1 Derivatives Pricing, Hedging and Risk Management: The State of the Art

1.1 INTRODUCTION

The purpose of this chapter is to give a brief review of the basic concepts used in finance for the purpose of pricing contingent claims. As our book is focusing on the use of copula functions in financial applications, most of the content of this chapter should be considered as a prerequisite to the book. Readers who are not familiar with the concepts exposed here are referred for a detailed treatment to standard textbooks on the subject. Here our purpose is mainly to describe the basic tools that represent the state of the art of finance, as well as general problems, and to provide a brief, mainly non-technical, introduction to copula functions and the reason why they may be so useful in financial applications. It is particularly important that we address three hot issues in finance. The first is the non-normality of returns, which makes the standard Black and Scholes option pricing approach obsolete. The second is the incomplete market issue, which introduces a new dimension to the asset pricing problem – that of the choice of the right pricing kernel both in asset pricing and risk management. The third is credit risk, which has seen a huge development of products and techniques in asset pricing.

This discussion would naturally lead to a first understanding of how copula functions can be used to tackle some of these issues. Asset pricing and risk evaluation techniques rely heavily on tools borrowed from probability theory. The prices of derivative products may be written, at least in the standard complete market setting, as the discounted expected values of their future pay-offs under a specific probability measure derived from non-arbitrage arguments. The risk of a position is instead evaluated by studying the negative tail of the probability distribution of profit and loss. Since copula functions provide a useful way to represent multivariate probability distributions, it is no surprise that they may be of great assistance in financial applications. More than this, one can even wonder why it is only recently that they have been discovered and massively applied in finance. The answer has to do with the main developments of market dynamics and financial products over the last decade of the past century.

The main change that has been responsible for the discovery of copula methods in finance has to do with the standard hypothesis assumed for the stochastic dynamics of the rates of returns on financial products. Until the 1987 crash, a normal distribution for these returns was held as a reasonable guess. This concept represented a basic pillar on which most of modern finance theory has been built. In the field of pricing, this assumption corresponds to the standard Black and Scholes approach to contingent claim evaluation. In risk management, assuming normality leads to the standard parametric approach to risk measurement that has been diffused by J.P. Morgan under the trading mark of RiskMetrics since 1994, and is still in use in many financial institutions: due to the assumption of normality, the approach only relies on volatilities and correlations among the returns on the assets in the portfolio. Unfortunately, the assumption of normally distributed returns has been severely challenged by the data and the reality of the markets. On one hand, even evidence on the returns of standard financial products such as stocks and bonds can be easily proved to be at odds with this assumption. On the other hand, financial innovation has spurred the development of products that are specifically targeted to provide non-normal returns. Plain vanilla options are only the most trivial example of this trend, and the development of the structured finance business has made the presence of non-linear products, both plain vanilla and exotic, a pervasive phenomenon in bank balance sheets. This trend has even more been fueled by the pervasive growth in the market for credit derivatives and credit-linked products, whose returns are inherently non-Gaussian. Moreover, the task to exploit the benefits of diversification has caused both equity-linked and credit-linked products to be typically referred to baskets of stocks or credit exposures. As we will see throughout this book, tackling these issues of non-normality and non-linearity in products and portfolios composed by many assets would be a hopeless task without the use of copula functions.

1.2 DERIVATIVE PRICING BASICS: THE BINOMIAL MODEL

Here we give a brief description of the basic pillar behind pricing techniques, that is the use of risk-neutral probability measures to evaluate contingent claims, versus the objective measure observed from the time series of market data. We will see that the existence of such risk measures is directly linked to the basic pricing principle used in modern finance to evaluate financial products. This requirement imposes that prices must ensure that arbitrage gains, also called "free lunches", cannot be obtained by trading the securities in the market. An arbitrage deal is a trading strategy yielding positive returns at no risk. Intuitively, the idea is that if we can set up two positions or trading strategies giving identical pay-offs at some future date, they must also have the same value prior to that date, otherwise one could exploit arbitrage profits by buying the cheaper and selling the more expensive before that date, and unwinding the deal as soon as they are worth the same. Ruling out arbitrage gains then imposes a relationship among the prices of the financial assets involved in the trading strategies. These are called "fair" or "arbitrage-free" prices. It is also worth noting that these prices are not based on any assumption concerning utility maximizing behavior of the agents or equilibrium of the capital markets. The only requirement concerning utility is that traders "prefer more to less", so that they would be ready to exploit whatever arbitrage opportunity was available in the market. In this section we show what the no-arbitrage principle implies for the risk-neutral measure and the objective measure in a discrete setting, before extending it to a continuous time model.

The main results of modern asset pricing theory, as well as some of its major problems, can be presented in a very simple form in a binomial model. For the sake of simplicity, assume that the market is open on two dates, t and T, and that the information structure of the economy is such that, at the future time T, only two states of the world $\{H, L\}$ are possible. A risky asset is traded on the market at the current time t for a price equal to $S(t)$, while at time T the price is represented by a random variable taking values $\{S(H), S(L)\}$ in the two states of the world. A risk-free asset gives instead a value equal to 1 unit of currency at time T no matter which state of the world occurs: we assume that the price at time t of the risk-free asset is equal to B. Our problem is to price another risky asset taking values $\{G(H), G(L)\}$ at time T. As we said before, the price $g(t)$ must be consistent with the prices $S(t)$ and B observed on the market.

1.2.1 Replicating portfolios

In order to check for arbitrage opportunities, assume that we construct a position in Δ_g units of the risky security $S(t)$ and Π_g units of the risk-free asset in such a way that at time T

$$\Delta_g S(H) + \Pi_g = G(H)$$
$$\Delta_g S(L) + \Pi_g = G(L)$$

So, the portfolio has the same value of asset G at time T. We say that it is the "replicating portfolio" of asset G. Obviously we have

$$\Delta_g = \frac{G(H) - G(L)}{S(H) - S(L)}$$
$$\Pi_g = \frac{G(L)S(H) - G(H)S(L)}{S(H) - S(L)}$$

1.2.2 No-arbitrage and the risk-neutral probability measure

If we substitute Δ_g and Π_g in the no-arbitrage equation

$$g(t) = \Delta_g S(t) + B\Pi_g$$

we may rewrite the price, after naive algebraic manipulation, as

$$g(t) = B[QG(H) + (1 - Q)G(L)]$$

with

$$Q \equiv \frac{S(t)/B - S(L)}{S(H) - S(L)}$$

Notice that we have

$$0 < Q < 1 \Leftrightarrow S(L) < \frac{S(t)}{B} < S(H)$$

It is straightforward to check that if the inequality does not hold there are arbitrage opportunities: in fact, if, for example, $S(t)/B \leqslant S(L)$ one could exploit a free-lunch by borrowing and buying the asset. So, in the absence of arbitrage opportunities it follows that $0 < Q < 1$, and Q is a probability measure. We may then write the no-arbitrage price as

$$g(t) = BE_Q[G(T)]$$

In order to rule out arbitrage, then, the above relationship must hold for all the contingent claims and the financial products in the economy. In fact, even for the risky asset S we must have

$$S(t) = BE_Q[S(T)]$$

Notice that the probability measure Q was recovered from the no-arbitrage requirement only. To understand the nature of this measure, it is sufficient to compute the expected rate of return of the different assets under this probability. We have that

$$E_Q\left[\frac{G(T)}{g(t)} - 1\right] = E_Q\left[\frac{S(T)}{S(t)} - 1\right] = \frac{1}{B} - 1 \equiv i$$

where i is the interest rate earned on the risk-free asset for an investment horizon from t to T. So, under the measure Q all of the risky assets in the economy are expected to yield the same return as the risk-free asset. For this reason such a measure is called *risk-neutral* probability.

Alternatively, the measure can be characterized in a more technical sense in the following way. Let us assume that we measure each risky asset in the economy using the risk-free asset as numeraire. Recalling that the value of the riskless asset is B at time t and 1 at time T, we have

$$\frac{g(t)}{B(t)} = E_Q\left[\frac{G(T)}{B(T)}\right] = E_Q[G(T)]$$

A process endowed with this property (i.e. $z(t) = E_Q(z(T))$) is called a *martingale*. For this reason, the measure Q is also called an *equivalent martingale measure* (EMM).[1]

1.2.3 No-arbitrage and the objective probability measure

For comparison with the results above, it may be useful to address the question of which constraints are imposed by the no-arbitrage requirements on expected returns under the objective probability measure. The answer to this question may be found in the well-known *arbitrage pricing theory* (APT). Define the rates of return of an investment on assets S and g over the horizon from t to T as

$$i_g \equiv \frac{G(T)}{g(t)} - 1 \qquad i_S \equiv \frac{S(T)}{S(t)} - 1$$

and the rate of return on the risk-free asset as $i \equiv 1/B - 1$.

The rate of returns on the risky assets are assumed to be driven by a linear data-generating process

$$i_g = a_g + b_g f \qquad i_S = a_S + b_S f$$

where the risk factor f is taken with zero mean and unit variance with no loss of generality.

[1] The term *equivalent* is a technical requirement referring to the fact that the risk-neutral measure and the objective measure must agree on the same subset of zero measure events.

Of course this implies $a_g = E\left(i_g\right)$ and $a_S = E\left(i_S\right)$. Notice that the expectation is now taken under the original probability measure associated with the data-generating process of the returns. We define this measure P. Under the same measure, of course, b_g and b_S represent the standard deviations of the returns. Following a standard no-arbitrage argument we may build a zero volatility portfolio from the two risky assets and equate its return to that of the risk-free asset. This yields

$$\frac{a_S - i}{b_S} = \frac{a_g - i}{b_g} = \lambda$$

where λ is a parameter, which may be constant, time-varying or even stochastic, but has to be the same for all the assets. This relationship, that avoids arbitrage gains, could be rewritten as

$$E\left(i_S\right) = i + \lambda b_S \qquad E\left(i_g\right) = i + \lambda b_g$$

In words, the expected rate of return of each and every risky asset under the objective measure must be equal to the risk-free rate of return plus a risk premium. The risk premium is the product of the volatility of the risky asset times the market price of risk parameter λ. Notice that in order to prevent arbitrage gains the key requirement is that the market price of risk must be the same for all of the risky assets in the economy.

1.2.4 Discounting under different probability measures

The no-arbitrage requirement implies different restrictions under the objective probability measures. The relationship between the two measures can get involved in more complex pricing models, depending on the structure imposed on the dynamics of the market price of risk. To understand what is going on, however, it may be instructive to recover this relationship in a binomial setting. Assuming that P is the objective measure, one can easily prove that

$$Q = P - \lambda\sqrt{P\left(1 - P\right)}$$

and the risk-neutral measure Q is obtained by shifting probability from state H to state L.

To get an intuitive assessment of the relationship between the two measures, one could say that under risk-neutral valuation the probability is adjusted for risk in such a way as to guarantee that all of the assets are expected to yield the risk-free rate; on the contrary, under the objective risk-neutral measure the expected rate of return is adjusted to account for risk. In both cases, the amount of adjustment is determined by the market price of risk parameter λ.

To avoid mistakes in the evaluation of uncertain cash flows, it is essential to take into consideration the kind of probability measure under which one is working. In fact, the discount factor applied to expected cash flows must be adjusted for risk if the expectation is computed under the objective measure, while it must be the risk-free discount factor if the expectation is taken under the risk-neutral probability. Indeed, one can also check that

$$g\left(t\right) = \frac{E\left[G\left(T\right)\right]}{1 + i + \lambda b_g} = \frac{E_Q\left[G\left(T\right)\right]}{1 + i}$$

and using the wrong interest rate to discount the expected cash flow would get the wrong evaluation.

1.2.5 Multiple states of the world

Consider the case in which three scenarios are possible at time T, say $\{S(HH), S(HL), S(LL)\}$. The crucial, albeit obvious, thing to notice is that it is not possible to replicate an asset by a portfolio of only two other assets. To continue with the example above, whatever amount Δ_g of the asset S we choose, and whatever the position of Π_g in the risk-free asset, we are not able to perfectly replicate the pay-off of the contract g in all the three states of the world: whatever replicating portfolio was used would lead to some *hedging error*. Technically, we say that contract g is not *attainable* and we have an *incomplete market* problem. The discussion of this problem has been at the center of the analysis of modern finance theory for some years, and will be tackled in more detail below. Here we want to stress in which way the model above can be extended to this multiple scenario setting. There are basically two ways to do so. The first is to assume that there is a third asset, whose pay-off is independent of the first two, so that a replicating portfolio can be constructed using three assets instead of two. For an infinitely large number of scenarios, an infinitely large set of independent assets is needed to ensure perfect hedging. The second way to go is to assume that the market for the underlying opens at some intermediate time τ prior to T and the underlying on that date may take values $\{S(H), S(L)\}$. If this is the case, one could use the following strategy:

- Evaluate $g(\tau)$ under both scenarios $\{S(H), S(L)\}$, yielding $\{g(H), g(L)\}$: this will result in the computation of the risk-neutral probabilities $\{Q(H), Q(L)\}$ and the replicating portfolios consisting of $\{\Delta_g(H), \Delta_g(L)\}$ units of the underlying and $\{\Pi_g(H), \Pi_g(L)\}$ units of the risk-free asset.
- Evaluate $g(t)$ as a derivative product giving a pay-off $\{g(H), g(L)\}$ at time τ, depending on the state of the world: this will result in a risk-neutral probability Q, and a replicating portfolio with Δ_g units of the underlying and Π_g units of the risk-free asset.

The result is that the value of the product will be again set equal to its replicating portfolio

$$g(t) = \Delta_g S(t) + B\Pi_g$$

but at time τ it will be *rebalanced*, depending on the price observed for the underlying asset. We will then have

$$g(H) = \Delta_g(H)S(H) + B\Pi_g(H)$$
$$g(L) = \Delta_g(L)S(L) + B\Pi_g(L)$$

and both the position on the underlying asset and the risk-free asset will be changed following the change of the underlying price. We see that even though we have three possible scenarios, we can replicate the product g by a replicating portfolio of only two assets, thanks to the possibility of changing it at an intermediate date. We say that we follow a *dynamic* replication trading strategy, opposed to the *static* replication portfolio of the simple example above. The replication trading strategy has a peculiar feature: the value of the replicating portfolio set up at t and re-evaluated using the prices of time τ is, in any circumstances, equal to that of the new replicating portfolio which will be set up at time τ. We have in fact that

$$\Delta_g S(H) + \Pi_g = g(H) = \Delta_g(H) S(H) + B\Pi_g(H)$$

$$\Delta_g S(L) + \Pi_g = g(L) = \Delta_g(L) S(L) + B\Pi_g(L)$$

This means that once the replicating portfolio is set up at time t, no further expense or withdrawal will be required to rebalance it, and the sums to be paid to buy more of an asset will be exactly those made available by the selling of the other. For this reason the replicating portfolio is called *self-financing*.

1.3 THE BLACK–SCHOLES MODEL

Let us think of a multiperiod binomial model, with a time difference between one date and the following equal to h. The gain or loss on an investment on asset S over every period will be given by

$$S(t+h) - S(t) = i_S(t) S(t)$$

Now assume that the rates of return are serially uncorrelated and normally distributed as

$$i_S(t) = \mu^* + \sigma^* \varepsilon(t)$$

with μ^* and σ^* constant parameters and $\varepsilon(t) \sim N(0, 1)$, i.e. a series of uncorrelated standard normal variables. Substituting in the dynamics of S we get

$$S(t+h) - S(t) = \mu^* S(t) + \sigma^* S(t) \varepsilon(t)$$

Taking the limit for h that tends to zero, we may write the stochastic dynamics of S in continuous time as

$$\mathrm{d}S(t) = \mu S(t)\, \mathrm{d}t + \sigma S(t)\, \mathrm{d}z(t)$$

The stochastic process is called *geometric brownian motion*, and it is a specific case of a *diffusive* process. $z(t)$ is a Wiener process, defined by $\mathrm{d}z(t) \sim N(0, \mathrm{d}t)$ and the terms $\mu S(t)$ and $\sigma S(t)$ are known as the *drift* and *diffusion* of the process. Intuitively, they represent the expected value and the volatility (standard deviation) of instantaneous changes of $S(t)$.

