Correlation Risk in Finance

Over the last decade, financial markets have witnessed a progressive concentration of focus on correlation dynamics models. New terms such as *correlation trading* and *correlation products* have become the frontier topic of financial innovation. Correlation trading denotes the trading activity aimed at exploiting changes in correlation, or more generally in the dependence structure of assets and risk factors. Correlation products denote financial structures designed with the purpose of exploiting these changes. Likewise, the new term *correlation risk* in risk management is meant to identify the exposure to losses triggered by changes in correlation. Going long or short correlation has become a standard concept for everyone working in dealing rooms and risk management committees. This actually completes a trend that led the market to evolve from taking positions on the direction of prices towards taking exposures to volatility and higher moments of their distribution, and finally speculating and hedging on cross-moments. These trends were also accompanied by the development of new practices to transfer risk from one unit to others. In the aftermative of the recent crisis, these products have been blamed as one of the main causes. It is well beyond the scope of this book to digress on the economics of the crisis. We would only like to point out that the modular approach which has been typical of financial innovation in the structured finance era may turn out extremely useful to ensure the efficient allocation of risks among the agents. While on the one hand the use of these techniques without adequate knowledge may represent a source risk, avoiding them for sure represents a distortion and a source of cost. Of course, accomplishing this requires the use of modular mathematical models to split and transfer risk. This book is devoted to such models, which in the framework of dependence are called *dependence functions* or *copula* functions.

1.1 CORRELATION RISK IN PRICING AND RISK MANAGEMENT

In order to measure the distance between the current practice of markets and standard textbook theory of finance, let us consider the standard static portfolio allocation problem. The aim is to maximize the expected utility of wealth W at some final date T using a set of risky assets, S_i , i = 1, ..., m. Formally, we have

$$\mathbb{E}_{\mathbb{P}}\left[\mathbb{U}\left(R_f + \sum_{i=1}^m w_i(R_i - R_f)\right)\right],\$$

where $R_i = \ln(S_i(T)/S_i(0))$ are the log-returns on the risky assets and R_f is the risk-free rate. The asset allocation problem is completely described by two functions: (i) the utility function $\mathbb{U}(.)$, assumed strictly increasing and concave; (ii) the joint distribution function of the returns \mathbb{P} . While we could argue in depth about both of them, throughout this book the focus will be on the specification of the joint distribution function. In the standard textbook problem, this is actually kept in the background and returns are assumed to be jointly normally distributed, which leads to rewriting the expected utility in terms of a mean–variance problem.

Nowadays, real-world asset management has moved miles away from this textbook problem, mainly for two reasons: first, investments are no longer restricted to linear products, such as stocks and bonds, but involve options and complex derivatives; second, the assumption that the distribution of returns is Gaussian is clearly rejected by the data. As a result, the expected utility problem should take into account three different dimensions of risk: (i) directional movements of the market; (ii) changes in volatility of the assets; (iii) changes in their correlation. More importantly, there is also clear evidence that changes in both volatility and correlation are themselves correlated with swings in the market. Typically, both volatility and correlation increase when the market is heading downward (which is called the *leverage effect*). It is the need to account for these new dimensions of risk that has led to the diffusion of derivative products to hedge against and take exposures to both changes in volatility and changes in correlation. In the same setting, it is easy to recover the other face of the same problem encountered by the pricer. From his point of view, the problem is tackled from the first-order conditions of the investment problem:

$$\mathbb{E}_{\mathbb{P}}\left[\mathbb{U}'\left(R_f + \sum_{i=1}^m w_i(R_i - R_f)\right)(R_i - R_f)\right] = 0 = \mathbb{E}_{\mathbb{Q}}\left[R_i - R_f\right],$$

where the new probability measure \mathbb{Q} is defined after the Radon–Nikodym derivative

$$\frac{\partial \mathbb{Q}}{\partial \mathbb{P}} = \frac{\mathbb{U}'\left(R_f + \sum_{i=1}^m w_i(R_i - R_f)\right)}{\mathbb{E}_{\mathbb{P}}\left[\mathbb{U}'\left(R_f + \sum_{i=1}^m w_i(R_i - R_f)\right)\right]}.$$

