Application of Stochastic Programming to Management of Cash Flows with FX Exposure

A thesis submitted for the degree of Doctor of Philosophy

By

Konstantin Volosov

School of Information Systems, Computing and Mathematics, Brunel University

April 2006

Dedicated to Belen, Victor and my parents

Acknowledgement

I would like to thank Prof. Gautam Mitra for providing me with the valuable research guidance throughout the years. His support and encouragement made my time here very productive and pleasant. Being a prominent academic and a successful businessman he showed how the two can blend and made my way of thinking more commercially focused. I am also very grateful to Dr. Cormac Lucas for numerous discussions and comments on various topics other than football. I feel indebted to him for sharing his knowledge in mathematical programming and risk modelling. I should not forget Dr. Fabio Spagnolo for assistance and contribution in writing the paper and introducing me to the rather more complex that it seems world of time series and scenario generation.

The last but not least, I would like to thank my dear wife, Belen, for planting the seed of doing PhD in my head in the first place and then her continual support and inspiration in getting me to this point. The actual completion of this thesis I overwhelmingly owe to my wonderful san, Victor, who recently started walking and has shown much interest in pressing the buttons of the keyboard. This fact made me think of submitting before he finds the "delete" button. I also thank my parents Valentina and Nikolay for their support and enduring love over the years.

Abstract

In this thesis we formulate a model for foreign exchange (FX) exposure management and multi-currency cash management taking into consideration random fluctuations of exchange rates and net revenues of a multinational firm (MNF). The central decision model used in this thesis is a scenario-based stochastic programming (SP) recourse model. A critical review of alternative scenario generation methods is given followed by analysis of some desirable properties of the scenario tree. The application of matching statistical moments of a probability distribution to generate a multiperiod scenario tree for our problem is described in detail. A four-stage SP decision model is formulated using the random parameter values. This model evaluates currency / cash flows hedging strategies, which provide rolling decisions on the size and timing of the forward positions. We compute an efficient frontier from which an investor can choose an optimal strategy according to his risk and return preferences. The flexibility of the SP model allows an investor to analyse alternative risk-return trading strategies. The model decisions are investigated by making comparisons with decisions based purely on the expected value problem. The investigation shows that there is a considerable improvement to the "spot only" strategy and provides insight into how these decisions are made.

The contributions of the thesis are summarised below. (i) The FX forward scenario trees are derived using an arbitrage-free pricing strategy and is in line with modern principles of finance. (ii) Use of the SP model and forward contracts as a tool for hedging decisions is novel. (iii) In particular smoothing of the effects in exchange rates and the smoothing of account receivables are examples of innovative modelling approaches for FX management.

Table of Contents

List of Figures

Chapter 1. Introduction and Background

FX markets have gone through a turbulent period since 1973 (after the collapse of Bretton Woods). More recently since 1999 with the emergence of the euro as well as increased globalisation of trade a spectacular amount of currency movement has been recorded. In her recent book Taylor (2003) reports that more than 1.2 trillion US dollars (USD) change hands daily on the foreign exchanges. It is therefore only natural that FX management has become an important topic especially so over the last decade.

The FX participants can be grouped into four categories. *(i)* The first participants are domestic and international banks, which act on their own behalf and for their customers. *(ii)* The second group comprise the Central banks, which may intervene in the market in order to support or suppress the value of the domestic currency for reserve management purposes. *(iii)* The third group is made up of multinational firms (MNFs) who are the customers of banks and buy physical currency in the spot or forward FX market for the purposes of facilitating trade. These MNFs buy and sell foreign currency. *(iv)* The fourth group includes the individual or corporate speculators or traders. In general FX decisions can be seen from two perspectives, such as: (a) hedgers and (b) speculators or traders. In this thesis we use the term trader and speculator interchangeably from now on.

The currency management undertaken by MNFs constitutes only a small fraction (5% – 10%) of total FX transactions. Yet for the purpose of treasury management hedging and limited trading are of vital importance to the corporations and FX decisions can be categorised as shown in Figure 1-1, Taylor (2003). Whereas introducing some element of FX trader (speculator) approach may lead to a better FX decision making there are natural pitfalls for an MNF should it move too far to the right of the scale shown in Figure 1-1. The well-known case of Metallgesellschaft A.G. is one of a few notorious examples of the plight of MNFs who ventured into FX trading activities largely from the position of a speculator. In this thesis we are concerned with the risk exposure of a MNF and treasury risk management requirement in respect of FX exposure.

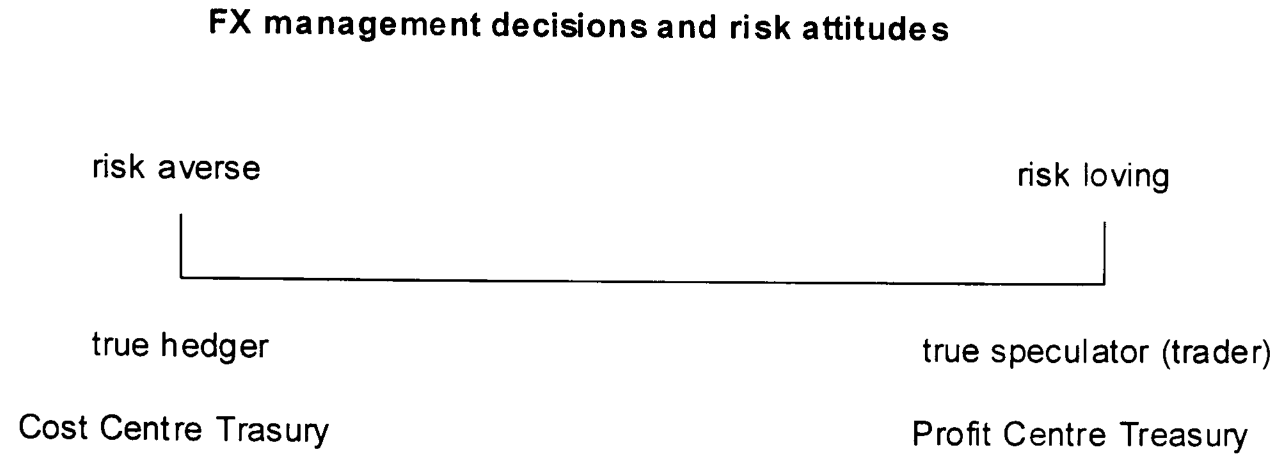

Figure 1-1 FX decisions and risk attitudes

The traditional foreign currency exposure represents a certain (known in advance) volume of foreign currency cash flows that need to be exchanged into the domestic currency at an uncertain future exchange rate. The optimal hedge ratio represents the ratio of the amount of foreign currency cash flow covered by forward contracts (FWD) to the uncovered future foreign currency cash flow, such that this ratio minimises the risk (measured by variance) of the portfolio formed by future cash flows and a position in FWD. The optimal hedge ratio can be calculated by creating a portfolio of two assets: an un-hedged future foreign currency cash flow and a position in a forward currency market. Then it can be shown that the minimum variance portfolio is achieved when the optimal hedge ratio takes the value $[-\mathrm{cov}(s_t, f_t)/\mathrm{var}(f_t)]$, where $s_{t,} f_t$ are the spot and forward exchange rates respectively. Provided the future cash flow stream is known with certainty it is very likely for the value of the optimal hedge ratio to be in the region of 0.9 or higher (Ederington (1979), Kwok (1987), and Swanson and Caples (1987)) for most of the currencies.

Adler & Dumas (1984), Eaker & Grant (1985), and Shapiro (1984) have addressed various implications of uncertain cash flows on hedging decisions. Eaker & Grant study the effect of new information on the optimal hedge, while Shapiro examines the case of multiple hedging tools. Adler & Dumas show that the optimal hedge ratio is the coefficient of a regression of the cash flow (expressed in home currency) on the exchange rate. First the treasury manager specifies a number of future states of nature regarding cash flows, exchange rates, and their respective probabilities. Then the regression coefficient is estimated from a linear regression across the states of nature. Rolfo (1980), Stiglitz (1983), Britto (1984), and Hirshleifer (1988) have examined the problem of hedging uncertain production and hedging in macro-market frameworks.

A more realistic setting, where an MNF has to hedge both uncertain FX exposure and uncertain future foreign currency cash flows simultaneously was investigated by Kerkvliet & Moffett (1991). They show that the optimal hedging decisions will be firm specific and depend on the extent of correlation among the cash flows, spot and futures exchange rates. Maurer & Valiani (2003) contrast effectiveness of hedging currency risk of internationally diversified portfolios using two hedging instruments: currency forwards and European put options. They also analysed hedging performance of in-the-money, at-the-money and out-of-the-money currency options.

FX risk hedging in a static, single-period framework is a straightforward decision problem. The variance-minimising hedge involves taking a position in forward FX market equal in size but opposite in sign to the particular future foreign currency cash flow exposure. It can be shown that this exposure represents the regression coefficient of the cash flow on the exchange rate.

In a multi-period setting optimal hedging is less straightforward. The hedging decision taken at an early stage may be revised many times due to new information being revealed to the market. These frequent revisions may themselves constitute additional risks to the MNF. Dumas (1994) investigates the timing when it is optimal to initialise a hedge. He examines the case of deliberately leaving the cash flows un-hedged for some time, initiating the hedge at some appropriate future time and then leaving the hedge unchanged until the cash flow is received or paid. He states that the appropriate timing of the optimal hedging decision depends on whether the cash flow to be hedged is correlated with the changes in the exchange rates or with its level.

Sharda & Musser (1986) used a multi-objective goal-programming model for bond portfolios. Their approach is to dynamically hedge interest rate risk using futures contracts. In 1993 Sharda & Wingender (1993) reapplied the same model with some modifications to hedging foreign currency accounts receivables using FX futures. Wingender & Sharda (1995) in their later paper modified their original model in several ways. They examined a portfolio of Treasury Notes, incorporation of priorities and the previous week's futures position. The above three studies improve on the static framework by allowing the treasury manager to re-estimate and re-adjust the optimal hedging decisions every time period of the multi-period time horizon. Although these are otherwise comprehensive optimum decision models, the main shortcomings of these studies are that they consider neither stochastic cash flows nor stochastic future exchange rates.

In many real world problems, the uncertainty relating to one or more parameters can be modelled by means of probability distributions. In essence, every uncertain parameter is represented by a random variable over some canonical probability space; this in turn quantifies the uncertainty. Stochastic Programming (SP) enables modellers to incorporate this quantifiable uncertainty into an underlying optimisation model. SP models combine the paradigm of dynamic linear programming with modelling of random parameters, providing optimal decisions which hedge against future uncertainties, see Dominguez-Ballesteros (2001) for a review of SP models and applications. SP has proven very popular in many areas of financial management because it makes it possible to incorporate multiple correlated sources of risk for various assets classes in a common framework, accommodates long-term horizons, provides for risk aversion and allows for dynamic decisions (e.g. portfolio) rebalancing while satisfying regulatory or policy requirements (constraints).

Two-stage and multistage SP frameworks provides a logical extension of the deterministic approach to optimum decision models. SP incorporates uncertain parameters into the model, and the optimal decisions recommended by the model take into account a multi-period time horizon. There have been numerous applications of SP methodology to real life problems over the last two decades. Kusy & Ziemba (1986) formulated a multistage SP to balance a bank's revenues from a set of assets against a set of liabilities. The assets consist of investments and loans with uncertain returns and varying risk levels, whereas the liabilities represent depositor's withdrawals from demand accounts. Klaassen *et al.* (1990) use a multistage SP model to select a minimal cost currency option portfolio to hedge FX exposure faced by an MNF. The portfolio guarantees an acceptable level of USD revenues subject to a certain (known) quantity of a foreign currency to be exchanged in the future. Carino *et al.* (1994) modelled a problem of asset management for a property insurance company as a multistage linear SP model. Golub *et al.* (1995) developed a two-stage SP model for money management using mortgage-backed securities. Beltratti *et al.* (1999) formulated an SP model for portfolio management in the international bond markets. Kouwenberg (2001) developed a multi-stage SP model for pension fund asset liability management using rolling horizon simulations. The use of two-stage SP model to determine the natural oil buying policy of an MNF taking up a forward position is discussed in Poojari *et al.* (2004). Infanger (2006) presents a novel approach to asset allocation based on stochastic dynamic programming and Monte Carlo sampling that permits one to consider many rebalancing periods, many asset classes, dynamic cash flows, and a general representation of investor risk preference.

Perold & Schulman (1988) argue that 100% of FX exposure of international portfolio should be hedged. Eun & Resnik (1994) show that risk-return characteristics of international portfolios where FX risk is hedged by FWD is superior to the unhedged portfolios. Wu & Sen (2000) used SP approach to develop currency option hedging models, which addresses a problem with multiple random factors in imperfect markets.

Until recently the majority of empirical studies considered hedging market risk of investment in internationally diversified portfolios and currency hedging risk as two separate risk management activities. Jorion (1994) shows that strategies when currency hedging polices (hedge ratios) are determined in advance, i.e. before assets allocation, or when portfolio construction and currency risk hedging are conducted sequentially are clearly suboptimal. In Beltratti *et al.* (2004) the authors develop a scenario based optimisation model that simultaneously makes optimal asset allocation and hedging decisions. In their model the hedge ratio can change across currencies and take any value between zero and one. They contrast selective hedging with complete hedging and no-hedging strategies. In Topaloglou *et al.* (2002) an integrated simulation and optimisation framework for multicurrency asset allocation problem is reported, where CVaR is used as a risk metric to account for asymmetric return distribution. The authors examine empirically the benefits of international diversification and the impact of hedging policies on risk-return profiles of portfolios. Topaloglou *et al.* (2004 (a)) apply multistage SP model in portfolio management context, which provides for later portfolio rebalancing decisions. Their model incorporates both market and currency risk hedging decisions into the framework of optimal international portfolio management decisions. FWD are used as currency risk hedging instruments. Their results show that inclusion of hedging instruments improves the performance of international portfolio and provides an efficient and effective way to control risk. Topaloglou *et al.* (2004 (b)) extend the range of available hedging instrument to include simple stock index and quanto options. The results indicate that the performance of the international portfolio is improved as the progressively integrated approach is taken with regards to controlling total risk, i.e. the more risk factors are controlled via inclusion of additional hedging instrument the better the risk-return characteristics of the portfolio.

A number of different hedging instruments are available to the treasury managers (see Abdullah & Wingender (1987)) but in the case study of this thesis we only consider FWD since they are the simplest and one of the most popular hedging products available to MNFs. The specification of the contract can be tailored to the requirements of the customer such as maturity date and size of the contract. Also the forward FX market is very liquid for major currencies and for maturities under two years, which makes it a perfect choice for the problem at hand. In Chapter 5 we illustrate how to formulate and apply a four-stage SP model with recourse to the problem at hand. By using an SP framework one can take into account both time and uncertainty in our ex ante decision model.

The rest of this thesis is organised in the following way. We introduce applications of SP in financial planning in chapter 2. A key aspect of these models are the random parameter values, that represent the uncertainty. Most of these parameters are assumed to follow continuous probability distributions. For discrete planning problems it is necessary to take discrete samples of the uncertain parameters. Chapter 3 provides a detailed description of a number of methods reported in the literature for constructing discrete samples of random variables from a continuous probability distribution. Having captured the discrete samples of the random parameters it is necessary to represent them in a tree structure for exploitation in a SP decision model. In building such a tree of the random parameter values, which represent asset prices (exchange rates in our case), the principle of "no arbitrage" is imposed. Chapter 4 describes different methods of constructing and evaluating these trees. A case study in FX hedging is presented in Chapter 5. The application is developed by first modelling the uncertainty of the relevant parameters after which an arbitrage-free scenario tree is constructed. The SP optimisation model is then formulated and solved. Various investigations into risk-return trade offs are presented and results analysed in this chapter. Finally, we present our conclusions in chapter 6.

Chapter 2. SP as Ex-ante Decision Model

SP problems are mathematical programming models characterised by uncertain future outcomes for some parameters. For decisions made under uncertainty it is a natural extension of the LP model (see Birge & Louveaux (1997) for a comprehensive treatment of the subject)

Consider the deterministic LP problem

$$z = \textit{Minimize} \quad f_0(x) \tag{2.1}$$

$$\textit{Subject to} \quad Ax = b \tag{2.2}$$

$$x \geq 0 \tag{2.3}$$

where $A \in \Re^{m0 \times n0}$ is the matrix of constraint coefficient, $f_0 = cx$, $x, c \in \Re^{n0}$ are the vectors of first stage decision variables and their objective function coefficients respectively, and $b \in \Re^{m0}$ is the vector of available recourses at the first stage. Inequality constraints can be easily converted to the form (2.2) by adding slack or surplus variables. *(2.4)*

Let $(\Omega, \Im, P)$ be the probability space, $\omega \in \Omega$ are realisations of uncertain elements (elementary events) of sample space, $\Im$ is a σ-field and $P(\omega)$ the probability of such elementary events ω, and let $\xi(\omega)$ = $(A,b,c)_\omega$ denote the vector of random model parameters which depends on the realisation of ω, also called a scenario. Let $C^\omega = \{x \mid Ax = b,\ x \geq 0\}$ for $(A,b,c)_\omega$ define the *feasibility set* for a given scenario (realisation) ω.

The two special cases of stochastic problem formulation are anticipative and adaptive models (see Kouwenberg & Zenios (2001) for more details). In an *anticipative* model the decision x must be made in an uncertain environment and is independent of future observations on random parameters. In anticipative models feasibility is expressed via *probabilistic (chance) constraints* and / or objective function. If α, where $0 < \alpha \leq 1$, is the reliability level then the constraints are expressed as follows:

$$P\{\omega \mid f_j(x,\omega) = 0, \quad j = 1,2,\dots,m_1\} \geq \alpha, \qquad (2.5)$$

where x is the n_0-dimensional vector of decision variables and $f_j : \Re^{(n0+n1)} \times \Omega \rightarrow \Re, \quad j = 1,2,\dots,m_1$. In the above formulation the probability of violation of a constraint is limited to a pre-specified level α. The precise value of α could depend on the application and the penalty attached to constrain violation. The objective function can also account for reliability level as follows:

$$P\{\omega \mid f_0(x,\omega) \leq \upsilon\}, \qquad (2.6)$$

where $f_0 : \Re^{(n0+n1)} \times \Omega \rightarrow \Re \cup \{+\infty\}$ and υ is a constant.

Adaptive model assumes partial release of information at the time the decision is taken. At one extreme it coincides with the anticipative model when no information is released and on the other extreme is becomes a deterministic model when all the uncertainty is released before the decision is taken.

Sigma algebra $\Im_t$ on Ω at time t represents all possible events, generated from the sample space (scenario set) Ω of the random vector ω at time t. $\Im_t$ corresponds to all available information at time t. In this case decisions x depend on the information revealed up-to time t, and called $\Im_t$-adapted or $\Im_t$-measurable. Using conditional expectation with respect to $\Im_t$, $E[\cdot \mid \Im_t]$ (see Pliska (1997)), an adaptive stochastic model can be represented as follows:

$$\text{Minimise} \quad E[f_0(x(\omega),\omega) \mid \Im_t] \qquad (2.7)$$

$$\text{Subject to} \quad E[f_j(x(\omega),\omega) \mid \Im_t] = 0, \quad j = 1,2,\dots,m_1; \ x(\omega) \in X. \qquad (2.8)$$

Recourse models combine both anticipative and adaptive models in that decisions are made anticipating uncertainty via scenarios of realisations of random variables and conditioning on (adapting to) the information revealed up to that time period. The decisions taken conditional on the partially revealed information are called *recourse* decisions. Asset allocation in the face of future uncertainty related to anticipation and the portfolio rebalancing at some future date when some information is revealed is related to adaptation.

The two-stage stochastic linear program with recourse makes the dynamic nature of SP explicit, by separating the model's decision variables into two vectors to separate between anticipative policy and adaptive policy.

$x \in \Re^{n0}$ represents the vector of first stage strategic (anticipative) decisions, which are taken before uncertainty is revealed.

$y \in \Re^{n1}$ represents the vector of second stage recourse (adaptive) actions, taken once the uncertainty is revealed.

The formulation of the two-stage SP model with recourse is as follows:

$$Z = \text{Minimise } f_0(x) + \text{E}\left[Q(x,\omega)\right] \tag{2.9}$$

$$\text{subject to} \quad \text{Ax} = \text{b}, \tag{2.10}$$

$$x \in \Re^{n0}, \tag{2.11}$$

where:

$$\text{Q(x,}\omega\text{)} = \quad \text{Minimize } q(y,\omega) \tag{2.12}$$

$$\text{subject to} \quad \text{W(}\omega\text{)y=d(}\omega\text{) - T(}\omega\text{)x}, \tag{2.13}$$

$$y \in \Re^{n1}. \tag{2.14}$$

The matrix A and the vector b are known with certainty. The function $q(y,\omega)$ refers to the second stage cost function, $Q(x,\omega)$ represents the optimal value of the second stage problem for scenario ω. The *recourse matrix* W(ω) of dimension $m_1 \times n_1$. the right-hand side resource m$_1$-vector $d(\omega)$ and the *technology matrix* T(ω) of dimension $m_1 \times n_0$ are random. For a given set of first stage decisions x and a given realisation ω, the corresponding second stage problem seeks to optimise the cost of second stage decisions y. The formulation (2.12)-(2.14) represents the *adaptation* model, equation (2.12) is called a *recourse* function and the second stage decision variables are the *recourse* (corrective) decisions.

In general, a two stage SP problem optimises "Here and Now" (HN) first stage (ex-ante) decisions and the expected cost of second stage decisions.

When ω has a discrete and finite distribution in the scenario set Ω the expected value of the second stage problem can be represented as follows:

$$E[Q(x,\omega)] = \sum_{\omega=1}^{|\Omega|} p(\omega)Q(x,\omega), \tag{2.15}$$

where $p(\omega)$ represents the probability of the scenario $\omega \in \Omega$, $p(\omega) > 0$ and $\sum_{\omega=1}^{|\Omega|} p(\omega) = 1$. For each scenario ω the second stage problem can be depicted as:

$$\text{Minimise} \quad q(y(\omega),\omega) \tag{2.16}$$

$$\text{subject to} \quad W(\omega)y(\omega) = d(\omega) - T(\omega)x, \tag{2.17}$$

where $y(\omega) \in \Re^{n1}$ is the second stage decision taken if the scenario ω is realised. Combining equations (2.15)-(2.17) we can formulate the SP problem in a *deterministic equivalent* form:

$$\text{Minimise} \quad f_0(x) + \sum_{\omega=1}^{|\Omega|} p(\omega)q(y(\omega),\omega) \tag{2.18}$$

$$\text{subject to} \quad \text{Ax} = \text{b}, \tag{2.19}$$

$$T(\omega)x + W(\omega)y(\omega) = h(\omega) \text{ for each } \quad \omega \in \Omega \tag{2.20}$$

$$\text{where} \quad x \in \Re^{n0},\ y(\omega) \in \Re^{n1}. \tag{2.21}$$

When the information is released and corresponding decisions are made at more than one future time period then the two-stage SP program can be extended to a multi-stage program. In this case, the information release over time can be modelled via *filtration*, which represents the nested sequence of sigma algebras: $F = \{\Im_t; \quad t = 0,1,\ldots,T\}$, where T is final time period (see Pliska (1997)). At stage t of the multi-stage SP program the decision is conditional of the information revealed up to time t, characterised by $\Im_t$.

