were involved in securitization, even though they were not legally required to do so.

Central Bank Borrowing Central banks (e.g., the Federal Reserve Board in the United States, the Bank of England in the UK, or the European Central Bank) are often referred to as "lenders of last resort." When commercial banks are experiencing financial difficulties, central banks are prepared to lend money to maintain the health of the financial system. Collateral has to be posted by the borrowers and the central bank typically applies a haircut (i.e., it lends less than 100% of the value of the collateral) and may charge a relatively high rate of interest. In March 2008, after the failure of Bear Stearns (which was taken over by JPMorgan Chase), the Federal Reserve Board extended its borrowing facility to investment banks as well as commercial banks.[4] Later, it also made the facility available to Fannie Mae and Freddie Mac (which were taken over by the government in September 2008).

[4] Central banks are concerned about the failure of investment banks because of systemic risk (see Business Snapshot 15.1). Investment banks have derivatives contracts with other investment banks and with commercial banks. There is a danger that, because of the huge amount of trading between banks, a failure by an investment bank will have a ripple effect throughout the financial sector leading to a failure by commercial banks.

BUSINESS SNAPSHOT 24.2

Ashanti Goldfields

Ashanti Goldfields, a West African gold-mining company based in Ghana, experienced problems resulting from its hedging program in 1999. It had sought to protect its shareholders from gold price declines by selling gold forward. On September 26, 1999, 15 European central banks surprised the market with an announcement that they would limit their gold sales over the following five years. The price of gold jumped up over 25%. Ashanti was unable to meet margin calls and this resulted in a major restructuring, which included the sale of a mine, a dilution of the interest of its equity shareholders, and a restructuring of its hedge positions.

Different central banks apply different rules. Following the credit crisis of August 2007, the haircuts used by the European Central Bank (ECB) were lower than those of other central banks. As a result, some British banks preferred to borrow from the European Central Bank (ECB) rather than the Bank of England. (There are even stories of North American banks contemplating the setting up of subsidiaries in Ireland to access the ECB.) By September 2008, the ECB had lent 467 billion euros and announced that it would apply larger haircuts in the future.

Banks try to keep their borrowing from a central bank a secret. There is a danger that the use of central bank borrowings will be interpreted by the market as a sign that the bank is experiencing financial difficulties with the result that other sources of liquidity dry up. As Business Snapshot 24.1 discusses, news that Northern Rock required emergency borrowing led to an immediate run on the bank, exacerbating its liquidity problems.

Hedging Issues Liquidity problems are liable to arise when companies hedge illiquid assets with contracts that are subject to margin requirements. As indicated in Business Snapshot 8.1, gold-mining companies often hedge their risks by entering into agreements with financial institutions to sell gold forward for two or three years. Often the gold-mining company is required to post margin and the amount of the margin required is calculated every day to reflect the value of its forward agreements. If the price of gold rises fast, the forward agreements lose money and result in big margin calls being made by the financial institution on the gold-mining company. The losses on the forward agreements are offset by increases in the value of the gold in the ground—but this is an illiquid asset. As discussed in Business Snapshot 24.2, Ashanti Goldfields was forced to undertake a major restructuring when it could not meet margin calls after a sharp rise in the price of gold.

BUSINESS SNAPSHOT 24.3

Metallgesellschaft

In the early 1990s, Metallgesellschaft (MG) sold a huge volume of 5- to 10-year heating oil and gasoline fixed-price supply contracts to its customers at six to eight cents above market prices. It hedged its exposure with long positions in short-dated futures contracts that were rolled forward. As it turned out, the price of oil fell and there were margin calls on the futures positions. MG's trading was made more difficult by the fact that its trades were very large and were anticipated by others.

Considerable short-term cash flow pressures were placed on MG. The members of MG who devised the hedging strategy argued that these short-term cash outflows were offset by positive cash flows that would ultimately be realized on the long-term fixed-price contracts. However, the company's senior management and its bankers became concerned about the huge cash drain. As a result, the company closed out all the hedge positions and agreed with its customers that the fixed-price contracts would be abandoned. The outcome was a loss to MG of $1.33 billion.

Another extreme example of a liquidity funding problem caused by hedging is provided by a German company, Metallgesellschaft, that entered into profitable fixed-price oil and gas contracts with its customers (see Business Snapshot 24.3). The lesson from the Ashanti and Metallgesellschaft episodes is not that companies should not use forward and futures contracts for hedging, but rather that they should ensure that they have access to funding to handle the cash flow mismatches that might arise in extreme circumstances.

24.2.2 Reserve Requirements

In some countries there are *reserve requirements* that require banks to keep a certain percentage of deposits as cash in the bank's vault or on deposit with the central bank. The reserve requirement applies only to transaction deposits (in essence, those made to a checking account). For large banks in the United States, the reserve requirement is currently about 10%. Some countries, such as Canada and the United Kingdom, have no compulsory reserve requirements. Others have higher compulsory reserve requirements than the United States.

In addition to ensuring that banks keep a minimum amount of liquidity, reserve requirements affect the money supply. When the reserve requirement is 10%, a $100 deposit leads to $90 of lending, which leads to a further $90 of deposits in the banking system, which leads to further $81 of lending, and so on. As this process continues, the total money supply (M1) that is created is $90 + 81 + 72.9 + \dots$ or $900. If the reserve requirement is 20%, a \$100 deposit leads to \$80 of lending, which leads to \$64 of lending, and so on. The total increase in the money supply is 80 + 64 + 51.2 + … or \$400. Most countries do not use the reserve requirement as a way of controlling the money supply. An exception appears to be China, where the reserve requirement is changed frequently.

24.2.3 Regulation

As explained in Chapter 16, Basel III introduced two liquidity risk requirements: the liquidity coverage ratio (LCR) and the net stable funding ratio (NSFR).

The LCR requirement is

$$\frac{\text{High-quality liquid assets}}{\text{Net cash outflows in a 30-day period}} \geq 100\%$$

The 30-day period considered in the calculation of LCR is one of acute stress involving a downgrade of three notches (e.g., from AA+ to A+), a partial loss of deposits, a complete loss of wholesale funding, increased haircuts on secured funding, and drawdowns on lines of credit. LCR was implemented in stages between 2015 and 2019. (The required ratio was 60% in 2015 and 100% in 2019.)

The NSFR requirement is

$$\frac{\text{Amount of stable funding}}{\text{Required amount of stable funding}} \geq 100\%$$

The numerator is calculated by multiplying each category of funding (capital, wholesale deposits, retail deposits, etc.) by an available stable funding (ASF) factor, reflecting their stability (see Table 16.4). The denominator is calculated from the assets and off-balance-sheet items requiring funding. Each category of these is multiplied by a required stable funding (RSF) factor to reflect the permanence of the funding (see Table 16.5). The implementation date for the NSFR requirement is January 1, 2018.

Following the liquidity crisis of 2007, bank regulators issued a revised set of principles on how banks should manage liquidity.[5] These are as follows:

1. A bank is responsible for the sound management of liquidity risk. A bank should establish a robust liquidity risk management framework that ensures it maintains sufficient liquidity, including a cushion of unencumbered, high-quality liquid assets, to withstand a range of stress events, including those involving the loss or impairment of both unsecured and secured funding sources. Supervisors should assess the adequacy of both a bank's liquidity risk management framework and its liquidity position and should take prompt action if a bank is deficient in either area in order to protect depositors and to limit potential damage to the financial system.

[5] See Bank for International Settlements, "Principles for Sound Liquidity Risk Management and Supervision," September 2008.

2. A bank should clearly articulate a liquidity risk tolerance that is appropriate for its business strategy and its role in the financial system.
3. Senior management should develop a strategy, policies, and practices to manage liquidity risk in accordance with the risk tolerance and to ensure that the bank maintains sufficient liquidity. Senior management should continuously review information on the bank's liquidity developments and report to the board of directors on a regular basis. A bank's board of directors should review and approve the strategy, policies, and practices related to the management of liquidity at least annually and ensure that senior management manages liquidity risk effectively.
4. A bank should incorporate liquidity costs, benefits, and risks in the internal pricing, performance measurement, and new product approval process for all significant business activities (both on- and off-balance-sheet), thereby aligning the risk-taking incentives of individual business lines with the liquidity risk exposures their activities create for the bank as a whole.
5. A bank should have a sound process for identifying, measuring, monitoring, and controlling liquidity risk. This process should include a robust framework for comprehensively projecting cash flows arising from assets, liabilities, and off-balance-sheet items over an appropriate set of time horizons.
6. A bank should actively monitor and control liquidity risk exposures and funding needs within and across legal entities, business lines, and currencies, taking into account legal, regulatory, and operational limitations to the transferability of liquidity.
7. A bank should establish a funding strategy that provides effective diversification in the sources and tenor of funding. It should maintain an ongoing presence in its chosen funding markets and strong relationships with funds providers to promote effective diversification of funding sources. A bank should regularly gauge its capacity to raise funds quickly from each source. It should identify the main factors that affect its ability to raise funds and monitor those factors closely to ensure that estimates of fund-raising capacity remain valid.
8. A bank should actively manage its intraday liquidity positions and risks to meet payment and settlement obligations on a timely basis under both normal and stressed conditions and thus contribute to the smooth functioning of payment and settlement systems.
9. A bank should actively manage its collateral positions, differentiating between encumbered and unencumbered assets. A bank should monitor the legal entity and physical location where collateral is held and how it may be mobilized in a timely manner.
10. A bank should conduct stress tests on a regular basis for a variety of short-term and protracted institution-specific and market-wide stress scenarios (individually and in combination) to identify sources of potential liquidity strain and to ensure that current exposures remain in accordance with a bank's established liquidity risk tolerance. A bank should use stress test outcomes to adjust its liquidity risk management strategies, policies, and positions and to develop effective contingency plans.

11. A bank should have a formal contingency funding plan (CFP) that clearly sets out the strategies for addressing liquidity shortfalls in emergency situations. A CFP should outline policies to manage a range of stress environments, establish clear lines of responsibility, include clear invocation and escalation procedures, and be regularly tested and updated to ensure that it is operationally robust.

12. A bank should maintain a cushion of unencumbered, high-quality liquid assets to be held as insurance against a range of liquidity stress scenarios, including those that involve the loss or impairment of unsecured and typically available secured funding sources. There should be no legal, regulatory, or operational impediment to using these assets to obtain funding.

13. A bank should publicly disclose information on a regular basis that enables market participants to make an informed judgment about the soundness of its liquidity risk management framework and liquidity position.

Recommendations for banks supervisors are:

14. Supervisors should regularly perform a comprehensive assessment of a bank's overall liquidity risk management framework and liquidity position to determine whether they deliver an adequate level of resilience to liquidity stress given the bank's role in the financial system.

15. Supervisors should supplement their regular assessments of a bank's liquidity risk management framework and liquidity position by monitoring a combination of internal reports, prudential reports, and market information.

16. Supervisors should intervene to require effective and timely remedial action by a bank to address deficiencies in its liquidity risk management processes or liquidity position.

17. Supervisors should communicate with other supervisors and public authorities, such as central banks, both within and across national borders, to facilitate effective cooperation regarding the supervision and oversight of liquidity risk management. Communication should occur regularly during normal times, with the nature and frequency of the information sharing increasing as appropriate during times of stress.

24.3 Liquidity Black Holes

It is sometimes argued that technological and other developments have led to an improvement in the liquidity of financial markets. This is questionable. It is true that bid–offer spreads have on average declined. But there has also been an increasing tendency for situations to develop where almost everyone wants to do the same type of trade at the same time. The result has been that what are referred to as "liquidity black holes" occur with increasing frequency.[6] A liquidity black hole describes a situation where liquidity

[6] See A. D. Persaud, ed., *Liquidity Black Holes: Understanding, Quantifying and Managing Financial Liquidity Risk* (London: Risk Books, 1999).

has dried up in a particular market because everyone wants to sell and no one wants to buy, or vice versa. It is sometimes also referred to as a "crowded exit."[7]

In a well-functioning market, the market may change its opinion about the price of an asset because of new information. However, the price does not overreact. If a price decrease is too great, traders will quickly move in and buy the asset and a new equilibrium price will be established. A liquidity black hole is created when a price decline causes more market participants to want to sell, driving prices well below where they will eventually settle. During the sell-off, liquidity dries up and the asset can be sold only at a fire-sale price.[8]

24.3.1 Positive and Negative Feedback Traders

Changes in the liquidity of financial markets are driven by the behavior of traders. There are two sorts of traders in the market: negative feedback traders and positive feedback traders.[9] Negative feedback traders buy when prices fall and sell when prices rise; positive feedback traders sell when prices fall and buy when prices rise.

In liquid markets, negative feedback traders dominate the trading. If the price of an asset gets unreasonably low, traders will move in and buy. This creates demand for the asset that restores the price to a more reasonable level. Similarly, if the price of an asset gets unreasonably high, traders will sell. This creates supply of the asset that also restores the price to a more reasonable level. The result is that the market is liquid with reasonable prices and a good balance of buyers and sellers.

When positive feedback traders dominate the trading, market prices are liable to be unstable and the market may become one-sided and illiquid. A reduction in the price of an asset causes traders to sell. This results in prices falling further and more selling. An increase in the price of an asset causes traders to buy. This results in the price of the asset increasing further and more buying.

There are a number of reasons why positive feedback trading exists. For example:

1. *Trend trading*. Trend traders attempt to identify trends in an asset price. They buy when the asset price appears to be trending up and sell when it appears to be trending down. A related strategy is *breakout trading*, which involves trading when an asset's price moves outside an established range. For example, if a stock has been trading in the range of $25 to $30 for the last six months, traders might be inclined to buy it as soon as it moves above $30 and sell it as soon as it moves below $25.
2. *Stop-loss rules*. Traders often have rules to limit their losses. When the price of an asset that is owned falls below a certain level, they automatically sell to limit their

[7] See, for example, J. Clunie, *Predatory Trading and Crowded Exits: New Thinking on Market Volatility* (Petersfield, UK: Harriman House, 2010).
[8] Liquidity black holes tend to be associated with price decreases, but it is in theory also possible for them to occur when there are price increases.
[9] This is a simplification of reality to help understand the dynamics of markets. Some traders follow complicated strategies that cannot be classified as positive feedback or negative feedback.

losses. These rules are known as "stop-loss" rules and are a source of positive feedback trading that is always present in the market.

3. *Dynamic hedging*. Chapter 8 explains how options traders maintain a delta-neutral position using dynamic hedging. In particular, Tables 8.2 and 8.3 show how a trader would hedge a short position in an option over a 20-week period. Hedging a short option position (call or put) involves buying after a price rise and selling after a price decline. This is positive feedback trading that has the potential to reduce liquidity. (By contrast, dynamically hedging a long position in a call or put option involves selling after a price rise and buying after a price decline. This is negative feedback trading and should not interfere with market liquidity.) Any situation where banks have a large short option position has the potential to destabilize the market and cause illiquidity. As discussed in Business Snapshot 3.1, at one point banks sold a huge volume of options on long-term interest rates to British insurance companies. As the banks hedged their risks, the behavior of long-term interest rates in the UK was dramatically affected.
4. *Creating options synthetically*. Hedging a short position in an option is equivalent to creating a long position in the same option synthetically. It follows that a financial institution can create a long option position synthetically by doing the same sort of trading as it would do if it were hedging a short option position. This leads to positive feedback trading that can cause market instability and illiquidity. The classic example here is the stock market crash of October 1987. In the period leading up to the crash, the stock market had done very well. Increasing numbers of portfolio managers were using commercially available programs to synthetically create put options on their portfolios. These programs told them to sell part of their portfolio immediately after a price decline and buy it back immediately after a price increase. The result, as indicated in Business Snapshot 24.4, was a liquidity black hole where prices plunged on October 19, 1987. In this case, the liquidity black hole was relatively short-lived. Within four months the market recovered to close to its pre-crash level.
5. *Margins*. A big movement in market variables, particularly for traders who are highly leveraged, may lead to margin calls that cannot be met. This forces traders to close out their positions, which reinforces the underlying move in the market variables. It is likely that volatility increases. This may exacerbate the situation because it leads to exchanges increasing their margin requirements.
6. *Predatory trading*. This was mentioned in Section 24.1. If traders know that an investor must sell large quantities of a certain asset, they know that the price of the asset is likely to decrease. They therefore short the asset. This reinforces the price decline and results in the price falling even further than it would otherwise do. To avoid predatory trading, large positions must usually be unwound slowly.
7. *LTCM*. The failure of the hedge fund Long-Term Capital Management (LTCM), as outlined in Business Snapshot 22.1, provides an example of positive feedback trading. One type of LTCM's trade was "relative value fixed income." LTCM would take a short position in a liquid bond and a long position in a similar illiquid bond, and wait for the prices to move close together. After the Russian default in 1998, the prices of

BUSINESS SNAPSHOT 24.4

The Crash of 1987

On Monday, October 19, 1987, the Dow Jones Industrial Average dropped by more than 20%. Portfolio insurance played a major role in this crash. In October 1987, portfolios involving over $60 billion of equity assets were being managed with trading rules that were designed to synthetically create put options on the portfolios. These trading rules involved selling equities (or selling index futures) when the market declined and buying equities (or buying equity futures) when the market rose.

During the period Wednesday, October 14, 1987, to Friday, October 16, 1987, the market declined by about 10%, with much of this decline taking place on Friday afternoon. The portfolio insurance rules should have generated at least $12 billion of equity or index futures sales as a result of this decline. In fact, portfolio insurers had time to sell only $4 billion and they approached the following week with huge amounts of selling already dictated by their models. It is estimated that on Monday, October 19, sell programs by three portfolio insurers accounted for almost 10% of the sales on the New York Stock Exchange, and that portfolio insurance sales amounted to 21.3% of all sales in index futures markets. It is likely that the decline in equity prices was exacerbated by investors other than portfolio insurers selling heavily because they anticipated the actions of portfolio insurers.

Because the market declined so fast and the stock exchange systems were overloaded, many portfolio insurers were unable to execute the trades generated by their models and failed to obtain the protection they required. Needless to say, the popularity of portfolio insurance schemes has declined significantly since 1987. One of the lessons from this story is that it is dangerous to follow a particular trading strategy—even a hedging strategy—when many other market participants are doing the same thing. To quote from a report on the crash, "Liquidity sufficient to absorb the limited selling demands of investors became an illusion of liquidity by massive selling, as everyone showed up on the same side of the market at once. Ironically, it was this illusion of liquidity which led certain similarly motivated investors, such as portfolio insurers, to adopt strategies which call for liquidity far in excess of what the market could supply."

illiquid instruments declined relative to similar liquid instruments. LTCM (and other companies that were following similar strategies to LTCM) were highly leveraged and unable to meet margin calls. They were forced to close out their positions. This involved buying the liquid bonds and selling the illiquid bonds. This reinforced the flight to quality and made the prices of illiquid and liquid bonds diverge even further.

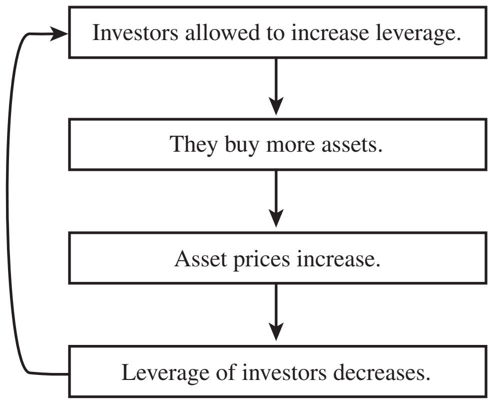


Figure 24.2 Leveraging

24.3.2 Leveraging and Deleveraging

A phenomenon in the market is leveraging and deleveraging. This is illustrated in Figures 24.2 and 24.3. When banks are awash with liquidity (e.g., because they have developed ways of securitizing assets or because deposit levels are higher than usual), they make credit easily available to businesses, investors, and consumers. Credit spreads decrease. The easy availability of credit increases demand for both financial and nonfinancial assets and the prices of these assets rise. Assets are often pledged as collateral for the loans that are used to finance them. When the prices of the assets rise, the collateral underlying loans (when measured at market prices) is greater and borrowing can increase further. This leads to further asset purchases and a repeat of the cycle. This cycle is referred to as "leveraging" because it leads to more borrowing throughout the economy.

Deleveraging is the opposite process to leveraging. Banks find themselves less liquid for some reason (e.g., because there is less demand for the products of securitization). They become more reluctant to lend money. Credit spreads increase. There is less demand for both nonfinancial and financial assets and their prices decrease. The value of the

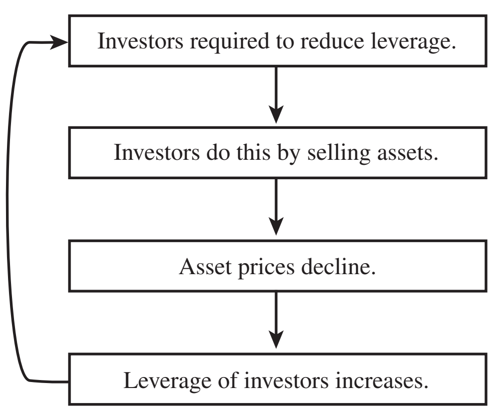


Figure 24.3 Deleveraging

collateral supporting loans decreases and banks reduce lines of credit. This leads to asset sales being necessary and a further reduction in asset prices.

The period leading up to 2007 was characterized by leveraging for many of the world's economies. Credit spreads declined and it was relatively easy to borrow money for a wide range of different purposes. From the middle of 2007 onward, the situation changed and the deleveraging process shown in Figure 24.3 started. Credit spreads increased, it became much less easy to borrow money, and asset prices decreased.

Hedge funds are particularly affected by the leveraging–deleveraging cycle. Consider a hedge fund that is able to borrow 20 times its equity during the pre-2007 period. Soon after the middle of 2007, the hedge fund might get a call from its prime broker telling it to reduce leverage to, say, five times equity. It can only do this by selling assets. Asset prices decrease as a result of what this hedge fund, and other hedge funds, are doing. The hedge fund's equity declines and further sales are necessary.

24.3.3 Irrational Exuberance

The term "irrational exuberance" was used by Alan Greenspan, Federal Reserve Board chairman, in a speech in December 1996 when, in reference to the stock market, he asked, "How do we know when irrational exuberance has unduly escalated asset values?" (The phrase has been remembered because the speech was followed by declines in stock prices worldwide.) Most liquidity black holes can be traced to irrational exuberance of one sort or another. What happens is that traders working for many different financial institutions become irrationally exuberant about a particular asset class or a particular market variable. The balance sheets of financial institutions then become overextended through the accumulation of exposure to this asset class or market variable. Often the process is self-reinforcing. When many financial institutions choose to take a particular position, prices increase, making the position look profitable. This creates extra desire on the part of financial institutions to take the position and yet more profits. Risk managers working for the financial institution should (and probably will) complain about the risks being taken, but in many instances senior management are likely to ignore their concerns because high profits are being made. To quote Chuck Prince, ex-CEO of Citigroup, on July 10, 2007: "When the music stops, in terms of liquidity, things will be complicated. But as long as the music is playing, you've got to get up and dance. We're still dancing."

At some stage the bubble must burst. Many traders then try to get out of their positions at the same time, causing illiquid markets and huge losses. Volatility increases and the risk management procedures used within the financial institution (e.g., the calculation of market VaR from historical data) can cause many financial institutions to try to unwind a wide range of risky positions at the same time. This can lead to further losses and more serious illiquidity problems. There may be failures (or rumors of failures) by some banks. Most banks are likely to experience liquidity funding problems and as a result lending may be curtailed.

The classic example of what has been described above is the subprime crisis that started in 2007. Other examples are the 1987 stock market crash, the 1994 bond market crash, the 1997–1998 Asian monetary crisis, and the 1998 Long-Term Capital Management failure. Irrational exuberance is part of human nature and to some extent is inevitable. As discussed in Chapter 6, it is exacerbated by the way traders are paid. A large part of the compensation comes from a bonus at year-end, which depends on performance during the year. A trader may be savvy enough to know that a market is irrationally exuberant and that there will be a correction. However, if there is a good chance that the correction will be delayed until next year, the trader is motivated to continue building up his or her position to maximize short-term compensation.

24.3.4 The Impact of Regulation

In many ways it is a laudable goal on the part of regulators to seek to ensure that banks throughout the world are regulated in the same way. As explained in Chapter 15, capital requirements and the extent to which they were enforced varied from country to country prior to Basel I. Banks were competing globally and as a result a bank subject to low capital requirements, or capital requirements that were not strictly enforced, found it easier to take risks and was therefore able to be more competitive in the pricing of some products.

However, a uniform regulatory environment comes with costs. All banks tend to respond in the same way to external events. Consider for example market risk. When volatilities and correlations increase, market risk VaR and the capital required for market risks increase. Banks then take steps to reduce their exposures. Since banks often have similar positions to each other, they try to do similar trades. A liquidity black hole can develop.

There is a similar issue as far as credit risk is concerned. During the low point of the economic cycle, default probabilities are relatively high and capital requirements for loans under the Basel II internal-ratings-based models tend to be high. As a result, banks may be less willing to make loans, creating problems for small and medium-sized businesses. During the high point of the business cycle, default probabilities are relatively low and banks may be too willing to grant credit. (This is similar to the phenomenon described in Figures 24.2 and 24.3.) The Basel Committee has recognized this problem and has dealt with it by asserting that the probability of default should be an average of the probability of default through the economic or credit cycle, rather than an estimate applicable to one particular point in time.

Should other financial institutions such as life insurance companies and pension funds be regulated in the same way as banks? It is tempting to answer "yes" so that one financial institution is not given an advantage over others. But the answer should be "no." These financial institutions have longer time horizons than banks. They should not be penalized for investing in illiquid assets. Also, they should not be required to adjust their portfolios when volatilities and correlations increase. These parameters tend to be mean reverting and so they eventually decrease again.

24.3.5 The Importance of Diversity

Models in economics usually assume that market participants act independently of each other. We have argued that this is often not the case. It is this lack of independence that causes liquidity black holes. Traders working for financial institutions tend to want to do the same trades at the same time. To solve the problem of liquidity black holes, we need more diversity in financial markets. One way of creating diversity is to recognize that different types of financial institutions have different types of risks and should be regulated differently.

Hedge funds have become important market participants. They are much less regulated than banks or insurance companies and can follow any trading strategy they like. To some extent they do add diversity (and therefore liquidity) to the market. But, as mentioned earlier, hedge funds tend to be highly leveraged. When liquidity tightens as it did in the second half of 2007, all hedge funds have to unwind positions accentuating the liquidity problems.

One conclusion from the arguments we have put forward is that a contrarian investment strategy has some merit. If markets overreact for the reasons we have mentioned, an investor can do quite well by buying when everyone else is selling and there is very little liquidity. However, it can be quite difficult for a financial institution to follow such a strategy if it is subject to short-term VaR-based risk management.

Summary

There are two types of liquidity risk: liquidity trading risk and liquidity funding risk. Liquidity trading risk is concerned with the ease with which positions in the trading book can be unwound. The liquidity trading risk of an asset depends on the nature of the asset, how much of the asset is to be traded, how quickly it is to be traded, and the economic environment. The credit crisis of 2007 emphasizes the importance of transparency. Assets that are not well defined or well understood are unlikely to trade in a liquid market for long. The liquidity of an asset at a particular time can be measured as the dollar bid–offer spread or as the proportional bid–offer spread. The latter is the difference between the bid and offer price divided by the average of the bid and offer price. The cost of unwinding a position in the asset is half of the bid–offer spread. Financial institutions should monitor the cost of unwinding the whole trading book in both normal market conditions and stressed market conditions.

A trader, when faced with the problem of unwinding a large position in an asset, has a trade-off between the bid–offer spread and market risk. Unwinding quickly leads to high bid–offer spreads, but low market risk. Unwinding slowly leads to lower bid–offer spreads, but more market risk. The optimal trading strategy depends on (a) the dollar bid–offer spread as a function of the quantity traded in a day and (b) the probability distribution for daily changes in the asset price. For any particular unwind strategy, the trader can choose a confidence level and calculate the unwind cost that will not be exceeded with the confidence level. The unwind strategy that minimizes this cost can then be determined.

Liquidity funding risk management is concerned with being able to meet cash needs as they arise. It is important for a financial institution to forecast its cash needs in both normal market conditions and stressed market conditions to ensure that they can be met with almost total certainty. Cash needs depend on depositor withdrawals, drawdowns on lines of credit, guarantees that have been made, defaults by counterparties, and so on. Sources of cash are instruments that can be readily converted into cash, borrowings in the wholesale market, asset securitizations, new depositors, cash itself, and (as a last resort) borrowings from a central bank. In June 2008, bank regulators issued a list of 17 principles describing how banks should manage their liquidity and indicated that they would be monitoring the liquidity management procedures of banks more carefully in the future.

The most serious liquidity risks arise from what are sometimes termed liquidity black holes. These occur when all traders want to be on the same side of the market at the same time. This may be because they have similar positions and manage risks in similar ways. It may also be because they become irrationally exuberant, overexposing themselves to particular risks. What is needed is more diversity in the trading strategies followed by market participants. Traders who have long-term objectives should avoid allowing themselves to be influenced by the short-term overreaction of markets.

Further Reading

Almgren, R., and N. Chriss. "Optimal Execution of Portfolio Transactions." *Journal of Risk* 3 (Winter 2001): 5–39.

Bangia, A., F. Diebold, T. Schuermann, and J. Stroughair. "Liquidity on the Outside." *Risk* 12 (June 1999): 68–73.

Bank for International Settlements. "Liquidity Risk Management and Supervisory Challenges," February 2008.

Bank for International Settlements. "Principles for Sound Liquidity Risk Management and Supervision," September 2008.

Brunnermeier, M. K., and L. H. Pedersen. "Market Liquidity and Funding Liquidity." *Review of Financial Studies* 22, no. 6 (2009): 2201–2238.

Brunnermeier, M. K., and L. H. Pedersen. "Predatory Trading." *Journal of Finance* 60, no. 4 (2005): 1825–1863.

Clunie, J. *Predatory Trading and Crowded Exits: New Thinking on Market Volatility.* Petersfield, UK: Harriman House, 2010.

Persaud, A. D., ed. *Liquidity Black Holes: Understanding, Quantifying and Managing Financial Liquidity Risk.* London: Risk Books, 1999.

Practice Questions and Problems (Answers at End of Book)

24.1 What was the transparency problem in the subprime crisis of 2007?

24.2 An asset is quoted bid 50, offer 55. What does this mean? What is the proportional bid–offer spread?

24.3 Suppose that an investor has shorted shares worth \$5,000 of Company A and bought shares worth \$3,000 of Company B. The proportional bid–offer spread for Company A is 0.01 and the proportional bid–offer spread for Company B is 0.02. What does it cost the investor to unwind the portfolio?

24.4 Suppose that in Problem 24.3 the bid–offer spreads for the two companies are normally distributed. For Company A the bid–offer spread has a mean of 0.01 and a standard deviation of 0.01. For Company B the bid–offer spread has a mean of 0.02 and a standard deviation of 0.03. What is the cost of unwinding that the investor is 95% confident will not be exceeded?

24.5 A trader wishes to unwind a position of 60 million units in an asset over 10 days. The dollar bid–offer spread as a function of daily trading volume, q, is $a + be^{cq}$ where $a = 0.2$, $b = 0.1$, and $c = 0.08$ and q is measured in millions. The standard deviation of the price change per day is \$0.1. What is the optimal strategy that minimizes the 95% confidence level for the costs?

24.6 Explain the difference between the liquidity coverage ratio (LCR) and the net stable funding ratio (NSFR).

24.7 Why is it risky to rely on wholesale deposits for funding?

24.8 What was the nature of the funding risk problems of Ashanti Goldfields and Metallgesellschaft?

24.9 What is meant by (a) positive feedback trading and (b) negative feedback trading? Which is liable to lead to liquidity problems?

24.10 What is meant by liquidity-adjusted VaR?

24.11 Explain how liquidity black holes occur. How can regulation lead to liquidity black holes?

24.12 Why is it beneficial to the liquidity of markets for traders to follow diverse trading strategies?

Further Questions

24.13 Discuss whether hedge funds are good or bad for the liquidity of markets.

24.14 Suppose that a trader has bought some illiquid shares. In particular, the trader has 100 shares of A, which is bid \$50 and offer \$60, and 200 shares of B, which is bid \$25 and offer \$35. What are the proportional bid–offer spreads? What is the impact of the high bid–offer spreads on the amount it would cost the trader to unwind the portfolio? If the bid–offer spreads are normally distributed with mean \$10 and standard deviation \$3, what is the 99% worst-case cost of unwinding in the future as a percentage of the value of the portfolio?

24.15 A trader wishes to unwind a position of 200,000 units in an asset over eight days. The dollar bid–offer spread, as a function of daily trading volume q, is $a + be^{cq}$ where $a = 0.2$, $b = 0.15$, and $c = 0.1$ and q is measured in thousands. The standard deviation of the price change per day is \$1.50. What is the optimal trading strategy for minimizing the 99% confidence level for the costs? What is the average time the trader waits before selling? How does this average time change as the confidence level changes?

Chapter 25

Model Risk Management

The use of models by financial institutions is growing fast. Models are employed for making credit decisions, liquidity management, the evaluation of credit exposures, the valuation of derivatives, the calculation of risk measures such as value at risk (VaR) and expected shortfall (ES), the management of portfolios of financial assets for clients, the assessment of capital adequacy, customer relationship management, fraud detection, the identification of money laundering, and so on. Complex models, with advanced analytic techniques such as machine learning, are being used to automate some activities previously carried out by humans. (See Chapter 28 for a discussion of machine learning.) It is now recognized that large financial institutions need a model risk management function to ensure that models are appropriate for their intended purpose and are used in the correct way.

Models are approximations to reality. The art of building a model is to capture the key aspects of reality for a particular application without allowing the model to become so complicated that it is almost impossible to use. Nearly always, a model relies on some assumptions about the phenomenon being modeled. It is important to understand those assumptions and know when they are no longer appropriate.

It used to be the case that models could be developed by a quant on a spreadsheet and used with minimal oversight. This is no longer the case. Models must now be well documented and reviewed periodically to determine whether they are still appropriate and are working as intended. When changes are made to models, approval is usually necessary. These aspects of model risk management are typically not popular with model builders, who tend to enjoy the creative aspects of model development and dislike spending time justifying their models and documenting them.

Earlier in this book, we have made the point that regulators have, since the 2007–2008 crisis, become less enamoured with the use of internal models to determine capital requirements. They have also become more concerned about the way models are used for other purposes within financial institutions. Model risk is a component of operational risk. Regulators want banks to have a system for managing the risks presented by models. Models must be developed, validated, and used in accordance with this system.

This chapter starts by reviewing regulatory requirements on model risk management. It describes the role of the model validation group and then moves on to compare the way in which models are used in finance with the way they are used by physicists and other scientists. After that, it focuses on the use of models for valuation and risk management and examines some of the lessons that can be learned from the mistakes made in that area.

25.1 Regulatory Requirements

In April 2011, the U.S. Board of Governors of the Federal Reserve System published SR 11-7, which provides guidance to banks on effective model risk management.[1] This document has proved to be very influential, and U.S. banks are expected to ensure that their practices are consistent with the principles outlined in the document. Bank supervisors elsewhere have indicated that they expect to see a similar systematic approach to understanding and managing model risk. If a bank supervisor feels that a bank's model risk management systems are inadequate, it may request the bank to hold additional capital.

SR 11-7 defines model risk as "the potential for adverse consequences from decisions based on incorrect or misused model outputs and reports." Model risk occurs primarily for two reasons:

1. The model may have fundamental errors when viewed from the perspective of its objectives. These errors may concern the data it uses, the way calculations are done, numerical procedures, assumptions made, and so on.
2. The model may be used incorrectly or inappropriately. It is important to understand the limitations of models. A model may have been developed assuming a particular set of market conditions or a particular type of customer behavior. When these change, the model may no longer be appropriate.

Financial institutions should attempt to identify the sources of model risk and their magnitude. Complex models generally lead to more model risk than simple models. When models are linked or when several models are based on the same underlying set of assumptions, there is liable to be a greater degree of model risk. It is important to monitor model performance to determine as early as possible when a model is, for whatever reason, not performing as well as expected.

[1] See Board of Governors of the Federal Reserve System, Office of the Comptroller of the Currency, "Supervisory Guidance on Model Risk Management," SR 11-7, April 2011.

25.1.1 Model Development

When a model is developed, documentation is important. There are several reasons for this. If the model developer leaves the organization and there has been no good documentation, it might be very difficult for others to assume responsibility for the model, and, as a result, the work of the developer might be lost. Regulators frequently ask to see model documentation when they evaluate the procedures used by a financial institution. Also, model validation and periodic review are facilitated if the model has been well documented.

An important aspect of model documentation is a statement of the purpose of the model, the underlying theory, an evaluation of industry practice, and a review of published research. It is important that the documentation clearly explains the underlying mathematics, numerical procedures used, assumptions made, and so on. The documentation should be sufficiently detailed that someone who is new to the area can understand what has been done. As indicated earlier, model documentation will often be viewed as a boring chore by the model developer. But in practice it often proves to be a useful discipline, forcing the model developer to step back and think about his or her work in a way not done before. Sometimes changes to the model result.

Models should be tested as thoroughly as possible. This should include careful testing of all parts of the computer code and evaluating the model's behavior for a range of input data. The testing should identify situations where the model performs well and situations where it performs poorly. This can be done by observing the performance of the model with progressively more extreme values for the inputs. In some situations, it is appropriate to compare the output from the proposed model with the output from a more elaborate model that would be computationally too slow to be used in practice.

The way in which the performance of a model is tested will depend on the nature of the model. Models developed from historical data should be tested out of sample. For example, if a model for making loan decisions is developed from a particular set of data collected over a particular period, it would clearly not be appropriate to test it using those data. Similarly, if a portfolio management strategy for a client is developed by analyzing the behavior of asset prices during a particular period, it should be tested using data from another period.

We have discussed in earlier chapters how back-testing can be used for VaR models. It can be used in other situations as well. Suppose a model based on a particular set of assumptions is used to price barrier options. It should be possible to determine how well it would have worked in the past. Would a user of the model have made a profit on average? If so, would the profit have been sufficiently great to compensate for the risks being taken?

As a model is used in practice, a financial institution should collect information on whether it is reflecting economic and business realities and working in the way intended. There is a danger that the existence of a model may lead to a change in behavior. For example, a model for managing customer relationships may lead to customers behaving differently; a model for detecting money laundering might work well until would-be money launderers learn (perhaps by trial and error) how the model works and adjust their behavior accordingly. Even model users may find ways of adjusting inputs to a model or the way the model is used to make themselves look good.

25.1.2 Model Validation

One result of the increasing attention being devoted to model risk management is the formation of model validation groups. These groups are responsible for verifying that all the models used by the organization are performing as expected. They provide a check on the work of the model developer and critically review what has been produced and how the model is performing.

The person charged with validating a model should be independent of the model developer and the model user. Ensuring independence is critical to the success of the model risk management function. The reporting lines and incentives should be chosen to promote independence. The success of the model validation function should be judged by the objectivity of the model reviews, the issues raised, and the actions taken by management to address the issues. Corporate culture is important. If objective thinking and the challenging of decisions are not encouraged, the model validation process is likely to be ineffective. Regulators are likely to ask to see instances where models and their use have been appropriately changed as a result of validation.

Models are sometimes used informally in a small way prior to initial validation. However, before a model becomes part of a financial institution's way of operating it is important that a first validation be carried out. The extent and rigor of the validation should depend on the risks presented by the model. Consider the situation where a model has been developed to automate billions of dollars of loan decisions. Here it is clearly important that the validation be thorough. Even if the model is performing well and in accordance with its objectives, the model validation group might insist on a staged implementation with close monitoring to reduce model risk.

Model validation should continue on an ongoing basis after the model has been approved for use. This ensures that the impact of changes in markets and business practices on the performance of the model is monitored. SR 11-7 recommends that models be reviewed by the validation group at least annually. Users of the model (particularly if they are different from the developers) can provide input, and the validation group can determine whether the performance of the model using recent data is as good as it was on the data available when the model was first produced. The outcomes from using the model (e.g., credit decisions made, return on investments from using the model, fraud attempts thwarted) should be documented and compared with what was expected.

Business Snapshot 25.1 describes the London Whale story, which is an example of a situation where the model validation group was under pressure to approve a model. This was because the model provided justification for reducing VaR, and therefore regulatory capital, when a huge position in one product was hedged with an opposite position in a similar product. If the model validation group had more time, it would have carried out a more careful review of the model. It might have argued that the correlation between

BUSINESS SNAPSHOT 25.1

The London Whale

The London Whale was a trader, Bruno Iksil, who worked for the Chief Investment Office (CIO) at JPMorgan Chase (JPM). His nickname arose from his huge positions in credit default swaps on CDX and iTraxx indices. (As explained in Section 19.4, these are products that provide protection against defaults for a portfolio of companies.)

The CIO's responsibility was to invest the bank's surplus cash. At the end of 2011, the bank's positions in the credit indices were moderately large and it was net long (i.e., it was buying more protection against defaults than it was selling). This net long position hedged the CIO's positions in fixed-income products. JPM wanted to reduce its risk-weighted assets to comply with the new Basel regulations. It also wanted to move toward a more neutral position in the credit indices because it felt that the economy was improving. Unwinding its existing positions in the indices would lead to a loss, partly because the positions had a negative mark-to-market value and partly because the positions were sufficiently large that unwinding them would tend to move the market against JPM. As a result, the CIO chose to hedge its net long position in credit indices by selling protection on different indices.

The CIO expected the long and short positions to offset each other. In fact, this did not happen. Market participants realized that someone was taking big positions. (JPM, like other dealers, tries to keep its positions secret. However, market participants realized that there was a "whale" in the market and it eventually became apparent that the whale was a trader for JPM.) Realizing that the positions would have to be unwound at some future time and that this would move the market, some market participants engaged in predatory trading, taking opposite positions to JPM. This moved the market against JPM and meant that the values of JPM's long and short positions were not as closely correlated as expected.

The CIO increased the size of its positions. This was partly for technical hedging reasons and partly to defend against predators. Its huge position was theoretically well balanced (i.e., models showed that its value should not be affected to any great extent by credit spread movements). It had long and short positions in CDX investment-grade (IG) indices and in the CDX high-yield (HY, non-investment-grade) indices. It sold large amounts of credit protection on the 10-year CDX IG 9 index and bought large amounts of protection on the five-year CDX IG 9 index. (This was an index created from 125 names in September 2007. Although not the current CDX IG index, it still traded reasonably actively; 121 of the original 125 names had not defaulted.) Prices did not move in the way that was expected, and a loss of about $6 billion was recognized by April/May 2012. Deferred bonuses to CIO employees were clawed back, and Jamie Dimon, JPM's chairman and CEO, took a pay cut for the year. Allegations that the value of the portfolio was misreported to shareholders were investigated by U.S. government agencies, and JPM ended up paying a fine of almost $1 billion.

Why was the CIO allowed to take such huge risks? In January 2012, the CIO exceeded its VaR limit and the bank as whole exceeded its VaR limit as a result of the CIO's trading. However, a new VaR model was being developed. This model, which was finished in the second half of January and quickly reviewed and approved by the model validation group, reduced the VaR of the CIO's portfolio by 50% so that pre-January VaR limits were adhered to. JPM's January 2013 internal report on the London Whale losses (particularly the appendix) indicates that a number of mistakes were made. The quant who developed the model and the model validation group were under pressure to develop and approve the model. The quant reported to traders in the CIO, and the model validation group had time to carry out only limited back-testing of the model. In May 2012, the model was re-evaluated. A number of serious errors were discovered and its use was discontinued.

This story emphasizes the importance of the model validation function and why it needs to be independent.

the two products was assumed to be too large. It could also have made the point that hedges that work well for small positions do not always work well for large positions. The bank's use of the model for its large position had the potential to lead to predatory trading, rendering the historical comparisons of the prices of the products being traded less relevant.

SR 11-7 argues that comprehensive validation should have three core elements:

1. Evaluation of conceptual soundness, including developmental evidence
2. Ongoing monitoring, including process verification and benchmarking
3. Outcomes analysis, including back-testing

A sound developmental process should produce documented evidence in support of all model assumptions. There is a danger that the model developer has a bias toward a particular type of model—or even an interest in showing off his or her quant skills by developing an unnecessarily complex model. The validation work should ensure that this has not happened. Stability is an important property of models. If a small change in an input produces a large change in the output, the model may be unstable and inappropriate for its intended use. Model validation should include model stress testing where the model is subjected to extreme inputs to test its range of applicability.

As already mentioned, ongoing monitoring is a core part of model validation. This includes ensuring that all changes made to a model are tested and documented. If the model user and the model developer are the same person (as is sometimes the case), there is a temptation for the model user to tweak the code to improve model performance. This is unsatisfactory and can lead to a situation where the model is nothing more than a collection of tweaks with no unifying theory. If a model is not performing as well as it could, a review should be undertaken and possibly a new model with better foundations developed. It is important for a model validation function to investigate situations where a model is not used because managers feel the output is not consistent with their experience and judgment. This can be an indication that improvements or a new approach is necessary.

Outcomes analysis, such as back-testing, is important for monitoring a model's performance. Some work in this area should be carried out by the model developer and documented. This work should be checked by the model validation group and continued on an ongoing basis. The reasons for any differences between expected outcomes and actual outcomes should be analyzed. For example, a particular liquidity management model might be found to work well in low-interest-rate environments, but less so when interest rates increase. This type of analysis can lead to model improvements.

25.1.3 Vendor Models

Not all models are developed internally. Some are purchased from model vendors. Sometimes the purpose of a model that has been purchased is to benchmark a similar model developed internally. In this case, the vendor model can be a useful validation tool. In other situations, the vendor model may be used because it is less expensive than developing a similar model internally. When this is the case, vendor models should be subject to the same validation as internal models. However, they present particular challenges for the model validation group, as the modeling expertise is external to the organization and some aspects of the model may be proprietary.

Financial institutions should have processes in place for selecting vendor models. The vendor should be required to provide documentation explaining the model design, testing results, and model limitations. Financial institutions should insist that vendors conduct ongoing performance monitoring and outcomes analysis, and that they provide the results of this to their clients.

The validation group should be responsible for approving the use of a vendor model. It may not have access to the computer code, but it can conduct extensive analysis on the results produced by the model for a variety of inputs. It should insist that the vendor provide full information about the data used to develop the model. There should be a contingency plan for a situation in which the vendor goes out of business or decides to stop supporting the model. One approach here is to insist that the computer code and all other proprietary material be placed in escrow and made available to the client in circumstances in which the vendor is no longer able to provide support.

25.2 Models in Physics and Finance

Many individuals who were well trained in physics now work in finance developing models. Some models in finance are structurally similar to those in physics. For example, the differential equation that leads to the famous Black–Scholes–Merton option pricing model is the heat-exchange equation that has been used by physicists for many years. However, as Derman has pointed out, there is an important difference between the models of physics and the models of finance.[2] The models of physics describe physical processes and are highly accurate. By contrast, the models of finance describe the behavior of market variables, customers, borrowers, and so on. The phenomena being modeled are a result of the actions of human beings. As a result, the models are at best approximate descriptions of the phenomena being modeled.

One important difference between the models of physics and the models of finance concerns model parameters. The parameters of models in physics generally do not change. For example, the gravitational pull on the surface of the earth is always 32 feet per second per second. By contrast, parameters in finance models change daily. The volatility used to price an option might be 20% one day, 22% the next day, and 19% the following day. Indeed, as we will see later, the parameters used on a day are chosen to be consistent with market prices, not to determine them.

25.3 Simple Models: Expensive Mistakes

Models do not have to be highly complex to lead to losses. Business Snapshot 25.2 provides a simple example of this. Kidder Peabody's computer system did not account correctly for funding costs when a linear product was traded. As a result, the system indicated that one of the company's traders was making a large profit when in fact he was taking a huge loss. The model used by Kidder Peabody should have been subject to independent validation. Even if the theoretical problem with the model was not picked up, the difference between model prices and market prices would have indicated to the model validation group that something was seriously wrong.

The Kidder Peabody modeling mistake would have been easy to identify if there was a culture within the organization that allowed traders to be challenged. Indeed, Barry Finer, risk manager for the government bond desk, did point out the difficulty of making large arbitrage profits from a market as efficient as the U.S. government bond market. But his concerns were dismissed out of hand.

---

[2] See E. Derman, *My Life as a Quant: Reflections on Physics and Finance* (Hoboken, NJ: John Wiley & Sons, 2004), and E. Derman, "Model Risk," *Risk* 9, no. 2 (May 1996): 139–145.

BUSINESS SNAPSHOT 25.2

Kidder Peabody's Embarrassing Mistake

Investment banks have developed a way of creating a zero-coupon bond, called a *strip*, from a coupon-bearing Treasury bond by selling each of the cash flows underlying the coupon-bearing bond as a separate security. Joseph Jett, a trader working for Kidder Peabody, had a relatively simple trading strategy. He would buy strips and sell them in the forward market. The forward price of the strip was always greater than the spot price, so it appeared that he had found a money-making machine! In fact, the difference between the forward price and the spot price represents nothing more than the cost of funding the purchase of the strip. Suppose, for example, that the three-month interest rate is 4% per annum and the spot price of a strip is \$70. The three-month forward price of the strip is $70e^{0.04\times3/12}$ = \$70.70. This is because, for someone who wants to own the strip in three months, the forward contract saves \$0.70 of funding costs.

Kidder Peabody's computer system reported a profit on each of Jett's trades equal to the excess of the forward price over the spot price (\$0.70 in our example). By rolling his contracts forward, Jett was able to prevent the funding cost from accruing to him. The result was that the system reported a profit of \$100 million on Jett's trading (and Jett received a big bonus) when in fact there was a loss in the region of \$350 million. This shows that even large financial institutions can get relatively simple things wrong!

Sometimes the assumptions made by model developers are inaccurate but less easy to identify than in the Kidder Peabody case. An example here is provided by the interest rate swap market. A plain vanilla interest rate swap, such as the one described in Section 5.5.3, can be valued by assuming that forward interest rates will be realized. This is explained in Appendix D. For example, if the forward interest rate for the period between 2 and 2.5 years is 4.3% per annum, the swap can be correctly valued on the assumption that the floating rate that is exchanged for a fixed rate at the 2.5-year point is calculated using this rate. It is tempting to generalize from this and argue that any swap agreement to exchange cash flows in the future can be valued on the assumption that forward rates are realized. This is not so. Consider, for example, what is known as a *LIBOR-in-arrears swap*. In this instrument, the floating rate that is observed on a particular date is paid on that date, not one accrual period later as is the case for a plain vanilla swap. A LIBOR-in-arrears swap should be valued on the assumption that the realized interest rate equals the forward interest rate plus a "convexity adjustment." As indicated in Business Snapshot 25.3, financial institutions that did not understand this lost money in the mid-1990s.

BUSINESS SNAPSHOT 25.3

Exploiting the Weaknesses of a Competitor's Model

A LIBOR-in-arrears swap is an interest rate swap where the floating interest rate is paid on the day it is observed, not one accrual period later. Whereas a plain vanilla swap is correctly valued by assuming that future rates will be today's forward rates, a LIBOR-in-arrears swap should be valued on the assumption that the future rate is today's forward interest rate plus a "convexity adjustment."

In the mid-1990s, sophisticated financial institutions understood the correct approach for valuing a LIBOR-in-arrears swap. Less sophisticated financial institutions used the naive "assume forward rates will be realized" approach. The result was that by choosing trades judiciously, sophisticated financial institutions were able to make substantial profits at the expense of their less sophisticated counterparties.

The derivatives business is one where traders do not hesitate to exploit the weaknesses of their competitors' models!

A similar example to this is provided by the interest rate futures markets. In most situations, it is reasonable to assume that a futures price is the same as the corresponding forward price. For example, the futures price in a futures contract to deliver gold in one year should be very close to the forward price in a forward contract to deliver gold at the same time. There are some differences between the two contracts. For example, a futures contract is settled daily and the party with the short position has some flexibility on the precise delivery date. But these differences do not create a material difference between the futures and forward prices. In the mid-1990s, many market participants chose to assume that the futures rate in an interest rate futures contract is the same as the corresponding forward rate. In fact, the approximate result that is true for most other futures contracts does not apply to interest rate futures. This is because interest rate futures last a long time so that the impact of daily settlement cannot be ignored. Also, in the case of interest rates, futures and forwards are settled at different times. Futures are settled at the beginning of the period covered by the underlying rate whereas forwards are settled at the end.[3]

25.3.1 Monitoring Trading Patterns

The situations we have just outlined suggest that an important way of identifying model risk in the valuation of products is to monitor trading patterns. In particular,

[3] See J. Hull, *Options, Futures, and Other Derivatives*, 10th edition (Upper Saddle River, NJ: Pearson, 2017), for a more detailed discussion.

the risk management function within a financial institution should keep track of the following:

1. The type of trading the financial institution is doing with other financial institutions
2. How competitive it is in bidding for different types of structured transactions
3. The profits being recorded from the trading of different products

Getting too much of a certain type of business, or making huge profits from relatively simple trading strategies, can be a warning sign. If a financial institution finds that its prices are out of line with the market, it must make adjustments to its valuation procedures to bring them into line.