Technically, a stochastic process in continuous time $S(t), t \leqslant T$, is defined with respect to a filtered probability space $\{\Omega, \Im_t, P\}$, where $\Im_t = \sigma(S(u), u \leqslant t)$ is the smallest σ-field containing sets of the form $\{a \leqslant S(u) \leqslant b\}$, $0 \leqslant u \leqslant t$: more intuitively, $\Im_t$ represents the amount of information available at time t.

The increasing σ-fields $\{\Im_t\}$ form a so-called filtration F:

$$\Im_0 \subset \Im_1 \subset \cdots \subset \Im_T$$

Not only is the filtration increasing, but $\Im_0$ also contains all the events with zero measure; and these are typically referred to as "the usual assumptions". The increasing property corresponds to the fact that, at least in financial applications, the amount of information is continuously increasing as time elapses.

A variable observed at time t is said to be measurable with respect to $\Im_t$ if the set of events, such that the random variable belongs to a Borel set on the line, is contained in $\Im_t$, for every Borel set: in other words, $\Im_t$ contains all the amount of information needed to recover the value of the variable at time t. If a process $S(t)$ is measurable with respect to $\Im_t$ for all $t \geqslant 0$, it is said to be adapted with respect to $\Im_t$. At time t, the values of a variable at any time $\tau > t$ can instead be characterized only in terms of the last object, i.e. the probability measure P, conditional on the information set $\Im_t$.

In this setting, a diffusive process is defined, assuming that the limit of the first and second moments of $S(t+h) - S(t)$ exist and are finite, and that finite jumps have zero probability in the limit. Technically,

$$\lim_{h\to 0} \frac{1}{h} E\left[S(t+h) - S(t) \mid S(t) = S\right] = \mu(S,t)$$

$$\lim_{h\to 0} \frac{1}{h} E\left[[S(t+h) - S(t)]^2 \mid S(t) = S\right] = \sigma^2(S,t)$$

and

$$\lim_{h\to 0} \frac{1}{h} \Pr\left(|S(t+h) - S(t)| > \varepsilon \mid S(t) = S\right) = 0$$

Of course the moments in the equations above are tacitly assumed to exist. For further and detailed discussion of the matter, the reader is referred to standard textbooks on stochastic processes (see, for example, Karlin & Taylor, 1981).

1.3.1 Ito's lemma

A paramount result that is used again and again in financial applications is Ito's lemma. Say $y(t)$ is a diffusive stochastic process

$$\mathrm{d}y(t) = \mu_y \,\mathrm{d}t + \sigma_y \,\mathrm{d}z(t)$$

and $f(y,t)$ is a function differentiable twice in the first argument and once in the second. Then f also follows a diffusive process

$$\mathrm{d}f(y,t) = \mu_f \,\mathrm{d}t + \sigma_f \,\mathrm{d}z(t)$$

with drift and diffusion terms given by

$$\mu_f = \frac{\partial f}{\partial t} + \frac{\partial f}{\partial y}\mu_y + \frac{1}{2}\frac{\partial^2 f}{\partial y^2}\sigma_y^2$$

$$\sigma_f = \frac{\partial f}{\partial y}\sigma_y$$

Example 1.1 Notice that, given

$$dS(t) = \mu S(t)\ dt + \sigma S(t)\ dz(t)$$

we can set $f(S,t) = \ln S(t)$ to obtain

$$d\ln S(t) = (\mu - \tfrac{1}{2}\sigma^2)\, dt + \sigma dz(t)$$

If μ and σ are constant parameters, it is easy to obtain

$$\ln S(\tau) \mid \Im_t \sim N(\ln S(t) + (\mu - \tfrac{1}{2}\sigma^2)(\tau - t), \sigma^2(\tau - t))$$

where $N(m,s)$ is the normal distribution with mean m and variance s. Then, $\Pr(S(\tau) \mid \Im_t)$ is described by the lognormal distribution.

It is worth stressing that the *geometric brownian motion* assumption used in the Black–Scholes model implies that the log-returns on the asset S are normally distributed, and this is the same as saying that their volatility is assumed to be constant.

1.3.2 Girsanov theorem

A second technique that is mandatory to know for the application of diffusive processes to financial problems is the result known as the Girsanov theorem (or Cameron–Martin–Girsanov theorem). The main idea is that given a Wiener process $z(t)$ defined under the filtration $\{\Omega, \Im_t, P\}$ we may construct another process $\widetilde{z}(t)$ which is a Wiener process under another probability space $\{\Omega, \Im_t, Q\}$. Of course, the latter process will have a drift under the original measure P. Under such measure it will be in fact

$$d\widetilde{z}(t) = dz(t) + \gamma\, dt$$

for γ deterministic or stochastic and satisfying regularity conditions. In plain words, changing the probability measure is the same as changing the drift of the process.

The application of this principle to our problem is straightforward. Assume there is an opportunity to invest in a *money market mutual fund* yielding a constant instantaneous risk-free yield equal to r. In other words, let us assume that the dynamics of the investment in the risk-free asset is

$$dB(t) = rB(t)$$

where the constant r is also called the interest rate intensity ($r \equiv \ln(1+i)$). We saw before that under the objective measure P the no-arbitrage requirement implies

$$E\left[\frac{dS(t)}{S(t)}\right] = \mu\, dt = (r + \lambda\sigma)\ dt$$

where λ is the market price of risk. Substituting in the process followed by $S(t)$ we have

$$\begin{aligned} \mathrm{d}S(t) &= (r + \lambda\sigma) S(t)\ \mathrm{d}t + \sigma S(t)\ \mathrm{d}z(t) \\ &= S(t)(r\,\mathrm{d}t + \sigma(\mathrm{d}z(t) + \lambda\,\mathrm{d}t)) \\ &= S(t)(r\,\mathrm{d}t + \sigma\,\mathrm{d}\widetilde{z}(t)) \end{aligned}$$

where $\mathrm{d}\widetilde{z}(t) = \mathrm{d}z(t) + \lambda\,\mathrm{d}t$ is a Wiener process under some new measure Q. Under such a measure, the dynamics of the underlying is then

$$\mathrm{d}S(t) = rS(t)\ \mathrm{d}t + \sigma S(t)\ \mathrm{d}\widetilde{z}(t)$$

meaning that the instantaneous expected rate of the return on asset $S(t)$ is equal to the instantaneous yield on the risk-free asset

$$E_Q\left[\frac{\mathrm{d}S(t)}{S(t)}\right] = r\,\mathrm{d}t$$

i.e. that Q is the so-called risk-neutral measure. It is easy to check that the same holds for any derivative written on $S(t)$. Define $g(S, t)$ the price of a derivative contract giving pay-off $G(S(T), T)$. Indeed, using Ito's lemma we have

$$\mathrm{d}g(t) = \mu_g g(t)\ \mathrm{d}t + \sigma_g g(t)\ \mathrm{d}z(t)$$

with

$$\begin{aligned} \mu_g g &= \frac{\partial g}{\partial t} + \frac{\partial g}{\partial S}(r + \lambda\sigma) S(t) + \frac{1}{2}\frac{\partial^2 g}{\partial S^2}\sigma^2(t) S^2 \\ \sigma_g g &= \frac{\partial g}{\partial S}\sigma \end{aligned}$$

Notice that under the original measure we then have

$$\mathrm{d}g(t) = \left[\frac{\partial g}{\partial t} + \frac{\partial g}{\partial S}\mu S(t) + \frac{1}{2}\frac{\partial^2 g}{\partial S^2}\sigma^2(t) S^2\right]\mathrm{d}t + \frac{\partial g}{\partial S}\sigma\ \mathrm{d}z(t)$$

However, the no-arbitrage requirement implies

$$\mu_g g = \frac{\partial g}{\partial t} + \frac{\partial g}{\partial S}(r + \lambda\sigma) S(t) + \frac{1}{2}\frac{\partial^2 g}{\partial S^2}\sigma^2(t) S^2 = rg + \lambda\frac{\partial g}{\partial S}\sigma$$

so it follows that

$$\frac{\partial g}{\partial t} + \frac{\partial g}{\partial S} rS(t) + \frac{1}{2}\frac{\partial^2 g}{\partial S^2}\sigma^2(t) S^2 = rg$$

This is the fundamental partial differential equation (PDE) of the Black–Scholes model. Notice that by substituting this result into the risk-neutral dynamics of g under measure Q we get

$$\mathrm{d}g(t) = rg(t)\ \mathrm{d}t + \frac{\partial g}{\partial S}\sigma\ \mathrm{d}\widetilde{z}(t)$$

and the product g is expected to yield the instantaneous risk-free rate. We reach the conclusion that under the risk-neutral measure Q

$$E_Q\left[\frac{\mathrm{d}S(t)}{S(t)}\right] = E_Q\left[\frac{\mathrm{d}g(t)}{g(t)}\right] = r\,\mathrm{d}t$$

that is, all the risky assets are assumed to yield the instantaneous risk-free rate.

1.3.3 The martingale property

The price of any contingent claim g can be recovered solving the fundamental PDE. An alternative way is to exploit the martingale property embedded in the measure Q. Define Z as the value of a product expressed using the riskless money market account as the numeraire, i.e. $Z(t) \equiv g(t)/B(t)$. Given the dynamics of the risky asset under the risk-neutral measure Q we have that

$$\mathrm{d}S(t) = rS(t)\,\mathrm{d}t + \sigma S(t)\,\mathrm{d}\widetilde{z}(t)$$
$$\mathrm{d}B(t) = rB(t)\,\mathrm{d}t$$

and it is easy to check that

$$\mathrm{d}Z(t) = \sigma Z(t)\,\mathrm{d}\widetilde{z}(t)$$

The process $Z(t)$ then follows a martingale, so that $E_Q(Z(T)) = Z(t)$. This directly provides us with a pricing formula. In fact we have

$$Z(t) = \frac{g(S,t)}{B(t)} = E_Q(Z(T)) = E_Q\left(\frac{G(S,T)}{B(T)}\right)$$

Considering that $B(T)$ is a deterministic function, we have

$$g(S,t) = \frac{B(t)}{B(T)}E_Q(G(S,T)) = \exp(-r(T-t))\,E_Q(G(S,T))$$

The price of a contingent claim is obtained by taking the relevant expectation under the risk-neutral measure and discounting it back to the current time t. Under the assumption of log-normal distribution of the future price of the underlying asset S, we may recover for instance the basic Black–Scholes formula for a plain vanilla call option

$$\mathrm{CALL}(S,t;K,T) = S(t)\,\Phi(d_1) - \exp[-r(T-t)]\,K\Phi(d_2)$$
$$d_1 = \frac{\ln(S(t)/K) + \left(r+\sigma^2/2\right)(T-t)}{\sigma\sqrt{T-t}}$$
$$d_2 = d_1 - \sigma\sqrt{T-t}$$

where $\Phi(x)$ is the standard normal distribution function evaluated at x

$$\Phi(x) = \frac{1}{2\pi}\int_{-\infty}^{x}\exp\left[-\frac{u^2}{2}\right]\mathrm{d}u$$

The formula for the put option is, instead,

$$\mathrm{PUT}(S, t; K, T) = -S(t)\,\Phi(-d_1) + \exp[-r(T-t)]\,K\Phi(-d_2)$$

Notice that a long position in a call option corresponds to a long position in the underlying and a debt position, while a long position in a put option corresponds to a short position in the underlying and an investment in the risk-free asset. As $S(t)$ tends to infinity, the value of a call tends to that of a long position in a forward and the value of the put tends to zero; as $S(t)$ tends to zero, the value of the put tends to the value of a short position in a forward and the price of the call option tends to zero.

The sensitivity of the option price with respect to the underlying is called *delta* (Δ) and is equal to $\Phi(d_1)$ for the call option and $\Phi(d_1) - 1$ for the put. The sensitivity of the delta with respect to the underlying is called *gamma* (Γ), and that of the option price with respect to time is called *theta* (Θ). These derivatives, called the *greek letters*, can be used to approximate, in general, the value of any derivative contract by a Taylor expansion as

$$\begin{aligned} g(S(t+h), t+h) &\simeq g(S(t), t) + \Delta_g (S(t+h) - S(t)) \\ &\quad + \tfrac{1}{2}\Gamma_g (S(t+h) - S(t))^2 + \Theta_g h \end{aligned}$$

Notice that the *greek letters* are linked one to the others by the *fundamental PDE* ruling out arbitrage. Indeed, this condition can be rewritten as

$$\Theta_g + \Delta_g r S(t) + \tfrac{1}{2}\Gamma_g \sigma^2(t)\, S^2 - rg = 0$$

1.3.4 Digital options

A way to understand the probabilistic meaning of the Black–Scholes formula is to compute the price of digital options. Digital options pay a fixed sum or a unit of the underlying if the underlying asset is above some strike level at the exercise date. Digital options, which pay a fixed sum, are called *cash-or-nothing* (CoN) options, while those paying the asset are called *asset-or-nothing* (AoN) options. Under the log-normal assumption of the conditional distribution of the underlying held under the Black–Scholes model, we easily obtain

$$\mathrm{CoN}(S, t; K, T) = \exp[-r(T-t)]\,\Phi(d_2)$$

The asset-or-nothing price can be recovered by arbitrage observing that at time T

$$\mathrm{CALL}(S, T; K, T) + K\ \mathrm{CoN}(S, T; K, T) = \mathbf{1}_{\{S(T)>K\}} S(T) = \mathrm{AoN}(S, T; K, T)$$

where $\mathbf{1}_{\{S(T)>K\}}$ is the indicator function assigning 1 to the case $S(T) > K$. So, to avoid arbitrage we must have

$$\mathrm{AoN}(S, t; K, T) = S(t)\,\Phi(d_1)$$

Beyond the formulas deriving from the Black–Scholes model, it is important to stress that this result – that a call option is the sum of a long position in a digital *asset-or-nothing* option and a short position in K *cash-or-nothing* options – remains true for all the option pricing models. In fact, this result directly stems from the no-arbitrage requirement imposed in the asset pricing model. The same holds for the result (which may be easily verified) that

$$-\exp\left[r\left(T-t\right)\right]\frac{\partial \text{CALL}(S,t;K,T)}{\partial K}=\Phi\left(d_2\right)=\Pr\left(S\left(T\right)>K\right)$$

where the probability is computed under measure Q. From the derivative of the call option with respect to the strike price we can then recover the risk-neutral probability of the underlying asset.

1.4 INTEREST RATE DERIVATIVES

The valuation of derivatives written on fixed income products or interest rates is more involved than the standard Black–Scholes model described above, even though all models are based on the same principles and techniques of arbitrage-free valuation presented above. The reason for this greater complexity is that the underlying asset of these products is the curve representing the discounting factors of future cash-flows as a function of maturity T. The discount factor $D(t,T)$ of a unit cash-flow due at maturity T, evaluated at current time t, can be represented as

$$D\left(t,T\right)=\exp\left[-r\left(t,T\right)\left(T-t\right)\right]$$

where $r\left(t,T\right)$ is the continuously compounded spot rate or yield to maturity. Alternatively, the discount factor can be characterized in terms of forward rates, as

$$D\left(t,T\right)=\exp\left[-\int_t^T f\left(t,u\right)\,\mathrm{d}u\right]$$

Term structure pricing models are based on stochastic representations of the spot or forward yield curve.