Pricers face the problem of evaluating financial products using measure \mathbb{Q} , which is called the risk-neutral measure (because all the risky assets are expected to yield the same return as the risk-free asset), or the *equivalent martingale measure* (EMM, because \mathbb{Q} is a measure equivalent to \mathbb{P} with the property that prices expressed using the risk-free asset as *numeraire* are martingale). An open issue is whether and under what circumstances volatility and correlation of the original measure \mathbb{P}_{arc} preserved under this change of measure. If this is not the case, we say that volatility and correlation risks are priced in the market (that is, a risk premium is required for facing these risks). Under this new measure, the pricers face problems which are similar to those of the asset manager, that is evaluating the sensitivity of financial products to changes in the direction of the market (long/short the asset), volatility (long/short volatility) and correlation (long/short correlation). They face a further problem, though, that is going to be the main motivation of this book: they must ensure that prices of multivariate products are consistent with prices of univariate products. This consistency is part of the so-called arbitrage*free* approach to pricing, which leads to the martingale requirement presented above. In the jargon of statisticians, this consistency leads to the term *compatibility*: the risk-neutral joint distribution \mathbb{Q} has to be compatible with the marginal distributions \mathbb{Q}_i .

Like the asset manager and the pricer, the risk manager also faces an intrinsically multivariate problem. This is the issue of measuring the exposure of the position to different risk factors. In standard practice, he transforms the financial positions in the different assets and markets into a set of exposures (*buckets*, in the jargon) to a set of risk factors (*mapping process*). The problem is then to estimate the joint distribution of losses on these exposures and define a risk measure on this distribution. Typical measures are *Value-at-Risk* (VaR) and *Expected Shortfall* (ES). These measures are multivariate in the sense that they must account for correlation among the losses, but there is a subtle point to be noticed here, which makes this practice obsolete with respect to structured finance products, and correlation products in particular.

3

A first point is that these products are non-linear, so that their value may change even though market prices do not move but their volatilities do. As for volatility, the problem can be handled by including a *bucket* of volatility exposures for every risk factor. But there is a crucial point that gets lost if correlation products are taken into account. It is the fact that the value of these products can change even if neither the market prices nor their volatilities move, but simply because of a change in correlation. In fact, this exposure to correlation among the assets included in the specific product is lost in the mapping procedure. Correlation risk then induces risk managers to measure this dimension of risk on a product-by-product basis, using either historical simulation or stress-testing techniques.

1.2 IMPLIED VS REALIZED CORRELATION

A peculiar feature of applications of probability and statistics to finance is the distinction between historical and implied information. This duality, that is found in many (if not all) applications in univariate analysis, shows up in the multivariate setting as well. On the one side, standard time series data from the market enable us to gauge the relevance of market comovements for investment strategies and risk management issues. On the other side, if there exist derivative prices which are dependent on market correlation, it is possible to recover the degree of co-movement credited by investors and financial intermediaries to the markets, and this is done by simply inverting the prices of these derivatives. Of course, recovering implied information is subject to the same flave as those that are typical of the univariate setting. First, the possibility of neatly backing out this information may be limited by the market incompleteness problem, which has the effect of introducing a source of noise into market prices. Second, the distribution backed out is the risk-neutral one and a market price of risk could be charged to allow for the possibility of correlation changes. These problems are indeed compounded and in a sense magnified in the multivariate setting, in which the uncertainty concerning the dependence structure among the markets adds to that on the shape of marginal distributions.)

Unfortunately, there are not many cases in which correlation can be implied from the market. An important exception is found in the FOREX market, because of the so-called *triangular arbitrage* relationship. Consider the Dollar/Euro ($e_{US,E}$), the Euro/Yen ($e_{E,Y}$) and the Dollar/Yen ($e_{US,Y}$) exchange rates. Triangular arbitrage requires that

$$e_{US,E} = e_{E,Y}e_{US,Y}.$$

Taking logs and denoting by $\sigma_{US,E}$, $\sigma_{E,Y}$, and $\sigma_{US,Y}$ the corresponding implied volatilities, we have that

$$\sigma_{US,E}^2 = \sigma_{E,Y}^2 + \sigma_{US,Y}^2 + 2\rho\sigma_{E,Y}\sigma_{US,Y},$$

from which

$$\rho = \frac{\sigma_{US,E}^2 - \sigma_{E,Y}^2 - \sigma_{US,Y}^2}{2\sigma_{E,Y}\sigma_{US,Y}}$$

is the implied correlation between the Euro/Yen and the Dollar/Yen priced by the market.