Let ω_t denote the random elementary outcome at time t, such that $\omega_t \in \Omega_t$, where Ω_t is the scenario set at time t. Assume that it is adapted to filtration $F = \{\Im_t; \quad t = 0,1,\ldots,T\}$, which means ω_t is $\Im_t$-measurable for every t = 0, 1,..., T. Now, recourse variables, $y_t \in \Re^{nt}$, cost function, $q_t(y_t, \omega_t)$, and random parameters, $T_t(\omega_t), W_t(\omega_t), d_t(\omega_t)$, have time index. With this notation a *multi-stage* SP problem can be represented as follows:

$$\text{Minimise} \quad f_0(x) + E\left[\underset{y_1 \in \Re^{n1}}{Min}\, q_1(y_1, \omega_1) + \cdots E\left[\underset{y_T \in \Re^{nT}}{Min}\, q_T(y_T, \omega_T)\right]\cdots\right] \tag{2.22}$$

$$\text{Subject to} \quad T_1(\omega_1)x + W_1(\omega_1)y_1 = h_1(\omega_1),$$

$$\vdots \tag{2.23}$$

$$T_T(\omega_T)y_{T-1} + W_{T1}(\omega_T)y_T = h_T(\omega_T)$$

The stochastic properties of the recourse model are characterised by analysing three alternative problems (see Infanger (1994) for more details):

a. The first problem is defined using the expected value over the set Ω. In this approach the stochastic parameters, ξ, are substituted by their expected values, $\bar{\xi}$. The "***Expected-Value***" model becomes:

   $$z_{ev} = \min f(x, \bar{\xi}) \ , \tag{2.24}$$

   where x represents a vector of first-stage decisions.

   The expectation of the expected value problem over the given scenarios is defined as:

   $$z_{eev} = \mathrm{E}_{\omega}\left[f(\bar{x}(\bar{\xi}), \xi(\omega))\right] \tag{2.25}$$

b. A second approach that relies on perfect information is called the "***Wait-and-See***" model:

$$z^{\omega} = \min f(x,\xi(\omega)),\ x \in C^{\omega}, \forall \omega \tag{2.26}$$

where C^{ω} represents a feasible set when all uncertain parameters belong to the scenario ω.

$$z_{ws} = \mathrm{E}(z^{\omega}) = \sum_{\omega \in \Omega} p(\omega) z^{\omega} \tag{2.27}$$

c. A third approach so-called "***Here-and-Now***" where the decision-maker makes the decision "now":

$$z_{hn} = \min \mathrm{E}[f(x,\xi)] = \min \mathrm{E}[cx],\ x \in C = \bigcap_{\omega \in \Omega} C^{\omega}, \tag{2.28}$$

where the optimal solution $x^* \in C$ hedges against all possible (known) contingencies $\omega \in \Omega$ that may occur in the future.

To assume the expected value scenario will occur and accept the solution of the expected value problem, is not always the right decision, since the expected value might be far from the scenario that actually takes place. The Wait-and-See problem cannot be implemented in reality, as the decision-maker must wait to take the decision only when the uncertainty is resolved. This is not realistic, since the decision-maker needs to decide before hand. Therefore, to consider at a time all the possible (known) scenarios, the so-called Here-and-Now approach, is the most appropriate, since it hedges against all the uncertain future outcomes. The solution that this model provides is not optimum for any one outcome, but is the best for many outcomes considered altogether.

To verify whether the stochastic approach is better than any other, some SP analysis must be carried out. The analysis of SP models requires that we

(i) investigate the underlying expected value (EV) problem, and

(ii) compute stochastic information, such as the Expected Value of Perfect Information (EVPI), and the Value of the Stochastic Solution (VSS) defined below:

(a) *Expected Value of Perfect Information (EVPI)*:

EVPI measures the maximum amount a decision-maker would be ready to pay in return for complete (and accurate) information about the future.

Let $z_{hn} = \min_x \mathrm{E}_\xi f(x,\xi)$, and $z_{ws} = \mathrm{E}_\xi \left[\min_x f(x,\xi)\right]$ then:

$$\mathrm{EVPI} = z_{hn} - z_{ws} \tag{2.29}$$

(b) *Value of the Stochastic Solution (VSS)*:

Let $EV = \min_x z(x,\bar{\xi})$, where $\bar{\xi} = \mathrm{E}(\xi)$, and $\bar{x}(\bar{\xi})$ an optimal solution to the EV problem. Let

$$EEV = \mathrm{E}_\xi \left[f\left(\bar{x}(\bar{\xi}),\xi\right)\right] \tag{2.30}$$

Then, VSS measures the cost of ignoring uncertainty in choosing a decision and is defined as:

$$\mathrm{VSS} = z_{eev} - z_{hn}. \tag{2.31}$$

It can be shown that the three objective function values z_{eev}, z_{hn}, z_{ws} are connected by the following relationship:

$$z_{ws} \le z_{hn} \le z_{eev} \tag{2.32}$$

The inequality:

$$z_{hn} \le z_{eev} \tag{2.33}$$

can be argued in the following way: any feasible solution of the average value approximation is already considered in the Here and Now model, therefore the optimal Here and Now objective must be better.

Bounds on EVPI and VSS

For functions $f(x,\xi)$ convex in both arguments some useful bounds on the EVPI and VSS are presented below:

$$0 \leq EVPI \leq z_{hn} - z_{ev} \leq z_{eev} - z_{ev} \tag{2.34}$$

$$0 \leq VSS \leq z_{eev} - z_{ev} \tag{2.35}$$

These can help in estimating the relative benefit of implementing the computationally costly SP solution, as opposed to approximate solutions obtained by processing the Expected Value LP problem.

Chapter 3. Representation of Uncertainty

In order to solve SP problems, which are formulated in Chapter 2, stochastic processes representing evolution of data are discretised via scenarios. Below we give the motivation for scenario generation and then move on to review some of the most popular approaches used in SP context.

Mulvey and Thorlacius (1998) specify the following three main purposes of scenario simulations based on the aim of the simulation:

Prediction. Methods that generate a point forecast, a single most likely scenario. They are mainly used for short-term market movements, hourly, daily. As the time horizon increases the single scenario approach becomes more risky.

Pricing. The value of the security at the end of the time horizon is evaluated on a set of scenarios. Here the no-arbitrage condition plays a prominent role. If the value of the security priced on the scenario tree is not consistent with its current market price then it indicates the existence of arbitrage.

Risk Analysis. Stochastic Asset Liability Models (ALM) fall within this category. Theses models evaluate potential risks of not meeting liabilities and rewards of different investment strategies on a set of scenarios. The riskiness of a particular strategy depends on the investor's liability structure. For example, long-term bonds may be too risky for some short-term investors but could be perfectly adequate for a pension fund with a long-term time horizon.

Single-point forecasts, which represent the first category above, have been widely criticized when applied in the context of efficient asset allocation. For example, mean-variance optimisation models, based on the work of Harry Markowitz, are driven by forecasted (expected) return, volatility and correlations, which are very sensitive to forecasts. Portfolio allocations tend to swing to extreme values due to some small changes in the forecasted parameters.

The problem is compounded by the fact that assets returns, volatilities and correlations are very difficult to estimate accurately and such estimates usually depend on rather arbitrarily chosen historical samples, which may not be representative of current conditions. Besides, there is no room for investors to express their own views with regards to the future, as well as a degree of confidence in those views. Standard models cannot distinguish between strong and week believes in the forecasts and treat them the same; see Koskosidis & Duarte (1997) for a thorough discussion of the subject.

The above-mentioned shortcomings of application of point forecasts to optimal asset allocation problem motivate the use of scenarios, which provide flexibility in modelling historical assets returns. Scenarios also let the investor "manipulate" the historical patterns according to his beliefs and his degree of confidence in those beliefs. Instead of relying on forecasted parameter values, which in essence represent a single scenario, the model can be solved on a set of plausible scenarios of future assets returns, and hence diversify the portfolio to take into accounts a large number of potential return outcomes. By doing this investors can structure a portfolio, which takes into consideration a variety of market conditions.

In stochastic optimisation problems scenarios represent the discretisation of continuous joint distribution of uncertain parameters, which take into account the co-movements of these uncertain parameters. In a multistage stochastic program a scenario tree represents the evolution of random parameters. The scenarios are not restricted to any particular probability distribution or stochastic process and provide the flexibility to model any distribution e.g. asymmetric or fat tailed distribution. Over the last twenty years there have been a number of scenario generation methods proposed in the literature. Most of these methods are problem-specific. Also, within the SP framework scenario simulation can be used for both building scenario trees for further input to the SP problem and for ex poste simulation. In the latter case, the model solutions are evaluated via out-of-sample simulations after the SP problem is solved and first-stage decisions are implemented.

Usually the scenario generation process for SP problems takes the following steps:

- Choose and calibrate a data process, which governs the dynamics of random variables.
- Sample random variables from their probability distributions. Some exceptions include the MM method where scenarios are not sampled but represent optimal solutions to an optimisation problem.
- Combine sampled scenarios in a certain way for further input in the SP problem.

In what follows in the rest of this chapter we review some of the most common methods used in each step of the scenario tree construction process.

3.1 Data Processes

In this section we give a brief overview of some of the data processes governing the dynamics of random variables (assets returns).

3.1.1 Stochastic differential equations (SDE) model

This method consists of specifying continuous time SDE (in the tradition of Merton (1990)) for the dynamics of the economic and financial variables of interest. Then discretise the time parameter in order to obtain the corresponding system of difference equations; calibrate the output of simulations to the historical data using various (ad hoc) methods to adjust parameters, see Dempster & Thorlacius (1998).

Mulvey & Thorlacius (1998) describe the scenario generation system, called CAP:Link, developed by Towers Perrin, actuarial-consulting company. The system uses an assumption that the asset returns in any other country are highly correlated to main economic indicators such as treasury yield, price inflation and dividend yield among others from three major world economies: United States, Japan, Germany. The projection of an asset return in a smaller country is related by means of SDEs to the main economic variables in the three largest economies at the previous time period as well as to the stochastic elements of the equation and to other explanatory variables and factors. The following example shows the SDEs used to generate scenarios of interest rates. Since the scenarios are generated for a number of countries simultaneously all the elements in the equations are indexed as vectors.

Short rate: $dr_t = f_1(r_u - r_t)dt + f_2(r_u, r_t, l_u, l_t, p_u, p_t)dt + f_3(r_t)dZ_1$ (3.1)

Long rate: $dl_t = f_4(l_u - l_t)dt + f_5(r_u, r_t, l_u, l_t, p_u, p_t)dt + f_6(r_t)dZ_2$, (3.2)

where

r_u is the normative level of short interest rates,

r_t is the level of interest rates at time t,

l_u is the normative level of long interest rates,

l_t is the level of rates at time t,

p_u is the normative level of inflation and p_t is the level at time t, and

$f_1,..., f_6$ are vector functions that depend upon various economic factors up to period t.

The random coefficient vectors - dZ_1 and dZ_2 - depict correlated Wiener terms.

Scenarios are generated by sampling from the stochastic term of the stochastic difference equations, where each scenario represents a particular sequence of realisations of the white noise term over the planning horizon. Variance reduction methods such as antithetic sampling were used.

Darius, Ilhan, Mulvey, Simsek & Sircar (2002) use SDE to generate scenarios of asset returns in the context of dynamic multi-period assets allocations to hedge funds.

3.1.2 Econometric model calibrated to historical data

Another widely used approach first introduced by Sims (1980) is vector autoregressive (VAR) modelling. This method applies on a rolling forward basis.

where current and past values of the variables of interest predict future realisations of these variables. This approach has proven to be a success in many ALM and pricing applications due to its adaptability to changing economic conditions, see Carino *et al.* (1994) and Berkelaar *et al.* (1999) among others. A typical model is represented as follows:

$$y_t = \mu + B_1 y_{t-1} + \ldots + B_k y_{t-k} + \varepsilon_t \tag{3.3}$$

where μ, y_t, and ε_t are $n \times 1$ vectors, B_i are $n \times n$ matrices, and $\varepsilon_t \sim IIDN(0, \Sigma)$, $t = 1, \ldots, T$. Such models are usually estimated by generalised least squares using Zellner's (1962) seemingly unrelated regression techniques, since all single models (estimated simultaneously) are related to one another through the variance-covariance matrix, Σ.

Once the model in (3.3) is estimated all the terms, except for ε_t, on the right-hand-side of the equation become known at time t – 1. Thus, in order to generate a scenario of y_t emanating from the current tree node one draws a random sample from the distribution of error term, ε_t and adds it to the rest of the terms on the right-hand-side of equation (3.3). In order to generate M scenarios the same process is repeated M times. Due to random sampling all the scenarios emanating from the current node are equiprobable with probability $\frac{1}{M}$.

Sometimes, when used for long-term horizon predictions the forecasts may diverge from an equilibrium level. This can be rectified by inclusion of an equilibrium condition in the VAR model, the resulting type of models is called *Error Correction Models*, see Boender *et al.* (1998) and Volosov *et al.* (2005).

Russell-Yasuda Kasai model (see Carino *et al.* 1994) is the first genuine commercial application of asset liability management methodology developed for a Japanese insurance company. The scenario-generating module of the model provides three different methods for scenario construction of asset returns.

The *first* method builds scenarios assuming inter-temporal independence of asset returns. The user has to estimate the joint probability distribution of asset returns for each time stage, for example, by using historical data. Then a required number of scenarios are sampled from the estimated distribution. In order to keep the decision model computationally tractable and to conform to the scenario tree structure the number of scenarios is reduced. The scenario reduction technique pairs up sampled scenarios, calculates the probability-weighted mean value of asset returns for the new consolidated scenario. The probability of the new scenario will be the sum of the probabilities of the original two scenarios. Further reduction can be achieved by applying the same technique to already consolidated scenarios.

By applying the above scenario reduction technique the mean value is preserved while the variable may differ from that of the target distribution. In this case variance-adjustment is used by simultaneously moving all the asset returns from or towards the mean value until the desired variance is achieved. These movements of asset returns are proportional to their distances from the mean. This way the shape of the asset return distribution is preserved.

One of the drawbacks of the variance adjustment is that the procedure is applied to each random variable individually. As a result, the correlations among the variables are not taken into account, thus rendering scenarios, which do not correspond to the joint probability distribution.

The *second* method of scenario generation provided by the Russell-Yasuda Kasai model uses statistical factor analysis to capture inter-temporal dependence of the asset returns. It assumes that the three factors relate to interest rates, equity return and exchange rates. Time series analysis is used to model time evolution of the factors.

Boender (1997) and Kouwenberg (2001) applied a vector autoregression approach to build scenarios of asset returns as well as wage growth rates for an ALM simulation system of a Dutch pension fund. First, a time series model was fitted to historical data and a distribution of resulting error terms were estimated. Then sampling from the error distribution was used to construct the scenarios.

The *third* method of scenario generation provided by the Russell-Yasuda Kasai model uses user-defined scenarios (qualitative methods).

3.2 Sampling Methods

In this section we review some of the sampling techniques, which are commonly used in scenario generation process.

3.2.1 Pure random sampling

Random sampling is performed as follows. Assuming a vector autoregression model was fitted to the vector of random variables of interest, the probability distribution of the vector of error terms is estimated. If N scenarios are required by the end of the first time period then N random samples are taken from the distribution of errors, each corresponding to a particular tree node. At each of the N (predecessor) nodes a further M (successor) scenarios (nodes) could be generated in a similar fashion. The scenarios for the second time period will be conditional on the state of the node at the end of the first time period. In order to take conditional distribution random variables into account values of random variables realised at the end of the first time period are added to the data sample and the vector autoregression model and probability distribution of the error terms are re-estimated. Thus, the scenario tree is constructed recursively, stage by stage, using a conditional distribution of next period random variables based on the current node of the tree.

The scenario tree built using random sampling can be regarded as a single random drawing from the underlying distribution and hence is itself random. Since the optimal solution of the problem depends on random scenario trees they also become random. One of the drawbacks of the random sampling is that if the sparse branching structure of the event tree is used, i.e. the number of successors is relatively small, then the optimal investment decisions may become unstable and change significantly from one tree to another. This could be attributed to the fact that in sampling a small number of error terms we may incorrectly represent statistical moments of the multivariate continuous distribution of random variables and as a result the investment strategy is chosen based on erroneous approximation of the distribution function. Sampling uncertainty of optimal solutions can be assessed by sampling many scenario trees, solving a SP model on each of the trees and hence obtaining a distribution of optimal solutions.

As mentioned above, Kouwenberg (2001) used random sampling to construct a scenario tree of asset returns and liabilities for a Dutch pension fund. He sampled randomly from the error term distribution of the VAR model fitted to historical data. The 5-year horizon model had 1-10-6-6-4-4 branching structure and hence $10\times6\times6\times4\times4 = 5760$ nodes at time period 5. The optimal solution to the ALM problem indicates excessive changes in asset mix over time. The instability in the optimal solution can be caused, as mentioned above, by the sparse branching structure of the scenario tree. At each time period there are no more than 10 states to represent the conditional joint distribution of random variables rendering a poor approximation. The problem is aggravated by the fact that these states are sampled randomly thus introducing sampling errors in statistical characteristics (e.g. moments) at most nodes on the tree. As a result, the optimal investment strategy at every node of the tree is based on limited and incorrect information.

One obvious remedy to reduce approximation error is to increase the amount of nodes in the tree. However, the SP problem may become computationally intractable since the size of the tree will grow exponentially with time when the sample (number of successor nodes) is increased.

3.2.2 Antithetic random sampling

Antithetic random sampling can remedy some of the problems of pure random sampling method, e.g. match odd statistical moments of symmetric error distribution. Antithetic sampling is conducted as follows. Suppose that N scenarios are required and N is an even number. Then the first $N/2$ scenarios are sampled randomly and the remaining $N/2$ scenarios are produced from the first half by multiplying their values by -1. This way all the odd moments are preserved.

In order to match the variance the scenario values are rescaled by moving their values from / to the mean of the distribution until the target variance is achieved, see Carino *et al.* (1994). The value of the shift of each error term is proportional to the distance of the error value from the mean. Thus adjusted error terms are substituted back into the estimated vector autoregressive model to obtain nodes of the tree (scenarios).

Kouwenberg (2001) used antithetic random sampling to construct a scenario tree of asset returns and liabilities for a Dutch pension fund. As expected, the optimal solution to the SP problem is more stable over time than that when pure random sampling was used to generate the scenario tree. Besides, the spurious profits resulted from pure random sampling were reduced.

3.2.3 Moment matching methods

Let $Z = (Z_1, \ldots, Z_N)'$ be a random sample from a standard normal distribution. The sample moments of Z will not exactly match statistical moments of a standard normal distribution. The idea of the methods proposed by Barraquand (1993) is to use certain transformations to convert the original sample $Z = (Z_1, \ldots, Z_N)'$ to a transformed sample, say, $\tilde{Z} = (\tilde{Z}_1, \ldots, \tilde{Z}_N)$, such that the finite number of its moments match those of the underlying population.

For example, in order to match mean values the following transformation is made:

$$\tilde{Z}_i = Z_i - \bar{Z}, \qquad i = 1, \ldots, N, \tag{3.4}$$

where $\bar{Z} = \sum_{i=1}^{N} Z_i / N$ is the sample mean of Z_i. $\tilde{Z}_i$ are normally distributed if Z_i are normal.

If the underlying distribution has a mean value other than zero the following transformation is necessary:

$$\tilde{Z}_i = Z_i - \bar{Z} + \mu_Z, \qquad i = 1, \ldots, N, \tag{3.5}$$

where μ_Z is the mean of the population.

This idea can be extended to match moments higher than one. For example, in order to match the first two statistical moments the following transformation of sampled standard normal variates is necessary:

$$\widetilde{Z}_i = (Z_i - \overline{Z})\frac{\sigma_Z}{s_Z} + \mu_Z, \qquad i = 1,\ldots,N, \tag{3.6}$$

where s_Z is the sample standard deviation of Z_i and σ_Z is the population standard deviation.

Boyle *et al.* (1997) provide illustration of the method in the context of option pricing.

3.2.4 Stratified sampling and Latin hypercube sampling

Stratified sampling is a variance reduction technique, which seeks to make sampling less random and more regular. It ensures that certain empirical probabilities match their theoretical counterparts.

For example, if one samples 100 times from a normal distribution, the resulting sample will not be exactly normal and the tails of the sample distribution will most likely be underrepresented. To mitigate the discrepancies *stratified sampling* can be used to force exactly one observation lie between the (i-1)-th and the i-th percentile, $i = 1,\ldots,100$ and hence to represent a better match of the normal distribution. The algorithm proceeds as follows:

1. 100 independent random variates, $U_1,\ldots,U_{100}$, from a uniform distribution on $[0,1]$ are sampled;

2. the sampled uniform variates are transformed as follows: $\widetilde{Z}_i = \Psi^{-1}((i + U_i - 1)/100)$, $i = 1,\ldots,100$, where Ψ^{-1} is the inverse of the cumulative normal distribution.

The above algorithm works because $(i + U_i - 1)/100$ falls between the (i-1)-th and the i-th percentiles of the uniform distribution and the inverse transformation preserves the percentiles.

The obtained $\widetilde{Z}_i,\ldots,\widetilde{Z}_{100}$, are not independent, which complicates computation of standard errors and hence confidence intervals. In order to compute confidence intervals simulation runs should be batched, see Boyle *et al.* (1997) for more details.

Stratified sampling can be applied to higher dimensions. To generate a stratified sample from a d-dimensional unit hypercube, with n strata in each coordinate, the sequence of vectors $U_j = \left(U_j^{(1)},\ldots,U_j^{(d)}\right)$, $j = 1,2,\ldots$, is generated and then the stratified variate is $V_j = \dfrac{U_j + (i_1,\ldots,i_d)}{n}$, $\quad i_k = 0,\ldots,n-1, \quad k = 1,\ldots,d$. Exactly one V_j will lie in each of the n^d cubes, which represent a product of n strata in each dimension.

It is very difficult to apply stratified sampling in higher dimensions unless n is small. Latin hypercube sampling, which was first introduced by McKay *et al.* (1979), can overcome some of the difficulties with stratified sampling. The method is summarised as follows:

1. sample d independent random permutations, $\pi_1,\ldots,\pi_d$, of $\{1,\ldots,n\}$, where each sample is uniformly distributed over all *n!* possible permutations.

2. Set $V_j^{(k)} = \dfrac{U_j^{(k)} + \pi_k(j) - 1}{n}$, $\quad k = 1,\ldots,d$, $j = 1,\ldots,n$

The resulting $V_j^{(k)}$ are uniformly distributed over the d-dimensional hypercube and perfectly stratified, that is exactly one of $V_1^{(k)},\ldots,V_n^{(k)}$ falls between $(j-1)/n$ and j/n, $j = 1,\ldots,n$, for each dimension $k = 1,\ldots,d$. As before, resulting sample observations are not independent, hence run batching is required in order to estimate standard errors, see Boyle *et al.* (1997).

3.2.5 Bootstrapping historical data

Bootstrapping represents sampling from the historical distribution. Mulvey & Vladimirou (1989), Koskosides & Duarte (1997) and Beltratti *et al.* (2004) are some of the examples where bootstrapping was used to generate scenarios. The main assumption of this approach is that the events that occurred in the past will reoccur in the future with the same probabilities though not necessarily in the same order. Hence the intertemporal dependence of returns is not preserved.

This method of scenario generation relies only on the historical data. If, for example, monthly data is used, then a month is sampled randomly from the historical sample. Then the asset returns and / or the risk factors associated with that month are used as a particular realisation of random asset returns for the future time period (month). If the time horizon consists of several months then in order to generate a scenario path the same procedure is repeated several times, ones for every consecutive month. Following this approach the correlations among asset classes are preserved.