1.4.1 Affine factor models

The classical approach to interest rate modeling is based on the assumption that the stochastic dynamics of the curve can be represented by the dynamics of some risk factors. The yield curve is then recovered endogenously from their dynamics. The most famous models are due to Vasicek (1977) and Cox, Ingersoll and Ross (1985). They use a single risk factor, which is chosen to be the intercept of the yield curve – that is, the instantaneous interest rate. While this rate was assumed to be constant under the Black–Scholes framework, now it is assumed to vary stochastically over time, so that the value of a European contingent claim g, paying $G(T)$ at time T, is generalized to

$$g\left(t\right)=E_Q\left[\exp\left[-\int_t^T r\left(u\right)\,\mathrm{d}u\right]G\left(T\right)\mid \Im_t\right]$$

where the expectation is again taken under the risk-neutral measure Q. Notice that for the discount factor $D\left(t,T\right)$ we have the pay-off $D\left(T,T\right)=1$, so that

$$D\left(t,T\right)=E_Q\left[\exp\left[-\int_t^T r\left(u\right)\,\mathrm{d}u\right]\middle|\,\Im_t\right]$$

We observe that even if the pay-off is deterministic, the discount factor is stochastic, and it is a function of the instantaneous interest rate $r(t)$. Let us assume that the dynamics of $r(t)$ under the risk-neutral measure is described by the diffusion process

$$\mathrm{d}r(t) = \mu_r \,\mathrm{d}t + \sigma_r \,\mathrm{d}\widetilde{z}(t)$$

and let us write the dynamics of the discount factor, under the same measure Q, as

$$\mathrm{d}D(t,T) = r(t)\, D(t,T)\,\mathrm{d}t + \sigma_T D(t,T)\,\mathrm{d}\widetilde{z}(t)$$

where σ_T is the volatility of instantaneous percentage changes of the discount factor. Applying Ito's lemma we have

$$r(t)\, D(t,T) = \frac{\partial D(t,T)}{\partial t} + \mu_r \frac{\partial D(t,T)}{\partial r} + \frac{1}{2}\sigma_r^2 \frac{\partial^2 D(t,T)}{\partial r^2}$$

which is a partial differential equation ruling out arbitrage opportunities.

It may be proved that in the particular case in which

$$\mu_r = \alpha + \beta r$$
$$\sigma_r^2 = \gamma + \zeta r$$

that is, in the case in which both the drift and the instantaneous variance are linear in the risk factor, the solution is

$$D(t,T) = \exp\left[A(T-t) - M(T-t)\, r(t)\right]$$

These models are called *affine factor models*, because interest rates are affine functions of the risk factor.

The general shape of the instantaneous drift used in one-factor affine models is $\mu_r = k(\theta - r)$, so that the interest rate is recalled toward a long run equilibrium level θ: this feature of the model is called *mean reversion*. Setting $\zeta = 0$ and $\gamma > 0$ then leads to the Vasicek model, in which the conditional distribution of the instantaneous interest rate is normal. Alternatively, assuming $\zeta > 0$ and $\gamma = 0$ then leads to the famous Cox, Ingersoll and Ross model: the stochastic process followed by the instantaneous interest rate is a *square root* process, and the conditional distribution is non-central chi-square. The case in which $\zeta > 0$ and $\gamma > 0$ is a more general process studied in Pearson and Sun (1994). Finally, the affine factor model result was proved in full generality with an extension to an arbitrary number of risk factors by Duffie and Kan (1996).

Looking at the solution for the discount factor $D(t,T)$, it is clear that the function $M(T-t)$ is particularly relevant, because it represents its sensitivity to the risk factor $r(t)$. In fact, using Ito's lemma we may write the dynamics of $D(t,T)$ under the risk-neutral measure as

$$\mathrm{d}D(t,T) = r(t)\, D(t,T)\,\mathrm{d}t + \sigma_T M(T-t)\, D(t,T)\,\mathrm{d}\widetilde{z}(t)$$

1.4.2 Forward martingale measure

Consider now the problem of pricing a contingent claim whose pay-off is a function of the interest rate. Remember that, differently from the Black–Scholes framework, the discount factor to be applied to the contingent claim is now stochastic and, if the underlying is an interest rate sensitive product, it is not independent from the pay-off. The consequence is that the discount factor and the expected pay-off under the risk-neutral measure cannot be factorized. To make a simple example, consider a call option written on a zero coupon bond maturing at time T, for strike K and exercise time τ. We have:

$$\begin{aligned} \text{CALL}\left(D\left(t,T\right),t;\tau,K\right) &= E_Q\left[\exp\left[-\int_t^\tau r\left(u\right)\,\mathrm{d}u\right]\max\left[D\left(\tau,T\right)-K,0\right]\middle|\,\Im_t\right] \\ &\neq E_Q\left[\exp\left[-\int_t^\tau r\left(u\right)\,\mathrm{d}u\right]\middle|\,\Im_t\right] \\ &\quad E_Q\left[\max\left[D\left(\tau,T\right)-K,0\right]\middle|\,\Im_t\right] \end{aligned}$$

and the price cannot be expressed as the product of the discount factor $D\left(t,\tau\right)$ and the expected pay-off. Factorization can, however, be achieved through a suitable change of measure.

Consider the discount factors evaluated at time t for one unit of currency to be received at time τ and T respectively, with $\tau < T$. Their dynamics under the risk-neutral measure are

$$\mathrm{d}D\left(t,T\right) = r\left(t\right)D\left(t,T\right)\,\mathrm{d}t + \sigma_T D\left(t,T\right)\,\mathrm{d}\widetilde{z}\left(t\right)$$
$$\mathrm{d}D\left(t,\tau\right) = r\left(t\right)D\left(t,\tau\right)\,\mathrm{d}t + \sigma_\tau D\left(t,\tau\right)\,\mathrm{d}\widetilde{z}\left(t\right)$$

We can define $D\left(t,\tau,T\right)$ as the *forward price* set at time t for an investment starting at time τ and yielding one unit of currency at time T. A standard no-arbitrage argument yields

$$D\left(t,\tau,T\right) = \frac{D\left(t,T\right)}{D\left(t,\tau\right)}$$

The dynamics of the forward price can be recovered by using Ito's division rule.

Remark 1.1 [*Ito's division rule*] Assume two diffusive processes $X\left(t\right)$ and $Y\left(t\right)$ following the dynamics

$$\mathrm{d}X\left(t\right) = \mu_X X\left(t\right)\,\mathrm{d}t + \sigma_X X\left(t\right)\mathrm{d}z\left(t\right)$$
$$\mathrm{d}Y\left(t\right) = \mu_Y Y\left(t\right)\,\mathrm{d}t + \sigma_Y Y\left(t\right)\mathrm{d}z\left(t\right)$$

Then, the process $F\left(t\right) \equiv X\left(t\right)/Y\left(t\right)$ follows the dynamics

$$\mathrm{d}F\left(t\right) = \mu_F F\left(t\right)\,\mathrm{d}t + \sigma_F F\left(t\right)\mathrm{d}z\left(t\right)$$

with

$$\sigma_F = \sigma_X - \sigma_Y$$
$$\mu_F = \mu_X - \mu_Y - \sigma_F\sigma_Y$$

Applying this result to our problem yields immediately

$$dD(t,\tau,T) = -\sigma_F\sigma_\tau D(t,\tau,T)\,dt + \sigma_F D(t,\tau,T)\,d\widetilde{z}(t)$$
$$\sigma_F = \sigma_T - \sigma_\tau$$

We may now use the Girsanov theorem to recover a new measure Q_τ under which $d\widehat{z} = d\widetilde{z} - \sigma_\tau dt$ is a Wiener process. We have then

$$dD(t,\tau,T) = \sigma_F D(t,\tau,T)\,d\widehat{z}(t)$$

and the forward price is a martingale. Under such a measure, the forward price of any future contract is equal to the expected spot value. We have

$$D(t,\tau,T) = E_{Q_\tau}\left[(D(\tau,\tau,T)) \mid \Im_t\right] = E_{Q_\tau}\left[(D(\tau,T)) \mid \Im_t\right]$$

and the measure Q_τ is called the *forward martingale measure*. This result, which was first introduced by Geman (1989) and Jamshidian (1989), is very useful to price interest rate derivatives. In fact, consider a derivative contract g, written on $D(t,T)$, promising the pay-off $G(D(\tau,T),\tau)$ at time τ. As $g(t)/D(t,\tau)$ is a martingale, we have immediately

$$g(D(t,\tau,T),t) = P(t,\tau)\,E_{Q_\tau}\left[G(D(\tau,T),\tau) \mid \Im_t\right]$$

and the factorization of the discount factor and expected pay-off is now correct.

To conclude, the cookbook recipe emerging from the forward martingale approach is that the forward price must be considered as the underlying asset of the derivative contract, instead of the spot.

1.4.3 LIBOR market model

While the standard classical interest rate pricing models are based on the dynamics of instantaneous spot and forward rates, the market practice is to refer to observed interest rates for investment over discrete time periods. In particular, the reference rate mostly used for short-term investments and indexed products is the 3-month LIBOR rate. Moreover, under market conventions, interest rates for investments below the one-year horizon are computed under simple compounding. So, the LIBOR interest rate for investment from t to T is defined as

$$L(t,T) = \frac{1}{T-t}\left(\frac{1}{D(t,T)} - 1\right)$$

The corresponding forward rate is defined as

$$L(t,\tau,T) = \frac{1}{T-\tau}\left(\frac{1}{D(t,\tau,T)} - 1\right) = \frac{1}{T-\tau}\left(\frac{D(t,\tau)}{D(t,T)} - 1\right)$$

Notice that, under the forward martingale measure Q_T, we have immediately

$$L(t,\tau,T) = E_{Q_T}\left[L(\tau,\tau,T) \mid \Im_t\right] = E_{Q_T}\left[L(\tau,T) \mid \Im_t\right]$$

The price of a floater, i.e. a bond whose coupon stream is indexed to the LIBOR, is then evaluated as

$$\begin{aligned}\mathrm{FLOATER}(t, t_N) &= \sum_{j=1}^{N} \delta_i E_Q[D(t, t_j)L(t_{j-1}, t_j) \mid \Im_t] + P(t, t_N) \\ &= \sum_{j=1}^{N} \delta_i D(t, t_j)E_{Q_{t_j}}[L(t_{j-1}, t_j) \mid \Im_t] + P(t, t_N) \\ &= \sum_{j=1}^{N} D(t, t_j)\delta_i L(t, t_{j-1}, t_j) + P(t, t_N)\end{aligned}$$

where the set $\{t, t_1, t_2, \ldots, t_N\}$ contains the dates at which a coupon is reset and the previous one is paid and $\delta_j = t_j - t_{j-1}$. Consider now a stream of call options written on the index rate for each coupon period. This product is called a *cap*, and the price is obtained, assuming a strike rate L_{CAP}, from

$$\begin{aligned}\mathrm{CAP}(t, t_1, t_N) &= \sum_{j=2}^{N} \delta_i E_Q[D(t, t_j) \max(L(t_{j-1}, t_j) - L_{\mathrm{CAP}}), 0 \mid \Im_t] \\ &= \sum_{j=2}^{N} \delta_i D(t, t_j)E_{Q_{t_j}}[\max(L(t_{j-1}, t_j) - L_{\mathrm{CAP}}), 0 \mid \Im_t]\end{aligned}$$

and each call option is called *caplet*. By the same token, a stream of put options are called *floor*, and are evaluated as

$$\mathrm{FLOOR}(t, t_1, t_N) = \sum_{j=1}^{N} \delta_2 D(t, t_j)E_{Q_{t_j}}[\max(L_{\mathrm{FLOOR}} - L(t_{j-1}, t_j)), 0 \mid \Im_t]$$

where L_{FLOOR} is the strike rate. The names *cap* and *floor* derive from the results, which may be easily verified

$$L(t_j, t_{j-1}) - \mathrm{CAPLET}(t_j, t_{j-1}) = \min(L(t_{j-1}, t_j), L_{\mathrm{CAP}})$$
$$L(t_j, t_{j-1}) + \mathrm{FLOORLET}(t_j, t_{j-1}) = \max(L(t_{j-1}, t_j), L_{\mathrm{FLOOR}})$$

Setting a cap and a floor amounts to building a *collar*, that is a band in which the coupon is allowed to float according to the interest rate. The price of each caplet and floorlet can then be computed under the corresponding forward measure. Under the assumption that each forward rate is log-normally distributed, we may again recover a pricing formula largely used in the market, known as Black's formula.

$$\begin{aligned}\mathrm{CAPLET}(t; t_j, t_{j-1}) &= D(t, t_j)E_{Q_{t_j}}[\max(L(t_{j-1}, t_j) - L_{\mathrm{CAP}}), 0 \mid \Im_t] \\ &= D(t, t_j)\{E_{Q_{t_j}}[L(t_{j-1}, t_j) \mid \Im_t]N(d_1) - L_{\mathrm{CAP}}N(d_2)\} \\ &= D(t, t_j)L(t, t_{j-1}, t_j)N(d_1) - D(t, t_j)L_{\mathrm{CAP}}N(d_2)\end{aligned}$$

$$d_1 = \frac{\ln(L(t, t_{j-1}, t_j)/L_{\mathrm{CAP}}) + \sigma_j^2(t_j - t)}{\sigma_j\sqrt{t_j - t}}$$

$$d_2 = d_1 - \sigma_j\sqrt{t_j - t}$$

where σ_j is the instantaneous volatility of the logarithm of the forward rate $L\left(t, t_{j-1}, t_j\right)$.

The floorlet price is obtained by using the corresponding put option formula

$$\mathrm{FLOORLET}(t; t_j, t_{j-1}) = -D(t, t_j)L(t, t_{j-1}, t_j)N(-d_1) + D(t, t_j)L_{\mathrm{CAP}}N(-d_2)$$

$$d_1 = \frac{\ln(L(t, t_{j-1}, t_j)/L_{\mathrm{FLOOR}}) + \sigma_j^2(t_j - t)}{\sigma_j\sqrt{t_j - t}}$$

$$d_2 = d_1 - \sigma_j\sqrt{t_j - t}$$

1.5 SMILE AND TERM STRUCTURE EFFECTS OF VOLATILITY

The Black–Scholes model, which, as we saw, can be applied to the pricing of contingent claims on several markets, has been severely challenged by the data. The contradiction emerges from a look at the market quotes of options and a comparison with the implied information, that is, with the dynamics of the underlying that would make these prices consistent. In the Black–Scholes setting, this information is collected in the same parameter, volatility, which is assumed to be constant both across time and different states of the world. This parameter, called *implied* volatility, represents a sufficient statistic for the risk-neutral probability in the Black–Scholes setting: the instantaneous rate of returns on the assets are in fact assumed normal and with first moments equal to the risk-free rate. Contrary to this assumption, implied volatility is typically different both across different strike prices and different maturities. The first evidence is called the *smile* effect and the second is called the *volatility term structure*.

Non-constant implied volatility can be traced back to market imperfections or it may actually imply that the stochastic process assumed for the underlying asset is not borne out by the data, namely that the rate of return on the assets is not normally distributed. The latter interpretation is indeed supported by a long history of evidence on non-normality of returns on almost every market. This raises the question of which model to adopt to get a better fit of the risk-neutral distribution and market data.

1.5.1 Stochastic volatility models

A first approach is to model volatility as a second risk factor affecting the price of the derivative contract. This implies two aspects, which may make the model involved. The first is the dependence structure between volatility and the underlying. The second is that the risk factor represented by volatility must be provided with a market price, something that makes the model harder to calibrate.

A model that is particularly easy to handle, and reminds us of the Hull and White (1987) model, could be based on the assumption that volatility risk is not priced in the market, and volatility is orthogonal to the price of the underlying. The idea is that conditional on a given volatility parameter taking value s, the stochastic process followed by the underlying asset follows a geometric brownian motion. The conditional value of the call would then yield the standard Black–Scholes solution. As volatility is stochastic and is not known at the time of evaluation, the option is priced by integrating the Black–Scholes formula times the volatility density across its whole support. Analytically, the pricing formula for a call option yields, for example,

$$\text{CALL}\,(S(t), t, \sigma(t); K, T) = \int_0^\infty \text{CALL}_{\text{BS}}\,(S, t; \sigma(t) = s, K, T)\, q_\sigma\,(s \mid \Im_t)\,\mathrm{d}s$$

where CALL_{BS} denotes the Black–Scholes formula for call options and $q_\sigma\,(s \mid \Im_t)$ represents the volatility conditional density.