The high profits being recorded for Joseph Jett's trading at Kidder Peabody (see Business Snapshot 25.2) should have indicated that something was amiss. Likewise, if in the mid-1990s a financial institution's risk management team discovered that traders were entering into a large number of LIBOR-in-arrears swaps with other financial institutions (see Business Snapshot 25.3) where they were receiving fixed and paying floating, they could have alerted modelers to a potential problem and directed that trading in the product be temporarily stopped.

25.4 Models for Pricing Actively Traded Products

When a financial instrument trades actively, we do not need a model to know what its price is. The market tells us this. Suppose, for example, that a certain option on a stock index trades actively and is quoted by market makers as bid \$30, offer \$30.5. Our best estimate of its current value is the mid-market price of \$30.25.

A situation that is common in the over-the-counter market is one where a financial instrument that has to be valued is a standard product such as an option, but not exactly the same as one that trades in the market. For example, it might be an option with a strike price or time to maturity (or both) different from the options whose prices can be observed. The model is then used as a tool to ensure that the way an instrument is priced is consistent with the observed market prices of other similar instruments. A good example of how this is done is provided by the way the Black–Scholes–Merton model is used in practice. (See Appendix E for a description of the Black–Scholes–Merton model.)

The Black–Scholes–Merton model has been successful because all its inputs except one are known with a high degree of certainty. The one input that cannot be observed is volatility. This means that there is a one-to-one correspondence between prices and volatilities. When the volatility for an option has been specified, the price can be calculated. Similarly, when the price is known, a volatility can be determined. The volatility determined from a market price is known as an *implied volatility*. Traders responsible for options on a particular asset keep track of what is known as the *volatility surface*. This is the set of implied volatilities calculated for options with different strike prices and times to maturity.

Table 25.1 Volatility Surface
Volatilities for different strike prices and maturities are shown as % per annum.

	Strike Price				
	0.90	**0.95**	**1.00**	**1.05**	**1.10**
1-month	14.2	13.0	12.0	13.1	14.5
3-month	14.0	13.0	12.0	13.1	14.2
6-month	14.1	13.3	12.5	13.4	14.3
1-year	14.7	14.0	13.5	14.0	14.8
2-year	15.0	14.4	14.0	14.5	15.1
5-year	14.8	14.6	14.4	14.7	15.0

Table 25.1 provides an example of the volatility surface that could be constructed for options on an exchange rate. (For simplicity we assume that the current exchange rate is close to 1.00.) Note that, if the Black–Scholes–Merton model were a perfect description of reality, the volatilities in the table would be constant and not vary through time. In practice, implied volatilities for options are not the same for different strike prices and times to maturity, and they do vary through time. (This is true for options on all assets.) It is the trader's job to understand the current volatility surface and how it could change through time.

Suppose that a trader wishes to value a nine-month option with a strike price of 1.05 on the day Table 25.1 is produced. A six-month option with this strike price has a volatility of 13.4% and a one-year option with this strike price has a volatility of 14.0%. Interpolating between these, a trader would use a volatility of 13.7% in the Black–Scholes–Merton model to value the option. Suppose next that both the time to maturity and the strike price are different from those in the table. A two-dimensional interpolation is then necessary. This involves first interpolating between strike prices and then between times to maturity. (The same answer is obtained if one first interpolates between maturities and then between strike prices. See Problem 25.12.) Consider a 1.5-year option with a strike price of 0.92. A one-year option with this strike price has an interpolated implied volatility of 14.42%; a two-year option with the strike price has an interpolated volatility of 14.76%. The interpolated volatility for a 1.5-year option is therefore 14.59%.

25.4.1 Sources of Model Risk for Actively Traded Products

One issue for the model validation group is whether the prices determined for actively traded instruments depend on the model used. If, in the example just considered, we replaced the Black–Scholes–Merton model by another model, would the prices be similar? The answer to this question is that, when the nature of the product being valued is exactly the same as actively traded instruments (i.e., only parameters like the strike price and time to maturity are different), the model does not usually affect prices materially.[4]

[4]This is demonstrated by S. Figlewski, "Assessing the Incremental Value of Option Pricing Theory Relative to an Informationally Passive Benchmark," *Journal of Derivatives* (Fall 2002): 80–96.

This is because the model is being used as nothing more than a sophisticated interpolation tool and two different interpolation tools usually give very similar valuations.

There is liable to be model risk in extreme situations. The model validation group should check this out. For example, it should examine the operation of different models when the volatility surface is very steep in either the strike price or time-to-maturity dimension. Also, the model validation group should check on the reliability of the pricing data that are available in the market and how the model would work when less data than usual are available.

25.4.2 Hedging

In the case of actively traded instruments, the main model risk concerns hedging. As we have discussed, switching from one model to another does not usually affect prices to any great extent, but it can have an effect on hedging performance.

We can distinguish between *within-model hedging* and *outside-model hedging*. Within-model hedging is designed to deal with the risk of changes in variables that are assumed to be uncertain by the model. Outside-model hedging deals with the risk of changes in variables that are assumed to be constant (or deterministic) by the model. For the Black–Scholes–Merton model, hedging against movements in the underlying stock price (delta and gamma hedging) is within-model hedging, because the model assumes that stock price changes are uncertain. However, hedging against volatility changes (vega hedging) is outside-model hedging because the model assumes that volatility is constant.

In practice, traders almost invariably do outside-model hedging as well as within-model hedging. This is because, as we have explained, the calibration process results in parameters such as volatilities (which are assumed by a model to be constant) changing daily. A good options trader will monitor the exposure of a trading book to the sorts of shifts in the volatility surface that are typically seen.[5]

If a model is designed to be always consistent with market prices and hedging is carried out continuously for all the variables that can change (both those that are assumed to be constant by the model and those that are assumed to be stochastic), the value of the hedger's position should in theory not change. However, in practice hedging cannot be carried out in this idealized way. Consider again the Black–Scholes–Merton model to value a position in options on an exchange rate. Each option price is a deterministic function of the exchange rate, its implied volatility, and time.[6] This means that the trader has an exposure to the exchange rate and the implied volatilities of all the options in the portfolio. The exposure to the exchange rate can be hedged without difficulty by trading the foreign currency. This is known as delta hedging and is usually done at least once a day (see Chapter 8). Hedging the risk of volatility changes is not as easy. It requires trades

[5] A principal components analysis, as described in Chapter 9, can be used. If exposure to each of the major principal components is small, the exposure to the shifts that have been observed in practice should also be small.

[6] To keep this example simple, it is assumed that interest rates and dividend yields do not change.

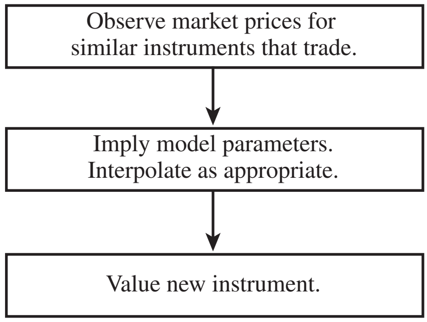


Figure 25.1 How Models Are Usually Used for Valuing a Financial Instrument

in other options (which can be expensive) and assumptions about how the volatility surface can change through time.[7] As a result, volatility hedging cannot be anything near as effective as delta hedging.

In practice, delta hedging is often the only hedging that a trader carries out on a regular basis. The delta of an option position does depend on the model used. This means that the hedge position taken by a trader varies from one model to the next. An important activity for the model validation group is to test the effectiveness of delta hedging for alternative models.

25.4.3 P&L Decomposition

One analysis that can be carried out is known as *P&L decomposition.* Risk managers divide day-to-day changes in the value of a portfolio into:

1. A change resulting from risks that were unhedged
2. A change resulting from the hedging model being imperfect
3. A change resulting from new trades done during the day

If the day-to-day change is unacceptable, the analysis will indicate the areas where more effort should be expended. The good news here is that, on average, the total gain or loss to a financial institution from inaccurate hedging is often small because it is largely diversified away across the portfolio of a large financial institution.

25.5 Models for Less Actively Traded Products

The general procedure for pricing derivatives is to build a model and then imply the parameters of the model from the prices that can be observed in the market. This is illustrated in Figure 25.1. One or more other financial instruments whose prices can be

[7] If perfect hedging against volatility surface movements were possible, differences in delta hedging would be offset by differences in volatility hedging. But volatility hedging is in practice far from perfect.

observed in the market are chosen as the "calibrating instruments." These instruments are chosen so that they are as similar as possible to the financial instrument of interest. They are then used to imply parameters for the model being used for the instrument under consideration.

We have argued that there is very little model risk as far as valuation is concerned when this approach is used for standard instruments that trade actively. The same is not true for non-standard instruments. Here the model used is likely to have a significant effect on valuations as well as on hedging. An example of model risk in the pricing of structured products is provided by Business Snapshot 6.1. This makes the point that the assumption that the BBB-rated tranche of an asset-backed security (ABS) is similar to a BBB bond is not a good one when tranches of an ABS CDO are being priced.

An important point to note is that we are nearly always concerned with what might be termed *relative valuation*. The objective is to value the financial instrument consistently with other similar financial instruments whose prices can be observed. It is very rare to use a model to value a derivative in absolute terms. This is another difference between finance models and the models of physics.

A financial institution should not rely on a single model for pricing non-standard products. Instead it should, whenever possible, use several different models. This leads to a price range for the instrument and a better understanding of the model risks being taken.

Suppose that three different models give prices of $6 million, $7.5 million, and $8.5 million for a particular product that a financial institution is planning to sell to a client. Even if the financial institution believes that the first model is the best one and plans to use that model as its standard model for daily repricing and hedging, it should ensure that the price it charges the client is at least $8.5 million. Moreover, it should be conservative about recognizing profits. If the product is sold for $9 million, it is tempting to recognize an immediate profit of $3 million ($9 million less the believed-to-be-accurate price of $6 million). However, this is overly aggressive. A better, more conservative, practice is to put the $3 million into a reserve account and transfer it to profits slowly during the life of the product.[8]

25.5.1 Weighted Monte Carlo Simulation

When different models are being used, they should be calibrated to the prices of actively traded standard products that are observed in the market. For example, when a barrier option is being priced, the models tested should produce prices for regular call and put options that are equal to those observed in the market. Rather than testing a small number of different models, a sophisticated analysis can be attempted to answer the question: "What range of model prices is possible for models that price actively traded products correctly?" This analysis, if carried out successfully, will give a worst possible price and a best possible price for the non-standard instrument. One approach is known as *weighted*

[8] This is also likely to have sensible implications for the way bonuses are paid.

Monte Carlo simulation.[9] This involves carrying out a Monte Carlo simulation for one model and then applying weights to the sampled paths. Constraints are placed on the weights so that standard instruments are priced correctly by the weighted Monte Carlo simulation. An optimization procedure is used to find the weights that produce maximum and minimum values for the non-standard derivative consistent with the constraints.

25.6 Accounting

Products that are held in the trading book must be valued daily. This process is referred to by traders as *marking to market*. The valuations are used by accountants to calculate quarterly financial statements. Accountants refer to marking to market as "fair-value accounting."

Accountants recognize that some models produce more reliable prices than others. The Financial Accounting Standards Board in the United States (FAS 157) and the International Accounting Standards Board in Europe (IAS 39) require banks to report three categories of valuations separately:

Level 1 valuations: Valuations for instruments where there are quoted prices in active markets.
Level 2 valuations: Valuations for instruments where there are quoted prices for similar instruments in active markets or quoted prices for the same instrument in markets that are not active.
Level 3 valuations: Valuations for instruments where some modeling assumptions are required. (This is sometimes referred to as *marking to model*.)

The valuations we discussed in Section 25.4 where the model is used as a sophisticated interpolation tool would be classified as Level 2 valuations. The valuations in Section 25.5 where the calibrating instruments are somewhat different from the instrument being valued would be classified as Level 3 valuations.

25.7 What Makes a Successful Pricing Model?

It might be argued that any model where the values of input parameters change from day to day is not helpful. Certainly this would be true for models in physics. But, as has been explained, models for valuing financial instruments are used in a quite different way from the models of physics.

The Black–Scholes–Merton model is popular for three reasons:

1. It can be used in the way we have described for interpolation in conjunction with a volatility surface, and ensures that prices are consistent with observed market prices.

[9] See, for example, A. Elices and E. Gimenez, "Weighted Monte Carlo," *Risk* (May 2006).

2. It is a communication tool. Traders often prefer quoting implied volatilities to quoting option prices, because implied volatilities are more stable than option prices and, as a result, their quotes do not change as frequently. When the price of the underlying asset changes or the interest rate changes, the option price changes but the implied volatility may not.
3. The model is sufficiently simple that a trader can develop intuition about the model and use it to structure his or her thinking about option markets.

Models in other areas that have these properties have also proved to be popular. Model developers and the model validation group should understand that the ultimate purpose of a model is often to enable the model user to understand the market and exercise his or her judgment more efficiently. When alternative models are compared, this is an important consideration and may lead to simpler models being preferred to more complex ones.

25.8 Model Building Missteps

The art of model building is to capture what is important for valuing and hedging an instrument without making the model more complex than it needs to be. Sometimes models have to be quite complex to capture the important features of a situation, but this is not always the case.

One potential misstep in model building is *over-fitting*. Consider the problem posed by the volatility surface in Table 25.1. We can exactly match the volatility surface with a single model by extending the Black–Scholes–Merton model so that volatility is a complex function of the underlying asset price and time.[10] This can be desirable for some applications, but the model can become difficult to use. We may find that some properties of the model are less reasonable than those of simpler models. In particular, the joint probability distribution of the asset price at two or more times might be less than ideal, and, in some circumstances, future volatility surfaces given by the model might look quite different from those observed in the market today.[11]

---

[10] This is the implied volatility function model proposed by B. Dupire, "Pricing with a Smile," *Risk* 7 (February 1994): 18–20; E. Derman and I. Kani, "Riding on a Smile," *Risk* (February 1994): 32–39; M. Rubinstein, "Implied Binomial Trees," *Journal of Finance* 49, no. 3 (July 1994): 771–818.

[11] Models that fit the volatility surface at all future times accurately price instruments where there is a payoff at one future time that is dependent only on the asset price at that time. However, they are liable to be less accurate for instruments such as barrier options and compound options that depend on the joint probability distribution of the asset price at two or more times. Hull and Suo find that the implied volatility function model works reasonably well for compound options, but is less accurate for barrier options. See J. Hull and W. Suo, "A Methodology for the Assessment of Model Risk and Its Application to the Implied Volatility Function Model," *Journal of Financial and Quantitative Analysis* 37, no. 2 (June 2002): 297–318.

Another potential misstep in model building is *over-parameterization*. The Black–Scholes–Merton model can be extended to include features such as stochastic volatility or jumps in the asset price. This invariably introduces extra parameters that have to be estimated and makes the model more difficult to use. It is usually claimed that the parameters in complex models are more stable than those in simpler models and do not have to be adjusted very much from day to day. This may be true, but we should remember that we are not dealing with physical processes. The parameters in a complex model may remain relatively constant for a period of time and then change, perhaps because there has been what economists refer to as a *regime change*. A financial institution may find that a more complicated model is an improvement over a simple model until there is a regime change. The more complicated model may not then have the flexibility to cope with changing market conditions.[12]

As we have already mentioned, simple models are often the most successful. Often users are skeptical of complex models because they are so-called black boxes and it is very difficult to develop intuition about them. In some situations, users' skepticism is well founded for the reasons we have just mentioned.

Summary

Model risk is a form of operational risk. It is now recognized that it requires special attention. Regulators expect to see that a financial institution has comprehensive procedures for developing, documenting, validating, and using models. Models should be reviewed periodically to ensure that they are working as originally intended. The model validation group should be independent of model developers and model users. It should provide a check on the work of the model developer and reach conclusions about the circumstances under which the model should be used.

Models in finance are different from those used in the physical sciences, because they are ultimately models of human behavior. They can usually only provide an approximation to reality. Furthermore, from time to time there are regime shifts leading to fundamental shifts in the performance of models.

Valuation models are often nothing more than sophisticated interpolation tools to ensure that one instrument is priced consistently with other similar instruments. As the instrument being valued becomes less standard, the valuation model chosen becomes more important. It is then a good practice for the model validation group to test the use of different models to obtain a realistic range for pricing and to understand the accompanying model risk. The models should always be calibrated to the current prices of actively traded instruments. The hedging performance of a model, as well as valuation performance, should be evaluated carefully.

---

[12]The nature of this type of problem in the social sciences is discussed in the famous Lucas Critique, which was propounded by the economist Robert Lucas in 1976. See R. Lucas, "Economic Policy Evaluation: A Critique," *Carnegie-Rochester Conference Series on Public Policy* 1 (1976): 19–46.

Further Reading

Avellaneda, M., R. Buff, C. Friedman, N. Grandchamp, L. Kruk, and J. Newman. "Weighted Monte Carlo: A New Technique for Calibrating Asset Pricing Models." *International Journal of Theoretical and Applied Finance* 4, no. 1 (February 2001): 91–119.

Bates, D. S. "Post '87 Crash Fears in the S&P Futures Market." *Journal of Econometrics* 94 (January/February 2000): 181–238.

Board of Governors of the Federal Reserve System, Office of the Comptroller of the Currency. "Supervisory Guidance on Model Risk Management," April 2011.

Cont, R. "Model Uncertainty and Its Impact on the Pricing of Derivative Instruments." *Mathematical Finance* 16, no. 3 (July 2006): 519–547.

Derman, E. "Model Risk," *Risk* 9, no. 5 (May 1996): 139–145.

Derman, E. *My Life as a Quant: Reflections on Physics and Finance*. Hoboken, NJ: John Wiley & Sons, 2004.

Elices, A., and E. Gimenez. "Weighted Monte Carlo." *Risk* (May 2006).

Glasserman, P., and X. Xu. "Robust Risk Measurement and Model Risk." *Quantitative Finance* 14, no. 1 (January 2014): 29–58.

Hull, J. C., and W. Suo. "A Methodology for the Assessment of Model Risk and Its Application to the Implied Volatility Function Model." *Journal of Financial and Quantitative Analysis* 37, no. 2 (June 2002): 297–318.

JPMorgan Chase. "Report of JPMorgan Chase & Co. Management Task Force Regarding 2012 CIO Losses," January 2013.

Practice Questions and Problems (Answers at End of Book)

25.1 Explain what is meant by (a) marking to market and (b) marking to model.

25.2 What are the two categories of model risk identified in SR 11-7?

25.3 Why is it important that a model be fully documented?

25.4 "The Black–Scholes–Merton model is nothing more than a sophisticated interpolation tool." Discuss this viewpoint.

25.5 Using Table 25.1, calculate the volatility a trader would use for an eight-month option with a strike price of 1.04.

25.6 What is the key difference between the models of physics and the models of finance?

25.7 How is a financial institution likely to find that it is using a model different from its competitors for valuing a particular type of derivatives product?

25.8 Distinguish between within-model and outside-model hedging.

25.9 What is meant by P&L decomposition?

25.10 What is meant by "Level 1," "Level 2," and "Level 3" valuations in accounting?

25.11 What is meant by over-fitting and over-parameterization in model building?

25.12 Section 25.4 calculates the implied volatility of an option with strike price 0.92 and time to maturity 1.5 years using the data in Table 25.1 as 14.59%. It does this by interpolating first along the strike price dimension and then along the time to maturity dimension. Show that the same answer is obtained by interpolating first along the time to maturity dimension and then along the strike price dimension.

25.13 "For structured products, traders mark to model. They do not mark to market." Explain this remark.

Further Questions

25.14 Suppose that all options traders decide to switch from Black–Scholes–Merton to another model that makes different assumptions about the behavior of asset prices. What effect do you think this would have on (a) the pricing of standard options and (b) the hedging of standard options?

25.15 Using Table 25.1, calculate the volatility a trader would use for an 11-month option with a strike price of 0.98.

25.16 Suppose that a financial institution uses an imprecise model for pricing and hedging a particular type of structured product. Discuss how, if at all, it is likely to realize its mistake.

Chapter 26

Economic Capital and RAROC

Up to now, we have focused on the development of procedures for evaluating different components of a financial institution's risk (credit risk, market risk, operational risk, and so on). We now consider how risks can be aggregated and allocated to different business units.

Economic capital (sometimes referred to as *risk capital*) is a financial institution's own internal estimate of the capital it needs for the risks it is taking. It is different from regulatory capital, which in the case of banks is based on one-size-fits-all rules determined by the Basel Committee. Economic capital can be regarded as a "currency" for risk-taking within a financial institution. A business unit can take a certain risk only when it is allocated the appropriate economic capital for that risk. The profitability of a business unit is measured relative to the economic capital allocated to the unit.

In this chapter, we discuss the approaches a financial institution uses to arrive at estimates of economic capital for particular risk types and particular business units, and how these estimates are aggregated to produce a single economic capital estimate for the whole financial institution. We also discuss risk-adjusted return on capital, or RAROC. This is the return earned by a business unit on the capital assigned to it. RAROC can be used to assess the past performance of business units. It can also be used to forecast future performance of the units and decide on the most appropriate way of allocating capital in the future. It provides a basis for determining whether some activities should be discontinued and others expanded.

26.1 Definition of Economic Capital

Economic capital is usually defined as the amount of capital a financial institution needs in order to absorb losses over one year with a certain confidence level. The confidence level is therefore the probability that the bank's losses will not exceed the economic capital in one year. The confidence level depends on the financial institution's objectives. A common objective for a large international bank is to maintain a AA credit rating. Corporations rated AA have a one-year probability of default of about 0.02%. This suggests that the confidence level should be set as high as 99.98% for economic capital to be a guide as to what is necessary to maintain a AA rating. The probability that capital will be insufficient to absorb losses so that a default results is then the required 0.02%. For a bank wanting to maintain a BBB credit rating, the confidence level can be lower. A BBB-rated corporation has a probability of about 0.2% of defaulting in one year, so that a confidence level of 99.8% gives the capital necessary to maintain a BBB rating.

Capital is required to cover unexpected loss. This is defined as the difference between the actual loss and the expected loss. As explained in Chapter 15, the idea here is that expected losses should be taken account of in the way a financial institution prices its products so that only unexpected losses require capital. As indicated in Figure 26.1, the economic capital is the difference between expected loss and the X percentile point on the probability distribution of the loss, where X% is the confidence level.

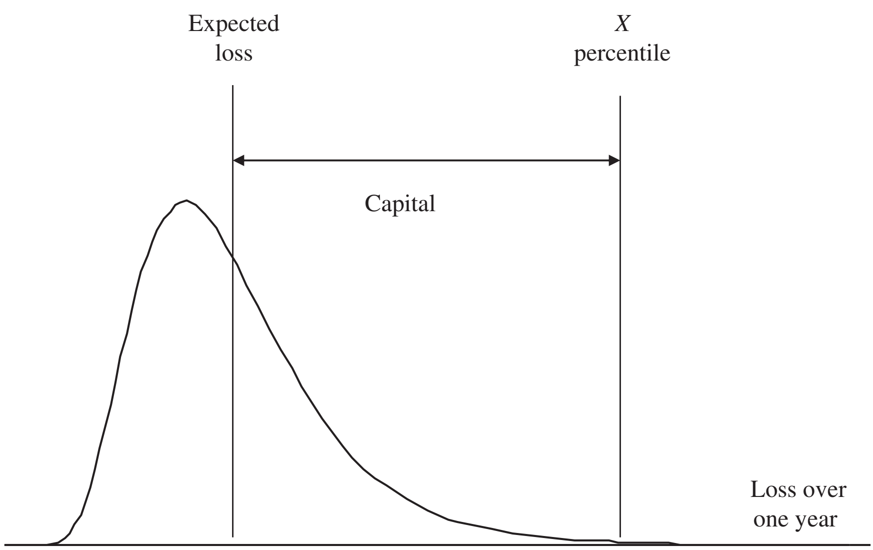


Figure 26.1 Calculation of Economic Capital from One-Year Loss Distribution
X% is the confidence level.

Example 26.1

Suppose that the main business of an AA-rated bank is lending in a certain region of the world. The bank estimates its losses as 1% of outstanding loan principal per year on average. The one-year 99.98% VaR is estimated as 5% of the outstanding loan principal. The economic capital required per $100 of loans made is therefore $4. (This is the difference between the one year 99.98% VaR and the expected loss.)

26.1.1 Approaches to Measurement

There are two broad approaches to measuring economic capital: the "top-down" and "bottom-up" approaches. In the top-down approach, the volatility of the financial institution's assets is estimated and then used to calculate the probability that the value of the assets will fall below the value of the liabilities by the end of the time horizon. A theoretical framework that can be used for the top-down approach is Merton's model, which was discussed in Section 19.8.

The approach most often used is the bottom-up approach where loss distributions are estimated for different types of risk and different business units and then aggregated. The first step in the aggregation can be to calculate probability distributions for losses by risk type or losses by business unit. A final aggregation gives a probability distribution of total losses for the whole financial institution.

The various risks facing a bank or other financial institution are summarized in Figure 26.2. As we saw in Chapter 23, bank regulators have chosen to define operational risk as: "The risk of loss resulting from inadequate or failed internal processes, people, and systems or from external events." Operational risk includes model risk and legal risk, but it does not include risk arising from strategic decisions and reputational risk. We will collectively refer to the latter risks as *business risk*. Regulatory capital is not required for business risk, but some banks do assess economic capital for it.

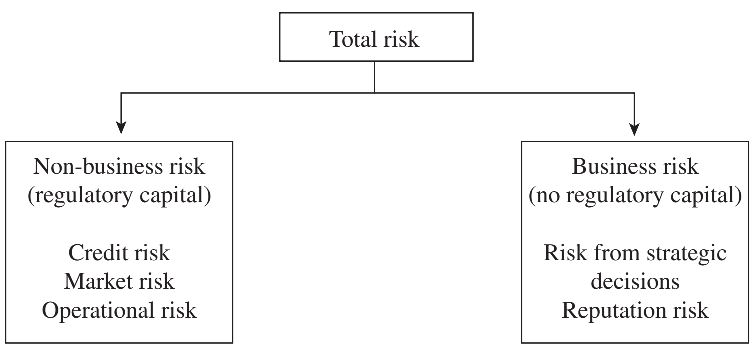


Figure 26.2 Categorization of Risks Faced by a Bank and Regulatory Capital Requirements

26.2 Components of Economic Capital

In earlier chapters, we covered approaches used to calculate loss distributions for different types of risk. This section reviews some of the key issues.

26.2.1 Market Risk Economic Capital

In Chapters 13 and 14, we discussed the historical simulation and model-building approaches for estimating the probability distribution of losses or gains from market risk. As explained, this distribution is usually calculated in the first instance with a one-day time horizon. Regulatory capital for market risk in Basel I and Basel II.5 is calculated from 10-day 99% VaR estimates. In FRTB, as described in Chapter 18, ES is used with time horizons determined by liquidity.

When calculating economic capital, we usually want to use the same time horizon and confidence level for all risks. The time horizon is usually one year and, as already discussed, the confidence level can be as high as 99.98%. Simple assumptions are (a) that the probability distributions of gains and losses for each day during the year are the same and (b) that the distributions are independent. The one-year loss or gain distribution is then approximately normal. Assuming 252 business days in the year, the standard deviation of the one-year loss or gain equals the standard deviation of the daily loss or gain multiplied by $\sqrt{252}$. If the average loss is zero, the 99.98% VaR is then 3.54 times the standard deviation of the one-year loss or gain. The 99.8% VaR is 2.88 times the standard deviation of the one-year loss or gain. Note that we are not assuming that the daily losses or gains are normal. All we are assuming is that they are independent and identically distributed. The central limit theorem of statistics tells us that the sum of many independent identically distributed variables is approximately normal.

Example 26.2

Suppose that the one-day standard deviation of market risk losses or gains for a bank is \$5 million. The one-year 99.8% VaR is $2.88 \times \sqrt{252} \times 5 = 228.6$ or \$228.6 million.

In practice, the losses (gains) on different days are not perfectly independent and identically distributed. The independence assumption can be relaxed by assuming a constant autocorrelation between the returns on successive days. This allows the result in equation (12.5) to be used and, providing the autocorrelation is not too large, the loss (gain) in a year can still be assumed to be approximately normal. Sometimes, it may be appropriate to develop a more elaborate model of how the standard deviation of the daily loss (gain) evolves through time. Monte Carlo simulation can then be used to aggregate the daily distributions and the result may not then be a normal distribution.

26.2.2 Credit Risk Economic Capital

Although Basel II gives banks that use the internal-ratings-based approach for regulatory capital a great deal of freedom, it does not allow them to choose their own credit correlation model and correlation parameters. When calculating economic capital, banks are free to make the assumptions they consider most appropriate for their situation. CreditMetrics, which is explained in Section 21.4, is a popular approach for calculating credit risk economic capital.

Another approach that has been proposed is Credit Risk Plus, which is described in Section 21.3. This approach borrows a number of ideas from actuarial science to calculate a probability distribution for losses from defaults. Whereas CreditMetrics calculates the loss from downgrades and defaults, Credit Risk Plus calculates losses from defaults only.

In calculating credit risk economic capital, a financial institution can choose to adopt a conditional or unconditional model. In a conditional (cycle-specific) model, the expected and unexpected losses take account of current economic conditions. In an unconditional (cycle-neutral) model, they are calculated by assuming economic conditions that are in some sense an average of those experienced through the cycle. Rating agencies aim to produce ratings that are unconditional. In addition, when regulatory capital is calculated using the internal-ratings-based approach, the PD and LGD estimates should be unconditional. Obviously, it is important to be consistent when economic capital is calculated. If expected losses are conditional, unexpected losses should also be conditional. If expected losses are unconditional, the same should be true of unexpected losses.

Whatever the approach used, a Monte Carlo simulation is usually necessary to calculate the probability distribution of credit losses. As explained in Chapters 20 and 21, derivatives complicate the calculations because of the uncertainty about the exposure at the time of a default or downgrade.

26.2.3 Operational Risk Economic Capital

As explained in Chapter 23, bank regulators are switching from the Advanced Measurement Approach (AMA) for determining operational risk regulatory capital to a simpler approach based on a bank's loss history. It seems likely that some sophisticated banks will continue to use the AMA for determining economic capital as it is based on the model in Figure 26.1.

26.2.4 Business Risk Economic Capital

As mentioned earlier, business risk includes strategic risk (relating to a bank's decision to enter new markets and develop new products) and reputational risk. Business risk is even more difficult to quantify than operational risk and estimates are likely to be largely subjective. However, attempting to quantify business risk can be a useful exercise, as it can lead to risk managers within a financial institution acquiring a good understanding

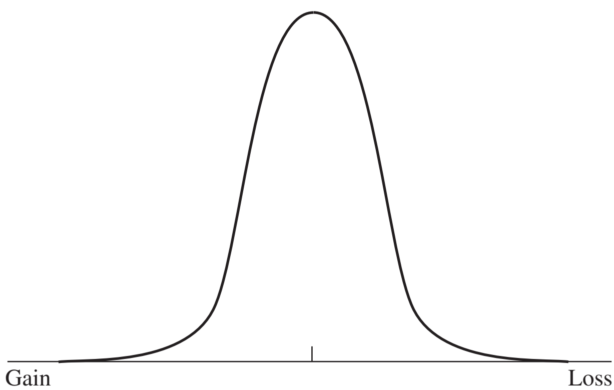


Figure 26.3 Loss Probability Density Function for Market Risk

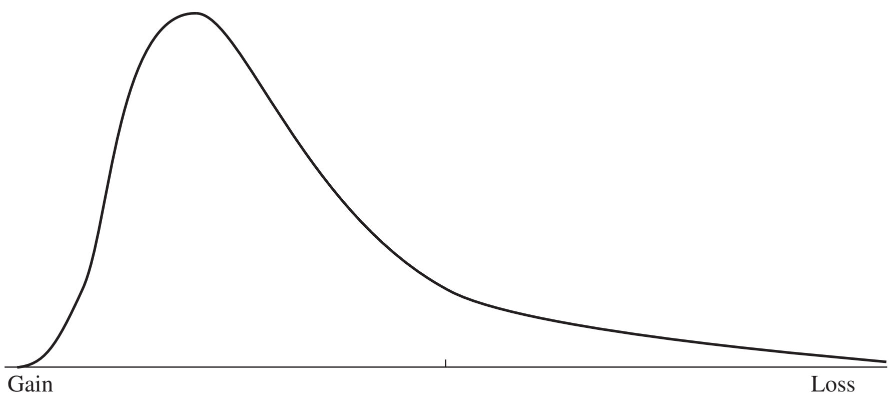


Figure 26.4 Loss Probability Density Function for Credit Risk

of the portfolio of business risks being taken. This enables them to assess the marginal impact on total risk of new strategic initiatives that are being contemplated.

26.3 Shapes of the Loss Distributions

The loss probability distributions for market, credit, and operational risk are very different. Rosenberg and Schuermann (2004) used data from a variety of different sources to estimate typical shapes for these distributions.[1] These are shown in Figures 26.3 to 26.5. The market risk loss distribution (see Figure 26.3) is symmetrical but not perfectly normally distributed. A *t*-distribution with 11 degrees of freedom provides a good fit. The credit risk loss distribution in Figure 26.4 is quite skewed, as one would expect. The operational risk distribution in Figure 26.5 has a quite extreme shape. Most of the time, losses are modest, but occasionally they are very large.

[1] See J. V. Rosenberg and T. Schuermann, "A General Approach to Integrated Risk Management with Skewed, Fat-Tailed Risks," Federal Reserve Bank of New York, Staff Report no. 185, May 2004.

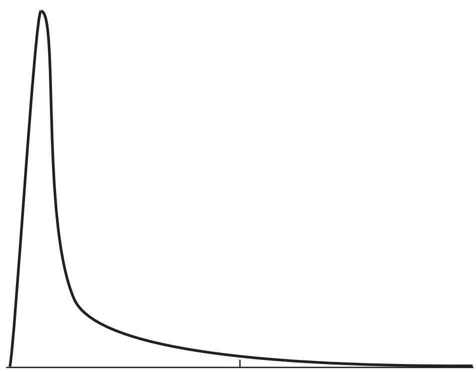


Figure 26.5 Loss Probability Density Function for Operational Risk

We can characterize a distribution by its second, third, and fourth moments. Loosely speaking, the second moment measures standard deviation (or variance), the third moment measures skewness, and the fourth moment measures kurtosis (i.e., the heaviness of tails). Table 26.1 summarizes the properties of typical loss distributions for market, credit, and operational risk.

26.4 Relative Importance of Risks

The relative importance of different types of risk depends on the business mix. Credit risk is important in commercial lending, retail lending, and a financial institution's derivatives business. Market risk is important in trading and some investment banking activities. But as mentioned in Chapter 23, operational risk (particularly cyber risk, legal risk, and compliance risk) is now considered by many observers to be the most important risk for financial institutions.

Operational risk has always been considered to be the most important risk in asset management. If the asset manager is found to be negligent in some way, there are liable to be expensive investor lawsuits. Business Snapshot 26.1 gives one example of this. Another high-profile example is provided by Unilever's pension plan. Mercury Asset Management, owned by Merrill Lynch, pledged not to underperform a benchmark index by more than 3%. Between January 1997 and March 1998, it underperformed the index by 10.5%. Unilever sued Merrill Lynch for $185 million and the matter was settled out of

Table 26.1 Characteristics of Loss Distributions for Different Risk Types

	Second Moment (standard deviation)	**Third Moment (skewness)**	**Fourth Moment (kurtosis)**
Market risk	High	Zero	Low
Credit risk	Moderate	Moderate	Moderate
Operational risk	Low	High	High

BUSINESS SNAPSHOT 26.1

The EGT Fund

In 1996, Peter Young was fund manager at Deutsche Morgan Grenfell, a subsidiary of Deutsche Bank. He was responsible for managing a fund called the European Growth Trust (EGT). It had grown to be a very large fund and Young had responsibilities for managing over one billion pounds of investors' money.

Certain rules applied to EGT. One of these was that no more than 10% of the fund could be invested in unlisted securities. Peter Young violated this rule in a way that it can be argued benefited him personally. When the facts were uncovered, he was fired and Deutsche Bank had to compensate investors. The total cost to Deutsche Bank was over 200 million pounds.

court. The Spanish bank Santander incurred operational risk losses in 2009 on funds it managed for investors and placed with Bernard Madoff, who, it transpired, was running a $50 billion Ponzi scheme.

26.4.1 Interactions between Risks

There are interactions between the different types of risks. If a financial institution's counterparty defaults, credit risk exists only if market variables have moved so that the uncollateralized value of the derivative to the financial institution is positive. Another interaction is that the probability of default by a counterparty may depend on the value of a financial institution's contract (or contracts) with the counterparty. This is the wrong-way issue discussed in Chapter 20. If the counterparty has entered into the contract for hedging purposes, the dependence should be small. However, if the contract has been entered into for speculative purposes and the contract is large in relation to the size of the counterparty, the dependence is likely to be important.

As the Long-Term Capital Management saga clearly shows, there can be interactions between liquidity risks and market risks (see Business Snapshot 22.1). There are also interactions between operational risks and market risks. It is unlikely that we would know about the activities of Jérôme Kerviel at Société Générale if he had guessed right about market movements (see Business Snapshot 5.5). Similarly, we are unlikely to hear about a violation of the rules for a fund (such as the one in Business Snapshot 26.1) if the violation leads to a gain rather than a loss.

26.5 Aggregating Economic Capital

A financial institution typically calculates market, credit, operational, and (possibly) business risk loss distributions for a number of different business units. It is then faced with the problem of aggregating the loss distributions to calculate a total economic capital for the whole enterprise.

The simplest approach is to assume that the total economic capital for a set of n different risks is the sum of the economic capital amounts for each risk considered separately, so that

$$E_{\text{total}} = \sum_{i=1}^{n} E_i \tag{26.1}$$

where E_{total} is the total economic capital for the financial institution facing n different risks and E_i is the economic capital for the ith risk considered on its own. This is what the Basel Committee does for regulatory capital. The total regulatory capital a bank is required to keep is the sum of the regulatory capital amounts for credit, market, and operational risks.

Equation (26.1) is clearly a very conservative assumption. It assumes perfect correlation. In the context of economic capital calculations where the confidence level is, say, 99.9%, it would mean that, if the market risk loss for a financial institution is in the extreme 0.1% tail of the distribution, the same is true for the credit risk loss and the operational risk loss. Rosenberg and Schuermann estimate the correlation between market risk and credit risk to be approximately 50% and the correlation between each of these risks and operational risk to be approximately 20%. They estimate that equation (26.1) when used as a way of aggregating market, credit, and operational risks overstates the total capital required by about 40%.

26.5.1 Assuming Normal Distributions

A simple assumption when aggregating loss distributions is that they are normally distributed. The standard deviation of the total loss from n sources of risk is then

$$\sigma_{\text{total}} = \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n} \sigma_i \sigma_j \rho_{ij}} \tag{26.2}$$

where σ_i is the standard deviation of the loss from the ith source of risk and ρ_{ij} is the correlation between risk i and risk j. The capital requirement can be calculated from this. For example, the excess of the 99.9% VaR over the expected loss is 3.09 times the number calculated in equation (26.2).

This approach tends to underestimate the capital requirement because it takes no account of the skewness and kurtosis of the loss distributions. Rosenberg and Schuermann estimate that, when the approach is applied to aggregating market, credit, and operational risks, the total capital is underestimated by about 40%.

26.5.2 Using Copulas

A more sophisticated approach to aggregating loss distributions is by using copulas. Copulas were discussed in Chapter 11. Each loss distribution is mapped on a percentile to percentile basis to a standard well-behaved distribution. A correlation structure between the standard distributions is defined and this indirectly defines a correlation structure between the original distributions.

Many different copulas can be defined. The simplest is the Gaussian copula, where the standard distributions are assumed to be multivariate normal. Other alternatives, with more tail correlation, are explained in Section 11.6.

26.5.3 The Hybrid Approach

The approach in Section 12.9 for aggregating VaR is a popular approach for aggregating economic capital estimates. It is sometimes referred to as the *hybrid approach*. It involves calculating the economic capital for a portfolio of risks from the economic capital for the individual risks using

$$E_{\text{total}} = \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n} E_i E_j \rho_{ij}} \tag{26.3}$$

When the distributions are normal, this approach is exactly correct. When distributions are non-normal, the hybrid approach gives an approximate answer—but one that reflects any heaviness in the tails of the individual loss distributions. Rosenberg and Schuermann find that the answers given by the hybrid approach are reasonably close to those given by copula models.

Example 26.3

Suppose that the estimates for economic capital for market, credit, and operational risk for two business units are as shown in Table 26.2. The correlations between the losses are shown in Table 26.3. The correlation between credit risk and market risk within the same business unit is 0.5, and the correlation between operational risk and either credit or market risk within the same business unit is 0.2. (These correspond to the estimates of Rosenberg and Schuermann mentioned earlier.) The correlation between two different

Table 26.2 Economic Capital Estimates for Example 26.3

	Business Units	
	1	**2**
Market Risk	30	40
Credit Risk	70	80
Operational Risk	30	90

Table 26.3 Correlations between Losses in Example 26.3

	MR-1	CR-1	OR-1	MR-2	CR-2	OR-2
MR-1	1.0	0.5	0.2	0.4	0.0	0.0
CR-1	0.5	1.0	0.2	0.0	0.6	0.0
OR-1	0.2	0.2	1.0	0.0	0.0	0.0
MR-2	0.4	0.0	0.0	1.0	0.5	0.2
CR-2	0.0	0.6	0.0	0.5	1.0	0.2
OR-2	0.0	0.0	0.0	0.2	0.2	1.0

MR, CR, and OR refer to market risk, credit risk, and operational risk; 1 and 2 refer to business units.

risk types in two different business units is zero. The correlation between market risks across business units is 0.4. The correlation between credit risks across business units is 0.6. The correlation between operational risks across business units is zero.

We can aggregate the economic capital using the hybrid approach. The total market risk economic capital is

$$\sqrt{30^2 + 40^2 + 2 \times 0.4 \times 30 \times 40} = 58.8$$

The total credit risk economic capital is

$$\sqrt{70^2 + 80^2 + 2 \times 0.6 \times 70 \times 80} = 134.2$$

The total operational risk economic capital is

$$\sqrt{30^2 + 90^2} = 94.9$$

The total economic capital for business unit 1 is

$$\sqrt{30^2 + 70^2 + 30^2 + 2 \times 0.5 \times 30 \times 70 + 2 \times 0.2 \times 30 \times 30 + 2 \times 0.2 \times 70 \times 30} = 100.0$$

The total economic capital for business unit 2 is

$$\sqrt{40^2 + 80^2 + 90^2 + 2 \times 0.5 \times 40 \times 80 + 2 \times 0.2 \times 40 \times 90 + 2 \times 0.2 \times 80 \times 90} = 153.7$$

The total enterprise-wide economic capital is the square root of

$$\begin{aligned}&30^2 + 40^2 + 70^2 + 80^2 + 30^2 + 90^2 + 2 \times 0.4 \times 30 \times 40 + 2 \times 0.5 \times 30 \times 70\\&\quad + 2 \times 0.2 \times 30 \times 30 + 2 \times 0.5 \times 40 \times 80 + 2 \times 0.2 \times 40 \times 90\\&\quad + 2 \times 0.6 \times 70 \times 80 + 2 \times 0.2 \times 70 \times 30 + 2 \times 0.2 \times 80 \times 90\end{aligned}$$

or 203.2.

There are clear diversification benefits. The sum of the economic capital estimates for market, credit, and operational risk is $58.8 + 134.2 + 94.9 = 287.9$ and the sum of the economic capital estimates for two business units is $100 + 153.7 = 253.7$. Both of these are greater than the total economic capital estimate of 203.2.

26.6 Allocation of Economic Capital

Suppose that the sum of the economic capital for the business units in a bank, $\sum_{i=1}^{n} E_i$, is \$2 billion and that the total economic capital for the whole bank, after taking less-than-perfect correlations into account, is \$1.3 billion (= 65% of the sum of the E_i). The \$0.7 billion is a diversification gain to the bank. How should the total economic capital be allocated to the business units?

A simple approach is to allocate $0.65E_i$ to business unit i. However, this is probably not the best approach. Consider a situation where there are 50 business units and that two particular units both have an economic capital of \$100 million. Suppose that, when the first business unit is excluded from the calculations, the bank's economic capital decreases by \$60 million and that, when the second business unit is excluded from the calculation, the bank's economic capital decreases by \$10 million. Arguably, the first business unit should have more economic capital than the second, because its incremental impact on the bank's total economic capital is greater.

The issues here are analogous to those we discussed in Section 12.8 for allocating VaR. One approach is to calculate incremental economic capital for each business unit and then allocate economic capital to business units in proportion to their incremental capital. (Incremental capital is the difference between the total economic capital with and without the business unit.) A popular approach is to work with the component economic capital. This involves allocating

$$x_i \frac{\partial E_{\text{total}}}{\partial x_i}$$

to the ith business unit, where x_i is the investment in the ith business unit. As we pointed out in Section 12.8, a result known as Euler's theorem then ensures that the total of the allocated capital is the total economic capital E_{total}.

Define Q_i as the increase in the total economic capital when we increase x_i by a small amount Δx_i. A discrete approximation for the component economic capital for business unit i is

$$\frac{Q_i}{\gamma_i} \tag{26.4}$$

where $\gamma_i = \Delta x_i / x_i$.

Example 26.4

Consider again Example 26.3. The total economic capital is 203.2. The economic capital calculated for Business Unit 1 is 100 and that calculated for Business Unit 2 is 153.7.

A simple approach would allocate 100/253.7 of the total economic capital to Business Unit 1 and 153.7/253.7 of the economic capital to Business Unit 2. This would result in 80.1 for Business Unit 1 and 123.1 to Business Unit 2.

The incremental effect of Business Unit 1 on the total economic capital is $203.2 - 153.7 = 49.5$. Similarly, the incremental effect of Business Unit 2 on the total economic capital is $203.2 - 100 = 103.2$. The two incremental capitals do not add up to the total capital (as is usually the case). However, we could use them as a basis for allocating the total capital. We would then allocate $49.5/(49.5 + 103.2)$ of the capital to Business Unit 1 and $103.2/(49.5 + 103.2)$ of it to Business Unit 2. This would result in 65.9 for Business Unit 1 and 137.3 for Business Unit 2.

To apply equation (26.4) and allocate component economic capital to each unit, we could calculate the partial derivative analytically. Alternatively, we can use a numerical calculation. When we increase the size of Business Unit 1 by 1%, its economic capital amounts for market, credit, and operational risk in Table 26.2 increase to 30.3, 70.7, and 30.3, respectively. The total economic capital becomes 203.906 so that $Q_1 = 203.906 - 203.224 = 0.682$.

When we increase the size of Business Unit 2 by 1%, its economic capital amounts for market, credit, and operational risk in Table 26.2 increase to 40.4, 80.8, and 90.9, respectively. The total economic capital becomes 204.577 so that $Q_2 = 204.577 - 203.224 = 1.353$.

In this case, because we are considering 1% increases in the size of each unit, $\gamma_1 = \gamma_2 = 0.01$. From equation (26.4), the economic capital allocations to the two business units are 68.2 and 135.3. (These do not quite add up to the total economic capital of 203.2 because we approximated the partial derivative.)

26.7 Deutsche Bank's Economic Capital

Deutsche Bank publishes the results of its economic capital calculations in its annual financial statements. Table 26.4 summarizes the economic capital and regulatory capital reported at the end of 2016. Capital is calculated for credit risk, market risk, operational risk, and business risk. Deutsche Bank calculated a diversification benefit reflecting the lack of perfect correlation between the different risk types. The total economic capital was about 35.4 billion euros.

The actual capital held was about 42.3 billion euros of core Tier 1 capital (i.e., common equity), 4.6 billion euros of additional Tier 1 capital, and about 12.7 billion euros of Tier 2 capital. Table 26.4 shows these capital amounts as a percentage of risk-weighted

Table 26.4 Deutsche Bank's Economic Capital and Regulatory Capital, December 2016 (millions of euros)

Credit risk	13,105
Market risk	14,593
Operational risk	10,488
Business risk	5,098
Diversification benefit	(7,846)
Total economic capital	35,438
Total risk-weighted assets	357,518
Common Equity Tier 1 capital (% of risk-weighted assets)	11.8%
All Tier 1 capital (% of risk-weighted assets)	13.1%
Total (Tier 1 plus Tier 2) capital (% of risk-weighted assets)	16.6%

assets. Deutsche Bank exceeds the capital requirements in Basel III, including the G-SIB additional capital, which in the case of Deutsche Bank was 2% of risk-weighted assets.

Table 26.5 shows how this capital was allocated to business units. The largest user of economic capital by far was Global Markets, which Deutsche Bank states is dominated by trading in support of origination, structuring, and market making. Credit risk, market risk, operational risk, and business risk were all greatest for this division.

26.8 RAROC

Risk-adjusted performance measurement (RAPM) has become an important part of how business units are assessed. There are many different approaches, but all have one thing in common. They compare return with capital employed in a way that incorporates an adjustment for risk.

The most common approach is to compare expected return with economic capital. This is usually referred to as RAROC (risk-adjusted return on capital). The formula is

$$\text{RAROC} = \frac{\text{Revenues} - \text{Costs} - \text{Expected losses}}{\text{Economic capital}} \tag{26.5}$$

Table 26.5 Allocation of Deutsche Bank's Economic Capital to Divisions, December 2016 (millions of euros)

Division	Allocated Economic Capital
Global Markets	15,587
Corporate and Investment Banking	5,015
Private Wealth and Commercial Clients	2,473
Deutsche Asset Management	2,480
Postbank	3,976
Non-Core Operations Unit	735
Consolidations and Adjustments and Other	5,172
Total	35,438

The numerator may be calculated on a pre-tax or post-tax basis. Sometimes, a risk-free rate of return on the economic capital is calculated and added to the numerator.

Example 26.5

A bank rated AA estimates its losses on loans as 1% of the outstanding principal on average. The 99.98% VaR (i.e., the loss exceeded only 0.02% of the time) is 5% of outstanding loans. As indicated in Example 26.1, the economic capital required per $100 of loans is $4, which is the difference between the 99.98% VaR and the expected loss. (This ignores diversification benefits that would in practice be allocated to the lending unit as a result of less than perfect correlation with the bank's other activities.) The spread between the cost of funds and the interest charged is 2.5%. Subtracting from this the expected loan loss of 1%, the expected contribution per $100 of loans is $1.50. Assuming that the lending unit's administrative costs total 0.7% of the amount loaned, the expected profit is reduced to $0.80 per $100 in the loan portfolio. RAROC is therefore

$$\frac{0.80}{4} = 20\%$$

An alternative calculation would add the interest on the economic capital to the numerator. Suppose that the risk-free interest rate is 2%. Then $0.02 \times 4 = 0.08$ is added to the numerator so that RAROC becomes

$$\frac{0.88}{4} = 22\%$$

As pointed out by Matten (2000), it is more accurate to refer to the approach in equation (26.5) as RORAC (return on risk-adjusted capital) rather than RAROC, because it is the capital (i.e., the denominator) not the return that reflects risk.[2] In theory, RAROC should involve adjusting the return (i.e., the numerator) for risk. In practice, RAROC is used to denote a wide range of return on capital calculations.

RAROC can be calculated ex-ante (before the start of the year) or ex-post (after the end of the year). Ex-ante calculations are based on estimates of expected profit. Ex-post calculations are based on actual profit results. Ex-ante calculations are typically used to decide whether a particular business unit should be expanded or contracted. Ex-post calculations are typically used for performance evaluation and bonus calculations.

It is usually not appropriate to base a decision to expand or contract a particular business unit on an ex-post analysis (although there is a natural temptation to do this). It may be that results were bad for the most recent year because of a one-time event

---

[2] See C. Matten, *Managing Bank Capital: Capital Allocation and Performance Measurement*, 2nd ed. (Chichester, UK: John Wiley & Sons, 2000).

(e.g., an unusually high credit loss or a particularly high operational risk loss). Key strategic decisions should be based on expected long-term results.

Summary

Economic capital is the capital that a bank or other financial institution deems necessary for the risks it is bearing. When calculating economic capital, a financial institution is free to adopt any approach it likes. It does not have to use the one proposed by regulators. Typically, it estimates economic capital for credit risk, market risk, operational risk, and (possibly) business risk for its business units and then aggregates the estimates to produce an estimate of the economic capital for the whole enterprise. The risks are usually assumed to be less than perfectly correlated. The benefits of diversification are estimated and allocated to business units. Usually the approach used is designed to reflect the incremental impact of the business units on the total economic capital.

The one-year loss distributions for market risk, credit risk, and operational risk are quite different. The loss distribution for market risk is symmetrical. For credit risk it is skewed, and for operational risk it is highly skewed with very heavy tails.

Financial institutions calculate RAROC, the risk adjusted return on economic capital, for each business unit by dividing the profit of the business unit by its allocated economic capital. The expected future RAROC is used to decide which areas of the business should be expanded or contracted. The actual RAROC achieved by a business unit is used for performance evaluation.

Further Reading

Dev, A. *Economic Capital: A Practitioner's Guide*. London: Risk Books, 2004.

Matten, C. *Managing Bank Capital: Capital Allocation and Performance Measurement*. 2nd ed. Chichester, UK: John Wiley & Sons, 2000.