3.2.6 Sampling from lattice models

Zenios and Shtilman (1993) address the problem of sampling from a binomial lattice of term structure of interest rates. The practical application of the developed method is the valuation of various interest rate contingencies with path-dependent cash flows, namely, mortgage pass through securities and single premium annuities. They proposed a sampling method of interest rates paths on a lattice, which satisfies some optimality conditions such as these paths can be used to estimate the mean value of discount functions within user-specified limits. The second feature of their method is that it does not require random sampling and is more efficient compared to the Monte Carlo method, since it required fewer samples to achieve a specified level of accuracy.

If Ω represents a scenario set (all possible paths on a binomial lattice) then it consists of a huge number of scenarios, for instance, if the time horizon is 30 years and the time step is one month then the set Ω will consist of 2^{360} elements. It implies that the calculation of expected value over such a large set is practically impossible. Zenios and Shtilman (1993) suggest substituting the original set Ω by a smaller set $\Omega' \subset \Omega$ with an "acceptable" number of scenarios in Ω', which will depend on the required level of accuracy.

The idea of the approach is similar to that of stratified sampling. It partitions the interval (set), on which samples are taken, into mutually disjoint sub-intervals and sampling is taken on each sub-interval. The difference of their method from stratified sampling is that they sample only one point from each sub-interval in a certain non-random manner (see Zenios & Shtilman (1993) for more details on the procedure for constructing optimal samples).

3.2.7 Sampling correlated random variables and conditional sampling

RiskMetrics (1996) methodology is based on estimation of volatilities and correlations of asset classes. The major assumption is that a random vector of asset returns has a jointly normal distribution with the following probability density function:

$$f(\xi) = (2\pi)^{-n/2} |\Sigma|^{-1/2} \exp\left[-\frac{1}{2} (\xi - E(\xi))' \Sigma^{-1} (\xi - E(\xi)) \right], \tag{3.7}$$

where ξ is a vector of random variables such as asset returns, n is the cardinality of ξ, $E(\xi)$ is expectation of ξ and Σ is the estimated covariance matrix of random variables.

Once parameters of the multivariate normal distribution are estimated, say, using historical data, *Monte Carlo simulation* can be used to randomly sample from this distribution. In order to take correlation among random variables into account one can use Cholesky decomposition (approach used in Riskmetrics) or *principal components* analysis (see Jamshidian & Zhu (1997)).

Loretan (1997) and Topaloglou *et al.* (2002) first find principal components, which are independent of each other by definition and then sample these components instead of the original variables. Apart from preserving correlation among the random variables this approach has an additional advantage of reducing dimensionality of the problem (usually the first three principal components are enough to explain most of the variability in random variables) and hence reducing the number of scenarios.

Cholesky decomposition is applied when using Monte Carlo simulation in order to preserve the covariance structure of asset returns. To illustrate the application of Cholesky decomposition, assume that the vector of N independent standard normal variables, $Z = (Z_1, \ldots, Z_N)'$, is sampled. The covariance matrix of vector Z is the identity matrix, i.e. $\Sigma(Z) = I$. Cholesky decomposition of the covariance matrix of asset returns, Σ, is a triangular matrix, C, such that $\Sigma = CC'$. Now, vector Z can be transformed into the vector of random normal correlated asset returns, R, by multiplying it by C, such that $R = CZ$. Since $\Sigma(R) = C\Sigma(Z)C' = CIC' = \Sigma$, the vector R reflects the targeted covariance structure.

Halling *et al.* (2005) use Simple Monte Carlo (SMC) and Improved Monte Carlo (IMC) techniques to construct scenarios of correlated asset returns.

- SMC scenario generation of a vector of correlated asset rates of returns, R, is carried out as follows. Let us assume that we need to simulate N scenarios for end of the period rates of assets returns and there are K assets in the portfolio. First, Monte Carlo simulate the K-dimensional vector of independent standard normal random variables, Z. Then multiply by the Cholesky decomposition of the covariance matrix Σ in order to preserve correlations among the variables comprising the vector, $R = CZ$. Then the vector of next period returns can be obtained by adding the mean rate of return vector, μ, such as $R^* = 1 + \mu + Z$. If the time interval between decision dates is τ time periods then the vector of simulated asset returns equals $R^* = (1+\mu)^\tau + Z \cdot \sqrt{\tau}$.

  The main drawback of the SMC method is that statistical moments of the sample may differ substantially from the theoretical (target) moments. One of the remedies such as antithetic sampling is described above.

- IMC scenario generation technique applies the quadratic resampling method developed by Barraquand (1993). Unlike SMC, which preserves only the first statistical moment, IMC preserves both mean values and variance-covariance structure of asset returns. It starts with sampling a random vector $\tilde{Z}$ from a multivariate normal distribution. Let $\tilde{\mu}$ and $\Sigma(\tilde{Z})$ be the sample mean and sample covariance matrix of $\tilde{Z}$ respectively. Cardinality of vector $\tilde{Z}$ corresponds to the number of assets in the portfolio, K in our case. Then we can obtain a vector Z, such that $Z = C^{-1}\tilde{Z}$, where C is the Cholesky decomposition of $\Sigma(\tilde{Z})$. Vector Z follows a K-dimensional standard normal distribution with zero mean and identity covariance matrix. The final step is to multiply Z by the Cholesky decomposition of the theoretical covariance matrix Σ, such that $\breve{Z} = \tilde{C}Z$, where $\tilde{C}$ is the Cholesky decomposition of Σ. To generate scenarios of correlated asset returns over τ time periods one can use the following equation: $R^* = (1+\mu)^\tau + \tilde{C} \cdot Z \cdot \sqrt{\tau}$, where μ is the vector of theoretical mean rates of return.

The above approaches, where all random variables are sampled simultaneously, can be extended to the case when one set of random variables is simulated conditional on the values taken by the other set of random variables. This idea was applied by Beltratti, Consiglio & Zenios (1999) where exchange rates were simulated conditional on the values of interest rates using multivariate normal distribution of both exchange rates and interest rates. Scenarios of interest rates were constructed beforehand using a binomial lattice approach.

The conditional simulation from a multivariate normal distribution is conducted as follows. First, a vector ξ of random variables with cardinality N is divided into two subvectors: K-dimensional vector ξ_1 and $(N-K)$-dimensional vector ξ_2. The expected values and covariance matrix can be partitioned as follows:

$$\bar{\xi} = \begin{bmatrix} \bar{\xi}_1 \\ \bar{\xi}_2 \end{bmatrix} \text{ and } \Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \tag{3.8}$$

The conditional probability density function of ξ_2 given ξ_1 is

$$f\left(\xi_2 \mid \xi_1 = \xi_1^*\right) = \left(2\pi\right)^{-n2/2} \left| \Sigma_{22.1} \right|^{-1/2} \exp\left[-\frac{1}{2}\left(\xi_2 - \bar{\xi}_{2.1}\right) \Sigma_{22.1}^{-1} \left(\xi_2 - \bar{\xi}_{2.1}\right) \right] \quad (3.9)$$

where the conditional expectation and conditional covariance matrix are given by $\bar{\xi}_{2.1}\left(\xi_1^*\right) = \left(\bar{\xi}_2 - \Sigma_{21}\Sigma_{11}^{-1}\bar{\xi}_1\right) + \Sigma_{21}\Sigma_{11}^{-1}\xi_1^*$ and $\Sigma_{22.1} = \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}$ respectively.

Using equation (3.9) one can generate scenarios of ξ_2 at some future time period t, conditional on ξ_1 taking value ξ_1^* as follows:

$$\xi_{2i}^{t} = \xi_{2i}^{0} \exp\left[\sigma_i \sqrt{t}\, \xi_{2i}\right], \quad (3.10)$$

where ξ_{2i}^{0} is the current (time t = 0) value of the i-th component of vector ξ_2, σ_i is the one period standard deviation of the i-th component of vector ξ_2 and ξ_{2i} is the i-th component of the conditionally sampled vector ξ_2, distributed according to equation (3.9).

3.3 Scenario Tree Construction Methods for SP Problems

Let us assume that we have estimated parameters of a discrete-time continuous-space stochastic process, $\left(\tilde{\xi}_t\right)_{t=1,2,\ldots}$, which governs the dynamics of our random variables using methods examined in section 3.1. The problem here is that multiperiod optimisation problems formulated on continuous-state stochastic processes cannot be numerically solved because such decisions are functions, which makes the problem a functional optimisation problem. In order to make the problem solvable one has to restrict the problem to a discrete-state optimisation problem, hence to find a good approximation ξ_t, which will be as close to the original stochastic process, $\left(\tilde{\xi}_t\right)_{t=1,2,\ldots}$, as possible. For a discrete-time stochastic process the history process, i.e. $\left(\left(\tilde{\xi}_1\right)\left(\tilde{\xi}_1, \tilde{\xi}_2\right), \left(\tilde{\xi}_1, \tilde{\xi}_2, \tilde{\xi}_3\right), \ldots\right)$, can be represented as a tree with certain predetermined tree structure. A tree, which approximates $\left(\tilde{\xi}_t\right)_{t=1,2,\ldots}$ and is used as a basis for decision-making is called a scenario tree (see Hochreiter & Pflug (2002)).

An SP model is based on a scenario tree of random variables, such as asset returns, liabilities, etc. The scenario tree shows how uncertainty is revealed over time and this affects the realisation of random variables. The complexity of the tree is determined by the "bushiness" of the tree, which represents the number of successor at each tree node (branching factor). The branching factor could be different at each node is it is common to have a higher branching factor at earlier stages than at later stages of the tree. A binary tree has a branching factor 2, a ternary tree has a branching factor 3 etc.

Figure 3-1 shows a scenario tree, which is a fan with flat out-of-sample scenarios. Every node, other than the root node, of such a tree has a branching factor 1. All the uncertainty is "concentrated" at the initial (current) stage (node) of the tree and this initial decision should incorporate future knowledge along scenarios.

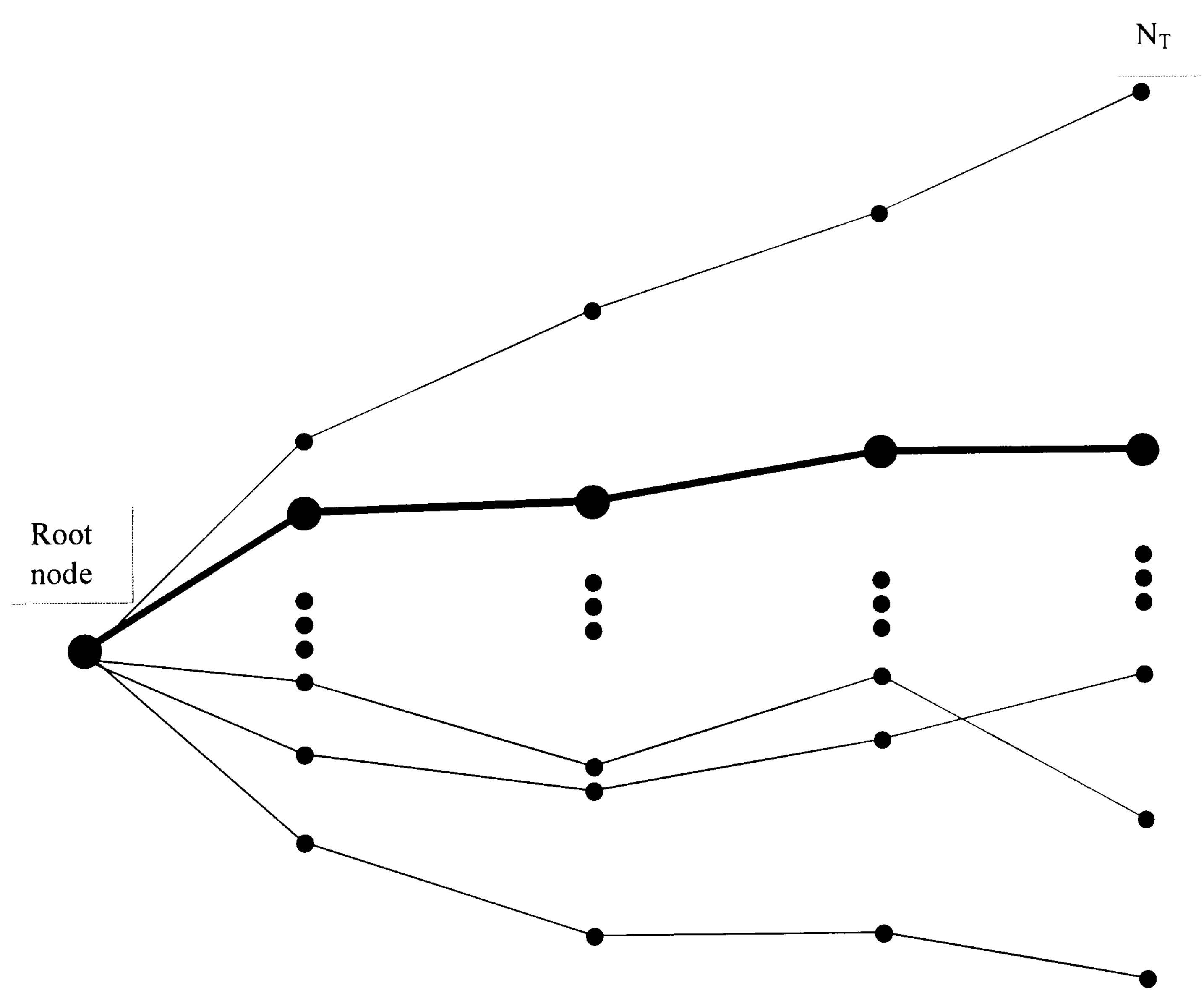

Figure 3-1 Scenario fan

The problem of generating independent scenarios as illustrated in Figure 3-1 is then such a structure is inappropriate for a multistage programming. In a tree for multistage programming information should be revealed gradually and not everything after the first stage as it happens in a “fan”.

A typical multistage scenario tree with the branching structure is depicted in Figure 3-2. The number of stages corresponds to the number of time periods t = 0, 1, …, T, when the decisions are or can be made. The root node (denoted as n = 0) corresponds to the initial t = 0, current time period when the information is certain and the actual decisions are implemented (initial asset allocation in the portfolio). The branches emanating from a node represent realisation of uncertainty and the successor nodes correspond to different states of the world and hence different realisations of random parameters conditional on the information (state) available at a predecessor node. Each node in the tree corresponds to a discrete realisation of the joint probability distribution of all random variables conditional on the state of the immediate predecessor node. Decisions at all nodes except the root and leaf nodes show the hypothetical decisions conditional on the specific path leading up to that node. The leaf nodes correspond to the terminal nodes of the tree, t = T at which no decisions are made and the terminal values, e.g. terminal wealth of the portfolio, are calculated.

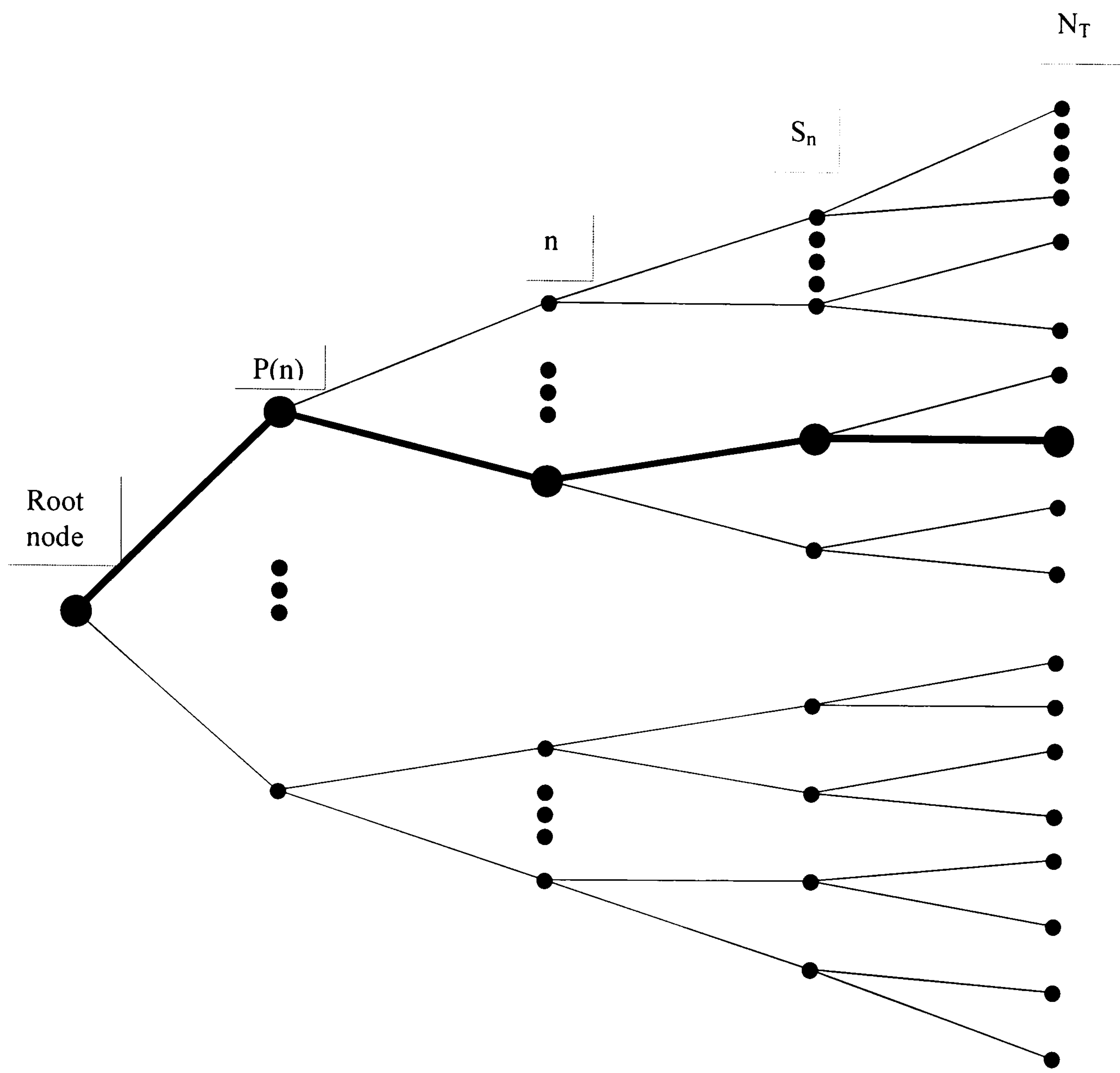

Figure 3-2 Scenario Tree

Each path in the tree, from the root node to the leaf node represents a particular trajectory of a stochastic process underlying the joint evolution of the random variables. The scenario tree need not have a symmetric or binomial structure. In Figure 3-2 we highlight one possible path in the tree for illustrative purposes.

For illustrative purposes we employ the notation used in Topaloglou *et al.* (2004(a)) as shown below.

N is the set of nodes of the scenario tree,

$n \in N$ is a typical node of the scenario tree ($n = 0$ denotes the root node at $t = 0$),

$N_t \subset N$ is the set of distinct nodes of the tree at time period $t = 0,1,...,T$,

$N_T \subset N$ is the set of leaf (terminal) nodes at the last period T, that uniquely identify the scenarios.

$p(n) \in N$ is the unique predecessor node of node $n \in N$,

$S_n \subset N$ is the set of immediate successor nodes of node $n \in N \setminus N_T$. This set of nodes represents the discrete distribution of the random variables at the respective time period, conditional on the state of node n.

p_t^n is the conditional probability for the outcome associated with the transition from the predecessor node $p(n)$ to node $n \in N$,

p_n is the probability of the state associated with node $n \in N$.

The probability of a particular scenario path is the product of the conditional probabilities of all nodes visited by the path. The sum of probabilities of all the constituent paths of the scenario tree equals 1. The sum of probabilities of all nodes at any distinct time period should equal 1, i.e., $\sum_{n \in N_t} p_n = 1, t = 0,1,...,T$. Also the probability of any node is equal to the sum of probabilities of all immediate successor nodes, i.e., $p_n = \sum_{m \in S_n} p_m, \forall n \in N \setminus N_T$.

Scenario trees used in SP problems should have a *non-anticipativity* property. This means that if certain paths coincide up to a particular time stage them the decisions on all the nodes of this path up to that particular time stage should coincide too.

Generally, SP problems are based on *non-recombining* scenario trees. To illustrate this point, suppose the objective function of the SP problem maximises terminal wealth of the portfolio. Then, as shown in Figure 3-3, an increase in the share price during the previous time period may be followed be an increase in the share holding of this asset in the portfolio. At the same time if the share price arrived at the same point after its decline during the previous time period then the objective function might suggest a decrease in the share holding of that asset in the portfolio.

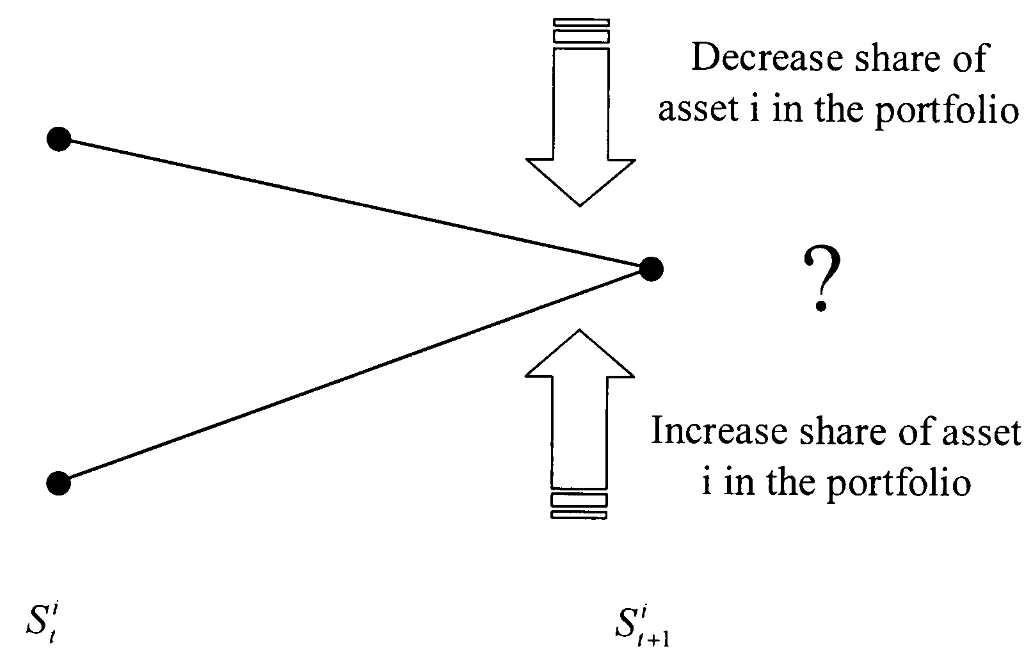

Figure 3-3 Recombining tree of asset prices

As a result, an SP model based on a recombining scenario tree may lead to contradictory decisions. To avoid this problem it is customary to use non-recombining scenario tree for input in SP problems. Non-recombining trees cause exponential growth in variables and constraints for every additional time stage in the SP problem. It is therefore crucial for successful implementation to build sparse and efficient non-recombining scenario trees.

3.3.1 Optimisation-based moment matching methods

We could classify all MM methods into two categories: optimisation-based methods and transformation-based methods. This section gives an overview of optimisation-based method.

Smith (1993), Keefer & Bodily (1983), and Keefer (1994) were the first who suggested MM method for generating scenarios for a static problem. Later, Hoyland & Wallace (2001) extended the idea to a multistage problem. Since then there has been considerable interest in furthering the improvement and application of the method.