Extensions of this model account for a dependence structure between volatility and the underlying asset. A good example could be to model instantaneous variance as a square root process, to exploit its property to be defined on the non-negative support only and the possibility, for some parameter configurations, of making zero volatility an *inaccessible barrier*. Indeed, this idea is used both in Heston (1993) and in Longstaff and Schwartz (1992) for interest rate derivatives.

1.5.2 Local volatility models

A different idea is to make the representation of the diffusive process more general by modeling volatility as a function of the underlying asset and time. We have then, under the risk-neutral measure

$$\mathrm{d}S\,(t) = rS\,(t)\,\mathrm{d}t + \sigma\,(S, t)\,S\,(t)\,\mathrm{d}\widetilde{z}\,(t)$$

The function $\sigma\,(S, t)$ is called the *local volatility surface* and should then be calibrated in such a way as to produce the smile and volatility term structure effects actually observed on the market. A long-dated proposal is represented by the so-called *constant elasticity of variance* (CEV) models, in which

$$\mathrm{d}S\,(t) = rS\,(t)\,\mathrm{d}t + \sigma S\,(t)^\alpha\,\mathrm{d}\widetilde{z}\,(t)$$

Alternative local volatility specifications were proposed to comply with techniques that are commonly used by practitioners in the market to fit the smile. An idea is to resort to the so-called *mixture of log-normal* or *shifted log-normal* distributions. Intuitively, this approach leads to closed form valuations. For example, assuming that the risk-neutral probability distribution Q is represented by a linear combination of n log-normal distributions Q_j

$$Q\,(S\,(T) \mid \Im_t) = \sum_{j=1}^{n} \lambda_j Q_j\left(X_j\,(T) \mid \Im_t\right)$$

where X_j are latent random variables drawn from log-normal distributions Q_j, corresponding to geometric brownian motions with volatility σ_j. It may be checked that the price of a call option in this model can be recovered as

$$\text{CALL}\,(S(t), t; K, T) = \sum_{j=1}^{n} \lambda_j \text{CALL}_{\text{BS}}\left(X_j, t; K, T\right)$$

Brigo and Mercurio (2001) provide the corresponding local volatility specification corresponding to this model, obtaining

$$\mathrm{d}S(t) = rS(t)\,\mathrm{d}t + \sqrt{\frac{\sum_{j=1}^{n} \sigma_j^2 \lambda_j q_j (X_j(T) \mid \Im_t)}{\sum_{j=1}^{n} \lambda_j q_j (X_j(T) \mid \Im_t)}}\, S(t)\,\mathrm{d}\widetilde{z}(t)$$

where $q_j\left(X_j\left(T\right) \mid \Im_t\right)$ are the densities corresponding to the distribution functions Q_j. Once the mixture weights λ_j are recovered from observed plain vanilla option prices, the corresponding dynamics of the underlying asset under the risk-neutral measure can be simulated in order to price exotic products.

1.5.3 Implied probability

A different idea is to use non-parametric techniques to extract general information concerning the risk-neutral probability distribution and dynamics implied by observed options market quotes. The concept was first suggested by Breeden and Litzenberger (1978) and pushes forward the usual implied volatility idea commonly used in the Black–Scholes framework. This is the approach that we will use in this book.

The basic concepts stem from the martingale representation of option prices. Take, for example, a call option

$$\mathrm{CALL}\left(S, t; K, T\right) = \exp\left[-r\left(T - t\right)\right] E_Q\left[\max\left(S\left(T\right) - K, 0\right)\right]$$

By computing the derivative of the pricing function with respect to the strike K we easily obtain

$$\frac{\partial \mathrm{CALL}\left(S, t; K, T\right)}{\partial K} = -\exp\left[-r\left(T - t\right)\right]\left(1 - Q\left(K \mid \Im_t\right)\right)$$

where $Q(K \mid \Im_t)$ is the conditional distribution function under the risk-neutral measure. Defining

$$\overline{Q}\left(K \mid \Im_t\right) \equiv 1 - Q\left(K \mid \Im_t\right)$$

that is, the probability corresponding to the complementary event $\Pr\left(S\left(T\right) > K\right)$, we may rewrite

$$\overline{Q}\left(K \mid \Im_t\right) = -\exp\left[r\left(T - t\right)\right] \frac{\partial \mathrm{CALL}\left(S, t; K, T\right)}{\partial K}$$

So, the risk-neutral probability of exercise of the call option is recovered from the forward value of the derivative of the call option, apart from a change of sign. The result can be immediately verified in the Black–Scholes model, where we easily compute $\overline{Q}\left(K \mid \Im_t\right) = N\left(d_2\left(K\right)\right)$.

Remark 1.2 Notice that by integrating the relationship above from K to infinity, the price of the call option can also be written as

$$\mathrm{CALL}\left(S, t; K, T\right) = \exp\left[-r\left(T - t\right)\right] \int_K^{\infty} \overline{Q}\left(u \mid \Im_t\right)\,\mathrm{d}u$$

where we remark that the cumulative probability, rather than the density, appears in the integrand. As we will see, this pricing representation will be used again and again throughout this book.

Symmetric results hold for put prices which, in the martingale representation, are written as

$$\mathrm{PUT}\,(S, t; K, T) = \exp\left[-r\,(T-t)\right] E_Q\left[\max\left(K - S\,(T)\,, 0\right)\right]$$

Computing the derivative with respect to the strike and reordering terms we have

$$Q\,(K \mid \Im_t) = \exp\left[r\,(T-t)\right] \frac{\partial \mathrm{PUT}\,(S, t; K, T)}{\partial K}$$

that is, the implied risk-neutral distribution. Again, we may check that, under the standard Black–Scholes setting, we obtain

$$Q\,(K \mid \Im_t) = N\,(-d_2\,(K)) = 1 - N\,(d_2\,(K))$$

Furthermore, integrating from zero to K we have

$$\mathrm{PUT}\,(S, t; K, T) = \exp\left[-r\,(T-t)\right] \int_0^K Q\,(u \mid \Im_t)\,\mathrm{d}u$$

Finally, notice that the density function can be obtained from the second derivatives of the put and call prices. We have

$$q\,(K \mid \Im_t) \equiv \frac{\partial Q\,(K \mid \Im_t)}{\partial K} = \exp\left[r\,(T-t)\right] \frac{\partial^2 \mathrm{PUT}\,(S, t; K, T)}{\partial K^2}$$
$$q\,(K \mid \Im_t) \equiv -\frac{\partial \overline{Q}\,(K \mid \Im_t)}{\partial K} = \exp\left[r\,(T-t)\right] \frac{\partial^2 \mathrm{CALL}\,(S, t; K, T)}{\partial K^2}$$

The strength of these results stems from the fact that they directly rely on the no-arbitrage requirement imposed by the martingale relationship. In this sense, they are far more general than the assumptions underlying the Black–Scholes setting. Indeed, if the assumptions behind the Black–Scholes model were borne out by the data, the results above would be of little use, as all the information sufficient to characterize the risk-neutral distribution would be represented by the volatility implied by the prices. If the price distribution is not log-normal, these results are instead extremely useful, enabling one to extract the risk-neutral probability distribution, rather that its moments, directly from the option prices.

1.6 INCOMPLETE MARKETS

The most recent challenge to the standard derivative pricing model, and to its basic structure, is represented by the *incomplete market* problem. A brief look over the strategy used to recover the *fair price* of a derivative contract shows that a crucial role is played by the assumption that the future value of each financial product can be *exactly* replicated by some trading strategy. Technically, we say that each product is *attainable* and the market is *complete*. In other words, every contingent claim is endowed with a *perfect hedge*. Both in the binomial and in the continuous time model we see that it is this assumption that leads to two strong results. The first is a unique risk-neutral measure and, through that, a unique price for each and every asset in the economy. The second is that this price is obtained with no reference to any preference structure of the agents in the market, apart from the very weak (and realistic) requirement that they "prefer more to less".

Unfortunately, the completeness assumption has been fiercely challenged by the market. Every trader has always been well aware that *no perfect hedge exists*, but the structure of derivatives markets nowadays has made consideration of this piece of truth unavoidable. Structured finance has brought about a huge proliferation of customized and exotic products. Hedge funds manufacture and manage derivatives on exotic markets and illiquid products to earn money from their misalignment: think particularly of long–short and relative value hedge fund strategies. Credit derivatives markets have been created to trade protection on loans, bonds, or mortgage portfolios. All of this has been shifting the core of the derivatives market away from the traditional underlying assets traded on the organized markets, such as stocks and government bonds, toward contingent claims written on illiquid assets. The effect has been to make the problem of finding a *perfect hedge* an impossible task for most of the derivative pricing applications, and the assumption of complete markets an unacceptable approximation. The hot topic in derivative pricing is then which hedge to choose, facing the reality that no hedging strategy can be considered completely safe.

1.6.1 Back to utility theory

The main effect of accounting for market incompleteness has been to bring utility theory back in derivative pricing techniques. Intuitively, if no perfect hedge exists, every replication strategy is a lottery, and selecting one amounts to defining a preference ranking among them, which is the main subject of utility theory. In a sense, the ironic fate of finance is that the market incompleteness problem is bringing it back from a preference-free paradigm to a use of utility theory very similar to early portfolio theory applications: this trend is clearly witnessed by terms such as "minimum variance hedging" (Follmer & Schweitzer, 1991). Of course, we know that the minimum variance principle is based on restrictive assumptions concerning both the preference structure and the distributional properties of the hedging error. One extension is to use more general expected utility representations, such as exponential or power preferences, to select a specific hedging strategy and the corresponding martingale measure (Frittelli, 2000).

A question that could also be useful to debate, even though it is well beyond the scope of this book, is whether the axiomatic structure leading to the standard expected utility framework is flexible enough and appropriate to be applied to the hedging error problem. More precisely, it is well known that standard expected utility results rest on the so-called independence axiom, which has been debated and criticized in decision theory for decades, and which seems particularly relevant to the problem at hand. To explain the problem in plain words, consider you prefer hedging strategy A to another denoted B ($A \succeq B$). The independence axiom reads that you will also prefer $\alpha A + (1 - \alpha) C$ to $\alpha B + (1 - \alpha) C$ for every $\alpha \in [0, 1]$, and for whatever strategy C. This is the crucial point: the preference structure between two hedging strategies is preserved under a mixture with any other third strategy, and if this is not true the expected utility results do not carry over. It is not difficult to argue that this assumption may be too restrictive, if, for example, one considers a hedging strategy C counter-monotone to B and orthogonal to A. Indeed, most of the developments in decision theory were motivated by the need to account for the possibility of hedging relationships among strategies, that are not allowed for under the standard expected utility framework. The solutions proposed are typically the restriction of the independence axiom to a subset of the available strategies. Among them, an interesting choice is to restrict C to the set of so-called *constant acts*, which in our application means a strategy yielding a risk-free return. This was proposed by Gilboa and Schmeidler (1989) and leads to a decision strategy called *Maximin Expected Utility* (MMEU). In intuitive terms, this strategy can be described as one taking into account the worst possible probability scenario for every possible event. As we are going to see in the following paragraph, this worst probability scenario corresponds to what in the mathematics of incomplete market pricing are called *super-replication* or *super-hedging* strategies.

1.6.2 Super-hedging strategies

Here we follow Cherubini (1997) and Cherubini and Della Lunga (2001) in order to provide a general formal representation of the incomplete market problem, i.e. the problem of pricing a contingent claim on an asset that cannot be exactly replicated. In this setting, a general contingent claim $g(S,t)$ with pay-off $G(S,T)$, can be priced computing

$$g(S,t) = \exp[-r(T-t)]\, E_Q\left[G(S,T);\, Q \in \wp \mid \Im_t\right]$$

where E_Q represents the expectation with respect to a conditional risk-neutral measure Q. Here and in the following we focus on the financial meaning of the issue and assume that the technical conditions required to ensure that the problem is well-defined are met (the readers are referred to Delbaen & Schachermayer, 1994, for details). The set $\wp$ contains the risk-neutral measures and describes the information available on the underlying asset. If it is very precise, and the set $\wp$ contains a single probability measure, we are in the standard complete market pricing setting tackled above. In the case in which we do not have precise information – for example, because of limited liquidity of the underlying – we have the problem of choosing a single probability measure, or a pricing strategy. Therefore, in order to price the contingent claim g in this incomplete market setting, we have to define: (i) the set of probability measures $\wp$ and (ii) a set of rules describing a strategy to select the appropriate measure and price. As discussed above, one could resort to expected utility to give a preference rank for the probabilities in the set, picking out the optimal one. As an alternative, or prior to that, one could instead rely on some more conservative strategy, selecting a range of prices: the bounds of this range would yield the highest and lowest price consistent with the no-arbitrage assumption, and the replicating strategies corresponding to these bounds are known as *super-replicating* portfolios. In this case we have

$$g^-(S,t) = \exp[-r(T-t)] \inf E_Q\left[G(S,T);\, Q \in \wp \mid \Im_t\right]$$
$$g^+(S,t) = \exp[-r(T-t)] \sup E_Q\left[G(S,T);\, Q \in \wp \mid \Im_t\right]$$

More explicitly, the lower bound is called the *buyer price* of the derivative contract g, while the upper bound is denoted the *seller price*. The idea is that if the price were lower than the buyer price, one could buy the contingent claim and go short a replicating portfolio ending up with an arbitrage gain. Conversely, if the price were higher than the maximum, one could short the asset and buy a replicating portfolio earning a safe return. Depending on the definition of the set of probability measures, one is then allowed to recover different values for long and short positions. Notice that this does not hold for models that address the incomplete market pricing problem in a standard expected utility setting, in which the selected measure yields the same value for long and short positions.

Uncertain probability model

The most radical way to address the problem of super-replication is to take the worst possible probability scenario for every event. To take the simplest case, that of a call digital option paying one unit of currency at time T if the underlying asset is greater than or equal to K, we have

$$\begin{aligned}
DC^{-}(S,t) &= \exp[-r(T-t)] \inf E_Q\left[\mathbf{1}_{S(T)\geqslant K};\ Q \in \wp \mid \Im_t\right] \\
&= \exp[-r(T-t)] \inf\left[\overline{Q}(K);\ Q \in \wp \mid \Im_t\right] \equiv B(t,T)\,\overline{Q}^{-} \\
DC^{+}(S,t) &= \exp[-r(T-t)] \sup E_Q\left[\mathbf{1}_{S(T)\geqslant K};\ Q \in \wp \mid \Im_t\right] \\
&= \exp[-r(T-t)] \sup E_Q\left[\overline{Q}(K);\ Q \in \wp \mid \Im_t\right] \equiv B(t,T)\,\overline{Q}^{+}
\end{aligned}$$

where we recall the definition $\overline{Q}(K) \equiv 1 - Q(K)$ and where the subscripts '+' and '−' stand for the upper and lower value of $\overline{Q}(K)$.

Having defined the pricing bounds for the digital option, which represents the pricing kernel of any contingent claim written on asset S, we may proceed to obtain pricing bounds for call and put options using the integral representations recovered in section 1.5.3. Remember in fact that the price of a European call option C under the martingale measure Q may be written in very general terms as

$$\text{CALL}(S,t;K,T) = \exp[-r(T-t)] \int_K^{\infty} \overline{Q}(u \mid \Im_t)\,\mathrm{d}u$$

We know that if the kernel were the log-normal distribution, the equation would yield the Black–Scholes formula. Here we want instead to use the formula to recover the pricing bounds for the option. The buyer price is then obtained by solving the problem

$$\text{CALL}^{-}(S,t;K,T) = \exp[-r(T-t)] \int_K^{\infty} \overline{Q}^{-}(u)\,\mathrm{d}u$$

By the same token, the seller price is obtained from

$$\text{CALL}^{+}(S,t;K,T) = \exp[-r(T-t)] \int_K^{\infty} \overline{Q}^{+}(u)\,\mathrm{d}u$$

and represents the corresponding upper bound for the value of the call option in the most general setting.