1.3 BOTTOM-UP VS TOP-DOWN MODELS

For all the reasons above, estimating correlation, either historical or implied, has become the focus of research in the last decade. More precisely, the focus has been on the specification of the joint distribution of prices and risk factors. This has raised a first strategic choice between two opposite classes of models, that have been denoted *top-down* and *bottom-up* approaches. In all applications, pricing of equity and credit derivatives, risk management aggregation and allocation, the first choice is then to fit all markets and risk factors with a joint distribution and to specify in the process both the marginal distributions of the risk factors and their dependence structure. The alternative is to take care of marginal distributions first, and of the dependence structure in a second step. It is clear that copula functions represent the main tool of the latter approach. It is not difficult to gauge what the pros and cons of the two alternatives might be. Selecting a joint distribution fitting all risks may not be easy, beyond the standard choices of the normal distribution for continuous variables and the Poisson distribution for discrete random variables. If one settles instead for the choice of non-parametric statistics, even for a moderate number of risk factors, the implementation runs into the so-called *curse of dimensionality*. As for the advantages, a top-down model would make it fairly easy to impose restrictions that make prices consistent with the equilibrium or no-arbitrage restrictions. Nevertheless, this may come at the cost of marginal distributions that do not fit those observed in the market. Only seldom (to say never) does this poor fit correspond to arbitrage opportunities, while more often it is merely a symptom of *model risk*. On the opposite side, the bottom-up model may ensure that marginal distributions are properly fitted, but it may be the case that this fit does not abide by the consistency relationships may must exist among prices: the most well known example is the no-arbitrage restriction requiring that prices of assets in speculative markets follow martingale processes. The main goal of this book is actually to show how to impound restrictions like these in a bottom up framework.

1.4 COPULA FUNCTIONS

Copula functions are the main tool for a bottom-up approach. They are actually built on purpose with the goal of pegging a multivariate structure to prescribed marginal distributions. This problem was first addressed and solved by Abe Sklar in 1959. His theorem showed that any joint distribution can be written as a function of marginal distributions:

$$F(x_1, x_2, \ldots, x_n) = C(F_1(x_1), F_2(x_2), \ldots, F_n(x_n))$$

and that the class of functions C(.), denoted copula functions, may be used to extend the class of multivariate distributions well beyond those known and usually applied. To quote the dual approach above, the former result allows us to say that any top-down approach may be written in the formalism of copula functions, while the latter states that copulas can be applied in a bottom-up approach to generate infinitely many distributions. A question is whether this multiplicity may be excessive for financial applications, and this whole book is devoted to that question.

Often a more radical question is raised, whether there is any advantage at all to working with copulas. More explicitly, one could ask what can be done with copulas that cannot be done with other techniques. The answer is again the essence of the bottom-up philosophy. The crucial point is that in the market, we are used to observing marginal distributions. All the information that we can collect is about marginals: the time series of this and that price, and

5

the implied distribution of the underlying asset of an option market for a given exercise date. We can couple time series of prices or of distributions together and study their dependence, but only seldom can we observe multivariate distributions. For this reason, it is mandatory that any model be made consistent with the univariate distributions observed in the market: this is nothing but an instance of that procedure pervasively used in the markets and called *calibration*.

1.5 SPATIAL AND TEMPORAL DEPENDENCE

To summarize the arguments of the previous section, it is of the utmost importance that multivariate models be consistent with univariate observed prices, but this consistency must be subject to some rules and cannot be set without limits. These limits were not considered in standard copula functions applications to finance problems. In these applications the term multivariate was used with the meaning that several different risk factors at a given point in time were responsible for the value of a position at that time. This concept is called *spatial* dependence in statistics and is also known as cross-section dependence in econometrics. Copula functions could be used in full flexibility to represent the consistency between the price of a multivariate product at a given date and the prices of the constituent products observed in the market. However, the term multivariate could be used with a different meaning, that would make things less easy. It could in fact refer to the dependence structure of the value of the same variable observed at different points in time: this is actually defined as a *stochastic* process. In the language of statistics, the dependence among these variables would be called temporal dependence. Curiously, in econometric applications copula functions have mainly been intended in this sense. If copular are used in the same sense in derivative pricing problems, the flexibility of copulas incidential becomes a problem: for example, one would like to impose restrictions on the dynamics to have Markov processes and martingales, and only a proper specification of copulas could be selected to satisfy these requirements.