The essence of the method is to minimize the least squares of deviations of the targeted moments and co-moments from the moments and co-moments implied by the generated discrete scenarios. The estimated decision variables in this minimization problem represent a vector of end of the period asset returns and associated probabilities, where each element of the vector corresponds to a scenario. The scenario generation model can be formulated as follows:

$$\min_{\xi(\omega),p(\omega)} \sum_{k\in K} w_k \left(f_k\left(\xi(\omega),p(\omega)\right)-SV_k\right)^2 \tag{3.11}$$

$$s.t. \sum_{\omega=1}^{|\Omega_t|} p(\omega)=1, \qquad p(\omega)\geq 0\,,\ \omega\in\Omega_t \tag{3.12}$$

where K denotes the set of all specified statistical properties; SV_k denotes statistical property k, $k\in K$; $\xi(\omega)$ and $p(\omega)$ denote vectors of asset returns and scenario probabilities respectively; $f_k\left(\xi(\omega),\, p(\omega)\right)$ is the mathematical specification of the k-th statistical property formulated as a function of a vector of asset returns $\xi(\omega)$ and a vector of scenario probabilities $p(\omega)$; w_k is the weight of k-th statistical property and Ω_t is the scenario set at each node at stage t, with $\omega\in\Omega_t$.

One should treat the results with caution since generally the knowledge of all moments does not determine the probability distribution uniquely. The following example taken from Grimmett and Walsh (2000) shows that there are continua of different distributions, which have all identical moments.

Log-normal distribution. *If X has the normal distribution with mean 0 and variance 1, then* $Y = e^X$ *has the log-normal distribution with density function*

$$f(y)=\begin{cases}\dfrac{1}{y\sqrt{(2\pi)}}\exp\left[-\frac{1}{2}(\log y)^2\right] & if\ y>0,\\ 0 & if\ y\leq 0.\end{cases}$$

Suppose that $-1 \leq a \leq 1$ *and define*

$$f_a(y) = [1 + a \sin(2\pi \log y)]f(y)$$

It is possible to show that

- f_a *is a density function,*
- f *has finite moments of all orders*
- f_a *and* f *have equal moments of all orders, in that*

$$\int_{-\infty}^{\infty} y^k f(y)\, dy = \int_{-\infty}^{\infty} y^k f_a(y)\, dy \qquad \text{for } k = 1,2,\ldots$$

Thus $\{f_a : -1 \leq a \leq 1\}$ *is a collection of density functions, each different from all the others but all having the same moments*

There are several ways to apply the MM method, see Gulpinar *et al.* (2004) for a more detailed review of MM methods.

Sequential optimisation, shown in Figure 3-4, starts at a root node of the tree, generates scenarios over the next time period, and then at newly generated tree nodes the process is repeated again using conditional next period distribution properties. This process continues in a sequential manner until each scenario path spans the whole planning horizon of the model.

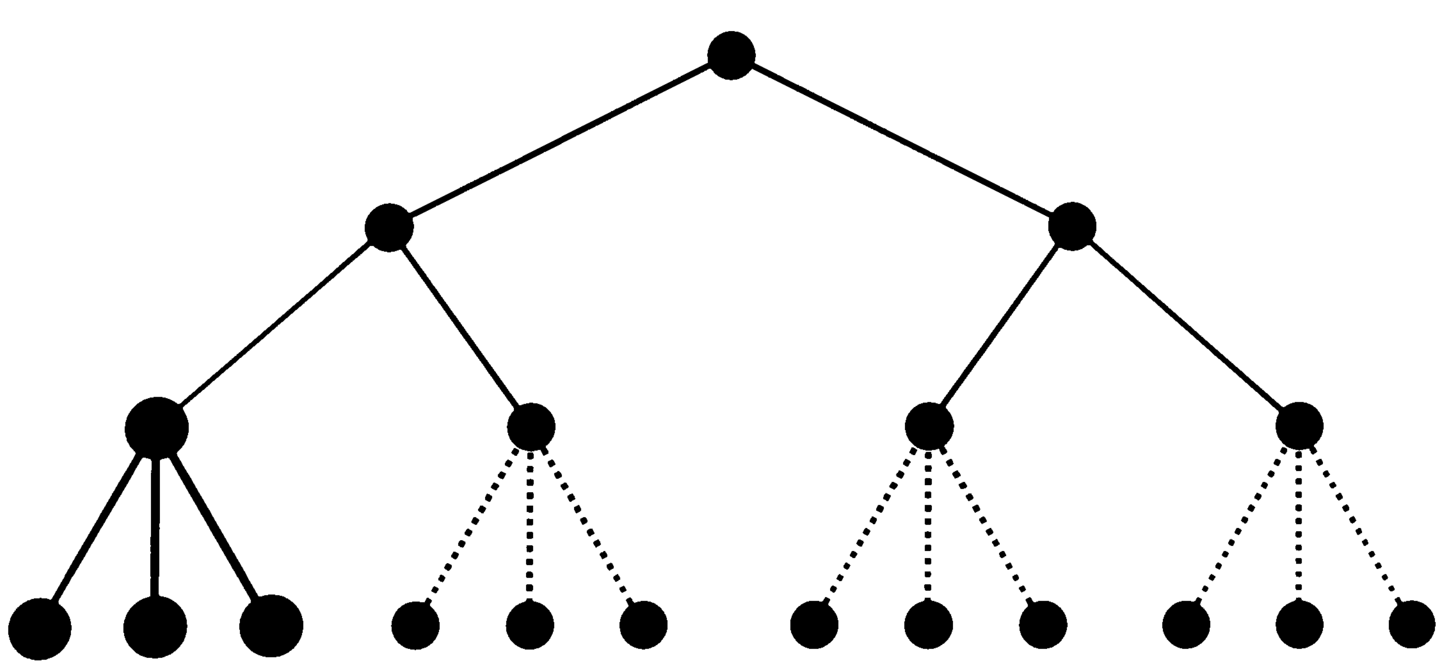

Figure 3-4 Sequential Optimisation

Mathematically, sequential optimisation method that matches the first four statistical moments and covariance of asset returns can be formulated as follows:

$$\min_{\xi(\omega),p(\omega)} \sum_{i\in I}\sum_{k=1}^{4} w_{ik}\left(m_{ik}-M_{ik}\right)^2+\sum_{i,j\in I,i<j} w_{i,j}\left(c_{ij}-C_{ij}\right)^2 \tag{3.13}$$

$$s.t. \quad \sum_{\omega=1}^{|\Omega_t|} p(\omega)=1, \tag{3.14}$$

$$m_{i1}=\sum_{\omega\in\Omega_t}^{|\Omega_t|}\xi(\omega)_i\, p(\omega), \quad i\in I, \tag{3.15}$$

$$m_{ik}=\sum_{\omega\in\Omega_t}^{|\Omega_t|}\left(\xi(\omega)_i-m_{i1}\right)^k p(\omega), \qquad i\in I,\ k=2,3,4, \tag{3.16}$$

$$c_{ij}=\sum_{\omega\in\Omega_t}^{|\Omega_t|}\left(\xi(\omega)_i-m_{i1}\right)\left(\xi(\omega)_j-m_{j1}\right)p(\omega), \qquad i,\ j\in I,\ i<j, \tag{3.17}$$

$$p(\omega)\geq 0,\ \omega\in\Omega_t, \tag{3.18}$$

where $I=\{1,2,\ldots\}$ is the set of assets; M_{ik} for k = 1, 2, 3, 4 is the targeted k-th central statistical moment for asset I. C_{ij} is the targeted covariance between assets i and j. M_{ik} and C_{ij} can be subjectively specified by the user or estimated on a historical sample. Ω_t is the scenario set at each node at stage t, with $\omega\in\Omega_t$. $\xi(\omega)_i$ represents a return on the i-th asset under scenario ω, where $i\in I$ and $\omega\in\Omega_t$; $p(\omega)$ represents a probability of scenario ω, where $\omega\in\Omega_t$; the weight w_{ik} represents the relative importance of k-th statistical property of asset i.

In the formulation above, constraint (3.14) requires that the probabilities of scenarios emanating from the same node sum to one. Constraints (3.15) and (3.16) define the first four central moments, and constraint (3.17) defines co-moments. Constraint (3.18) guarantees positive probabilities. If the moments and co-moments are substituted into the objective function the problem can be restated as a non-linear optimisation problem with linear constraints. It should be pointed out that the problem (3.13)-(3.18) is a non-convex problem and the solution is sensitive to initial values

This approach is computationally simple since it optimises one single-period submodel at a time as well as it can generate subtrees that perfectly match the target moments for each single-period model. The drawback of this approach is that the resulting distribution parameters over a time interval comprising two or more decision stages may not match the target distribution parameters, since the optimisation model matches only one-period distributional properties. Sequential optimisation can also lead to the trees where a perfect match of first period specifications causes poor matching of conditional second period specifications. An alternative approach, which takes the above criticism of the Sequential Optimisation method, is the Overall Optimisation method.

Overall Optimisation approach to apply MM methods constructs the whole tree in one large optimisation model. Here scenarios emanating from all nodes in the tree are derived simultaneously. The example below, taken from Gulpinar *et al.* (2004), with the same notation as for the Sequential Optimisation, shows a typical non-linear problem formulation for scenario tree generation that matches the first four central statistical moments and covariances of asset returns.

$$\min_{\xi(\omega),p(\omega)} \sum_{t=0}^{T-1} \sum_{n\in N_t} \left(\sum_{i\in I} \sum_{k=1}^{4} w_{ikn} \left(m_{ikn} - M_{ikn}\right)^2 + \sum_{i,j\in I, i<j} w_{ijn} \left(c_{ijn} - C_{ijn}\right)^2 \right) \tag{3.19}$$

$$s.t. \quad \sum_{\omega=1}^{|\Omega_t|} p(\omega)_n = 1, \qquad n \in N_t, \quad t = 0,1,\ldots,T-1, \tag{3.20}$$

$$m_{i1n} = \sum_{\omega=1}^{|\Omega_t|} \xi(\omega)_{in} \, p(\omega)_n, \qquad i \in I, \; n \in N_t, \quad t = 0,1,\ldots,T-1, \tag{3.21}$$

$$m_{ikn} = \sum_{\omega=1}^{|\Omega_t|} \left(\xi(\omega)_{in} - m_{i1n}\right)^k p(\omega)_n, \tag{3.22}$$

$$i \in I, k = 2,3,4, \; n \in N_t, \quad t = 0,1,\ldots,T-1,$$

$$c_{ijn} = \sum_{\omega=1}^{|\Omega_t|} \left(\xi(\omega)_{in} - m_{i1n}\right)\left(\xi(\omega)_{jn} - m_{j1n}\right)p(\omega)_n \,, \tag{3.23}$$

$$i,\ j \in I\,,\ i < j\,,\ \ n \in N_t,\quad t = 0,1,\ldots,T-1,$$

$$p(\omega)_n \geq 0\,,\ \omega \in \Omega_t\,,\ n \in N_t,\quad t = 0,1,\ldots,T-1, \tag{3.24}$$

where N_t denotes a set of tree nodes at time t and n denotes a tree node such that $n \in N_t, \quad t = 0,1,\ldots,T-1$.

The model (3.19)-(3.24) suffers from additional computational burden. The size of the problem is increased in terms of decision variables and number of constraints. The degree of non-convexity of the model is also increased creating additional problems with finding a global optimal solution. It may also become infeasible if some one-period subtrees have distributional properties that lead to multi-period trees, which do not match the target distributional parameters. It is less easy to update conditional targeted statistical specifications when applying Overall Optimisation since conditional moments at later time periods become functions of decision variables (generated scenarios) at earlier time periods. Despite the above-mentioned criticism, Hoyland & Wallace have found that the stability of the solution to ALM problem has improved when generating the whole scenario tree in one big optimisation model.

The implementation of the method becomes more complicated in a multi-period framework, since now one has to account for inter-temporal dependencies such as mean-reversion and volatility clumping. Hoyland & Wallace (2001) model volatility clumping as follows:

$$\sigma_{it} = c_i R_{i,t-1} - \overline{R}_{i,t-1} + (1 - c_i)\overline{\sigma}_{it} \tag{3.25}$$

where $c_i \in [0,1]$ is a volatility clumping parameter (the higher the more clumping). r_{it} is the realized return with expectation $\overline{R}_{it}$ and $\overline{\sigma}_{it}$ is the average standard deviation of asset i in period t. The mean-reversion of the bond classes is modelled as follows:

$$\overline{R}_{it} = MRF_i MRL_i + (1 - MRF_i)R_{i,t-1} \tag{3.26}$$

where $MRF_i \in [0,1]$ is the mean reversion factor (the higher the more mean reversion), MRL_i is the mean reversion level and R_{it} the interest rate for bond class i in period t. Mean-reversion implies that there is a long-term equilibrium level to which the asset reverts. Volatility clumping assumes higher volatility after large shocks in the asset markets.

When generating scenarios for the second and third time period Hoyland & Wallace assume state-dependency for the first two statistical moments and state-independency for the third and fourth moments. The correlation matrix estimated for the first time period is assumed to be the same for the second and third periods.

Hoyland & Wallace (2001) also analyse how to specify relevant statistical properties and how to avoid possible pitfalls while doing this. One requirement is that the *derived statistical specifications*, i.e. those properties derived from realised properties in earlier time periods, should not be contradictory or implausible. One example of this would be a conditional distribution when one asset becomes first-order stochastically dominant to another asset, and hence creating arbitrage opportunity. Some statistical properties may be specified implicitly by other statistical properties. An examples of this case is when a mean value in one time period is dependent on the mean value in the previous time period, thus the correlation between these two time periods is specified implicitly. If the *implicit specification* does not correspond to the explicit specification of the same property then the property will become *inconsistent*. Therefore understanding and reconciliation of implicit and explicit distributional specifications is paramount to construction of a consistent scenario tree.

Hoyland & Wallace (2001) also analyse and provide the guidelines for the minimum number of scenarios to match the statistical specifications. They show that an *overspecification* occurs when the total number of scenarios (outcomes) is too small relative to the number of the statistical properties to be matched. On the contrary, too many scenarios cause *underspecification*. If the scenario probabilities are defined as variables and the tree is underspecified than the extra degrees of freedom will cause some scenarios to have zero probabilities. If the scenario probabilities are defined as parameters and the tree is underspecified then the unnecessary scenarios will have the asset values very close to the mean values, thus causing problems when large trees are to be generated. The above discussion shows that when building a scenario tree the number of scenarios should be *balanced* with the number of statistical specifications to be matched.

In order to illustrate how to generate a balanced scenario tree, let us consider a one-period problem where the goal is to approximate a continuous distribution of a single variable by a discrete distribution. As shown by Miller & Rice (1983) (see Appendix A), the first $(2N - 1)$ moments plus the requirement that the sum of the probabilities equals one can be perfectly matched by N scenarios. According to this result the number of variables (number of scenarios plus their probabilities) equals the number of constraints. This method gives a rule of thumb about how many scenarios are needed given the number of degrees of freedom. Also, all resulting probabilities will be positive.

As an example of finding the minimum number of scenarios, taken from Hoyland & Wallace (2001), consider a five-dimensional case, i.e. there are five assets. The first four moments and correlations are to be matched for each asset return. Therefore, there are 5 by 4 moments matchings plus 10 correlations matchings = 30 constraints. The number of variables in the tree is $(D + 1)N - 1$, where D is the dimensionality of the problem (five in this example) and N is the number of scenarios. In order not to have an overspecified tree we need at least 30 variables (since we have 30 constraints). Thus, to obtain the 30 variables we need at least 6 scenarios ($N = 6$) according to the above formula. Though not perfect, this rule of thumb can provide a starting point for selecting the number of scenarios.

Generally it is not always straightforward to determine what statistical properties should be matched when constructing the scenario tree. However, Hoyland & Wallace (2001) postulate that if all the relevant statistical properties are captured than the objective function values of the decision model solved for different scenario trees conforming to these statistical properties should be approximately the same. For example, for a single-period mean-variance model (see Markowitz (1959) for more details) with quadratic objective function all the relevant statistical properties are captured by the first two statistical moments. With such objective function different scenario trees with the same first two moments will yield identical solutions.

Apart from some simple examples such as the above mean-variance model it is often less obvious what statistical properties are relevant and should be included as constraints in the optimisation problem. One way to tackles this dilemma is to solve the problem on a number of different scenario trees. If the objective function values are not stable on this set of trees then add an additional statistical specification to the problem. This process can be carried on by adding more statistical properties until the objective value becomes stable.

The stability of the objective function value can be enhanced even further if sampling is used to generate a tree. First a number of small scenario trees are sampled, where each tree satisfies the same statistical specifications. Then the small trees are aggregated into a large tree while preserving the statistical properties. This large tree is used as an input to the decision optimisation problem. The advantage of sampling small trees is that the noise of the statistical properties of the continuous distribution, not included as constraints, is reduced.

As was shown in the case of the mean-variance problem, it is sometimes easy to figure out what statistical properties to include as constraints. However, quite often constraints of the decision problem may also require certain properties in order to achieve stability. For example, if the capital adequacy constraints are added to the portfolio management problem with quadratic utility function then different optimal solutions and hence different objective function values may be yielded for different scenario trees with only the same first two statistical moments. In this case in order to achieve stability one needs the trees to have identical third and fourth statistical moments as well.

Kouwenberg & Vorst (1998) extend the MM method to build arbitrage-free multiperiod scenario trees, which include contingent claims maturing beyond the first time period. Their proposed scheme consists of the following steps:

Step 1. Fit the first few statistical moments of asset returns and other random variables at the current node of the tree.

Step 2. Enforce arbitrage-free condition

Step 3. Translate scenario tree of asset returns to the scenario tree of asset prices.

Step 4. Calculate derivative prices at successor nodes of the current node using external derivative pricing model from the option pricing literature, such as implied binomial tree (e.g. Rubinstein 1994) or stochastic volatility model (Hull and While 1987).

Step 5. Build the scenario tree by solving one-period models recursively.

Step 4 warrants some additional explanation. As Kouwenberg & Vorst (1998) point out the alternative to the external derivative pricing method is the internal derivative pricing method, which has the following drawback. When using the internal pricing method, first the arbitrage-free scenario tree of the underlying asset prices is built (Steps 1 to 3) and risk-neutral probabilities consistent with the initial asset prices are calculated. Then using the terminal derivative payoff function at the leaf nodes of the tree and risk-neutral probabilities derivative prices at each predecessor node are calculated. The process is repeated recursively. The problem appears when the market is incomplete and hence there will be many risk-neutral probability measures fitting the initial underlying asset prices. Each risk-neutral probability measure will result in a different price process for the derivative. Moreover, by defining prices of derivatives with longer maturities some economic assumptions and empirical fact may be violated. Therefore, using external derivative pricing models helps to enforce known theoretical and empirical properties on derivative price processes. Then, in order to ensure no arbitrage, the "new" set of risk neutral probabilities is calculated for the models, which include both underlying asset prices and externally computed derivative prices.

Topaloglou *et al.* (2004(a)) use the method to generate the scenario tree of joint realisations of asset prices and exchange rates. Whereas the method used captures the desired moments, calculated from historical data, it does not account for the inter-temporal dependencies of random variables such as mean-reversion.

One of the shortcomings of the MM method is that it cannot perfectly replicate the desired distribution. Figure 3-5, taken from Hochreiter & Pflug (2002), shows two different distributions with identical four moments. Despite this one of the key advantages of the MM is that it generates scenarios with some additional distributional features such as asymmetric or / and with heavy tails by means of matching the first four moments and correlations among random variables.

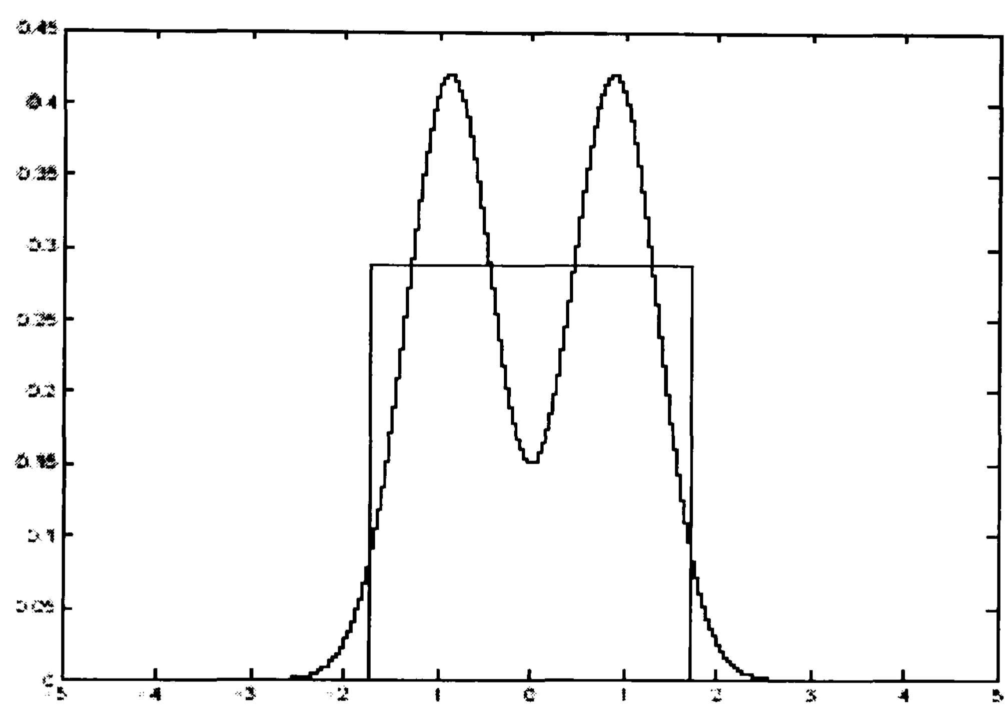

Figure 3-5 Four moments

Kouwenberg (2001) used MM method to construct a scenario tree of asset returns and liabilities for a Dutch pension fund. He applied sequential optimisation at every tree node thus generating the scenario tree recursively taking into account conditional distribution of random variables at every node. The first four statistical moments were fitted for the first three time periods and only the mean and variance were fitted for the fourth and fifth time period of the tree. Kouwenberg (2001) compared random sampling, adjusted random sampling and MM methods to generate scenario trees. Rolling horizon simulations show that MM method is the most stable resulting in the least asset switching over time. SP ALM model solved on a tree fitted by MM method also outperforms fix-mix benchmark model.

One additional benefit in using moments as a summarizing measure of probability distribution is that if we accurately represent the moments of the input distribution (distribution of the random vector) then we can easily and accurately compute the moments of the output distribution (the function of the random vector, e.g.: portfolio of assets). The usefulness of moments as a summarizing measure is related to the effectiveness of polynomial approximation. As an example, consider a simple function of a random variable $g(\xi)$. If we can approximate this function as a polynomial of degree M such as:

$$g(\xi) \approx \sum_{m=0}^{M} a_m \xi^m , \tag{3.27}$$

then the expectation of $g(\xi)$ will be well approximated by the expansion in the moments of the distribution of ξ:

$$E[g(\xi)] \approx \sum_{m=0}^{M} a_m E[\xi^m] \tag{3.28}$$

Therefore, if $g(\xi)$ is well approximated by a polynomial then $E[g(\xi)]$ is accurately computed if the moments of ξ, $E[\xi^m]$, are accurately calculated.

3.3.2 Transformation-based moment matching methods

Whereas the optimisation-based method of moments matching proposed by Hoyland & Wallace (2001) works well for relatively small scenario trees, it becomes very computationally inefficient for reasonably large problems with many asset classes. This drawback is caused by the fact that the method generates scenarios for all assets simultaneously and hence it becomes very slow when the number of assets increases. In order to overcome very long solution times Hoyland *et al.* (2003) proposed a heuristic algorithm, which will be outlined in this section.