Rosenberg, J. V., and T. Schuermann. "A General Approach to Integrated Risk Management with Skewed, Fat-Tailed Risks." Federal Reserve Bank of New York, Staff Report no. 185, May 2004.

Practice Questions and Problems (Answers at End of Book)

26.1 What is the difference between economic capital and regulatory capital?

26.2 What determines the confidence level used by a AA-rated financial institution in its economic capital calculations?

26.3 What types of risk are included in business risk?

26.4 In what respects are the models used to calculate economic capital for market risk, credit risk, and operational risk likely to be different from those used to calculate regulatory capital?

26.5 Suppose that the credit loss in a year has a lognormal distribution. The logarithm of the loss is normal with mean 0.5 and standard deviation 4. What is the economic capital requirement if a confidence level of 99.97% is used?

26.6 Suppose that the economic capital estimates for two business units are as follows:

	Business Units	
	1	**2**
Market Risk	20	40
Credit Risk	40	30
Operational Risk	70	10

The correlations are as in Table 26.3. Calculate the total economic capital for each business unit and the two business units together.

26.7 In Problem 26.6, what is the incremental effect of each business unit on the total economic capital? Use this to allocate economic capital to business units. What is the impact on the economic capital of each business unit increasing by 0.5%? Show that your results are consistent with Euler's theorem.

26.8 A bank is considering expanding its asset management operations. The main risk is operational risk. It estimates the expected operational risk loss from the new venture in one year to be \$2 million and the 99.97% VaR (arising from a small chance of an large investor lawsuit) to be \$40 million. The expected fees it will receive from investors for the funds under administration are \$12 million per year, and administrative costs are expected to be \$3 million per year. Estimate the before-tax RAROC.

26.9 RAROC can be used in two different ways. What are they?

Further Questions

26.10 Suppose that daily gains (losses) are normally distributed with a standard deviation of \$5 million.

(a) Estimate the minimum regulatory capital the bank is required to hold. (Assume a multiplicative factor of 4.0.)

(b) Estimate the economic capital using a one-year time horizon and a 99.9% confidence level assuming that there is a correlation of 0.05 between gains (losses) on successive days.

26.11 Suppose that the economic capital estimates for two business units are

	Business Units	
	1	**2**
Market Risk	10	50
Credit Risk	30	30
Operational Risk	50	10

The correlation between market risk and credit risk in the same business unit is 0.3. The correlation between credit risk in one business unit and credit risk in another is 0.7. The correlation between market risk in one business unit and market risk in the other is 0.2. All other correlations are zero. Calculate the total economic capital. How much should be allocated to each business unit?

26.12 Suppose that a bank's sole business is to lend in two regions of the world. The lending in each region has the same characteristics as in Example 26.5 of Section 26.8. Lending to Region A is three times as great as lending to Region B. The correlation between loan losses in the two regions is 0.4. Estimate the total RAROC.

26.13 Suppose daily losses (gains) from trading are independent and normally distributed with mean zero. Calculate in terms of the standard deviation of the daily losses (gains) (a) the basic Basel I regulatory capital requirement calculated as 3 times the 10-day 99% VaR and (b) the economic capital calculated using a 99.97% confidence level and a one-year time horizon. Would you expect the economic and regulatory capital to become closer together or farther apart if daily losses/gains are generated by a distribution with much heavier tails than the normal distribution? What would you expect to be the impact of the daily losses/gains exhibiting positive autocorrelation?

Chapter 27

Enterprise Risk Management

Much of this book has focused on understanding and quantifying different types of risk: market risk, credit risk, operational risk, and so on. An important part of a risk manager's job is to ensure that these risks are correctly evaluated. But it is also important for the risk management function to take a holistic, big-picture view of risk. It should identify potential adverse events and their full consequences. The total exposure to an adverse event can be greater than (or less than) the result obtained by considering each risk type separately. Enterprise risk management (ERM) is the name given to this holistic approach to risk management.

In understanding ERM, it is important to distinguish top-down and bottom-up approaches to risk management. Bottom-up approaches are concerned with assessing the different types of risk borne by different business units and combining them. We discussed how this can be done in the previous chapter. In top-down approaches, the overall risk appetite of the organization is defined, and this is then used to define risk limits for different parts of the organization. In practice, a financial institution needs to use both top-down and bottom-up approaches. A top-down approach is necessary to define the overall risk appetite, and a bottom-up approach is necessary to evaluate whether the risks being taken by business units are consistent with this risk appetite.

In 2004, the Committee of Sponsoring Organizations of the Treadway Commission (COSO) issued the Enterprise Risk Management Integrated Framework (see www.coso.org) and proposed this definition of ERM:

> *Enterprise risk management is a process, effected by an entity's board of directors, management, and other personnel, applied in strategy setting and across the enterprise, designed to*

identify potential events that may affect the entity, and manage risk to be within its risk appetite, to provide reasonable assurance regarding the achievement of entity objectives.

This definition makes a number of points. First, there should be board involvement in ERM. Second, ERM is a component of a company's strategy. Third, it involves identifying potential adverse events. Fourth, as part of ERM the company must identify its risk appetite and manage risks in a way that is consistent with that appetite. Fifth, ERM should help a company achieve its objectives and be a core part of strategic planning and strategic execution processes.

Many financial institutions (and their boards) have historically had a compliance orientation as far as top-down risk management is concerned. Their aim was to satisfy regulators and stay within the letter of the law. ERM is moving financial institutions away from this mind-set. It is leading to the risk management function having a strategic orientation and being value enhancing. A company that manages risk effectively has a competitive advantage over another company that does not do so.

This chapter considers the essential elements of ERM. It considers risk appetite, risk culture, how potential adverse events are identified, and how ERM is integrated into a financial institution's strategy.

27.1 Risk Appetite

An essential part of ERM is concerned with specifying the organization's risk appetite. How much risk is the organization prepared to take in order to achieve its strategic objectives and business plans? One dimension of the risk appetite of a financial institution is likely to be concerned with the loss it is prepared to incur if a worst-case scenario happens. This can be expressed in terms of an enterprise-wide value at risk (VaR) or expected shortfall (ES) measure.

For a fund manager, this dimension of risk appetite is relatively easy to quantify—at least in theory. Chapter 1 explained the theoretical trade-off between expected return and risk. The mix of assets influences the beta of the portfolio, and this in turn determines the theoretical expected return and risk.[1] Figure 27.1, which is similar to Figure 1.4 in Chapter 1, shows the trade-off between risk and expected return for a portfolio. As shown in Section 1.2, if the beta of the portfolio is β, the expected return μ_P is

$$\mu_P = (1 - \beta)R_F + \beta R_M$$

and the standard deviation of the return is

$$\sigma_P = \beta\sigma_M$$

[1] In practice, as explained in Chapter 1, the theoretical return for a particular value of beta is often used as a benchmark. Fund managers who beat the benchmark are adding value.

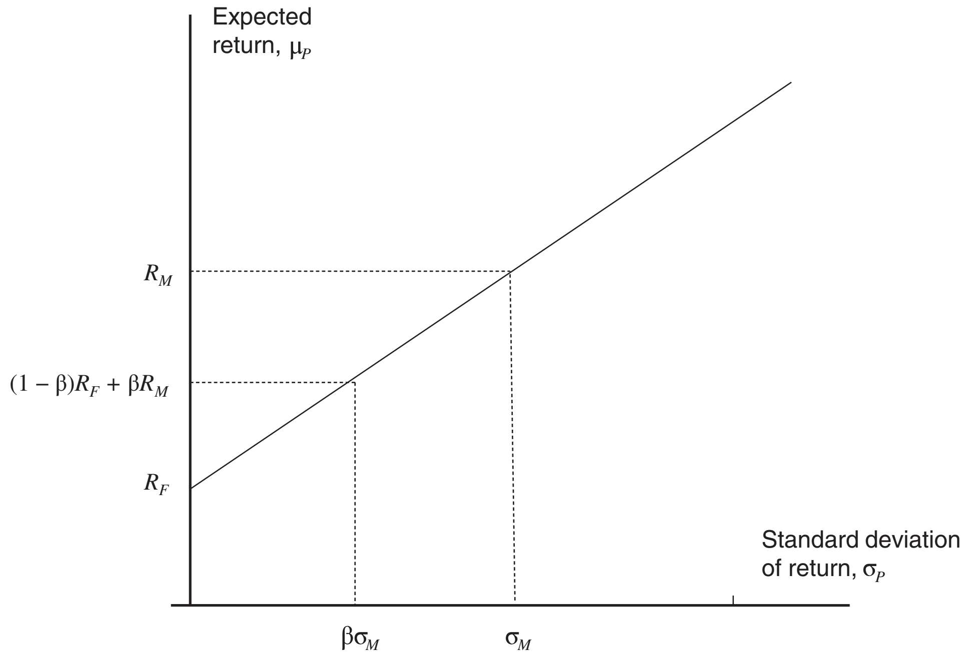


Figure 27.1 Dependence of Mean and Standard Deviation of a Portfolio on Beta, β

where σ_M is the standard deviation of the market return, R_F is the risk-free return, and R_M is the expected return on the market. Assume returns are normally distributed.[2] The return, R, that has a probability p of being exceeded in one year is given by

$$N\left(\frac{R - \mu_P}{\sigma_P}\right) = (1 - p)$$

where N is the cumulative normal distribution function. Substituting for μ_P and σ_P we obtain

$$\frac{R - (1 - \beta)R_F - \beta R_M}{\beta\sigma_M} = N^{-1}(1 - p)$$

or

$$\beta = \frac{R - R_F}{R_M - R_F + N^{-1}(1 - p)\sigma_M} \tag{27.1}$$

This formula allows us to determine the beta that is consistent with a particular VaR objective. Table 27.1 shows returns including dividends from the S&P 500 over the 47 years between 1970 and 2016. The mean and standard deviation of returns over

[2] This assumption is made to provide a simple illustration. The returns from the market have heavier tails than the normal distribution, so in practice the returns would be fitted to a different distribution, perhaps a Student's t-distribution with a relatively small number of degrees of freedom.

Table 27.1 Total Annual Return from the S&P 500 (including dividends) between 1970 and 2016

Year	Return (%)	Year	Return (%)	Year	Return (%)
1970	4.01	1986	18.67	2002	–22.10
1971	14.31	1987	5.25	2003	28.68
1972	18.98	1988	16.61	2004	10.88
1973	–14.66	1989	31.69	2005	4.91
1974	–26.47	1990	–3.10	2006	15.79
1975	37.20	1991	30.47	2007	5.49
1976	23.84	1992	7.62	2008	–37.00
1977	–7.18	1993	10.08	2009	26.46
1978	6.56	1994	1.32	2010	15.06
1979	18.44	1995	37.58	2011	2.11
1980	32.50	1996	22.96	2012	16.00
1981	–4.92	1997	33.36	2013	32.39
1982	21.55	1998	28.58	2014	13.69
1983	22.56	1999	21.04	2015	1.38
1984	6.27	2000	–9.10	2016	18.86
1985	31.73	2001	–11.89		

Table 27.2 Return Statistics from S&P 500 Data in Table 27.1

Period	Mean Annual Return (%)	Standard Deviation of Annual Returns (%)
2007 to 2016	8.75	17.11
1997 to 2016	9.38	18.50
1970 to 2016	11.74	18.86

the final 10 years, the final 20 years, and the whole 47 years are as shown in Table 27.2. Suppose that the risk-free rate of interest that can be earned is 2% ($R_F = 0.02$) and that the past 10 years are used to estimate R_M and σ_M so that $R_M = 0.0875$ and $\sigma_M = 0.1711$. Suppose further that a fund's risk appetite is such that it wants to have a 95% probability of not losing more than 10% in one year. Substituting $R = -0.1$ and $p = 0.95$ into equation (27.1), the fund manager's beta should be

$$\beta = \frac{-0.1 - 0.02}{0.0875 - 0.02 + N^{-1}(0.05) \times 0.1711} = 0.56$$

This means that to achieve the VaR objective the risk manager should invest a little more than half a portfolio in the market and the rest a risk-free asset.[3]

[3] In practice, a fund manager is likely to want to incorporate his or her judgment about returns into the choice of a portfolio. A way this can be done is provided by F. Black and R. Litterman, "Asset Allocation: Combining Investor Views with Market Equilibrium," *Journal of Fixed Income* 1, no. 2 (September 1991): 7–18.

Unfortunately, the available risk-return alternatives for a financial institution such as a bank or an insurance company cannot usually be quantified as easily as for a fund manager. When markets are efficient, we can use the arguments in Chapter 1 to show that for stock market and similar investments an efficient frontier similar to that in Figure 27.1 should exist. But there is no reason why this efficient frontier should define trade-offs between risk and expected return for major strategic investments undertaken by a bank, insurance company, or other type of financial institution. Some strategic investments will be above the efficient frontier and some will be well below.

For any company, strategic investments that are above the efficient frontier are usually those that take advantage of the company's competitive advantages. Companies, whether financial or non-financial, should therefore define their competitive advantages and look for strategic investments that exploit those advantages. There is some evidence that they have not always been very successful at doing this. A well-known study by Bowman uncovered a negative relationship between risk and return for companies in most sectors.[4] This is sometimes termed the *Bowman paradox*. In a further study, Bowman showed that more troubled firms often take less justifiable risks.[5] One explanation for this might be that a company experiencing difficulties often has few competitive advantages. It may be tempted to undertake a major project that has only a small probability of success if success would ensure survival.

The market in which financial institutions operate is highly competitive, and competitive advantages are not as common as they are in some other industries. Typically, attempts by a financial institution to increase returns involve an increase in systematic (beta) risk, and the financial institution's shareholders can often achieve an equivalent (and sometimes better) risk-return trade-off for themselves. The main competitive advantage of a financial institution is the trust of its customer base (retail and commercial) that it has built up. Providing services to that customer base should give a satisfactory risk-return trade-off. For regulated financial institutions, the risk-return trade-off is impacted by increasingly stringent capital and liquidity requirements.

As mentioned, the part of a financial institution's risk appetite concerned with the worst-case loss, or equivalently the worst-case erosion of equity capital, can be expressed by specifying VaR or expected shortfall with a particular time horizon and confidence level. Other parts of the risk appetite are less easy to quantify. They are likely to be concerned with reputational risk, credit rating risk, the risk of not complying with regulations, legal risks, and so on. Some of these are likely to be expressed qualitatively rather than quantitatively. For example, a bank might state that any project that could appreciably reduce its reputation among its core customers, and how trustworthy it is perceived to be by those customers, is unacceptable.

Business Snapshot 27.1 describes how Wells Fargo put employees under a huge amount of pressure to meet unrealistic sales targets in order to increase sales. This strategy

---

[4] Bowman, E. H., "A Risk-Return Paradox for Strategic Management," *Sloan Management Review* 21 (1980): 17–31.

[5] Bowman, E. H., "Risk Seeking by Troubled Firms," *Sloan Management Review* 23 (1982): 33–42.

BUSINESS SNAPSHOT 27.1

Cross-Selling at Wells Fargo

Wells Fargo is a bank that survived the 2007–2008 crisis well. It avoided the risks taken by other banks during the period leading up to the crisis and gained a reputation for stability. Unfortunately, subsequent events tarnished its reputation and led to a loss of trust from potential clients.

Cross-selling involves attempting to sell multiple products to clients. A client with a checking account might be encouraged to take out a credit card or open a savings account. Wells Fargo was considered to be one of the best banks at cross-selling, and indeed there is nothing wrong with cross-selling. But if it is carried out too aggressively it can be bad for a bank's reputation and even be fraudulent.

Wells Fargo considered branch employees to be "salespeople" and clients to be "customers." Branch personnel were given very aggressive quotas for the number of products they were expected to sell to clients, and they were under a huge amount of pressure. As a result, they ordered credit cards for clients without the clients' consent. They also created checking and savings accounts for clients who had not asked for them. In many cases, these activities led to clients paying additional fees to the bank. Quotas were sometimes even satisfied by enrolling homeless people in fee-accruing products, even though they would almost certainly fail to make payments as they became due.

Wells Fargo's dysfunctional sales culture was first documented by the *Wall Street Journal* in 2011 and further described in an article in the *Los Angeles Times* in 2013. Some reforms were then made, but the bank was fined $185 million in September 2016 for creating more than 1.5 million unauthorized deposit accounts and over 0.5 million unauthorized credit-card accounts between 2011 and 2016. Later estimates in May 2017 suggested that the number of unauthorized accounts was higher than this, closer to 3.5 million in total.

Senior executives claimed that they were unaware of what was going on as a result of the company's pressure-cooker sales culture. This seems unlikely given that employees made calls to the company's ethics hotline and complaints from customers had to be dealt with. The bank fired 5,300 employees between 2011 and 2016 as a result of the fraudulent sales practices. The CEO, John Stumpf, agreed to forgo $41 million in stock options and resigned in October 2016. Later, $28 million of his earnings were clawed back. Other executives also lost earnings. The bank paid $110 million to clients who had accounts opened without their permission. Investigations were commenced by the Senate Banking Committee, the Securities and Exchange Commission, and prosecutors in a number of jurisdictions.

Clearly, the aggressive sales culture at Wells Fargo was misguided. It alienated employees, many of whom were under unreasonable stress and found it difficult to find new jobs when they were fired. It also destroyed the trust and good reputation that Wells Fargo had built up with its clients over many years. If Wells Fargo's risk appetite had included a desire to avoid doing activities that could adversely affect its reputation, it would surely have reduced the sales pressure and emphasized the importance of developing good long-term relationships with clients.

Wells Fargo took steps to change its sales culture under a new CEO in late 2016. It stopped the practice of giving employees sales quotas. Some employees who had been fired were rehired. It will take time for the bank's relationships with employees and clients to recover, but Wells Fargo is a financially strong bank and we can expect that it will eventually bounce back.

led to employees engaging in questionable and sometimes fraudulent activities in order to meet the targets. The strategy should have been rejected (or at least monitored more carefully) because of its potential adverse effect on how customers perceived the bank and because of its adverse effect on the bank's relationship with its employees.

The quantification of risk has been a major theme running through this book. Up to now we have implicitly tended to take a bottom-up approach. We have looked at different ways risk measures can be calculated by business units and then considered how the risk measures can be combined to produce a risk measure for the whole enterprise. In discussing economic capital, for example, we considered how the different types of risk taken by different business units can be aggregated to produce an overall economic capital measure. The specification of a risk appetite framework is a top-down activity that should involve the board of directors. Risk appetite is first defined for the institution as a whole. It is then used to define risk limits for business lines. This can involve some experimentation where bottom-up calculations are used to see if the risk appetite of the organization is being met. For example, a financial institution wanting to preserve its AA rating might decide that the probability of losses exceeding $50 billion in one year must be 0.03% or less. It then needs to determine by trial and error what the VaR limits for its trading activities, its lending activities, and so on must be so that when aggregated they are consistent with this aspect of the risk appetite. Once risk limits have been set, it is important to monitor the decisions taken in the business units to ensure that the limits are being adhered to.

One important element of risk appetite is risk concentration. With the benefit of hindsight we can say that some large banks had portfolios that were too concentrated in subprime mortgages in 2006 and 2007. The risk appetite statement should directly or indirectly have implications for concentration risk. It should limit credit exposure to any one company, limit credit exposure to any one sector of the economy, limit credit exposure to movements in any one market variable, and so on.

Not surprisingly, regulators are interested in encouraging banks to develop risk appetite frameworks. Indeed, bank supervisors are required to ensure that effective risk appetite statements have been developed and implemented. In 2013, the Financial Stability Board published a document that contained definitions of relevant terminology and the principles on which a risk appetite framework should be based.[6] The document spells out the responsibilities of the board of directors, the chief executive officer, the chief risk officer, the chief financial officer, and business line managers.

27.2 Risk Culture

A financial institution's risk culture is all about how decisions are made. If the risk culture is good, all decisions are made in a disciplined way with a careful consideration of possible outcomes, weighing risks and rewards. It should be emphasized that this does not mean that no risks are taken. It means that risks are considered in relation to potential rewards and the risk appetite of the organization. The Financial Stability Board has suggested the way bank supervisors can assess risk culture at large systemically important banks.[7]

Decisions often involve trade-offs between short-term profits and medium- to long-term risks. A decision that is certain to lead to an immediate profit sometimes has the disadvantage that it is liable to give rise to serious problems later. Consider the situation of a salesman who works for a financial institution and is deciding whether to sell a sophisticated financial product to a customer when he knows it does not meet the customer's true needs. The sale would be profitable to the financial institution and would increase the year-end bonuses of the salesman and the group that developed the product. But it might have adverse future consequences. For example, the reputation of the financial institution could suffer if the customer incurs losses as a result of the product. Expensive lawsuits might follow.

The product described in Business Snapshot 5.4, which was sold by Bankers Trust to Procter & Gamble, is a classic example of what can go wrong. The product had a fairly high probability of saving Procter & Gamble 75 basis points on its funding and a small probability of costing the company a huge amount of money. The scenario where there is a huge cost to Procter & Gamble may or may not have been fully explained by the Bankers Trust salesman. But, if there had been a good risk culture within Bankers Trust, management would have questioned whether it made any sense for it to sell the product

[6] See Financial Stability Board, "FSB Principles for an Effective Risk Appetite Framework," Consultative Document, July 2013.

[7] See Financial Stability Board, "Guidance on Supervisory Interaction with Financial Institutions on Risk Culture: A Framework for Assessing Risk Culture," April 2014.

at all. Indeed, the quants working for Bankers Trust would have been told not to develop the product in the first place.

Procter & Gamble was one of many companies that entered into transactions similar to the one described in Business Snapshot 5.4 with Bankers Trust. As it happens, the low-probability outcome materialized and the transactions proved to be very costly for Bankers Trust's clients. There were numerous lawsuits, all of which were settled out of court. Bankers Trust, a company that had built an enviable reputation in financial markets, was ruined and later acquired by Deutsche Bank. Of course things could have worked out well and Bankers Trust's clients could have been grateful to the bank for saving them 75 basis points on their funding. But the risk to Bankers Trust and its reputation of huge client losses was clearly too great. The risk culture was one where employees thought only about short-term profits and not about the longer-term repercussions.

A more recent example is provided by the transaction between the Spanish bank Santander and the Portuguese state-backed rail operator Metro do Porto in 2007, described in Business Snapshot 27.2. The transaction allowed the rail operator to reduce its funding costs for the first two years, but this was at the risk of very high funding costs for later years. From Metro do Porto's perspective, the trade-off being made between short-term gain and long-term risk was highly questionable. But the same was true of Santander. This was one of many similar products sold by Santander to Portuguese state-owned entities. No doubt the products were viewed as profitable by Santander at the time it entered into them, but, as later settlements and lawsuits show, subsequent costs and the negative impact on Santander's reputation have the potential to outweigh the short-term profits.

Yet another example of the conflicts that can exist between short-term profits and long-term risks is provided by Business Snapshot 27.3. This describes a deal, ABACUS, that Goldman Sachs sold to one of its clients without, it is alleged, fully explaining the risks and how it had been created. This is another example of short-term profit being booked without regard for longer-term reputational risk and costs.

The ABACUS deal and one or two other events have led Goldman Sachs to embark on a program to change its corporate culture. Everyone at the top agreed that change was necessary if Goldman was to win back the allegiance of its corporate and investment clients. The CEO, Lloyd Blankfein, traveled to 23 countries during an 18-month period during which he stressed that ethics and reputation were as important as making money (i.e., short-term profits should not be taken if there are potential adverse long-term implications). Gary Cohn was appointed head of the Firmwide Client and Business Standards Committee. One thing this committee did was assess "whether our clients had the background experience and capacity to understand the range of outcomes from transactions they execute with us." The Firmwide Client and Business Standards Committee took steps to ensure that Goldman's bankers and traders were interfacing honestly with clients and avoiding any questionable activities. Consensus is that the reputation of Goldman Sachs has improved as a result of the changes.

BUSINESS SNAPSHOT 27.2

The Santander Rail Deal

In 2007, the Portuguese rail company Metro do Porto (MdP) was looking for a way to reduce its 4.76% per annum funding costs. After consulting a number of banks, it agreed to a 14-year swap with the Spanish bank Santander. In this swap, Santander paid the 4.76% funding cost and in return MdP paid 1.76% plus a spread. For the first two years, the spread was zero, saving MdP 3% on its funding. After that, the spread was calculated each quarter as

$$\max[0\%,\ \text{Previous Spread} + 2 \times \max(2\% - R,\ 0) + 2 \times \max(R - 6\%,\ 0) - D]$$

where R is the three-month Euribor rate and D, referred to as the DigiCoupon, is 0.5% if R lies between 2% and 6%, and zero otherwise.

At the start of the deal, three-month Euribor was about 4%. If it had stayed between 2% and 6% for the whole 14 years, MdP would have saved 3% per annum on its funding for the whole period. As it happened, three-month Euribor had dropped below 2% by 2009 when the deal was only two years old, and it stayed below 2% for several years. As a result, the spread increased rapidly, reaching 40.6% in September 2013. The problem for MdP was that the spread in one quarter is dependent on the spread in the previous quarter. Suppose that three-month Euribor is 0.5% from year 2 onward. MdP's spread would increase by 3% per quarter or 12% per year. If subsequently three-month Euribor increases so that it is in the 2% to 6% range, the spread declines, but only at the rate of 0.5% per quarter or 2% per year.

The deal was subject to a lawsuit. (Disputes over other similar deals entered into by Santander with other Portuguese state-owned entities were settled out of court.) MdP should of course have run scenario analyses to understand the potential costs of the transaction. Whether it did this and whether it understood the cumulative nature of the spread adjustments is not known. Possibly, it overweighted the attractiveness of saving 3% per year for the first two years. According to an article in *Risk* magazine in 2014, one London-based corporate hedging adviser has described this deal as a contender for the worst trade of all time!

In financial institutions, compensation plays a key role in determining the trade-offs between short-term profits and long-term risks. Traditionally, bonuses have been paid once a year to employees based on their performance during the year. This meant that the time horizon for decision making was usually the next bonus date (always less than a year). If a deal could lead to adverse consequences but not before the year end, the

BUSINESS SNAPSHOT 27.3

ABACUS

In 2007, Goldman Sachs created a product called ABACUS from a synthetic ABS CDO. As explained in Chapter 6, an ABS CDO is formed from a portfolio of tranches (often mezzanine tranches) of ABSs. The word *synthetic* means that Goldman defined the ABS CDO, but it did not actually buy the underlying tranches. It then acted as an intermediary between two parties. Party A bought a credit default swap providing protection against what the losses on the senior tranche of the ABS CDO would have been if it had been created. Party B sold the credit default swap, agreeing to provide the protection. Party A paid Party B an insurance premium (the CDS spread).

So far, so good. This is what investment banks such as Goldman Sachs do all the time. They act as intermediaries between buyers and sellers. In this particular case, Party A was a hedge fund controlled by John Paulson. This was one of the few hedge funds to make big bets against the performance of subprime mortgages and the housing market generally. It transpired that Paulson paid Goldman Sachs $15 million for creating the structure and that his hedge fund had a hand in choosing the ABSs that went into the structure. Party B was IKB Deutsche Industriebank AG (IKB), a bank based in Dusseldorf, Germany, that specializes in lending to small and medium-size companies. The portfolio manager was ACA, which at the time had a good reputation for managing ABS CDOs.

It is alleged that Goldman Sachs represented to IKB that the ABS CDO had been defined by ACA when the ABSs in the portfolio had in fact been chosen by John Paulson's hedge fund as ones that were likely to incur huge losses. Specifically, it is alleged that Paulson constructed the synthetic ABS CDO with tranches formed from portfolios of mortgages where the borrowers had low FICO scores and came from states where the rate of house appreciation had been the highest.

Both IKB and ACA took huge losses on ABACUS, and Paulson's hedge fund made a huge profit. (IKB had to be bailed out because of the massive losses it sustained from the subprime meltdown.) The deal was investigated by the Securities and Exchange Commission (SEC), and Goldman Sachs ended up paying a fine of $550 million, the largest in the history of the SEC at that time. Goldman Sachs's position was hurt when it was revealed that an executive working on the deal, Fabrice Tourre, had sent the following e-mail to a friend: "More and more leverage in the system. The whole building is about to collapse anytime now.... Only potential survivor, the fabulous Fab ... standing in the middle of all these complex, highly leveraged, exotic trades he created without necessarily understanding all the implications of those monstrosities!!!"

possible adverse consequences were likely to be ignored. There has been an attempt in many financial institutions to change compensation arrangements to address this short-term focus. The payment of the bonus for a year is now often spread out over several subsequent years, and bonuses can be clawed back (i.e., not paid) if subsequent events indicate that the year was not as profitable as thought.[8]

27.3 Identifying Major Risks

An important part of ERM is identifying key exposures for both a financial institution's existing business and any proposed strategic investments. The identification of these exposures is in many ways similar to the choice of scenarios for stress testing. History can provide some guidelines. For example, a financial institution should consider how it would fare if a recession as bad as that experienced during 2007 to 2009 were to occur. The impact of current and emerging trends should also be carefully analyzed. Economists and senior management working for the financial institution can provide useful insights. As explained in earlier chapters, the environment is changing for banks. One of the results of the credit crisis that started in 2007 is that banks are becoming more tightly regulated. Some activities such as the trading of derivatives are less profitable than before, and other activities such as proprietary trading can (at least in the United States) no longer be undertaken at all. Capital requirements and liquidity requirements are higher. Some financial intermediation business is moving from banks to shadow banks such as money market funds, mortgage companies, securitization vehicles, and so on. In many cases, shadow banks are less heavily regulated than banks and, as a result, can offer more competitive services.

27.3.1 Cognitive Biases

Our ability to identify risks effectively is affected by what are termed *cognitive biases*. These describe the tendencies for human beings to think in certain ways and be less than perfectly rational. One common cognitive bias is wishful thinking. It is sometimes difficult to distinguish between what we want to happen (e.g., a project to be a success) and what we think will happen. (Try asking a Manchester United supporter to estimate the chance of Manchester United winning the FA cup next year!) When we want something to happen, we are liable to think only of reasons why it will happen.

Over 100 cognitive biases have been listed by psychologists. Much of the pioneering work was done by Daniel Kahneman and Amos Tversky.[9] Kahneman won the Nobel

[8] Legal issues can make it difficult for financial institutions to try to recover bonuses that have already been paid. Hence, clawback schemes have to be combined with bonus deferral.

[9] See, for example, D. Kahneman, P. Slovic, and A. Tversky, *Judgment Under Uncertainty: Heuristics and Biases* (New York: Cambridge University Press, 1982).

prize for economics in 2002 for his work with Tversky on prospect theory, which is concerned with the way people choose between risky alternatives. (Tversky had died a few years earlier.)

One important bias is *anchoring*. When evaluating a potential outcome (e.g., the revenue resulting from a major new venture), we are liable to anchor onto the first estimate that is made. We tend to make relatively small adjustments to that estimate (this is referred to as "anchoring and adjustment") and often never consider the full range of possible outcomes. In particular, important adverse outcomes may implicitly be considered to have no chance of occurring. To illustrate anchoring, one could ask a group of people to make a best guess of something that is unknown to them such as, perhaps the population of Iceland. They can then be asked to provide a range consisting of the 5th percentile to 95th percentile of their subjective probability distribution. If their estimates are good, the true population of Iceland should lie outside the range only 10% of the time. In practice, it is found that this happens much more frequently. Anchoring causes people to behave as though they know more than they do.

Another cognitive bias is *availability*. This is where recent information is given undue weight. Sadly, enterprise risk management—and risk management generally—can suffer from availability. Prior to the credit crisis, risk managers in some financial institutions were often not listened to because recent experience had been good. After the crisis, risk managers have had more influence, but as memories of the crisis fade, the "good times will last forever" attitude may return.

Another cognitive bias is known as *representativeness*. This is where an individual categorizes a situation based on a pattern of previous experiences or beliefs about the underlying scenario. It can be useful when trying to make a quick decision, but it can also be limiting because it leads to close-mindedness and stereotyping. Based on previous experience, a senior manager at a financial institution might consider it almost impossible for any other financial institution to successfully compete with it in a particular market. However, if the manager's past experience is limited, the previous situations might not be representative of future scenarios.

A more subtle mistake sometimes made in estimating probabilities is inverting the conditionality. Suppose that 1 in 10,000 people have a particular disease. A test that is 99% accurate gives you a positive result (suggesting that you have the disease). What is the chance that you have the disease? Your immediate response is likely to be 99%. However, the true answer is actually about 1%! We are interested in

$$\text{Prob}(D|TP)$$

where D indicates that you have the disease and TP indicates a positive result from the test.

It is true that

$$\text{Prob}(TP|D) = 0.99$$

but

$$\text{Prob}(\text{D}|\text{TP}) \neq \text{Prob}(\text{TP}|\text{D})$$

Out of 10,000 people there will be about 100 positive results on average but only one person with the disease.[10] Hence the probability we are interested in is about 0.01. This is an application of the result in probability theory known as Bayes' theorem.

Yet another bias is the *sunk costs bias*. Suppose that a financial institution has already spent $1 billion trying to enter a new market. Things are not going well and there seems to be very little prospect of success. Should the $1 billion influence the financial institution's decision making? The answer is the $1 billion is what accountants refer to as a sunk cost. Regardless of the decisions taken now, it cannot be recovered. The key issue is whether future profits will be sufficiently high to justify future expenditures. In practice, many people are reluctant to admit mistakes, and they continue with projects that are clearly failures for too long. Irrationally, they want to try to get back money already spent, even when the chance of this is very small.

Understanding these biases may assist decision making and the identification of key risks. It should be noted, however, that experiments have shown that it is extremely difficult to eliminate biases. Even when biases such as anchoring are carefully explained and subjects are given financial incentives to make good estimates, the biases persist.

The challenge for ERM is identifying tail risks and trying to estimate the probabilities associated with the adverse scenarios giving rise to the tail risks as well as possible. The cognitive biases we have discussed (and many others have been documented) suggest that the risks will be underestimated. Nassim Taleb makes this point in a best-selling book.[11] He is particularly critical of the use of normal distributions for calculating risk measures (as we have been in this book) and argues that extreme events such as the crash of 1987 or the credit crisis of 2007 to 2009 are more likely than many people think.

Using committees of senior management and economists to define adverse scenarios can be useful. Encouraging employees to play the role of devil's advocate when ideas are discussed can also be important. Risks are obviously less likely to be considered if there is an authoritarian CEO who is convinced that he or she is right on major strategic decisions and does not encourage contrary views to be expressed.

27.4 Strategic Risk Management

Once a company's strategy has been defined, strategic risk management involves considering what assumptions have been made and what can go wrong. What will competitors

[10] If there are 100 positive results, the test is correct for 9,901 people out of 10,000 so that the accuracy is about 99%.

[11] See N. N. Taleb, *The Black Swan: The Impact of the Highly Improbable* (New York: Random House, 2007).

do? How will customers react? How could the strategy be impacted by adverse external events? To quote Michael Porter:[12] "Risk is a function of how poorly a strategy will perform if the 'wrong' scenario occurs."

Once a particular risk has been identified, either in the existing business or in a new strategic investment, there are a number of alternative courses of action:

Avoid: Exit the business/investment.
Reduce: Take action to reduce the probability of the adverse event or its impact.
Adjust: Modify plans to reduce risks.
Share or insure: Transfer or share a portion of the risk.
Accept: Take no action.

Which of these actions is taken is likely to depend on the situation, but it is important that the risk is identified and carefully considered, rather than being ignored.

Suppose that a successful Canadian bank is considering an expansion of its retail operations by buying a smaller bank in the United States. It has prepared plans indicating that the expansion will increase shareholder value and provide a springboard for increasing its customer base in the United States. There are a number of risks that it needs to consider. Will it be able to transfer its success in Canada to a more competitive banking environment? Will it be able to retain the customers of the bank it is acquiring in the United States? Will it be able to retain key managers? What will happen if there is a severe two-year recession soon after the acquisition is completed? The Canadian bank needs to make sure the acquisition is consistent with its risk appetite. The risk appetite might require it to ensure that if a worst-case scenario happens it will survive without its core Canadian business being appreciably affected. It needs to ensure that the risks can be managed and monitored and that if things work out badly there is a good exit strategy.

Summary

Enterprise risk management is an attempt to take a holistic approach to risk management rather than addressing risk in silos. The central idea is that business units within the financial institution should be part of an integrated, strategic, and enterprise-wide risk management system. The risk appetite for the organization should be defined at the top and approved by the board. Steps should then be taken to ensure that the management of different types of risks in different business units is consistent with the overall risk appetite framework.

A financial institution's risk appetite statement should define how much risk the organization is prepared to take in pursuit of its objectives. Some aspects of the risk appetite, such as the amount of capital the financial institution is prepared to risk losing, can be expressed quantitatively. Others, such as risks to reputation, are likely to be

[12]See M. E. Porter, *Competitive Advantage* (New York: Free Press, 1985), 476.

expressed qualitatively. The risk appetite must then be converted into directives for the risks taken by business units. Appropriate limits are then set on trading risk, how concentrated the financial institution's portfolio can become, how much credit risk can be taken, and so on. The risks being taken should be continually monitored to ensure that they are consistent with the financial institution's risk appetite.

Developing a good risk culture is an important part of ERM. A key element of the risk culture is ensuring that medium- to long-term risks are considered when opportunities that offer short-term profits are evaluated. There are many examples of financial institutions aggressively pursuing short-term profits when the activity being undertaken has led several years later to serious problems such as lawsuits, fines, or a loss of reputation. The bonus structure within financial institutions has in the past encouraged employees to use the maximization of short-term profits as their sole criterion for decision making. Since the 2008 credit crisis, there have been attempts to change this culture by introducing deferred bonuses and clawbacks.

It is important for a financial institution to develop procedures for identifying worst-case outcomes. Unfortunately, there are many biases that are liable to cause managers to misjudge future outcomes and their probabilities of occurring. Once the risks being taken have been identified, it is important to actively manage the risks. Sometimes it will be necessary to discontinue certain activities; sometimes risks can be shared (e.g., with insurance or joint ventures); sometimes steps can be taken to reduce the impact of the risks; and sometimes risks can be accommodated within the organization's risk appetite.

Further Reading

Fraser, J., and B. J. Simkins. *Enterprise Risk Management: Today's Leading Research and Best Practices for Tomorrow's Executives*. Hoboken, NJ: John Wiley & Sons, 2010.

Los Angeles Times. "Wells Fargo's Pressure-Cooker Sales Culture Comes at a Cost," December 21, 2013.

Osborn, T. "Worst Trade of All Time Pits Santander Against Portuguese Client." *Risk* (May 2014).

Porter, M. E. *Competitive Advantage*. New York: Free Press, 1985.

Practice Questions and Problems (Answers at End of Book)

27.1 Explain the difference between top-down and bottom-up approaches to risk management. Why are both necessary in ERM?

27.2 What is the Bowman paradox?

27.3 "The risk culture at a bank should lead it to consider the longer-term implications of decisions as well as short-term profits." Discuss.

27.4 What steps can a bank take to encourage employees to consider more than short-term profits?

27.5 What distinguishes enterprise risk management from more traditional approaches to risk management?

27.6 A fund's risk appetite is such that it wants to be 90% certain that it will not lose more than 20% in any one year. Using the performance of the S&P 500 between 1970 and 2016 (see Table 27.2), determine the beta the fund should have. Assume a risk-free rate of 3% per annum.

27.7 Calculate the interest rate paid by MdP six years after the beginning of the deal in Business Snapshot 27.2 if the Euribor rate proves to be 8% from year 2 onward.

27.8 Give three examples of cognitive biases.

Further Questions

27.9 Assume that in Business Snapshot 27.2 the change in the three-month Euribor rate in each quarter is normally distributed with mean zero and a standard deviation equal to x basis points. Use Monte Carlo simulation (500 trials) to calculate a probability distribution for the average interest rate paid by MdP over the four-year period for values of x equal to 10, 20, and 50.

27.10 A fund's risk appetite is such that it wants to be 97.5% certain it will not lose more than 25% in any one year. Using the performance of the S&P 500 between 1997 and 2016 (see Table 27.2), determine the beta the fund should have. Assume a risk-free rate of 2.5% per annum.

Chapter 28

Financial Innovation

In the 1980s, when booking a flight from New York to London, a customer would contact a travel agent, who would check availability, present options, receive instructions, and then finally make a reservation. Today, that customer is likely to go to the airline's website and make the booking directly with the airline. What has happened is a process known as *disintermediation*. The travel agent as an intermediary is no longer needed.

This does not mean that there is no need at all for intermediaries in the travel business. Online services such as Expedia and Travelocity have sprung up to be of assistance when customers do not know which airline to use or want to quickly compare prices offered by different airlines. But the nature of the intermediaries in the travel business has changed, and the human involvement in the process of making straightforward flight arrangements has largely disappeared. The creation of new technology-based intermediaries like Expedia and Travelocity is referred to as *reintermediation*. Disintermediation followed by reintermediation is a common pattern in technological change.

Banks and other financial companies are intermediaries, like the travel agents of the 1980s, and they are similarly in danger of having the services they provide disrupted by technology. A commercial bank is an intermediary that channels money from depositors to borrowers. It also is an intermediary in the payments system, clearing checks and transferring funds as instructed by clients. Investment banks are intermediaries that facilitate the issue of stocks and bonds. Banks and other financial companies are intermediaries when they perform wealth management services for clients, channeling their savings to appropriate investment products.

FinTech is concerned with the application of technology to financial services. Many applications of technology such as ATMs, online banking, and online trading have already taken place. This chapter looks at how financial services are likely to be impacted by FinTech in the future, which activities are most likely to be digitized, how customers will respond, and how financial services companies can restructure themselves to adjust to rapid technological changes. These are important questions for today's risk management professionals.

28.1 Technological Advances

The increase in computer processing power in the first 50 years of the computer age has been truly astounding, and this has made possible many societal advances. In 1965 Gordon Moore, co-founder of Intel, correctly predicted that computer processing power would double every two years. A single smartphone now has more processing power than NASA had when it landed a man on the moon in 1969!

Many activities formerly carried out by humans have been automated. An interesting development is *robotics process automation* (RPA). This is a software application that replicates the actions of human beings when they interact with a company's systems. It is an alternative to automating by changing the systems themselves. The software robot is a virtual worker that can be trained to perform certain tasks in much the same way as a human would be trained. The tasks could for example be concerned with opening a new customer account or hiring a new employee. Typically, data are sent to a number of different systems within the organization.

28.1.1 Machine Learning

One important result of the increase in computing power has been the growth of *machine learning*. This is a branch of artificial intelligence that allows computers to learn without being explicitly programmed.

Suppose we wish to teach a computer to translate between English and Japanese. One approach is to program all the grammatical rules of English and definitions contained in an English dictionary, do the same for Japanese, and then supply a Japanese–English dictionary. An alternative, pioneered by Google, that has proved to be much more successful is to use machine learning. The computer is fed many texts in English with Japanese translations produced by bilingual human beings. The computer then uses various tools to search for patterns and develop its own complex translation rules.

Machine learning has many potential applications in finance, such as fraud recognition and the automation of lending decisions. There are three types of machine learning:

1. *Supervised learning:* The computer is presented with data consisting of inputs and desired outputs, and the goal is to develop rules for mapping inputs to outputs. The translation example we have just given falls into this category.
2. *Unsupervised learning:* The computer is asked to learn by finding structure or hidden patterns in data.

3. *Reinforcement learning:* A computer program interacts with a dynamic environment in which it must perform a certain goal, such as driving a vehicle.

There are many different statistical and other tools used in machine learning to reveal patterns in data. Supervised learning can involve linear regression and logistic regression. Unsupervised learning can involve clustering techniques such as the k-means algorithm. Sometimes dimensionality reduction algorithms such as principal components analysis are used before processing data.

Human beings are very good at some pattern recognition tasks that computers have traditionally found difficult, for example, distinguishing one person's face from another. A study of how we do this has led to some important breakthroughs in machine learning. *Neural networks* are tools designed to mimic the way humans recognize patterns. They rely on a mathematical procedure known as backpropagation to iteratively calculate weights on nodes to produce a desired output.

A large amount of data is necessary for a machine learning algorithm to train on. Without parallel processing, cloud computing, and other advancements in hardware and software, today's applications would not be possible.

28.1.2 Blockchain

Often ledgers recording the ownership of assets involve the possibility of fraud, lack of trust by users, and computer hacking. Many readers will have heard stories about title fraud where someone who is not the owner of a house manages to claim ownership and obtain a mortgage. Sometimes ledgers recording the same items are kept by different people from different organizations. There may then be inconsistencies in the records, and it can be expensive and time-consuming to sort these out.

Blockchain is a distributed ledger where many people can add entries and a community of users control how it is updated. Problems involving inconsistent records, fraud, or hacking are either impossible or extremely unlikely. Identical records are on many different computers and can be inspected at any given time. A blockchain was first developed in 2009 in connection with bitcoin, but many other applications of the technology are now being explored. The ledger defining who owns what is a long list of blocks. When a new block of transactions is created (every 10 minutes in the case of bitcoin), a new block is added to the chain. A constantly updated list of blocks is given to everyone who participates.

An important aspect of blockchain is what is known as hashing. A hash is a string of 64 characters. Any text can be converted into a hash, but the process cannot be reversed so that the hash is converted back into text. Even a small change in the text will totally change the hash. To illustrate this, we can go to the SHA 256 hash calculator at:

http://www.xorbin.com/tools/sha256-hash-calculator

Hashing the title of this book, *Risk Management and Financial Institutions* (without italics), gives

1dcc48387a27cd95378b08ab26261b161a97c51a7c9146f3d3ff73710d656a3f

Adding the edition number so that *Risk Management and Financial Institutions 5* is input (again without italics), produces a totally different hash:

117e1e23121f8db4b75d3e2a63d37ef052b11a63cf448721825aadb882492c6b

It is theoretically possible for two sets of records to give the same hash, but in practice this is to all intents and purposes impossible. The hash we have just looked at is formed from the numerals 0 to 9 and the first six letters of the alphabet. This means that there are $16^{64} - 1$ different hashes.[1]

The records in each block of a blockchain are sealed and made tamper proof with a hash. One of the inputs to the hash is the hash of the immediately preceding block. This means that any attempt to tamper with a block will be immediately spotted because it will affect that block and all subsequent blocks.

In order for the blockchain to be free from human intervention, participants must be given an incentive to check new blocks and hash them. Bitcoin does this by inviting participants, referred to as *miners*, to search for something called a *nonce*, which when added to the block will lead to a hash that starts with several zeros. The participant who finds the nonce first is rewarded with a predetermined number of bitcoins. The new block is then hashed and distributed to the computers of all participants.

There are two sorts of blockchains: permissioned and permissionless. Permissionless blockchains such as bitcoin have no restrictions on participation and there is no need for participants to know or trust one another. Permissioned blockchains are privately shared systems, and the participants may know one another and may (at least to some extent) trust one another. Permissioned blockchains are a way financial institutions can share data to speed the settlement of transactions.

Distributed ledger technologies (DLTs) such as blockchain are now receiving a great deal of attention by financial institutions and the big four accounting firms, as well as many other large businesses. In August 2016, the World Economic Forum reported six developments:[2]

1. There had been $1.4 billion of venture capital investments in DLT since 2013.
2. More than 2,500 patents had been filed since 2013, many by financial institutions.
3. More than 24 countries were currently investing in DLT.
4. More than 90 central banks were engaged in DLT discussions.
5. More than 90 corporations had joined DLT consortia.
6. Some 80% of banks were predicted to initiate a DLT project by 2017.

Whether blockchain and other DLT ideas live up to the hype remains to be seen. IBM in 2017 announced it was working with seven European banks to use blockchain

[1] This is approximately equal to the number of atoms in the known observable universe. (In our count, we exclude a hash consisting of 64 zeros.)

[2] See World Economic Forum, *The Future of Financial Infrastructure: An Ambitious Look at How Blockchain Can Reshape Financial Services*, August 2016.

technology to make it easier for small and medium-size businesses to trade with one another. Other potential applications in finance are to the trading of securities, settlement, clearing, and other back-office functions. DLTs have the potential to speed up these processes. Indeed, almost real-time settlement of trades becomes a possibility. Costs should be reduced and there should be less need for human intervention. Operational risks should be reduced because the possibility of fraud or cyber-attacks should be lower. This may in turn lead to a reduction in regulatory capital.

28.2 Payment Systems

Technology has already had a huge effect on the way payments are made. As a society we have moved from cash and checks to credit and debit cards to the use of mobile wallets. Developing countries have in some respects progressed further in this direction than developed countries, in part because traditional payment systems were not as well established. Many FinTech start-ups are offering new services. Some, such as PayPal, Apple Pay, and Google Wallet in the United States, and Alipay (part of Alibaba) and Tencent in China, are large, well-established companies.

The key attributes of a payments method are speed, convenience, security, simplicity, and cost. Many services such as PayPal transfer funds almost immediately. Storing credit-card-like information in an iPhone or similar device adds to the convenience for many consumers. Indeed, given the dominant position of Apple in the smartphone market, it has been natural for Apple to expand into payments, and some have speculated that it will not be long before Apple offers a wider range of banking services. One can imagine that wearables such as watches or bracelets, or even implants, may be used in the same way as smartphones to add to the convenience of the way payments are made.

Security is a big issue for all forms of payment. Tens of billions of dollars are lost each year from credit card fraud. Embedding chips as well as magnetic strips in cards helps but does not eliminate this problem. We are likely to see big changes in the way fraud is avoided in the future. Digital wallets are considered to be more secure than credit cards. Many payment providers, including banks, are investigating the use of biometric authorization. (Biometrics often involves machine learning, which has already been discussed.) In the future, it may be the case that a purchase can be completed by waving the palm of one's hand over a reader. Retinal scanning, facial recognition, voice authentication, and even heartbeat monitoring are all being considered. The costs of fraud are usually borne by the payments system provider, and they are of course passed on to merchants in the form of the fee charged. (PayPal, for example, charged 2.9% plus $0.30 per transaction in mid-2017.) The merchants in turn pass the fee on to consumers. Everyone therefore has an interest in reducing fraud, and the approaches we have just discussed for making more secure payments should be welcomed.

If biometric authorization seems far-fetched, bear in mind that India has issued biometric IDs (involving fingerprints and retinal scans) to over one billion people and is laying the groundwork for a cashless society. Making more people part of the financial ecosystem is an important objective for development of the country. The biometric IDs also have the advantage that some government benefits can be distributed with less involvement from intermediaries. Of course, some would argue that the provision of biometric information is an unacceptable violation of a person's privacy, and this may slow down its acceptance in developed countries.

Some payment systems allow users to borrow money. The interest rates charged by credit card companies are very high, but it should be remembered that users do get free credit for the period of time between a purchase and the next monthly due date. (PayPal competes with this by offering 14 days of free credit.) It is likely that more convenient credit facilities, tailored to the needs of users, will be developed. Alibaba through its subsidiaries such as Alipay and Mybank is already offering many of the same services as banks.

What other services can be offered to make payment systems attractive? Many individuals remit funds on a regular basis to family members in another country. The foreign exchange services associated with those transactions are likely to get more convenient and more competitive. For businesses, easy-to-use foreign exchange hedging services that compete with those offered by banks are likely to be developed. FinTechs may carry out sophisticated analyses of sales to help a company understand its customers better or provide accounting services.

Of course, people who prefer cash will continue to exist for some time. Some people have bad credit histories and do not qualify for credit cards, while others are too risk averse to give their credit card information to third parties. Amazon has recognized this and allows customers to open an account at selected retailers by depositing cash. When goods are purchased the account is debited.

One aspect of the digitization of payments is that it becomes much easier to collect data on a person's spending habits. This could be useful to banks when making credit decisions. (Knowing how a potential borrower spends money can be almost as important as knowing how much he or she earns.) Of course it presents huge privacy issues. Sometimes payment systems sell anonymized data. Often the data are analyzed using techniques such as machine learning to better understand consumer preferences.

28.2.1 Cryptocurrencies

The use of bitcoin as a means of payment has grown rapidly since it was first introduced in 2009. The number of bitcoin transactions per day was 60,000, 120,000, and 220,000 in August 2014, August 2015, and August 2016, respectively. In the future, can we expect dollars, euros, and yen to be replaced by bitcoin or other cryptocurrencies? Many people are bullish on bitcoin, as indicated by the 2017 growth of its value when expressed in U.S. dollars (see Figure 28.1).