The same could be done for the European put option with the same strike and maturity. In this case we would have

$$\text{PUT}(S,t;K,T) = \exp[-r(T-t)] \int_0^{K} Q(u \mid \Im_t)\,\mathrm{d}u$$

for any conditional measure $Q \in \wp$ and the pricing bounds would be

$$\mathrm{PUT}^{-}(S,t;K,T) = \exp\left[-r(T-t)\right]\int_0^K Q^{-}(u \mid \Im_t)\,\mathrm{d}u$$

$$\mathrm{PUT}^{+}(S,t;K,T) = \exp\left[-r(T-t)\right]\int_0^K Q^{+}(u \mid \Im_t)\,\mathrm{d}u$$

where $Q^{-}(u)$ and $Q^{+}(u)$ have the obvious meanings of the lower and upper bound of the probability distribution for every u. Notice that whatever pricing kernel, Q in the $\wp$ set has to be a probability measure, so it follows that $Q(u) + \overline{Q}(u) = 1$. This implies that we must have

$$Q^{-}(u) + \overline{Q}^{+}(u) = 1$$

$$Q^{+}(u) + \overline{Q}^{-}(u) = 1$$

In the case of incomplete markets, in which the set $\wp$ is not a singleton, we have $Q^{-}(u) < Q^{+}(u)$, which implies

$$Q^{-}(u) + \overline{Q}^{-}(u) = Q^{-}(u) + \left[1 - Q^{+}(u)\right] < 1$$

and the measure Q^{-} is sub-additive. In the same way, it is straightforward to check that

$$Q^{+}(u) + \overline{Q}^{+}(u) > 1$$

and the measure Q^{+} is super-additive.

So, if we describe the probability set as above, the result is that the buyer and seller prices are integrals with respect to non-additive measures, technically known as *capacities*. The integrals defined above are well defined even for non-additive measures, in which case they are known in the literature as *Choquet integrals*. This integral is in fact widely used in the modern decision theory trying to amend the standard expected utility framework: lotteries are ranked using capacities instead of probability measures and expected values are defined in terms of Choquet integrals rather than Lebesgue integrals, as is usual in the standard expected utility framework.

Example 1.2 [*Fuzzy measure model*] A particular parametric form of the approach above was proposed by Cherubini (1997) and Cherubini and Della Lunga (2001). The idea is drawn from fuzzy measure theory: the parametric form suggested is called Sugeno fuzzy measure. Given a probability distribution Q and a parameter $\lambda \in \Re_{+}$, define

$$\overline{Q}^{-}(u) = \frac{1 - Q(u)}{1 + \lambda Q(u)} \qquad \overline{Q}^{+}(u) = \frac{1 - Q(u)}{1 + \lambda^{*} Q(u)}$$

with

$$\lambda^{*} = -\frac{\lambda}{1+\lambda}$$

It may be easily checked that the measure $\overline{Q}^{-}$ is sub-additive, and $\overline{Q}^{+}$ is the dual super-additive measure in the sense described above.

The pricing bounds for call options are then recovered as discussed above based on any choice of the reference probability distribution Q. If the pricing kernel is chosen to be log-normal, we obtain

$$\mathrm{CALL}^{-}(S, t; K, T) = \exp\left[-r(T-t)\right] \int_{d_2}^{\infty} \frac{\Phi(u)}{1+\lambda\Phi(-u)} \mathrm{d}u$$

$$\mathrm{CALL}^{+}(S, t; K, T) = \exp\left[-r(T-t)\right] \int_{d_2}^{\infty} \frac{\Phi(u)}{1+\lambda^{*}\Phi(-u)} \mathrm{d}u$$

Notice that in the case $\lambda = \lambda^{*} = 0$ the model yields the Black–Scholes formula. For any value $\lambda > 0$, the model yields buyer and seller prices. The discount (premium) applied to buyer (seller) prices is higher the more the option is *out-of-the-money*.

Uncertain volatility model

An alternative strategy to address the incomplete market problem would be to define a set of risk-neutral dynamics of the underlying asset, rather than the set of risk-neutral measures. A typical example is to assume that the volatility parameter is not known exactly, and is considered to be included in a given interval. Assume further that the stochastic process followed by the underlying asset is a geometric brownian motion. Under any risk-neutral measure Q we have

$$\mathrm{d}S(t) = rS(t)\,\mathrm{d}t + \sigma S(t)\,\mathrm{d}\widetilde{z}$$

and we assume that $\sigma \in \left[\sigma^{-}, \sigma^{+}\right]$. This model is called the *uncertain volatility model* (UVM) and is due to Avellaneda, Levy and Paràs (1995) and Avellaneda and Paràs (1996).

Working through the solution as in the standard Black–Scholes framework, assume to build a dynamic hedged portfolio. Notice that if we knew the exact value of the σ parameter, the delta hedging strategy could be designed precisely, enabling perfect replication of the contingent claim. Unfortunately, we are only allowed to know the interval in which the true volatility value is likely to be located, and we are not aware of any probability distribution about it. Assume that we take a conservative strategy designing the hedging policy under the worst possible volatility scenario. Avellaneda, Levy and Paràs (1995) show that this leads to the pricing formula

$$\frac{\partial g}{\partial t} + \frac{1}{2}\sigma^{2} S^{2}(t) \left[\frac{\partial^{2} g}{\partial S^{2}}\right]^{+} + rS\frac{\partial g}{\partial S} - rg = 0$$

with

$$\left[\frac{\partial^{2} g}{\partial S^{2}}\right]^{+} = \begin{Bmatrix} \sigma^{-} & (\partial^{2} g/\partial S^{2}) > 0 \\ \sigma^{+} & (\partial^{2} g/\partial S^{2}) < 0 \end{Bmatrix}$$

Notice that the partial differential equation is a modified non-linear version of the Black–Scholes no-arbitrage condition. The non-linearity is given by the fact that the multiplicative term of the second partial derivative is a function of the sign of the second partial derivative. This equation was denoted the BSB (Black, Scholes & Barenblatt) fundamental equation. The solution has to be carried out numerically except in trivial cases in which it may be proved that the solution is globally convex or concave, when it obviously delivers the same results as the standard Black–Scholes model. Notice also that in this approach, as in the previous uncertain probability model, the result yields different values for long and short positions.

1.7 CREDIT RISK

The recent developments of the market have brought about a large increase of credit risk exposures and products. On the one hand, this has been due to the massive shift of the investment practices from standard stocks and bonds products toward the so-called *alternative investments*. This shift has been motivated both by the quest for portfolio diversification and the research of higher returns in a low interest rate period. Moreover, another face of credit risk has become increasingly relevant along with the progressive shift from the classical standard intermediation business toward structured finance products, and the need to resort to *over-the-counter* (OTC) transactions to hedge the corresponding exposure. Contrary to what happens in derivatives transactions operated in *futures-style* organized markets, OTC deals involve some credit risk, as the counterparty in the contract may default by the time it has to honor its obligations. The credit-risk feature involved in derivative contracts is known as *counterparty risk*, and has been getting all the more relevant in the risk management debate.

A very general way to represent the pay-off of a *defaultable* contingent claim – that is, a contract in which the counterparty may go bankrupt – is

$$G\left(S,T\right)\left[1-\mathbf{1}_{\{\mathrm{DEF}\}}(T)\mathrm{LGD}\right]$$

where $\mathbf{1}_{\{\mathrm{DEF}\}}$ is the indicator function denoting the default of the counterparty by time T and LGD is the *loss given default* figure, also defined as $\mathrm{LGD} \equiv 1 - \mathrm{RR}$, that is one minus the *recovery rate*. In very general terms, the value of the contract at time t is computed under the risk-neutral measure as

$$E_Q\left[\exp\left[-\int_t^T r\left(u\right)\,\mathrm{d}u\right]G\left(S,T\right)\left[1-\mathbf{1}_{\{\mathrm{DEF}\}}\mathrm{LGD}\right]\right]$$

Notice that there are three risk factors involved in this representation: (i) market risk due to fluctuations of the underlying asset S; (ii) interest rate risk due to changes in the discount factor; and (iii) credit risk due to the event of default of the counterparty. We will see that evaluating defaultable contingent claims in this framework crucially involves the evaluation of the dependence structure among the sources of risk involved. Fortunately, we know that one of the sources may be made orthogonal by the change in measure corresponding to the bond numeraire (*forward martingale measure*). In this case we have

$$D\left(t,T\right)E_{Q_T}\left[G\left(S,T\right)\left[1-\mathbf{1}_{\{\mathrm{DEF}\}}\mathrm{LGD}\right]\right]$$

and the credit risk problem is intrinsically bivariate, involving the dependence structure between the underlying dynamics and default of the counterparty.

The standard credit risk problem that we are used to think of is only the simplest case in this general representation. Setting in fact $G(S,T) = 1$ we have the standard defaultable bond pricing problem. In the discussion below, we will first address this topic, before extending it to the case in which the defaultable security is a derivative contract. Dealing with the simplest case will enable us to stress that credit risk itself is similar to an exposure in the derivative market. Curiously enough, it can be seen as a position in an option, following the so-called structural approach, or as a position with the same features as an interest rate derivative, according the so-called reduced form approach.

1.7.1 Structural models

Structural models draw the main idea from the pioneering paper by Merton (1974). Assume that an entrepreneur is funding a project whose value is $V(t)$ with debt issued in the form of a zero coupon bond with a face value of $\overline{DD}$. The debt is reimbursed at time T. If at that date the value of the asset side of the firm is high enough to cover the value of debt, the nominal value is repaid and equityholders get the remaining value. If instead the value of the firm is not sufficient to repay the debt, it is assumed that the debtholders take over the firm at no cost, and stockholders get zero (a feature called *limited liability*). The pay-off value of debt at maturity is then $\min(\overline{DD}, V(T))$, while the value of equity is what is left after bondholders have been repaid (we say stockholders are *residual claimants*).

$$C(T) = \max(V(T) - \overline{DD}, 0)$$

The value of equity capital is then the value of a call option written on the asset value of the firm for a strike equal to the face value of the debt. Notice that the value of debt at the same date can be decomposed alternatively as

$$DD(T) = V(T) - \max(V(T) - \overline{DD}, 0)$$

or

$$DD(T) = \overline{DD} - \max(\overline{DD} - V(T), 0)$$

The latter representation is particularly instructive. The value of defaultable debt is the same as that of default-free debt plus a short position in a put option written on the asset value of the firm for a strike equal to the face value of debt. Notice that if put–call parity holds we have

$$V(T) = \max(V(T) - \overline{DD}, 0) + \overline{DD} - \max(\overline{DD} - V(T), 0)$$

and the value of the firm is equal to the value of equity, the call option, plus the value of debt, in turn decomposed into default-free debt minus a put option. This result is known in the corporate finance literature as the *Modigliani–Miller* theorem. Let us remark that it is not a simple accounting identity, but rather a separation result: it means that the value of the asset side of a firm is invariant under different funding policies; to put it another way, following an increase in the amount of nominal debt its value increases exactly by the same amount as the decrease in the value of equity. It is well known that this is only true under very restrictive assumptions, such as the absence of taxes and bankruptcy costs, or agency costs. Accounting for all of these effects would imply a break-up of the relationship above, and the choice of the amount of debt would have a feedback effect on the output of the firm.

Apart from such possible complications, it is clear that option theory could be applied to recover both the value of debt and equity, and to decompose debt into the default-free part and the credit risk premium.

Assume that the asset side of the firm $V(t)$ follows a geometric brownian motion, so that under the risk-neutral measure we have

$$\mathrm{d}V(t) = rV(t)\,\mathrm{d}t + \sigma_V V(t)\,\mathrm{d}\widetilde{z}(t)$$

Then, the standard Black–Scholes formula can be applied to yield the value of equity C

$$C(t) = V(t)\,\Phi(d_1) - \exp(-r(T-t))\,\overline{DD}\Phi(d_2)$$
$$d_1 = \frac{\ln(V(t)/\overline{DD}) + (r + \sigma_V^2/2)(T-t)}{\sigma_V\sqrt{T-t}}$$
$$d_2 = d_1 - \sigma_V\sqrt{T-t}$$

and the value of debt DD is recovered as

$$\begin{aligned} DD(t) &= V(t) - [V(t)\Phi(d_1) - \exp(-r(T-t))\overline{DD}\Phi(d_2)] \\ &= \Phi(-d_1)V(t) + \exp(-r(T-t))\overline{DD}\Phi(d_2) \end{aligned}$$

Notice that, by adding and subtracting $\exp(-r(T-t))\,\overline{DD}$ we can rewrite the value as

$$DD(t) = \exp(-r(T-t))\,\overline{DD} - \left[-V(t)\,\Phi(-d_1) + \exp(-r(T-t))\,\overline{DD}\Phi(-d_2)\right]$$

and we recognize the short position in the put option representing credit risk.

The result could be rewritten by defining the underlying asset of the option in percentage terms, rather than in money amounts. For this reason, we introduce

$$d = \frac{\exp(-r(T-t))\,\overline{DD}}{V(t)}$$

which is called by Merton *quasi-debt-to-firm-value ratio* or *quasi-leverage*. The *quasi* term is motivated by the fact that the debt is discounted using the risk-free rate rather than the defaultable discount factor. We have

$$DD(t) = \exp(-r(T-t))\,\overline{DD}\left\{1 - \left[-\frac{1}{d}\Phi(-d_1) + \Phi(-d_2)\right]\right\}$$

$$d_1 = \frac{\ln(1/d) + \sigma_V^2/2\,(T-t)}{\sigma_V\sqrt{T-t}}$$
$$d_2 = d_1 - \sigma_V\sqrt{T-t}$$

Remembering that the probability of exercise of a put option is equal to $\Phi(-d_2)$, a modern way to rewrite the formula above would be

$$DD(t) = \exp(-r(T-t))\,\overline{DD}\left\{1 - \Phi(-d_2)\left[1 - \frac{1}{d}\frac{\Phi(-d_1)}{\Phi(-d_2)}\right]\right\}$$
$$= \exp(-r(T-t))\,\overline{DD}\{1 - \mathrm{Dp} * \mathrm{LGD}\}$$

where Dp stands for default probability and LGD is the *loss given default* figure in this model.

$$\mathrm{Dp} = \Phi(-d_2)$$
$$\mathrm{LGD} = 1 - \frac{1}{d}\frac{\Phi(-d_1)}{\Phi(-d_2)}$$

Notice that both the default probability and the loss given default are dependent on the quasi leverage d.

Finally, in order to account for different maturities, credit risk can be represented in terms of credit spreads as

$$r^*(t,T) - r = -\frac{\ln\left\{1 - \Phi(-d_2)\left[1 - \frac{1}{d}\frac{\Phi(-d_1)}{\Phi(-d_2)}\right]\right\}}{T-t}$$
$$= -\frac{\ln\{1 - \mathrm{Dp} * \mathrm{LGD}\}}{T-t}$$

where $r^*(t,T)$ is the yield to maturity of the defaultable bond.

While the original model is based on very restrictive assumptions, some extensions have been proposed to make it more realistic. In particular, the extension to defaultable coupon bond debt was handled in Geske (1977), while the possibility of default events prior to maturity as well as the effects of debt seniority structures was tackled in Black and Cox (1976). Finally, the effects of bankruptcy costs, strategic debt servicing behavior and absolute priority violations were taken into account in Anderson and Sundaresan (1996) and Madan and Unal (2000).