The main difference of the method described here from the method by Hoyland & Wallace (2001) is that marginal distributions of each asset are generated individually and then used to create a joint distribution of assets. All marginal distributions are generated with the same number of outcomes and the probabilities of each outcome are the same for all marginal distributions. The i-th scenario from the joint distribution is then created by using the i-th outcome from each marginal distribution and its corresponding probability using some transformations described below.

The algorithm proposed by Hoyland *et al.* (2003) is intimately related to articles by Fleishman (1978), Vale & Maurelli (1983), and Lurie & Goldberg (1998).

Fleishman (1978) observed that a real-life distribution is quite often described by the first four statistical moments. He then proposed a procedure for generation of non-normal random numbers with specified first four statistical moments using a liner combination of a random number drawn from a normal distribution, its square and its cube.

Vale & Maurelli (1983) extended the method of Fleishman to generation of multivariate non-normal random numbers that conform to a specified correlation structure of the variables while preserving the target univariate (marginal) means, variances, skews and kurtosis. Their algorithm starts with generation of (correlated) multivariate normal random numbers with certain "intermediate" (specially adjusted) correlation structure. Then cubic transformation (see Appendix B for more details) is used to convert the normal variables of the first step to the non-normal random variables with pre-specified first four statistical moments. The cubic transformation converts the intermediate correlation matrix to the final (target) correlation matrix. Vale & Maurelli also provide a system of non-linear equation for calculation of the intermediate correlation matrix from the required (target) correlation matrix.

The algorithm of Lurie & Goldberg (1998) is similar to that proposed (and outlined below) by Hoyland *et al.* (2003) They also generate random numbers from marginal distributions independently and then combine them into a joint multivariate distribution using some transformations. There are two main differences, however. The first one is in the way both algorithms handle the change in the distribution when the individual marginal distributions are combined into a multivariate distribution. Lurie & Goldberg adjust the intermediate correlation matrix so that after the transformations the (final) multivariate distribution is as desired. Hoyland *et al.* instead start by adjusting the moments of marginal distributions. The other difference is that Lurie & Goldberg require parametric specifications of marginal distributions while Hoyland *et al.* require only the specifications of marginal moments. The latter approach seems to be more flexible since the marginal moments can always be derived from parametric distributions.

The essence of the algorithm is to generate n independent random variables with desired first four statistical moments. Using Cholesky decomposition transforms these random variables to variables satisfying certain correlation structure. This transformation will distort the marginal moments of order higher than 2. Therefore one needs to start with random variables having different marginal moments so that after the transformation the final moments match the (target) desired ones.

The proposed algorithm will produce the desired moments and correlations only if the individual random variables are independent. In order to achieve independence one has to generate very large number of samples and all scenarios should be equally probable. Since the number of generated scenarios is limited and their probabilities could be distinct the method does not lead to exact target marginal moments and correlations. However the algorithm outlined below is used in an iterative fashion in order to achieve errors within a specified interval.

In the rest of this section we use the same notation as that used in Hoyland *et al.* (2003).

Notation

n number of random variables;

s number of scenarios;

$\tilde{X}$ general (continuous or discrete) n-dimensional random vector: $\tilde{X} = (\tilde{X}_1, \tilde{X}_2, \ldots, \tilde{X}_n)$;

X matrix of s scenario outcomes. X has dimension $n \times s$;

X_i row vector of outcomes of the i-th random variable. X_i has size s;

P row vector of scenario probabilities – given by the user;

$\tilde{\chi}$ discrete n-dimensional random variable given by X and P;

$E[\tilde{X}]$ and $E[\tilde{\chi}]$ vector of means of a random variable (general and discrete);

$RV(mom; corr)$ the set of all random variables with moments $mom = mom_1, \ldots, mom_4$ and a correlation matrix $corr$, such that

$$RV(mom; corr) = \begin{pmatrix} mom_{11} & mom_{21} & mom_{31} & mom_{41} & & corr_{11} & \cdots & corr_{1n} \\ \vdots & \vdots & \vdots & \vdots & ; & \vdots & \ddots & \vdots \\ mom_{1n} & mom_{2n} & mom_{3n} & mom_{4n} & & corr_{n1} & \cdots & corr_{nn} \end{pmatrix};$$

$TARMOM$ matrix of target moments $(4 \times n)$;

R target correlation matrix (4×4).

Since the scenario probabilities P are given, generating a discrete random vector $\tilde{\chi}$ is associated with generating a matrix of its outcomes X.

There are two assumptions of the correlation matrix R:

1. R is symmetric positive semi-definite with 1's on the main diagonal. This assumption can be checked by applying Cholesky decomposition to the correlation matrix R. If R is not positive semi-definite then Cholesky decomposition will fail. This in turn will indicate some internal inconsistency in the data.

2. The random variables (asset returns) are not collinear. This condition is attained when the correlation matrix R is positive definite. To check for this condition one can use again Cholesky decomposition, $R = LL^T$. If the random variables are collinear then the lower-triangular matrix L will have zero(s) on its main diagonal.

Algorithm

The goal of the algorithm is to generate scenarios of random vector $\tilde{Z}$. The scenarios are defined by the $n \times s$ matrix of outcomes Z and a vector of probabilities P. The generated scenarios of $\tilde{Z}$ should have target moments equal to $TARMOM$ and correlation matrix equal to R, i.e. $\tilde{z} \in RV(TARMOM;R)$.

The algorithm uses two transformations: cubic transformation and matrix transformation. Cubic transformation is used in order to generate a univariate distribution with specified statistical moments and matrix transformation is used to transform a multivariate distribution in order to obtain required correlation structure of random variables.

Cubic transformation

The cubic transformation can be expressed as:

$$\tilde{Y}_i = a + b\tilde{X}_i + c\tilde{X}_i^2 + d\tilde{X}_i^3 \qquad i = 1,\dots,n, \tag{3.29}$$

where $\tilde{Y}_i$ is a non-normal random vector with specified (target) four statistical moments; $\tilde{X}_i$ is any arbitrary random vector. For more detail on the cubic transformation see Appendix B.

In the context of the algorithm proposed by Hoyland *et al.* the cubic transformation is used to generate discrete approximation $\tilde{y}_i$ of $\tilde{Y}_i$ as shown below:

- Select some arbitrary discrete random variables $\tilde{\chi}_i$ with the same number of outcomes as $\tilde{y}_i$.
- Calculate the first 12 statistical moments of $\tilde{\chi}_i$.
- Calculate the coefficients a,b,c,d of the equation (3.29) using the system of equations described in Appendix B.

- Using the cubic transformation compute the vector of outcomes (scenarios) Y_i for each discrete random variable $\tilde{y}_i$ as $Y_i = a + b X_i + c X_i^2 + d X_i^3$.

Matrix transformation

The matrix transformation (see Appendix C for more details) is applied to a vector of random independent variables, X, to obtain another random vector, Y, with required (target) correlations, R, among the elements of the resulting random vector and can be expressed as follows:

$$Y = LX, \tag{3.30}$$

where L is the lower-triangular matrix obtained from decomposing the correlation matrix such that: $R = LL^T$.

The algorithm consists of two stages: input stage and output stage. During the input stage the desired (target) properties are collected and transformed into a form necessary for the algorithm. During the output stage the distributions (scenarios) are generated and transformed to match the required (target) properties. Thus during the input stage all the transformations are performed on (target) moments and correlations whereas in the output stage the transformations are conducted on the discrete outcomes (realizations of random variables).

The input stage

The input stage can be divided into 3 steps:

1. Specify the target moments $TARMOM$ and target correlation matrix R of $\tilde{Z}$. This information can be obtained from stochastic processes governing the data dynamics of random variables or specified directly by the user.

2. Normalise the original random vector $\tilde{Z}$ such that the resulting normalized random variables in vector $\tilde{Y}$ have zero mean and variance equal 1. This normalization is done for convenience when applying Cholesky decomposition.

3. Lower-triangular matrix, L, obtained from decomposing the correlation matrix such that: $R = LL^T$, is calculated. Also, the moments, $TRSFMOM$, of the normalized independent random variables $\tilde{X}_i$ (for $i = 1,\ldots,n$) are calculated.

The input stage can also be summarized by the following diagram below.

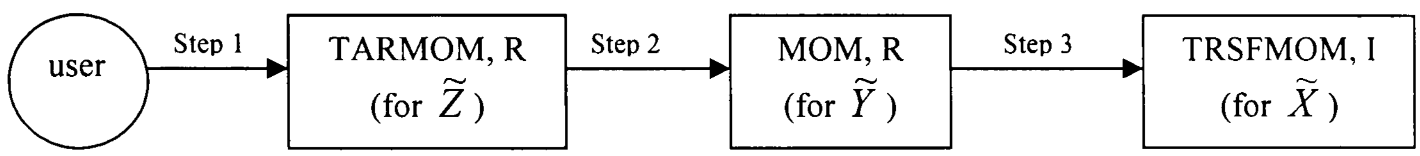

Figure 3-6 Input Stage

The Output Stage

The output stage can be divided into the following 3 steps:

4. Having the target moments, $TRSFMOM$, obtained at step 3, one can generate outcomes of independent random variables $\tilde{\chi}_i$ (for $i = 1,\ldots,n$) by sampling independently from $N(0,1)$. Then applying the cubic transformation the univariate independent random outcomes X_i with desired moments are obtained.

5. Transform $\tilde{\chi}$ to target correlations using the following linear transformation: $Y = LX$, hence $\tilde{y} \in RV(MOM;R)$.

6. Transform $\tilde{y}$ to the original moments by $Z = aY + b$, hence $\tilde{z} \in RV(TARMOM;R)$.

The output stage can also be summarised by the following diagram below.

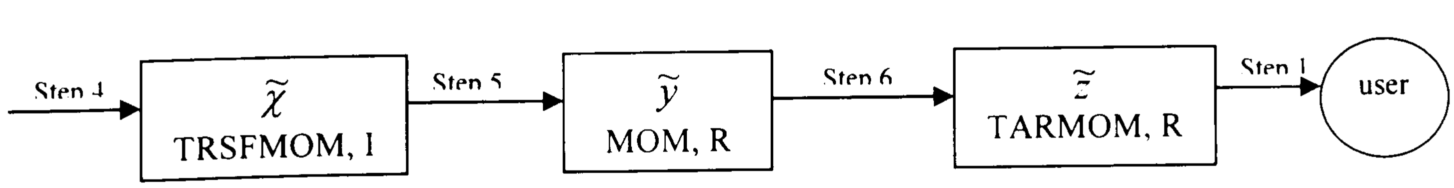

Figure 3-7 Output Stage

The above algorithm produces exact results only if the random variables $\widetilde{\chi}_i$ are independent. Since the finite number of samples were used the resulting correlations are most likely to be different from zero and hence the algorithm will be only approximate. In order to overcome this obstacle Hoyland *et al.* (2003) propose an iterative algorithm, which allows keeping the approximation error within a specified interval.

In step 4 of the algorithm the goal of obtaining an independent random variable is replaced by the goal of obtaining uncorrelated random variables $\widetilde{\chi}_i \in RV(TRSFNMOM; I)$, since independence is very hard to achieve. For this purpose an iterative approach is introduced in step 4 as shown below.

<u>Step 4</u>

4.i. Generate n univariate random variables $\widetilde{\chi}_i$ with moments $TRSFMOM$ (independently). $\widetilde{\chi}_i$ will have correlation matrix R_1 close but not equal to I due to finite samples.

4.ii. *let* $k = 1$ and $\widetilde{\chi}_1 = \widetilde{\chi}$, where k is an iteration counter

4.iii. while $dist(R_k; I) > \varepsilon_x$ do

4.iv. do Cholesky decomposition: $R_k = L_k L_k^T$

4.v. do backward transform $X_k^* = L^{-1} X_k$, which results in $\widetilde{\chi}^*$ having zero correlations and wrong moments

4.vi. do cubic transform of χ_k^* with $TRSFMOM$ as the target moments; store results as χ_{k+1}^*. Now χ_{k+1}^* has the right moments and wrong correlations.

4.vii. Compute correlation matrix R_{k+1}

4.viii. $let\ k = k + 1$

4.ix. $let\ \tilde{\chi} = \tilde{\chi}_k$, where $\tilde{\chi} \in RV(TRSFMOM; R_k)$ with correlation error $dist(R_k; I) \le \varepsilon_x$

In the above iterative procedure $dist(\ )$ represents the root-mean-squared-error. There are two possible outcomes from step 4: random variables $\tilde{\chi}^*_{k-1}$ with zero correlations but with the slightly distorted moments and random variables $\tilde{\chi}_k$ with slightly distorted correlations but with the right moments. If step 5 of the algorithm starts with the former outcome then the iterative procedure is carried out as follows.

Step 5

5.i. Apply matrix transformation as follows: $\tilde{y}_1 = L\ \tilde{\chi}^*$. Here $\tilde{y}_1$ will have both moments and correlations incorrect due to higher co-moments different from zero.

5.ii. $let\ k = 1$ and let R_1 be the correlation matrix of $\tilde{y}_1$

5.iii. while $dist(R_k; R) > \varepsilon_y$ do

5.iv. do Cholesky decomposition: $R_k = L_k L_k^T$

5.v. do backward transform $\mathrm{Y}^*_k = L_k^{-1}\ \mathrm{Y}_k$. Now $\tilde{y}_k$ has zero correlations, incorrect moments

5.vi. do forward transform $\mathrm{Y}^{**}_k = L\ \ \mathrm{Y}^*_k$. Now $\tilde{y}_k$ has correct correlations (R) and incorrect moments

5.vii. do cubic transform of $\tilde{y}^{**}_k$ with MOM as the target moments; store results as $\tilde{y}_{k+1}$. Now $\tilde{y}_{k+1} \in RV(MOM; R_{k+1})$

5.viii. $let\ k = k + 1$

5.ix. *let* $\tilde{y} = \tilde{y}_k$, where $\tilde{y} \in RV(MOM; R_k)$ with correlation error $dist(R_k; R) \leq \varepsilon_y$

3.3.3 Optimal discretisation

Pflug (2001) investigates the problem of approximating continuous-state stochastic processes by scenario trees (discrete-state approximations) in the optimal way, i.e. with the smallest approximation error. By the approximation error he means the difference in the optimal objective values between the original (continuous-state) problem and the optimal objective value when the solution of the discretised (solved on the tree) problem is substituted into the original problem.

Keefer (1994) points out that matching moment of the continuous and discrete probability distribution does not result in a good approximation of objective values. Pflug (2001) and Hochreiter & Pflug (2002) argue that the optimal discretisation could be achieved by minimising a transportation metric.

The main idea of their method can be illustrated on a simple one-stage SP. We start with a few definitions along the lines presented in Pflug (2001). Suppose the SP problem is to minimise

$$F(x) = \int f(x, \xi) dG(\xi) \tag{3.31}$$

$$x \in X$$

where $f(x, \xi)$ is the cost function, G is some continuous distribution function on R and $X \subseteq R$ is the feasible set. Denote by $x^* = \arg\min_x F(x)$ its solution (assume for simplicity it is unique). In order to solve the problem one has to optimise it on a scenario tree, hence approximate the discrete distribution $\tilde{G}$ instead of the true distribution function G, i.e. one has to minimise

$$\tilde{F}(x) = \int f(x, \xi) d\tilde{G}(\xi) \tag{3.32}$$

$$x \in X$$

The approximation error $e\left(F, \tilde{F}\right)$ is defined as the price (measured in the objective function) one has to pay when $\tilde{F}$ is optimised instead of the true F, i.e.,

$$e\left(F, \tilde{F}\right) = F\left(\arg\min_x \tilde{F}(x)\right) - F\left(\arg\min_x F(x)\right) \tag{3.33}$$

Let $L_1(f)$ be the Lipschitz-constant of f, i.e.

$$L_1(f) = \inf\{L : |f(\xi) - f(\zeta)| \leq L|\xi - \zeta| \ for\ all\ \xi, \zeta\} \tag{3.34}$$

The Wasserstien-distance d_1 between G and $\tilde{G}$ is defined as

$$d_1\left(G, \tilde{G}\right) = \sup\left\{\int f(x, \xi) dG(\xi) - \int f(x, \xi) d\tilde{G}(\xi) : L_1(f) \leq 1\right\} \tag{3.35}$$

A transportation metric is defined as the Wasserstein-distance d_1, which in turn is related to the mass transportation problem (see Rachev (1991)).

Suppose that the cost functions $\xi \mapsto f(x, \xi)$ are uniformly Lipschitz, i.e. for all $x \in X$, $L_1(f(x, \cdot)) = \inf\{L : |f(x, \xi) - f(x, \zeta)| \leq L|\xi - \zeta|\} \leq \bar{L}_1$. Then,

$$\sup_x \left|F(x) - \tilde{F}(x)\right| \leq \bar{L}_1 \cdot d_1\left(G, \tilde{G}\right) \tag{3.36}$$

and we see that the original problem of minimising $\sup_x \left|F(x) - \tilde{F}(x)\right|$ can be approximated by the problem of minimising the distance $d_1\left(G, \tilde{G}\right)$, which is equivalent to finding the discrete $\tilde{G}$, which is closest to G in the mass transportation sense.

3.3.4 Clustering algorithm

Clustering algorithms reduce a large number of samples (nodes of the tree) to a smaller number while preserving the statistical distributional properties of the original (larger) sample. For examples of applications of clustering techniques to scenario generation see, for instance, Birge & Mulvey (1996) and Canestrelli & Giove (1999).

In order to take interstage correlations *multi-level clustering scheme* could be used (see Dupacova *et al.* (2000) for a discussion of the subject). The scheme exploits relationships among simulated data paths (scenarios) over the whole time horizon. $\xi(\omega) = \{\xi(\omega_1), \ldots, \xi(\omega_T)\}$, where $\xi(\omega_i)$ represents a random variable at time i. The algorithm could be summarised as follows:

- Evaluate a dissimilarity measure for every pair of scenarios, $\xi(\omega^i), \xi(\omega^j)$, e.g.,

$$d(\xi(\omega^i), \xi(\omega^j)) = \sum_{t=1}^{T} w_t \left\| \xi(\omega_t^i) - \xi(\omega_t^j) \right\|, \tag{3.37}$$

  where $\xi(\omega^i)$ represents a particular scenario over the whole time horizon, $\xi(\omega_t^i)$ is a value of scenario i at time t and w_t are selected non-increasing weights, which allow to give more importance to the differences at the beginning of the sequence.

- Measure of dissimilarity between pairs of scenarios is used in definitions of measures of dissimilarity between clusters (see Hansen & Jaumard (1997) for a discussion of the topic). The result is K_1 clusters, $C_1^1, \ldots, C_1^{K1}$, represented by mean or modal values, $\xi(\tilde{\omega}_1^k)$, of the first components $\xi(\omega_1)$ of scenarios included into cluster C_1^k, where $k = 1, \ldots, K_1$. Probability of $\xi(\tilde{\omega}_1^k)$ equal the sum of probabilities of scenarios $\xi(\omega^i)$ contained in cluster C_1^k, $k = 1, \ldots, K_1$.

- The clustering procedure continues for each cluster C_1^k individually, starting with the second component, $\xi(\omega_2)$ of scenarios included in cluster C_1^k with the first component $\xi(\omega_1)$ replaced by $\xi(\tilde{\omega}_1^k)$ and so on. The required structure of the scenario tree is taken into account.

Gulpinar *et al.* (2004) introduce a *randomised clustering algorithm*, which can be repeated until an acceptable clustering is found. The randomised clustering algorithm can be summarised in the following steps:

Step 1. (Initialisation): Initialise a root node, e.g. a root node relates to the current asset prices.

Step 2. (Simulation): Simulate asset returns over the next time period.

Step 3. (Randomised seeds): Randomly choose a number of distinct scenarios around which to cluster the rest of scenarios. This number should correspond to the number of branches that will emanate from the current node.

Step 4. (Clustering): Group each scenario produced by Step 2 around the scenarios generated in Step 3 so as to minimise the cumulative seed difference in each group. If the resulting groups are not satisfactory Step 3 is repeated.

Step 5. (Centroid selection): In each group of Step 4 find the scenario closest to the centre of the group and designate it as the centroid.

Step 6. Create branches emanating from the current node so that each centroid corresponds to a child node. The probability of the child node is equal to the sum of probabilities of scenarios included in the group.

The two methods to simulate scenarios for subsequent clustering are parallel simulation and sequential simulation, shown in Figure 3-8 and Figure 3-9 respectively.

Parallel Simulation. All scenarios spanning the whole time horizon are generated at the initial time period, t_0. The root node corresponds to the current asset prices. The randomised clustering algorithm is performed by dividing the scenarios into clusters and creating corresponding tree nodes at t_1. The clusters generated at t_1 are divided into sub-clusters at t_2 and corresponding t_2 nodes are created. The process is repeated till leaf nodes are achieved. With this method the total number of scenarios at every time period is the same.

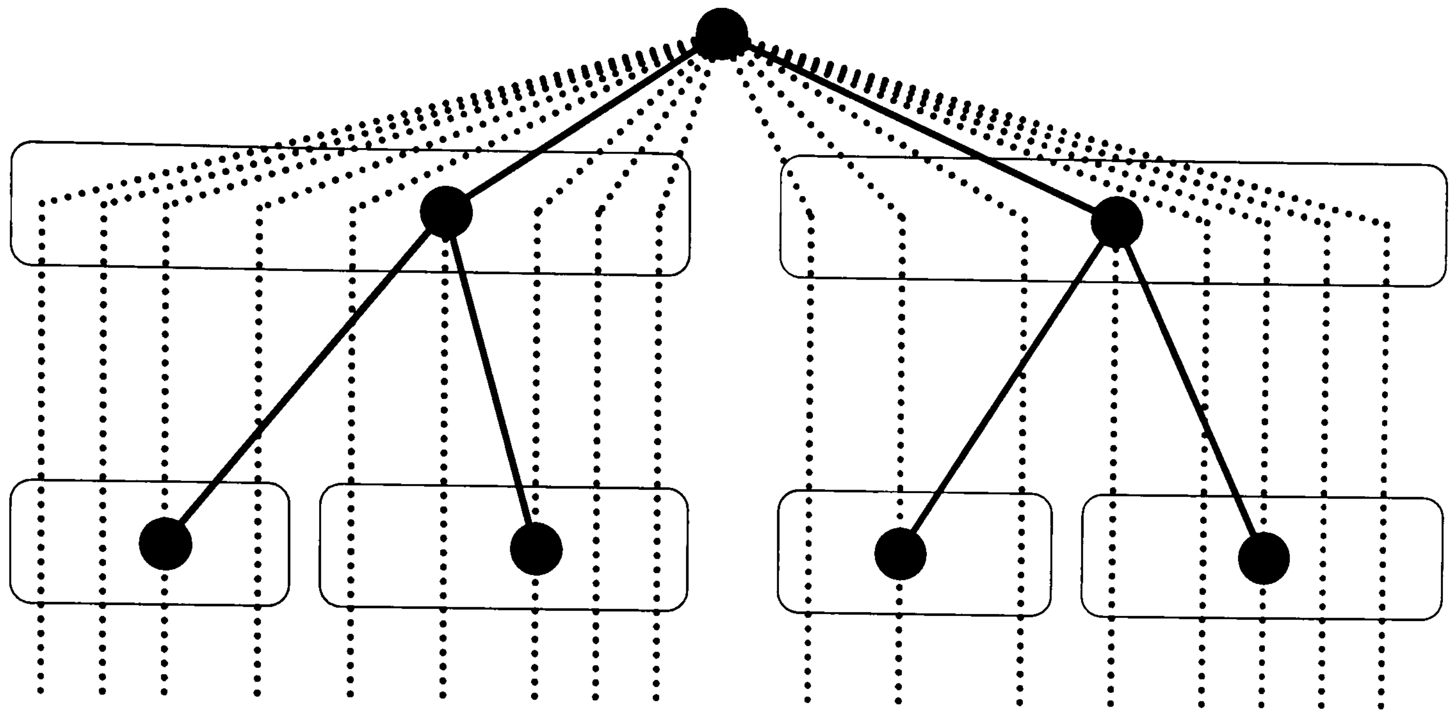

Figure 3-8 Parallel simulation

Sequential simulation. In sequential simulation scenarios, emanating from the root node, are generated only one time period ahead. Then clustering, centroids designation and successor nodes at t_1 selection are performed. In order to create nodes at t_2 a new set of scenarios conditional on nodes at t_1 is generated and the process of clustering and centroids designation is repeated.