The advantage of using a cryptocurrency such as bitcoin as a means of payment is that it is outside the control of any one government and therefore its value cannot be

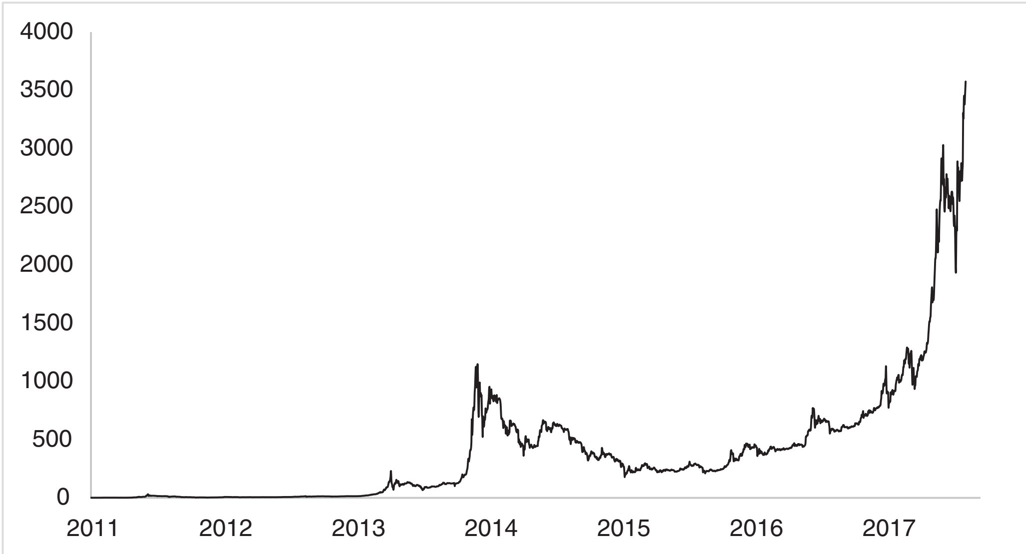


Figure 28.1 Value of Bitcoin in U.S. Dollars, January 2011 to August 2017

debased by a government's money supply policies.[3] This can be particularly attractive for individuals living in developing countries such as Venezuela or Zimbabwe where the value of the domestic currency is highly volatile. Payments are secure, private, and very low cost. An individual can remain anonymous while using a cryptocurrency.

The disadvantage of a cryptocurrency is that its value is more volatile than other fiat (traditional) currencies.[4] It pays no interest. It is often associated (fairly or unfairly) with illegal activities such as drug dealing. Jamie Dimon, CEO of JPMorgan, said in September 2017, "Bitcoin is a fraud that will ultimately blow up." China introduced regulations concerning cryptocurrencies in 2017 and other governments may do so as well. Users of cryptocurrencies are reliant on the way a cryptocurrency has been set up. If a flaw were found, the currency would be worthless. The bitcoin protocol has not been successfully hacked, but bitcoin exchanges have been. (See Business Snapshot 28.1.)

There are now many cryptocurrency alternatives to bitcoin. Examples are litecoin, ripple, dogecoin, nxt, monero, ethereum, and zcash. Some have procedures for making transactions secure that are quite different from those used by bitcoin. Some cryptocurrencies such as ethereum facilitate *smart contracts*. These are contracts that execute automatically when all conditions are met, thereby eliminating the need for intermediaries.

[3] The increase in the supply of bitcoin is determined by the amounts paid to miners (participants) when they verify transactions, as described in Section 28.1.2.

[4] The volatility of the U.S. dollar value of bitcoin calculated from the data between August 2006 and August 2007 is about 57% per year. The volatility of a fiat currency such as the euro, British pound, or Japanese yen when expressed in terms of the U.S. dollar is typically less than 20% per year.

BUSINESS SNAPSHOT 28.1

Mt. Gox

Bitcoin is a secure means of payment, but some bitcoin exchanges, which allow individuals to trade bitcoins for conventional (fiat) money, have had problems. There have been a number of reports of exchanges being hacked. The most high-profile one concerns Mt. Gox, which is based in the Shibuya district of Tokyo. In February 2014, it announced that bitcoins worth more than \$450 million were missing and likely stolen. The reason for the loss, which happened over a number of years, is unclear. It appears that bad management and poorly controlled computer code allowed hackers to steal bitcoins. Some have blamed the CEO for embezzling the funds, but this has been denied. Whatever the reason, Mt. Gox filed for bankruptcy in April 2014. This damaged the image of cryptocurrencies, but, as Figure 28.1 indicates, the value of bitcoin recovered and then improved dramatically in 2017.

How can you avoid these types of losses? Bitcoins should not be stored on an exchange. They should be transferred to a trustworthy online exchange only when you want to trade them.

28.2.2 Central Banks

Central banks are researching the use of cryptocurrencies. There may be an element of "if you can't beat them, join them" about this. But a digital currency issued by a central bank could make it easier for the central bank to control the money supply. In the current system, banks create money. (A bank lends money to an entity; that entity puts the money on deposit with a bank or uses it to buy something from another entity that puts it on deposit with a bank; the bank receiving the funds relends it; and so on.) In a digital currency regime, the total amount of currency could be fixed unless the central bank makes an explicit decision to increase it. There could also be less need to guarantee bank deposits through programs like those offered by the Federal Deposit Insurance Corporation. The payment systems in countries such as the United States that have many small banks could be simplified. Another advantage, particularly in developing countries, could be financial inclusion. Individuals who, for whatever reason, do not have bank accounts could become part of the financial ecosystem.

One can imagine a country where all citizens have a digital currency account at the central bank and private companies compete to provide payment services. The central bank can increase the money supply by adding to the holdings of all citizens in an equitable way. (It might do this to increase aggregate demand.)

Some of the implications of all this are of course uncertain. Replacing banknotes with digital currencies would make it much easier for a resident of one country to transact in the digital currency of another country. This could lead to a new form of competition between the currencies of different countries and could actually end up making it more difficult for any one country to manage its economy by controlling the money supply.

28.3 Lending

As all students of banking know, there are a number of *C*s associated with lending: character, capacity, capital, collateral, and so on. Can technology automate lending decisions? Some banks think it can. Loan officers are being replaced by systems involving machine learning in some large banks. Given enough data about a bank's lending experience, it is considered that a machine learning algorithm can sort good loans from bad loans as well as or better than a human being can. In principle, a machine learning program can be more objective and exhibit less bias than a human being.[5]

28.3.1 P2P Lending

A bank's net interest margin is determined by the difference between the interest it pays on liabilities and the interest it earns on assets. If it obtains deposits on which it is paying 2% per annum and then lends the funds out at 5% per annum, its net interest margin is 3%. Some of this (perhaps as much as 1%) is to compensate for expected loan losses, but if the depositor could lend money directly to the borrower at, say, 4%, both sides could potentially be better off. Peer-to-peer (P2P) lending involves disintermediation followed by reintermediation. Banks are no longer intermediaries, but new intermediaries are set up to provide services such as:

1. Providing online platforms for borrowers and lenders to contact each other.
2. Verifying the borrower's identity, bank account, employment, income, and so on.
3. Assessing the borrower's credit risk and, if the borrower is approved, determining the appropriate interest rate.
4. Processing cash flows between borrowers and lenders.
5. Attempting to collect payments from borrowers who are in default.

Many borrowers who currently use P2P platforms have already been refused by banks, so the interest rates can be quite high compared with conventional loans (but lower than the rates on credit card balances and other sources of credit for moderate- to high-risk borrowers). Two well-known P2P lending platforms in the United States are Prosper and Lending Club, both based in San Francisco.[6] The P2P lending platform assigns a

[5] But it has been shown that machine learning is not completely free from bias. The data that are used by the machine learning algorithm may incorporate biases. The machine learning programmer may consciously or unconsciously incorporate biases into the way algorithms are used and trade-offs are made.

[6] For regulatory reasons, the P2P company enters into an arrangement with a bank so that the loan is passed through the bank and the lender obtains a security in the loan.

Table 28.1 Published Interest Rates Paid by Borrowers and Returns Received by Lenders on Lending Club Loans, June 30, 2017

Grade	A	B	C	D	E	F + G
Average interest rate (%)	7.26	10.79	14.01	17.20	19.94	24.04
Average net annual return (%)	4.86	6.29	6.69	6.32	5.59	4.52

credit rating to a borrower in much the same way that a bank does. Lending Club, for example, categorizes borrowers by assigning a letter grade between A and G. The interest rate charged to the less creditworthy borrowers is higher than to A-grade borrowers, but the expected loss from defaults is also higher. Statistics published by Lending Club in June 2017 are summarized in Table 28.1. This shows that both interest rates and loan losses are higher than on most loans made by banks. However, the net annual returns that investors receive on average are quite attractive compared with other opportunities.

The fees at P2P lenders can be quite high. At Lending Club, for example, the borrower pays an origination fee typically between 1% and 5% of the amount borrowed. The lender pays a service fee (typically about 1%) on payments received and may also have to pay costs associated with collections on delinquent accounts.

It clearly makes sense for lenders to be diversified. Typically, they can either choose a portfolio of loans themselves or ask the P2P lending platform to construct a portfolio that is consistent with their risk preferences. The risk of borrowers is often idiosyncratic so that correlations between the performances of different loans are low, making diversification easier.

Some lending platforms can be criticized because they have no skin in the game. If loans do not perform as well as expected, the lender bears the entire cost. A exception is Upstart (started by former Google employees in 2014). It has a different model from Lending Club and Prosper. It charges borrowers an origination fee but does not charge lenders a fee. It uses the origination fee to reimburse lenders if a loan defaults, giving it a stake in the performance of the loan. Its credit assessments, which look at the borrower's educational background as well as FICO scores and other data, have proved to be quite accurate and it has grown fast.

P2P lending has not been immune to scandal. The founder of Lending Club (which used an IPO to become a public company in 2014) had to step down in 2016 as a result of a governance scandal (but the company seems to have bounced back). In China, retail investors have lost billions of dollars in incidents where P2P platform operators have simply disappeared with investors' cash. (This has led to a crackdown on the industry by Chinese regulators.) All financial innovations are liable to have teething troubles of this sort. Indeed, banks over their long history have had their fair share of scandals. The real question for P2P lending is whether it will make inroads into traditional bank lending. Will P2P lending become a widely used option for financing the purchase of cars and houses? Will P2P between corporations become more common? Because P2P platforms are relatively new, it will be interesting to see how they perform in an economic downturn or when interest rates increase.

Some P2P platforms allow lenders to sell a loan before it has been fully repaid, but this depends on the willingness of another client of the P2P lending platform to take over the loan. There is likely to be a natural assumption that if a lender wants to sell a loan there is something wrong with it.[7] However, it may just be that the lender has run into liquidity problems so that he or she needs cash. Creating a better secondary market for loans is a way the market can evolve. An idea that can be used for increasing liquidity is to create tranches from portfolios of loans as described in Chapter 6 and sell securities created from them to lenders.

28.3.2 Crowdfunding

Crowdfunding is another activity that has the potential to be a disruptive force in finance. It involves obtaining funds for a project from a large number of different people. It is estimated that over $34 billion was raised in this way in 2015. Some crowdfunding projects are donation-based with the purpose of funding projects that donors identify with and want to see funded. Kickstarter is a well-known platform for handling such projects.

Other crowdfunding platforms are designed to provide a return to investors. Funding Circle is a successful British crowdfunding platform where investors lend to small businesses. Companisto is a crowdfunding platform designed to provide funds for start-ups and high-growth companies. Investors receive either a fixed rate of interest or an equity investment. SyndicateRoom provides equity investments in early-stage businesses.

Crowdfunding has opened up equity investing in private companies to a much wider range of investors than venture capitalists and angel investors. The valuation of the company is a critical part of equity crowdfunding. If a company is valued at $500,000, then shares representing 1% of the company should require an investment of $5,000; if it is valued at $1 million, the required investment is $10,000. Sometimes the company determines the value and sometimes investors do so. In the latter case, some platforms use experienced professionals to carry out due diligence and ensure that an investment is properly priced. Ideally, the professional is an angel investor who has some skin in the game. Of course, he or she will typically require to be compensated in some way, and it should be recognized that, for even an experienced professional, valuations are highly subjective.[8]

The growth of equity crowdfunding has been slowed by regulation. Most countries have strict rules on the issuance of debt and equity securities. In the United States, the Jumpstart Our Business Startups (JOBS) Act of 2012 provides exemptions from the rules for certain crowdfunding activities. Under the JOBS Act, the maximum amount of funding a company can raise in a year is limited to about $1 million, and the maximum funds an investor can provide is subject to a limit based on the investor's income.

[7] Economists refer to this as the "market for lemons" problem following a 1970 paper by George Akerlof with that name.

[8] As entertaining television shows such as *Shark Tank* illustrate, experts do not always agree when valuing early-stage businesses.

Initial coin offerings (ICOs) are a form of crowdfunding involving cryptocurrencies. Typically a company issues to investors its own cryptocurrency in exchange for an established one such as bitcoin or ethereum. The investors have no claim on the company, but if the company does well, they expect the cryptocurrency it has issued to increase in value. This is because use of the new cryptocurrency is integral to the operations of the company. ICOs have attracted regulatory attention. China has banned them, and other countries are considering whether they should be regulated in the same way as IPOs.

28.4 Wealth Management

Wealth management has traditionally been very profitable for banks. Fees are often in the 1% to 1.5% range of the amount invested per year and can be much more when hidden fees associated with mutual fund investments and trading costs are taken into account. Once a client's risk appetite has been assessed, wealth management involves finding appropriate investments for the client. Often similar investments are recommended to all clients with similar risk appetites, easing the job of being a wealth manager. It is not surprising that wealth management is an industry that is in the process of being disrupted. As mentioned in Chapter 4, John Bogle took the first step toward reducing the costs of investing with the first index fund in 1975. Index funds have since become very popular and charge fees as low as 0.15% with no human intervention in the form of a wealth manager being necessary.

Robo-advisers first appeared in about 2010. In most countries they must register with the authorities and are subject to regulation. They provide digital platforms where investors express their risk preferences. A portfolio, usually involving ETFs tracking an equity index and safe investments, is then chosen, and the portfolio is automatically rebalanced as necessary. There is very little human intervention. Fees are lower than those charged by traditional wealth managers, typically 0.50% to 0.75% of the amount invested per year. Well-known companies include Wealthfront and Betterment. Some banks and other wealth managers are now responding to the competition by offering their own automated wealth management services. Indeed, those that fail to do this are unlikely to survive. Providers of index mutual funds, such as Vanguard, are also active in this space.

Robo-advisers are making investment advice available to a much wider range of individuals than previously. Often investors can start with as little as \$500 or \$1,000 whereas a traditional wealth manager might require a minimum investment of \$50,000. Robo-advisers make it easy for clients to add to the funds under management on a regular basis. Arguably they serve an important role in society by encouraging people to save when they might not otherwise do so. Robo-advising tended to attract young investors with small amounts to invest in its early days, but a much wider range of investors, including those classified as "high net worth" and HENRYs (high earners not rich yet) are now using the services.

Up to now the main innovation underlying robo-advising has been the delivery of services in a cheaper, novel way that many investors find appealing. The investment strategies underlying the advice given are usually similar to those that have been used by the investment industry for many years (see Chapter 1). Tax-related strategies (such as tax loss harvesting) are often incorporated into the advice that is given.[9]

There is plenty of scope for the strategies to become more sophisticated. Investments can be better diversified internationally and across sectors. They can be better targeted to the goals of the investor, taking into account the investor's age, retirement plans, plans to pay for the higher education of children, and so on.

In 1992, Fischer Black and Robert Litterman at Goldman Sachs published a widely used way of incorporating the views of investors in the selection of a portfolio.[10] Robo-advisers may find ways of expanding the range of alternatives offered to investors using this technology. Alternative sets of views with rationales could be presented, with investors being invited to choose between them. It might even be possible to let the views of the investor be a less structured direct input to the determination of the portfolio.

In Chapter 27, we mentioned that a number of cognitive biases influence decision making. Investors are subject to these biases. For example, they are reluctant to sell losers, they chase trends, and they get disillusioned and exit equity markets when they should stay for the long term. It is often the ability to avoid these biases that distinguishes a professional investor from an amateur. Robo-advisers can try to stop investors falling victim to the biases by developing innovative ways of explaining them. Finally, robo-advising can be combined with other financial innovations so that a percentage of a client's funds is allocated to P2P lending or equity crowdfunding.

Robo-advising has already become an important part of the financial landscape and is likely to become more widely used as the millennial generation accumulates wealth. For this generation, it is much cooler to invest with an iPhone than make a trip to the bank. However, it is perhaps worth sounding one note of caution. Equity markets performed really well in the years following the start of robo-advising in 2010. The appeal of robo-advising may decline when there is a downturn and the clients of robo-advisers, who never invested before, complain about losing money. It is hoped that the advisers will be able to educate investors on the importance of staying focused on the long term.

28.5 Insurance

The application of technology in the insurance industry is referred to as *insurtech*. In some respects insurance companies are less likely to be disrupted by technology innovations than banks are. This is because the nature of insurance is such that consumers require a conservatively run, stable company. For example, it does not make sense to buy house

[9] Tax loss harvesting involves selling stocks with capital losses at the year end to reduce the capital gains tax payable.

[10] See F. Black and R. Litterman, "Global Portfolio Optimization," *Financial Analysts Journal* 48, no. 5 (1992): 28–43.

insurance if there is a non-negligible probability that the insurance company will not be around in the event of the house being destroyed by fire. Regulators monitor the financial condition of insurance companies carefully, and it is very difficult for a new company to offer insurance unless it has a large amount of capital to cover claims. However, some applications of technology in insurance are straightforward. For example, it is now easy to compare the quotes of many different insurance companies online.

Usually customers do not deal directly with insurance companies. They have to go through brokers or agents. These intermediaries are sometimes tied to just one insurance company or to a small number of insurance companies, rather than having a free hand to find the best deal for clients. A natural development would be disintermediation where insurance companies create websites for dealing directly with customers. However, this seems to be happening fairly slowly. Insurance companies are in many cases reluctant to upset the brokers and agents with whom they have worked closely for many years. But eventually as the market develops, brokers and agents are likely to either disappear or become digitized providers of price comparisons.

One important potential insurtech application is the collection of data by insurance companies so that risks are assessed more accurately. Auto insurance is a good example. By placing a "black box" in a car, an insurance company can use telematics (the branch of information technology that deals with the long-distance transmission of computerized information) to accumulate data on a client's driving. The black box can record driving speeds, distances traveled, the time of day or night that the car is being driven, driving style (e.g., frequency of braking and the cornering speed), and the type of road usually used. This enables the insurance company to build up an accurate profile of the driver. If the driver is less risky than average, the insurance premium charged will go down. Similarly, if the driver is more risky than average, it will go up. Black box insurance is still in its infancy, although it has made some inroads into the UK insurance market. As the cost of car insurance rises, it may become more popular, particularly with young people. The latter have traditionally faced very high premiums when they have little or no driving record. A black box could provide a quicker way for them to convince insurance companies that they are good drivers. Another point to note is that the cost of a black box (which is typically paid for by the driver) has come down.

The black box can give the driver information on road conditions. It can also be used for driver education. The driver may be able to access a secure website to find out how well he or she is performing. This can suggest changes to driving style and provide tips on bringing down the cost of insurance. In principle, this could improve the average ability of drivers and make roads safer. Of course not everyone can expect a reduced premium as a result of a black box, and some drivers may regard them as an invasion of privacy. One can speculate that as black boxes become more widely accepted, only good drivers will use them and the premiums for non-users will increase.

To the insurance company, the key advantage of a black box is that it is able to collect data allowing risks to be assessed more accurately. In an ideal world, an insurance premium charged should reflect miles driven and the risks incurred by the driver per mile.[11] Another potential advantage of black boxes to the insurance company is that they may be able to pinpoint blame in the event of an accident and make claims simpler.

A general point is that insurance companies now have opportunities to collect more data than ever before to assess the risks they are taking. These data can potentially allow them to improve estimates of an individual's life expectancy, the probability of a particular house burning down, or the chance of a hurricane in a certain part of the world. Machine learning provides the tools to analyze the data. The insurance industry is by nature conservative (perhaps rightly so), and it is difficult to imagine insurtech radically changing the structure of the industry. (In many cases, new entrants have to partner with established insurers.) But the companies that prosper are likely to be those that take advantage of the opportunities created by what is sometimes referred to as "big data."

28.6 Regulation and Compliance

There are two aspects of the impact of technology on regulation. One is how regulators should handle companies that use technology in innovative ways and have the potential to be a disruptive force in finance. The other, known as *RegTech*, is concerned with how regulation itself can use technology. We will consider both in this section.

28.6.1 Regulation of Financial Innovators

As we have seen in this book, financial institutions are highly regulated. They have to deal with many different regulatory agencies. This is a barrier to entry for small FinTech start-ups. P2P lenders in the United States must enter into an arrangement with a bank to comply with regulations. Robo-advisers must satisfy know-your-client rules, even though there is limited or no human interaction in their services, and register with the authorities. If a FinTech ignores regulations, it is likely to be closed down by supervisory authorities, at least temporarily, and suffer reputational damage.

Regulators do not want to stifle financial innovation. Indeed, as we have pointed out, financial innovation can bring many benefits to society by developing innovative ways of saving, investing, and borrowing so that more people are encouraged to be part of the financial ecosystem. But regulators must tread a fine line. They do not want *shadow banks* (i.e., non-banks that are offering similar services to banks) to subject clients to unreasonable risks.

The Financial Conduct Authority (FCA) in the United Kingdom has developed a "sandbox" program.[12] It defines a sandbox as a safe space in which businesses can test

[11] Metromile, a San Francisco company, is an example of an insurance company moving in this direction. It uses a gadget to monitor mileage driven and sends a monthly invoice for insurance to policy holders to reflect the records of miles driven. The policies are actually written by a large well-capitalized and highly regulated insurance company, National General Insurance.

[12] See Financial Conduct Authority, "Regulatory Sandbox," November 2015.

innovative products, services, business models, and delivery mechanisms without immediately incurring all the normal regulatory consequences of pilot activities. Innovators must fill out an application form explaining what they are planning to do. The FCA encourages virtual testing before real consumers are used and has considered four other ways in which consumers can be protected:

1. Consumers must give informed consent to be included in testing. They are notified of risks and available compensation.
2. The FCA agrees, on a case-by-case basis, to the disclosure and compensation.
3. Customers have the same rights as customers of other firms that are established and authorized by the FCA.
4. The businesses undertaking sandbox trials must compensate individuals for any losses and must demonstrate that it has the resources to do this.

The FCA seems to prefer the second approach, which can include one or more of the other approaches.

Despite pressure from technology companies and from forward-looking established financial institutions, the United States does not have a sandbox program. Some contend that this could put the country at a competitive disadvantage. As mentioned earlier, it does have the JOBS Act, which eases the regulatory burden on crowdfunding activities.

Regulating FinTech companies is likely to be a challenge for regulators in the future. The main asset of a FinTech company is often its computer code, and it is not at all easy for supervisors to evaluate this in the same way that they would evaluate how a bank does its risk management. High-frequency traders (HFTs) illustrate the problem. Their trading activities are determined entirely by their computer code. On May 6, 2010, largely as a result of the activities of HFTs, a "flash crash" occurred in the United States where the Dow Jones Industrial Average declined by about 9% in a matter of minutes before partially recovering. On August 1, 2012, a software error in the program used by an HFT caused a major disruption in the prices of 148 stocks listed on the New York Stock Exchange. The Securities and Exchange Commission has taken some steps to avoid similar problems in the future, but it is difficult for it to guard against all the ways in which software can (intentionally or unintentionally) interfere adversely with the functioning of financial markets.

Another issue for regulators is that the jurisdiction may be difficult to determine for some FinTech companies. This is particularly likely to be the case if all of its transactions are carried out in digital currencies.

28.6.2 RegTech

RegTech is the application of technology to regulation and compliance. Some of the applications are straightforward. Regulatory reporting has traditionally involved thousands of pages of documents being provided to regulators at the end of each reporting period (e.g., every quarter). Moving this online and providing data in real time have the potential to simplify the job of both the regulator and the regulated financial institution, while making regulation more timely and efficient.

Fines for failing to comply with regulations can be steep.[13] For example, in Chapter 23 we mentioned that a French bank, BNP Paribas, was fined an amount equal to its profits for a whole year for violating U.S. sanctions. No doubt this happened because the activities of a small number of employees were not being monitored closely by the bank's compliance group. Technology can make it easier for a bank to ensure that it complies with regulations and the laws of the different countries it operates in. For example, technology can make it possible for it to screen both new customers and new transactions in real time to identify problems. Compliance violations involving issues such as drug money laundering, terrorist financing, and sanctions can be effectively eliminated if a real-time RegTech application, approved by regulators and reflecting the latest rules, is used.

Preventing staff from engaging in questionable behavior is important to financial institutions, as the behavior can prove to be very costly in terms of regulatory fines, legal costs, and reputation. Some innovations allow companies to keep track of their employees' conversations across a number of communications platforms. For example, Digital Reasoning, a company headquartered in Nashville, with offices in London, New York, and Washington, has created surveillance software used by some of the largest banks and asset management companies. It can monitor English conversations (handling six different dialects). It can monitor employee behavior by analyzing millions of e-mails, chat logs, and phone conversations to detect suspicious or unusual activity using machine learning. If machine learning detects an employee whose behavior is quite different from the usual pattern, further investigation might be warranted. In an asset management firm, the behavior might prove to be indicative of insider trading, which, if allowed to continue, could lead to a huge fine. In a bank, it might prove to be indicative of rogue trading or a failure to treat subordinates with respect.

As mentioned in Chapter 23, cyber risk is the biggest operational risk for banks. It is also becoming a regulatory issue. For example, the General Data Protection Regulation in Europe begins in May 2018 and will see financial institutions charged hefty fines if they lose data meant to remain confidential. Some RegTech applications are helping banks keep their data secure and comply with regulations like this one.

Just knowing all the regulations that exist is a major problem for financial institutions, as the regulatory environment it operates in throughout the world changes almost daily. IBM's Watson Regulatory Compliance gives relevant employees access to a library of regulatory requirements in which materials can be filtered geographically, by business type, by product, and by compliance area. This application can automatically highlight the parts of a new regulation that are relevant to a particular issue and compare the regulations related to a particular activity in different parts of the world.

[13]Between the end of the 2007–2008 crisis and 2017, banks throughout the world have paid about $321 billion in fines. Technology to improve a bank's compliance can therefore have a huge payoff.

As explained in Part Three of this book, regulation has moved from being unidimensional to being multidimensional. Regulators used to focus on only a single capital measure. They now have two capital measures (one based on risk-weighted assets and the other based on a simpler leverage calculation) and two liquidity ratios (the liquidity coverage ratio and the net stable funding ratio). Different activities undertaken by the bank have different effects on the ratios. This leads to the potential for technology firms to come up with ways of optimizing bank strategies or providing convenient scenario analysis tools.

28.7 How Should Financial Institutions Respond?

FinTech innovations pose a strategic risk for large banks and other financial institutions. Many of the services they offer will be disrupted. How should they respond? Clearly they should embrace technological change, rather than hoping it is a fad that will disappear. But they should also carefully evaluate how consumer behavior is being affected by technological change and take steps to change their business model accordingly.

Business Snapshot 28.2 describes Eastman Kodak, a company that did not survive technology changes well. A key point is that Kodak was not unaware of the changes taking place in its industry. Indeed, the first digital camera was created in 1975 by an engineer working for Kodak, and Kodak invested billions in the new technology. So where did Kodak go wrong? It understood the new technology but did not appreciate the way it was changing consumer behavior until too late. In the course of a short period of time, a large segment of the population (not just young people) became comfortable sharing photographs and storing them on computers. They no longer felt the need to print photographs.

Kodak coined the term "Kodak moment," which it used extensively in its advertising to convince people that they should have a Kodak camera loaded with Kodak film ready to use at all times. It is easy to be wise after the event, but strategy professionals would argue that the company could have extrapolated from its sales pitch to recognize the business it was in. It was in the imaging or moment-sharing business rather than the film business. We will never know how well it would have survived if it had more enthusiastically embraced electronic sharing and the storing of digital images when the market was in its early stages. (As indicated in Business Snapshot 28.2, Fuji, which was a similar company to Kodak but based in Japan, fared better than Kodak.) It was its implicit belief that demand for hard-copy photographs would continue that ultimately doomed the company.

The disruption of large financial institutions does not seem to be happening as quickly as the disruption of Kodak. They have a number of competitive advantages. They are well capitalized (although the same is true of Apple, Google, and Alibaba). They understand how to deal with the highly regulated environment they operate in (something many FinTech start-ups find difficult). They have a huge customer base

BUSINESS SNAPSHOT 28.2

Eastman Kodak and the Digitization of Photography

In the 1980s and 1990s, the Eastman Kodak company was a successful business selling cameras and film for cameras. It had almost single-handedly transformed photography in the United States from something that was the exclusive preserve of professional photographers to something that almost everyone did. It had successfully navigated technological changes in its industry. It moved from dry-plate to black-and-white film and then sensibly invested in color film even when it was demonstrably inferior to the black-and-white film from which it derived most of its revenue. Its market capitalization in 1997 was about $30 billion.

Steve Sasson, an engineer at Kodak, invented the first digital camera in 1975. According to Sasson, the company's response was "That's cute—but don't tell anyone about it." In fact, Kodak did not ignore the trend to digital cameras. It invested heavily and in 1995 it brought its first digital camera to market, the DC40. In 2001, it launched the EasyShare line of point-and-shoot cameras. Kodak had a first-rate research capability. It had the ability to develop state-of-the-art cameras that could enhance colors, automatically adjust lighting, organize portfolios, share photos, detect smiles to get the best shots, and so on.

Unfortunately, Kodak exhibited some reluctance to continually improve products and innovate, a mistake in an industry where technology changes fast. Cannibalizing existing products in the name of progress is painful, but necessary (even when the new products are less profitable than the old ones). How well Kodak would have survived the smartphone revolution if it had adopted a different strategy is debatable. All we can say is that some of its competitors such as Fuji have fared better. In 2017, Fuji's market capitalization was more than $15 billion. Kodak filed for bankruptcy at the end of 2012, and its successor company had a market capitalization of about $300 million in 2017.

and are still trusted by many customers (although it is important to recognize that the 2007–2008 financial crisis and events since then such as the Wells Fargo debacle described in Business Snapshot 27.1 have served to erode this trust). One can speculate that financial institutions are not as vulnerable as Kodak in that many people are less inclined to experiment with the way their money is handled than with the way they do photography. Also, many start-ups need established financial institutions to offer their products.

However, there are some warning signs that banks should respond to. The Millennial Disruption Index survey indicated that 71% of millennials would rather visit the dentist than listen to what banks are saying, while 73% would rather handle their financial services needs through Google, Amazon, Apple, PayPal, or Square.[14] Millennials also voted four leading U.S. banks among their least loved brands. Millennials are individuals born between 1981 and 2000 and are important future consumers of bank services. By 2022 they will make up 40% of the workforce and by 2050 they will have inherited about $30 trillion. They may feel less radically inclined about financial services as they age, but banks should not rely on this.

Kodak was ultimately rendered largely irrelevant by the digital cameras incorporated in smartphones. Financial institutions naturally do not want to become similarly irrelevant. They have recognized the need to offer mobile apps for payments, wealth management, and a host of other services. But it is important for them to embrace technological change, not just to pay lip service to it. Technological change in the financial sector will continue at an accelerating rate. In many cases it will erode the profits banks previously relied upon (as was the case for Kodak). Being flexible enough to adjust will be a continuing challenge.

The new services developed by banks need to be convenient and designed so that young people classify them as "cool" while older people find them easy to use. There are a number of different approaches being adopted. Some financial institutions have developed new services in-house; some have bought start-ups that have already developed the services; and some have entered into partnerships with start-ups. The first alternative, although the least expensive and most appealing to many working in the financial services sector, can be quite difficult given the somewhat complacent culture that often permeates large companies. The second and third alternatives can be used as a way of disrupting the culture and accelerating change. Some banks have found it useful to create an organizationally distinct unit that has the ability to bring in outside talent when necessary and can partner with start-ups. Whatever approach is adopted, strong support from the CEO and the board is necessary for success.

The financial institutions that prosper will be the ones that manage to change their culture so that they are able to provide services in a fast, flexible way that appeals to consumers. Lou Gerstner was appointed CEO at IBM in the 1990s when the company was losing money. He successfully changed the culture and transformed it into a profitable company. He is quoted as saying, "Culture isn't just one aspect of the game. It is the game." He successfully changed IBM from being a supplier of mainframe computers to being a business problem solver where teams of people worked well together. Banks are similar to IBM in that they are huge organizations with many bright employees. The job of the CEO is to create a culture where individuals are rewarded for embracing technology and finding innovative ways of making their services more convenient and valuable for clients.

Summary

What will the financial services sector look like in, say, 2035? Of course, the pace of technological change is so fast that accurate forecasts are impossible. However, offering

[14]See www.millennialdisruptionindex.com.

a few predictions is a convenient way of summarizing some of the points made in this chapter.

Checks and credit cards will largely disappear. Mobile wallets on smartphones and wearables will become the norm. Biometrics will be used to make payments more secure. Some central banks will choose to switch from paper currencies to digital currencies, and transacting in digital currencies will become more widespread. Machine learning will be able to undertake many tasks such as credit assessment and fraud detection much better than human beings. Record keeping will be more secure and faster using blockchains and other DLTs.

Large insurance companies will continue to exist, but brokers and agents will not. Insurers will have access to more data than ever before and will use machine learning and other techniques to assess risks more precisely. There will be partnerships formed between insurance companies and technology firms to deliver insurance services in innovative ways.

Investing and borrowing will see big changes. Crowdfunding and P2P lending will become more widely used. Some of today's large banks will find it necessary to offer these services to savers and borrowers. It will be unusual for funds to remain idle for any length of time, and the ease of borrowing and investing will make more people part of the financial ecosystem.

Some banks that fail to adjust to new technology enthusiastically will fail and be acquired by other banks. (Indeed, too-big-to-fail may no longer be considered a relevant descriptor of very large banks.) Some large technology firms will become banks and offer a full range of banking services. Most bank services will be offered more cheaply and more efficiently than before so that banks will lose some of their traditional sources of profit. The banks that survive will have to cut costs by making big reductions in the number of branches they have and the number of people they employ. Banks will have to form partnerships with many different technology firms, to keep the services they offer up to date.

Further Reading

Chishti, S., and J. Barberis. *The FinTech Book: The Financial Technology Handbook for Investors, Entrepreneurs, and Visionaries*. Chichester, UK: John Wiley & Sons, 2016.

Christensen, C. M., and M. E. Raynor. *The Innovator's Solution*. Boston: Harvard Business Review Press, 2003.

Harvey, C. R. "Cryptofinance." Working Paper, SSRN 2438299.

IOSCO. *Research Report on Financial Technologies*, February 2017.

Sironi, P. *FinTech Innovation: From Robo-Advisors to Goal-Based Investing and Gamification*. Chichester, UK: John Wiley & Sons, 2016.

Practice Questions and Problems (Answers at End of Book)

28.1 Explain the terms *disintermediation* and *reintermediation*.

28.2 What is meant by (a) machine learning and (b) distributed ledger technologies?

28.3 What is meant by biometric authorization? Give some examples.

28.4 Why might a central bank switch from fiat money to digital money?

28.5 Explain how P2P lending and equity crowdfunding work.

28.6 "When thinking about the impact of FinTech, financial institutions should be thankful for regulation." Explain.

28.7 "Index funds are a great idea, but if everyone chose them there would be no price discovery." Do you agree with this argument?

28.8 What is meant by RegTech? Give some examples.

28.9 "Banks should aim to respond to change like IBM, not like Kodak." Discuss this statement.

Further Questions

28.10 What is meant by a bubble? Consider whether the increase in the price of bitcoin in 2017 is a bubble.

28.11 Look at the data in Table 28.1. Is Lending Club good at assessing risk? Is there a reasonable trade-off between risk and return for lenders? What risks are lenders taking?

28.12 With the benefit of hindsight, we can say that Kodak was in the imaging and moment-sharing business. What business are banks in?

Chapter 29

Risk Management Mistakes to Avoid

Since the mid-1980s, there have been some spectacular losses in financial markets. This chapter uses them to review some of the key points made in this book. The losses that we will consider are listed in Business Snapshot 29.1.

One remarkable aspect of the list in Business Snapshot 29.1 is the number of times huge losses were caused by the activities of a single person. In 1995, Nick Leeson's trading brought a 200-year-old British bank, Barings, to its knees; in 1994, Robert Citron's trading led to Orange County, a municipality in California, losing about $2 billion. Joseph Jett's trading for Kidder Peabody caused losses of $350 million. John Rusnak's losses of $700 million at Allied Irish Bank came to light in 2002. Jérôme Kerviel lost over $7 billion for Société Générale in 2008. Kweku Adoboli lost $2.3 billion for UBS in 2011.

The losses emphasize the importance of internal controls. Many of the losses we will consider occurred because systems were inadequate so that the risky positions being taken were simply not known. It is also important for risk managers to continually "think outside the box" about what could go wrong so that as many potential adverse events as possible are identified.

29.1 Risk Limits

The first and most important lesson from the losses concerns risk limits. It is essential that all companies (financial and nonfinancial) define in a clear and unambiguous way limits

BUSINESS SNAPSHOT 29.1

Big Losses

Allied Irish Bank: This bank lost about $700 million from the unauthorized speculative activities of one of its foreign exchange traders, John Rusnak, which lasted for a number of years. Rusnak covered up his losses by creating fictitious options trades.

Barings: This 200-year-old British bank was wiped out in 1995 by the activities of one trader, Nick Leeson, in Singapore. The trader's mandate was to arbitrage between Nikkei 225 futures quotes in Singapore and Osaka. Instead he made big bets on the future direction of the Nikkei 225 using futures and options. The total loss was close to $1 billion.

Enron's Counterparties: Enron managed to conceal its true situation from its shareholders with some creative contracts. Several financial institutions that allegedly helped Enron do this have each had to settle shareholder lawsuits for over $1 billion.

Hammersmith and Fulham (see Business Snapshot 23.1): This British Local Authority lost about $600 million on sterling interest rate swaps and options in 1988. The two traders responsible for the loss knew surprisingly little about the products they were trading.

Kidder Peabody (see Business Snapshot 25.2): The activities of a single trader, Joseph Jett, led to this New York investment dealer losing $350 million trading U.S. government securities. The loss arose because of a mistake in the way the company's computer system calculated profits.

Long-Term Capital Management (see Business Snapshot 22.1): This hedge fund lost about $4 billion in 1998 carrying out convergence arbitrage strategies. The loss was caused by a flight to quality after Russia defaulted on its debt.

National Westminster Bank: This British bank lost about $130 million from using an inappropriate model to value swap options in 1997.

Orange County (see Appendix B): The activities of the treasurer, Robert Citron, led to this California municipality losing about $2 billion in 1994. The treasurer was using derivatives to speculate that interest rates would not rise.

Procter & Gamble (see Business Snapshot 5.4): The treasury department of this large U.S. company lost about $90 million in 1994 trading highly exotic interest rate derivatives contracts with Bankers Trust. It later sued Bankers Trust and settled out of court.

Société Générale (see Business Snapshot 5.5): Jérôme Kerviel, an equity trader in the Paris office, lost over $7 billion speculating on movements in equity indices in January 2008. He is alleged to have concealed his exposure by creating fictitious trades. Like Leeson at Barings, his mandate was to do arbitrage trades.

Subprime Mortgage Losses (see Chapter 6): In 2007, investors lost confidence in the structured products created from U.S. subprime mortgages. This led to a "credit crunch" with losses of tens of billions of dollars by financial institutions and the worst recession since the 1930s.

UBS: In 2011, Kweku Adoboli lost $2.3 billion taking unauthorized speculative positions in stock market indices.

to the financial risks that can be taken. They should then set up procedures for ensuring that the limits are adhered to. Ideally, overall risk limits should be set at board level. These should then be converted to limits applicable to the individuals responsible for managing particular risks. Daily reports should indicate the gain or loss that will be experienced for particular movements in market variables. These should be checked against the actual gains and losses that are experienced to ensure that the valuation procedures underlying the reports are accurate.

It is particularly important that companies monitor risks carefully when derivatives are used. This is because derivatives can be used for hedging or speculation or arbitrage. Without close monitoring, it is impossible to know whether a derivatives trader has switched from being a hedger to a speculator or switched from being an arbitrageur to being a speculator. The Barings, Société Générale, and UBS losses are classic examples of what can go wrong. In each case, the trader's mandate was to carry out low-risk arbitrage and hedging. Unknown to their superiors, they switched from being arbitrageurs and hedgers to taking huge bets on the future direction of market variables. The systems within the banks were inadequate and did not detect what was going on.

The argument here is not that no risks should be taken. A trader in a financial institution or a fund manager should be allowed to take positions on the future direction of relevant market variables. What we are arguing is that the sizes of the positions that can be taken should be limited and the systems in place should accurately report the risks being taken.

29.1.1 A Difficult Situation

What happens if an individual exceeds risk limits and makes a profit? This is a tricky issue for senior management. It is tempting to ignore violations of risk limits when profits result. However, this is shortsighted. It leads to a culture where risk limits are not taken seriously, and it paves the way for a disaster. The classic example here is Orange County. Robert Citron's activities in 1991–1993 had been very profitable for Orange County, and the municipality had come to rely on his trading for additional funding. People chose to ignore the risks he was taking because he had produced profits. Unfortunately, the losses made in 1994 far exceeded the profits from previous years.

The penalties for exceeding risk limits should be just as great when profits result as when losses result. Otherwise, traders who make losses are liable to keep increasing their bets in the hope that eventually a profit will result and all will be forgiven.

29.1.2 Do Not Assume You Can Outguess the Market

Some traders are quite possibly better than others. But no trader gets it right all the time. A trader who correctly predicts the direction in which market variables will move 60% of the time is doing well. If a trader has an outstanding track record (as Robert Citron did in the early 1990s), it is likely to be a result of luck rather than superior trading skill. As our discussion of mutual fund performance in Chapter 4 shows, it appears that fund managers usually produce superior returns as a result of luck rather than skill.

Suppose that a financial institution employs 16 traders and one of those traders makes profits in every quarter of a year. Should the trader receive a good bonus? Should the trader's risk limits be increased? The answer to the first question is that the trader will inevitably receive a good bonus. The answer to the second question should be no. The chance of making a profit in four consecutive quarters from random trading is 0.5^4 or 1 in 16. This means that just by chance one of the 16 traders will "get it right" every single quarter of the year. We should not assume that the trader's luck will continue and we should not increase the trader's risk limits.

29.1.3 Do Not Underestimate the Benefits of Diversification

When a trader appears good at predicting a particular market variable, there is a tendency to increase the trader's risk limits. We have just argued that this is a bad idea because it is quite likely that the trader has been lucky rather than clever. However, let us suppose that we are really convinced that the trader has special talents. How undiversified should we allow ourselves to become in order to take advantage of the trader's special skills? As indicated in Section 1.1, the benefits from diversification are large. A trader has to be very good for it to be worth foregoing these benefits to speculate heavily on just one market variable.

An example will illustrate the point here. Suppose that there are 20 stocks, each of which has an expected return of 10% per annum and a standard deviation of return of 30%. The correlation between the returns from any two of the stocks is 0.2. By dividing an investment equally among the 20 stocks, an investor has an expected return of 10% per annum and standard deviation of returns of 14.7%. Diversification enables the investor to reduce risks by over half. Another way of expressing this is that diversification enables an investor to double the expected return per unit of risk taken. The investor would have to be very good at stock picking to consistently get a better risk-return trade-off by choosing undiversified stock portfolios.

29.1.4 Carry Out Scenario Analyses and Stress Tests

As discussed in Chapter 22, the calculation of risk measures such as VaR should always be accompanied by scenario analyses and stress testing to obtain an understanding of what can go wrong. Without the discipline of stress testing, human beings have an unfortunate tendency to anchor on one or two scenarios when evaluating decisions. In 1993 and 1994, for example, Procter & Gamble was so convinced that interest rates would remain low that it ignored the possibility of a 100-basis-point increase in its decision making.

Once stress-testing results have been produced, they should become inputs to a financial institution's strategic decision making. All too often, particularly when times are good, the results of stress testing are ignored. This happened at a number of financial institutions prior to July 2007.

29.2 Managing the Trading Room

In trading rooms there is a tendency to regard high-performing traders as "untouchable" and to not subject their activities to the same scrutiny as other traders. Apparently Joseph Jett, Kidder Peabody's star trader of Treasury instruments, was often "too busy" to answer questions and discuss his positions with the company's risk managers.

All traders—particularly those making high profits—should be fully accountable. It is important for the financial institution to know whether the high profits are being made by taking unreasonably high risks. It is also important to check that the financial institution's computer systems and pricing models are correct and are not being manipulated in some way.

29.2.1 Separate the Front, Middle, and Back Office

The *front office* in a financial institution consists of the traders who are executing trades, taking positions, and so on. The *middle office* consists of risk managers who are monitoring the risks being taken. The *back office* is where the record keeping and accounting takes place. Some of the worst derivatives disasters have occurred because these functions were not kept separate. Nick Leeson controlled both the front and back office for Barings in Singapore and was, as a result, able to conceal the disastrous nature of his trades from his superiors in London for some time.

29.2.2 Do Not Blindly Trust Models

We discussed model risk in Chapter 25. Some of the large losses experienced by financial institutions arose because of the models and computer systems being used. Kidder Peabody was misled by its own systems. Another example of an incorrect model leading to losses is provided by National Westminster Bank. This bank had an incorrect model for valuing swap options that led to a large loss.

If large profits are reported when relatively simple trading strategies are followed, there is a good chance that the models underlying the calculation of the profits are wrong. Similarly, if a financial institution appears to be particularly competitive on its quotes for a particular type of deal, there is a good chance that it is using a different model from other market participants—which almost certainly means that, before too long, it will have to change its model and report a loss. Getting too much business of a certain type can be just as worrisome to the head of a trading room as getting too little business of that type.

29.2.3 Be Conservative in Recognizing Inception Profits

When a financial institution sells a highly exotic instrument to a nonfinancial corporation, the valuation can be highly dependent on the underlying model. For example, instruments with long-dated embedded interest rate options can be highly dependent on the interest rate model used. In these circumstances, a phrase used to describe the daily marking to market of the deal is *marking to model*. This is because there are no market prices for similar deals that can be used as a benchmark.

Suppose that a financial institution manages to sell an instrument to a client for $10 million more than it is worth—or at least $10 million more than its model says it is worth. The $10 million is known as an *inception profit*. When should it be recognized? There appears to be a lot of variation in what different derivatives dealers do. Some recognize the $10 million immediately, whereas others are much more conservative and recognize it slowly over the life of the deal.

Recognizing inception profits immediately is very dangerous. It encourages traders to use aggressive models, take their bonuses, and leave before the model and the value of the deal come under close scrutiny. It is much better to take reserves and to recognize inception profits slowly, so that traders are motivated to investigate the impact of several different models and several different sets of assumptions before committing themselves to a deal.

29.2.4 Do Not Sell Clients Inappropriate Products

It is tempting to sell corporate clients inappropriate products, particularly when they appear to have an appetite for the underlying risks. But this is shortsighted. A dramatic illustration is provided by the activities of Bankers Trust (BT) in the period leading up to the spring of 1994. BT's clients, such as Procter & Gamble, were persuaded to buy high-risk and totally inappropriate products. A typical product would give a client a good chance of saving a few basis points on its borrowings and a small chance of losing a large amount of money. The products worked well for BT's clients in 1992 and 1993, but blew up in 1994 when interest rates rose sharply. The bad publicity that followed hurt BT greatly. The years it had spent building up trust among corporate clients and developing an enviable reputation for innovation in derivatives were largely lost as a result of the activities of a few overly aggressive salesmen. BT was forced to pay large amounts of money to its clients to settle lawsuits out of court. It was taken over by Deutsche Bank in 1999.

29.2.5 Beware of Easy Profits

Enron provides an example of how overly aggressive deal makers can cost their banks billions of dollars. Doing business with Enron seemed very profitable and banks competed with each other for this business. But the fact that many banks push hard to get a certain type of business should not be taken as an indication that the business will ultimately be profitable. The business that Enron did with banks resulted in many very expensive shareholder lawsuits. In general, transactions where high profits seem easy to achieve should be looked at closely for potential operational, credit, or market risks.

Investing in the AAA-rated tranches of ABSs and ABS CDOs that were created from subprime mortgages (see Chapter 6) seemed like a money-making machine for many banks. The promised returns were higher than the returns normally earned on AAA-rated instruments. Most traders did not stop to ask whether this was because there were risks that had not been taken into account by rating agencies.

29.3 Liquidity Risk

We discussed liquidity risk in Chapter 24. Financial engineers usually base the pricing of exotic instruments and other instruments that trade relatively infrequently on the prices of actively traded instruments. For example:

1. A trader often calculates a zero curve from actively traded government bonds (known as on-the-run bonds) and uses it to price bonds that trade less frequently (off-the-run bonds).
2. A trader often implies the volatility of an asset from actively traded options and uses it to price less actively traded options.
3. A trader often implies information about the behavior of interest rates from actively traded interest rate caps and swap options and uses it to price products that are highly structured.

These practices are not unreasonable. However, it is dangerous to assume that less actively traded instruments can always be traded at close to their theoretical prices. When financial markets experience a shock of one sort or another, liquidity black holes may develop (see Section 24.3). Liquidity then becomes very important to investors, and illiquid instruments often sell at a big discount to their theoretical values.

An example of liquidity risk is provided by Long-Term Capital Management (LTCM), which is the subject of Business Snapshot 22.1. This hedge fund followed a strategy known as *convergence arbitrage*. It attempted to identify two securities (or portfolios of securities) that should in theory sell for the same price. If the market price of one security was less than that of the other, it would buy that security and sell the other. The strategy is based on the idea that if two securities have the same theoretical price, their market prices should eventually be the same.

In the summer of 1998, LTCM took a huge loss. This was largely because a default by Russia on its debt caused a flight to quality. LTCM tended to be long illiquid instruments and short the corresponding liquid instruments. (For example, it was long off-the-run bonds and short on-the-run bonds.) The spreads between the prices of illiquid instruments and the corresponding liquid instruments widened sharply after the Russian default. LTCM was highly leveraged. It experienced huge losses and there were margin calls on its positions that it had difficulty meeting.

The LTCM story reinforces the importance of carrying out scenario analyses and stress testing to look at what can happen in extreme scenarios. As discussed in Chapter 22, it is important to consider not only immediate losses but also losses created by knock-on effects.

29.3.1 Beware When Everyone Is Following the Same Trading Strategy

It sometimes happens that many market participants are following essentially the same trading strategy. This creates a dangerous environment where there are liable to be big market moves, liquidity black holes, and large losses for the market participants.

We gave one example of this in Business Snapshot 24.4, which discussed portfolio insurance and the market crash of October 1987. In the months leading up to the crash, increasing numbers of portfolio managers were attempting to insure their portfolios by creating synthetic put options. This involved buying stocks or stock index futures after a rise in the market and selling them after a fall. The result was an unstable market. A relatively small decline in stock prices could lead to a wave of selling by portfolio insurers. The latter would lead to a further decline in the market, which could give rise to another wave of selling, and so on. There is little doubt that, without portfolio insurance, the crash of October 1987 would have been much less severe.

Another example is provided by LTCM in 1998. Its position was made more difficult by the fact that many other hedge funds were following similar convergence arbitrage strategies. After the Russian default and the flight to quality, LTCM tried to liquidate part of its portfolio to meet margin calls. Unfortunately, other hedge funds were facing similar problems to LTCM and trying to do similar trades. This exacerbated the situation, causing liquidity spreads to be even higher than they would otherwise have been and reinforcing the flight to quality. Consider, for example, LTCM's position in U.S. Treasury bonds. It was long the illiquid off-the-run bonds and short the liquid on-the-run bonds. When a flight to quality caused spreads between yields on the two types of bonds to widen, LTCM had to liquidate its positions by selling off-the-run bonds and buying on-the-run bonds. Other large hedge funds were doing the same. As a result, the price of on-the-run bonds rose relative to off-the-run bonds and the spread between the two yields widened even more.

A further example is provided by British insurance companies in the late 1990s. This is discussed in Business Snapshot 3.1. All insurance companies decided to hedge their exposure to a fall in long-term rates at about the same time. The result was a fall in long-term rates!

The key lesson to be learned from these stories is that it is important to see the big picture of what is going on in financial markets and to understand the risks inherent in situations where many market participants are liable to follow the same trading strategy.

29.3.2 Do Not Make Excessive Use of Short-Term Funding for Long-Term Needs

All financial institutions finance long-term needs with short-term sources of funds to some extent. But a financial institution that relies too heavily on short-term funds is likely to expose itself to unacceptable liquidity risks.

During the period leading up to the credit crisis of 2007, there was a tendency for subprime mortgages and other long-term assets to be financed by commercial paper while they were in a portfolio waiting to be packaged into structured products. Conduits and special purpose vehicles had an ongoing requirement for this type of financing. The commercial paper would typically be rolled over every month. For example, the purchasers of commercial paper issued on April 1 would be redeemed with the proceeds of a new commercial paper issue on May 1; this new commercial paper issue would in turn be redeemed with another new commercial paper issue on June 1; and so on. When investors lost confidence in subprime mortgages in August 2007, it became impossible to roll over commercial paper. In many instances, banks had provided guarantees and had to provide financing. This led to a shortage of liquidity. As a result, the credit crisis was more severe than it would have been if longer-term financing had been arranged.

Many of the failures of financial institutions during the crisis (e.g., Lehman Brothers and Northern Rock) were caused by excessive reliance on short-term funding. Once the market (rightly or wrongly) becomes concerned about the health of a financial institution, it can be impossible to roll over the financial institution's short-term funding. The Basel Committee has recognized the importance of liquidity risks by introducing liquidity requirements in Basel III.

29.3.3 Market Transparency Is Important

One of the lessons from the credit crisis of 2007 is that market transparency is important. During the period leading up to 2007, investors had been trading highly structured products without any real knowledge of the underlying assets. All they knew was the credit rating of the security being traded. With hindsight, we can say that investors should have demanded more information about the underlying assets and should have more carefully assessed the risks they were taking.