Structural models represent a particularly elegant approach to defaultable bond evaluation and convey the main idea that credit risk basically amounts to a short position in an option. Unfortunately, the hypothesis that both the recovery rate and default probability depend on the same state variable, i.e. the value of the firm, may represent a serious drawback to the flexibility of the model, overlooking other events that may trigger default. As a result, the credit spreads that are generated by this model consistently with reasonable values of asset volatility turn out to be much smaller than those actually observed on the market. Furthermore, the fact that the value of the asset is modeled as a diffusive process observed in continuous time gives a typical hump-shaped credit spread curve (in the usual case with $d < 1$) with zero intercept: technically speaking this is due to the fact that default is a *predictable* event with respect to the information set available at any time t. Three different ways have been suggested to solve this problem: the first is to include a jump in the process followed by the value of assets (Zhou, 1996); the second is to assume that the value of the underlying is not observable in continuous time (Duffie & Lando, 2001); the third is to assume that the default barrier is not observed at any time t (the CreditGrades approach followed by Finger et al., 2002).

1.7.2 Reduced form models

A more radical approach to yield a flexible parametric representation for the credit spreads observed in the market is to model default probability and loss given default separately. By contrast with structural models, this approach is called the reduced form.

Assuming the recovery rate to be exogenously given, the most straightforward idea is to model the default event as a Poisson process. We know that the probability distribution of this process is indexed by a parameter called *intensity* (or hazard rate): for this reason, these models are also called *intensity based*. If γ is the intensity of the Poisson process representing default, the probability that this event will not occur by time T is described by the function

$$\Pr(\tau > T) = \exp\left[-\gamma (T - t)\right]$$

where we assume $\tau > t$, that is, the firm is not in default as of time t. Assume that under the risk-neutral measure Q we have

$$E_Q[1 - \mathbf{1}_{\mathrm{DEF}}] \equiv \Pr(\tau > T) = \exp\left[-\gamma (T - t)\right]$$

and that the default event is independent of interest rate fluctuations. Furthermore, let us assume that the recovery rate RR is equal to zero, so that the whole principal is lost in case of default. Under these assumptions, the price of a defaultable zero-coupon bond maturing at time T is simply

$$\begin{aligned} DD(t, T; \mathrm{RR} = 0) &= D(t, T)\, E_Q\left[1 - \mathbf{1}_{\mathrm{DEF}}\right] \\ &= D(t, T) \exp\left[-\gamma (T - t)\right] \end{aligned}$$

and the credit spread is obtained as

$$r^*(t, T; \mathrm{RR} = 0) - r(t, T) \equiv \left(-\frac{\ln DD(t, T; \mathrm{RR} = 0)}{T - t}\right) - \left(-\frac{\ln D(t, T)}{T - t}\right) = \gamma$$

In this special case the credit spread curve is flat and equal to the intensity figure of the default process.

In the more general case of a positive recovery rate $\mathrm{RR} \equiv 1 - \mathrm{LGD}$, assumed to be non-stochastic, we have instead

$$\begin{aligned} DD(t, T; \mathrm{RR}) &= D(t, T)\, E_Q\left[1 - \mathbf{1}_{\mathrm{DEF}}\mathrm{LGD}\right] \\ &= D(t, T)\, E_Q\left[(1 - \mathbf{1}_{\mathrm{DEF}}) + \mathrm{RR}\mathbf{1}_{\mathrm{DEF}}\right] \\ &= D(t, T)\left\{\mathrm{RR} + (1 - \mathrm{RR})\, E_Q\left[(1 - \mathbf{1}_{\mathrm{DEF}})\right]\right\} \\ &= D(t, T)\,\mathrm{RR} + (1 - \mathrm{RR})\, D(t, T)\, E_Q\left[(1 - \mathbf{1}_{\mathrm{DEF}})\right] \\ &= D(t, T)\,\mathrm{RR} + (1 - \mathrm{RR})\, DD(t, T; \mathrm{RR} = 0) \end{aligned}$$

So, the value of the defaultable bond is recovered as a portfolio of an investment in the default-free bond, and one in a defaultable bond with the same default probability and recovery rate zero.

In terms of credit spreads we have

$$r^*(t, T; \mathrm{RR}) - r(t, T) = -\frac{\ln\{\mathrm{RR} + (1 - \mathrm{RR}) \exp[-\gamma(T - t)]\}}{T - t}$$

Notice that in this case the term structure of the credit spreads is not flat, even though the intensity is still assumed constant.

A natural extension of the model is to assume the intensity to be stochastic. In this case, the default event is said to follow what is called a *Cox process*. The survival probability of the obligor beyond time T is determined as

$$E_Q\left[1 - \mathbf{1}_{\mathrm{DEF}}\right] \equiv \Pr(\tau > T) = E_Q\left[\exp\left[-\int_t^T \gamma(u)\, \mathrm{d}u\right]\right]$$

It is easy to see that, from a mathematical point of view, the framework is much the same as that of interest rate models. These techniques can then be directly applied to the evaluation of credit spreads.

Affine intensity

As an example, assume that the instantaneous intensity $\gamma(t)$ follows a diffusive process dynamics under the risk neutral measure Q

$$\mathrm{d}\gamma(t) = k(\overline{\gamma} - \gamma(t))\, \mathrm{d}t + \sigma\gamma^\alpha\, \mathrm{d}\widetilde{w}$$

For $\alpha = 0,1$ we know that the model is affine and we know that the solution to

$$E_Q\left[1 - \mathbf{1}_{\mathrm{DEF}}\right] = E_Q\left[\exp\left[-\int_t^T \gamma(u)\, \mathrm{d}u\right]\right]$$

is

$$E_Q\left[\exp\left[-\int_t^T \gamma(u)\, \mathrm{d}u\right]\right] = \exp\left[A(T - t) + M(T - t)\gamma(t)\right]$$

The value of a defaultable discount bond is then

$$\begin{aligned} DD(t, T; \mathrm{RR}) &= D(t, T)\mathrm{RR} + (1 - \mathrm{RR})DD(t, T; \mathrm{RR} = 0) \\ &= D(t, T)\{\mathrm{RR} + (1 - \mathrm{RR}) \exp[A(T - t) + M(T - t)\gamma(t)]\} \end{aligned}$$

Notice that using the framework of the forward martingale measure we can easily extend the analysis to the case of correlation between interest rate and credit risk. In fact, we leave the reader to check that the dynamics of the default intensity under such measure, which we denoted Q_T, is

$$\mathrm{d}\gamma(t) = \left[k(\overline{\gamma} - \gamma(t)) - \sigma_T\sigma\gamma^\alpha\right] \mathrm{d}t + \sigma\gamma^\alpha \mathrm{d}w^*$$

where we recall that σ_T is the instantaneous volatility of the default free zero-coupon bond with maturity T. Using the dynamics above one can compute or simulate the price from

$$DD(t, T; \mathrm{RR}) = D(t, T)\,\mathrm{RR} + (1 - \mathrm{RR})\, D(t, T)\, E_{Q_T}\left[\exp\left[-\int_t^T \gamma(u)\, \mathrm{d}u\right]\right]$$

A final comment is in order concerning the recovery rate. Extensions of the model refer to a stochastic recovery rate. Of course, the extension is immediate as long as one is willing to assume that the recovery rate is independent of the default intensity and interest rate. In this case the expected value is simply substituted for the deterministic value assumed in the analysis above. Obviously, as the support of the recovery rate is in the unit interval, one has to choose a suitable probability distribution, which typically is the Beta. Accounting for recovery risk, however, has not been investigated in depth.

Finally, consider that the choice of the amount with respect to which the recovery rate is computed may be relevant for the analysis. There are three possible choices. The first is to measure recovery rate with respect to the nominal value of the principal, as supposed in Jarrow and Turnbull (1995) and Hull and White (1995). The second choice is to compute it with respect to the market value of debt right before default, as in Duffie and Singleton (1998). The last one, which is much more common in practice, is to compute it with respect to principal plus accrued interest. Notice that with the last choice, we get the unfortunate result that the value of a coupon bond cannot be decomposed into a stream of defaultable zero-coupon bonds, and the analysis may turn out to be much more involved.

1.7.3 Implied default probabilities

A look at the models above shows that credit risk is evaluated drawing information from different markets, in particular the equity market, for structural models, and the corporate bond market, for reduced form models. Nowadays more information is implied in other markets, such as the credit derivatives markets. A question is how to extract and combine information from all of these markets to determine the *implied* risk-neutral default probability concerning a particular obligor. Here we give a brief account of the different choices available.

Stock markets

A first choice, implicit in structural models, is to draw information from the equity market. Taking the standard Merton model we have that

$$C(t) = V(t)\,\Phi(d_1) - \exp(-r(T - t))\,\overline{DD}\Phi(d_2)$$

where $C(t)$ is the value of equity. As we know, this is a standard application of the Black–Scholes formula, and we are interested in recovering the probability of exercise $\Phi(d_2)$. Let us remark that this probability is referred to the event that the option representing equity ends up in the money, so that the company does not default. Default probability is then $1 - \Phi(d_2) = \Phi(-d_2)$. The main difference with respect to the Black–Scholes framework is that in this case not only the volatility of the underlying asset σ_V, but also its current value $V(t)$, cannot be observed on the market. What we observe instead is the value of equity $C(t)$. Some other piece of information is needed to close the model. A possible solution is to resort to some estimate of the volatility of equity, σ_C, which can be obtained from the historical time series of prices or from the options traded on the stock. From Ito's lemma, we know that volatility of equity must satisfy

$$\sigma_C = \sigma_V \Phi(d_1) \frac{V(t)}{C(t)}$$

This equation, along with the Black–Scholes formula above, constitutes a non-linear system of two equations in two unknowns that can be solved to yield the values of $V(t)$ and σ_C implied by market prices. The default probability is then recovered, under the risk-neutral measure, as

$$\Phi(-d_2) = \Phi\left(-\frac{\ln\left(V(T)/\overline{DD}\right) + \left(r - \sigma_V^2/2\right)(T-t)}{\sigma_V\sqrt{T-t}}\right)$$

The default probability under the objective measure can be recovered by simply substituting the actual drift μ_V of the asset value of the firm. The latter can be estimated either from historical data or by resorting to the no-arbitrage relationship $\mu_V = r + \lambda\sigma_V$, where λ is the market price of risk.

The solution described above is used in the very well known application of structural models employed by KMV, a firm specialized in supplying default probability estimates about many companies across the world, and recently purchased by Moody's. We know that a serious flaw of the Merton approach is that it underestimates the default probability. The key KMV idea is to apply the argument of the default probability function, which they denote *distance to default*

$$-\frac{\ln\left(V(T)/\overline{DD}\right) + \left(\mu_V - \sigma_V^2/2\right)(T-t)}{\sigma_V\sqrt{T-t}}$$

to fit the empirical distribution of actual historical defaults.

Example 1.3 Based on Standard and Poor's statistics for the year 2001, the leverage figures of AA and BBB industrial companies were equal to 26.4% and 41%. Using these figures, an interest rate equal to 4% and a volatility of the asset side equal to 25% for both firms, we compute a risk-neutral default probability over five years equal to 0.69% and 4.71% respectively. Assuming a market price of risk equal to 6%, the corresponding objective probabilities are 0.29% and 2.42% for the AA and the BBB firm.

Corporate bond markets

Reduced form models suggest that the information about default is in the observed prices of corporate bonds. Given the zero-coupon-bond yield curve of debt issues from a single obligor, and given a recovery rate figure RR we know that

$$DD(t,T;\mathrm{RR}) = D(t,T)\,\mathrm{RR} + (1-\mathrm{RR})\,DD(t,T;\mathrm{RR}=0)$$

Furthermore, we know that the value of the zero-coupon with recovery rate zero implied in this price is

$$DD\left(t,T;\mathrm{RR}=0\right)=D\left(t,T\right)\Pr\left(\tau>T\right)$$

where again we assume that interest rate risk and default risk are orthogonal. The implied survival probability is then obtained from defaultable and non-defaultable bond prices as

$$\Pr\left(\tau>T\right)=\frac{DD\left(t,T;\mathrm{RR}\right)/D\left(t,T\right)-\mathrm{RR}}{1-\mathrm{RR}}$$

Alternatively, a common practice in the market is to refer to asset swap spreads as representative of the credit spread of a specific issue. To get the main idea behind this practice, consider a defaultable coupon bond issued at par with coupon equal to r^*. We know that if the bond issued were default-free, it could be swapped at the swap rate SR. We remind the reader that the swap rate is defined as the coupon that equals the value of the fixed leg to that of the floating one in a plain vanilla swap. So, the defaultable cash flow r^* can be swapped against a stream of floating payments plus a spread equal to the difference between the coupon and the swap rate. The advantage of using the asset swap spread is that it conveys information on the riskiness of the individual bond, rather than a whole set of bonds issued by the same entity, while the main flaw is that it may represent other sources of risk, beyond that of default, linked to specific features of the issue, particularly its liquidity. Furthermore, by its very nature it is not well suited to represent the term structure of default risk and credit spreads. Typically, then, the asset swap spread is used to represent a flat credit spread and default intensity curve.

Credit default swap markets

The process of financial innovation that has characterized the recent development credit market has offered new tools to extract market information on the default risk of the main obligors. Credit derivative products, which are used to transfer credit risk among financial institutions, represent a natural source of information concerning the default risk. In particular, credit default swaps represent a very liquid market to extract such information. A credit default swap is a contract in which one counterparty *buys protection* from the other against default of a specific obligor, commonly denoted *name*. The buyer of protection promises periodic payments of a fixed coupon until the end of the contract or default of the *name*. The seller of protection agrees to refund the loss on the assets of a *name* if default occurs, either by buying its obligations at par (*physical settlement*) or by cash refund of the loss on them (*cash settlement*). As in a standard swap, its value at origin is zero.

Assuming, for the sake of simplicity, that no payment is made in case of default for the coupon period in which the credit event occurs, the credit default swap coupon for maturity t_N is defined from

$$\mathrm{LGD}\sum_{i=1}^{N-1}D\left(t,t_i\right)\left(\overline{Q}\left(t_i\right)-\overline{Q}\left(t_{i+1}\right)\right)=c_N\sum_{i=1}^{N-1}D\left(t,t_i\right)\overline{Q}\left(t_{i+1}\right)$$

where c_N are the credit default swap spreads observed on the market, $\overline{Q}\left(t_i\right)$ is the survival probability of the obligor beyond time t_i and the loss given default figure is supposed to be non-stochastic. Notice that the term structure of survival probabilities can be recovered by means of a *bootstrap* algorithm. The credit default swap rates are sorted from short to long maturities. Then, for maturity t_1 we have

$$\overline{Q}(t_1) = \frac{\text{LGD}}{c_1 + \text{LGD}}$$

and for any other maturity t_N, $N \geqslant 2$, one can compute

$$\overline{Q}(t_N) = \frac{c_{N-1} - c_N}{D(t, t_N)(c_N + \text{LGD})} \sum_{i=1}^{N-1} D(t, t_i)\,\overline{Q}(t_i) + \overline{Q}(t_{N-1}) \frac{\text{LGD}}{c_N + \text{LGD}}$$

Alternatively, one can assume that the coupon of the period in which the underlying credit defaults is paid at the end of the period. In this case, the credit default swap is defined as

$$\text{LGD} \sum_{i=1}^{N-1} D(t, t_i)\left(\overline{Q}(t_i) - \overline{Q}(t_{i+1})\right) = c_N \sum_{i=1}^{N-1} D(t, t_i)\,\overline{Q}(t_i)$$

The bootstrap procedure now yields

$$\overline{Q}(t_1) = 1 - \frac{c_1}{\text{LGD}}$$

and

$$\overline{Q}(t_N) = \frac{c_{N-1} - c_N}{D(t, t_N)\,\text{LGD}} \sum_{i=1}^{N-1} D(t, t_i)\,\overline{Q}(t_{i-1}) + \overline{Q}(t_{N-1}) \left(1 - \frac{c_N}{\text{LGD}}\right)$$

1.7.4 Counterparty risk

Credit risk is not only a feature of standard corporate or defaultable bonds. It is also an element that should be taken into account in the evaluation of any contractual exposure to a counterparty. Derivative contracts may generate such credit risk exposures, particularly in transactions on the OTC market, that, as we have noticed above, represent the main development of the derivative industry.