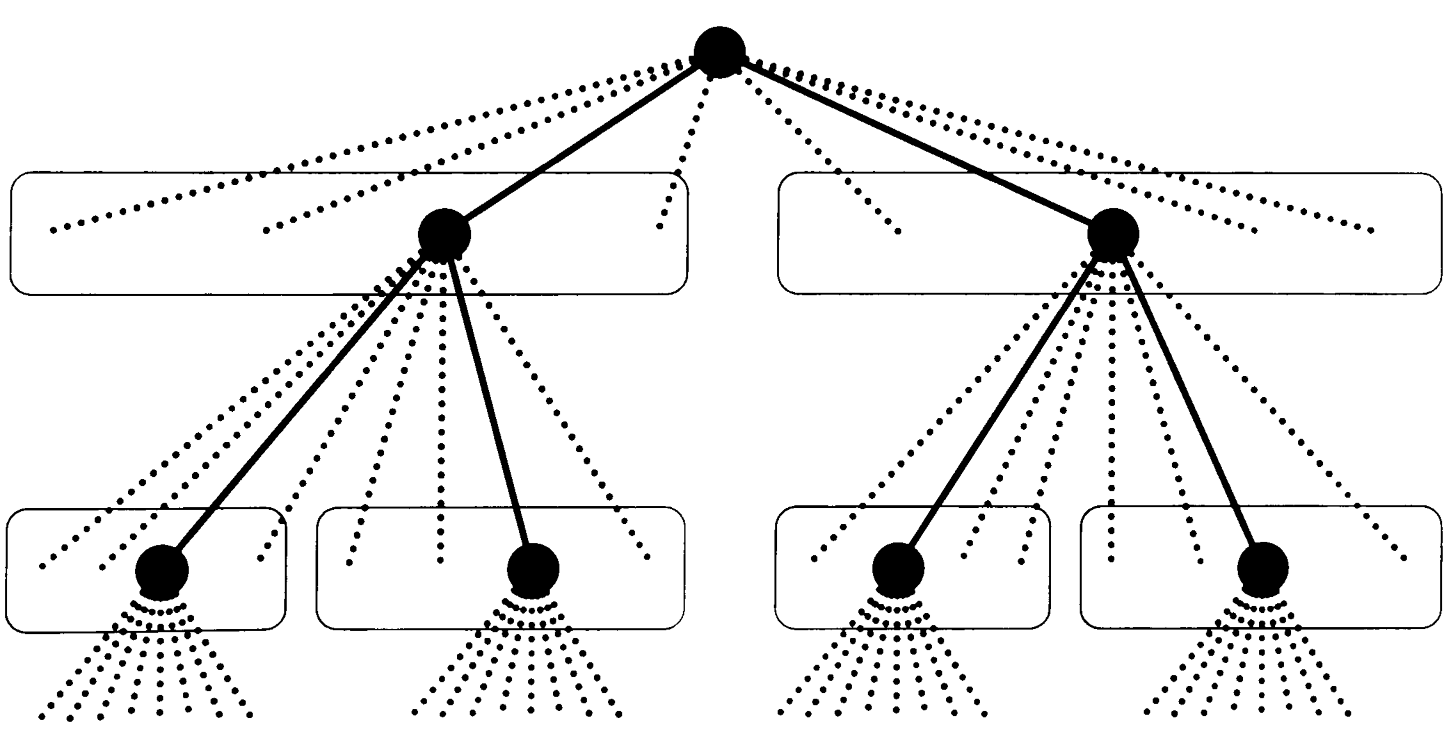

Figure 3-9 Sequential simulation

The main difference in the two methods of scenario simulation stems from the fact that in parallel simulation a large number of scenarios is simulated at a root node. This is necessary since as we move down the tree the clusters become smaller and smaller. With sequential simulation, since the root node scenarios are not "re-used" for subsequent nodes, there is no need to generate more scenarios at the root node that at nodes further down the tree. In sequential simulation it is possible to simulate the same (new) set of scenarios at every node of the tree. Because of this feature of not remembering the original, root node scenario, of the sequential method generally is less memory intensive than the parallel method.

The sequential method would also produce more homogeneous trees since as we move down the tree scenarios simulated at a predecessor node are discarded and only centroid (centre of the cluster) is used to designate the current node. Therefore, successor scenarios are simulated from the centroid and all the extreme scenarios are excluded from the process. Another reason for a less homogeneous tree when using the parallel method is when we move down the tree clusters become smaller and centroids are chosen from smaller sample sizes thus increasing the influences of extreme scenarios. This last feature of the tree can generally be justified in the context of multistage SP problems since more scenarios and hence accuracy is required at the first stage when the actual decisions are made.

3.3.5 Some other methods of scenario generation

In this section we give a brief overview, along the line of Smith (1993), of some additional discretisation (scenario generation) methods such as: "bracket-median" method, "bracket-mean" method and "extended Pearson-Tukey" method. A comprehensive review of a number of discretisation methods can be found in Keefer & Bodily (1983), and Keefer & Verdini (1990).

"Bracket-median" method (see Clemen (1991) for more details) begins by dividing the support of the continuous cumulative probability distribution into N equally-likely intervals. Then each interval is approximated by its median point and the respective probability of such a point is $\frac{1}{N}$.

"Bracket-mean" method is similar to the "Bracket-median" method but instead of using the median value to approximate the interval one has to use the mean value (see McNamee & Celona (1987) for more detailed discussion of the method). For three-point approximations, McNamee & Celona recommend using the interval (of the support of the continuous cumulative distribution function) with probabilities 0.25, 0.50, and 0.25 rather than equally-likely probabilities typically used for "bracket-median" method.

"Extended Pearson-Tukey" (EP-T) method approximates the continuous distribution by a distribution with three discrete points only (see Keefer & Bodily (1983) for more details). It assigns 0.185, 0.630, and 0.185 probabilities to the 5-th, 50-th, and 95-th percentiles of the distribution.

Chapter 4. Desirable Tree Properties

4.1 Measures of quality of a scenario tree

There are a number of criteria that measure the quality of discrete approximations. One of them is accuracy of the approximation. One of the measures of accuracy is the Kolmogorov-Smirnov statistic (see Keefer & Bodily (1983), Miller & Rice (1983), and Smith (1993) for examples). The *Kolmogorov-Smirnov statistic* measures the maximum distance between the true (continuous) distribution, $F(x)$, and the discrete approximation, $F_A(x)$: $\max|F(x)-F_A(x)|$.

Some additional measures of accuracy of a discrete approximation include errors in mean values and errors in higher moments between the true distribution and its discrete approximation.

An alternative measure of accuracy is to look at the errors in the value lottery and certainty equivalent of the discrete approximation. These errors can be assessed by computing the value lottery and certainty equivalent of the decision tree, and comparing them to the value lottery and certainty equivalent of the true distribution. The exact value lottery and certainty equivalent can be assessed by using Monte Carlo simulations.

It should be noted that most of the errors in the discrete approximations could be reduced by taking more points in the discrete approximations. Unfortunately the tree grows very rapidly with the number of the points (successor nodes) in the discrete approximation. For example, if there are 10 random variables then the three-point approximation of each of the variables would yield $3^{10}=59049$ successor nodes. If the distribution of each variable were approximated by 4 discrete points then the tree would have $4^{10}=1048576$ successor nodes, which would exceed the capacity of most decision tree programs.

Kaut & Wallace (2003) expose the characteristics of a "good" scenario-generation method for a given decision model. They specifically make an emphasis on the link between the scenario-generation technique and the decision model at hand and postulate that there is no "universal" scenario-generation method that performs well for all the decision models and therefore the same scenario-generation method may be good for one problem and bad for another. For example, there could be a method that converges to the true distribution when the number of scenarios is increased to "infinity" but at the same time it does not perform well when only a small number of scenarios is to be selected for input into the decision model. On the contrary, there could be a scenario-generation method that does not converge to the true distribution but it performs well in real-life problems.

Let $\tilde{\xi}(\omega)$ define a vector of random variables and $\xi(\omega)$ define a vector of discrete random variables. Then $\{\tilde{\xi}_t(\omega)\}$ will define a stochastic process, where $t \in \{0,1,\ldots,T\}$ points to the time stages in the decision model. Similarly, $\{\xi_t(\omega)\}$ denotes a discrete multivariate stochastic process. When constructing a scenario tree we approximate a "true" stochastic process $\tilde{\xi}(\omega)$ by a discrete stochastic process $\{\xi_t(\omega)\}$. We also define the decision model by

$$\min_{x \in X} F\left(x;\tilde{\xi}_t\right), \tag{4.1}$$

where x is a vector of decision variables, $\tilde{\xi}_t$ is understood as $\{\tilde{\xi}_t(\omega)\}$. When the "true" stochastic process $\{\tilde{\xi}_t(\omega)\}$ is approximated by the scenario tree $\{\xi_t(\omega)\}$ the objective function becomes $F(x;\xi_t)$.

Since the problem cannot be solved directly with $\{\tilde{\xi}_t\}$ we approximate it by the scenario tree $\{\xi_t\}$. Therefore, we will judge the success of the approximation $\{\xi_t\}$ by the quality of decisions the problem solved on $\{\xi_t\}$ gives and not by the "closeness" of the approximation to the "true" stochastic process $\{\tilde{\xi}_t\}$ in the statistical sense.

The error of approximation of $\{\tilde{\xi}_t\}$ by $\{\xi_t\}$ is defined by the difference between the objective functions estimated at optimal solutions of the "true" and the approximated problems:

$$\begin{aligned} e_f(\tilde{\xi}_t, \xi_t) &= F\left(\arg\min_x F(x; \xi_t); \tilde{\xi}_t\right) - F\left(\arg\min_x F(x; \tilde{\xi}_t); \tilde{\xi}_t\right) \\ &= F\left(\arg\min_x F(x; \xi_t); \tilde{\xi}_t\right) - \min_x F(x; \tilde{\xi}_t) \end{aligned} \tag{4.2}$$

In the above definition of the error of approximation $e_f(\tilde{\xi}_t, \xi_t) \geq 0$ since the second term on the right-hand-side of (4.2) is the true minimum of the objective function while the first term in (4.2) is the minimum of the objective function at the approximate solution. Also the definition of the approximation error related only objective functions and not the optimal solutions, x, of the true and the approximated problems. This is due to the fact that objective functions of SP problems are typically flat, that is there could be different solutions with similar objective function values. As such a scenario-generation method may be seen as a heuristic for minimizing the error $e_f(\tilde{\xi}_t, \xi_t)$.

Both terms on the right-hand-side of (4.2) present some problems. The second term is virtually impossible to estimate since it implies the solution of a decision problem on a continuous stochastic process $\{\tilde{\xi}_t\}$. The second term can be estimated for example via simulation.

Kaut & Wallace (2003) propose two characteristics that can be used to qualify a given scenario tree as good for a given decision model. Thus the decision-maker can test different scenario-generation techniques and select the best one for the given model based on the two proposed criteria. The first criterion is the stability condition. Since most of scenario-generation methods involve randomness any two trees (estimated on the same data) should lead to approximately the same optimal objective function values when the optimisation model is solved on those trees. The second condition is that the tree should be unbiased when compared to the true solution.

Stability condition. This condition can be defined as follows. If we approximate the true stochastic process $\{\tilde{\xi}_t\}$ by several scenario trees $\{\xi_t\}$ and solve the SP model on these trees then the optimal objective function values should be approximately the same.

Let us assume we have generated K scenario trees $\{\xi_t\}$ and solved the SP model on each of them. Denote the optimal solutions on each tree as x_k^*, $k = 1,\dots,K$. Then the *in-sample stability* can be defined as:

$$F(x_k^*;\xi_{tk}) \approx F(x_l^*;\xi_{tl}), \qquad k,l \in 1,\dots,K, \tag{4.3}$$

The *out-of-sample stability* can be defined as:

$$F(x_k^*;\tilde{\xi}_t) \approx F(x_l^*;\tilde{\xi}_t) \qquad k,l \in 1,\dots,K. \tag{4.4}$$

We can equivalently represent the above stability conditions as:

in-sample:
$$\min_x F(x;\xi_{tk}) \approx \min_x F(x;\xi_{tl}), \tag{4.5}$$

out-of-sample:
$$F\left(\arg\min_x F(x;\xi_{tk});\tilde{\xi}_t\right) \approx F\left(\arg\min_x F(x;\xi_{tl});\tilde{\xi}_t\right), \tag{4.6}$$

out-of-sample, using equation (4.2):

$$e_f(\tilde{\xi}_t,\xi_{tk}) \approx e_f(\tilde{\xi}_t,\xi_{tl}). \tag{4.7}$$

The difference between the in-sample and out-of-sample stabilities is in the fact that in order to evaluate the in-sample stability one only needs to solve the SP model on several scenario trees whereas in order to evaluate out-of-sample stability one needs to evaluate the "true" objective function $F(x;\tilde{\xi}_t)$, i.e. the objective function evaluated on the true stochastic process $\{\tilde{\xi}_t\}$ and not on the scenario tree $\{\xi_t\}$. In order to do this one needs to know the full knowledge of the distribution of $\{\tilde{\xi}_t\}$.

Let us now look at some of the difference between the two types of stability and compare the importance of each of them. First, let us assume that there is an out-of-sample stability, then the "real" performance of the optimal solution x_i^* obtained on a particular scenario tree $\{\xi_{ti}\}$ does not change much when we re-evaluate the true objective function already for another optimal solution x_j^* obtained from another scenario tree $\{\xi_{tj}\}$. However, if there is no in-sample stability then we will not know whether an optimal solution on a particular tree is good or bad unless there is a possibility to test for an out-of-sample stability. On the other hand, when there is an in-sample stability and no out-of-sample stability the real performance of the solutions will depend on which scenario tree was used to calculate the optimal solutions. At the same time, due to in-sample stability, there is no possibility of learning, which scenario tree will give an out-of-sample stable solution and which not.

If, when solved on several scenario trees, both the in-sample stability in the objective function and in-sample stability in the solutions are achieved than the out-of-sample stability is guaranteed. Thus, when in-sample stability in the objective function is detected one should also check for the stability in the solutions. However, it does not work the other way around. Having out-of-sample stability does not guarantee the in-sample stability in solutions since the objective functions of SP problems are generally flat.

As pointed out by Kaut & Wallace (2003) in most practical situations there are either both stabilities present or none, so the in-sample stability test should be sufficient to check for potential instability. Also, if there is a way to perform out-of-sample stability test one should try to do so.

There are a number of options for out-of-sample testing (evaluation of $F(x_k; \tilde{\xi}_t)$ for a give optimal solution x_k on the k-th scenario tree ξ_k). If the true stochastic process $\{\tilde{\xi}_t\}$ is known then the Monte-Carlo-like simulation can be used to evaluate the true objective function. If the scenario tree was constructed using some historical sample then back-testing can be used to evaluate out-of-sample stability. Also, if there is a stable scenario-generation method, which one can use as a reference than the out-of-sample stability of optimal solutions x_i^* and x_j^* can be evaluated on this reference scenario tree. This reference scenario tree can be quite big since it is not used to solve an SP problem on it but rather to evaluate an objective function on it for a given optimal solution x_i^*.

To summarise, in order to check for the appropriateness of a particular scenario-generation method for a particular optimisation problem one has to test the method (in conjunction with the optimisation problem) for in-sample and if possible for out-of-sample stability.

Unbiasedness condition. In addition to satisfying stability conditions (both in-sample and out-of-sample) the scenario tree should not introduce any bias in the solution of the SP problem. It means that the optimal solution to the SP problem obtained on the tree $\{\xi_t\}$,

$$\breve{x}^* = \arg\min_x F(x;\, \xi_t), \tag{4.8}$$

should be (almost) optimal when the problem is solved for the true stochastic process $\{\tilde{\xi}_t\}$. Therefore the value of the true objective function obtained when the scenario-based solution is substituted $F(\breve{x}^*;\tilde{\xi}_t)$ should be almost equal to the true optimal objective function value $\min_x F(x;\tilde{\xi}_t)$:

$$F(\breve{x}^*;\tilde{\xi}_t) = F\left(\arg\min_x F(x;\xi_t);\tilde{\xi}_t\right) \approx \min_x F(x;\tilde{\xi}_t). \tag{4.9}$$

Using the definition of approximation error in (4.2) the unbiasedness condition can be formulated as:

$$e_f(\tilde{\xi}_t, \xi_t) \approx 0. \tag{4.10}$$

It is practically impossible to test the unbiasedness condition since it required the solution of the optimisation problem using the true (continuous) stochastic process $\{\tilde{\xi}_t\}$. At the same time if one could solve the SP problem using the true stochastic process then there would be no need to approximate the true process $\{\tilde{\xi}_t\}$ with the scenario tree $\{\xi_t\}$.

In certain situation one can use the *reference tree*, which is known to be unbiased, in order to test for existence of a bias in the tree one is testing. Such a reference tree is built to approximate the true stochastic process $\{\tilde{\xi}_t\}$ and should be as big as possible. Then as before in order to test for a bias one compares the objective function of a problem $F(\breve{x}^*;\tilde{\xi}_t)$ using scenario-based solution $\breve{x}^*$ and the minimum objective function of a problem $\min_x F(x;\tilde{\xi}_t)$, with $\tilde{\xi}_t$ representing a reference tree.

Most of the time the reasons of instability or biasedness lie in the type of the scenario-generation method used. For example, for sampling methods the instability or bias is most likely caused by the insufficient number of scenarios used. As we know, increasing the number of scenarios improves convergence of the scenario tree to the true (continuous) distribution stochastic process. Therefore individual scenario trees with many scenarios will be closer to the true distribution and hence closer to each other and thus reducing the instability and the bias. Furthermore, as discussed above, improving the sampling method may also improve the quality of the scenario tree.

If the MM methods are used for scenario-generation then increasing the number of scenarios will not help since these methods do not typically guarantee convergence. If we assume that the required statistical properties (e.g. statistical moments and co-moments) have been matched than the instability and bias may be caused by other statistical properties, which were not required to be matched when constructing the scenario tree. In this context the first thing to check is still the adequate number of scenarios For example, both a one-period scenario tree with just three successor nodes and a one-period scenario tree with a thousand successor nodes may match the required five statistical moments and the correlations but rank differently in terms of stability and biasedness. As pointed out by Kaut & Wallace (2003) this is caused by the differences in smoothness of the distribution. However, it is important to understand that not all MM methods show increase in smoothness as a result of in crease in the number of scenarios. While this is normally the case for transformation-based MM methods it is not true for the optimisation-based methods (see the section on the transformation-based and optimisation-based MM methods above for a discussion of these methods).

Another important factor affecting the quality of MM methods is whether one matches the right properties. In some circumstances including higher statistical moments and co-moments in the required properties to match might improve the scenario tree.

4.2 Arbitrage-free Condition

4.2.1 Arbitrage and its implications

When generating scenarios of portfolios of many financial assets such as bonds, financial derivatives and term structures of interest and exchange rates it is important to satisfy a no-arbitrage property. Existence of arbitrage opportunities is intrinsically related to existence of *replicating* strategies. In essence it means that if a security's payoff at the end of the time period can be replicated by holding a portfolio of some assets then the price of the security and that of the portfolio at the beginning of the time period should be equal. An example of the existence of arbitrage opportunity is a violation of a put-call parity relation. If the parity is violated then one can construct a portfolio of a stock and a call option with the same end of time period payoffs as a portfolio of a bond and a put option (with the same strike as the call) but with different beginning of the period values. In order to exploit the arbitrage opportunity an investor can short-sell the more expensive portfolio, buy the cheaper portfolio and thus generate immediate profit not facing any future risk.

An *arbitrage* is a trading strategy that

- has a positive initial cash flow and no risk of a loss in the future (*type-A* arbitrage), or
- has no initial cash investment, has no risk of a loss, and has a positive probability of making profit in the future (*type-B* arbitrage).

If the scenario tree provides an arbitrage opportunity, i.e. availability of riskless profit, then the optimisation model will exploit it resulting in increased objective function value. It is often argued that the models allowing arbitrage opportunities are not realistic. This argument is questionable since in real world arbitrage opportunities do exist and are exploited. In many cases these opportunities are limited by transaction costs, size of a line of credit and exploiting arbitrage drives the prices in the direction, which eliminates the arbitrage opportunities. These issues suggest the ways to measure arbitrage and the ways to design arbitrage-free scenario tree as discussed below. Therefore it is prudent to build a scenario tree, which does not allow arbitrage.

One situation where arbitrage opportunities can arise is due to approximation errors. Klaassen (1997) was the first who analysed this problem. This problem arises when a small number of scenarios (nodes) of asset returns is sampled to approximate continuous probability distribution of future returns. The resulting approximation is usually poor and as a result it may not comply with the distribution implied with the currently observed market prices (values) of various financial assets, hence providing spurious arbitrage opportunities.

When scenarios are generated only for broad asset classes, such as bond index, stock index, real estate index then approximation errors are unlikely to occur. However, if scenarios are generated for derivative securities such as options, interest rate derivatives etc., then approximation errors may result in arbitrage opportunities thus creating spurious profit opportunities for the SP optimisation model.

4.2.2 Arbitrage-free scenario trees

Thorlacius (1998) investigates the situations where arbitrage arises and gives some remedies to prevent it. First, he considers the "string" scenario structure, which was depicted in Figure 3-1 of section 3.3, where branching occurs only at the root node. In this tree structure knowledge of the state after the initial time period determines which scenario will persist till the final time period. As a result, since there is no branching, all the uncertainty is resolved. Hence an investor can use the perfect foresight to construct an arbitrage trading strategy. Thorlacius postulates that in order to avoid arbitrage opportunities in a string tree structure all assets should have identical returns after the initial time period for every scenario.

One way to tackle the arbitrage problem with the string scenarios is to introduce branching at every node in the tree so that the investor does not know with certainty which state will occur by the end of the next time period. This effectively creates a tree-like scenario structure as was depicted in Figure 3-2 of section 3.3.

Harrison & Kreps (1979) prove the necessary and sufficient condition for the absence of arbitrage opportunities:

Theorem 1. Suppose that there are (n + 1) primitive securities traded in a one-time period market model. Then, there are no arbitrage opportunities if there exist a strictly positive probability measure $p(\omega)$, such that:

$$\sum_{\omega\in\Omega} p(\omega)\frac{S_1^i(\omega)+D_1^i(\omega)}{S_0^i} = \sum_{\omega\in\Omega} p(\omega)\frac{S_1^0(\omega)+D_1^0(\omega)}{S_0^0}, \quad i = 0, 1, \dots, n, \qquad (4.11)$$

where S_0^i represents the initial price of the i-th asset, $S_1^i(\omega)$ and $D_1^i(\omega)$ denote respectively the final price and payoff (e.g. dividend) of the i-th asset, for i = 0, 1,..., n and $\omega \in \Omega$, where Ω is the scenario set, characterising all the tree nodes at the end of the time period. S_0^0 and S_1^0 represent initial and final prices for a risk-free asset.