The subprime meltdown of August 2007 caused investors to lose confidence in all structured products and withdraw from that market. This led to a market breakdown where tranches of structured products could only be sold at prices well below their theoretical values. There was a flight to quality and credit spreads increased. If there had been market transparency, so that investors understood the asset-backed securities they were buying, there would still have been subprime losses, but the flight to quality and disruptions to the market would have been less pronounced.

29.4 Lessons for Nonfinancial Corporations

Here are some lessons applicable primarily to nonfinancial corporations.

29.4.1 Make Sure You Fully Understand the Trades You Are Doing

Corporations should never undertake a trade or a trading strategy that they do not fully understand. This is a somewhat obvious point, but it is surprising how often a trader working for a nonfinancial corporation will, after a big loss, admit to not really understanding what was going on and claim to have been misled by investment bankers. Robert Citron, the treasurer of Orange County did this. So did the traders working for Hammersmith and Fulham, who in spite of their huge positions were surprisingly uninformed about how the swaps and other interest rate derivatives they traded really worked.

If a senior manager in a corporation does not understand a trade proposed by a subordinate, the trade should not be approved. A simple rule of thumb is that if a trade and the rationale for entering into it are so complicated that it cannot be easily understood, it is almost certainly inappropriate for the corporation. The trades undertaken by Procter & Gamble would have been vetoed under this criterion.

One way of ensuring that you fully understand a financial instrument is to value it. If a corporation does not have the in-house capability to value an instrument, it should not trade it. In practice, corporations often rely on their investment bankers for valuation advice. This is dangerous, as Procter & Gamble found out. When it wanted to unwind its transactions, it found it was facing prices produced by Bankers Trust's proprietary models, which it had no way of checking.

29.4.2 Make Sure a Hedger Does Not Become a Speculator

One of the unfortunate facts of life is that hedging is relatively dull, whereas speculation is exciting. When a company hires a trader to manage foreign-exchange risk, commodity price risk, or interest rate risk, there is a danger that the following happens. At first the trader does the job diligently and earns the confidence of top management. He or she assesses the company's exposures and hedges them. As time goes by, the trader becomes convinced that he or she can outguess the market. Slowly the trader becomes a speculator. At first things go well, but then a loss is made. To recover the loss, the trader doubles down. Further losses are made, and so on. The result is a disaster.

As mentioned earlier, clear limits to the risks that can be taken should be set by senior management. Controls should be put in place to ensure that the limits are obeyed. The trading strategy for a corporation should start with an analysis of the risks facing the corporation in foreign exchange, interest rate, and other markets. A decision should then be taken on how the risks are to be reduced to acceptable levels. It is a clear sign that something is wrong within a corporation if the trading strategy is not derived in a very direct way from the company's exposures.

29.4.3 Be Cautious about Making the Treasury Department a Profit Center

In the last 30 years, there has been a tendency to make the treasury department within a corporation a profit center. This seems to have much to recommend it. The treasurer is motivated to reduce financing costs and manage risks as profitably as possible. The problem is that the potential for the treasurer to make profits is limited. When raising funds and investing surplus cash, the treasurer is facing an efficient market. The treasurer can usually improve the bottom line only by taking additional risks. The company's hedging program gives the treasurer some scope for making shrewd decisions that increase profits. But it should be remembered that the goal of a hedging program is to reduce risks, not to increase expected profits. The decision to hedge will lead to a worse outcome than the decision not to hedge roughly 50% of the time. The danger of making the treasury department a profit center is that the treasurer is motivated to become a speculator. An Orange County or Procter & Gamble type of outcome is then liable to occur.

29.5 A Final Point

Most of the risks we have considered in this book are what are termed *known risks*. They are risks such as market risks and credit risks that can be quantified using historical data. Two other types of risk are important to financial institutions: *unknown risks* and *unknowable risks*.

Unknown risks are risks where the event that could cause a loss is known, but its probability of occurrence cannot easily be determined. Operational risks and business risks include many different types of unknown risks. What is the probability of a rogue trader loss? What is the probability of a loss from a major lawsuit? What is the probability that operations in a particular country will be expropriated by the government? These probabilities cannot usually be estimated using historical data. As discussed in Chapter 22, subjective probabilities are often used. In a widely referenced book, Knight (1921) uses the term "risk" to refer to known risks and the term "uncertainty" to refer to unknown risks.[1]

Unknowable risks are risks where even the event that could cause a loss is not known. Unknowable risks are in many ways the most insidious because they come as a complete surprise and often lead to dramatic losses. An unknowable risk is sometimes referred to as a *black swan*. (Black swans were not considered possible until they were discovered in Australia.) As pointed out by Taleb (2007), once it has occurred, a black swan event is often considered to be obvious.[2] Did the producers of multivolume encyclopedias in 1970 consider the possibility that technological developments would render their product worthless? Probably not, but ex post it seems a fairly obvious risk.

[1] See F. H. Knight, *Risk, Uncertainty and Profit* (Boston: Houghton Mifflin, 1921).

[2] See N. N. Taleb, *The Black Swan: The Impact of the Highly Improbable* (New York: Random House, 2007).

How can companies manage unknown and unknowable risks? A key tool is flexibility. Companies should avoid excessive leverage and try to ensure that their costs are variable rather than fixed as far as possible. Diversification across products and markets also increases flexibility. In the future, insurance companies may well develop more products to handle unknown and unknowable risks. As discussed in Chapter 23, products have already been developed to provide protection against some operational (unknown) risks. Handling unknowable risk is a challenging (but not totally impossible) contract design problem.

Further Reading

Diebold, F. X., N. A. Doherty, and R. J. Herring. *The Known, the Unknown, and the Unknowable in Financial Risk Management*. Princeton, NJ: Princeton University Press, 2010.

Dunbar, N. *Inventing Money: The Story of Long-Term Capital Management and the Legends Behind It*. Chichester, UK: John Wiley & Sons, 2000.

Gomory, R. "The Known, the Unknown and the Unknowable." *Scientific American*, June 1995.

Jorion, P. *Big Bets Gone Bad: Derivatives and Bankruptcy in Orange County*. New York: Academic Press, 1995.

Jorion, P. "How Long-Term Lost Its Capital." *Risk* (September 1999): 31–36.

Ju, X., and N. Pearson. "Using Value at Risk to Control Risk Taking: How Wrong Can You Be?" *Journal of Risk* 1 (1999): 5–36.

Persaud, A. D., ed. *Liquidity Black Holes: Understanding Quantifying and Managing Financial Liquidity Risk*. London: Risk Books, 2003.

Sorkin, A. R. *Too Big to Fail*. New York: Penguin, 2009.

Tett, G. *Fool's Gold: How the Bold Dream of a Small Tribe at JPMorgan Was Corrupted by Wall Street Greed and Unleashed a Catastrophe*. New York: Free Press, 2009.

Thomson, R. *Apocalypse Roulette: The Lethal World of Derivatives*. London: Macmillan, 1998.

Zhang, P. G. *Barings Bankruptcy and Financial Derivatives*. Singapore: World Scientific Publishing, 1995.

Part Six

APPENDICES

Appendix A

Compounding Frequencies for Interest Rates

A statement by a bank that the interest rate on one-year deposits is 10% per annum sounds straightforward and unambiguous. In fact, its precise meaning depends on the way the interest rate is measured.

If the interest rate is measured with annual compounding, the bank's statement that the interest rate is 10% means that \$100 grows to

$$\$100 \times 1.1 = \$110$$

at the end of one year. When the interest rate is measured with semiannual compounding, it means that we earn 5% every six months, with the interest being reinvested. In this case, \$100 grows to

$$\$100 \times 1.05 \times 1.05 = \$110.25$$

at the end of one year. When the interest rate is measured with quarterly compounding, the bank's statement means that we earn 2.5% every three months, with the interest being reinvested. The \$100 then grows to

$$\$100 \times 1.025^4 = \$110.38$$

at the end of one year. Table A.1 shows the effect of increasing the compounding frequency further.

Table A.1 Effect of the Compounding Frequency on the Value of $100 at the End of One Year When the Interest Rate is 10% per Annum

Compounding Frequency	Value of $100 at End of Year ($)
Annually ($m = 1$)	110.00
Semiannually ($m = 2$)	110.25
Quarterly ($m = 4$)	110.38
Monthly ($m = 12$)	110.47
Weekly ($m = 52$)	110.51
Daily ($m = 365$)	110.52

The compounding frequency defines the units in which an interest rate is measured. A rate expressed with one compounding frequency can be converted into an equivalent rate with a different compounding frequency. For example, from Table A.1 we see that 10.25% with annual compounding is equivalent to 10% with semiannual compounding. We can think of the difference between one compounding frequency and another to be analogous to the difference between kilometers and miles. They are two different units of measurement.

To generalize our results, suppose that an amount A is invested for n years at an interest rate of R per annum. If the rate is compounded once per annum, the terminal value of the investment is

$$A(1 + R)^n$$

If the rate is compounded m times per annum, the terminal value of the investment is

$$A\left(1 + \frac{R}{m}\right)^{mn} \tag{A.1}$$

When $m = 1$ the rate is sometimes referred to as the *equivalent annual interest rate*.

A.1 Continuous Compounding

The limit as the compounding frequency, m, tends to infinity is known as *continuous compounding*.[1] With continuous compounding, it can be shown that an amount A invested for n years at rate R grows to

$$Ae^{Rn} \tag{A.2}$$

where $e = 2.71828$. The function e^x, which is also written $\exp(x)$, is built into most calculators, so the computation of the expression in equation (A.2) presents no problems.

[1] Actuaries sometimes refer to a continuously compounded rate as the *force of interest*.

In the example in Table A.1, $A = 100$, $n = 1$, and $R = 0.1$, so that the value to which A grows with continuous compounding is

$$100e^{0.1} = \$110.52$$

This is (to two decimal places) the same as the value with daily compounding. For most practical purposes, continuous compounding can be thought of as being equivalent to daily compounding. Compounding a sum of money at a continuously compounded rate R for n years involves multiplying it by e^{Rn}. Discounting it at a continuously compounded rate R for n years involves multiplying by e^{-Rn}.

Suppose that R_c is a rate of interest with continuous compounding and R_m is the equivalent rate with compounding m times per annum. From the results in equations (A.1) and (A.2), we have

$$Ae^{R_c n} = A\left(1 + \frac{R_m}{m}\right)^{mn}$$

or

$$e^{R_c} = \left(1 + \frac{R_m}{m}\right)^{m}$$

This means that

$$R_c = m \ln\left(1 + \frac{R_m}{m}\right) \tag{A.3}$$

and

$$R_m = m(e^{R_c/m} - 1) \tag{A.4}$$

These equations can be used to convert a rate with a compounding frequency of m times per annum to a continuously compounded rate and vice versa. The function ln is the natural logarithm function and is built into most calculators. It is defined so that if $y = \ln x$, then $x = e^y$.

Example A.1

Consider an interest rate that is quoted as 10% per annum with semiannual compounding. From equation (A.3), with $m = 2$ and $R_m = 0.1$, the equivalent rate with continuous compounding is

$$2 \ln\left(1 + \frac{0.1}{2}\right) = 0.09758$$

or 9.758% per annum.

Example A.2

Suppose that a lender quotes the interest rate on loans as 8% per annum with continuous compounding, and that interest is actually paid quarterly. From equation (A.4), with $m = 4$ and $R_c = 0.08$, the equivalent rate with quarterly compounding is

$$4(e^{0.08/4} - 1) = 0.0808$$

or 8.08% per annum. This means that on a \$1,000 loan, interest payments of \$20.20 would be required each quarter.

Appendix B

Zero Rates, Forward Rates, and Zero-Coupon Yield Curves

The n-year zero-coupon interest rate is the rate of interest earned on an investment that starts today and lasts for n years. All the interest and principal is realized at the end of n years. There are no intermediate payments. The n-year zero-coupon interest rate is sometimes also referred to as the n-year *spot interest rate*, the n-year *zero rate*, or just the n-year zero. The zero rate as a function of maturity is referred to as the *zero curve*. Suppose a five-year zero rate with continuous compounding is quoted as 5% per annum. (See Appendix A for a discussion of compounding frequencies.) This means that \$100, if invested for five years, grows to

$$100 \times e^{0.05 \times 5} = 128.40$$

A forward rate is the future zero rate implied by today's zero rates. Consider the zero rates shown in Table B.1. The forward rate for the period between six months and one year is 6.6%. This is because 5% for the first six months combined with 6.6% for the next six months gives an average of 5.8% for the first year. Similarly, the forward rate for the period between 12 months and 18 months is 7.6%, because this rate, when combined

Table B.1 Zero Rates

Maturity (years)	**Zero Rate (%) (cont. comp.)**
0.5	5.0
1.0	5.8
1.5	6.4
2.0	6.8

with 5.8% for the first 12 months, gives an average of 6.4% for the 18 months. In general, the forward rate F for the period between times T_1 and T_2 is

$$F = \frac{R_2 T_2 - R_1 T_1}{T_2 - T_1} \tag{B.1}$$

where R_1 is the zero rate for maturity T_1 and R_2 is the zero rate for maturity T_2. This formula is exactly true when rates are measured with continuous compounding and approximately true for other compounding frequencies. The results from using this formula on the rates in Table B.1 are in Table B.2. For example, substituting $T_1 = 1.5$, $T_2 = 2.0$, $R_1 = 0.064$, and $R_2 = 0.068$ gives $F = 0.08$ showing that the forward rate for the period between 18 months and 24 months is 8%.

Investors who think that future interest rates will be markedly different from forward rates have no difficulty in finding trades that reflect their beliefs. Consider an investor who can borrow or lend at the rates in Table B.1. Suppose the investor thinks that the six-month interest rates will not change much over the next two years. The investor can borrow six-month funds and invest for two years. The six-month borrowings can be rolled over at the end of 6, 12, and 18 months. If interest rates do stay about the same, this strategy will yield a profit of about 1.8% per year because interest will be received at 6.8% and paid at 5%. This type of trading strategy is known as a *yield curve play*. The investor is speculating that rates in the future will be quite different from the forward rates shown in Table B.2.

Robert Citron, the treasurer at Orange County, used yield curve plays similar to the one we have just described very successfully in 1992 and 1993. The profit from Mr. Citron's trades became an important contributor to Orange County's budget and he was re-elected. In 1994, he used the same strategy more aggressively. If short-term interest rates had remained the same or declined, he would have done very well. As it

Table B.2 Forward Rates for Zero Rates in Table B.1

Period (years)	**Forward Rate (%) (cont. comp.)**
0.5 to 1.0	6.6
1.0 to 1.5	7.6
1.5 to 2.0	8.0

happened, interest rates rose sharply during 1994. On December 1, 1994, Orange County announced that its investment portfolio had lost $1.5 billion and several days later it filed for bankruptcy protection.

B.1 Bond Pricing

Most bonds provide coupons periodically. The bond's principal (which is also known as its par value or face value) is received at the end of its life. The theoretical price of a bond can be calculated as the present value of all the cash flows that will be received by the owner of the bond. The most accurate approach is to use a different zero rate for each cash flow. To illustrate this, consider the situation where zero rates are as in Table B.1. Suppose that a two-year bond with a principal of $100 provides coupons at the rate of 6% per annum semiannually. To calculate the present value of the first coupon of $3, we discount it at 5% for six months; to calculate the present value of the second coupon of $3, we discount it at 5.8% for one year; and so on. The theoretical price of the bond is therefore:

$$3e^{-0.05\times0.5} + 3e^{-0.058\times1.0} + 3e^{-0.064\times1.5} + 103e^{-0.068\times2.0} = 98.39$$

or $98.39.

B.2 Bond Yields

A bond's yield is the discount rate that, when applied to all cash flows, gives a bond price equal to its market price. Suppose that the theoretical price of the bond we have been considering, $98.39, is also its market value (i.e., the market's price of the bond is in exact agreement with the data in Table B.1). If y is the yield on the bond, expressed with continuous compounding, we must have:

$$3e^{-y\times0.5} + 3e^{-y\times1.0} + 3e^{-y\times1.5} + 103e^{-y\times2.0} = 98.39$$

This equation can be solved using Excel's Solver or in some other way to give $y = 6.76\%$.

B.3 Treasury Rates

Treasury rates are the rates an investor earns on Treasury bills and Treasury bonds. These are the instruments used by a government to borrow in its own currency. Japanese Treasury rates are the rates at which the Japanese government borrows in yen; U.S. Treasury rates are the rates at which the U.S. government borrows in U.S. dollars; and so on.

B.4 Determining Treasury Zero Rates

One way of determining Treasury zero rates such as those in Table B.1 is to observe the yields on "strips." These are zero-coupon bonds that are artificially created by traders when they sell the coupons on a Treasury bond separately from the principal.

Table B.3 Rates After Two-Year Rate Has Been Determined Using the Bootstrap Method

Maturity (years)	Zero Rate (%) (cont. comp.)
0.5	5.00
1.0	5.80
1.5	6.40
2.0	6.80
2.5	??

Another way of determining Treasury zero rates is from regular Treasury bills and bonds. The most popular approach is known as the *bootstrap method*. This involves working from short maturities to successively longer maturities matching prices. Suppose that Table B.3 gives the Treasury rates determined so far and that a 2.5-year bond providing a coupon of 8% sells for \$102 per \$100 of principal. We would determine the 2.5-year zero-coupon interest rate as the rate which, when used in conjunction with the rates in Table B.3, gives the correct price for this bond. This involves solving

$$4e^{-0.05\times0.5} + 4e^{-0.058\times1.0} + 4e^{-0.064\times1.5} + 4e^{-0.068\times2.0} + 104e^{-R\times2.5} = 102$$

which gives $R = 7.05\%$. The zero curve is usually assumed to be linear between the points that are determined by the bootstrap method. (In our example, the 2.25-year zero rate would be 6.9025%.) It is also assumed to be constant prior to the first point and beyond the last point. The zero curve for our example is shown in Figure B.1.

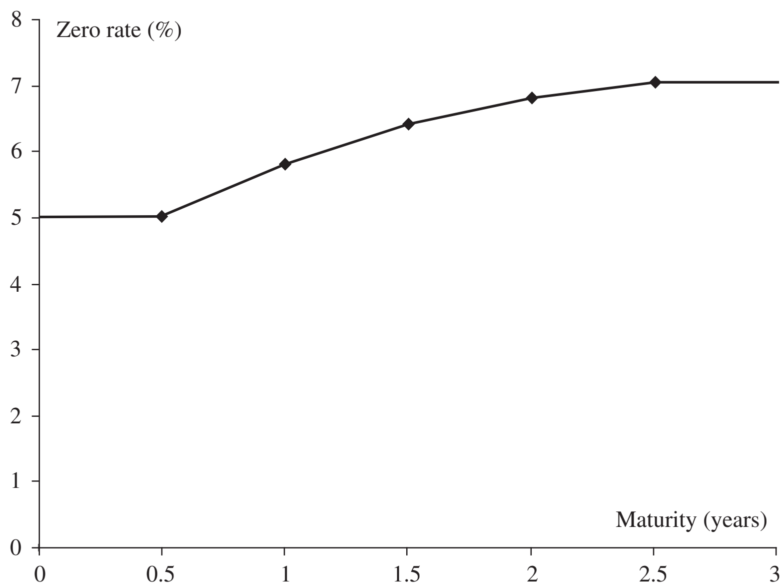


Figure B.1 Zero Curve for Data in Table B.3

B.5 OIS Zero Rates

The OIS zero curve is determined using a similar bootstrap method to that used for determining the Treasury zero curve. The OIS rates for maturities out to one year involve a single exchange at maturity. They therefore give immediate information about the zero rates for those maturities. OIS rates for maturities longer than one year are typically for contracts where there are exchanges every three months. They are the yields on bonds that sell for par and provide quarterly payments.

Suppose that the 3-month, 6-month, 9-month, and 12-month OIS zero rates have been calculated as 2%, 2.3%, 2.5%, and 2.7%, respectively (with continuous compounding). Suppose further that the 1.25-year OIS rate is 2.8%. A 1.25-year bond with a principal of 100 providing a coupon of 0.7 every three months is worth par (i.e., 100). If R is the 1.25-year OIS zero rate, we must have

$$0.7e^{-0.02\times0.25} + 0.7e^{-0.023\times0.5} + 0.7e^{-0.025\times0.75} + 0.7e^{-0.027\times1} + 100.7e^{-R\times1.25} = 100$$

This gives $R = 2.794\%$.

Appendix C

Valuing Forward and Futures Contracts

The forward or futures price of an investment asset that provides no income is given by

$$F_0 = S_0 e^{rT}$$

where S_0 is the spot price of the asset today, T is the time to maturity of the forward or futures contract, and r is the continuously compounded risk-free rate for maturity T. When the asset provides income during the life of the contract that has a present value I, this becomes

$$F_0 = (S_0 - I)e^{rT}$$

When it provides a yield at rate q, it becomes

$$F_0 = S_0 e^{(r-q)T}$$

A foreign currency can be regarded as an investment asset that provides a yield equal to the foreign risk-free rate, so that the forward or futures price for a foreign currency is

$$F_0 = S_0 e^{(r-r_f)T}$$

where r_f is the foreign risk-free rate (continuously compounded) and S_0 is the spot exchange rate. The value of a forward contract where the holder has the right to buy the asset for a price of K is, in all cases,

$$(F_0 - K)e^{-rT}$$

where F_0 is the forward price, given by one of the formulas above. The value of a forward contract where the holder has the right to sell the asset for a price of K is similarly

$$(K - F_0)e^{-rT}$$

Example C.1

Consider a six-month futures contract on the S&P 500. The current value of the index is 1,200, the six-month risk-free rate is 5% per annum, and the average dividend yield on the S&P 500 over the next six months is expected to be 2% per annum (both rates continuously compounded). The futures price is $1{,}200e^{(0.05-0.02)\times 0.5}$ or 1,218.14.

Example C.2

The current forward price of a commodity for a contract maturing in nine months is \$550 per ounce. A company has a forward contract to buy 1,000 ounces of the commodity for a delivery price of \$530 in nine months. The nine-month risk-free rate is 4% per annum continuously compounded. The value of the forward contract is $1{,}000 \times (550 - 530)e^{-0.04\times 9/12}$, or \$19,409.

In most situations forward and futures prices can be assumed to be the same. An exception is provided by contracts on interest rates. The quote for a Eurodollar futures contract should not be assumed to be the same as the corresponding forward interest rate, particularly when the contract has a long time to maturity.

Appendix D

Valuing Swaps

A plain vanilla interest rate swap can be valued by assuming that the interest rates that are realized in the future equal today's forward interest rates. As an example, consider an interest rate swap that has 14 months remaining and a notional principal of \$100 million. A fixed rate of 5% per annum is received and LIBOR is paid, with exchanges taking place every six months. Assume that (a) four months ago, the six-month LIBOR rate was 4%, (b) the forward LIBOR interest rate for a six-month period starting in two months is 4.6%, and (c) the forward LIBOR for a six-month period starting in eight months is 5.2%. All rates are expressed with semiannual compounding. Assuming that forward rates are realized, the cash flows on the swap are as shown in Table D.1. (For example, in eight months the fixed-rate cash flow received is $0.5 \times 0.05 \times 100$, or \$2.5 million; the floating-rate cash flow paid is $0.5 \times 0.046 \times 100$, or \$2.3 million.) The value of the swap is the present value of the net cash flows in the final column.[1] The discount rates used to discount the cash flows are usually OIS zero rates. (See Appendix B for their calculation.)

D.1 LIBOR Forward Rates

LIBOR forward rates can be calculated using a bootstrap method. Suppose that the 6-month, 12-month, 18-month, and 24-month OIS zero rates are 3.8%, 4.3%, 4.6%, and 4.75% (with continuous compounding). Suppose that in a two-year swap where cash

[1] Note that the precise timing of cash flows depends on day count conventions and holiday calendars.

Table D.1 Valuing an Interest Rate Swap by Assuming Forward Rates Are Realized

Time	Fixed Cash Flow ($ million)	Floating Cash Flow ($ million)	Net Cash Flow ($ million)
2 months	2.5	−2.0	0.5
8 months	2.5	−2.3	0.2
14 months	2.5	−2.6	−0.1

flows are exchanged every six months the swap rate is 5%. Suppose further that the 0.5-year LIBOR rate is 4%, the forward rate for the period between 6 months and 12 months is 5%, and the forward rate for the period between 12 months and 18 months is 5.5% (all semiannually compounded). The value of a two-year swap where LIBOR is received and 5% is paid is zero. The value of the first exchange is

$$0.5 \times (0.04 - 0.05) \times 100 \times e^{-0.038 \times 0.5} = -0.4906$$

The value of the second exchange is

$$0.5 \times (0.05 - 0.05) \times 100 \times e^{-0.043 \times 1} = 0$$

The value of the third exchange is

$$0.5 \times (0.055 - 0.05) \times 100 \times e^{-0.046 \times 1.5} = 0.2333$$

The total value of the first three payments is $-0.4906 + 0 + 0.2333 = -0.2573$ Suppose that the (assumed unknown) forward rate for the final payment is F. For the swap to be worth zero, we must have

$$0.5 \times (F - 0.05) \times 100 \times e^{-0.0475 \times 2} = 0.2573$$

This gives $F = 0.05565$ or 5.565%.

D.2 Currency Swaps

A currency swap can be valued by assuming that exchange rates in the future equal today's forward exchange rates. As an example, consider a currency swap in which 4% will be received in GBP and 6% will be paid in USD once a year. The principals in the two currencies are 10 million USD and 5 million GBP. The swap will last for another three years. The swap cash flows are shown in the second and third columns of Table D.2. The forward exchange rates are (we assume) those shown in the fourth column. These are used to convert the GBP cash flows to USD. The final column shows the net cash flows. The value of the swap is the present value of these cash flows.

Table D.2 Valuing a Currency Swap by Assuming Forward Exchange Rates Are Realized (all cash flows in millions)

Time	USD Cash Flow	GBP Cash Flow	Forward Exchange Rate	USD Value of GBP Cash Flow	Net Cash Flow in USD
1	−0.6	0.2	1.8000	0.360	−0.240
2	−0.6	0.2	1.8400	0.368	−0.232
3	−0.6	0.2	1.8800	0.376	−0.224
3	−10.0	5.0	1.8800	9.400	−0.600

An alternative approach (which gives the same valuation) is to regard the swap as a long position in a GBP bond and a short position in a USD bond. Each bond can be valued in its own currency in the usual way and the current exchange rate can be used to convert the value of the GBP bond from GBP to USD.

Appendix E

Valuing European Options

The Black–Scholes–Merton formulas for valuing European call and put options on an investment asset that provides no income are

$$c = S_0 N(d_1) - Ke^{-rT} N(d_2)$$

and

$$p = Ke^{-rT} N(-d_2) - S_0 N(-d_1)$$

where

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

$$d_2 = \frac{\ln(S_0/K) + (r - \sigma^2/2)T}{\sigma\sqrt{T}} = d_1 - \sigma\sqrt{T}$$

The function $N(x)$ is the cumulative probability distribution function for a standardized normal distribution (see tables at the end of the book or Excel's NORMSDIST function). The variables c and p are the European call and European put price, S_0 is today's asset price, K is the strike price, r is the continuously compounded risk-free rate, σ is the stock price volatility, and T is the time to maturity of the option.

Table E.1 Greek Letters for Options on an Asset That Provides a Yield at Rate q

Greek Letter	Call Option	Put Option
Delta	$e^{-qT}N(d_1)$	$e^{-qT}[N(d_1)-1]$
Gamma	$\dfrac{N'(d_1)e^{-qT}}{S_0\sigma\sqrt{T}}$	$\dfrac{N'(d_1)e^{-qT}}{S_0\sigma\sqrt{T}}$
Theta (per yr)	$-S_0N'(d_1)\sigma e^{-qT}/(2\sqrt{T})$ $+qS_0N(d_1)e^{-qT}$ $-rKe^{-rT}N(d_2)$	$-S_0N'(d_1)\sigma e^{-qT}/(2\sqrt{T})$ $-qS_0N(-d_1)e^{-qT}$ $+rKe^{-rT}N(-d_2)$
Vega (per %)	$\dfrac{S_0\sqrt{T}N'(d_1)e^{-qT}}{100}$	$\dfrac{S_0\sqrt{T}N'(d_1)e^{-qT}}{100}$
Rho (per %)	$\dfrac{KTe^{-rT}N(d_2)}{100}$	$-\dfrac{KTe^{-rT}N(-d_2)}{100}$

When the underlying asset provides a cash income, the present value of the income during the life of the option should be subtracted from S_0. When the underlying asset provides a yield at rate q, the formulas become:

$$c = S_0e^{-qT}N(d_1) - Ke^{-rT}N(d_2)$$

and

$$p = Ke^{-rT}N(-d_2) - S_0e^{-qT}N(-d_1)$$

where

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

$$d_2 = \frac{\ln(S_0/K) + (r - q - \sigma^2/2)T}{\sigma\sqrt{T}} = d_1 - \sigma\sqrt{T}$$

Options on a foreign currency can be valued by setting q equal to the foreign risk-free rate. Options on a futures or forward price can be valued using these formulas by setting S_0 equal to the current forward or futures price, $r = q$, and σ equal to the volatility of the forward or futures price.

Table E.1 gives formulas for the Greek letters for European options on an asset that provides income at rate q. $N'(x)$ is the standard normal density function:

$$N'(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}$$

The implied volatility is the volatility, σ, which when substituted into a Black–Scholes–Merton pricing formula gives the market price. (See Section 10.2.) When Greek letters are calculated, σ is set equal to the implied volatility.

Example E.1

Consider a six-month European call option on a stock index. The current value of the index is 1,200, the strike price is 1,250, the risk-free rate is 5%, the dividend yield on the index is 2%, and the index volatility is 20%. In this case, $S_0 = 1{,}200$, $K = 1{,}250$, $r = 0.05$, $q = 0.02$, $\sigma = 0.2$, and $T = 0.5$. The value of the option is 53.44, the delta of the option is 0.45, the gamma is 0.0023, the theta is –0.22, the vega is 3.33, and rho is 2.44. Note that the formula in Table E.1 gives theta per year. The theta quoted here is per calendar day.

The calculations in this appendix can be done with the RMFI software, available on the author's website, by selecting Option Type: Black–Scholes European. Option valuation is described more fully in Hull (2018).[1]

[1] See J. C. Hull, *Options, Futures, and Other Derivatives*, 10th ed. (Upper Saddle River, NJ: Pearson, 2018).

Appendix F

Valuing American Options

To value American-style options, we divide the life of the option into n time steps of length Δt. Suppose that the asset price at the beginning of a step is S. At the end of the time step it moves up to Su with probability p and down to Sd with probability $1 - p$. For an investment asset that provides no income the values of u, d, and p are given by

$$u = e^{\sigma\sqrt{\Delta t}} \qquad d = \frac{1}{u} \qquad p = \frac{a - d}{u - d}$$

with

$$a = e^{r\Delta t}$$

where r is the risk-free rate and σ is the volatility.

Figure F.1 shows the tree constructed for valuing a five-month American put option on a non-dividend-paying stock where the initial stock price is 50, the strike price is 50, the risk-free rate is 10%, and the volatility is 40%. In this case, there are five steps so that $\Delta t = 0.08333$, $u = 1.1224$, $d = 0.8909$, $a = 1.0084$, and $p = 0.5073$. The upper number at each node is the stock price and the lower number is the value of the option.

At the final nodes of the tree, the option price is its intrinsic value. For example, at node G, the option price is $50 - 35.36 = 14.64$. At earlier nodes, we first calculate a value assuming that the option is held for a further time period of length Δt and then check to see whether early exercise is optimal. Consider first node E. If the option is

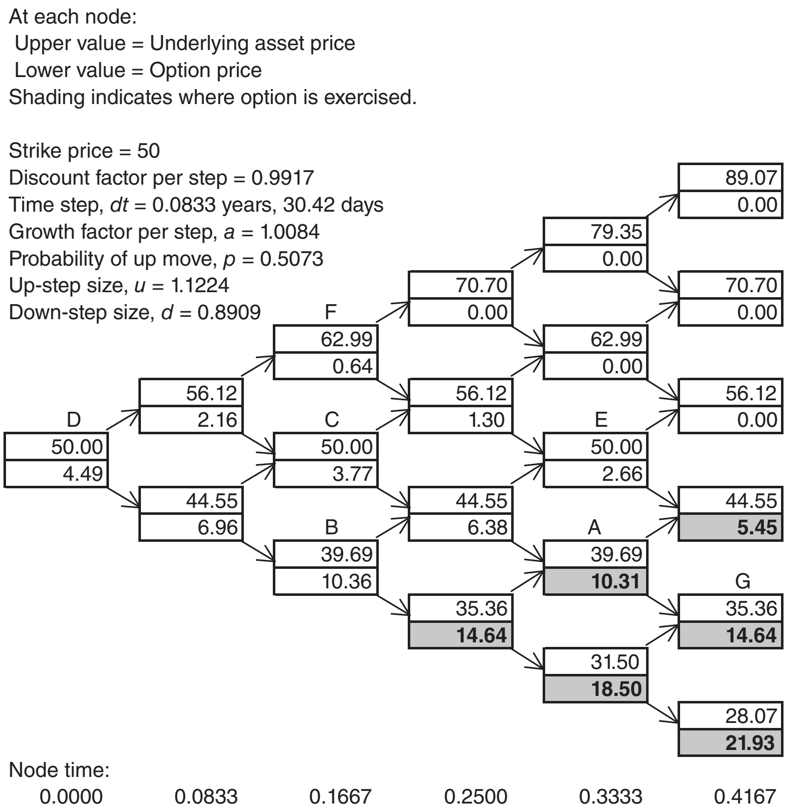


Figure F.1 Binomial Tree for American Put on Non-Dividend-Paying Stock

held for a further time period, it will be worth 0.00 if there is an up move (probability: p) and 5.45 if there is a down move (probability: $1 - p$). The expected value in time Δt is therefore $0.5073 \times 0 + 0.4927 \times 5.45$ or 2.686, and the 2.66 value at node E is calculated by discounting this at the risk-free rate of 10% for one month. The option should not be exercised at node E as the payoff from early exercise would be zero. Consider next node A. A calculation similar to that just given shows that, assuming it is held for a further time period, the option's value at node A is 9.90. If exercised, its value is $50 - 39.69 = 10.31$. In this case, it should be exercised and the value of being at node A is 10.31.

Continuing to work back from the end of the tree to the beginning, the value of the option at the initial node D is found to be 4.49. As the number of steps on the tree is increased, the accuracy of the option price increases. With 30, 50, and 100 time steps, we get values for the option of 4.263, 4.272, and 4.278, respectively.

To calculate delta, we consider the two nodes at time Δt. In our example, as we move from the lower node to the upper node, the option price changes from 6.96 to 2.16 and the stock price changes from 44.55 to 56.12. The estimate of delta is the change in the option price divided by the change in the stock price:

$$\text{Delta} = \frac{2.16 - 6.96}{56.12 - 44.55} = -0.41$$

To calculate gamma we consider the three nodes at time $2\Delta t$. The delta calculated from the upper two nodes (C and F) is -0.241. This can be regarded as the delta for a stock price of $(62.99 + 50)/2 = 56.49$. The delta calculated from the lower two nodes (B and C) is -0.639. This can be regarded as the delta for a stock price of $(50 + 39.69)/2 = 44.84$. The estimate of gamma is the change in delta divided by the change in the stock price:

$$\text{Gamma} = \frac{-0.241 - (-0.639)}{56.49 - 44.84} = 0.034$$

We estimate theta from nodes D and C as

$$\text{Theta} = \frac{3.77 - 4.49}{2 \times 0.08333}$$

or –4.30 per year. This is –0.0118 per calendar day. Vega is estimated by increasing the volatility, constructing a new tree, and observing the effect of the increased volatility on the option price. Rho is calculated similarly.

When the asset underlying the option provides a yield at rate q, the procedure is exactly the same except that $a = e^{(r-q)\Delta t}$ instead of $e^{r\Delta t}$ in the equation for p. (When the asset is a foreign currency q is set equal to the foreign risk-free rate.) When the option is on the forward or futures price, a is set equal to one and the tree shows the forward or futures price at each node. The calculations we have described can be done using the RMFI software, available on the author's website, by selecting Option Type: Binomial American. Binomial trees and other numerical procedures are described more fully in Hull (2018).[1]

[1] See J. C. Hull, *Options, Futures, and Other Derivatives*, 10th ed. (Upper Saddle River, NJ: Pearson, 2018).

Appendix G

Taylor Series Expansions

Consider a function $z = F(x)$. When a small change Δx is made to x, there is a corresponding small change Δz in z. A first approximation to the relationship between Δz and Δx is

$$\Delta z = \frac{dz}{dx}\Delta x \tag{G.1}$$

This relationship is exact if z is a linear function of x and approximate in other situations. A more accurate approximation is

$$\Delta z = \frac{dz}{dx}\Delta x + \frac{1}{2}\frac{d^2 z}{dx^2}(\Delta x)^2 \tag{G.2}$$

This relationship is exact if z is a quadratic function of x and approximate in other situations. By adding more terms in the series, we can increase accuracy. The full expansion is

$$\Delta z = \frac{dz}{dx}\Delta x + \frac{1}{2!}\frac{d^2 z}{dx^2}(\Delta x)^2 + \frac{1}{3!}\frac{d^3 z}{dx^3}(\Delta x)^3 + \frac{1}{4!}\frac{d^4 z}{dx^4}(\Delta x)^4 + \cdots$$

where ! denotes the factorial function: $2! = 2 \times 1 = 2$; $3! = 3 \times 2 \times 1 = 6$; $4! = 4 \times 3 \times 2 \times 1 = 24$; and so on.

Example G.1

Consider the function $z = \sqrt{x}$ so that

$$\frac{dz}{dx} = \frac{1}{2x^{1/2}} \qquad \frac{d^2z}{dx^2} = -\frac{1}{4x^{3/2}} \qquad \frac{d^3z}{dx^3} = \frac{3}{8x^{5/2}}$$

Suppose that $x = 2$ and $\Delta x = 0.1$ so that $\Delta z = \sqrt{2.1} - \sqrt{2} = 0.034924$. When $x = 2$,

$$\frac{dz}{dx} = 0.35355 \qquad \frac{d^2z}{dx^2} = -0.08839 \qquad \frac{d^3z}{dx^3} = 0.06629$$

The first order approximation to Δz, given by equation (G.1), is

$$\Delta z = 0.35355 \times 0.1 = 0.035355$$

The second order approximation, given by equation (G.2), is

$$\Delta z = 0.35355 \times 0.1 + \frac{1}{2} \times (-0.08839) \times 0.1^2 = 0.034913$$

The third order approximation is

$$\Delta z = 0.35355 \times 0.1 + \frac{1}{2} \times (-0.08839) \times 0.1^2 + \frac{1}{6} \times 0.06629 \times 0.1^3 = 0.034924$$

It can be seen that the series expansion quickly converges to the correct answer of 0.034924.

G.1 *Functions of Two Variables*

Consider next a function of two variables, $z = F(x, y)$. Suppose that Δx and Δy are small changes in x and y, respectively, and that Δz is the corresponding small change in z. In this case, the first order approximation is

$$\Delta z = \frac{\partial z}{\partial x}\Delta x + \frac{\partial z}{\partial y}\Delta y \tag{G.3}$$

The second order approximation is

$$\Delta z = \frac{\partial z}{\partial x}\Delta x + \frac{\partial z}{\partial y}\Delta y + \frac{1}{2}\frac{\partial^2 z}{\partial x^2}(\Delta x)^2 + \frac{1}{2}\frac{\partial^2 z}{\partial y^2}(\Delta y)^2 + \frac{\partial^2 z}{\partial x \partial y}(\Delta x \Delta y) \tag{G.4}$$

Example G.2

Consider the function $z = \sqrt{xy}$ so that

$$\frac{\partial z}{\partial x} = \frac{y^{1/2}}{2x^{1/2}} \qquad \frac{\partial z}{\partial y} = \frac{x^{1/2}}{2y^{1/2}}$$

$$\frac{\partial^2 z}{\partial x^2} = -\frac{y^{1/2}}{4x^{3/2}} \qquad \frac{\partial^2 z}{\partial y^2} = -\frac{x^{1/2}}{4y^{3/2}} \qquad \frac{\partial^2 z}{\partial x \partial y} = \frac{1}{4(xy)^{1/2}}$$

Suppose that $x = 2$, $y = 1$, $\Delta x = 0.1$, and $\Delta y = 0.1$ so that

$$\Delta z = \sqrt{2.1 \times 1.1} - \sqrt{2 \times 1} = 0.10565$$

When $x = 2$ and $y = 1$

$$\frac{\partial z}{\partial x} = 0.35355 \qquad \frac{\partial z}{\partial y} = 0.70711$$

$$\frac{\partial^2 z}{\partial x^2} = -0.08839 \qquad \frac{\partial^2 z}{\partial y^2} = -0.35355 \qquad \frac{\partial^2 z}{\partial x \partial y} = 0.17678$$

The first order approximation to Δz, given by equation (G.3), is

$$\Delta z = 0.35355 \times 0.1 + 0.70711 \times 0.1 = 0.10607$$

The second order approximation, given by equation (G.4), is

$$\begin{aligned}\Delta z &= 0.35355 \times 0.1 + 0.70711 \times 0.1 \\ &\quad + \frac{1}{2} \times (-0.08839) \times (0.1)^2 + \frac{1}{2} \times (-0.35355) \times (0.1)^2 \\ &\quad + 0.17678 \times 0.1 \times 0.1 = 0.10562\end{aligned}$$

The series solution is converging to the correct answer of 0.10565.

G.2 General Result

For a function z of n variables, $x_1, x_2, \ldots, x_n$, the Taylor series expansion of Δz is

$$\Delta z = \sum_{m_1=0}^{m_1=\infty} \cdots \sum_{m_n=0}^{m_n=\infty} \frac{1}{m_1! \cdots m_n!} \frac{\partial^m z}{\partial x_1^{m_1} \cdots \partial x_n^{m_n}} \Delta x_1^{m_1} \cdots \Delta x_n^{m_n}$$

where $m = m_1 + \cdots m_n$ and the term where all m_i are zero is zero.

Appendix H

Eigenvectors and Eigenvalues

Consider an $n \times n$ matrix A and suppose that $\mathbf{x}$ is an $n \times 1$ vector. Consider the equation

$$A\mathbf{x} = \lambda\mathbf{x} \tag{H.1}$$

The equation can be written

$$(A - \lambda I)\mathbf{x} = 0$$

where I is the $n \times n$ identity matrix (which is the $n \times n$ matrix with diagonal elements equal to 1 and all other elements equal to zero). Clearly, $\mathbf{x} = 0$ is a solution to equation (H.1). Under what circumstances are there other solutions? A theorem in linear algebra tells us that there are other solutions when the determinant of $A - \lambda I$ is zero. The values of λ that lead to solutions of equation (H.1) are therefore the values of λ that we get when we solve the equation that sets the determinant of $A - \lambda I$ equal to zero. This equation is an nth order polynomial in λ. In general, it has n solutions. The solutions are the *eigenvalues* of the matrix, A. The vector $\mathbf{x}$ that solves equation (H.1) for a particular eigenvalue is an *eigenvector*. In general, there are n eigenvectors, one corresponding to each eigenvalue.

As a simple example, suppose that

$$A = \begin{pmatrix} 1 & -1 \\ 2 & 4 \end{pmatrix}$$

In this case

$$A - \lambda I = \begin{pmatrix} 1-\lambda & -1 \\ 2 & 4-\lambda \end{pmatrix}$$

The determinant of this matrix is

$$(1-\lambda)(4-\lambda) - (-1) \times 2 = \lambda^2 - 5\lambda + 6$$

The solutions to this equation are $\lambda = 3$ and $\lambda = 2$. These are the two eigenvalues of the matrix.

To determine the eigenvectors corresponding to $\lambda = 3$, we solve equation (H.1):

$$\begin{pmatrix} 1 & -1 \\ 2 & 4 \end{pmatrix} \mathbf{x} = 3\mathbf{x}$$

Setting

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

the equation to be solved becomes

$$\begin{pmatrix} 1 & -1 \\ 2 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 3 \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

The simultaneous equations corresponding to this are

$$x_1 - x_2 = 3x_1$$

and

$$2x_1 + 4x_2 = 3x_2$$

Both these equations are equivalent to

$$x_2 + 2x_1 = 0$$

It follows that any pair of numbers x_1 and x_2 solve the equation when $\lambda = 3$ providing $x_2 = -2x_1$. By convention, the values of x_1 and x_2 that are reported are those for which the length of the vector $\mathbf{x}$ is 1. This means that $x_1^2 + x_2^2 = 1$. In this case, the solution where

$\mathbf{x}$ has a length of 1 is $x_1 = \sqrt{0.2} = 0.447$ and $x_2 = -2\sqrt{0.2} = -0.894$. (An alternative is $x_1 = -0.447$ and $x_2 = 0.894$.) The solution

$$\mathbf{x} = \begin{pmatrix} 0.447 \\ -0.894 \end{pmatrix}$$

is the eigenvector corresponding to the first eigenvalue ($\lambda = 3$).

A similar calculation shows that equation (H.1) is satisfied for $\lambda = 2$ when $x_1 + x_2 = 0$. The solution where $\mathbf{x}$ has a length of 1 is $x_1 = \sqrt{0.5} = 0.707$ and $x_2 = -\sqrt{0.5} = -0.707$. (An alternative is $x_1 = -0.707$ and $x_2 = 0.707$.) The eigenvector corresponding to the second eigenvalue ($\lambda = 2$) is therefore

$$\mathbf{x} = \begin{pmatrix} 0.707 \\ -0.707 \end{pmatrix}$$

For matrices of larger size, a numerical procedure must be used to determine eigenvectors and eigenvalues. One such numerical procedure is provided by Press et al. (2007).[1]

Applications of eigenvalues and eigenvectors are in Appendices I and J. Software for calculating eigenvectors and eigenvalues is on the author's website.

[1] See W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes: The Art of Scientific Computing*, 3rd ed. (Cambridge, UK: Cambridge University Press, 2007).

Appendix I

Principal Components Analysis

A principal components analysis is concerned with understanding the structure of data on n correlated variables. The aim of the analysis is to replace the n variables by a smaller number of uncorrelated variables. In the example in Section 9.7, there are eight variables. These are the daily changes in the 1-year, 2-year, 3-year, 4-year, 5-year, 7-year, 10-year, and 30-year swap rates.

The first step in the analysis is to calculate a covariance matrix from the data. As explained in Section 11.3, this is an $n \times n$ matrix where the (i, j) entry is the covariance between variable i and variable j. The entries on the diagonal (where $i = j$) are variances.

The next step is to calculate the eigenvalues and eigenvectors for this matrix (see Appendix H). The eigenvectors are chosen to have length 1. (As explained in Appendix H, this means that the sum of the squares of their elements is 1.) The eigenvector corresponding to the highest eigenvalue is the first principal component; the eigenvector corresponding to the second highest eigenvalue is the second principal component; and so on. The principal components for the example in Section 9.7 are shown in Table 9.6.

The eigenvalue for the ith principal component, as a percentage of the sum of all the eigenvalues, is the percentage of the overall variance explained by the ith principal component. The square root of the ith eigenvalue is the standard deviation of the ith factor score (see Table 9.7).

Software for carrying out a principal components analysis is on the author's website.

Appendix J

Manipulation of Credit Transition Matrices

Suppose that A is an $n \times n$ matrix of credit rating changes in one year. This is a matrix such as the one shown in Table 21.1. Assuming that rating changes in successive time periods are independent, the matrix of credit rating changes in m years is A^m. If m is an integer, this can be readily calculated using the normal rules for matrix multiplication.

Consider next the problem of calculating the transition matrix for $1/m$ years where m is an integer. (For example, when we are interested in one-month changes, $m = 12$.) This is a more complicated problem because we need to calculate the mth root of a matrix. We first calculate eigenvectors $\mathbf{x_1}, \mathbf{x_2}, \ldots, \mathbf{x_n}$ and the corresponding eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ of the matrix A. These are explained in Appendix H. They have the property that

$$A\mathbf{x_i} = \lambda_i \mathbf{x_i} \tag{J.1}$$

Define X as an $n \times n$ matrix whose ith column is $\mathbf{x_i}$ and Λ as an $n \times n$ diagonal matrix (i.e., a matrix that has zero values everywhere except on the diagonal) where the ith diagonal element is λ_i. From equation (J.1), we have

$$AX = X\Lambda$$

so that

$$A = X\Lambda X^{-1}$$

Define Λ^* as a diagonal matrix where the ith diagonal element is $\lambda_i^{1/m}$. Then

$$(X\Lambda^* X^{-1})^m = (X\Lambda^* X^{-1})(X\Lambda^* X^{-1})(X\Lambda^* X^{-1}) \dots (X\Lambda^* X^{-1}) = X(\Lambda^*)^m X^{-1}$$
$$= X\Lambda X^{-1} = A$$

showing that the mth root of A, and therefore the transition matrix for time $1/m$ years, is $X\Lambda^* X^{-1}$.

Some authors, such as Jarrow, Lando, and Turnbull (1997), prefer to handle this problem in terms of what is termed a *generator matrix*.[1] This is a matrix Γ such that the transition matrix for a short period of time, Δt, is $I + \Gamma\Delta t$ and the transition matrix for a longer period of time, t, is

$$\exp(t\Gamma) = \sum_{k=0}^{\infty} \frac{(t\Gamma)^k}{k!}$$

where I is the identity matrix (with ones on the diagonal and zeros elsewhere).

Software for manipulating credit transition matrices is on the author's website.

[1] See R. A. Jarrow, D. Lando, and S. M. Turnbull, "A Markov Model for the Term Structure of Credit Spreads," *Review of Financial Studies* 10 (1997): 481–523.

Appendix K

Valuation of Credit Default Swaps

Credit default swaps (CDSs) are described in Chapter 19. They can be valued using (risk-neutral) default probability estimates.

Suppose that the hazard rate of a reference entity is 2% per annum for five years. Table K.1 shows survival probabilities and unconditional default probabilities (that is, default probabilities as seen at time zero) for each of the five years. From equation (19.2), the probability of survival to time t is $e^{-0.02t}$. The probability of default during a year is the probability of survival to the beginning of the year minus the probability of survival to the end of the year. For example, the probability of survival to time 2 years is $e^{-0.02\times 2} = 0.9608$ and the probability of survival to time 3 years is $e^{-0.02\times 3} = 0.9418$. The probability of default during the third year is $0.9608 - 0.9418 = 0.0190$.

We will assume that defaults always happen halfway through a year and that payments on a five-year credit default swap are made once a year, at the end of each year. We also assume that the risk-free interest rate is 5% per annum with continuous compounding and the recovery rate is 40%. There are three parts to the calculation. These are shown in Tables K.2, K.3, and K.4.

Table K.2 shows the calculation of the present value of the expected payments made on the CDS assuming that payments are made at the rate of s per year and the notional principal is \$1. For example, there is a 0.9418 probability that the third payment of s is made. The expected payment is therefore $0.9418s$ and its present value is $0.9418se^{-0.05\times 3} = 0.8106s$. The total present value of the expected payments in Table K.2 is $4.0728s$.

Table K.1 Unconditional Default Probabilities and Survival Probabilities

Time (years)	Probability of Surviving to Year End	Probability of Default During Year
1	0.9802	0.0198
2	0.9608	0.0194
3	0.9418	0.0190
4	0.9321	0.0186
5	0.9048	0.0183

Table K.2 Calculation of the Present Value of Expected Payments (payment = s per annum)

Time (years)	Probability of Survival	Expected Payment	Discount Factor	PV of Expected Payment
1	0.9802	$0.9802s$	0.9512	$0.9324s$
2	0.9608	$0.9608s$	0.9048	$0.8694s$
3	0.9418	$0.9418s$	0.8607	$0.8106s$
4	0.9231	$0.9231s$	0.8187	$0.7558s$
5	0.9048	$0.9048s$	0.7788	$0.7047s$
Total				$4.0728s$

Table K.3 Calculation of the Present Value of Expected Payoff (notional principal = $1)

Time (years)	Probability of Default	Recovery Rate	Expected Payoff ($)	Discount Factor	PV of Expected Payoff ($)
0.5	0.0198	0.4	0.0119	0.9753	0.0116
1.5	0.0194	0.4	0.0116	0.9277	0.0108
2.5	0.0190	0.4	0.0114	0.8825	0.0101
3.5	0.0186	0.4	0.0112	0.8395	0.0094
4.5	0.0183	0.4	0.0110	0.7985	0.0088
Total					0.0506

Table K.4 Calculation of the Present Value of Accrual Payment

Time (years)	Probability of Default	Expected Accrual Payment	Discount Factor	PV of Expected Accrual Payment
0.5	0.0198	$0.0099s$	0.9753	$0.0097s$
1.5	0.0194	$0.0097s$	0.9277	$0.0090s$
2.5	0.0190	$0.0095s$	0.8825	$0.0084s$
3.5	0.0186	$0.0093s$	0.8395	$0.0078s$
4.5	0.0183	$0.0091s$	0.7985	$0.0073s$
Total				$0.0422s$

Table K.3 shows the calculation of the present value of the expected payoff assuming a notional principal of \$1. As mentioned earlier, we are assuming that defaults always happen halfway through a year. For example, there is a 0.0190 probability of a payoff halfway through the third year. Given that the recovery rate is 40%, the expected payoff at this time is $0.0190 \times 0.6 \times 1 = 0.0114$. The present value of the expected payoff is $0.0114e^{-0.05\times2.5} = 0.0101$. The total present value of the expected payoffs in Table K.3 is \$0.0506.