Even more restrictions would apply in an even more general setting, in which a multivariate process would be considered as a collection of random variables representing the value of each asset or risk factor at different points in time. In the standard practice of econometrics, in which it is often assumed that relationships are linear, this would give rise to the models called *vector autoregression* (VAR). Copula functions allow us to extend these models to a general setting in which relationships are allowed to be non-linear and non-Gaussian, which is the rule rather than the exception of portfolios of derivative products.

1.6 LONG-RANGE DEPENDENCE

These models that extend the traditional VAR time series with a specification in terms of copula functions are called *semi-parametric* and the most well known example is given by the so-called SCOMDY model (Semiparametric Copula-based Multivariate Dynamic). By taking the comparison with linear models one step further, these models raise the question of the behavior of the processes over long time horizons. We know from the theory of time series that a univariate process y_t modeled with the dynamics

$$y_t = \omega + \epsilon_t + \alpha_1 \epsilon_{t-1} + \ldots + \alpha_q \epsilon_{t-q} + \beta_1 y_{t-1} + \ldots + \beta_p y_{t-p},$$

where ω , $\alpha_1, \ldots, \alpha_p$ and β_1, \ldots, β_q are constant parameters, $\epsilon_{t-i} \sim N(0, \sigma)$ and $Cov(\epsilon_{t-i}, \epsilon_{t-j}) = 0, i \neq j$ is called an ARMA(*p*,*q*) model (Autoregressive Moving Average

Process), that could be extended to a multiple set of processes called VARMA. In particular, the MA part of the process is represented by the dependence of y_t on the past q innovations ϵ_{t-i} and the AR part is given by its dependence on the past p values of the process itself. If we focus on the autoregressive part, we know that in cases in which the characteristic equation of the process

$$z - \beta_1 z - \ldots - \beta_p z^p = 0$$

has solutions strictly inside the unit circle, the process is said to be *stationary* in mean. To make the meaning clear, let us just focus on the simplest AR(1) process:

$$y_t = \omega + \epsilon_t + \beta_1 y_{t-1}.$$

It is easy to show that if $\beta_1 < 1$ by recursive substitution of y_{t-i-1} into y_{t-i} , we have

$$\mathbb{E}(y_t) = \omega \sum_{i=0}^{\infty} \beta_1^i + \sum_{i=0}^{\infty} \beta_1^i \mathbb{E}(\epsilon_{t-i}) = \underbrace{\omega}_{1-\beta_1}$$

and

$$VAR(y_t) = \sum_{i=0}^{\infty} \mathbb{E}(\beta_1^i, \beta_i)^2 = \frac{\sigma^2}{1 - \beta_1^2}$$

where we have used the moments of the discribution of ϵ_{t-i} . Notice that if instead it is $\beta_1 = 1$, the dynamics of y_t is defined by

$$y_t = \omega + \epsilon_t + y_{t-1}$$

and neither the mean nor the variance of the unconditional distribution are defined. In this case the process is called *integrated* (of order 1) or *difference stationary*, or we say that the process contains a *unit root*. The idea is that the first difference of the process is stationary (in mean). The distinguishing feature of these processes is that any shock affecting a variable remains in its history forever, a property called *persistence*. As an extension, one can conceive that several processes may be linear combinations of the same persistent shock y_t , that is also called the *common stochastic trend* of the processes. In this case we say that the set of processes constitutes a *co-integrated* system. More formally, a set of processes is said to constitute a co-integrated system if there exists at least one linear combination of the processes that is stationary in mean.

In another stream of literature, another intermediate case has been analyzed, in which the process is said to be *fractionally* integrated, so that the process is made stationary by taking fractional differences: the long-run behavior of these processes is denoted *long memory*. In Chapter 4 we shall give a formal definition of long memory (due to Granger, 2003) and we will discuss the linkage with a copula-based stochastic process. As for the contribution of copulas to these issues, notice that while most of the literature on unit roots and persistence vs stationary models has developed under the maintained assumption of Gaussian innovations, the use of copula functions extends the analysis to non-Gaussian models. Whether these models can represent a new specification for the long-run behavior of time series remains an open issue.