If there is a traded riskless asset in the model then equation (4.11) can be reformulated as:

$$S_0^i = e^{-r} \sum_{\omega \in \Omega} p(\omega)\left(S_1^i(\omega) + D_1^i(\omega)\right), \text{ for all i = 0, 1,..., n.} \tag{4.12}$$

We have to solve a system of n equations, one for each asset. For this system to have a solution and preserve linear independence of asset returns, there should be at least (n + 1) unknowns, $p(\omega)$, which means there should be at least (n + 1) scenarios and hence leaf nodes at the end of the time period. The implication for the arbitrage-free scenario tree construction is the following. Each node of the tree should have at least n successor nodes, otherwise two things happen:

1. If the number of scenarios, $m < (n+1)$ and the system (4.12) is solvable then some of the asset returns are linearly dependent on others. This is due to the fact that the rank of the coefficient matrix in the system of (n + 1) linear equations (4.12) is less than (n + 1).

2. If the number of scenarios, $m < (n+1)$ and the system (4.12) is not solvable then some of the (n + 1 – m) redundant asset returns cannot be replicated by the first m asset. This means that there are some assets that have the same final returns but different initial prices, therefore creating an arbitrage opportunity.

Kouwenberg & Vorst (1998) apply the MM method to build a multiperiod arbitrage-free scenario tree for joint evolution of asset returns, economic variables and contingent claims. They analyse the conditions, resulting in arbitrage opportunities on the tree. As pointed out in their paper, in order to track down an arbitrage opportunity one has to analyse the dual of the linear system (the primal problem includes equations (4.11) or (4.12) as constraints). See Pliska 1997, p.7, or Ingersoll 1987, p.55 for an application of LP duality theory to verify the arbitrage-free condition. The dual formulation will pinpoint the portfolio that violates the no-arbitrage condition. By studying the cause of the infeasibility of the dual solution one could get a clue as to how to eliminate it. Some other remedies of arbitrage could be to reduce the amount of moment matching constraints in the model or to increase the amount of successor nodes.

Now, we prove Theorem 1 using LP duality principals. This proof, which follows below, is taken from Cornuejols and Tutuncu (2005).

Proof of Theorem 1: We assume that the state space $\Omega = \{\omega_1, \omega_2, \ldots, \omega_m\}$ is finite. S_0^i represents the current price of asset i and $S_1^i(\omega_j)$ represents the future price of asset i if scenario ω_j is realised, where i = 0 to n and j = 1 to m. Next, consider the following linear programming (LP) problem with variable x_i denoting the holding in asset i:

$$\min_x \quad \sum_{i=0}^{n} S_0^i x_i \qquad (4.13)$$

$$\sum_{i=0}^{n} S_1^i(\omega_j) x_i \geq 0, \quad j = 1, \ldots, m.$$

$$x_i \geq 0, \quad i = 1, \ldots, n.$$

The formulation in (4.13) can be used to check for existence of type A-arbitrage, which would correspond to a feasible solution of this LP problem with a negative objective value. Since $x_i \equiv 0$ is always a feasible solution for this problem, the optimal objective value is always non-positive. Moreover, since all the constraints are homogeneous, if there exists a feasible solution such that $S_0^i x_i < 0$ (this indicated type-A arbitrage), the problem is unbounded. Therefore, there is no type-A arbitrage if and only if the optimal objective value of (4.13) is 0.

Assume that there is no type-A arbitrage. Then, there is no type-B arbitrage if and only if all constraints are tight for all optimal solutions of (4.13), where these optimal solutions must have objective value 0. The dual of (4.13) is:

$$\max_p \sum_{j=1}^{m} 0 p_j$$

$$\sum_{j=1}^{m} S_1^i(\omega_j) p_j = S_0^i, \quad i = 0,\dots,n, \tag{4.14}$$

$$p_j \geq 0, \quad j = 1,\dots,m$$

Since the dual has a constant (0) objective function, any dual feasible solution will also be dual optimal. As stated above, when there is no type-A arbitrage (4.13) has an optimal solution and it, in turn, implies due to the Strong Duality Theorem (see, for example, Chvatal (1980) for more details) that the dual must have a feasible solution. If there is no type-B arbitrage also, Goldman & Tucker's theorem indicates that, there exists a feasible (and therefore optimal) dual solution p* such that p* > 0 (from strict complementarity with tight primal constraints $\sum_{i=0}^{n} S_1^i(\omega_j) x_i \geq 0$). From the dual constraint corresponding to i = 0, it follows that $\sum_{j=1}^{m} p_j^* = \frac{1}{R}$, where R is a return on a riskless asset. When both sides of this equation are multiplied by R one obtains $\sum_{j=1}^{m} p_j^* R = 1$. Each term in the summation on the left-hand-side of this equation represents a risk-neutral probability. Therefore, no arbitrage assumption implies the existence of risk-neutral probabilities.

The converse direction, i.e. existence of risk-neutral probabilities implies no arbitrage, can be proved in a similar manner. It follows that the existence of risk-neutral probabilities implies that (4.14) is feasible and it in turn implies that its dual (4.13) is bounded. Boundedness of (4.13) indicates that there is no type-A arbitrage. Moreover, strictly feasible (and therefore optimal) solutions of the dual problem (4.14) imply that the constraints are tight for any optimal solution of the primal problem (4.13), which indicated that there is no type-B arbitrage.

Naik (1995) adapts Theorem 1 to the market with bid and ask prices (transaction costs).

Theorem 2. Suppose P_{0i}^{a} and P_{0i}^{b} are the initial bid and ask prices for asset i, such that $P_{0i}^{b} < P_{0i}^{a}$. Then, there are no arbitrage opportunities if there exist a strictly positive probability measure $p(\omega)$, such that:

$$P_{0i}^{b} \le e^{-r} \sum_{\omega \in \Omega} p(\omega)\left(P(\omega)_{1i} + D(\omega)_{1i}\right) \le P_{0i}^{a}, \text{ for all i} = 1, 2, \ldots, \text{n}. \qquad (4.15)$$

Once the scenario tree is built, in order to check it for existence of arbitrage one has to solve equations (4.12) or (4.15) at each node of the tree. If the solution exists, i.e. there is a strictly positive probability measure $p(\omega)$ at every node of the tree, then the tree is arbitrage-free.

When constructing a tree using, for example, a MM method one can include equations (4.12) or (4.15) as constraints in order to satisfy no-arbitrage condition. The resulting optimal solution will represent "regular" probabilities, asset returns as well as risk-neutral probabilities since the equations (4.12) and (4.15) enforce risk-neutral probability measure.

One can use the LP formulations (4.13) and (4.14) to detect the assets, which create arbitrage opportunities where the state space is *finite*. The same concept can be extended to the *infinite* state spaces. Moving from finite to infinite state spaces we have to restrict the scope of possible investing assets to derivative securities on the same underlying security with the same maturity date. The discussion below is based on Herzel (2000).

Let S_0 and S_1 denote the time 0 (current) and time 1 (random) prices of underlying security respectively. We assume that there are n derivative securities, all with maturity at time 1, written on this underlying. The payoff function, $\psi_i(S_1)$, of each of the i-th derivative security is piecewise linear with the breakpoint corresponding to the strike price K_i. For example, if the i-th derivative security is a European call option then $\psi_i(S_1) = (S_1 - K_i)^+$. Let S_0^i denote the current price of the i-th derivative security. We consider a portfolio $x = (x_1, \ldots, x_n)$ of derivative securities 1 to n. Then the payoff function on a portfolio is:

$$\psi^x(S_1) = \sum_{i=1}^n \psi_i(S_1)x_i \,, \tag{4.16}$$

and the cost of constructing such a portfolio is:

$$\sum_{i=1}^n S_0^i x_i \tag{4.17}$$

To check for existence of arbitrage in the current prices of derivative securities, S_0^i, we follow the same logic as we did when checking for arbitrage in security prices on a finite state space. We minimise the cost of constructing a portfolio of derivatives 1 to n while ensuring that the payoff function $\psi^x(S_1)$ is nonnegative for all $S_1 \in [0,\infty)$. It can be seen that underlying security and hence the derivative securities are defined on a continuous interval (infinite state space) in contrast to the finite discrete state space, $\Omega = \{\omega_1, \omega_2, \ldots, \omega_m\}$, considered earlier. If the optimal objective function is negative then, by definition, we have type-A arbitrage.

Since every payoff function $\psi_i(S_1)$ is piecewise linear then the payoff function on the portfolio, $\psi^x(S_1)$ is linear too with breakpoints in K_1 through K_n. The piecewise linear function $\psi^x(S_1)$ is nonnegative over any $S_1 \in [0,\infty)$ if and only if it is nonnegative at 0, all the breakpoints and its slope to the right of the highest breakpoint, K_n, is nonnegative. These conditions can be formulated as follows:

1. $\psi^x(0) \geq 0$,
2. $\psi^x(K_j) \geq 0, \quad \forall j$,
3. $\left[(\psi^x)'_+(K_n)\right] \geq 0$.

Therefore, in order to check for type-A arbitrage opportunities among derivative securities defined on an infinite state space one has to solve the following LP problem.

$$\min_x \quad \sum_{i=1}^{n} S_0^i x_i \tag{4.18}$$

$$\sum_{i=1}^{n} \psi_i(0) x_i \geq 0 \tag{4.19}$$

$$\sum_{i=1}^{n} \psi_i(K_j) x_i \geq 0, \quad j = 1, \ldots, n \tag{4.20}$$

$$\sum_{i=1}^{n} (\psi_i(K_n + 1) - \psi_i(K_n)) x_i \geq 0, \tag{4.21}$$

where $\psi_i(K_n + 1) - \psi_i(K_n)$ denotes right-derivative of $\psi_i(S_1)$ at K_n since all $\psi_i(S_1)$ are piecewise linear. The left-hand-side of constraint (4.21) denotes the right-derivative of $\psi^x(S_1)$ at K_n.

Proposition 1.1 *If the optimal objective function (4.18) is negative then there is type-A arbitrage in the current prices of derivative securities,* S_0^i.

The proposition below can be proved in a similar way as it was done for finite state space.

Proposition 1.2 *Assume that there is no type-A arbitrage in the current derivative securities prices,* S_0^i. *Then, there is no type-B arbitrage if and only if the dual of the problem (4.18)-(4.21) has a strictly feasible solution.*

Using the earlier results the following theorem can be proved. It provides the necessary and sufficient conditions to prevent arbitrage opportunities in the prices of a set of call options.

Theorem 1.2 *Assume,* $K_1 < K_2 < \cdots < K_n$ *denote strike prices on the set of* n *European call options with the same maturity written on the same underlying security. There are no arbitrage opportunities if the current option prices* S_0^i *satisfy the following conditions:*

$S_0^i > 0, \quad i = 1, \ldots, n$

$S_0^i > S_0^{i+1}, \quad i = 1, \ldots, n-1$

The function $C(K_i) = S_0^i$ *defined on the set* $\{K_1, K_2, \ldots, K_n\}$ *is a strictly convex function.*

Thorlacius (1998) also addresses the problem of measuring the arbitrage. The first measure is the level of uniform (the same percentage for every trade) transaction costs. We keep increasing the level of transaction costs until all arbitrage in the model is eliminated. The higher this level the larger will be the arbitrage gains.

The second measure of arbitrage is related to short-selling, since short-selling is required to exploit arbitrage opportunity. We can set certain short-selling limit and search for a strategy, which realises the most riskless profits. The greater the amount of this profits the greater the arbitrage. On the other hand, assume short-selling of any asset is limited to 100%, then if the additional gains from exploiting arbitrage are relatively small then it is reasonable to assume the arbitrage does not bias the results.

Berkelaar *et al.* (1999) investigates the effect of sampling uncertainty on the existence of arbitrage. He finds that even when the number of successor nodes is bigger that the number of assets but when random sampling is used to generate scenarios there could be arbitrage opportunities. Generally the number of arbitrage opportunities declines with increase in the amount of samples drawn. Berkelaar shows that if for a one-stage tree of 8 years the number of scenarios is reduced from 625 to 25 the probability of generating a tree with arbitrage increases from 0% to 2.9% as expected.

Berkelaar also finds that the risk of having arbitrage in the tree is increased if the number of decision stages in the tree is increased while holding the planning horizon and total number of scenarios fixed. For instance, he reports no arbitrage in a one-stage tree of 8 years when 625 scenarios were randomly sampled. At the same time he finds that the tree consisting of two stages, 4 years each, where each node has 25 successors (and hence the tree has 625 leaf nodes) there is a probability of 29.2% of sampling a tree with arbitrage opportunities. Nevertheless, the percentage of arbitrage of 29.2% for a two-stage tree is lower than that if the two-stage tree were decomposed into 26 one-stage (4years each) problems. One would expect the probability of arbitrage of about $1-(1-0.029)^{26} = 1-0.971^{26} \approx 53\% > 29.2\%$. The difference is due to the time length, i.e. probability of arbitrage of 2.9% was calculated for a one-stage problem where the stage consists of 8 years whereas the 26 one-stage problems use a 4-year stage. Therefore, keeping the same number of successor nodes and decreasing the time length of a stage will decrease the probability of arbitrage in the tree.

When a binomial lattice is used as a scenario tree, the number of scenarios grows exponentially (e^T) with the amount of time periods, hence rendering the problem computationally intractable. Klaassen (1998) proposes an aggregation method for lattice trees. He uses example of a fine-grained risk-neutral binomial lattice of the bond prices. Such a lattice will represent a recombining scenario tree with multitude of short time periods, which will make it unsuitable for most SP problems. The proposed aggregation method reduces the amount of scenarios as well as the amount of decision stages in the model, while keeping the tree arbitrage-free. A drawback of the method of Klaassen (1998) is that it ignores the "real-world" distributions (probabilities) of asset returns. Also it is not at all trivial to construct a risk-neutral scenario tree of multiple asset returns, which are correlated with (non-traded) random economic variables.

Gondzio, Kouwenberg & Vorst (2003) propose another aggregation method for fine-grained grids (trees). The essence of the method is first to partition the grid into groups at a small number of dates, which correspond to decision stages in SP model. Each partition of points on the grid is represented by a single aggregated node. An asset price corresponding to an aggregated node should equal the conditional expected value of asset prices on the grid included in that partition under risk-neutral probability measure. This way the aggregated scenario tree will be arbitrage-free.

Chapter 5. Case Study

5.1 The Problem Setting

In this chapter we use the case study in order to illustrate a particular application of SP framework to the problem of minimising volatility (risk) of future foreign currency revenues. The problem setting in this section is similar to that used in Volosov *et al.* (2005). The situation naturally arises when a company expects to receive or pay some uncertain amount of funds in foreign currency at a future date.

It should be noted that minimising variance of the "portfolio" of accounts receivables and FWD could produce sub-optimal FWD positions. It can be explained as follows: if there is no "maximising return" driver incorporated into the objective function there may be scenarios where the model "throws away" money in order to minimize the variance. The proposed objective function could be improved by applying the notion of utility function, see Luenberger (1998). One possibility is to subtract from the variance of the year-end cash flows the expectation of the year-end cash flows times the constant (coefficient of risk aversion), thus approximating the objective function by the quadratic utility function. The drawback of this quadratic utility function is that it is not monotonically decreasing. Some alternatives may include exponential or power utility functions.

Another issue worth mentioning is the choice of variance as the measure of risk. If the decision-maker is only interested in minimising losses (returns below the pre-specified level) then asymmetric risk measures such as Expected Downside Risk (Domar and Musgrave, 1944), Value at Risk (VaR) (JP Morgan, 1996), Conditional Value at Risk (CVaR) (Uryasev and Rockafellar, 1999) among others would be more appropriate.

In this case study we consider the case where a MNF is expecting USD inflows in 3, 6, 9 and 12 months from the current time period. Being a British company it reports its revenues in GBP. The uncertainty in the amount of future GBP revenues stems from two sources: 1) the MNF does not know exactly how much USD it will receive in the future and 2) the future USD to GBP exchange rate is uncertain at present. The current strategy of the MNF is to convert received USD revenues into GBP using the available spot rate. Although the spot rate is uncertain for future time periods the MNF has not engaged in using FWD.

The problem under investigation is to determine the strategy for employing FWD to minimise the volatility of the GBP-converted revenues received over the next 12 months. In this case the MNF can use FWD for two purposes: 1) to hedge against fluctuations in the spot rate between USD and GBP and 2) provided there is a non-zero correlation between future USD revenues and future forward exchange rates, FWD can be used as any other asset in a portfolio allocation problem to diversify portfolio risk. Thus, taking the structure of correlations among future USD revenues and forward exchange rates FWD could be used to minimize the overall volatility of the "portfolio" consisting of future revenue stream and FWD. An extreme case of this "portfolio"- type strategy would be to use FWD in the portfolio with the GBP revenues. Here there is no inherent FX risk involved, but by exploiting the correlations among GBP revenues and FWD, the portfolio may have smaller volatility than each of the assets individually.

In what follows we wish to determine a policy, which would minimise volatility of the GBP-converted revenue over the next 12 months by allowing the MNF to engage in FWD. Given the inherent risks in speculative trading in FX we include limits to reduce the risks of speculation on forward exchange rates.

5.2 Scenario Tree Generation

The uncertainties involving forward and spot exchange rates and future USD revenues have been modelled as a discrete set of scenarios. Conditional distribution of exchange rates can be modelled using various statistical techniques such as vector error correction model (VECM), which we used in Volosov *et al.* (2005). Modelling future revenue stream is generally company-specific. In order to keep most of the emphasis on the method of building a scenario tree with required properties we use very simple statistical models to describe evolution of exchange rates and future revenues. Obviously, more sophisticated models should be used in practice. The two main assumptions used here are: 1) exchange rates follow a Random Walk process with drift and 2) Only cash flows in 3 months are assumed to be random and independent of exchange rates; cash flows in 6, 9 and 12 months from the current time period are assumed to be deterministic and equal to $1,000,000 per month. This assumption could be partially justified by noting that the MNF would be more concerned with modelling uncertainty of more immediate parameters, i.e. cash flows that are 3 months away. These 3 months away cash flows have a normal distribution with mean $ 1,000,000 and standard deviation $ 150,000, that is USD revenues $\sim N(1{,}000{,}000; 150{,}000)$.

The exchange rates data set used in this case study consists of 21 quarterly observations of the spot and forward GBP/USD exchange rate from DATASTREAM for the period from January 1998 to January 2003. Exchange rates used and their first differences can be found in Appendix D.

The scenario tree is constructed in two steps: 1) we build the scenario tree of spot and forward exchange rates and 2) at each node of the tree for exchange rates 3 months into the future from the current time period we take two random samples from the probability distribution of USD revenues. At all the tree nodes 6, 9 and 12 months ahead USD revenues are assumed constant at $1,000,000. The justification for this comes from the earlier assumption of independence of probability distribution of revenues from that of exchange rates. If the assumption of independence of distributions is removed then revenue realisations should be sampled from conditional probability distribution conditioned on the realisations of exchange rates at each node of the scenario tree. In real-life commercial applications the number of scenarios should be considerably higher in order to give a good approximations of the underlying continuous distribution (see section 3.2. for a discussion of drawbacks of small samples).

Recent empirical research (see, for example, Topaloglou *et al.* 2004(b)) suggests that monthly (and quarterly) exchange rates are skewed, exhibit excess kurtosis and fail Jacque-Bera null hypothesis of normality. These results motivate generation of scenarios based on the empirical distribution and not imposing any specific distributional form on random variables. In this context MM method renders itself as a natural candidate.

The MM method, which was discussed in section 3.3.1, is used to construct the scenario tree[1]. Since for the problem at hand decisions are taken and random parameters are observed at four points in time, i.e. in 3, 6, 9 and 12 months into the future, the four-period (quarters) scenario tree is needed.

In this case study we use an optimisation model in order to generate a scenario tree, which probability distribution will have statistical moments as close as possible to the target or desired moments. For the purposes of this study we use the first four statistical moments and co-moments of random variables as a criteria for distribution fitting. Besides, at each node other than the root node the first two conditional statistical moments are being matched to the targets. Thus, the objective of moments matching here is two-fold: 1) to match the first four unconditional moments and co-moments and 2) to match the first two conditional moments at each node. These two sub-objectives are placed together in the objective function and hence a weighting coefficient could be attached to each of the sub-objectives in order to set the relative importance of each particular sub-objective.

Matching the first two conditional moments is administered as follows: because of the earlier assumption that exchange rates follow a random walk process the mean value of the successor nodes should equal the exchange rate at the current node. Also, in order to simplify calculation we assumed constant volatility (variances) of exchange rate at all nodes of the tree. This volatility equals the historical variance. Both assumptions could be easily relaxed when necessary.

The optimisation model in AMPL format (see Fourer, Gay and Kernighan (2003) for more details on AMPL modelling language) used sequentially to generate scenarios in 3, 6, 9 and 12 months from the initial time period is depicted in Figure 5-1. In what follows we will use AMPL notation since it is more intuitive and the notation

[1] A possible suitable alternative to MM could be, for example, a parsimonious vector autoregressive model, say, of lag one, VAR(1).

becomes self-explanatory. The algebraic formulation of the model is presented in Appendix E.

```
### SETS          ###
set predecessors ordered;
set scenarios ordered;
set assets ordered;
set links := {a1 in assets, a2 in assets: ord(a1) < ord(a2)};
### PARAMETERS       ###
param T;
param mean_tar {assets};
param variance_tar {assets};
param skewness_tar {assets};
param kurtosis_tar {assets};
param covariance_tar {links};
param Prob_ind;
param Prob ;
param mean_tar_ind {assets, predecessors};
param variance_tar_ind {assets, predecessors};
### VARIABLES       ###
var x{assets, predecessors, scenarios};
var mean {a1 in assets} = sum{p in predecessors, s in scenarios} x[a1, p, s] * Prob;
var variance {a1 in assets} =
        sum{p in predecessors, s in scenarios} ((x[a1, p, s]-mean[a1])^2)*Prob;
var skewness {a1 in assets} =
        sum{p in predecessors, s in scenarios} ((x[a1, p, s]-mean[a1])^3)*Prob;
var kurtosis {a1 in assets} =
```

```
        sum{p in predecessors, s in scenarios} ((x[a1, p, s]-mean[a1])^4)*Prob;

var covariance {(a1, a2) in links} =

        sum{p in predecessors, s in scenarios} (x[a1, p, s]-mean[a1])*(x[a2, p, s]-
mean[a2])*Prob;

var mean_ind {a1 in assets, p in predecessors} = sum{s in scenarios} x[a1, p, s] * Prob_ind;

var variance_ind {a1 in assets, p in predecessors} =

        sum{s in scenarios} ((x[a1, p, s] - mean_ind[a1, p])^2) * Prob_ind;

### OBJECTIVES      ###

minimize Diff_MomCov_initial :

        sum {a1 in assets} (mean[a1] - mean_tar[a1])^2

        + sum{a1 in assets} (variance[a1] - variance_tar[a1])^2

        + sum {a1 in assets} (skewness[a1] - skewness_tar[a1])^2

        + sum {a1 in assets} (kurtosis[a1] - kurtosis_tar[a1])^2

        + sum {(a1, a2) in links} (covariance[a1, a2] - covariance_tar[a1, a2])^2 ;


minimize Diff_MomCov_ind :

        sum {a1 in assets} (mean[a1] - mean_tar[a1])^2

        + sum{a1 in assets} (variance[a1] - variance_tar[a1])^2

        + sum {a1 in assets} (skewness[a1] - skewness_tar[a1])^2

        + sum {a1 in assets} (kurtosis[a1] - kurtosis_tar[a1])^2

        + sum {(a1, a2) in links} (covariance[a1, a2] - covariance_tar[a1, a2])^2


        + sum {a1 in assets, p in predecessors} (mean_ind[a1, p] - mean_tar_ind[a1, p])^2

        + sum {a1 in assets, p in predecessors} (variance_ind[a1, p] - variance_tar_ind[a1,
p])^2;


minimize Diff_Mom_final :

        sum {a1 in assets} (mean[a1] - mean_tar[a1])^2

        + sum{a1 in assets} (variance[a1] - variance_tar[a1])^2
```

```
        + sum {a1 in assets} (skewness[a1] - skewness_tar[a1])^2

        + sum {a1 in assets} (kurtosis[a1] - kurtosis_tar[a1])^2


        + sum {a1 in assets, p in predecessors} (mean_ind[a1, p] - mean_tar_ind[a1, p])^2

        + sum {a1 in assets, p in predecessors} (variance_ind[a1, p] - variance_tar_ind[a1,
p])^2;
```

Figure 5-1 AMPL format for moment matching scenario generating optimisation model

Most of the notation in the AMPL representation is self-explanatory, so comments on just some identifiers will be made. The generated scenario tree is depicted in Figure 5-2.