As a final step, we evaluate in Table K.4 the accrual payment made in the event of a default. These accrual payments arise because the spread payments s are made in arrears so that, when a default occurs, a portion of a spread payment is owed. Consider the third year. There is a 0.0190 probability of a default halfway through the year. An accrual payment of $0.5s$ is due if there is a default. The expected accrual payment halfway through the third year is therefore $0.0190 \times 0.5s = 0.0095s$. Its present value is $0.0095se^{-0.05\times2.5} = 0.0084s$. The total present value of the expected accrual payments is $0.0422s$.

From Tables K.2 and K.4, the present value of the expected payments is

$$4.0728s + 0.0422s = 4.1150s$$

From Table K.3, the present value of the expected payoff is 0.0506. Equating the two, the CDS spread for a new CDS is given by

$$4.1150s = 0.0506$$

or $s = 0.0123$. The mid-market spread should be 0.0123 times the principal or 123 basis points per year. (This is roughly what we would expect from the relationship in equation (19.3): the recovery rate is 40% and the hazard rate is 2% so that equation (19.3) predicts a spread of $0.02 \times 0.6 = 0.012$.)

This example is designed to illustrate the calculation methodology. In practice, we are likely to find that the calculations are more extensive than those in Table K.2 to K.4 because (a) payments are often made more frequently than once a year and (b) we want to assume that defaults can happen more frequently than once a year.

K.1 Marking to Market a CDS

At the time it is negotiated, a CDS, like most other swaps, is worth close to zero. At later times it may have a positive or negative value. Suppose, for example, the credit default swap in our example had been negotiated some time ago for a spread of 150 basis points, the present value of the payments by the buyer would be $4.1150 \times 0.0150 = 0.0617$ and the present value of the payoff would be 0.0506, as given by Table K.3. The value of the swap to the seller would therefore be $0.0617 - 0.0506$ or 0.0111 times the principal. Similarly, the mark-to-market value of the swap to the buyer of protection would be -0.0111 times the principal.

The RMFI software that accompanies this book and can be downloaded from the author's website includes a worksheet for carrying out the above calculations.

Appendix L

Synthetic CDOs and Their Valuation

Synthetic collateralized debt obligations (CDOs) consist of tranches where one party (Party A) agrees to make payments to another party (Party B) that are equal to those losses on a specified portfolio of debt instruments that are in a certain range. In return, Party B agrees to make payments to Party A that are a certain proportion of the amount of principal that is being insured.

Suppose that the range of losses for a particular tranche is from α_L to α_H. The variables α_L and α_H are known as the attachment point and detachment point, respectively. If α_L is 8% and α_H is 18%, Party A pays to Party B the losses on the portfolio, as they are incurred, in the range 8% to 18% of the total principal of the portfolio. The first 8% of losses on the portfolio does not therefore affect the tranche. The tranche is responsible for the next 10% of losses and its notional principal (initially $18 - 8 = 10\%$ of the portfolio principal) reduces as these losses are incurred. The tranche is wiped out when losses exceed 18%. The payments that are made by Party B to Party A are made periodically at a specified rate applied to the remaining notional tranche principal. This specified rate is known as the *tranche spread*.

The usual assumption is that all the debt instruments in the portfolio have the same probability distribution for the time to default. Define $Q(t)$ as the probability of a debt instrument defaulting by time t. The one-factor Gaussian copula model of time to default presented in Section 11.6 has become the standard market model for valuing a tranche of a collateralized debt obligation (CDO). From equation (11.13)

$$Q(t|F) = N\left\{\frac{N^{-1}[Q(t)] - \sqrt{\rho}F}{\sqrt{1-\rho}}\right\} \tag{L.1}$$

where $Q(t|F)$ is the probability of the ith entity defaulting by time t conditional on the value of the factor, F. In the calculation of $Q(t)$ it is usually assumed that the hazard rate for a company is constant. When a CDS spread or other credit spread is available, it can be used to determine the hazard rate using calculations similar to those in Appendix K in conjunction with a search procedure.

Suppose that the hazard rate is λ. Then:

$$Q(t) = 1 - e^{-\lambda t} \tag{L.2}$$

From the properties of the binomial distribution, the probability of exactly k defaults by time t, conditional on F is

$$P(k, t|F) = \frac{n!}{(n-k)!k!}Q(t|F)^k[1 - Q(t|F)]^{n-k} \tag{L.3}$$

Define

$$n_L = \frac{\alpha_L n}{1-R} \qquad \text{and} \qquad n_H = \frac{\alpha_H n}{1-R}$$

where R is the recovery rate (assumed constant). Also, define $m(x)$ as the smallest integer greater than x. The tranche suffers no losses when the number of defaults, k, is less than $m(n_L)$. It is wiped out when k is greater than or equal to $m(n_H)$. Otherwise the tranche principal at time t is a proportion

$$\frac{\alpha_H - k(1-R)/n}{\alpha_H - \alpha_L}$$

of the initial tranche principal. These results can be used in conjunction with equations (L.1), (L.2), and (L.3) to calculate the expected tranche principal at all times conditional on F. We can then integrate over F to find the (unconditional) expected tranche principal. This integration is usually accomplished with a procedure known as Gaussian quadrature. (The author's website provides the tools for integrating over a normal distribution using Gaussian quadrature.)

It is usually assumed that defaults happen at the midpoint of the intervals between payments. Similarly to Appendix K, we are interested in the following quantities:

1. The present value of the expected spread payments received by Party A,
2. The present value of the expected payments for tranche losses made by Party A, and
3. The present value of accrual payments received by Party A.

The spread payments received by Party A at a particular time are linearly dependent on the tranche principal at that time. The tranche loss payments made by Party A (assumed to be at the midpoint of an interval) is the change in the principal during the interval. The accrual payment received by Party A is proportional to the tranche loss payments. For any assumption about spreads, all three quantities of interest can therefore be calculated from the expected tranche principal on payment dates. The break-even spread can therefore be calculated analogously to the way it is calculated for CDSs in Appendix K.

Derivatives dealers calculate an implied copula correlation, ρ, from the spreads quoted in the market for tranches of CDOs and tend to quote these rather than the spreads themselves. This is similar to the practice in options markets of quoting Black–Scholes–Merton implied volatilities rather than dollar prices. Two correlations are used. The compound (tranche) correlation for a tranche is the correlation consistent with the tranche's market value. The base correlation for x% is the correlation consistent with the market value of a tranche with attachment point at zero and detachment point at x%. There is a correlation smile phenomenon in CDO markets similar to the volatility smile phenomenon in options markets.

The RMFI software that accompanies this book and can be downloaded from the author's website includes a worksheet for carrying out the above calculations.[1]

[1] More details on the calculation can be found in J. C. Hull, *Options, Futures, and Other Derivatives*, 10th ed. (Upper Saddle River, NJ: Pearson, 2018).

Answers to Questions and Problems

Chapter 1

1.1 The expected return is 12.5%. The standard deviation of returns is 17.07%.
1.2 From equations (1.1) and (1.2), expected return is 12.5%. SD of return is

$$\sqrt{0.5^2 \times 0.1707^2 + 0.5^2 \times 0.1707^2 + 2 \times 0.15 \times 0.5^2 \times 0.1707^2} = 0.1294$$

or 12.94%.

1.3

w_1	w_2	μ_P	σ_P ($\rho = 0.3$)	σ_P ($\rho = 1$)	σ_P ($\rho = -1$)
0.0	1.0	15%	24.00%	24.00%	24.00%
0.2	0.8	14%	20.39%	22.40%	16.00%
0.4	0.6	13%	17.42%	20.80%	8.00%
0.6	0.4	12%	15.48%	19.20%	0.00%
0.8	0.2	11%	14.96%	17.60%	8.00%
1.0	0.0	10%	16.00%	16.00%	16.00%

1.4 Nonsystematic risk can be diversified; systematic risk cannot. Systematic risk is more important to an equity investor. Either type of risk can lead to the bankruptcy of a corporation.

1.5 We assume that investors trade off mean return and standard deviation of return. For a given mean return, they want to minimize standard deviation of returns. All make the same estimates of means, standard deviations, and coefficients of correlation for returns on individual investments. Furthermore, they can borrow or lend at the risk-free rate. The result is that they all want to be on the "new efficient frontier" in Figure 1.4. They choose the same portfolio of risky investments combined with borrowing or lending at the risk-free rate.

1.6 (a) 7.2%, (b) 9%, (c) 14.4%.

1.7 The capital asset pricing model assumes that there is one factor driving returns. Arbitrage pricing theory assumes multiple factors.

1.8 In many jurisdictions, interest on debt is deductible to the corporation whereas dividends are not deductible. It can therefore be more tax-efficient for a company to fund itself with debt. However, as debt increases, the probability of bankruptcy increases.

1.9 Risk decomposition refers to a procedure where risks are handled one by one. Risk aggregation refers to a procedure where a portfolio of risks is considered. Risk decomposition requires an in-depth understanding of individual risks. Risk aggregation requires an understanding of the correlations between risks.

1.10 When potential losses are large, we cannot aggregate them and assume they will be diversified away. It is necessary to consider them one by one and handle them with insurance contracts, tighter internal controls, and the like.

1.11 This is the probability that profit is no worse than –4% of assets. This profit level is $4.6/1.5 = 3.067$ standard deviations from the mean. The probability that the bank will have a positive equity is therefore $N(3.067)$ where N is the cumulative normal distribution function. This is 99.89%.

1.12 Banks are allowed to accept deposits from the general public. Companies in retailing and manufacturing are not.

1.13 Professional fees ($10 million per month), lost sales (people are reluctant to do business with a company that is being reorganized), and loss of key senior executives (lack of continuity).

1.14 The return earned by the hedge fund manager with zero alpha would be $0.05 + 0.6 \times (0.10 - 0.05) = 0.08$ or 8%. Because the alpha equals 4%, the hedge fund manager's return was 8% plus 4%, or 12%.

Chapter 2

2.1 The banking system became more concentrated, with large banks having a bigger share of the market. The total number of banks decreased from 14,483 to 5,060.

2.2 In the early twentieth century, many states passed laws restricting banks from opening more than one branch. The McFadden Act of 1927 restricted banks from opening branches in more than one state.

2.3 The main risk is that interest rates will rise so that, when deposits are rolled over, the bank has to pay a higher rate of interest. The result will be a reduction in the bank's net interest income. There is also liquidity risk.

2.4 DLC's loss is more than its equity capital, and it would probably be liquidated. The subordinated long-term debt holders would incur losses on their $5 million investment. The depositors should get their money back.

2.5 The net interest income of a bank is interest received minus interest paid.

2.6 Credit risk primarily affects loan losses. Non-interest income includes trading gains and losses. Market risk therefore affects non-interest income. It also affects net interest income if assets and liabilities are not matched. Operational risk primarily affects non-interest expense.

2.7 A private placement is a new issue of securities that is sold to a small number of large institutional investors. A public offering is a new issue of securities that is offered to the general public. In a best efforts deal, the investment bank does as well as it can to place securities with investors, but does not guarantee that they can be sold. In a firm commitment deal, the investment bank agrees to buy the securities from the issuing company for a particular price and attempts to sell them in the market for a higher price.

2.8 The bidders when ranked from the highest price bid to the lowest are: H, C, F, A, B, D, E, and G. Bidders H, C, and F have bid for 140,000 shares. A has bid for 20,000. The price that clears the market is the price that was bid by A, or \$100. H, C, and F get their orders filled at this price. Half of A's order is filled at this price.

2.9 A Dutch auction potentially attracts a wide range of bidders. If all interested market participants bid, the price paid should be close to the market price immediately after the IPO. The usual IPO situation where the price turns out to be well below the market price should therefore be avoided. Also, investment banks are not able to restrict purchasers to their best current and potential clients. The Google IPO was different from a standard Dutch auction in that Google reserved the right to choose the number of shares that would be issued, and the percentage allocated to each bidder, when it saw the bids.

2.10 Poison pills can give management a negotiation tool, particularly if the board has the right to overturn a poison pill or make it ineffective. When it is confronted with a potential acquirer, the poison pill can buy the company time to bargain for a better purchase price or find other bidders. However, there is the danger that the poison pill will discourage potential buyers from approaching the company in the first place.

2.11 The brokerage subsidiary of a bank might recommend securities that the investment banking subsidiary is trying to sell. The commercial banking subsidiary might pass confidential information about its clients to the investment banking subsidiary. When a bank does business with a company (or wants to do business with the company), it might persuade the brokerage subsidiary to recommend the company's shares as a "buy." The commercial banking subsidiary might persuade a company to which it has lent money to do a bond issue because it is worried about its exposure to the client. (It wants the investment banking subsidiary to persuade its clients to take on the credit risk.) These conflicts of interest are handled by what are known as Chinese walls. They prevent the flow of information from one part of the bank to another.

2.12 The trading book includes the instruments that a bank has as a result of its trading activities. The banking book consists of instruments it does not expect to trade, such as loans to corporations and individuals.

2.13 As a result of IFRS 9 and similar accounting standards from FASB, banks have to record loans they have made in their financial statements at the amount owed less an estimate of expected losses.

2.14 In the originate-to-distribute model, a bank originates loans and then securitizes them so that they are passed on to investors. This was done extensively with household mortgages during the seven-year period leading up to July 2007. In July 2007, investors lost confidence in the securitized products, and banks were forced to abandon the originate-to-distribute model, at least temporarily.

Chapter 3

3.1 Term life insurance lasts a fixed period (e.g., five years or ten years). The policyholder pays premiums. If the policyholder dies during the life of the policy, the policyholder's beneficiaries receive a payout equal to the principal amount of the policy. Whole life insurance lasts for the whole life of the policyholder. The policyholder pays premiums (usually the same each year), and the policyholder's beneficiaries receive a payout equal to the principal amount of the policy when the policyholder dies. There is an investment element to whole life insurance because the premiums in early years are high relative to the expected payout in those years. (The reverse is true in later years.)

3.2 Variable life insurance is whole life insurance where the policyholder can specify how the funds generated in early years (the excess of the premiums over the actuarial cost of the insurance) are invested. There is a minimum payout on death, but the payout can be more than the minimum if the investments do well. Universal life insurance is whole life insurance where the premium can be reduced to a specified minimum level without the policy lapsing. The insurance company chooses the investments (generally fixed income) and guarantees a minimum return. If the investments do well, the return provided on the policyholder's death may be greater than the guaranteed minimum.

3.3 Annuity contracts have exposure to longevity risk. Life insurance contracts have exposure to mortality risk.

3.4 The lifetime annuity created from an accumulated value was calculated using an interest rate that was the greater of (a) the market interest rate and (b) a prespecified minimum interest rate.

3.5 The probability that the woman will die during the first year is 0.003182. The probability that the woman will die during the second year is $0.003473 \times (1 - 0.003182) = 0.003462$. Suppose that the break-even premium is X. We must have

$$1{,}000{,}000 \times (0.003182 + 0.003462) = X + (1 - 0.003182)X$$

so that $X = 3{,}327$. The break-even premium is therefore \$3,327.

3.6 The probability of a male surviving to 30 is 0.97519. The probability of a male surviving to 90 is 0.17735. The probability of a male surviving to 90 conditional that

30 is reached is therefore 0.17735/0.97519 = 0.18186. The probability of a female surviving to 90 conditional that 30 is reached is 0.29104/0.98635 = 0.29507.

3.7 The biggest risks are those arising from catastrophes such as earthquakes and hurricanes and those arising from liability insurance (e.g., claims related to asbestos in the United States). This is because there is no "law of large numbers" working in the insurance company's favor. Either the event happens and there are big payouts or the event does not happen and there are no payouts.

3.8 CAT bonds (catastrophe bonds) are an alternative to reinsurance for an insurance company that has taken on a certain catastrophic risk (e.g., the risk of a hurricane or an earthquake) and wants to get rid of it. CAT bonds are issued by the insurance company. They provide a higher rate of interest than risk-free bonds. However, the bondholders agree to forgo interest, and possibly principal, to meet any claims against the insurance company that are within a prespecified range.

3.9 The CAT bond has very little systematic risk. Whether a particular type of catastrophe occurs is independent of the return on the market. The risks in the CAT bond are to some extent diversified away by the other investments in the portfolio. A B-rated bond does have systematic risk so that less of its risks are diversified away. It is likely therefore that the CAT bond is a better addition to the portfolio.

3.10 In Canada and the United Kingdom, health care is provided by the government. In the United States, publicly funded health care is limited and most individuals buy private health care insurance of one sort or another. In the United Kingdom, a private health care system operates alongside the public system.

3.11 Both moral hazard and adverse selection are potential problems. The insurance might lead to an individual not trying to keep a job as much as he or she otherwise would. Indeed, an individual might purposely lose his or her job to collect the insurance payout! Also, individuals who are most at risk for losing their jobs would be the ones who would choose to buy the insurance.

3.12 The payouts of property–casualty insurers show more variability than the payouts of life insurers. This is because of the possibility of catastrophes such as earthquakes and hurricanes and liability insurance claims such as those related to asbestos in the United States.

3.13 The loss ratio is the ratio of payouts to premiums in a year. The expense ratio is the ratio of expenses (e.g., sales commissions and expenses incurred in validating losses) to premiums in a year. The statement is not true because investment income can be significant. Premiums are received at the beginning of a year, and payouts are made during the year or after the end of the year.

3.14 A defined contribution plan is a plan where the contributions of each employee (together with contributions made by the employer for that employee) are kept in a separate account and invested for the employee. When retirement age is reached, the accumulated amount is usually converted into an annuity. In a defined benefit plan, all contributions for all employees are pooled and invested. Employees receive a predefined pension that is based on their years of employment and final salary. At any given time, a defined benefit plan may be in surplus or in deficit.

3.15 The employee's wages are constant in real terms. Suppose that they are X per year. (The units for X do not matter for the purposes of our calculation.) The pension is $0.75X$. The real return earned is zero. Because employees work for 40 years, the present value of the contributions made by one employee is $40XR$ where R is the contribution rate as a percentage of the employee's wages. The present value of the benefits is $20 \times 0.75X = 15X$. The value of R that is necessary to adequately fund the plan must therefore satisfy

$$40XR = 15X$$

The solution to this equation is $R = 0.375$. The total of the employee contributions (and employer contributions, if any) should therefore be 37.5% of salary.

Chapter 4

4.1 The number of shares of an open-end mutual fund increases as investments in the fund increase and decreases as investors withdraw their funds. A closed-end fund is like any other corporation with a fixed number of shares that trade.

4.2 The net asset value (NAV) of an open-end mutual fund is calculated at 4 P.M. each day as the value of the assets held by the fund divided by the number of shares outstanding.

4.3 The investor is deemed to have made capital gains of $300 and $100 in 2018 and 2019, respectively. In 2020, the investor is deemed to have made a capital loss of $200.

4.4 An index fund is a fund that is designed so that its value tracks the performance of an index such as the S&P 500. It can be created by buying all the stocks (or a representative subset of the stocks) that underlie the index. Sometimes futures contracts on the index are used.

4.5 The front-end load is the amount an investor pays, as a percentage of his or her investment, when shares of the fund are purchased. The back-end load is the amount an investor pays, as a percentage of his or her investment, when shares of the fund are redeemed.

4.6 An exchange-traded fund (ETF) that tracks an index is created when an institutional investor deposits a portfolio of shares that is designed to track the index and receives shares in the ETF. Institutional investors can at any time exchange shares in the ETF for the underlying shares held by the ETF, or vice versa. The advantages over an open-end mutual fund that tracks the index are that the fund can be traded at any time, the fund can be shorted, and the fund does not have to be partially liquidated to accommodate redemptions. The advantage over a closed-end mutual fund is that there is very little difference between the ETF share price and the net asset value per share of the fund.

4.7 The arithmetic mean of a set of n numbers is the sum of the numbers divided by n. The geometric mean is the nth root of the product of the numbers. The arithmetic mean is always greater than or equal to the geometric mean. The return per year realized when an investment is held for several years is calculated using a geometric mean, not an arithmetic mean. (The procedure is to calculate the geometric mean of one plus the return in each year and then subtract one.)

4.8 Late trading is the illegal practice of putting in an order to buy or sell an open-end mutual fund at the 4 P.M. price after 4 P.M. Market timing is a practice where favored clients are allowed to buy and sell a mutual fund frequently to take advantage of the fact that some prices used in the calculation of the 4 P.M. net asset value are stale. Front running is the practice of trading by individuals ahead of a large institutional trade that is expected to move the market. Directed brokerage describes the situation where a mutual fund uses a brokerage house for trades when the brokerage house recommends the fund to clients.

4.9 Mutual funds must disclose their investment policies; their use of leverage is limited; they must calculate NAV daily; their shares must be redeemable at any time.

4.10 If a hedge fund is making money out of trading convertible bonds, it must be doing so at the expense of its counterparties. If most of the traders are hedge funds, they cannot all be making money.

4.11 Hurdle rate is the minimum return necessary for an incentive fee to be applicable. High-water mark refers to the previous losses that must be recouped before incentive fees are applicable. Clawback refers to investors being able to use some of the past incentive fees they have paid as an offset to current losses.

4.12 If the return is $X(> 2\%)$, the investors pay $0.02 + 0.2(X - 0.02)$ in fees. It must therefore be the case that

$$X - 0.02 - 0.2(X - 0.02) = 0.2$$

so that $0.8X = 0.216$ or $X = 0.27$. A return of 27% is necessary.

4.13 Short-term gains and losses do matter if the hedge fund is highly leveraged. Short-term losses can lead to margin calls that can destroy the hedge fund.

4.14 The leverage a hedge fund is allowed to take is limited by its prime broker. This in turn influences the risks that the hedge fund can take.

Chapter 5

5.1 When a trader enters into a long forward contract, she is agreeing to *buy* the underlying asset for a certain price at a certain time in the future. When a trader enters into a short forward contract, she is agreeing to *sell* the underlying asset for a certain price at a certain time in the future.

5.2 A trader is *hedging* when she has an exposure to the price of an asset and takes a position in a derivative to offset the exposure. In a *speculation*, the trader has no exposure to offset. She is betting on the future movements in the price of the asset. *Arbitrage* involves taking a position in two or more different markets to lock in a profit.

5.3 In the first case, the trader is obligated to buy the asset for \$50. (The trader does not have a choice.) In the second case, the trader has an option to buy the asset for \$50. (The trader does not have to exercise the option.)

5.4 Selling a call option involves giving someone else the right to buy an asset from you for a certain price. Buying a put option gives you the right to sell the asset to someone else.

5.5 **(a)** The investor is obligated to sell pounds for 1.3000 when they are worth 1.2900. The gain is $(1.3000 - 1.2900) \times 100{,}000 = \$1{,}000$.

(b) The investor is obligated to sell pounds for 1.3000 when they are worth 1.3200. The loss is $(1.3200 - 1.3000) \times 100{,}000 = \$2{,}000$.

5.6 **(a)** The trader sells for 50 cents per pound something that is worth 48.20 cents per pound. Gain $= (\$0.5000 - \$0.4820) \times 50{,}000 = \$900$.

(b) The trader sells for 50 cents per pound something that is worth 51.30 cents per pound. Loss $= (\$0.5130 - \$0.5000) \times 50{,}000 = \$650$.

5.7 You have sold a put option. You have agreed to buy 100 shares for \$40 per share if the party on the other side of the contract chooses to exercise the right to sell for this price. The option will be exercised only when the price of the stock is below \$40. Suppose, for example, that the option is exercised when the price is \$30. You have to buy at \$40 shares that are worth \$30; you lose \$10 per share, or \$1,000 in total. If the option is exercised when the price is \$20, you lose \$20 per share, or \$2,000 in total. The worst that can happen is that the price of the stock declines to almost zero during the three-month period. This highly unlikely event would cost you \$4,000. In return for the possible future losses, you receive the price of the option from the purchaser.

5.8 The over-the-counter (OTC) market is a market where financial institutions, fund managers, and corporate treasurers get in touch with each other directly and can enter into relatively large mutually acceptable transactions. An exchange-traded market is a market organized by an exchange where traders either meet physically or communicate electronically and the contracts that can be traded have been defined by the exchange. (a) OTC, (b) exchange, (c) both, (d) OTC, (e) OTC.

5.9 One strategy would be to buy 200 shares. Another would be to buy 2,000 options. If the share price does well, the second strategy will give rise to greater gains. For example, if the share price goes up to \$40, you gain $[2{,}000 \times (\$40 - \$30)] - \$5{,}800 = \$14{,}200$ from the second strategy and only $200 \times (\$40 - \$29) = \$2{,}200$ from the first strategy. However, if the share price does badly, the second strategy gives greater losses. For example, if the share price goes down to \$25, the first strategy leads to a loss of $200 \times (\$29 - \$25) = \$800$, whereas the second strategy leads to a loss of the whole \$5,800 investment. This example shows that options contain built-in leverage.

5.10 You could buy 5,000 put options (or 50 contracts) with a strike price of \$25 and an expiration date in four months. This provides a type of insurance. If, at the end of four months, the stock price proves to be less than \$25, you can exercise the options and sell the shares for \$25 each. The cost of this strategy is the price you pay for the put options.

5.11 A stock option provides no funds for the company. It is a security sold by one trader to another. The company is not involved. By contrast, a stock when it is first issued is a claim sold by the company to investors and does provide funds for the company.

5.12 Ignoring the time value of money, the holder of the option will make a profit if the stock price in March is greater than \$52.50. This is because the payoff to the holder of the option is, in these circumstances, greater than the \$2.50 paid for the option. The option will be exercised if the stock price at maturity is greater than \$50.00. Note that, if the stock price is between \$50.00 and \$52.50, and the option is exercised, the holder of the option takes a loss overall.

5.13 Ignoring the time value of money, the seller of the option will make a profit if the stock price in June is greater than \$56.00. This is because the cost to the seller of the option is in these circumstances less than the price received for the option. The option will be exercised if the stock price at maturity is less than \$60.00. Note that if the stock price is between \$56.00 and \$60.00, the seller of the option makes a profit even though the option is exercised.

5.14 A long position in a four-month put option can provide insurance against the exchange rate falling below the strike price. It ensures that the foreign currency can be sold for at least the strike price.

5.15 The company could enter into a long forward contract to buy 1 million Canadian dollars in six months. This would have the effect of locking in an exchange rate equal to the current forward exchange rate. Alternatively, the company could buy a call option giving it the right (but not the obligation) to purchase 1 million Canadian dollars at a certain exchange rate in six months. This would provide insurance against a strong Canadian dollar in six months while still allowing the company to benefit from a weak Canadian dollar at that time.

5.16 The payoff from an ICON is the payoff from:

(a) A regular bond

(b) A short position in call options to buy 169,000 yen with an exercise price of 1/169

(c) A long position in call options to buy 169,000 yen with an exercise price of 1/84.5

This is demonstrated by the following table:

	Terminal Value of Regular Bond	Terminal Value of Short Calls	Terminal Value of Long Calls	Terminal Value of Whole Position
$S_T > 169$	1,000	0	0	1,000
$84.5 \le S_T \le 169$	1,000	$-169{,}000\left(\frac{1}{S_T}-\frac{1}{169}\right)$	0	$2{,}000-\frac{169{,}000}{S_T}$
$S_T < 84.5$	1,000	$-169{,}000\left(\frac{1}{S_T}-\frac{1}{169}\right)$	$169{,}000\left(\frac{1}{S_T}-\frac{1}{84.5}\right)$	0

5.17 **(a)** The trader buys a 180-day call option and takes a short position in a 180-day forward contract.

(b) The trader buys 90-day put options and takes a long position in a 90-day forward contract.

5.18 It enters into a five-year swap where it pays 3.30% and receives LIBOR. Its investment is then at LIBOR minus 0.30%.

5.19 It enters into a five-year swap where it receives 3.26% and pays LIBOR. Its net cost of borrowing is LIBOR plus 1.74%.

5.20 It enters into a three-year swap where it receives LIBOR and pays 3%. Its net borrowing cost for the three years is then 4% per annum.

5.21 Suppose that the weather is bad and the farmer's production is lower than expected. Other farmers are likely to have been affected similarly. Corn production overall will be low and as a consequence the price of corn will be relatively high. The farmer is likely to be overhedged relative to actual production. The farmer's problems arising from the bad harvest will be made worse by losses on the short futures position. This problem emphasizes the importance of looking at the big picture when hedging. The farmer is correct to question whether hedging price risk while ignoring other risks is a good strategy.

5.22 It may well be true that there is just as much chance that the price of oil in the future will be above the futures price as that it will be below the futures price. This means that the use of a futures contract for speculation would be like betting on whether a coin comes up heads or tails. But it might make sense for the airline to use futures for hedging rather than speculation. The futures contract then has the effect of reducing risks. It can be argued that an airline should not expose its shareholders to risks associated with the future price of oil when there are contracts available to hedge the risks.

5.23 Microsoft is choosing an option on a portfolio of assets instead of the corresponding portfolio of options. The former is always less expensive because there is the potential for an increase in the price of one asset to be netted off against a decrease in the price of another asset. Compare (a) an option with a strike price of $20 on a portfolio of two assets each worth $10 and (b) a portfolio of two options with a strike price of $10, one on each of the assets. If both assets increase in price or both assets decrease in price, the payoffs are the same. But if one decreases and the other increases, the payoff from (a) is less than that from (b). Both the Asian feature and the basket feature in Microsoft's options help to reduce the cost of the options because of the possibility of gains and losses being netted.

5.24 It means that the price of the energy source can go up or down but will tend over time to get pulled back to its long-run average level. Electricity has the highest rate of mean reversion; oil has the lowest.

5.25 As we increase the frequency with which the asset price is observed, the asset price becomes more likely to hit the barrier and the value of a knock-out call goes down.

5.26 The average of the highest and lowest temperature each day is 75° Fahrenheit. The CDD each day is therefore 10 and the cumulative CDD for the month is

$10 \times 31 = 310$. The payoff from the call option is therefore $(310 - 250) \times 5{,}000 =$ \$300,000.

5.27 A 5×8 contract for May 2019 is a contract to provide electricity for five days per week during the off-peak period (11 P.M. to 7 A.M.). When daily exercise is specified, the holder of the option is able to choose each weekday whether he or she will buy electricity at the strike price. When there is monthly exercise, he or she chooses once at the beginning of the month whether electricity is to be bought at the strike price for the whole month. The option with daily exercise is worth more.

5.28 Two calculations are necessary to determine the initial margin. The first gives

$$500 \times (3.5 + 0.2 \times 57 - 3) = 5{,}950$$

The second gives

$$500 \times (3.5 + 0.1 \times 57) = 4{,}600$$

The initial margin is the greater of these, or \$5,950. Part of this can be provided by the initial amount of \$500 × 3.5 = \$1,750 received for the options.

5.29 The cost of the shares is \$500 × 50 or \$25,000. When shares are shorted, the proceeds of the sale form part of the margin. In this case, the total margin required is 1.6 × \$25,000, or \$40,000. The extra margin required is therefore \$15,000. This can be in the form of cash or marginable securities. When the share price rises to S, the value of the underlying position is $500S$. There is a margin call when

$$40{,}000 < 1.3 \times 500S \text{ or } S > 61.54.$$

5.30 A broker sets an initial margin and a maintenance margin. When the balance in the margin account falls below the maintenance margin level, the broker's client is required to bring it back up to the initial margin level. The exchange clearing house, when dealing with its members, sets the initial margin and maintenance margin to be the same. Every day the balance in the margin account must be at this level or more.

Chapter 6

6.1 Mortgages were frequently securitized. The only information that was retained during the securitization process was the applicant's FICO score and the loan-to-value ratio of the mortgage. The originators knew this and so this was the only information that they cared about.

6.2 There was a short-term imbalance between supply and demand because many people were persuaded to take out mortgages they could not afford.

6.3 When the loss rate on the mortgages is 5%, there are no losses on the mezzanine tranche of the ABS and therefore there are no losses on any of the tranches of the ABS CDO. When the loss rate on the mortgages is 12%, the loss on the mezzanine tranche of the ABS is 7/20 or 35%. The loss rate on the equity and mezzanine tranches of the ABS CDO is 100%. The loss rate on the senior tranche of the ABS CDO is 10/75 = 13.333%.

6.4 Often a tranche is thin and the probability distribution of the loss is quite different from that on the bond. If a loss occurs, there is a high probability that it will be 100%. A 100% loss is much less likely for a bond.

6.5 An ABS is a set of tranches created from a portfolio of loans, bonds, credit card receivables, and so on. An ABS CDO is an ABS created from particular tranches (e.g., the BBB-rated tranches) of a number of different ABSs.

6.6 Investors underestimated how high the default correlations between mortgages would be in stressed market conditions. Investors also did not always realize that the tranches underlying ABS CDOs were usually quite thin so that they were either totally wiped out or untouched. There was an unfortunate tendency to assume that a tranche with a particular rating could be considered to be the same as a bond with that rating. This assumption was not valid for the reasons just mentioned.

6.7 *Agency costs* is a term used to describe the costs in a situation where the interests of two parties are not perfectly aligned. The incentives of traders, originators, valuers, the creators of the structured products, and the rating agencies arguably created agency costs.

6.8 The waterfall defines how the interest and principal cash flows from the underlying portfolio are distributed to the tranches. In a typical arrangement, interest cash flows are first used to pay the most senior tranche its promised return on its outstanding principal. The cash flows (if any) that are left over are used to provide the next most senior tranche with its promised return on its outstanding principal, and so on. Principal cash flows are usually used first to repay the most senior tranche, then the next most senior tranche, and so on. The equity tranche receives principal and interest payments only when more senior tranches have been paid.

6.9 An ABS CDO is created from tranches of ABSs (e.g., BBB tranches). A motivation on the part of the originator was that it could be difficult to find investors for the tranches directly.

6.10 Mian and Sufi showed that regions where mortgage application denials were highest in the United States in 1996 were also regions where mortgage origination grew particularly fast during the 2000 to 2006 period.

6.11 The mezzanine tranche of an ABS or ABS CDO is a tranche that is in the middle as far as seniority goes. It ranks below the senior tranches and therefore absorbs losses before they do. It ranks above the equity tranche (so that the equity tranche absorbs losses before it does).

6.12 As default correlation increases, the senior tranche of an ABS becomes more risky because it is more likely to suffer losses. The equity tranche becomes less risky as the default correlation increases. To understand this, note that in the limit when there is perfect correlation (and assuming no recoveries), all tranches have the same loss experience because either (a) all companies default or (b) no companies default. As we move toward the perfect correlation situation, the tranches become more similar so that the senior tranche becomes more risky and the equity tranche becomes less risky.

6.13 The end-of-year bonus reflects performance during the year. Traders and other employees of banks are liable to be always focusing on their next end-of-year bonus (and not on possible later effects of their decisions) and therefore have a short-term time horizon for their decision making. A trend toward deferred bonuses that can be clawed back is changing this to some extent.

Chapter 7

7.1 From equation (7.3), the probability is $N(d_2)$ where

$$d_2 = \frac{\ln(50/70) + (0.12 - 0.2^2/2) \times 2}{0.2\sqrt{2}} = -0.4825$$

The probability is therefore $N(-0.4825) = 0.315$.

7.2 From equation (7.4), this is

$$50 \times \exp[(0.12 - 0.2^2/2) \times 2 - N^{-1}(0.05) \times 0.2 \times \sqrt{2}] = 97.25$$

or \$97.25.

7.3 Risk-neutral valuation states that, if we value a derivative on the assumption that investors are risk-neutral (i.e., investors do not require a higher expected return for bearing risks), we get the right answer for all worlds (including the real world).

7.4 The expected value is higher in the real world. The higher value reflects the compensation for risk required by the investor.

7.5 By buying both derivatives, one is certain to earn \$2 on an initial investment of \$98. The risk-free rate is therefore $2/98 = 2.04\%$ with annual compounding. The risk-neutral probability of default is π where $100\pi/(1 + 0.0204) = 3$ or $\pi = 3.0612\%$.

7.6 The risk-neutral probability of the payoff is given by equation (7.3) with $V = 30$, $S_0 = 25$, $\sigma = 0.3$, $\mu = 0.03$, and $T = 0.25$. It is

$$N\left[\frac{(25/30) + (0.03 - 0.3^2/2) \times 0.25}{0.3 \times \sqrt{0.25}}\right] = 0.1074$$

The value of the option is therefore $100 \times 0.1074 \times e^{-0.03\times0.25} = \$10.66$. For the real-world probability of a payoff, we set $\mu = 0.1$ instead of 0.03 in the equation to get 0.1305.

7.7 The real world is used to simulate possible values of the underlying variables in one year. The risk-neutral world is then used to calculate the value of the portfolio at the one-year point for each simulation trial.

7.8 *Mean reversion* is the term used to describe the tendency for the values of variables such as interest rates and commodity prices to be pulled back to a central value. Of course, an uncertain random movement is superimposed on this pull-back.

7.9 Girsanov's theorem tells us that when we move from the real world to the risk-neutral world or vice versa, the expected growth rates of market variables change but their volatilities remain the same.

Chapter 8

8.1 The value of the portfolio decreases by \$10,500.

8.2 The value of the portfolio increases by \$400.

8.3 In both cases, it increases by $0.5 \times 30 \times 2^2$ or \$60.

8.4 A delta of 0.7 means that, when the price of the stock increases by a small amount, the price of the option increases by 70% of this amount. Similarly, when the price of the stock decreases by a small amount, the price of the option decreases by 70% of this amount. A short position in 1,000 options has a delta of −700 and can be made delta neutral with the purchase of 700 shares.

8.5 A theta of −100 per day means that if one day passes with no change in either the stock price or its volatility, the value of the option position declines by \$100. If a trader feels that neither the stock price nor its implied volatility will change, she should write an option with as big a negative theta as possible so that her position has a high positive theta. Relatively short-life, at-the-money options have the biggest negative theta.

8.6 The gamma of an option position is the rate of change of the delta of the position with respect to the asset price. For example, a gamma of 0.1 would indicate that, when the asset price increases by a certain small amount, delta increases by 0.1 of this amount. When the gamma of an option writer's position is large and negative and the delta is zero, the option writer will lose significant amounts of money if there is a large movement (either an increase or a decrease) in the asset price.

8.7 To hedge an option position, it is necessary to create the opposite option position synthetically. For example, to hedge a long position in a put, it is necessary to create a short position in a put synthetically. It follows that the procedure for creating an option position synthetically is the reverse of the procedure for hedging the option position.

8.8 A long position in either a put or a call option has a positive gamma. From Figure 8.9, when gamma is positive, the hedger gains from a large change in the stock price and loses from a small change in the stock price. Hence the hedger will fare better in case (b). When the portfolio contains short option positions, the hedger will fare better in case (a).

8.9 The delta indicates that when the value of the euro exchange rate increases by \$0.01, the value of the bank's position increases by $0.01 \times 30{,}000 = \$300$. The gamma indicates that when the euro exchange rate increases by \$0.01, the delta of the portfolio decreases by $0.01 \times 80{,}000 = 800$. For delta neutrality, 30,000 euros should be shorted. When the exchange rate moves up to 0.93, we expect the delta of the portfolio to decrease by $(0.93 - 0.90) \times 80{,}000 = 2{,}400$ so that it becomes 27,600. To maintain delta neutrality, it is therefore necessary for the bank to unwind its short position by 2,400 euros so that a net 27,600 euros have been shorted. When a portfolio is delta neutral and has a negative gamma, a loss is experienced when there is a large movement in the underlying asset price. We can conclude that the bank is likely to have lost money.

8.10 When used in the way described in the text, it does assume volatility is constant. In theory, we could implement a static options replication strategy where there are three dimensions: time, the stock price, and volatility. Prices are then matched on a surface in the three-dimensional space.

8.11 Ten regular options are likely to be needed. This is because there are 10 equations to be satisfied, one for each point on the boundary.

8.12 The payoff from an Asian option becomes more certain with the passage of time. As a result, the amount of uncertainty that needs to be hedged decreases with the passage of time.

8.13 Consider a portfolio of options dependent on a single market variable. A single trade is all that is necessary to make the position delta neutral regardless of the number of options in the portfolio.

8.14 The price, delta, gamma, vega, theta, and rho are 0.0217, –0.396, 5.415, 0.00203, –0.0000625, and –0.00119. Delta predicts that the option price should decrease by approximately 0.000396 when the exchange rate increases by 0.001. This is what we find. When the exchange rate is increased to 0.751, the option price decreases to 0.0213.

Chapter 9

9.1 In this case, the interest rate mismatch is \$10 billion. The bank's net interest income declines by \$100 million each year for the next three years.

9.2 If long-term rates were simply a reflection of expected future short-term rates, we would expect long rates to be less than short rates as often as they are greater than short rates. (This is based on the assumption that half of the time investors expect rates to increase, and half of the time investors expect rates to decrease.) Liquidity preference theory argues that long-term rates are high relative to expected future short-term rates. This means that long rates are greater than short rates most of the time. When long rates are less than short rates, the market is expecting a relatively steep decline in rates.

9.3 There are two reasons:

1. The amount of capital a bank is required to hold to support an investment in Treasury bills and bonds (typically zero) is substantially smaller than the capital required to support a similar investment in other very-low-risk instruments.
2. In the United States, Treasury instruments are given a favorable tax treatment compared with most other fixed-income investments because they are not taxed at the state level.

9.4 In an overnight indexed swap, the geometric average of the federal funds rate for a period such as three months is exchanged for a prespecified fixed rate.

9.5 The LIBOR-OIS spread is a measure of the reluctance of banks to lend to each other.

9.6 Duration provides information about the effect of a small parallel shift in the yield curve on the value of a bond portfolio. The percentage decrease in the value of the portfolio equals the duration of the portfolio multiplied by the amount by which interest rates are increased in the small parallel shift. Its main limitation is that it applies only to parallel shifts in the yield curve that are small.

9.7 (a) The bond's price is 86.80, (b) the bond's duration is 4.256 years, (c) the duration formula shows that when the yield decreases by 0.2% the bond's price increases by 0.74, and (d) recomputing the bond's price with a yield of 10.8% gives a price of 87.54, which is approximately consistent with (a) and (c).

9.8 (a) The bond's price is 88.91, (b) the bond's modified duration is 3.843 years, (c) the duration formula shows that when the yield decreases by 0.2% the bond's price increases by 0.68, and (d) recomputing the bond's price with a yield of 10.8% (annually compounded) gives a price of 89.60, which is approximately consistent with (a) and (c).

9.9 The bond's price is 104.80. The duration of the bond is 5.35. The convexity is 30.60. The effect of a 1% increase in the yield is estimated by equation (9.4) as

$$104.80 \times (-0.01 \times 5.35 + 0.5 \times 30.60 \times 0.0001) = -5.44$$

The bond's price actually changes to 99.36, which is consistent with the estimate.

9.10 We can (a) perturb points on the yield curve (see Figure 9.5), (b) perturb sections of the yield curve (see Figure 9.7), and (c) perturb the market quotes used to create the yield curve.

9.11 The deltas (changes in portfolio value per unit of factor with factor loading being assumed to be in basis points) are –5.64 and 225.63.

9.12 The impact on the portfolio, measured as a proportion of the value of the portfolio, is

$$\begin{aligned}&-(0.2 \times 0.001 + 0.6 \times 0.0008 + 0.9 \times 0.0007 + 1.6 \times 0.0006\\&+2.0 \times 0.0005 - 2.1 \times 0.0003 - 3.0 \times 0.0001) = -0.00234\end{aligned}$$

The portfolio decreases by 0.234%.

9.13 Dollar duration is defined as the product of the duration of a portfolio and its value. Dollar convexity is defined as the product of the convexity of a portfolio and its value.

9.14 The partial durations add up to the total duration. The DV01 is the total duration times the portfolio value times 0.0001.

Chapter 10

10.1 $2 \times \sqrt{3}$, or 3.46%.

10.2 The standard deviation of the percentage price change in one day is $25/\sqrt{252}$ or 1.57%, and 95% confidence limits are from –3.09% to +3.09%.

10.3 Volatility is much higher when markets are open than when they are closed. Traders therefore measure time in trading days rather than calendar days.

10.4 Implied volatility is the volatility that leads to the option price equaling the market price when Black–Scholes–Merton assumptions are used. It is found by trial and error. Because different options have different implied volatilities, traders are not using the same assumptions as Black–Scholes–Merton. (See Chapter 25 for a further discussion of this.)

10.5 The standard formula for calculating standard deviation gives 0.547% per day. The simplified approach in equation (10.4) gives 0.530% per day.

10.6 The power law gives $0.01 = K \times 500^{-2}$ so that $K = 2{,}500$. (a) $2{,}500 \times 1{,}000^{-2} = 0.0025$ or 0.25%, and (b) $2{,}500 \times 2{,}000^{-2} = 0.000625$ or 0.0625%.

10.7 The variance rate estimate calculated at the end of day $n - 1$ for day n equals λ times the variance rate calculated at the end of day $n - 2$ for day $n - 1$ plus $1 - \lambda$ times the squared return on day $n - 1$. This is equivalent to assigning weights to squared returns that decline exponentially as we move back through time.

10.8 GARCH(1,1) adapts the EWMA model by giving some weight to a long-run average variance rate. Whereas the EWMA has no mean reversion, GARCH(1,1) is consistent with a mean-reverting variance rate model.

10.9 In this case, $\sigma_{n-1} = 0.015$ and $u_{n-1} = 0.5/30 = 0.01667$, so that equation (10.8) gives

$$\sigma_n^2 = 0.94 \times 0.015^2 + 0.06 \times 0.01667^2 = 0.0002281$$

The volatility estimate on day n is therefore $\sqrt{0.0002281} = 0.015103$ or 1.5103%.

10.10 Reducing λ from 0.95 to 0.85 means that more weight is given to recent observations of u_i^2 and less weight is given to older observations. Volatilities calculated with $\lambda = 0.85$ will react more quickly to new information and will bounce around much more than volatilities calculated with $\lambda = 0.95$.

10.11 With the usual notation $u_{n-1} = 20/1{,}040 = 0.01923$ so that

$$\sigma_n^2 = 0.000002 + 0.06 \times 0.01923^2 + 0.92 \times 0.01^2 = 0.0001162$$

This gives $\sigma_n = 0.01078$. The new volatility estimate is therefore 1.078% per day.

10.12 The proportional daily change is $-0.005/1.5000 = -0.003333$. The current daily variance estimate is $0.006^2 = 0.000036$. The new daily variance estimate is

$$0.9 \times 0.000036 + 0.1 \times 0.003333^2 = 0.000033511$$

The new volatility is the square root of this. It is 0.00579 or 0.579%.

10.13 The weight given to the long-run average variance rate is $1 - \alpha - \beta$ and the long-run average variance rate is $\omega/(1 - \alpha - \beta)$. Increasing ω increases the long-run average variance rate. Increasing α increases the weight given to the most recent data item, reduces the weight given to the long-run average variance rate, and increases the level of the long-run average variance rate. Increasing β increases the weight given to the previous variance estimate, reduces the weight given to the long-run average variance rate, and increases the level of the long-run average variance rate.

10.14 The long-run average variance rate is $\omega/(1 - \alpha - \beta)$ or $0.000004/0.03 = 0.0001333$. The long-run average volatility is $\sqrt{0.0001333}$ or 1.155%. The equation describing the way the variance rate reverts to its long-run average is

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

In this case,

$$E\left[\sigma_{n+k}^2\right] = 0.0001333 + 0.97^k\left(\sigma_n^2 - 0.0001333\right)$$

If the current volatility is 20% per year, $\sigma_n = 0.2/\sqrt{252} = 0.0126$. The expected variance rate in 20 days is

$$0.0001333 + 0.97^{20}(0.0126^2 - 0.0001333) = 0.0001471$$

The expected volatility in 20 days is therefore $\sqrt{0.0001471} = 0.0121$ or 1.21% per day.

10.15 The FTSE expressed in dollars is XY where X is the FTSE expressed in sterling and Y is the exchange rate (value of one pound in dollars). Define x_i as the proportional change in X on day i and y_i as the proportional change in Y on day i. The proportional change in XY is approximately $x_i + y_i$. The standard deviation of x_i is 0.018 and the standard deviation of y_i is 0.009. The correlation between the two is 0.4. The variance of $x_i + y_i$ is therefore

$$0.018^2 + 0.009^2 + 2 \times 0.018 \times 0.009 \times 0.4 = 0.0005346$$

so that the volatility of $x_i + y_i$ is 0.0231 or 2.31%. This is the volatility of the FTSE expressed in dollars. Note that it is greater than the volatility of the FTSE expressed in sterling. This is the impact of the positive correlation. When the FTSE increases, the value of sterling measured in dollars also tends to increase. This creates an even bigger increase in the value of FTSE measured in dollars. Similarly, when FTSE decreases, the value of sterling measured in dollars also tends to decrease, creating an even bigger decrease in the value of FTSE measured in dollars.

10.16 In this case, $V_L = 0.000003/0.02 = 0.00015$ and equation (10.14) gives the expected variance rate in 30 days as

$$0.00015 + 0.98^{30}(0.01^2 - 0.00015) = 0.000123$$

The volatility is $\sqrt{0.000123} = 0.0111$ or 1.11% per day.

10.17 In this case, $V_L = 0.000002/0.02 = 0.0001$. In equation (10.15), $V_L = 0.0001$, $a = 0.0202$, $T = 20$, and $V(0) = 0.000169$ so that the volatility is 19.88%.

Chapter 11

11.1 You need the standard deviations of the two variables.

11.2 Loosely speaking, correlation measures the extent of linear dependence. It does not measure other types of dependence. When $y = x^2$, there is clearly dependence between x and y. However, $E(xy) = E(x^3)$. Both $E(x)$ and $E(x^3)$ are zero when x is normal (or when it has any symmetrical distribution centered at zero). It follows that the coefficient of correlation between x and y is zero.

11.3 In a factor model, the correlation between two variables arises entirely because of their correlation with other variables known as factors. A factor model reduces the number of estimates that have to be made when correlations between large numbers of variables are being produced.

11.4 A positive-semidefinite matrix is a matrix that satisfies equation (11.4) for all vectors $\mathbf{w}$. If a correlation matrix is not positive-semidefinite, the correlations are internally inconsistent.

11.5 **(a)** The volatilities and correlation imply that the current estimate of the covariance is $0.25 \times 0.016 \times 0.025 = 0.0001$.

(b) If the prices of the assets at close of trading are \$20.50 and \$40.50, the proportional changes are $0.5/20 = 0.025$ and $0.5/40 = 0.0125$. The new covariance estimate is

$$0.95 \times 0.0001 + 0.05 \times 0.025 \times 0.0125 = 0.0001106$$

The new variance estimate for asset A is

$$0.95 \times 0.016^2 + 0.05 \times 0.025^2 = 0.00027445$$

so that the new volatility is 0.0166. The new variance estimate for asset B is

$$0.95 \times 0.025^2 + 0.05 \times 0.0125^2 = 0.000601562$$

so that the new volatility is 0.0245. The new correlation estimate is

$$\frac{0.0001106}{0.0166 \times 0.0245} = 0.272$$

11.6 The most recent returns for X and Y are $1/30 = 0.03333$ and $1/50 = 0.02$, respectively. The previous covariance is $0.01 \times 0.012 \times 0.50 = 0.00006$. The new estimate of the covariance is

$$0.000001 + 0.04 \times 0.03333 \times 0.02 + 0.94 \times 0.00006 = 0.0000841$$

The new estimate of the variance of X is

$$0.000003 + 0.04 \times 0.03333^2 + 0.94 \times 0.01^2 = 0.0001414$$

so that the new volatility of X is $\sqrt{0.0001414} = 0.01189$ or 1.189%. The new estimate of the variance of Y is

$$0.000003 + 0.04 \times 0.02^2 + 0.94 \times 0.012^2 = 0.0001544$$

so that the new volatility of Y is $\sqrt{0.0001544} = 0.01242$ or 1.242%. The new estimate of the correlation between the assets is therefore $0.0000841/(0.01189 \times 0.01242) = 0.569$.

11.7 Continuing with the notation in the answer to Problem 10.15, define z_i as the proportional change in the value of the S&P 500 on day i. The covariance between x_i and z_i is

$$0.7 \times 0.018 \times 0.016 = 0.0002016$$

The covariance between y_i and z_i is

$$0.3 \times 0.009 \times 0.016 = 0.0000432$$

The covariance between $x_i + y_i$ and z_i equals the covariance between x_i and z_i plus the covariance between y_i and z_i. It is

$$0.0002016 + 0.0000432 = 0.0002448$$

The volatility of $x_i + y_i$ is 2.31% from Problem 10.15. The correlation between $x_i + y_i$ and z_i is

$$\frac{0.0002448}{0.016 \times 0.0231} = 0.662$$

Note that the volatility of the S&P 500 drops out in this calculation.