Correlation Risk in Finance

7

1.7 MULTIVARIATE GARCH MODELS

Since copula functions are a general technique to address multivariate non-Gaussian systems, one of the alternative approaches that represents the fiercest challenge to them is the multivariate version of conditional heteroscedasticity models. The idea is to assume that the Gaussian dynamics, or a dynamics that can be handled with tractable distributions, can be maintained if it is modeled conditionally on the dynamics of volatility. So, in the univariate case a shock has the twofold effect of changing the return of a period and the volatility of the following period. Just like in the univariate case, a possible choice to model the non-normal joint distribution of returns is to assume that they could be conditionally normal. A major problem with this approach is that the number of parameters to be estimated can become huge very soon. Furthermore, restrictions are to be imposed to ensure that the covariance matrix is symmetric and positive definite. In particular, imposing the latter restriction is not very easy. The most general model of covariance matrix dynamics is specified by arranging all the coefficients in the matrix in a vector. The most well known specification is, however, that called BEKK (from the authors: Baba, Engle, Kraft, and Kroner, 1990). Calling V_t the eovariance matrix at time t, this specification reads

$$\mathbf{V}_t = \boldsymbol{\Omega}'\boldsymbol{\Omega} + \mathbf{A}'\boldsymbol{\epsilon}_{t-1}'\boldsymbol{\epsilon}_{t-1}\mathbf{A} + \mathbf{B}'\mathbf{V}\mathbf{B},$$

where Ω , **A** and **B** are *n*-dimensional matrices of coefficients. Very often, special restrictions are imposed on the matrices in order to reduce the dimension of the estimation problem. For example, typical restrictions are to assume the matrices **A** and **B** are diagonal, so limiting the flexibility of the representation.

In order to reduce the dimensionality of the problem, the typical recipe used in statistics is to resort to data compression methods: principal component analysis and factor analysis. Both these approaches have been applied to the multivariate GARCH problem. Engle, Ng and Rothschild (1990) resort to a factor GARCH representation in which the dynamics of the common factors are modered as a multivariate factor model. This way, the dimension of the estimation problem is drastically reduced. Alexander (2001) proposes the so-called orthogonal GARCH model. The idea is to use principal component analysis to diagonalize the covariance matrix and to estimate a GARCH model on the diagonalized model. Eigenvectors are then used to reconstruct the variance matrix. The maintained assumption in this model is of course that the same linear transformation diagonalizes not only the unconditional variance matrix, but also the conditional one.

Besides data compression methods, a computationally effective alternative would be to separate the specification of the marginal distributions from the dependence structure. This approach to the specification of the system, that very closely resembles copula functions, was proposed by Engle (2002). The model is known as *Dynamic Conditional Correlation* (DCC). The idea is to separate the specification of a multivariate GARCH model into a two-step procedure in which univariate marginal GARCH processes are specified in the first step, and the dependence structure in the second step. More formally, the approach is based on standardized returns, defined as

$$\eta_{j,t} = \frac{r_{j,t}}{\sigma_{j,t}},$$

which are assumed to have standard normal distribution. Now, notice that the pairwise conditional correlation of return j and k is

$$\rho_{jk,t} = \frac{\mathbb{E}_t[r_{j,t}r_{k,t}]}{\sqrt{\mathbb{E}_t[r_{j,t}^2]\mathbb{E}_t[r_{k,t}^2]}} = \mathbb{E}_t[\eta_{j,t}\eta_{k,t}]$$

Engle (2002) proposes to model such conditional correlations in an autoregressive framework:

$$\chi_{jk,t} = \bar{\rho}_{jk} + \alpha \left(\eta_{j,t} \eta_{k,t} - \bar{\rho}_{jk} \right) + \beta \left(\chi_{jk,t-1} - \bar{\rho}_{jk} \right),$$

where $\overline{\rho}_{jk}$ is the steady-state value of the correlation between returns *j* and *k*.