The pay-off of a defaultable, or as termed in the literature, *vulnerable* derivative, is defined as

$$G(S, T)\left[1 - \mathbf{1}_{\text{DEF}}(T)\,\text{LGD}\right]$$

Of course the dependence structure between the pay-off and the default event may be particularly relevant, and will be the object of some of the applications presented in this book. However, even assuming independence of the two risk factors, some important effects of counterparty risk on the evaluation of derivative contracts can be noticed.

The first, obvious, point is that accounting for counterparty risk leads to a discount in the value of the derivative, with respect to its default-free value. Even under independence, the value of the derivative contract is obtained under the risk-neutral valuation measure as

$$D(t, T)\,E_Q\left[G(S, T)\right] - D(t, T)\,E_Q\left[G(S, T)\right] E_Q\left[\mathbf{1}_{\text{DEF}}(T)\,\text{LGD}\right]$$

that is, the value of the default-free derivative minus the product of such value times the default probability and the loss given default figure.

Both the approaches described above to represent credit risk can be exploited to evaluate the discount to be applied to a derivative contract in order to account for counterparty risk. So, under a structural model one could have

$$D(t,T)\,E_Q[G(S,T)]\left[1-\Phi(-d_2)\left[1-\frac{1}{d}\frac{\Phi(-d_1)}{\Phi(-d_2)}\right]\right]$$

while an intensity based model would yield

$$D(t,T)\,E_Q[G(S,T)]\left[1-(1-\exp(-\gamma(T-t)))\,\mathrm{LGD}\right]$$

The second point to realize is that even though market and credit risk are orthogonal, they must be handled jointly in practice. If one overlooks counterparty risk in evaluating a vulnerable derivative, one obtains the wrong price and the wrong hedging policy, ending up with an undesired market risk.

The third point to notice is that counterparty risk generally turns linear derivatives into non-linear ones. To make the point clear, consider the simplest example of a linear vulnerable derivative, i.e. forward contract. Assume that counterparty A is long in the contract, and counterparty B is short. The delivery price is the forward price F: we remind the reader that the forward price is the delivery value that equals to zero the value of a forward contract at the origin. Assume now that the two counterparties have default probabilities $Q_A(T)$ and $Q_B(T)$ and zero recovery rates, and that the time of default is independent of the underlying asset of the forward contract. Notice that the default risk of counterparty A is relevant only if the contract ends up in the money for counterparty B, that is, if $S(T) < F$, while default of counterparty B is relevant only if the long counterparty ends up with a gain, that is, if $S(T) > F$. The value of the forward contract is then

$$E_Q\left[(S(T)-F)\,\mathbf{1}_{\{S(T)>F\}}\right]\overline{Q}_B(T)+E_Q\left[(S(T)-F)\,\mathbf{1}_{\{S(T)<F\}}\right]\overline{Q}_A(T)$$

where we remind that $\overline{Q}_A(T)$ and $\overline{Q}_B(T)$ are the survival probabilities beyond time T. Notice that linearity of the product is broken unless $\overline{Q}_A(T)=\overline{Q}_B(T)$. Even in the latter case, the delta of the contract would not be equal to 1, but would rather be equal to the survival probability of the two counterparties.

1.8 COPULA METHODS IN FINANCE: A PRIMER

Up to this point, we have seen that the three main frontier problems in derivative pricing are the departure from normality, emerging from the smile effect, market incompleteness, corresponding to hedging error, and credit risk, linked to the bivariate relationship in OTC transactions. Copula functions may be of great help to address these problems. As we will see, the main advantage of copula functions is that they enable us to tackle the problem of specification of marginal univariate distributions separately from the specification of market comovement and dependence. Technically, we will see in Chapter 3 that the term "dependence" is not rigorously correct, because, strictly speaking, dependence is a concept limited to positive comovement of a set of variables. However, we will stick to the term "dependence" throughout most of this book because it is largely diffused both among practitioners in the financial markets and academics in statistics and finance. We will instead insist again and again on the distinction between the concept of dependence, defined in this broad sense, and the concept of linear correlation, which is used by quantitative analysts in most of the financial institutions in the world. In fact, we will show that the concept of dependence embedded in copula functions is much more general than the standard linear correlation concept, and it is able to capture non-linear relationships among the markets.

1.8.1 Joint probabilities, marginal probabilities and copula functions

To give an intuitive grasp of the use of copula functions in finance, consider a very simple product, a bivariate digital option. This option pays one unit of currency if two stocks or indexes are above or below a pair of strike price levels. Options like these are very often used in structured finance, particularly index-linked products: examples are digital bivariate notes and, more recently, Altiplano notes.

As an example, assume a product written on the Nikkei 225 and S&P 500 indexes which pays, at some exercise date T, one unit if both are lower than some given levels K_{NKY} and K_{SP}. According to the basic pricing principles reviewed in this chapter, the price of this digital put option in a complete market setting is

$$\mathrm{DP} = \exp\left[-r\left(T-t\right)\right] Q\left(K_{\mathrm{NKY}}, K_{\mathrm{SP}}\right)$$

where $Q\left(K_{\mathrm{NKY}}, K_{\mathrm{SP}}\right)$ is the joint risk-neutral probability that both the Japanese and US market indexes are below the corresponding strike prices.

How can we recover a price consistent with market quotes? The first requirement that may come to mind is to ensure that the price is consistent with the market prices for plain vanilla options on each of the two indexes. Say, for example, we can recover, using some of the models or techniques described in this chapter, the risk-neutral probability Q_{NKY} that the Nikkei index at time T will be below the level K_{NKY}. We can do the same with the S&P 500 index, recovering probability Q_{SP}. In financial terms, we are asking what is the forward price of univariate digital options with strikes K_{NKY} and K_{SP}; in statistical terms, what we are estimating from market data are the marginal risk-neutral distributions of the Nikkei and the S&P indexes.

In order to compare the price of our bivariate product with that of the univariate ones, it would be great if we could write the price as

$$\mathrm{DP} = \exp\left[-r\left(T-t\right)\right] Q\left(K_{\mathrm{NKY}}, K_{\mathrm{SP}}\right) = \exp\left[-r\left(T-t\right)\right] C\left(Q_{\mathrm{NKY}}, Q_{\mathrm{SP}}\right)$$

with $C\left(x, y\right)$ a bivariate function.

Without getting involved in heavy mathematics, we can also discover the general requirements that the function $C\left(x, y\right)$ must satisfy in order to be able to represent a joint probability distribution. Beyond the basic requirement that the output of the function must be in the unit interval, as it must represent a probability, three requirements immediately come to mind. The first: if one of the two events has zero probability, the joint probability that both events occur must also be zero. So, if one of the arguments of $C\left(x, y\right)$ is equal to 0 the function must return 0. On the contrary, if one event will occur for sure, the joint probability that both the events will take place corresponds to the probability that the second event will be observed. This leads to the second technical requirement that if one of the arguments $C(x, y)$ is equal to 1 the function must yield the other argument. Finally, it is intuitive to require that if the probabilities of both the events increase, the joint probability should also increase, and for sure it cannot be expected to decrease. Technically, this implies a third requirement for the function $C(x, y)$, that must be increasing in the two arguments (2-increasing is approximately the correct term: you will learn more on this in Chapter 2). We have just described the three requirements that enable us to define $C(x, y)$ as a copula function.

If we go back to our pricing problem, that's where copula functions come in: they enable us to express a joint probability distribution as a function of the marginal ones. So, the bivariate product is priced consistently with information stemming from the univariate ones. Beyond the intuitive discussion provided here, this opportunity rests on a fundamental finding, known as Sklar's theorem. This result states that any joint probability distribution can be written in terms of a copula function taking the marginal distributions as arguments and that, conversely, any copula function taking univariate probability distributions as arguments yields a joint distribution.

1.8.2 Copula functions duality

Consider now a bivariate digital call option. Differently from the digital put option, it pays one unit of currency if both the Nikkei 225 and the S&P 500 indexes are above the strike levels K_{NKY} and K_{SP}. The relevant probability in this case is

$$\mathrm{DC} = \exp\left[-r\left(T-t\right)\right] \overline{Q}\left(K_{\mathrm{NKY}}, K_{\mathrm{SP}}\right)$$

Analogously to the approach above, the copula function method enables us to recover a copula function $\overline{C}(v, z)$ such that

$$\begin{aligned}\mathrm{DC} &= \exp\left[-r\left(T-t\right)\right] \overline{Q}\left(K_{\mathrm{NKY}}, K_{\mathrm{SP}}\right) \\ &= \exp\left[-r\left(T-t\right)\right] \overline{C}[\overline{Q}\left(K_{\mathrm{NKY}}\right), \overline{Q}\left(K_{\mathrm{SP}}\right)]\end{aligned}$$

The new copula function $\overline{C}(v, z)$ is known as *survival copula*. Readers will learn from the mathematical treatment in Chapter 2 that the survival copula is related to the copula function by the relationship

$$\overline{C}[\overline{Q}\left(K_{\mathrm{NKY}}\right), \overline{Q}\left(K_{\mathrm{SP}}\right)] = 1 - Q\left(K_{\mathrm{NKY}}\right) - Q\left(K_{\mathrm{SP}}\right) + C[Q\left(K_{\mathrm{NKY}}\right), Q\left(K_{\mathrm{SP}}\right)]$$

Readers can also check, as will be discussed in detail in Chapter 8, that the relationship above corresponds to a requirement to rule out arbitrage opportunities.

1.8.3 Examples of copula functions

Let us start with the simplest example of a copula function. This obviously corresponds to the simplest hypothesis corresponding to the comovements of the Japanese and the US markets. Assume, to keep things simple, that the two markets are independent. In this case we know from basic statistics that the joint probability corresponds to the product of the marginal probabilities, and we have

$$\mathrm{DP} = \exp\left[-r\left(T-t\right)\right] Q\left(K_{\mathrm{NKY}}, K_{\mathrm{SP}}\right) = \exp\left[-r\left(T-t\right)\right] Q_{\mathrm{NKY}} Q_{\mathrm{SP}}$$

So, $C(x, y) = xy$, also known as the product copula, is the first function we are able to build and use to price our bivariate option.

The next question is what would happen if the two markets were perfectly positively or negatively correlated. The answer to this question requires us to draw from more advanced statistics, referring to the so-called *Fréchet bounds*. The joint probability is constrained within the bounds

$$\max(Q_{\mathrm{NKY}} + Q_{\mathrm{SP}} - 1, 0) \leqslant Q(K_{\mathrm{NKY}}, K_{\mathrm{SP}}) \leqslant \min(Q_{\mathrm{NKY}}, Q_{\mathrm{SP}})$$

Moreover, the upper bound corresponds to the case of perfect positive dependence between the two markets and the lower bound represents perfect negative dependence. We can therefore check the impact of perfect positive dependence on the value of the bivariate product by computing

$$\mathrm{DP} = \exp[-r(T-t)]\, Q(K_{\mathrm{NKY}}, K_{\mathrm{SP}}) = \exp[-r(T-t)] \min(Q_{\mathrm{NKY}}, Q_{\mathrm{SP}})$$

So, $C(x, y) = \min(x, y)$ is another copula function, known as the maximum copula. The minimum copula will instead correspond to the case of perfect negative dependence and to the Fréchet lower bound $C(x, y) = \max(x + y - 1, 0)$ yielding

$$\begin{aligned}\mathrm{DP} &= \exp[-r(T-t)]\, Q(K_{\mathrm{NKY}}, K_{\mathrm{SP}}) \\ &= \exp[-r(T-t)] \max(Q_{\mathrm{NKY}} + Q_{\mathrm{SP}} - 1, 0)\end{aligned}$$

We have then recovered copula functions corresponding to the extreme cases of independence and perfect dependence. Moving one step forward, we could try to build a copula function accounting for imperfect dependence between the two markets. The first idea would be to try a linear combination of the three cases above, obtaining

$$\begin{aligned}C(Q_{\mathrm{NKY}}, Q_{\mathrm{SP}}) = \beta \max(Q_{\mathrm{NKY}} + Q_{\mathrm{SP}} - 1, 0) + (1 - \alpha - \beta)\, Q_{\mathrm{NKY}} Q_{\mathrm{SP}} \\ + \alpha \min(Q_{\mathrm{NKY}}, Q_{\mathrm{SP}})\end{aligned}$$

with $0 \leqslant \alpha, \beta \leqslant 1$ and $\alpha + \beta = 1$. Copula functions obtained in this way define the so-called Fréchet family of copula functions.

Other ways of obtaining copula functions are more involved and less intuitive. For example, consider taking a function $\varphi(.)$ satisfying some technical conditions that will be discussed in more detail throughout the book. If we define

$$C(Q_{\mathrm{NKY}}, Q_{\mathrm{SP}}) = \varphi^{[-1]}[\varphi(Q_{\mathrm{NKY}}) + \varphi(Q_{\mathrm{SP}})]$$

we obtain copula functions. Copulas constructed in this way are called Archimedean copulas and are largely used in actuarial science.

As a final idea, one could try to generalize and make more flexible the standard setting under which most of the results in finance were obtained under the Black–Scholes theory. This corresponds to normal distribution of the returns, which in this case is extended to multivariate normality. From this perspective, a particularly useful result is that the joint standard normal distribution computed in the inverse of the arguments satisfies the requirements of a copula function. We may then price our bivariate claim using

$$\mathrm{DP} = \exp[-r(T-t)]\Phi[\Phi^{-1}(Q_{\mathrm{NKY}}), \Phi^{-1}(Q_{\mathrm{SP}}); \rho]$$

where $\Phi(x, y; \rho)$ is the standard bivariate normal distribution with correlation parameter ρ. This example is particularly useful to highlight the main advantage from the use of copula functions. Notice in fact that in this way we may preserve the dependence structure typical of a multivariate normal distribution by modifying **only** the marginal distributions, which may be allowed to display skewness and fat-tails behavior consistently with the data observed from the market.

1.8.4 Copula functions and market comovements

As we have already seen from the examples, copula functions provide a way to represent the dependence structure between markets and risk factors, while preserving the specification of the marginal distribution of each and every one of them. Representing market comovements in a world in which the marginal distribution of returns is not normal raises problems that may be new for many scholars and practitioners in finance.

The main result is that linear correlation, which represents the standard tool used in the dealing rooms and risk management units to measure the comovement of markets may turn out to be a flawed instrument in the presence of a non-normal return. Linear correlation between the rate of returns r_{NKY} and r_{SP} in our two markets may be written as

$$\begin{aligned}\mathrm{corr}\,(r_{\mathrm{NKY}}, r_{\mathrm{SP}}) &= \frac{\mathrm{cov}\,(r_{\mathrm{NKY}}, r_{\mathrm{SP}})}{\sigma_{\mathrm{NKY}}\sigma_{\mathrm{SP}}} \\ &= \frac{1}{\sigma_{\mathrm{NKY}}\sigma_{\mathrm{SP}}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}[Q\,(x, y) - Q_{\mathrm{NKY}}Q_{\mathrm{SP}}]\,\mathrm{d}x\,\mathrm{d}y\end{aligned}$$

where σ_{NKY} and σ_{SP} represent volatilities. Notice that the correlation depends on the marginal distributions of the returns. The maximum value it can achieve can be computed by substituting the upper Fréchet bound in the formula

$$\begin{aligned}&\mathrm{corr}_{\max}\,(r_{\mathrm{NKY}}, r_{\mathrm{SP}}) \\ &\quad = \frac{1}{\sigma_{\mathrm{NKY}}\sigma_{\mathrm{SP}}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}[\min\,(Q_{\mathrm{NKY}}, Q_{\mathrm{SP}}) - Q_{\mathrm{NKY}}Q_{\mathrm{SP}}]\,\mathrm{d}x\,\mathrm{d}y\end{aligned}$$

and the value corresponding to perfect negative correlation is obtained by substituting the lower bound

$$\begin{aligned}&\mathrm{corr}_{\min}\,(r_{\mathrm{NKY}}, r_{\mathrm{SP}}) \\ &\quad = \frac{1}{\sigma_{\mathrm{NKY}}\sigma_{\mathrm{SP}}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}[\max\,(Q_{\mathrm{NKY}} + Q_{\mathrm{SP}} - 1, 0) - Q_{\mathrm{NKY}}Q_{\mathrm{SP}}]\,\mathrm{d}x\,\mathrm{d}y\end{aligned}$$

Of course everyone would expect these formulas to yield $\mathrm{corr}_{\max} = 1$ and $\mathrm{corr}_{\min} = -1$. The news is that this is not true in general. Of course, that is what we are used to expect in a world of normal returns. The result would also hold in the more general case of elliptic distributions, but not for other arbitrary choices. Looking at the problem from a different viewpoint, correlation is an effective way to represent comovements between variables if they are linked by linear relationships, but it may be severely flawed in the presence of non-linear links. Readers may check this in the simple case of a variable z normally distributed and z^2 which is obviously perfectly correlated with the first one, but has a chi-squared distribution.