11.8

	V_2		
V_1	0.25	0.5	0.75
0.25	0.095	0.163	0.216
0.50	0.163	0.298	0.413
0.75	0.216	0.413	0.595

11.9 The formulas are

$$\epsilon_1 = z_1 \, , \qquad \epsilon_2 = \rho_{12} z_1 + z_2 \sqrt{1 - \rho_{12}^2} \, , \qquad \epsilon_3 = \alpha_1 z_1 + \alpha_2 z_2 + \alpha_3 z_3$$

where

$$\alpha_1 = \rho_{13} \, , \qquad \alpha_1 \rho_{12} + \alpha_2 \sqrt{1 - \rho_{12}^2} = \rho_{23} \, , \qquad \alpha_1^2 + \alpha_2^2 + \alpha_3^2 = 1$$

This means that

$$\alpha_1 = \rho_{13} \, , \qquad \alpha_2 = \frac{\rho_{23} - \rho_{13}\rho_{12}}{\sqrt{1 - \rho_{12}^2}} \, , \qquad \alpha_3 = \sqrt{1 - \alpha_1^2 - \alpha_2^2}$$

11.10 Tail dependence is the tendency for extreme values for two or more variables to occur together. The choice of the copula affects tail dependence. For example, the Student's *t*-copula gives more tail dependence than the Gaussian copula.

11.11 Sample from a bivariate Student's *t* distribution as in Figure 11.5. Convert each sample to a normal distribution on a percentile-to-percentile basis.

11.12 The probability that $V_1 < 0.1$ is 0.05. The conditional probability that $V_2 < 0.1$ is $0.006/0.05 = 0.12$. The conditional probability that $V_2 < 0.2$ is $0.017/0.05 = 0.34$, and so on.

11.13 When $V_1 = 0.2$, $U_1 = -0.8416$. From the properties of the bivariate normal distribution, the median of U_2 is $-0.5 \times 0.8416 = -0.4208$. The percentile of the conditional median is therefore $N(-0.4208) = 0.3369$. This translates into a median value for V_2 of x where $2x^2 = 0.3369$. The conditional median value for V_2 is therefore 0.4105.

11.14 In this case,

$$\text{WCDR}(T, X) = N\left(\frac{N^{-1}(0.015) + \sqrt{0.2}N^{-1}(0.995)}{\sqrt{1 - 0.2}}\right) = 0.127$$

The loss rate that we are 99.5% certain will not be exceeded is 12.7%.

11.15 The maximum likelihood estimates for the probability of default and the copula correlation are 3.92% and 11.18%, respectively.

Chapter 12

12.1 VaR is the loss that is not expected to be exceeded with a certain confidence level. Expected shortfall is the expected loss conditional that the loss is worse than the VaR level. Expected shortfall has the advantage that it always satisfies the subadditivity (diversification is good) condition.

12.2 A spectral risk measure is a risk measure that assigns weights to the quantiles of the loss distribution. For the subadditivity condition to be satisfied, the weight assigned to the qth quantile must be a non-decreasing function of q.

12.3 There is a 5% chance that you will lose \$6,000 or more during a one-month period.

12.4 Your expected loss during a bad month is \$6,000. Bad months are defined as the months where returns are less than the five-percentile point on the distribution of monthly returns.

12.5 **(a)** \$1 million.

(b) The expected shortfall is $0.9 \times 10 + 0.1 \times 1$ or \$9.1 million.

(c) There is a probability of $0.009^2 = 0.000081$ of a loss of \$20 million, a probability of $2 \times 0.009 \times 0.991 = 0.017838$ of a loss of \$11 million, and a probability of $0.991^2 = 0.982081$ of a loss of \$2 million. The VaR when the confidence level is 99% is therefore \$11 million.

(d) The expected shortfall is $(0.000081 \times 20 + 0.009919 \times 11)/0.01 =$ \$11.07 million.

(e) Because $1 + 1 < 11$, the subadditivity condition is not satisfied for VaR. Because $9.1 + 9.1 > 11.07$, it is satisfied for expected shortfall.

12.6 (a) $2 \times 1.96 =$ \$3.92 million, (b) $\sqrt{5} \times 2 \times 1.96 =$ \$8.77 million, (c) $\sqrt{5} \times 2 \times 2.33 = 10.40$ million.

12.7 (b) becomes \$9.96 million and (c) becomes \$11.82 million.

12.8 Marginal VaR is the rate of change of VaR with the amount invested in the ith asset. Incremental VaR is the incremental effect of the ith asset on VaR (i.e., the difference between VaR with and without the asset). Component VaR is the part of VaR that can be attributed the ith asset (the sum of component VaRs equals the total VaR).

12.9 The probability of 17 or more exceptions is 1-BINOMDIST(16,1000,0.01, TRUE) or 2.64%. The model should be rejected at the 5% confidence level.

12.10 Bunching is the tendency for exceptions to be clustered together rather than occurring randomly throughout the time period considered.

12.11 We are interested in the variance of $\Delta P_1 + \Delta P_2 + \ldots + \Delta P_T$. This is $\sum_{i=1}^{T} \sigma_i^2 + 2\sum_{i>j} \rho_{ij}\sigma_i\sigma_j$ where σ_i is the standard deviation of ΔP_i and ρ_{ij} is the correlation between ΔP_i and ΔP_j. In this case, $\sigma_i = \sigma$ for all i and $\rho_{ij} = \rho^{i-j}$ when $i > j$. After further algebraic manipulations, this leads directly to equation (12.5).

12.12 The standard deviation in three months is $2\sqrt{3} =$ \$3.464 million. Also, $N^{-1}(0.98) = 2.054$. The 98% three-month VaR is therefore $3.464 \times 2.054 =$

\$7.11 million. The 98% three-month ES is $3.464 \times \exp(-2.054^2/2)/(\sqrt{2\pi} \times 0.02) =$ \$8.39 million.

Chapter 13

13.1 The assumption is that the statistical process driving changes in market variables over the next day is the same as that over the past 500 days.

13.2

$$\lambda^{i-1}(1-\lambda)/(1-\lambda^n) = \frac{\lambda^{i-1}}{1+\lambda+\lambda^2+\ldots+\lambda^{n-1}}$$

This shows that as λ approaches 1, the weights approach $1/n$.

13.3 The standard error of the estimate is

$$\frac{1}{0.01}\sqrt{\frac{0.05 \times 0.95}{1{,}000}} = 0.69$$

The standard error of the VaR estimate is \$0.69 million.

13.4 (a) The 95% one-day VaR is the 25th worst loss. This is \$156,511. (b) The 95% one-day ES is the average of the 24 highest losses. It is \$209,310. (c) The 97% one-day VaR is the 15th worst loss. This is \$172,224. (d) The 97% one-day ES is the average of the 14 highest losses. It is \$240,874.

13.5 In worksheet 2 (Scenarios), the portfolio investments are changed to 2,500 in cells L2:O2. The losses are then sorted from the largest to the smallest. The fifth worst loss is \$238,526. This is the one-day 99% VaR. The average of the four worst losses is \$372,872. This is the one-day 99% ES.

13.6 The value of λ in cell F2 of worksheet 4 (Scenarios with Weights) is changed from 0.995 to 0.99. The losses are then sorted from the largest to the smallest. The cumulative weight for the largest loss (\$477,841, Scenario 494) is 0.00948. The cumulative weight for the second largest loss (\$345,435, Scenario 339) is 0.01147. The one-day 99% VaR is therefore the second worst loss or \$345,435. The one-day ES is $0.948 \times 477{,}841 + 0.052 \times 345{,}435$ or \$470,917.

13.7 The value of λ in cell K1 of worksheet 9 (Losses Adjusted for Loss SD) is changed from 0.94 to 0.96. The adjusted losses are then sorted from the highest to the lowest. The fifth worst loss is \$541,337. This is the new one-day 99% VaR. The average of the four highest losses is \$707,146. This is the new one-day 99% ES.

13.8 This is

$$\frac{22}{500}\left[1+0.436\frac{400-160}{32.532}\right]^{-1/0.436}$$

or 0.001623.

13.9 The VaR ($000s) is

$$160 + \frac{32.532}{0.436}\left\{\left[\frac{500}{22}(1 - 0.97)\right]^{-0.436} - 1\right\} = 173.6$$

or $173,600.

13.10 The maximum likelihood estimates of ξ and β become 0.353 and 34.05. The one-day 99% VaR becomes $230,725 and the one-day 99.9% VaR becomes $452,831. The one-day 99% ES becomes $327,336 and the one-day 99.9% ES becomes $670,499.

13.11 The ranked losses from the volatility scaling procedure must be transferred to worksheet 11 (Extreme Val Theory). With $u = 400$, we have $n_u = 17$ and the maximum likelihood values of ξ and β are 0.438 and 82.838. The one-day 99% VaR is $534,100. The one-day 99.9% VaR is $1,096,661. The one-day ES with 99% confidence is $785,819. The one-day ES with 99.9% confidence is $1,786,335. The probability of a loss greater than $600,000 is

$$\frac{17}{500}\left[1 + 0.438\frac{600 - 400}{82.838}\right]^{-1/0.438} = 0.00655$$

Chapter 14

14.1 The standard deviation of the daily change in the investment in each asset is $1,000. The variance of the portfolio's daily change is

$$1{,}000^2 + 1{,}000^2 + 2 \times 0.3 \times 1{,}000 \times 1{,}000 = 2{,}600{,}000$$

The standard deviation of the portfolio's daily change is the square root of this, or $1,612.45. The five-day 97% value at risk is therefore $1.8808 \times \sqrt{5} \times 1{,}612.45 =$ $6,781. The five-day 97% ES is

$$\sqrt{5} \times 1{,}612.45 \times \frac{e^{-1.8808^2/2}}{\sqrt{2\pi} \times 0.03} = \$8{,}178$$

14.2 The two alternative procedures mentioned in the chapter for handling interest rates when the model-building approach is used to calculate VaR involve (a) the use of principal components analysis and (b) describing the term structure with 10 vertices.

14.3 The 10-year vertex has a sensitivity of $30,000 because a one-basis-point increase in that rate leads to an increase of 0.6 basis points in the 12-year rate. Similarly, the 15-year vertex has a sensitivity of $20,000. The other vertices have no sensitivity.

14.4 $\Delta P = 3.9\Delta x$. The standard deviation of ΔP is $3.9 \times 0.007 = 0.027$. It follows that the 10-day 99% VaR for the portfolio is $0.027 \times 2.326 \times \sqrt{10} = 0.20$.

14.5 The relationship is $\Delta P = 3.9\Delta x + 0.5 \times 4.3 \times \Delta x^2$ or $\Delta P = 3.9\Delta x + 2.15\Delta x^2$.

14.6 The variance of the change in the value of the portfolio in one day is (in \$000s)

$$(10 \times 7)^2 + (8 \times 8)^2 + 2 \times 10 \times 7 \times 8 \times 8 \times 0.8 = 16{,}164$$

Because $N^{-1}(0.98) = 2.054$, the ES is therefore

$$\frac{\sqrt{16{,}164 \times 5e^{-2.054^2/2}}}{\sqrt{2\pi} \times 0.02} = 688.2$$

or \$688,200.

14.7 Suppose that the risk measure (VaR or ES) for a portfolio is its standard deviation multiplied by β. The risk weight of a risk factor is its standard deviation multiplied by β. The risk sensitivity for a portfolio is the amount by which the portfolio changes for a 1% or one basis point change in the risk factor.

14.8 The daily variance of the portfolio is

$$6^2 \times 20^2 + 4^2 \times 8^2 = 15{,}424$$

and the daily standard deviation is $\sqrt{15{,}424} = \$124.19$. Since $N^{-1}(0.9) = 1.282$, the five-day 90% value at risk is

$$124.19 \times \sqrt{5} \times 1.282 = \$355.89$$

14.9 (a) 2.0, (b) 43.9.

14.10 The delta of the options is the rate of change of the value of the options with respect to the price of the asset. When the asset price increases by a small amount, the value of the options decreases by 30 times this amount. The gamma of the options is the rate of change of their delta with respect to the price of the asset. When the asset price increases by a small amount, the delta of the portfolio decreases by five times this amount. The delta and gamma with respect to proportional changes are $-30 \times 20 = -600$ and $-5 \times 20 \times 20 = -2{,}000$. In this case, $\sigma = 0.01$ so that $E(\Delta P) = -0.10$, $E(\Delta P^2) = 36.03$, and $E(\Delta P^3) = -32.415$. The mean change in the portfolio value in one day is -0.1 and the standard deviation of the change in one day is $\sqrt{36.03 - 0.1^2} = 6.002$. The skewness is

$$\frac{-32.415 - 3 \times 36.03 \times (-0.1) + 2 \times (-0.1)^3}{6.002^3} = -\frac{21.608}{216.180} = -0.10$$

Using only the first two moments, the one-day 99% value at risk is \$14.08. When three moments are considered in conjunction with a Cornish-Fisher expansion, it is \$14.53.

14.11 Define σ as the volatility per year, ΔS as the change in the asset price in one day, and $\Delta\sigma$ as the change in the volatility in one day. The relationship between ΔP, ΔS and $\Delta\sigma$ is

$$\Delta P = -30\Delta S - 0.5 \times 5 \times (\Delta S)^2 - 2\Delta\sigma$$

or

$$\Delta P = -30 \times 20\Delta x - 0.5 \times 5 \times 20^2(\Delta x)^2 - 2\Delta\sigma$$

(where $\Delta x = \Delta S/S$), which simplifies to

$$\Delta P = -600\Delta x - 1{,}000(\Delta x)^2 - 2\Delta\sigma$$

The change in the portfolio value now depends on two risk factors.

14.12 The change in the value of an option is not linearly related to the changes in the values of the underlying variables. When the changes in the values of underlying variables are normal, the change in the value of the option is non-normal. The linear model assumes that it is normal and is, therefore, only an approximation.

14.13 The contract is a long position in a sterling bond combined with a short position in a dollar bond. The value of the sterling bond is $1.53e^{-0.05\times0.5}$, or \$1.492 million. The value of the dollar bond is $1.5e^{-0.05\times0.5}$, or \$1.463 million. The variance of the change in the value of the contract in one day is

$$1.492^2 \times 0.0006^2 + 1.463^2 \times 0.0005^2 - 2 \times 0.8 \times 1.492 \times 0.0006 \times 1.463 \times 0.0005 = 0.000000288$$

The standard deviation is therefore \$0.000537 million. The 10-day 99% VaR is $0.000537 \times \sqrt{10} \times 2.33$ = \$0.00396 million.

14.14 The alphas should be changed to 2,500. This changes the one-day 99% VaR to \$226,836 and the one-day ES to \$259,878 when volatilities and correlations are estimated using the equally weighted model. It changes the one-day 99% VaR to \$487,737 and the one-day 99% ES to \$558,783 when EWMA with $\lambda = 0.94$ is used.

14.15 This changes the one-day 99% VaR from \$471,025 to \$389,290 and the one-day ES from \$539,637 to \$445,996.

14.16 One way is to assume that the correlation between a vertex of the first term structure and a vertex of the second term structure is the same for all vertices. Another is to calculate the impact on the portfolio of changes to each term structure separately and to use the correlation to aggregate the two impacts.

Chapter 15

15.1 The removal of a competitor may be beneficial. However, banks enter into many contracts with each other. When one bank goes bankrupt, other banks are liable to lose money on the contracts they have with the bank. Also, other banks will be adversely affected if the bankruptcy reduces the public's overall level of confidence in the banking system.

15.2 Deposit insurance means that depositors are safe regardless of the risks taken by their financial institution. It is liable to lead to financial institutions taking more risks than they otherwise would because they can do so without the risk of losing deposits. This in turn leads to more bank failures and more claims under the deposit insurance system. Regulation requiring the capital held by a bank to be related to the risks taken is necessary to avoid this happening.

15.3 The credit risk on the swap is the risk that the counterparty defaults at some future time when the swap has a positive value to the bank.

15.4 The value of a currency swap is liable to deviate further from zero than the value of an interest rate swap because of the final exchange of principal. As a result, the potential loss from a counterparty default is higher.

15.5 There is some exposure. If the counterparty defaulted now, there would be no loss. However, interest rates could change so that at a future time the swap has a positive value to the financial institution. If the counterparty defaulted at that time, there would be a loss to the financial institution. The capital under Basel I would, from Table 15.2, be 0.5% of the swap's principal.

15.6 The risk-weighted assets for the three transactions are (a) \$1.875 million, (b) \$2 million, and (c) \$3 million, for a total of \$6.875 million. The capital required is 0.08×6.875 or \$0.55 million.

15.7 The net replacement ratio (NRR) is $2.5/4.5 = 0.556$. The credit equivalent amount is $2.5 + (0.4 + 0.6 \times 0.556) \times 9.25$ or \$9.28 million. The risk-weighted assets are \$4.64 million, and the capital required is \$0.371 million.

15.8 In this case, there is no value to the netting provisions.

15.9 This converts the estimated capital requirement to estimated risk-weighted assets. Capital required equals 8% of risk-weighted assets.

15.10 The trading book consists of instruments that are actively traded and marked to market daily. The banking book consists primarily of loans that are held to maturity and not marked to market daily. The effect of the change is to move the client's borrowings from the banking book to the trading book. This typically reduces capital requirements. (However, the incremental risk charge in Basel II.5, which is discussed in Chapter 16, brings capital requirements back up to where they were before.)

15.11 Under Basel I, the capital charged for lending to a corporation is the same regardless of the credit rating of the corporation. This leads to a bank's return on capital being relatively low for lending to highly creditworthy corporations. Under Basel II, the capital requirements of a loan are tied much more carefully to the creditworthiness of the borrower. As a result, lending to highly creditworthy companies became more attractive.

15.12 Regulatory arbitrage involves entering into a transaction or series of transactions, without affecting the risks being taken, in order to reduce regulatory capital requirements.

15.13 EAD is the estimated exposure at default. LGD is the loss given default, which is the proportion of the exposure that will be lost if a default occurs. WCDR is the one-year probability of default in a bad year that occurs only one time in 1,000. PD is the probability of default in an average year. MA is the maturity adjustment, which allows for the fact that, in the case of instruments lasting longer than a year, there may be losses arising from a decline in the creditworthiness of the counterparty during the year as well as from a default during the year.

15.14 Under the simple approach, the risk weight of the counterparty is replaced by the risk weight of the collateral for the part of the exposure covered by the collateral. Under the comprehensive approach, the exposure is adjusted for possible increases and the collateral is adjusted for possible decreases in value. The counterparty's risk weight is applied to the excess of the adjusted exposure over the adjusted collateral.

15.15 The standardized approach uses external ratings to determine capital requirements (but in a more sophisticated way than in Basel I). In the IRB approach, the Basel II correlation model is used with PD being determined by the bank. In the advanced IRB approach, the Basel II correlation model is used with PD, LGD, EAD, and M being determined by the bank.

15.16 In the basic indicator approach, total capital is 15% of the average total annual gross income. In the standardized approach, gross income is calculated for different business lines, and capital as a percentage of gross income is different for different business lines. In the advanced measurement approach, the bank uses internal models to determine the 99.9% one-year VaR.

15.17 In this case, $\rho = 0.1216$, WCDR $= 0.0914$, and the capital requirement is $200 \times 0.7 \times (0.0914 - 0.01)$ or \$11.39 million. At least half of this must be Tier 1.

15.18 The probability of five or more exceptions is 1-BINOMDIST(4,250,0.01, TRUE) or 10.8%. It could be argued that regulators are using a confidence level of about 10% (rather than the more usual 5%) in choosing to reject a VaR model.

Chapter 16

16.1 The three major components of Basel II.5 are: the calculation of stressed VaR, a new incremental risk charge, and a comprehensive risk measure for instruments dependent on credit correlation.

16.2 The six major components of Basel III are: capital definitions and requirements, the capital conservation buffer, the countercyclical buffer, the leverage ratio, liquidity ratios, and counterparty credit risk.

16.3 VaR, as it is usually defined, is calculated from the most recent one to four years of daily data. Stressed VaR is calculated from a 250-day period in the past that would be particularly bad for the bank's current portfolio.

16.4 The incremental risk charge is calculated as the one-year 99.9% VaR for losses from credit instruments in the trading book. It takes account of rating changes and liquidity horizons. It was introduced because instruments in the trading book often attracted less capital than equivalent instruments in the banking book.

16.5 The capital requirements for the AAA-rated ABS are 1.6% of principal whereas the capital requirements for the AAA-rated ABS CDO are 3.2% of principal.

16.6 Tier 1 equity capital has increased from 2% to 7%, and the definition of equity capital has been tightened.

16.7 (a) 40%, (b) 20%.

16.8 In the leverage ratio, the denominator is not risk-weighted assets. It is total assets on the balance sheet without risk weighting plus derivatives exposures (calculated as in Basel I) and some off-balance-sheet items such as loan commitments.

16.9 The liquidity coverage ratio (LCR) is the ratio of high-quality liquid assets to net cash outflows during a stressed period of 30 days. The net stable funding ratio (NSFR) is the ratio of a weighted sum of the items on the "liabilities and net worth" side of the balance sheet divided by a weighted sum of the items on the "assets" side of the balance sheet.

16.10 The amount of stable funding would change to 81.6 and the NSFR would become $81.6/74.25 = 110\%$.

16.11 The credit value adjustment (CVA) is a charge to income reflecting expected losses from counterparty defaults in derivatives transactions. The new regulations require the exposure of CVA to credit spreads to be a component of market risk capital.

16.12 CoCo bonds are automatically converted into equity when a predefined trigger indicates that the bank's capital is low. They are attractive to banks because they do not affect return on equity prior to conversion. They are attractive to regulators because they are a source of capital that can absorb losses in stressed market conditions.

Chapter 17

17.1 When an ISDA master agreement is used, each side to an OTC derivatives transaction is taking the risk that the other side will default. When a CCP is used, the CCP acts as an intermediary so that each side has a transaction with the CCP.

17.2 Standard transactions between financial institutions must be cleared through CCPs. For transactions between financial institutions that are cleared bilaterally, both initial margin and variation margin are required. The initial margin must cover, with 99% confidence, market moves during a 10-day period of stressed market conditions.

17.3 The amount of collateral (typically cash or Treasury securities) that has to be provided by financial institutions for their derivatives books will increase.

17.4 The haircut applied to the market price of a financial instrument is the percentage reduction in its price that is made to determine its value for collateral purposes. For example, if an instrument is worth \$100 and is subject to a 10% haircut, it can be used to satisfy a \$90 collateral requirement.

17.5 An event of default occurs when one side to a derivatives transaction that is subject to an ISDA master agreement fails to post collateral or make a payment as required. An early termination follows an event of default and occurs when the nondefaulting party terminates all outstanding transactions with the defaulting party.

17.6 The exposure of A including its exposure to the CCP is reduced to 70. The exposure of A excluding its exposure to the CCP remains 0. The exposure of B including its exposure to the CCP is reduced to 100. The exposure of B excluding its exposure to the CCP is reduced to 70. The exposures of C are unaffected. The average exposure of the three parties, including the exposures to the CCP, is reduced from 110 to 86.7. The average exposure, excluding the exposures to the CCP, is reduced from 70 to 53.3.

17.7 The company will lose money if (a) the CCP defaults or (b) one of the other members defaults and not enough margin and default fund contribution has been provided by that member to cover losses when its positions are closed out.

17.8 Collateral requirements are normally calculated from the net mark-to-market value of all outstanding transactions. In determining settlement amounts, transactions with a positive value are netted against those with a negative value.

17.9 Rehypothecation occurs when collateral posted by A with B is used by B to meet collateral demands from C.

17.10 The activities of CCPs are much simpler than those of banks.

17.11 The mid-market values of transactions to the non-defaulting party at the time of the early termination adjusted upward by half the bid–offer spreads.

Chapter 18

18.1 In Basel I, market risk capital is calculated from the current 99% ten-day VaR, which is assumed to be $\sqrt{10}$ times the current 99% one-day VaR. In Basel II.5, there are two components of the capital. One is calculated from the current 99% VaR (as in Basel I); the other is calculated from the stressed 99% 10-day VaR. Each component is assumed to be $\sqrt{10}$ times the corresponding one-day VaR. In the FRTB, capital is calculated from the 97.5% stressed ES with the time horizon for a variable being dependent on its liquidity.

18.2 The 97.5% ES is $\mu + \sigma e^{-1.96^2/2}/(\sqrt{2\pi} \times 0.025) = \mu + 2.338\sigma$.

18.3 Consider how historical simulation works. The change from Day 0 to Day 10 and the change from Day 1 to Day 11 have nine days in common. They are therefore not independent. Indeed we would expect the two changes to be similar in most situations.

18.4 A theoretical advantage of ES over VaR is that it satisfies the subadditivity (diversification) condition in Section 12.5. A practical advantage is that is quantifies tail risk better.

18.5 The trading book consists of items that are held for trading and marked to market daily. The banking book consists of items that are expected to be held to maturity. Capital requirements are calculated in quite different ways for the two books, and regulators do not want to give banks a way of reducing capital by allowing them to choose the book they put an item in.

18.6 Credit spread risk is treated in the same way as other market risks. Jump-to-default risk is handled similarly to the credit risk associated with items in the banking book.

Chapter 19

19.1 The 10 investment grade ratings used by Moody's are: Aaa, Aa1, Aa2, Aa3, A1, A2, A3, Baa1, Baa2, and Baa3.

19.2 The 10 investment grade ratings used by S&P are: AAA, AA+, AA, AA–, A+, A, A–, BBB+, BBB, and BBB–.

19.3 From equation (19.2), the average hazard rate, $\bar{\lambda}$, satisfies $0.03573 = 1 - e^{-\bar{\lambda}\times 1}$ so that $\bar{\lambda} = -\ln(0.96427) = 0.0364$. The average hazard rate is 3.64% per year.

19.4 Conditional on no default by year 2, the probability of a default in year 3 is

$$(0.04492 - 0.02583)/(1 - 0.02583) = 0.01960$$

Average hazard rate for the third year, $\bar{\lambda}$, satisfies $1 - e^{-\bar{\lambda}\times 1} = 0.01960$. It is 1.98% per year.

19.5 The seller receives $300{,}000{,}000 \times 0.0060 \times 0.5 = \$900{,}000$ at times 0.5, 1.0, 1.5, 2.0, 2.5, 3.0, 3.5, and 4.0 years. The seller also receives a final accrual payment of about \$300,000 ($=\$300{,}000{,}000 \times 0.0060 \times 2/12$) at the time of the default (four years and two months). The seller pays $300{,}000{,}000 \times 0.6 = \$180{,}000{,}000$ at the time of the default.

19.6 Sometimes there is physical settlement of a credit default swap, and sometimes there is cash settlement. In the event of a default when there is physical settlement, the buyer of protection sells bonds issued by the reference entity for their face value. Bonds with a total face value equal to the notional principal can be sold. In the event of a default when there is cash settlement, a calculation agent or an auction determines the value of the cheapest-to-deliver bonds issued by the reference entity a specified number of days after the default event. The cash payoff is then based on the excess of the face value of these bonds over the estimated value.

19.7 Risk-neutral default probabilities are backed out from credit spreads. Real-world default probabilities are calculated from historical data. Risk-neutral default probabilities should be used for valuation. Real-world default probabilities should be used for scenario analysis and credit VaR calculations.

19.8 The payoff is $L(1 - R)$ where L is the notional principal and R is the recovery rate.

19.9 From equation (19.3), the average hazard rate over the three years is $0.0050/(1 - 0.3) = 0.0071$ or 0.71% per year.

19.10 From equation (19.3), the average hazard rate over five years is $0.0080/(1 - 0.4)$ or 1.333% per year. Similarly, the average hazard rate over three years is 1.1667% per year. This means that the average hazard rate for years 4 and 5 is $(5 \times 1.333 - 3 \times 1.1667)/2 = 1.58\%$.

19.11 Real-world probabilities of default should be used for calculating credit value at risk. Risk-neutral probabilities of default should be used for adjusting the price of a derivative for default.

19.12 The recovery rate for a bond is the value of the bond shortly after the issuer defaults as a percentage of its face value.

19.13 The first number in the second column of Table 19.5 is calculated as

$$-\frac{1}{7}\ln(1 - 0.00195) = 0.00028$$

or 0.028% per year. The second number in the second column of Table 19.5 is calculated as

$$-\frac{1}{7}\ln(1 - 0.00525) = 0.00075$$

and so on. The final number is

$$-\frac{1}{7}\ln(1 - 0.42132) = 0.07814$$

The numbers in the fourth column of Table 19.6 are calculated by multiplying the hazard rates (expressed in basis points) in Table 19.5 by 0.6. For Caa, this gives $781.4 \times 0.6 = 469$.

19.14 The no-default value of the bond

$$2e^{-0.03\times 0.5} + 2e^{-0.03\times 1.0} + \ldots + 102e^{-0.03\times 4.0} = 103.66$$

The market price is 96.16. An analysis similar to that in Table 19.3 shows that if Q is the default probability per year, the loss from defaults is $272.69Q$. The implied probability of default is therefore given by solving $103.66 - 96.16 = 272.69Q$. It is 2.74% per year.

19.15 If Q_1 is the default probability at times 0.5, 1.5, and 2.5 years, an analysis similar to that in Table 19.3 shows that the present value of the loss from defaults for the first bond is $178.31Q_1$ per \$100 of face value. The default-free value of the bond is

$$4 \times e^{-0.035\times 1} + 4 \times e^{-0.035\times 2} + 104 \times e^{-0.035\times 3} = 101.23$$

and the market value is 98.35. It follows that $178.31Q_1 = 101.23 - 98.35$ so that $Q_1 = 0.0157$ or 1.57%. If Q_2 is the probability of default at times 3.5 and 4.5, an analysis similar to that in Table 19.3 gives the present value of the loss from default for the second bond to be $180.56Q_1 + 108.53Q_2$. The default-free value of the second bond is

$$4e^{-0.035\times 1} + 4 \times e^{-0.035\times 2} + \ldots + 104 \times e^{-0.035\times 5} = 101.97$$

The market value is 96.24. It follows that $180.56Q_1 + 108.53Q_2 = 101.97 - 96.24$. Substituting for Q_1, we find that $Q_2 = 0.0260$, or 2.60%.

19.16 We can assume that the principal is paid and received at the end of the life of the swap without changing the swap's value. If the spread were zero, the present value of the floating payments per dollar of principal would be 1. The payment of LIBOR plus the spread therefore has a present value of $1 + V$. The payment of the bond cash flows has a present value per dollar of principal of B^*. The initial payment required from the payer of the bond cash flows per dollar of principal is $1 - B$. (This may be negative; an initial amount of $B - 1$ is then paid by the payer of the floating rate.) Because the asset swap is initially worth zero, we have

$$1 + V = B^* + 1 - B$$

so that

$$V = B^* - B$$

19.17 The value of the debt in Merton's model is $V_0 - E_0$ or

$$V_0 - V_0N(d_1) + De^{-rT}N(d_2) = De^{-rT}N(d_2) + V_0N(-d_1)$$

If the credit spread is s, this should equal $De^{-(r+s)T}$ so that

$$De^{-(r+s)T} = De^{-rT}N(d_2) + V_0N(-d_1)$$

Substituting $De^{-rT} = LV_0$, we get

$$Le^{-sT} = LN(d_2) + N(-d_1)$$

so that

$$s = -\ln[N(d_2) + N(-d_1)/L]/T$$

19.18 In this case $E_0 = 2$, $\sigma_E = 0.50$, $D = 5$, $r = 0.04$, and $T = 1$. Solving the simultaneous equations gives $V_0 = 6.80$ and $\sigma_V = 14.82$. The probability of default is $N(-d_2)$ or 1.15%.

19.19 At the end of each quarter for the first four years, the seller receives \$1 million. The seller pays \$70 million after four years and two months. The seller receives a final accrual payment of \$666,667.

19.20 A credit default swap insures a corporate bond issued by the reference entity against default. Its approximate effect is to convert the corporate bond into a risk-free bond. The buyer of a credit default swap has therefore chosen to exchange a corporate bond for a risk-free bond. This means that the buyer is long a risk-free bond and short a similar corporate bond.

19.21 Payoffs from credit default swaps depend on whether a particular company defaults. Arguably some market participants have more information about this than other market participants. (See Business Snapshot 19.2.)

19.22 Suppose that the principal is \$100. The present value of the bond if it were risk-free would be

$$2.5e^{-0.06\times0.5} + 2.5e^{-0.06\times1} + \ldots + 2.5e^{-0.06\times4.5} + 102.5e^{-0.06\times5} = 95.3579$$

The present value of the expected loss from defaults is therefore $95.3579 - 90 = 5.3579$. The asset swap is structured so that the \$10 is paid initially. After that, \$2.50 is paid every six months. In return, LIBOR plus a spread is received on the principal of \$100. The present value of the fixed payments is

$$\begin{aligned}&10 + 2.5e^{-0.06\times0.5} + 2.5e^{-0.06\times1} + \ldots + 2.5e^{-0.06\times4.5} + 102.5e^{-0.06\times5}\\&= 105.3579\end{aligned}$$

The spread over LIBOR must therefore have a present value of 5.3579. The present value of \$1 received every six months for five years is 8.5105. The spread received every six months must therefore be $5.3579/8.5105 = \$0.6296$. The asset swap spread is therefore $2 \times 0.6296 = 1.2592\%$ per annum. This problem provides an illustration of the result in Problem 19.16.

Chapter 20

20.1 The new transaction will increase the bank's exposure to the counterparty if it tends to have a positive value whenever the existing transactions have a net positive value and a negative value whenever the existing transactions have a negative value. However, if the new transaction tends to offset the existing transactions, it is likely to have the incremental effect of reducing credit risk.

20.2 A company's own equity would not be good collateral. When the company defaults on the transactions it has with you, its equity is likely to be worth very little.

20.3 The statements in (a) and (b) are true. The statement in (c) is not. Suppose that v_X and v_Y are the exposures to X and Y. The expected value of $v_X + v_Y$ is the expected value of v_X plus the expected value of v_Y. The same is not true of 95% confidence limits.

20.4 Assume that defaults happen only at the end of the life of the forward contract. In a default-free world, the forward contract is the combination of a long European call and a short European put where the strike price of the options equals the delivery price and the maturity of the options equals the maturity of the forward contract. If the no-default value of the contract is positive at maturity, the call has a positive value and the put is worth zero. The impact of defaults on the forward contract is the same as that on the call. If the no-default value of the contract is negative at maturity, the call has a zero value and the put has a positive value. In this case, defaults have no effect. Again, the impact of defaults on the forward contract is the same as that on the call. It follows that the contract has a value equal to a long position in a call that is subject to default risk and a short position in a default-free put.

20.5 The Black–Scholes–Merton price must be multiplied by $e^{-0.012\times 3} = 0.964$. Black–Scholes–Merton overstates the price by about 3.6%.

20.6 If many contracts entered into by Company X are subject to the same default trigger, the effect of the default trigger may be to increase risk. If the company is downgraded so that the default trigger is activated, those of its counterparties who have contracts with negative values to Company X will request collateral. As a result, Company X is likely to experience liquidity problems and may be forced into bankruptcy.

20.7 A dealer is likely to have wrong-way risk when the counterparty is entering into a CDS selling credit protection to the dealer or when the counterparty is speculating. The dealer is likely to have right-way risk when the counterparty is entering into a CDS buying credit protection from the dealer or when the counterparty is partially hedging an existing position.

20.8 The cure period is the period that is assumed between an event of default and an early termination in CVA calculations.

20.9 All outstanding transactions between Party A and Party B are considered to be a single transaction if Party B defaults. It is therefore the total value of all outstanding transactions than determines the loss, if any, to Party A.

20.10 When a bank is experiencing financial difficulties, its probability of default increases. This increases the expected cost to its counterparties from a default by the bank. The latter is known as DVA and is an accounting entry that increases the value of transactions to the bank.

20.11 The part of CVA relating to the credit spread changes of counterparties is a component of market risk under Basel III.

20.12 There will be losses in the CVA model if the value of the portfolio moves in the dealer's favor during a 15-day period immediately prior to the midpoint of a time interval. It is assumed that either the dealer will fail to receive additional collateral from the counterparty or the counterparty will fail to return collateral it has received from the dealer.

Chapter 21

21.1 In Vasicek's model and Credit Risk Plus, a credit loss is recognized when a default occurs. In CreditMetrics, both downgrades and defaults lead to credit losses. In Vasicek's model, a Gaussian copula model of time to default is used. In Credit Risk Plus, a probability distribution is assumed for the default rate per year. In CreditMetrics, a Gaussian copula model is used to define rating transitions.

21.2 The constant level of risk assumption assumes that after a certain period of time, t, an instrument, X, if it has deteriorated is replaced with an instrument, Y, that has the same risk as X had originally. After a further period of time of length t, the instrument if it has deteriorated is again replaced with an instrument and has the same risk as X had originally, and so on.

21.3 The probability of an Aaa rating staying Aaa over two years is 82.77%. The probability of it moving to Aa is 15.12%.

21.4 The probability of an Aaa rating staying Aaa over six months is 95.35%. The probability of it moving to Aa is 4.40%.

21.5 Movements in credit spreads for all companies over the next day could be assumed to be a random sample from their movements over the past 500 days. The disadvantages of this approach are that the companies have zero chance of defaulting and accurate daily credit spread data may not be available for all companies.

21.6 Using the binomial distribution, the probability of six or more defaults is 0.0005.

21.7 In this case, we must average the cumulative binomial distributions for 0.5% and 1.5% loss probabilities. The probability of six or more defaults is 0.0021. This shows that introducing some correlation increases the tail risk.

21.8 The autocorrelation is quite high at 0.546. This suggests that the credit VaR estimate should take account of recent default experience. If the default rate was high last year, it is more likely to be high this year.

Chapter 22

22.1 Scenarios can be generated by (a) making large changes to key variables such as interest rates, equity prices, and so on; (b) making percentage changes in all market variables that are the same as those that occurred on particular extreme days in the past; and (c) asking a committee of senior management to generate the scenarios.

22.2 Reverse stress testing involves using an algorithm to search for scenarios that lead to big losses. It is used to help identify appropriate scenarios for stress testing.

22.3 Financial institutions might consider that regulators will require more capital if scenarios leading to large losses are considered.

22.4 Traffic light options were options that provided payoffs when the scenarios considered by insurance company regulators occurred. The danger is that, when a financial institution buys traffic light options, it is protecting itself against a too narrow range of adverse scenarios. The financial institution might not be protected against a scenario that is similar to, but not exactly the same as, the scenarios covered by the traffic light options.

22.5 Senior managers are in the best position to develop scenarios for stress testing. Their involvement in the development of scenarios makes it more likely that they will take the results of stress testing seriously and incorporate stress testing in their strategic decision making.

22.6 The advantage is that the same scenarios are considered by different banks and systemic risks can be evaluated. The scenarios might be worse than those considered by the banks themselves (see Problem 22.3). The disadvantage is that it might lead to the banks themselves not spending as much time as they should on stress testing.

22.7 An objective probability is calculated from data. A subjective probability is a "degree of belief" and reflects a person's judgment.

22.8 The total probability of the stressed scenarios is 1.5%. The probability associated with the historical simulation scenarios is therefore 98.5%. Each historical simulation scenario has a probability of 0.197. When the scenarios and their probabilities are ranked, we see that the VaR with a 99% confidence limit is \$284,204. (For a loss of \$340,000, the cumulative probability is 0.00943. For a loss of \$284,204, the cumulative probability is 0.01141.)

22.9 In this case, the position values are 941.34, –164.39, –1,349.94, and –78.36. The worst-case scenario is where the asset price is 60 and its volatility is 30%. This leads to a loss of \$341.39.

Chapter 23

23.1 The definition includes all internal risks and external risks except those related to reputational risk and risks resulting from strategic decisions.

23.2 Based on the results reported by Shih, the loss would be $100 \times 3^{0.23}$ or \$128.7 million.

23.3 $\text{Prob}(v > x) = Kx^{-0.8}$. When $x = 20$ the probability is 0.1. This means that $K = 1.0986$. The probabilities of the specified losses being exceeded are (a) 5.74%, (b) 3.30%, and (c) 1.58%.

23.4 Moral hazard is handled by deductibles and by making premiums dependent on past claims. Adverse selection is handled by finding out as much as possible about a driver before insurance is granted and then modifying premiums as more information on the driver becomes available.

23.5 CEOs must prepare a statement asserting that the financial statements are accurate. They must return bonuses in the event that there is a restatement of financial statements.

23.6 If a trader operates within established risk limits and takes a loss, this is part of market risk. If risk limits are violated, the loss becomes classified as an operational risk.

23.7 (a) It is unlikely that an individual would not look after his or her health because of the existence of a life insurance contract. But it has been known for the beneficiary of a life insurance contract to commit murder to receive the payoff from the contract! (b) Individuals with short life expectancies are more likely to buy life insurance than individuals with long life expectancies.

23.8 External loss data are data relating to the losses of other banks. They are data obtained from sharing agreements with other banks or from data vendors. Vendor data are used to determine relative loss severity. They can be a useful indicator of the ratio of mean loss severity in Business Unit A to mean loss severity in Business Unit B or the ratio of the standard deviation of loss severity in Business Unit A to the standard deviation of loss severity in Business Unit B.

23.9 The Poisson distribution is often used for loss frequency. The lognormal distribution is often used for loss severity.

23.10 Examples of key risk indicators are staff turnover, number of failed transactions, number of positions filled by temps, ratio of supervisors to staff, number of open positions, and percentage of staff who did not take 10 days of consecutive leave in the past 12 months.

23.11 When the loss frequency is 3, the mean total loss is about 3.3 and the standard deviation is about 2.0. When the loss frequency is increased to 4, the mean loss is about 4.4 and the standard deviation of the loss is about 2.4.

23.12 The key inputs are the business indicator and (for banks with a business indicator over \$1 billion) the losses experienced over the previous 10 years. The business indicator is a measure of a bank's size that is relevant for operational risk purposes.

23.13 For Bank A the Loss Component (in euro millions) is $7 \times 20 + 7 \times 20 = 280$ while for Bank B it is $7 \times 20 + 7 \times 20 + 5 \times 20 = 380$. For Bank A the Internal Loss Multiplier is

$$\ln\left(e - 1 + \frac{280}{300}\right) = 0.975$$

while for Bank B it is

$$\ln\left(e - 1 + \frac{380}{300}\right) = 1.094$$

In millions of euros, Bank A's SMA capital is $110 + (300 - 110) \times 0.975 = 295$ while Bank B's SMA capital is $110 + (300 - 110) \times 1.094 = 318$.

Chapter 24

24.1 Investors did not know very much about the mortgages underlying the tranches that were created, and the waterfalls were complex.

24.2 The company or individual providing the quotes is prepared to buy at 50 and sell at 55. The mid-market quote is 52.5. The proportional bid–offer spread is 5.0/52.5 or 0.0952.

24.3 The bid–offer spread for the holding in Company A is $0.01 \times 5{,}000 = \$50$. The bid–offer spread for the holding in Company B is $0.02 \times 3{,}000 = \$60$. The cost of unwinding the portfolio is $(50 + 60)/2$ or \$55.

24.4 The bid–offer spread for the first holding that we are 95% confident will not be exceeded is $5{,}000 \times (0.01 + 1.645 \times 0.01) = \$132.24$. The bid–offer spread for the second holding that we are 95% confident will not be exceeded is $3{,}000 \times (0.02 + 1.645 \times 0.03) = \$204.04$. The total cost of unwinding that we are 95% certain will not be exceeded is $(132.24 + 204.04)/2 = \$170.14$.

24.5 The amounts traded on successive days should be 15.9, 12.9, 10.0, 7.4, 5.2, 3.4, 2.2, 1.4, 0.9, and 0.7. The bid–offer spread cost is \$13.4. The total price variance is 36.6 so that the VaR for the market risk with a 95% confidence level is $1.645 \times \sqrt{36.6} = \$9.9$. (The objective function that is minimized is the sum of these two, or \$23.3.)

24.6 LCR tests whether a bank has sufficient liquidity to survive 30 days of extreme stress (credit downgrades, loss of deposits, drawdowns on lines of credit, etc.). NSFR tests how well the maturities of assets and liabilities are matched and discourages overreliance on short-term sources of funds for long-term needs.

24.7 Wholesale deposits are more likely to disappear in stressed market conditions.

24.8 Their hedging led to a loss on the hedge and a gain on the position being hedged. The loss on the hedge gave rise to margin calls. Unfortunately, the position being hedged, although it increased in value to the company, was illiquid.

24.9 Positive feedback trading refers to situations where traders accentuate market movements. They buy when prices increase and sell when prices decrease. Negative feedback trading is when traders do the reverse; that is, they buy when prices decline and sell when prices increase. Positive feedback trading is liable to lead to a liquidity problem.

24.10 This is a VaR measure that includes an adjustment for the bid–offer spread costs that are incurred in a close-out of the position.

24.11 Liquidity black holes occur when most market participants want to be on one side of a market. Regulation is liable to lead to liquidity black holes because, when all financial institutions are regulated in the same way, they tend to want to respond to external economic events in the same way.

24.12 Liquidity black holes are typically caused by too many traders following the same trading strategy. If traders follow diverse trading strategies, liquidity black holes are less likely to occur.

Chapter 25

25.1 Marking to market involves valuing a position using the prices at which the same or similar positions are trading in the market. Marking to model occurs when a model plays a key role in determining the price of an instrument.

25.2 Models may have fundamental errors or they may be used inappropriately.

25.3 The model developer may leave. Also, other employees of the financial institution, as well as its regulators, can understand what has been done.

25.4 When plain vanilla call and put options are being priced, traders do use the Black–Scholes–Merton model as an interpolation tool. They calculate implied volatilities for the options that are actively traded. By interpolating between strike prices and between times to maturity, they estimate implied volatilities for other options. These implied volatilities are then substituted into the Black–Scholes–Merton model to calculate prices for these options. However, Black–Scholes–Merton is more than an interpolation tool when used for hedging.

25.5 13.45%. We get the same answer by (a) interpolating between strike prices of 1.00 and 1.05 and then between maturities six months and one year and (b) interpolating between maturities of six months and one year and then between strike prices of 1.00 and 1.05.

25.6 The models of physics describe the behavior of physical processes. The models of finance ultimately describe the behavior of human beings.

25.7 It might notice that it is getting a large amount of business of a certain type because it is quoting prices different from its competitors. The pricing differences might also become apparent if it decides to unwind transactions and approaches competitors for quotes. Also, it might subscribe to a service where once a month it obtains the average price quotes by dealers for particular transactions.

25.8 Within-model hedging involves hedging against changes in variables that the model assumes to be stochastic. Outside-model hedging involves hedging against parameters that the model assumes to be constant.

25.9 P&L decomposition analyzes day-to-day profits and losses to distinguish the effects of (a) unhedged risks, (b) the hedging model operating imperfectly, and (c) new trades.

25.10 Level 1 transactions are transactions where there are quoted prices for the same transaction in active markets. Level 2 transactions are transactions where there are quoted prices for similar transactions in active markets or for the same transaction in markets that are not active. Level 3 transactions are transactions where some modeling assumptions are required.

25.11 Over-fitting involves trying to match too much of what is observed. Over-parameterization involves making a model too complex and difficult to use by introducing extra parameters.

25.12 An option with a strike price of 0.90 and a time to maturity of 1.5 years has an implied volatility of 14.85%. An option with a strike price of 0.95 and a time to maturity of 1.5 years has an implied volatility of 14.20%. An option with a strike price of 0.92 and a maturity of 1.5 years has an implied volatility of $0.6 \times 14.85\% + 0.4 \times 14.20\%$, which is 14.59%.

25.13 The term *marking to market* refers to the practice of revaluing instruments (usually daily) so that they are consistent with the market. The prices calculated for actively traded products do reflect market prices. The model is used merely as an interpolation tool. The term *marking to market* is therefore accurate for these products. The prices for structured products depend on the model being used, hence the term *marking to model*.

Chapter 26

26.1 Economic capital is a bank's own estimate of the capital it requires. Regulatory capital is the capital it is required to keep by bank supervisors.

26.2 The probability of a AA-rated company defaulting in one year.

26.3 Business risk includes risks relating to strategic decisions and reputation.

26.4 The models used for economic capital are likely to be broadly similar to those used to calculate regulatory capital in the case of market risk and operational risk. When calculating credit risk economic capital, a bank may consider it appropriate to use a different credit correlation model and different correlation parameters from those used in regulatory capital calculations.

26.5 The 99.97% worst-case value of the logarithm of the loss is $0.5 + 4 \times 3.43 = 14.23$. The 99.97% worst-case loss is therefore \$1.510 million. From the properties of the lognormal distribution, the expected loss is $\exp(0.5 + 4^2/2)$ or \$4,915. The capital requirement is therefore \$1.505 million.

26.6 The economic capital for Business Unit 1 is 96.85. The economic capital for Business Unit 2 is 63.87. The total economic capital is 124.66.

26.7 The incremental effect of Business Unit 1 on total economic capital is 60.78. The incremental effect of Business Unit 2 on total economic capital is 27.81. This suggests that $60.78/(60.78 + 27.81)$ or 68.61% of economic capital should be allocated to Business Unit 1, and $27.81/(60.78 + 27.81)$ or 31.39% to Business Unit 2. The marginal effect of increasing the size of Business Unit 1 by 0.5% is 0.4182. The marginal effect of increasing the size of Business Unit 2 by 0.5% is 0.2056. Euler's theorem is satisfied because the total economic capital is approximately equal to the sum of 0.4182/0.005 and 0.2056/0.005.

26.8 The capital is \$38 million and the return before tax is \$7 million. The before-tax RAROC is therefore 18.4%. In practice, the allocation of diversification benefits to this venture might reduce capital and increase RAROC.

26.9 RAROC can be used to compare the past performance of different business units (for performance appraisal) or to project the expected future performance of business units (to decide how capital should be allocated).

Chapter 27

27.1 In top-down approaches, risk limits are set at the senior management or board level for the whole organization. In bottom-up approaches, risk limits are set for the different business units. Both approaches are necessary in ERM because it is necessary to ensure that risk limits for the business units, when aggregated, are consistent with the risk limit for the whole organization.

27.2 The Bowman paradox is based on empirical data showing that taking higher risks tends to lead to lower returns as far as major strategic investments are concerned.

27.3 There is a danger that decisions will be made that are profitable in the short term but lead to long-term problems (e.g., negative publicity, loss of reputation, and lawsuits). Business Snapshots 5.4, 27.1, 27.2, and 27.3 are examples.

27.4 Part of the bonus for a year can be deferred with the possibility of clawbacks.

27.5 ERM takes a holistic (big picture) approach to risk management rather than separately considering the management of different types of risk (credit risk, market risk, operational risk, etc.).

27.6 In equation (27.1), we set $R = -0.2$, $R_F = 0.03$, $R_M = 0.1174$, $\sigma_M = 0.1886$, and $p = 0.9$. The appropriate beta is 1.49.

27.7 Each quarter from year 2 onward, the interest rate would increase by 4%. After 16 quarters, the rate would increase to $1.76\% + 16 \times 4\% = 65.76\%$.

27.8 Examples of cognitive biases are (a) anchoring (making too small adjustments to an initial estimate), (b) availability (recent information given too much weight), (c) representativeness (placing too much reliance on experience when it is limited or not representative of a current situation), and (d) inverting the conditionality.

Chapter 28

28.1 Disintermediation occurs when two sides that previously interacted through an intermediary start to interact directly. Reintermediation occurs when, following disintermediation, a new technology intermediary (providing a different service from the previous intermediary) enters the market.

28.2 Machine learning is a branch of artificial intelligence where computers learn without being explicitly programmed. Distributed ledger technologies are technologies like blockchain where a database is consensually shared and synchronized across multiple computers.

28.3 Biometric authorization uses the unique biological characteristics of individuals to verify that they are who they say they are. Fingerprints, retinal scans, facial recognition, voice authentication, and heartbeat monitoring are examples of what can be used.

28.4 It might think a digital currency makes it easier to control the money supply and manage inflation.

28.5 In P2P lending, borrowers and lenders transact directly with each other using an online platform provided by a technology company. Equity crowdfunding is a mechanism that allows a broad group of investors to fund start-ups via an online platform.

28.6 Regulation is a barrier to entry for FinTechs. In order to conform with regulation, FinTechs sometimes have to offer their services through an established financial institution.

28.7 It is true that if everyone invested only in index funds there would be no price discovery. We need some investors and fund managers to be following news reports and actively searching for stocks that should be bought or sold.

28.8 RegTech refers to technology developments that address regulatory compliance challenges. Examples are systems for online real-time reporting, systems to screen customers and transactions to avoid compliance issues, and online libraries making it easy for financial institutions to retrieve information about regulations on different subjects in different countries.

28.9 IBM recognized that it could not rely on the sale of mainframes and changed its culture and business model to reflect changes in the market. Kodak did not change its culture and business model to the same extent as the photography market changed. Banks will need to change both their culture and their business model in order to survive.

Glossary

ABS See Asset-Backed Security.

ABS CDO A security created from the tranches of different ABSs.

Accrued Interest The interest earned on a bond since the last coupon payment date.

Additional Tier 1 Capital Items such as non-cumulative preferred stock that do not qualify as Tier 1 equity capital.

Add-On Factor When the credit equivalent amount for a derivatives transaction is being calculated, this is the percentage of principal added to the current exposure to allow for possible future changes in the value of the derivative.

Adjoint Differentiation A way of modifying a computer program so that Greek letters are calculated in a computationally efficient way.

Advanced Measurement Approach The way in which the most sophisticated banks will be allowed to calculate regulatory capital for operational risk under Basel II.

Adverse Selection The phenomenon that, if an insurance company offers the same premiums to everyone, it tends to end up providing coverage for the worst risks.