The estimator proposed for conditional correlation is

$$\hat{\rho}_{jk,t} = \frac{\mathbb{E}_t[\chi_{j,t}\chi_{k,t}]}{\sqrt{\mathbb{E}_t[\chi_{j,t}^2]\mathbb{E}_t[\chi_{k,t}^2]}}.$$

Once written in matrix form, the DCC model can be estimated in a two-stage procedure similar to the IFM (inference functions for margins) method that is typical of copula functions. The model is also well suited to generalization beyond the Gaussian dependence structure, applying copula function specifications as suggested in Fengler *et al.* (2010).

1.8 COPULAS AND CONVOLUTION

In this book, we propose an original approach to the use of copulas in financial applications. Our main task is to identify the limits that must be imposed on the freedom of selecting whatever copula for whatever application. The motivation behind this project is that the feeling of complete flexibility and freedom that from the start was considered to be the main advantage of this tool in the end turned out to be its major limitation. As we are going to see, there is something of a paredox in this: the major flaw in the use of copulas in finance, which in the first place were applied to allow for non-linear relationships, is that it is quite difficult to apply them to linear combinations of variables. We could say that a sort of *curse of linearity* is haunting financial econometrics. To put it in more formal terms, in almost all applications in finance we face the problem of determining the distribution of a sum of random variables. If random variables are assumed to be independent, this distribution is called the *convolution* of the distributions of those variables. In a setting in which the variables are not independent, this convolution concept has to be extended. The extension of this concept to the general case of dependent variables is the innovation at the root of this book: we define and call C-convolution the distribution of the sum of variables linked by a copula function C(.). We also show that once this convolution concept is established, we immediately obtain the copula functions that link the distribution of variables which are part of a sum and the convolution itself. It is our claim that this convolution restriction is the main limitation that must be imposed on the selection of copula functions if one wants to obtain well-founded financial applications. Throughout this book we provide examples of this consistency requirement in all typical fields of copula applications, namely multivariate equity and credit derivatives and risk aggregation and allocation.

As a guide to the reader, the book is made up of three parts. Chapter 2 recalls the main concepts of copula function theory and collects the latest results, limiting the review to the standard way in which copula functions are applied in finance. In Chapters 3 and 4 we construct

the theory and the econometrics of convolution-based copulas. In Chapters 5, 6 and 7 we show that limiting the selection of copulas to members of this class enables us to solve many of the consistency problems arising in financial applications. In Chapter 8 we cast a bridge to future research by addressing the problem of convergence of our copula model, which is developed in a discrete time setting, to continuous time: the issue is open, and for the time being we cannot do any better than review existing copula models that were built in continuous time or that are able to deal with the convolution restriction.

To provide further motivation to the reader, we anticipate here examples in which the convolution restriction arises in applied work, in equity, credit and risk measurement. Consider multivariate equity or FOREX derivatives expiring at different dates. One of the current arguments against the use of copula functions is that they make it difficult to set price consistency between different dates. Indeed, both the efficient market hypothesis and the martingale-based pricing techniques require specific restrictions to the dynamics of market prices. These restrictions, which for short mean that price increments must be unpredictable, imply a restriction in the copula functions that may be used to represent the temporal dependence of the price process: more explicitly, future levels cannot be predicted using the dependence between increment and level. This restriction is exactly of the convolution type: the dependence of the price at different points in time must be consistent with the dependence between a price and its increment in the period, and the latter must be designed to fulfill the unpredictability requirement above. Let us now come to credit derivatives, and consider a typical term structure problem: assume you buy insurance on a basket or credit exposures for five years and sell the same insurance on the same basket for seven years. Could the prices of the two insurance contracts be chosen arbitrarily and independently of each other? The answer is clearly no, because the difference in the two prices is the value today of selling insurance for two years starting five years from now: put in other terms, the distribution of losses in seven years is the convolution of the losses in the first five years and those in the following two years; ignoring this convolution restriction could leave room for arbitrage possibilities. We may conclude with the most obvious risk management application: your firm is made up of a desk that invests in equity products and another one that invests in credit products. For each of the desks you may come up with a risk measure based on the profit and loss distribution. How much capital would you require for the firm as a whole? One more time, the answer to this standard question is convolution.

P1: TIX

9

http://www.pbookshop.com