So, using linear correlation to measure the comovements of markets in the presence of non-linear relationships may be misleading because it may not cover the whole range from -1 to $+1$ even though two markets are moved by the same factor, and so are perfectly dependent.

The alternative offered by statistics to this shortcoming is the use of non-parametric dependence measures, such as Spearman's ρ_S and Kendall's τ. The non-parametric feature of these measures means that they do not depend on the marginal probability distributions. It does not come as a surprise, then, that these measures are directly linked to the copula function. In particular, it may be proved that the following relationships hold

$$\rho_S = 12 \int_0^1 \int_0^1 C(u, v) \, \mathrm{d}u \mathrm{d}v - 3$$

$$\tau = 4 \int_0^1 \int_0^1 C(u, v) \, \mathrm{d}C(u, v) - 1$$

Notice that the specific shape of the marginal probability distributions does not enter these relationships. Furthermore, it may be proved that substituting the maximum and minimum copulas in these equations gives values of 1 and -1 respectively. Differently from the linear correlation measure, then, if the two variables (markets in our case) are perfectly dependent we expect to observe figures equal to 1 for Spearman's ρ_S and Kendall's τ, while a score -1 corresponds to perfect negative dependence.

The relationship between non-parametric dependence measures and copula functions can also be applied to recover a first calibration technique of the copula function itself. In some cases the relationship between these non-parametric statistics and the parameters of the copula function may also be particularly easy. One of the easiest, that we report as an example, is the relationship between the copula functions of the Fréchet family and Spearman's ρ_S. We have in fact

$$\rho_S = \alpha - \beta$$

where the parameters α, β are reported in the definition of the Fréchet family given above.

1.8.5 Tail dependence

The departure from normality in a multivariate system and the need to represent the comovement of markets as closely as possible raises a second dimension of the problem. We know that non-normality at the univariate level is associated with skewness and leptokurtosis phenomena, and what is known as the *fat-tail* problem. In a multivariate setting, the *fat-tail* problem can be referred both to the marginal univariate distributions or to the joint probability of large market movements. This concept is called *tail dependence*. Intuitively, we may conceive markets in which the marginal distributions are endowed with *fat tails*, but extreme market movements are orthogonal, or cases in which the returns on each market are normally distributed, but large market movements are likely to occur together. The use of copula functions enables us to model these two features, *fat tails* and *tail dependence*, separately.

To represent tail dependence we consider the likelihood that one event with probability lower than v occurs in the first variable, given that an event with probability lower than v occurs in the second one. Concretely, we ask which is the probability to observe, for example, a crash with probability lower than $v = 1\%$ in the Nikkei 225 index, given that a crash with probability lower than 1% has occurred in the S&P 500 index. We have

$$\begin{aligned} \lambda(v) &\equiv \Pr(Q_{\mathrm{NKY}} \leqslant v \mid Q_{\mathrm{SP}} \leqslant v) \\ &= \frac{\Pr(Q_{\mathrm{NKY}} \leqslant v, Q_{\mathrm{SP}} \leqslant v)}{\Pr(Q_{\mathrm{SP}} \leqslant v)} \\ &= \frac{C(v, v)}{v} \end{aligned}$$

If we compute this dependence measure far in the lower tail, that is, for very small values of v, we obtain the so-called *tail index*, in particular the *lower* tail index

$$\lambda_{\mathrm{L}} \equiv \lim_{v \to 0^+} \frac{C(v, v)}{v}$$

It may be easily verified that the tail index is zero for the product copula and 1 for the maximum copula. Along the same lines, one can also recover the tail dependence for the *upper* tail index. Analogously, using the duality among copulas described above, we have

$$\begin{aligned} \lambda_{\mathrm{U}} = \lim_{v \to 1^-} \lambda_v &\equiv \lim_{v \to 1^-} \frac{\Pr(\overline{Q}_{\mathrm{NKY}} > v, \overline{Q}_{\mathrm{SP}} > v)}{\Pr(\overline{Q}_{\mathrm{SP}} > v)} \\ &= \lim_{v \to 1^-} \frac{1 - 2v + C(v, v)}{1 - v} \end{aligned}$$

and this represents the probability that price booms may occur at the same time in the US and Japanese markets.

1.8.6 Equity-linked products

Here we give a brief preview of applications to equity-linked products, beyond the simple multivariate digital options seen above. Consider a simple case of a rainbow option, such as, for example, a call option on the minimum between two assets. These derivatives are largely used in structured finance. An example is a class of products, known as Everest notes, whose coupon at the given time T is determined by computing, for example,

$$\mathrm{coupon}(T) = \max\left[\min\left(\frac{S_{\mathrm{NKY}}(T)}{S_{\mathrm{NKY}}(0)}, \frac{S_{\mathrm{SP}}(T)}{S_{\mathrm{SP}}(0)}\right) - 1, 0\right]$$

where S_{NKY} and S_{SP} are the values of the Nikkei 225 and the S&P 500 indexes and time 0 is the initial date of the contract. At any time $0 < t < T$, the value of the coupon will be computed as a call option with strike on the minimum between two assets whose initial value was 1. The strike price is set equal to 1. We will see in Chapter 8 that the price of options like these can be computed using copula functions. Here we just convey the intuition by working the argument backward. Assume that you have a price function for the rainbow option above

$$\begin{aligned}
&\mathrm{CALL}\left[s_{\mathrm{NKY}}(t), s_{\mathrm{SP}}(t); K, T\right] \\
&\quad = \exp\left[-r(T-t)\right] E_Q\left[\max\left(\min\left(s_{\mathrm{NKY}}(T), s_{\mathrm{SP}}(T)\right) - K, 0\right) \mid \Im_t\right]
\end{aligned}$$

where we have simplified the notation defining $s_{\mathrm{NKY}}(t)$ and $s_{\mathrm{SP}}(t)$, the values of the indexes rescaled with respect to their levels at time 0. Of course in our case we also have $K = 1$. Applying what we know about implied risk-neutral probability we have

$$\begin{aligned}
\Pr\left(\min\left(s_{\mathrm{NKY}}(T), s_{\mathrm{SP}}(T)\right) > 1 \mid \Im_t\right) &= \Pr\left(s_{\mathrm{NKY}}(T) > 1, s_{\mathrm{SP}}(T) > 1 \mid \Im_t\right) \\
&= \overline{Q}\left(1, 1 \mid \Im_t\right) \\
&= -\exp\left[r(T-t)\right] \frac{\partial \mathrm{CALL}}{\partial K}
\end{aligned}$$

Using copula functions we obtain

$$\overline{Q}\left(1, 1 \mid \Im_t\right) = \overline{C}\left(\overline{Q}_{\mathrm{NKY}}(1), \overline{Q}_{\mathrm{SP}}(1) \mid \Im_t\right) = -\exp\left[r(T-t)\right] \frac{\partial \mathrm{CALL}}{\partial K}$$

By integrating from the strike $K = 1$ to infinity we have

$$\mathrm{CALL}\left(s_{\mathrm{NKY}}(t), s_{\mathrm{SP}}(t); K, T\right) = \int_1^{\infty} \overline{C}\left(\overline{Q}_{\mathrm{NKY}}(\eta), \overline{Q}_{\mathrm{SP}}(\eta) \mid \Im_t\right) \mathrm{d}\eta$$

and the call option is written in terms of copula functions. Much more about applications and cases like these and techniques by which closed form solutions may also be recovered is reported in Chapter 8.

1.8.7 Credit-linked products

The vast majority of copula function applications have been devoted to credit risk and products whose pay-off depends on the performance of a basket of obligations from several obligors (*names*). In order to illustrate the main choices involved, we describe the application to a standard problem, that is the pricing of a *first-to-default swap*. This product is a credit derivative, just like the credit default swap described above, with the difference that the counterparty offering protection pays a sum, for example a fixed amount, at the first event of default out of a basket of credit exposures.

To see how the pricing problem of a first-to-default derivative leads to the use of copula functions consider a product that pays one unit of currency if at least one out of two credit exposures defaults by time T. It is clear that the risk-neutral probability of paying the protection is equal to that of the complement to the event that a credit exposure goes bankrupt – that is, the case that both names will survive beyond time T. Formally,

$$\mathrm{FTD} = \exp\left[-r\left(T-t\right)\right]\left[1-\Pr\left(\tau_1 > T, \tau_2 > T \mid \Im_t\right)\right]$$

where FTD denotes *first-to-default,* τ_i, $i = 1, 2$ denote the default times of the two names. It is then immediate to write the price in terms of copula functions

$$\mathrm{FTD} = \exp\left[-r\left(T-t\right)\right]\left[1-\overline{C}\left(\overline{Q}_1\left(T\right), \overline{Q}_2\left(T\right) \mid \Im_t\right)\right]$$

Using the duality relationship between a copula function and its survival copula we obtain

$$\mathrm{FTD} = \exp\left[-r\left(T-t\right)\right]\left[Q_1\left(T\right)+Q_2\left(T\right)-C\left(Q_1\left(T\right), Q_2\left(T\right) \mid \Im_t\right)\right]$$

and the value of the product is negatively affected by the dependence between the defaults of the two names. This diversification effect may be appraised computing the case of perfect positive dependence

$$\begin{aligned}\mathrm{FTD}_{\max} &= \exp\left[-r\left(T-t\right)\right]\left[Q_1\left(T\right)+Q_2\left(T\right)-\min\left(Q_1\left(T\right), Q_2\left(T\right) \mid \Im_t\right)\right]\\ &= \exp\left[-r\left(T-t\right)\right]\left[\max\left(Q_1\left(T\right), Q_2\left(T\right) \mid \Im_t\right)\right]\end{aligned}$$

and that corresponding to independence

$$\begin{aligned}\mathrm{FTD}_{\perp} &= \exp\left[-r\left(T-t\right)\right]\left[Q_1\left(T\right)+Q_2\left(T\right)-Q_1\left(T\right)Q_2\left(T\right) \mid \Im_t\right]\\ &= \exp\left[-r\left(T-t\right)\right]\left[Q_1\left(T\right)+\overline{Q}_1\left(T\right)Q_2\left(T\right), \mid \Im_t\right]\end{aligned}$$

So, as the value of the copula function increases with dependence, the value of the first-to-default product decreases.

Of course one could consider reconducting the analysis to the multivariate normal distribution, by using structural models to specify the marginal distributions and the Gaussian copula to represent dependence

$$C\left(Q_1\left(T\right), Q_2\left(T\right) \mid \Im_t\right) = \Phi[\Phi^{-1}\left(Q_1\left(T\right)\right), \Phi^{-1}\left(Q_2\left(T\right)\right); \rho]$$

where $\Phi^{-1}\left(Q_i\left(T\right)\right)$, $i = 1, 2$ denote the inverse of marginal default probabilities consistent with both the leverage figures of the names and volatilities of their assets, while ρ is the correlation between the assets. In this approach, which is used for example in CreditMetrics, the correlation figure is recovered either from equity correlation or by resorting to the analysis of industrial sector comovements.

Example 1.4 Consider a first-to-default option written on a basket of two names, rated AA and BBB. Under the contract, the counterparty selling protection will pay 1 million euros if one of the names defaults over a 5-year period. We saw in a previous example that the leverage figures of AA and BBB industrial companies were equal to 26.4% and 41%. Under the risk-neutral probability measure, assuming the risk-free rate flat at 4% and a volatility of the asset side equal to 25% for both the firms, we obtained 0.69% and 4.71% default probabilities respectively. The maximum value of the first-to-default can be immediately computed as

$$\text{FTD}_{\max} = 1\,000\,000 \exp\left[-0.04(5)\right]\left[0.0471\right] = 38\,562$$

If one assumes independence between the two credit risks, we obtain instead

$$\text{FTD}_{\perp} = 1\,000\,000 \exp\left[-0.04(5)\right]\left[0.0069 + 0.9931(0.0471)\right] = 43\,945$$

Finally, assuming a Gaussian copula with an asset correlation equal to 20%, the value "in fashion" in the market at the time we are writing, we obtain a joint default probability of 0.088636%. The price of the first-to-default swap is then

$$\text{FTD} = 1\,000\,000 \exp\left[-0.04(5)\right]\left[0.0471 + 0.0069 - 0.00088636\right] = 43\,486$$

Besides this case, it is easy to see why copula functions may be particularly useful in this case. If, for example, we choose to model the distribution of the time to default as in reduced form models, rather than the structure of the firm as in structural models, it is clear that the assumption of normality can no longer be preserved. In this case the marginal distributions are obviously non-Gaussian since they are referred to default times and are naturally defined on a non-negative support. Nevertheless, we may conceive applications that may involve features from both structural and reduced form models. For example, the joint default probability may be specified by using a reduced form model for the marginals and the structural model for the dependence structure. We may write

$$C(Q_1(T), Q_2(T) \mid \Im_t) = \Phi[\Phi^{-1}(1 - \exp(-\gamma_1(T-t))), \Phi^{-1}(1 - \exp(-\gamma_2(T-t))); \rho]$$

where γ_i, $i = 1, 2$ denote the default intensities of the two names and now, differently from the fully structural model quoted above, ρ is correlation between the default times. Notice that in this way we may mix information stemming from different sources, such as equity market for the specification of the dependence structure, and corporate bond or credit default bond markets for the marginal distributions. We now give a simple example of this flexibility, but, again, this has to be taken only as an appetizer to invite readers to get into the details of the matter, which will be covered in the rest of the book.

Example 1.5 Consider a 5-year first-to-default option written on a basket of two names, namely Deutsche Telecom and Dresdner Bank. The nominal value is 1 million euros. The information we have is that the default probability of DT, bootstrapped from a credit default swap, is 12.32%. As for Dresdner, we know that the asset swap spread for a 5-year bond is 75 bp. This allows us to compute a default probability of $[1 - \exp(-0.0075(5))] = 3.6806\%$. We assume that the correlation between the default times is 50% and that the copula is Gaussian. So, we first compute $\Phi^{-1}(12.32\%) = -1.15926$. Analogously, we have for Dresdner $\Phi^{-1}(3.6806\%) = -1.788967169$. The joint default probability is computed from

$$\begin{aligned} C\left(Q_1(T), Q_2(T) \mid \Im_t\right) &= \Phi(\Phi^{-1}(3.6806\%), \Phi^{-1}(12.32\%); 50\%) \\ &= \Phi\left(-1.788967169, -1.15926; 50\%\right) = 1.729\% \end{aligned}$$

The price of the first-to-default is then

$$\text{FTD} = 1\,000\,000 \exp\left[-0.04(5)\right]\left[0.03606 + 0.1232 - 0.01729\right] = 116\,240$$

Notice that the value obtained is very close to the case of perfect default dependence, which would obviously cost 123 200, and the basket of names of the first-to-default in this example is definitely undiversified.