Agency Costs Costs in a business relationship where the interests of the two parties are not perfectly aligned.

Alpha Return earned on a portfolio in excess of that predicted by the capital asset pricing model.

Alternative Investments See Hedge Funds.

American Option An option that can be exercised at any time during its life.

Analytic Result Result where the answer is in the form of an equation.

Arbitrage A trading strategy that takes advantage of two or more securities being mis-priced relative to each other.

Arbitrage Pricing Theory A theory where the return from an investment is assumed to depend on several factors.

Arbitrageur An individual engaging in arbitrage.

ASF Factor Available stable funding factor, a weighting factor used for sources of funding in the calculation of the Net Stable Funding Ratio.

Asian Option An option with a payoff dependent on the average price of the underlying asset during a specified period.

Ask Price The price that a dealer is offering to sell an asset.

Asked Price See Ask Price.

Asset Swap Exchanges the promised coupon on a bond for LIBOR plus a spread.

Asset-Backed Security A security created from the cash flows from bonds, mortgages, credit card receivables, or other instruments.

Asset Swap Exchanges the promised coupon on a bond for LIBOR plus a spread.

At-the-Money Option An option in which the strike price equals the price of the underlying asset.

Autocorrelation The correlation between the value of a variable and the value of the same variable k days later. (k is referred to as the time lag.)

Average Price Call Option An option giving a payoff equal to the greater of zero and the amount by which the average price of the asset exceeds the strike price.

Average Price Put Option An option giving a payoff equal to the greater of zero and the amount by which the strike price exceeds the average price of the asset.

Back-End Load Fee charged when an investment in a mutual fund is terminated.

Back Office Where the record keeping takes place.

Back-Testing Testing a value-at-risk or other model using historical data.

Backward Induction A procedure for working from the end of a tree to its beginning in order to value an option.

Banking Book Part of a bank's portfolio that consists of instruments that are expected to be held to maturity.

Bankruptcy Costs Costs such as lost sales, loss of key managers, and professional fees arising from a declaration of bankruptcy. These costs are not associated with the adverse events leading to bankruptcy.

Barrier Option An option whose payoff depends on whether the path of the underlying asset has reached a barrier (i.e., a certain predetermined level).

Basel I The first international agreement on the regulation of banks in 1988.

Basel II New international regulations for calculating bank capital introduced in 2007.

Basel II.5 Extra capital charges for items in the trading book, introduced following the credit crisis.

Basel III International banking regulations introduced in 2010 involving capital for the banking book and liquidity ratios.

Basel Committee The committee of bank regulators from around the world that produces standards intended to apply to banks in all countries.

Basic Indicator Approach The simplest way of calculating regulatory capital for operational risk under Basel II.

Basis The difference between the spot price and the futures price of a commodity.

Basis Point When used to describe an interest rate, a basis point is one-hundredth of 1% (= 0.01%).

Basis Risk The risk to a hedger arising from uncertainty about the basis at a future time.

Basket Credit Default Swap A credit default swap where there are several reference entities.

Basket Option An option on a portfolio of assets.

Bermudan Option An option that can be exercised on certain dates during its life.

Best Efforts A phrase used to describe the situation where an investment bank agrees that it will do the best it can to sell a new issue of securities at a certain price, but does not guarantee that it will be able to sell them.

Beta A measure of the systematic risk of an asset.

Bid-Ask Spread The amount by which the ask price exceeds the bid price.

Bid-Offer Spread See Bid-Ask Spread.

Bid Price The price that a dealer is prepared to pay for an asset.

Bilateral Clearing An arrangement between two parties to handle transactions in the OTC market, often involving an ISDA master agreement.

Binary Credit Default Swap An instrument where there is a fixed dollar payoff in the event of a default by a particular company.

Binary Option An option with a discontinuous payoff; for example, a cash-or-nothing option or an asset-or-nothing option.

Binomial Model A model where the price of an asset is monitored over successive short periods of time. In each short period it is assumed that only two price movements are possible.

Binomial Tree A tree that represents how an asset price can evolve under the binomial model.

BIS Accord An agreement reached in 1988 between the central banks of 12 countries concerning how banks should be regulated.

Bitcoin The first cryptocurrency.

Bivariate Normal Distribution A joint distribution for two correlated variables, each of which has a normal distribution.

Black's Model An extension of the Black–Scholes–Merton model for valuing European options on futures contracts. It is used extensively in practice to value European options when the distribution of the asset price at maturity is assumed to be lognormal.

Black–Scholes–Merton Model A model for pricing European options on stocks, developed by Fischer Black, Myron Scholes, and Robert Merton.

Blockchain Distributed ledger technology where a secure database is consensually shared and synchronized across multiple computers.

Bond Option An option where a bond is the underlying asset.

Bond Yield The discount rate which, when applied to all the cash flows of a bond, causes the present value of the cash flows to equal the bond's market price.

Bootstrap Method A procedure for calculating the zero-coupon yield curve from market data. Also a statistical procedure for calculating confidence levels when distributions are determined empirically.

Bunching A tendency for days when the loss is greater than the value at risk to be clustered close together.

Business Risk When used by a bank, this refers to strategic risk (related to a bank's decision to enter new markets and develop new products) and reputation risk.

Buy and Hold A trading strategy where a position is set up and there is no subsequent trading of the position.

Calendar Days Includes every day.

Calibration A method for implying a model's parameters from the prices of actively traded options.

Callable Bond A bond containing provisions that allow the issuer to buy it back at a predetermined price at certain times during its life.

Call Option An option to buy an asset at a certain price by a certain date.

Cancelable Swap A swap that can be canceled by one side on prespecified dates.

Cap See Interest-Rate Cap.

Capital Adequacy The adequacy of the capital held by a bank or other financial institution.

Capital Asset Pricing Model A model relating the expected return on an asset to its beta.

Capital Conservation Buffer Extra equity capital that, under Basel III, must be kept so that it is available to absorb losses during downturns.

Caplet One component of an interest rate cap.

Cap Rate The rate determining payoffs in an interest rate cap.

Cascade Approach Describes the calculation methodology in FRTB's internal models approach.

Case–Shiller Index An index of house prices in the United States.

Cash-Flow Mapping A procedure for representing an instrument as a portfolio of zero-coupon bonds for the purpose of calculating value at risk.

Cash Settlement A procedure for settling a contract in cash rather than by delivering the underlying asset.

CAT Bond A bond where the interest and, possibly, the principal paid are reduced if a particular category of "catastrophic" insurance claims exceeds a certain amount.

CCAR The Comprehensive Capital Analysis and Review is a stress test in the United States for banks with consolidated assets over $50 billion.

CCP See Central Clearing Party.

CDD See Cooling Degree Days.

CDO See Collateralized Debt Obligation.

CDO Squared An instrument in which the default risks in a portfolio of CDO tranches are allocated to new securities.

CDS See Credit Default Swap.

CDS–Bond Basis The excess of the CDS spread over the asset swap spread.

CDX An index of the credit quality of 125 North American investment grade companies.

Central Clearing The use of a central counterparty (CCP) for OTC transactions.

Central Counterparty Clearing house for over-the-counter transactions.

Central Limit Theorem A theorem that states that the sum of a large number of independent identically distributed variables is approximately normally distributed even if the variables themselves are not normally distributed.

Cheapest-to-Deliver Bond A bond that is cheapest to buy and then deliver in a futures contract or a credit default swap contract.

Chinese Walls A phrase used to describe the policies in place within a financial institution that prevent information flowing from one part of the financial institution to another when this would be disadvantageous to one or more of the financial institution's clients.

Cholesky Decomposition A method of sampling from a multivariate normal distribution.

Clawback In the case of a hedge fund, a clawback allows a percentage of previous incentive fees to be recouped to compensate investors for current losses. In the case of a deferred bonus, it leads to the bonus not being paid if financial performance deteriorates.

Clean Price of Bond The quoted price of a bond. The cash price paid for the bond (or dirty price) is calculated by adding the accrued interest to the clean price.

Clearing House A firm that guarantees the performance of the parties in an exchange-traded derivatives transaction. (Also referred to as a clearing corporation.)

Clearing Margin A margin posted by a member of a clearing house.

Closed-End Fund A mutual fund where there is a fixed number of shares.

CoCo See Contingent Convertible Bond.

Cognitive Bias Illogical pattern of judgment.

Coherent Risk Measure A risk measure satisfying a number of conditions.

Collar See Interest-Rate Collar.

Collateral Cash or marketable securities that must be posted by a party to a transaction.

Collateralization A system for posting collateral by one or both parties in a derivatives transaction.

Collateralized Debt Obligation A way of packaging credit risk. Several classes of securities (known as tranches) are created from a portfolio of bonds and there are rules for determining how the costs of defaults are allocated to classes.

Commercial Bank A bank that takes deposits and makes loans.

Component Value at Risk VaR corresponding to a component of a portfolio. Defined so that the sum of the component VaRs for the components of a portfolio equals the VaR for the whole portfolio.

Compounding Frequency This defines how an interest rate is measured.

Compounding Swap A swap where interest compounds instead of being paid.

Compound Option An option on an option.

Comprehensive Risk Measure A risk measure that calculates the capital charge for instruments dependent on credit correlation.

Conditional Tail Expectation See Expected Shortfall.

Conditional Value at Risk (C-VaR) See Expected Shortfall.

Confidence Level A probability indicating the level of confidence that an event will (or will not) happen. In value at risk calculations, the confidence level is the probability that the loss will not exceed the value at risk.

Confirmation A contract confirming a verbal agreement between two parties to a trade in the over-the-counter market.

Constant Level of Risk An assumption in the calculation of a risk measure that a financial institution will trade periodically so that its risk at the beginning of each period is the same as today. (Can be contrasted with Buy and Hold.)

Consumption Asset An asset held for consumption rather than investment.

Contagion See Credit Contagion.

Contingent Convertible Bond (CoCo) Bond that is automatically converted to equity when there is a trigger indicating that the issuer needs more equity capital.

Continuous Compounding A way of quoting interest rates. It is the limit as the assumed compounding interval is made smaller and smaller.

Convenience Yield A measure of the benefits from ownership of an asset that are not obtained by the holder of a long futures contract on the asset.

Conversion Factor A factor multiplied by principal to convert an off-balance-sheet item to its credit equivalent amount.

Convertible Bond A corporate bond that can be converted by the holder into a predetermined amount of the company's equity at certain times during its life.

Convexity A measure of the curvature in the relationship between bond prices and bond yields.

Convexity Adjustment An overworked term. For example, it can refer to the adjustment necessary to convert a futures interest rate to a forward interest rate. It can also refer to the adjustment to a forward rate that is sometimes necessary when instruments are valued.

Cooke Ratio A ratio of capital to risk-weighted assets under Basel I.

Cooling Degree Days The maximum of zero and the amount by which the daily average temperature is greater than 65 degrees Fahrenheit. The average temperature is the average of the highest and lowest temperatures (midnight to midnight).

Copula A way of defining the correlation between variables with known distributions.

Core Tier 1 Capital See Tier 1 Equity Capital.

Cornish-Fisher Expansion An approximate relationship between the fractiles of a probability distribution and its moments.

Correlation A measure of the extent to which there is a linear relation between two variables.

Correlation Matrix An $n \times n$ matrix where the $\{i,j\}$ element is the correlation between variable i and variable j.

Cost of Carry The storage costs plus the cost of financing an asset minus the income earned on the asset.

Countercyclical Buffer An extra capital charge, under Basel III, that is left to the discretion of national regulators.

Counterparty The opposite side in a financial transaction.

Coupon The interest payment made on a bond.

Covariance A measure of the linear relationship between two variables (equals the correlation between the variables times the product of their standard deviations).

Covariance Rate The covariance between daily returns of two variables.

Covered Call A short position in a call option on an asset combined with a long position in the asset.

Crashophobia The fear of a stock market crash similar to that in 1987 that some people claim causes market participants to increase the value of deep-out-of-the-money put options.

Credit Contagion The process whereby credit losses in one sector of the world economy lead to credit losses in other sectors.

Credit Default Swap An instrument that gives the holder the right to sell a bond for its face value in the event of a default by the issuer.

Credit Derivative A derivative whose payoff depends on the creditworthiness of one or more companies or countries.

Credit Equivalent Amount Size of loan that is considered equivalent to an off-balance-sheet transaction in Basel I.

Credit Event A default or other event that triggers a payout in a credit default swap.

Credit Indices Indices that track the cost of buying protection for companies in a specified portfolio.

CreditMetrics A procedure for calculating credit value at risk.

Credit Migration The movement of a company from one rating category to another.

Credit Rating A measure of the creditworthiness of a bond issue.

Credit Ratings Transition Matrix A table showing the probability that a company will move from one credit rating to another during a certain period of time.

Credit Risk The risk that a loss will be experienced because of a default by the counterparty in a derivatives transaction.

Credit Risk Plus A procedure for calculating credit value at risk.

Credit Spread The cost of buying credit protection. Also, the difference between the yield on a bond that might default and the yield on a risk-free bond.

Credit Support Annex Part of an ISDA master agreement specifying collateral arrangements.

Credit Transition Matrix See Credit Ratings Transition Matrix.

Credit Value Adjustment The adjustment to the value of derivatives transactions made by a dealer to allow for the possibility that the counterparty may default.

Credit Value at Risk The credit loss that will not be exceeded at some specified confidence level.

Crowdfunding Funding a project or a company using an online platform.

Cryptocurrency Digital currency designed to provide a secure medium of exchange.

CSA See Credit Support Annex.

Cumulative Distribution Function The probability that a variable will be less than x as a function of x.

Cure Period A period of time assumed to elapse between a default event and a termination in OTC derivatives transactions that are settled bilaterally.

Currency Swap A swap where interest and principal in one currency are exchanged for interest and principal in another currency.

Current Exposure Method A method for calculating exposure at default on derivatives in Basel I and Basel II. It sets exposure at default equal to current exposure plus an add-on factor.

Current VaR or ES VaR or ES calculated using most recent data, not data from a stressed period.

Curvature Gamma.

CVA See Credit Value Adjustment.

Day Count A convention for quoting interest rates.

Day Trade A trade that is entered into and closed out on the same day.

Debit Value Adjustment An adjustment to the value of derivatives transactions made by a dealer to allow for the possibility that the dealer may default.

Debt Value Adjustment See Debit Value Adjustment.

Default Correlation Measures the tendency of two companies to default at about the same time.

Default Intensity See Hazard Rate.

Deferred Annuity An annuity where time elapses between the payment of the funds that will provide the annuity and the start of the annuity.

Defined Benefit Plan A pension plan where there is a formula defining the pension that will be received. Typically, the formula depends on the number of years of service and the salary during the final years of service.

Defined Contribution Plan A pension plan where an employee's contributions plus the contributions made for the employee by the employer are kept in a separate account and invested. On retirement the funds in the account are usually converted into an annuity. Sometimes, they can be taken out as a lump sum.

Deleveraging The process of individuals and companies reducing their borrowing.

Delivery Price The price that will be paid or received in a forward contract.

Delta The rate of change of the price of a derivative with the price of the underlying asset.

Delta Hedging A hedging scheme that is designed to make the price of a portfolio of derivatives insensitive to small changes in the price of the underlying asset.

Delta-Neutral Portfolio A portfolio with a delta of zero so that there is no sensitivity to small changes in the price of the underlying asset.

Dependence Variable A is dependent on variable B if knowing the value of B affects the probability density function of A.

Deposit Insurance Government programs for providing restitution to the depositors of a bank in the event that the bank fails.

DerivaGem Software for valuing options, available on the author's website.

Derivative An instrument whose price depends on, or is derived from, the price of another asset.

Deterministic Variable A variable whose future value is known.

DFAST The Dodd–Frank Act Stress Test is a stress test in the United States for banks with consolidated assets over $10 billion.

Digital Currency Type of currency available only in digital form. It is not available as banknotes or coins.

Directed Brokerage A phrase used to describe the situation where a fund directs its trades to a brokerage house in return for the brokerage house advising its clients to buy the fund.

Dirty Price of Bond The cash price of a bond.

Discount Bond See Zero-Coupon Bond.

Discount Instrument An instrument, such as a Treasury bill, that provides no coupons.

Discount Rate The annualized dollar return on a Treasury bill or similar instrument expressed as a percentage of the final face value.

Disintermediation Change where less use is made of intermediaries.

Distance to Default The number of standard deviations that the value of a company's assets must move for a default to be triggered.

Diversification Reducing risk by dividing a portfolio between many different assets.

Dividend A cash payment made to the owner of a stock.

Dividend Yield The dividend as a percentage of the stock price.

Dodd–Frank Act An act in the United States that was introduced following the credit crisis and is designed to protect consumers and investors, avoid future bailouts, and monitor the functioning of the financial system more carefully.

Dollar Convexity The convexity of an interest rate–dependent portfolio multiplied by the value of the portfolio.

Dollar Duration The duration of an interest rate–dependent portfolio multiplied by the value of the portfolio.

Down-and-In Option An option that comes into existence when the price of the underlying asset declines to a prespecified level.

Down-and-Out Option An option that ceases to exist when the price of the underlying asset declines to a prespecified level.

Downgrade Trigger A clause in a contract that states that the contract can be terminated, or collateral demanded, by one side if the credit rating of the other side falls below a certain level.

Duration A measure of the average life of a bond. It is also an approximation of the ratio of the proportional change in the bond price to the absolute change in its yield.

Duration Matching A procedure for matching the durations of assets and liabilities.

Dutch Auction A process where investors provide bids indicating the number of shares they are prepared to buy and the price they are prepared to pay. Suppose there are N shares on offer. The price that clears the market is the highest price, P, that is such that investors are prepared to buy more than N shares at either P or a higher price than P. The investors who have bid more than P get their orders filled at a price of P. The investors who have bid P get part of their orders filled at P.

DVA See Debit Value Adjustment.

DV01 The impact of a one-basis-point increase in all interest rates.

Dynamic Hedging A procedure for hedging an option position by periodically changing the position held in the underlying asset. The objective is usually to maintain a delta-neutral position.

Dynamic Scenarios Scenarios involving adverse movements in market variables that consider the way companies will respond to the adverse movements.

EAD See Exposure at Default.

Early Exercise Exercise prior to the maturity date.

Early Termination Event The early termination of OTC derivative transactions by one side because of an event of default by the other side.

Economic Capital The capital that a bank's own calculation indicates it needs.

Efficient Frontier The optimal trade-offs for an investor between expected return and standard deviation of return.

Efficient Market Hypothesis A hypothesis that asset prices reflect relevant information.

Electronic Trading System of trading where a computer is used to match buyers and sellers.

Embedded Option An option that is an inseparable part of another instrument.

Empirical Research Research based on historical market data.

Endowment Life Insurance Insurance that pays out a predetermined amount on the maturity of the contract or on the policyholder's death, whichever is earlier.

Enterprise Risk Management A holistic approach to risk management.

Equity Swap A swap where the return on an equity portfolio is exchanged for either a fixed or a floating rate of interest.

Equity Tranche The tranche that is first to be affected by losses on the underlying portfolio.

ERM See Enterprise Risk Management.

ES See Expected Shortfall.

Eurocurrency A currency that is outside the formal control of the issuing country's monetary authorities.

Eurodollar A dollar held in a bank outside the United States.

Eurodollar Futures Contract A futures contract written on a Eurodollar deposit.

Eurodollar Interest Rate The interest rate on a Eurodollar deposit.

European Option An option that can be exercised only at the end of its life.

Event of Default An event such as a bankruptcy or a failure to make payments as they become due or a failure to post collateral as it becomes due.

EWMA Exponentially weighted moving average.

Exception A situation in back-testing where realized loss exceeds the value-at-risk estimate.

Excess Cost Layers The costs for an insurance company on a particular type of business that are within a certain range.

Excess Kurtosis A measure of the extent to which the tails of a probability distribution are fatter than those of a normal distribution.

Excess Spread A situation where the aggregate return promised to the tranches is less than the promised return on the underlying assets.

Exchange-Traded Fund A fund created in such a way that units can be exchanged for the underlying shares and vice versa by institutional investors.

Exchange-Traded Market A market organized by an exchange such as the New York Stock Exchange or Chicago Board Options Exchange.

Ex-Dividend Date When a dividend is declared, an ex-dividend date is specified. Investors who own shares of the stock just before the ex-dividend date receive the dividend.

Exercise Price The price at which the underlying asset may be bought or sold in an option contract. (Also called the strike price.)

Exotic Option A nonstandard option.

Expectations Theory The theory that forward interest rates equal expected future spot interest rates.

Expected Shortfall Expected loss during N days conditional on being in the $(100 - X)\%$ tail of the distribution of losses. The variable N is the time horizon and $X\%$ is the confidence level.

Expected Tail Loss See Expected Shortfall.

Expected Value of a Variable The average value of the variable obtained by weighting the alternative values by their probabilities.

Expense Ratio Ratio of expenses to value of assets under management for a fund.

Expiration Date The end of life of a contract.

Exponentially Weighted Moving Average Model A model to provide forecasts for a variable from historical data where the weights applied to data decline exponentially as we move back through time. It is sometimes applied to variances and covariances in value at risk calculations.

Exponential Weighting A weighting scheme where the weight given to an observation depends on how recent it is. The weight given to an observation t time periods ago is λ times the weight given to an observation $t - 1$ time periods ago where $\lambda < 1$.

Exposure at Default The maximum amount that could be lost (assuming no recovery) when a default occurs.

Extreme Value Theory A theory enabling the shape of the tails of a distribution to be estimated from data.

Factor Source of uncertainty.

Factor Analysis An analysis aimed at finding a small number of factors that describe most of the variation in a large number of correlated variables. (Similar to a principal components analysis.)

Factor Copula A copula involving several variables where a factor model describes the correlation structure of the transformed variables.

Factor Loadings The values of variables in a factor model when we have one unit of a particular factor and no units of other factors.

Factor Model A model where a set of correlated variables are assumed to depend linearly on a number of uncorrelated factors.

Factor Score In a factor model, this is the amount of a factor present in a particular observation.

Fair Value Accounting Involves showing financial instruments on the balance sheet at their market value.

FICO A credit score developed by Fair Isaac Corporation.

Financial Intermediary A bank or other financial institution that facilitates the flow of funds between different entities in the economy.

FinTech Application of technology in finance.

Firm Commitment Phrase used to describe the situation where an investment bank guarantees that a new issue of securities will be sold at a certain price. If investors do not want to buy the securities, the investment bank will have to do so.

Fixed Annuity An annuity (usually for life) where payments that will be received by the policyholder are fixed in advance.

Floor See Interest-Rate Floor.

Floor-Ceiling Agreement See Collar.

Floorlet One component of a floor.

Floor Rate The rate in an interest rate floor agreement.

Foreign Currency Option An option on a foreign exchange rate.

Forward Contract A contract that obligates the holder to buy or sell an asset for a predetermined delivery price at a predetermined future time.

Forward Exchange Rate The forward price of one unit of a foreign currency.

Forward Interest Rate The interest rate for a future period of time implied by the rates prevailing in the market today.

Forward Price The delivery price in a forward contract that causes the contract to be worth zero.

Forward Rate Can refer to a forward interest rate or a forward exchange rate.

Forward Rate Agreement (FRA) An agreement that a certain interest rate will apply to a certain principal amount for a certain time period in the future.

Front-End Load A fee charged when a mutual fund is purchased.

Front Office Where the trading takes place.

Front Running Trading a stock on a personal account before a fund trades a large volume of the stock.

FRTB See Fundamental Review of the Trading Book.

Fundamental Review of the Trading Book A new way of assigning capital for market risk proposed by the Basel Committee.

Fund of Funds A fund that invests in a portfolio of different hedge funds.

Futures Contract A contract that obligates the holder to buy or sell an asset at a predetermined delivery price during a specified future time period. The contract is settled daily.

Futures Option An option on a futures contract.

Futures Price The delivery price currently applicable to a futures contract.

G-30 Policy Recommendations A set of recommendations concerning derivatives issued by nonregulators in 1993.

Gamma The rate of change of delta with respect to the asset price.

Gamma-Neutral Portfolio A portfolio with a gamma of zero.

Gap Management A procedure for matching the maturities of assets and liabilities.

GARCH Model A model for forecasting volatility where the variance rate follows a mean-reverting process.

Gaussian Copula Model A copula model based on the multivariate normal distribution.

Girsanov's Theorem A theorem showing that when we move from a world with one set of risk preferences to a world with another set of risk preferences the drifts of stochastic variables may change but their volatilities remain the same.

Glass–Steagall Act An act passed in the United States separating commercial and investment banks.

Greek Letters See Greeks.

Greeks Hedge parameters such as delta, gamma, vega, theta, and rho.

Gross Income Net interest income plus non-interest income.

Group Life Insurance Life insurance arranged for a group of people, usually by an employer.

G-SIB Global systemically important bank.

Haircut The discount applied to the value of an asset when it is used as collateral.

Hazard Rate Measures probability of default in a short period of time conditional on no earlier default.

HDD See Heating Degree Days.

Heating Degree Days The maximum of zero and the amount by which the daily average temperature is less than 65 degrees Fahrenheit. The average temperature is the average of the highest and lowest temperatures (midnight to midnight).

Hedge A trade designed to reduce risk.

Hedge Funds Funds that are subject to less restrictions and less regulation than mutual funds, but are limited in the extent to which they can sell units in the fund to the general public.

Hedger An individual who enters into hedging trades.

Hedge Ratio The ratio of the size of a position in a hedging instrument to the size of the position being hedged.

High-Frequency Trading (HFT) Trading generated by the use of computer algorithms.

High-Water-Mark Clause Clause stating that gains equal to previous losses must be made before incentive fees apply.

Historical Default Probability Default probability estimated from historical default experience.

Historical Simulation A simulation based on historical data.

Historical Volatility A volatility estimated from historical data.

Holiday Calendar Calendar defining which days are holidays for the purposes of determining payment dates in a financial transaction.

Hurdle Rate The return that must be earned by a hedge fund before the incentive fee applies.

Hybrid Approach An approach to aggregating different types of economic capital.

IFRS 9 A new accounting standard requiring banks to value loans by taking expected default losses into account.

Implied Default Probability See Risk-Neutral Default Probability.

Implied Volatility Volatility implied from an option price using the Black–Scholes–Merton or a similar model.

Inception Profit Profit created by selling a derivative for more than its theoretical value.

Incremental Risk Charge An extra capital charge for credit-sensitive instruments in the trading book. The charge reflects the fact that capital requirements for credit instruments in the banking book are often higher than for equivalent-risk instruments in the trading book.

Incremental Value at Risk The difference between the value at risk with and without a particular component of the portfolio.

Independence Variable A is independent of variable B if knowing the value of variable B does not affect the probability density function for variable A.

Independent Amount Collateral required by a dealer in bilateral OTC trading that is independent of the value of outstanding transactions.

Index Fund A fund that is designed to match the performance of a particular stock index.

Initial Margin The cash required from a futures trader at the time of the trade.

Initial Public Offering An offering to investors of shares in a company for the first time. Prior to an IPO, shares are typically held only by the company's founders, its employees, and the providers of venture capital.

Instantaneous Forward Rate A forward rate for a very short period of time in the future.

Interest-Rate Cap An option that provides a payoff when a specified interest rate is above a certain level. The interest rate is a floating rate that is reset periodically.

Interest-Rate Collar A combination of an interest-rate cap and an interest-rate floor.

Interest-Rate Derivative A derivative whose payoffs are dependent on future interest rates.

Interest-Rate Floor An option that provides a payoff when an interest rate is below a certain level. The interest rate is a floating rate that is reset periodically.

Interest-Rate Option An option where the payoff is dependent on the level of interest rates.

Interest-Rate Swap An exchange of a fixed rate of interest on a certain notional principal for a floating rate of interest on the same notional principal.

Internal Credit Rating A credit rating produced by a financial institution rather than by an outside agency such as Moody's or Standard & Poor's.

In-the-Money Option Either (a) a call option where the asset price is greater than the strike price or (b) a put option where the asset price is less than the strike price.

Intrinsic Value For a call option, this is the greater of the excess of the asset price over the strike price and zero. For a put option, it is the greater of the excess of the strike price over the asset price and zero.

Investment Asset An asset held by a significant number of individuals for investment purposes.

Investment Bank A bank that helps companies issue debt and equity securities.

Investment Grade A bond or other instrument with a credit rating of BBB (Baa) or above.

IPO See Initial Public Offering.

IRB Approach Internal Ratings Based approach for assessing credit risk capital in Basel II.

IRC See Incremental Risk Charge.

ISDA International Swaps and Derivatives Association.

ISDA Master Agreement Agreement between two parties governing their OTC derivatives trades.

Isserlis' Theorem A theorem for calculating the mean of the product of zero-mean normally distributed variables.

ITraxx An index of the credit quality of 125 European investment-grade companies.

Jump-to-Default Risk The risk that a company will default. It can be contrasted with credit spread risk, which is the risk that the company's credit spread will increase without the company defaulting.

Junk Bond A non-investment-grade bond.

Key Risk Indicators Indicators to track the level of operational risk.

Kurtosis A measure of the heaviness of the tails of a distribution.

Late Trading Trading a mutual fund after 4 P.M. at the 4 P.M. price.

Leveraging Individuals and companies increasing their borrowings.

LGD See Loss Given Default.

Liar Loan A loan where the borrower did not tell the truth on the application form.

LIBOR See London Interbank Offered Rate.

LIBOR-in-Arrears Swap A swap where the interest paid on a date is determined by the interest rate observed on that date (not by the interest rate observed on the previous payment date).

LIBOR-OIS Spread The difference between the LIBOR rate of interest and the OIS rate of interest for a particular maturity.

LIBOR/Swap Zero Curve Zero rates that are calculated from LIBOR rates, eurodollar futures, and swap rates.

LIBOR Zero Curve See LIBOR/Swap Zero Curve.

Life Insurance Insurance where payouts depend on when a person dies.

Linear Model A model where the change in the portfolio value depends linearly on the returns of the underlying market variables.

Linear Product A derivative product whose price depends linearly on one or more underlying variables.

Liquidity-Adjusted VaR A value at risk calculation that takes account of the impact of the bid–offer spread when positions are unwound.

Liquidity Black Hole A situation where liquidity dries up because everyone wants to be on the same side of the market.

Liquidity Coverage Ratio The ratio of high-quality liquid assets to cash outflows during a stressed period of 30 days.

Liquidity Funding Risk The risk that sources of funding will dry up.

Liquidity Horizon The time horizon used for a risk factor in FRTB's expected shortfall calculations.

Liquidity Preference Theory A theory leading to the conclusion that forward interest rates are above expected future spot interest rates.

Liquidity Premium The amount that forward interest rates exceed expected future spot interest rates.

Liquidity Trading Risk The risk that it will not be possible to unwind positions in financial instruments at their theoretical price.

Living Will A plan for winding up a financial institution in an orderly way so that some parts possibly survive.

Lognormal Distribution A variable has a lognormal distribution when the logarithm of the variable has a normal distribution.

London Interbank Offered Rate The rate at which an AA-rated bank can borrow from other banks.

Longevity Bond A bond where a population of individuals is defined and the coupon paid at a particular time is proportional to the number of people in the population who are alive at that time.

Longevity Risk The risk that people will live longer than expected.

Long Position A position involving the purchase of an asset.

Long-Tail Risk The risk in property–casualty insurance that there will be claims made well after the end of the life of a policy.

Lookback Option An option whose payoff is dependent on the maximum or minimum of the asset price achieved during a certain period.

Loss Given Default The percentage of the exposure to a counterparty that is lost when a default by the counterparty occurs.

Macaulay's Duration The weighted average of the times when cash flows are received where the weight applied to a cash flow is proportional to the present value of the cash flow.

Machine Learning A branch of artificial intelligence that allows computers to learn without being explicitly programmed.

Maintenance Margin When the balance in a trader's margin account falls below the maintenance margin level, the trader receives a margin call requiring the margin account to be topped up.

Margin The amount of collateral required in cash or marketable securities from a trader.

Marginal Distribution The unconditional distribution of a variable.

Marginal Value at Risk The rate of change of the value at risk with the size of one component of the portfolio.

Margin Call A request for extra margin when the balance in the margin account falls below the maintenance margin level.

Margin Period of Risk See Cure Period.

Marginal Distribution The unconditional distribution of a variable.

Market Maker A trader who is willing to quote both bid and offer prices for an asset.

Market Model A model most commonly used by traders.

Market Portfolio A portfolio consisting of the universe of all possible investments.

Market Risk A risk relating to movements in market variables.

Market Timing Frequent trading of a mutual fund to benefit from stale prices or to do late trading.

Marking to Market The practice of revaluing a financial instrument to reflect the current values of the relevant market variables.

Marking to Model Use of a model to determine the current value of a financial instrument.

Maturity Date The end of the life of a contract.

Maximum Likelihood Method A method for choosing the values of parameters by maximizing the probability of a set of observations occurring.

Mean Reversion The tendency of a market variable (such as a volatility or an interest rate) to revert back to some long-run average level.

Merton's Model A model that uses equity prices to estimate default probabilities. (Other models developed by Merton are also sometimes referred to as Merton's model.)

Mezzanine Tranche A tranche that experiences losses after the equity tranche but before the senior tranche.

Middle Office Where risk management takes place.

Minimum Transfer Amount The minimum amount transferred in a collateralization agreement.

Model Building Approach The use of a model to estimate value at risk.

Model Risk The risk relating to the use of models to price derivative products.

Modified Duration A modification to the standard duration measure so that it more accurately describes the relationship between proportional changes in a bond price and actual changes in its yield. The modification takes account of the compounding frequency with which the yield is quoted.

Money Center Banks Banks with a global presence that fund themselves to a significant extent in wholesale markets.

Monte Carlo Simulation A procedure for randomly sampling changes in market variables.

Moral Hazard The possibility that the behavior of an insured entity will change because of the existence of an insurance contract.

Mortality Risk The risk that people will die earlier than expected.

Multibank Holding Company A holding company that owns several banks. This structure was used in the United States for getting around some bank regulations.

Multivariate Normal Distribution The joint distribution of many variables, each of which is normal.

Mutual Fund A vehicle for investing the funds of many small investors.

Naked Position A short position in a call option that is not combined with a long position in the underlying asset.

National Association of Insurance Commissioners A national body that provides services to the insurance regulators in each state in the United States.

Negative Feedback Trading Trading where assets are sold after a price increase and bought after a price decrease.

Net Asset Value The value of a fund's investments divided by the number of shares in the fund.

Net Interest Income The excess of interest earned over interest paid for a bank.

Net Replacement Ratio The ratio of current exposure with netting to current exposure without netting.

Net Stable Funding Ratio The weighted average of funding available divided by weighted average of funding required.

Netting The ability to offset contracts with positive and negative values in the event of a default by a counterparty.

Neural Networks A machine learning tool that allows computers to recognize patterns in a similar way to humans.

NINJA A term used to describe a poor credit risk: no income, no job, no assets.

Node Delta When the term structure is piecewise linear, this measures the impact on a portfolio of changing the interest rate corresponding to one of the vertices by one basis point while keeping the other vertices unchanged.

Non-Investment Grade A bond or other instrument with a credit rating below BBB (Baa).

Nonlife Insurance See Property–Casualty Insurance.

Nonlinear product A derivative product that is not linearly dependent on the underlying variables.

Nonperforming Loan A loan where interest is more than 90 days overdue.

Nonsystematic Risk Risk that can be diversified away.

Normal Distribution The standard bell-shaped distribution of statistics.

Normal Market A market where futures prices increase with maturity.

Notional Principal The principal used to calculate payments in a derivatives transaction. The principal is notional because it is neither paid nor received.

Numerical Procedure A method of calculation when no formula is available.

Objective Probability Probability based on data.

Offer Price See Ask Price.

OIS See Overnight Indexed Swap.

OIS Rate The fixed rate exchanged for geometric average of overnight interest rates in an OIS swap.

Open-End Fund A mutual fund where the number of shares outstanding increases and decreases as investors deposit and withdraw funds.

Open Interest The total number of long positions outstanding in a futures contract (equals the total number of short positions).

Open Outcry A system of trading where traders meet on the floor of the exchange.

Operational Risk The risk of loss arising from inadequate or failed internal processes, people, and systems or from external events.

Option The right to buy or sell an asset.

Organized Trading Facility Electronic platform for standard over-the-counter derivatives in Europe.

Originate-to-Distribute Model Phrase used to describe the practice where a bank originates loans, credit card receivables, and so on, and securitizes them so that the credit risks are distributed to other investors.

OTC Market See Over-the-Counter Market.

OTF See Organized Trading Facility.

Out-of-the-Money Option Either (a) a call option where the asset price is less than the strike price or (b) a put option where the asset price is greater than the strike price.

Outside-Model Hedging Hedging against parameters that are assumed to be constant in the model (e.g., hedging against volatility in Black–Scholes–Merton).

Overcollateralization A situation where total principal of the underlying assets is greater than the sum of the principals of the tranches that are created from the assets.

Overnight Indexed Swap An agreement to exchange the geometric average of overnight rates during a certain period for a fixed rate.

Over-the-Counter Market A market where traders deal directly with each other rather than through an exchange. The traders are usually financial institutions, corporations, and fund managers.

P2P Lending Peer-to-peer lending is lending money to a borrower using an online platform without a bank or similar intermediary being involved.

P&L Attribution The analysis of the reason for daily profit and loss.

Parallel Shift A movement in the yield curve where each point on the curve changes by the same amount.

Partial Duration The percentage change in value of a portfolio for a small change in one point on the zero-coupon yield curve.

Partial Expected Shortfall The expected shortfall for part of a portfolio.

Partial Simulation Approach An approach for calculating value at risk where Greek letters and Taylor series expansions are used to approximate the change in the portfolio value.

Par Value The principal amount of a bond.

Par Yield The coupon on a bond that makes its price equal the principal.

Payoff The payment received by the holder of an option or other instrument.

PD Probability of default.

Peak Exposure A high percentile (e.g., 97.5) of the estimated exposure distribution at a particular future time.

Performing Loan A loan where interest is not more than 90 days overdue.

Permanent Life Insurance See Whole Life Insurance.

Physical Default Probability See Historical Default Probability.

Plain Vanilla A term used to describe a standard deal.

Poison Pill Action taken by a company to make a takeover of the company more difficult.

Poisson Distribution A distribution for number of events in a certain time period in a Poisson process.

Poisson Process A process describing a situation where events happen at random. The probability of an event in time Δt is $\lambda \Delta t$ where λ is the intensity of the process.

Policyholder The holder of an insurance policy.

Portfolio Immunization Making a portfolio relatively insensitive to interest rates.

Portfolio Insurance Entering into trades to ensure that the value of a portfolio will not fall below a certain level.

Positive Feedback Trading Trading where assets are sold after a price decline and bought after a price increase.

Positive Semidefinite A condition that must be satisfied by a variance–covariance matrix for it to be valid.

Power Law Law describing the tails of many probability distributions that are encountered in practice.

Predatory Trading Trading by someone who thinks another market participant will do a similar trade in the near future and that the similar trade, if it happens, will move the market. May or may not be legal, depending on the circumstances.

Premium The price of an option.

Prime Broker A bank that clears a hedge fund's trades, lends the hedge fund money, and provides other services.

Principal The par or face value of a debt instrument.

Principal Components Analysis An analysis aimed at finding a small number of factors that describe most of the variation in a large number of correlated variables. (Similar to factor analysis.)

Private Placement Selling a new issue of securities to a small number of large financial institutions without offering it to the general public.

Property–Casualty Insurance Property insurance provides a payout in the event of loss of property or damage to property. Casualty insurance provides protection against legal liabilities or damage to the property of others.

Proprietary Trading Trading for a financial institution's own account rather than the account of one of the financial institution's clients.

Public Offering Attempting to sell a new issue of securities to the general public.

Put-Call Parity The relationship between the price of a European call option and the price of a European put option when they have the same strike price and maturity date.

Put Option An option to sell an asset for a certain price by a certain date.

Puttable Bond A bond where the holder has the right to sell it back to the issuer at certain predetermined times for a predetermined price.

Puttable Swap A swap where one side has the right to terminate early.

QIS See Quantitative Impact Study.

Quadratic Model A quadratic relationship between change in portfolio value and percentage changes in market variables.

Quantitative Impact Study A study by the Basel Committee of the effect of proposed new regulations.

RAROC Risk-adjusted return on capital.

RCSA Risk control and self-assessment. An approach to assessing operational risk.

Real World The world we live in. Can be contrasted with the risk-neutral world, which is an artificial world used for valuing derivatives.

Rebalancing The process of adjusting a trading position periodically. Usually the purpose is to maintain delta neutrality.

Recovery Rate The amount recovered in the event of a default as a percent of the face value.

Reference Entity A company or country whose default is being insured against in a credit default swap.

RegTech The use of technology by a financial institution to improve compliance with regulations.

Regulatory Arbitrage Transactions designed to reduce the total regulatory capital of the financial institutions involved.

Regulatory Capital Capital a financial institution is required by regulators to keep.

Rehypothecation The use of collateral posted by one counterparty to satisfy the collateral requirements of another counterparty.

Reinsurance A situation where an insurance company passes on some of its risks to another company. Of course, it has to pay the other company for taking on the risks.

Reintermediation The development of new intermediaries following disintermediation.

Repo Repurchase agreement. A procedure for borrowing money by selling securities to a counterparty and agreeing to buy them back later at a slightly higher price.

Repo Rate The rate of interest in a repo transaction.

Reserve Requirement Percentage of deposits that must be kept in cash or on deposit with the central bank.

Reset Date The date in a swap or cap or floor when the floating rate for the next period is set.

Retail Banking Taking relatively small deposits from retail customers and making relatively small loans to them.

Reverse Stress Testing Using an algorithm to search for scenarios that lead to large losses.

Reversion Level The level to which the value of a market variable (e.g., a volatility or interest rate) tends to revert.

Reversion Rate The rate at which a market variable is pulled to the reversion level.

Rho The rate of change of the price of a derivative with the interest rate.

Right-Way Risk A situation where a counterparty becomes less likely to default as the dealer's exposure to the counterparty increases.

Risk Appetite A statement by a financial institution describing the level of risk that can be tolerated.

Risk Culture How risks are taken into account within an organization. An important aspect of a company's risk culture is the extent to which its employees focus on short-term profits without considering longer-term risks.

Risk Factor A term used to describe a market variable that changes through time, such as an exchange rate, commodity price, or interest rate.

Risk Weights In the weighted sensitivity approach, risk weights are a multiple of the daily standard deviation or daily volatility of a risk factor.

Risk-Free Rate The rate of interest that can be earned without assuming any risks.

Risk-Neutral Default Probability The default probability implied from a credit spread.

Risk-Neutral Valuation The valuation of an option or other derivative assuming the world is risk neutral. Risk-neutral valuation gives the correct price for a derivative in all worlds, not just in a risk-neutral world.

Risk-Neutral World A world where investors are assumed to require no extra return on average for bearing risks.

Risk-Weighted Amount See Risk-Weighted Assets.

Risk-Weighted Assets Quantity calculated in Basel I and Basel II. Total capital must be at least 8% of risk-weighted assets.

Robo-Advisor Online provider of wealth management services.

Robotics Process Automation A software application where a computer is trained to interact with existing computer systems.

Roll Back See Backward Induction.

RSF Factor Required stable funding factor, a weighting factor applied to funding requirements in the calculation of the Net Stable Funding Ratio.

SA-CCR A method for calculating exposure at default when standardized approaches are used. It is more elaborate than the current exposure method.

Sarbanes–Oxley An act passed in the United States in 2002 increasing the responsibilities of directors, CEOs, and CFOs of public companies.

Scenario Analysis An analysis of the effects of possible future scenarios on the value of a portfolio. Also used to estimate possible operational risk losses.

Scorecard Approach An assessment procedure used for operational risk.

SEC Securities and Exchange Commission.

Securitization A process whereby securities are created from future cash flow streams.

SEF See Swap Execution Facility.

Senior Tranche A tranche that is the last to experience losses because of defaults on the underlying assets.

Shadow Banking Bank services provided by non-banks.

Short Position A position assumed when traders sell shares they do not own.

Short Selling Selling in the market shares that have been borrowed from another investor.

SIFI Systemically important financial institution.

SIMM See Standard Initial Margin Model.

Simulation See Monte Carlo Simulation or Historical Simulation.

Skewness The measure of the lack of symmetry in a probability distribution.

Solvency I The current regulatory framework for insurance companies in the European Union.

Solvency II A new regulatory framework for insurance companies proposed by the European Union.

Solvency Risk The risk that liabilities will be greater than assets.

Sovereign Wealth Funds Funds set up to make investments for a country.

Specific Risk Charge Capital requirement for idiosyncratic risks in the trading book.

Spectral Risk Measure Risk measure that assigns weights to the quantiles of the loss distribution.

Speculative Grade See Non-Investment Grade.

Speculator An individual who is taking a position in the market. Usually the individual is betting that the price of an asset will go up or that the price of an asset will go down.

Spot Interest Rate See Zero-Coupon Interest Rate.

Spot Price The price for immediate delivery.

Spot Volatilities The volatilities used to price a cap when a different volatility is used for each caplet.

SR 11-7 Guidance on model risk management provided by the Federal Reserve Board.

Standard Initial Margin Model A model developed by ISDA for calculating initial margin required by regulators on a portfolio of bilaterally cleared derivatives between two financial institutions.

Standard Normal Distribution A normal distribution with mean equal to zero and standard deviation equal to one.

Standardized Measurement Approach A new method for calculating operational risk capital proposed by the Basel Committee.

Static Hedge A hedge that does not have to be changed once it is initiated.

Static Options Replication A procedure for hedging a portfolio that involves finding another portfolio of approximately equal value on some boundary.

Stochastic Variable A variable whose future value is uncertain.

Stock Index An index monitoring the value of a portfolio of stocks.

Stock Index Futures Futures on a stock index.

Stock Index Option An option on a stock index.

Stock Option An option on a stock.

Stop–Loss Rule A trading rule where a position is unwound when losses on the position reach a certain level.

Storage Costs The costs of storing a commodity.

Stressed ES Expected shortfall calculated from a 250-day period of stressed market conditions that would be particularly bad for the financial institution.

Stressed VaR Value at risk calculated from a 250-day period of stressed market conditions that would be particularly bad for the financial institution.

Stress Testing Testing the impact of extreme but plausible market moves on the value of a portfolio.

Strike Price The price at which the asset may be bought or sold in an option contract. (Also called the exercise price.)

Structured Product A derivative designed by a financial institution to meet the needs of a client.

Student's *t*-Copula Copula based on the multivariate Student's *t*-distribution.

Student's *t*-Distribution Distribution with heavier tails than the normal distribution.

Subjective Probability A probability that reflects a person's opinion, rather than data.

Subprime Mortgage A mortgage that is more risky than the average mortgage.

Supplementary Capital See Tier 2 Capital.

Survivor Bond See Longevity Bond.

Swap An agreement to exchange cash flows in the future according to a prearranged formula.

Swap Execution Facility An electronic platform for standard over-the-counter derivatives in the United States.

Swap Rate The fixed rate in an interest rate swap that causes the swap to have a value of zero.

Swaption An option to enter into an interest rate swap where a specified fixed rate is exchanged for floating.

Swap Zero Curve See LIBOR/Swap Zero Curve.

Synthetic CDO A CDO where tranche payments depend on defaults in a bond portfolio, but the bond portfolio is not created.

Synthetic Option An option created by trading the underlying asset.

Systematic Risk Risk that cannot be diversified away.

Systemic Risk The risk that a default by one financial institution will lead to defaults by other financial institutions, creating risks for the whole financial system.

Tail Correlation A correlation between the tails of two distributions. Measures the extent to which extreme values tend to occur together.

Tail Dependence The measure of the frequency with which extreme outcomes occur for two variables at the same time.

Tail Loss See Expected Shortfall.

Taylor Series Expansion For a function of several variables, this relates changes in the value of the function to changes in the values of the variables when the changes are small.

Teaser Rate A low interest rate offered for the first two or three years on a mortgage.

Temporary Life Insurance See Term Life Insurance.

Terminal Value The value at maturity.

Term Life Insurance Life insurance that pays off if death occurs before a certain date.

Term Structure of Interest Rates The relationship between interest rates and their maturities.

Theta The rate of change of the price of an option or other derivative with the passage of time.

Threshold The value of outstanding transactions that leads to collateral being required in bilateral OTC trading.

Tier 1 Equity Capital Equity capital, redefined under Basel III (also called Tier 1 Core Capital).

Tier 2 Capital Subordinated debt (life greater than five years) and similar sources of capital.

Tier 3 Capital Short-term subordinated debt (life between two and five years).

Time Decay See Theta.

Time Value The value of an option arising from the time left to maturity (equals an option's price minus its intrinsic value).

Total Expense Ratio The ratio of total expenses to premiums for a property–casualty insurance company.

Total Return Swap A swap where the return on an asset such as a bond is exchanged for LIBOR plus a spread. The return on the asset includes income such as coupons and the change in value of the asset.

Tracking Error A measure of the error in a trading strategy designed to track a stock index.

Trading Book The part of a bank's portfolio that consists of instruments that are held for resale.

Trading Days Days when markets are open for trading.

Tranche One of several securities created from a portfolio. Examples are the tranches of an ABS.

Transaction Costs The cost of carrying out a trade (commissions plus the difference between the price obtained and the midpoint of the bid-offer spread).

Treasury Bill A short-term non-coupon-bearing instrument issued by the government to finance its debt.

Treasury Bond A long-term coupon-bearing instrument issued by the government to finance its debt.

Treasury Note A Treasury bond lasting less than 10 years.

Tree A representation of the evolution of the value of a market variable for the purposes of valuing an option or other derivative.

Uncleared Trades Trades that are cleared bilaterally rather than through a central counterparty.

Unconditional Default Probability The default probability for a future time period as seen today. Can be contrasted with a conditional default probability, where the probability is conditional on no earlier default.

Underlying Variable A variable that the price of an option or other derivative depends on.

Unit Trust See Mutual Fund.

Universal Life Insurance A form of whole life insurance where the premium payments can be varied from year to year. The final payout is determined from the premiums, the performance of the investment, and costs.

Unsystematic Risk See Nonsystematic Risk.

Up-and-In Option An option that comes into existence when the price of the underlying asset increases to a prespecified level.

Up-and-Out Option An option that ceases to exist when the price of the underlying asset increases to a prespecified level.

Value at Risk A loss that will not be exceeded at some specified confidence level.

VaR See Value at Risk.

Variable Annuity An annuity (usually for life) where the amount received depends on the performance of a portfolio.

Variable Life Insurance A form of whole life insurance where the policyholder decides how funds will be invested. If the investments do well, the proceeds can be used to reduce premiums. There is a guaranteed minimum payout on death, but the payout may be higher if the investments do well.

Variance-Covariance Matrix A matrix showing variances of, and covariances between, a number of different market variables.

Variance Rate The square of volatility.

Variation Margin The extra margin required to bring balance in the margin account up to the required level.

Vasicek's Model A model of default correlation based on the Gaussian copula. (Other models developed by Vasicek are also sometimes referred to as Vasicek's model.)

Vega The rate of change in the price of an option or other derivative with volatility.

Vega-Neutral Portfolio A portfolio with a vega of zero.

VIX Index An index of the implied volatilities of options on the S&P 500.

Volatility A measure of the uncertainty of the return realized on an asset.

Volatility Scaling A procedure for scaling a historical daily change up or down to reflect differences between the current volatility and the volatility at the time of the historical daily change.

Volatility Skew A term used to describe the volatility smile when it is nonsymmetrical.

Volatility Smile The variation of implied volatility with strike price.

Volatility Surface A table showing the variation of implied volatilities with strike price and time to maturity.

Volatility Term Structure The variation of implied volatility with time to maturity.

Volcker Rule Part of the Dodd-Frank Act in the United States that prohibits proprietary trading.

Waterfall Rules for determining how principal and interest cash flows from an underlying portfolio are distributed to tranches. In a typical arrangement, interest cash flows are first used to pay the senior tranche its promised return. The cash flows (if any) that are left over are used to provide the next-most-senior tranche with its promised return, and so on. Principal cash flows are usually used first to repay the senior tranche, then the next-most-senior tranche, and so on. The equity tranche is the last to receive both interest and principal cash flows.

Weather Derivative A derivative where the payoff depends on the weather.

Weighted Sensitivity Terminology used for delta in the weighted sensitivity approach.

Whole Life Insurance Life insurance that lasts for the whole life of the insured and so is certain to provide a payout.

Wholesale Banking Taking large deposits and making large loans.

Within-Model Hedging Hedging only against those variables that the model considers as stochastic (cf. Outside-Model Hedging).

Worst Case Default Rate (WCDR) The default rate that will not be exceeded at some confidence level.

Writing an Option Selling an option.

Wrong-Way Risk A situation where a counterparty becomes more likely to default as the dealer's exposure to the counterparty increases.

Yield A return provided by an instrument.

Yield Curve See Term Structure of Interest Rates.

Zero-Coupon Bond A bond that provides no coupons.

Zero-Coupon Interest Rate The interest rate that would be earned on a bond that provides no coupons.

Zero-Coupon Yield Curve A plot of the zero-coupon interest rate against time to maturity.

Zero Curve See Zero-Coupon Yield Curve.

Zero Rate See Zero-Coupon Interest Rate.

Z-Score A number indicating how likely a company is to default.