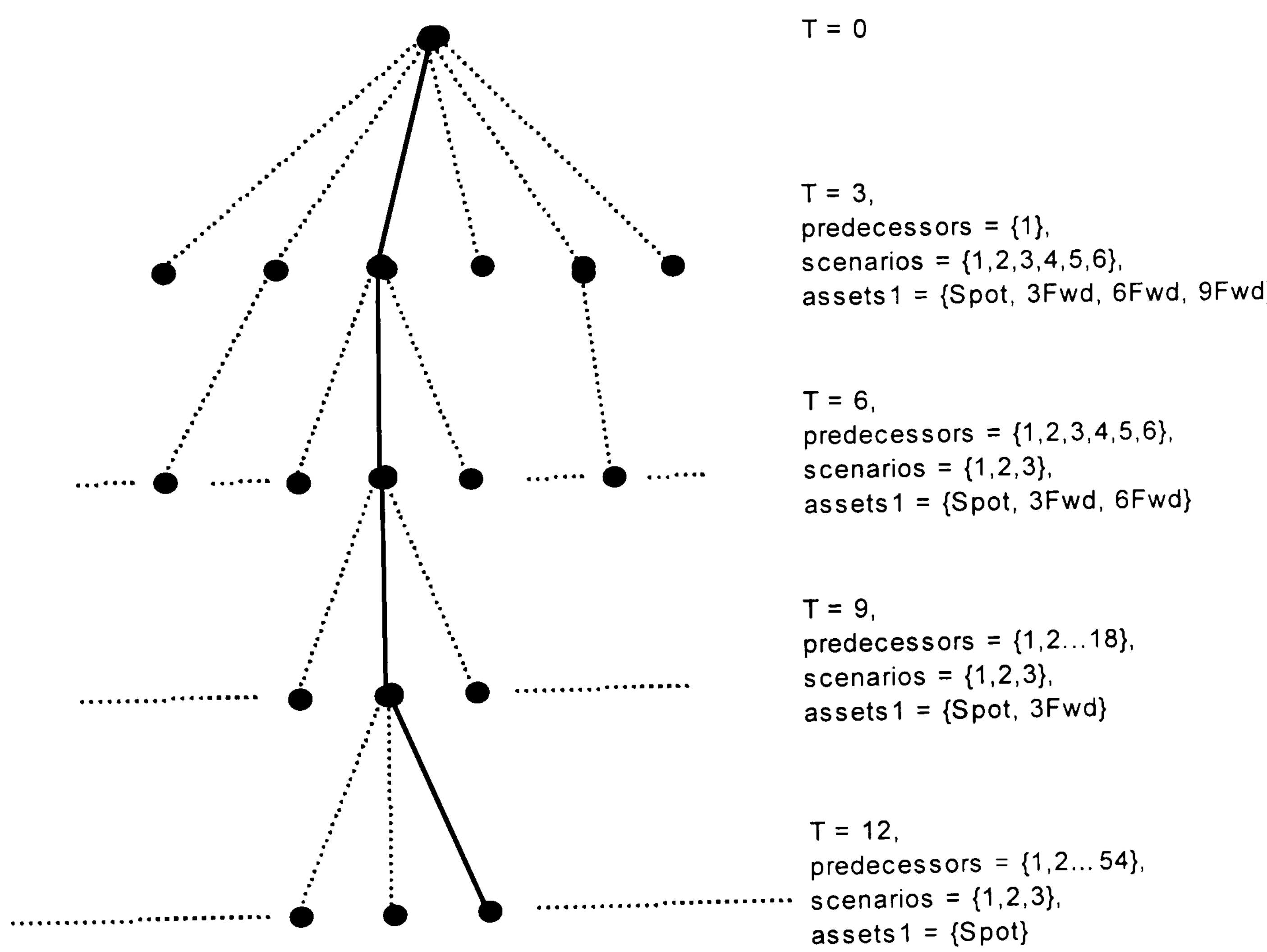

Figure 5-2 Scenario tree

The set "predecessors" represents the amount of nodes at a particular time period as opposed to the set "scenarios", which represent the number of scenarios emanating from each node at that time period. For instance, at the initial time period (root node) the set "predecessors" consists of just one element, $predecessors = \{1\}$ and the set "scenarios" consists of 6 elements, since there are in total six scenarios emanating from the root node, $scenarios = \{1,2,3,4,5,6\}$. In a similar fashion at time period 3 months into the future from the initial time period, t = 3, the set "predecessors" will be the same as the set "scenarios" at the initial time period, t = 0. Now the set "predecessors" consists of 6 elements, $predecessors = \{1,2,3,4,5,6\}$ and the set "scenario" consists of 3 element, $scenarios = \{1,2,3\}$, indicating that there are only 3 scenarios emanating from each node at t = 3. The same logic applies to t = 6 and t = 9.

The set "assets" consists of different types of exchange rates and is different at each time period. At time period 3 months ahead of the initial time period, i.e. t` = 3, $assets = \{Spot, 3Fwd, 6Fwd, 9Fwd\}$, for t' = 6 $assets = \{Spot, 3Fwd, 6Fwd\}$, for t' = 9 $assets = \{Spot, 3Fwd\}$ and for t' = 9 $assets = \{Spot\}$.

The set "links" is used to compute covariances between any two exchange rates. It represents a subset of Cartesian product of the set assets with itself, such that each element of "links" represents an ordered pair of elements from "assets", where the duplicated elements like (Spot, Spot) are removed from the subset and any two elements like (Spot, 3Fwd) and (3Fwd, Spot) are treated as just one element (Spot, 3Fwd).

Parameter T denotes the time period for which the tree nodes will be generated, for example, when T = 3 the model generates spot, 3-, 6- and 9-month forward exchange rates 3 month into the future from the initial time period (root node).

Parameter Prob refers to the probability of the node at a particular T. Since the scenario tree has the following branching structure: $6 \times 3 \times 3 \times 3$, the tree has 6 scenarios at T = 3, $18 = 6 \times 3$ scenarios at T = 6, $54 = 6 \times 3 \times 3$ scenarios at T = 9 and $162 = 6 \times 3 \times 3 \times 3$ scenarios at T = 12.

The suffix "_ind" in both parameters and variables refers to the conditional distribution at a particular tree node. As such, parameter Prob_ind refers to the conditional probabilities for each tree node given a particular time period. For instance, at time t = 0 (initial time period) each node (which is the only node – root node) has 6 successors therefore Prob_ind = 1/6 and at time t = 3 (3 months away from the initial time period) each node has 3 successors therefore Prob_ind = 1/3 at each node.

Variable x{assets, predecessors, scenarios} represents the value of element of the set "assets" for a particular scenario. Each scenario for a given time period is uniquely defined by a combination the set "predecessors" and set "scenarios" for that time period. For example, at time period T = 3, there are 6 scenarios since each scenario is built by combining only one predecessor, root node, with any of the six scenarios (successors) emanating from the root node. At time T = 6 there are 18 scenarios and each scenario is created by combining any of the 6 predecessors (representing tree nodes at T = 3) with any of the 3 scenarios (successors) emanating from any predecessor node.

As it can be seen from Figure 5-1, there are three distinct objective functions in the model, each selected for the relevant time period. The scenario generation model is run and hence the particular objective function is selected from the AMPL script file. The extract from the script for objective function selection is depicted in Figure 5-3.

```
if T = 3 then {
        objective Diff_MomCov_initial;}
else if T = 6 then {
        objective Diff_MomCov_ind:}
else if T = 9 then
```

```
        objective Diff_MomCov_ind;
else if T = 12 then
        objective Diff_Mom_final;
solve;
```

Figure 5-3 AMPL script to select an appropriate objective function for a particular time period.

The (moment matching) optimisation process is carried out sequentially and has four steps as outlined below:

1. *Generation of nodes for time, three months from now, i.e. t' = 3.* At the current (initial) time period we want to match the first four statistical moments and co-moments to their target counterparts. The random vector consists of spot exchange rates and 3-, 6- and 9-months forward exchange rates in three months from now, t' = 3. The target moments are estimated from historical sample, see Appendix D for the data sample used. In order to avoid arbitrage opportunities and hence unbounded solutions and infinite spurious profits inherent in the scenario tree we force the expected value of the spot exchange rates at newly generates tree nodes (in three months from now) be equal to the currently observed 3-months forward exchange rate. If we do not enforce this equality then the scenario tree will imply that "on average" one will be better off (worse off) by purchasing FWD (exchanging at spot rates), which would force trading in the market to bring the relationship to equilibrium. In this case study we assume equilibrium in the market, therefore equality of expected future spot rates with the currently observed forward rate for that time period thus complying with forward exchange rate parity relationship. Using the same reasoning, the expectation of generated 3-months forward rates at t' = 3 should equal the currently observed 6-months forward exchange rate, the expectation of 6-months forward rates at t' = 3 should equal the currently observed 9-months forward rate and the expectation of 9-months forward rates at t' = 3 should equal the currently observed 12-months forward rate. In

this step the objective function selected for minimisation from AMPL model formulation in Figure 5-1 is "Diff_MomCov_initial".

2. *Generation of nodes for time, six months from now, i.e. t' = 6.* At t' = 6 we repeat the algorithm used in step 1 and generate scenarios of rates for nodes distant 6 months from the initial (root node) time period. Since the overall model's time horizon is 12 months in this step the scenarios of spot, 3- and 6-months forward exchange rates are generated. Analogously, the expectation of spot exchange rates at t' = 6 should equal the currently observed 6-months forward rate, the expectation of 3-months forward rates at t' = 6 should equal the currently observed 9-months forward exchange rate and the expectation of 6-months forward rates at t' = 6 should equal the currently observed 12-months forward exchange rate. In this step the objective function selected for minimisation from AMPL model formulation in Figure 5-1 is "Diff_MomCov_ind".

3. *Generation of nodes for time, nine months from now, i.e. t' = 9.* At t' = 9 we repeat the algorithm used in step 1 and generate scenarios of rates for nodes distant 9 months from the initial (root node) time period. Since the overall model's time horizon is 12 months in this step the scenarios of spot, and 3-months forward exchange rates are generated. Analogously, the expectation of spot exchange rates at t' = 9 should equal the currently observed 9-months forward rate, the expectation of 3-months forward rates at t' = 6 should equal the currently observed 12-months forward exchange rate. In this step the objective function selected for minimisation from AMPL model formulation in Figure 5-1 is "Diff_MomCov_ind".

4. *Generation of nodes for time, twelve months from now, i.e. t' = 12.* At t' = 12 we repeat the algorithm used in step 1 and generate scenarios of rates for nodes distant 12 months from the initial (root node) time period. Since the overall model's time horizon is 12 months in this step only the scenarios of spot exchange rates are generated. The expectation of spot exchange rates at t' = 12 should equal the currently observed 12-months forward rate. In this step

the objective function selected for minimisation from AMPL model formulation in Figure 5-1 is "Diff_Mom_final".

In steps 1 to 3 the covariances among rates are taken into account so the objective function matches covariances for each time period to the target covariances. In step 4, since there is just one random variable left in the random vector, namely spot exchange rate, covariances matching is excluded from the objective.

In steps 2 to 4 when generating scenarios emanating from any particular node the objective function also matches the first two conditional moments to their conditional target counterparts. Higher moments conditional could be included in the same manner if desired. We include only the first two moments in order to keep the tree rather simple and small, which is appropriate to illustrate the concept in the current case study.

The first two target conditional moments, the first four target unconditional moments and target covariancies as well as the sets "predecessors" and "scenarios" and conditional and unconditional probabilities for each time period are included in Appendix F.

Since I do not include future realisations of interest rates in the scenario tree the interest rate parity relationship becomes redundant here. If, on the contrary, interest rates were included in the random vector at each tree node (other than the root node) then in order to generate an arbitrage-free scenario tree one would have to make sure that interest rate parity holds at each tree node.

The behaviour of the parameters $Spot, 3Fwd, 6Fwd, 9Fwd$ over time, is illustrated by the scenario tree illustrated in Figure 5-2. At any given time period each scenario is equiprobable. A probability of any "leaf" node is equal to the product of probabilities of scenarios, which lead from the root to that leaf node.

Once the scenario tree of exchange rates has been generated we add the remaining data parameter, namely the expected USD revenue at each tree node 3 months from the initial time period. Due to the assumption of independency of exchange rates from

USD revenues, we sample randomly at each tree node from the distribution function of USD revenues. We use just two realisation of random sampling for each tree node in order to keep the tree reasonably simple for this case study. In order to get a "good" approximation of USD revenues distribution function, however, one would need to use many realisations of random sampling at each node.

Now we will outline how the branching structure of the tree at each time period, i.e. the number of scenarios emanating from each node, was decided upon. The logic behind this decision is based on analysis outlined in Miller & Rice (1983) and is summarised in Appendix A. For an example of application of this logic see Hoyland & Wallace (2001), which is also briefly discussed in section 3.3.1 of this thesis.

The idea taken from Miller & Rice, manifests itself in solving a system of equations, where right-hand-sides are the target statistical moments and the left-hand-sides are the polynomials in variables, which represent particular realisations of a random variable. This is a special case of the system discussed by Miller & Rice, where probabilities of each realisation of a random variable are also variables. Since in this case study we enforce equiprobable scenarios the probabilities are eliminated from the vector of variables.

In order to solve such a system of equation the number of variable should be equal to the number of equations. If, for example, one wanted to match just the first two statistical moments and the random vector consisted of only one variable then the system would look like:

$$\begin{aligned} p_1x_1 &+ p_2x_2 &= E(x) \\ p_1x_1^2 &+ p_2x_2^2 &= E(x-E(x))^2 = \text{var}(x) \end{aligned} \tag{5.1}$$

Here, there are two equations in two unknowns and the minimum number of scenarios (realisations of the random variable) necessary to solve such a system is 2.

If the first four statistical moments were to be matched, as it is done in this case study, then the system of equations would look as follows:

$$\begin{array}{lllllllllll} p_1x_1 & + & p_2x_2 & + & p_3x_3 & + & p_4x_4 & = & E(x) & & \\ p_1x_1^2 & + & p_2x_2^2 & + & p_3x_3^2 & + & p_4x_4^2 & = & E(x-E(x))^2 & = & \mathrm{var}\,iance(x) \\ p_1x_1^3 & + & p_2x_2^3 & + & p_3x_3^3 & + & p_4x_4^3 & = & E(x-E(x))^3 & = & skewness(x) \\ p_1x_1^4 & + & p_2x_2^4 & + & p_3x_3^4 & + & p_4x_4^4 & = & E(x-E(x))^4 & = & kurtosis(x) \end{array} \qquad (5.2)$$

In order to solve such a system of four equations the number of variables should equal 4, hence there should be at least four scenarios.

The same principal applies to the multidimensional case, i.e. when the random vector has a dimension higher than 1. The calculations that follow were used to compute the minimum number of scenarios at each time period when designing the branching structure of the scenario tree.

- **T = 3 case**. Three months from the initial time period we generate scenario for the assets set $assets = \{Spot, 3Fwd, 6Fwd, 9Fwd\}$, which has dimensionality 4 (4 assets). For each asset we match 4 statistical moments, hence 4 constraints. For all 4 assets there would be $4 \times 4 = 16$ constraints. Besides, we match covariances among the 4 assets, which results in 6 additional constraints to satisfy the required co-moments. Therefore, the total number of constraints is 16 + 6 = 22. To find a solution of such a system there should be at least 22 variables. Since each node corresponds to 4 variables (each assets is a random variable) the minimum number of scenarios required to have at least 22 variables is 6.

- **T = 6 case**. Six months from the initial time period we generate scenario for the assets set $assets = \{Spot, 3Fwd, 6Fwd\}$, which has dimensionality 3 (3 assets). Here there would be two types of constraints: those that match unconditional and those that match conditional moments. When matching unconditional moments we match 4 statistical moments for each asset, hence 4 constraints. For all 3 assets there would be $3 \times 4 = 12$ constraints. Besides, we match covariances among the 3 assets, which results in 3 additional constraints to satisfy the required co-moments. Therefore, the total number of constraints is 12 + 3 = 15.

When matching conditional moments there would be 6 separate sub-trees, each corresponding to one of the 6 predecessors (six tree nodes at T = 3). For each such sub-tree we match the first 2 moments hence the number of constraints will be = $3 \times 2 = 6$. For all 6 sub-trees the number of constraints will be = $6 \times 6 = 36$. If we add the 15 unconditional constraints from the previous paragraph .the total number of all constraints will be = $36 + 15 = 51$.

The minimum number of variables need to solve a system of equations with 51 constrains is 51. It can be demonstrated that the minimum number of scenarios emanating from each of the 6 predecessor nodes should be 3, which results in 54 variables. Indeed, since each node at T = 6 corresponds to 3 variables and there 3 such nodes in each sub-tree there will be $3 \times 3 = 9$ variables. Since there are 6 such sub-trees the total number of variables will equal $9 \times 6 = 54$. It can be easily verified that if for each of the 6 predecessors the number of scenarios chosen is less than 3, e.g. is 2, it would result in the system where the number of equations is larger than the number of unknowns, hence there is no solution (unless the there some equations, which are the liner combinations of other equations in the system, which is not the case here).

- **<u>T = 9 case</u>**. Nine months from the initial time period we generate scenario for the assets set $assets = \{Spot, 3Fwd\}$, which has dimensionality 2. Here there would also be two types of constraints: those that match unconditional and those that match conditional moments. When matching unconditional moments we match 4 statistical moments for each asset, hence 4 constraints. For all two assets there would be $2 \times 4 = 8$ constraints. Besides, we match covariance between the two assets, which results in one additional constraint to satisfy the required co-moments. Therefore, the total number of constraints is 8 + 1 = 9.

When matching conditional moments there would be 18 separate sub-trees, each corresponding to one of the 18 predecessors (18 tree nodes at

T = 6). For each such sub-tree we match the first 2 moments hence the number of constraints will be = $2 \times 2 = 4$. For all 18 sub-trees the number of constraints will be = $4 \times 18 = 72$. If we add the 9 unconditional constraints from the previous paragraph the total number of all constraints will be = $72 + 9 = 81$.

The minimum number of variables need to solve a system of equations with 81 constrains is 81. It can be demonstrated that the minimum number of scenarios emanating from each of the 18 predecessor nodes should be 3, which results in 108 variables. Indeed, since each node at T = 9 corresponds to two variables and there are 3 such nodes in each sub-tree there will be $2 \times 3 = 6$ variables. Since there are 18 such sub-trees the total number of variables will equal $18 \times 6 = 108$. It can be easily verified that if for each of the 18 predecessors the number of scenarios chosen is less than 3, e.g. is 2, it would result in the system where the number of equations is larger than the number of unknowns, hence there is no solution.

- **<u>T = 12 case</u>**. Twelve months from the initial time period we generate scenario for the assets set $assets = \{Spot\}$, which has dimensionality 1. Here there would also be two types of constraints: those that match unconditional and those that match conditional moments. When matching unconditional moments we match 4 statistical moments, hence 4 constraints. Since there is just one assets, spot exchange rate, there is no co-moments matching requirements.

  When matching conditional moments there would be 54 separate sub-trees, each corresponding to one of the 54 predecessors (54 tree nodes at T = 9). For each such sub-tree we match the first 2 moments hence the number of constraints will be = 2. For all 54 sub-trees the number of constraints will be = $2 \times 54 = 108$. If we add the 4 unconditional constraints from the previous paragraph the total number of all constraints will be = $108 + 4 = 112$.

The minimum number of variables need to solve a system of equations with 112 constrains is 112. It can be demonstrated that the minimum number of scenarios emanating from each of the 54 predecessor nodes should be 3, which results in 162 variables. Indeed, since each node at T = 12 corresponds to 1 variable and there are 3 such nodes in each sub-tree there will be $1 \times 3 = 3$ variables. Since there are 54 such sub-trees the total number of variables will equal $54 \times 3 = 162$. It can be easily verified that if for each of the 54 predecessors the number of scenarios chosen is less than 3, e.g. is 2, it would result in the system where the number of equations is larger than the number of unknowns, hence there is no solution.

After generating the scenario tree for exchange rates applying the above algorithm, two random samples are drawn from the probability distribution of USD cash flows 3 months from the initial time period. When doing this the number of scenarios doubles from 162 to 324. The resulting tree is used as an input to the SP (decision) model described in the next section.

5.3 The SP Decision Model

In this section we develop a stochastic optimisation model for determining the best "hedged" investments in FWD. The SP model represents a four-stage SP model, where each stage corresponds to the beginning of the calendar quarter. Thus, the first stage is the first quarter; the second stage is the second quarter etc. The first stage decisions represent the contracts on the forward exchange rates that should be purchased or sold at the current time period (at the beginning of the first quarter); the second stage decisions represent the FWD that should be purchased or sold at the beginning of the second quarter, i.e. in 3 months from the current period; the third stage decisions represent the FWD that should be purchased or sold at the beginning of the third quarter, i.e. in 6 months from the current period and the fourth stage decisions represent the FWD that should be purchased or sold at the beginning of the fourth quarter, i.e. in 9 months from the current period.

The idea embedded in the decision model is similar to that of Sharda & Wingender (1991). They formulated a goal-programming model to dynamically hedge accounts receivable with futures currency contracts. Apart from using the stochastic structure of random parameters and incorporating uncertain future cash flows into the scenario tree we also extend the objective function of the model as follows.

Our objective function has two main components: (i) minimising volatility of GBP-converted total cash inflows over the year; (ii) minimising transaction costs. The objective function can be easily extended to the case where the MNF would like to maximise risk-adjusted return (cash inflows). It makes even more sense to maximise the risk-adjusted inflows if the set of available assets included apart from FWD (which essentially are risk-free assets) risky assets such as options on foreign exchange rates. Since such an extension of the assets set would significantly complicate construction of the scenario tree we postpone it to our future research. The treasury manager of the MNF can also specify the weights attached to each of these goals.

It should be pointed out that since we use FWD to hedge both FX and cash flow risk we effectively use speculative strategies to minimize volatility of GBP-converted accounts receivables. In order to prevent the model to take very big positions in FWD the treasury manager sets the upper limit on the size of the position in FWD that the MNF is allowed to take, which could reflect the MNF's internal views or rules towards hedging FX risk.

By varying the weights assigned to different goals and varying the maximum forward exposure limit, the MNF has the flexibility to choose its preferred strategy. The four-stage SP decision model in AMPL notation is formulated below. In what follows we use AMPL notation for ease of explanation of the model element. The algebraic formulation of the model is presented in Appendix G.

Indices

assets = {Spot, 3Fwd, 6Fwd, 9Fwd, 12Fwd}: the set of assets (type of exchange rates) available for investment.

time = {0, 3, 6, 9, 12}: time (month of the year) when decisions are taken or data parameters are observed.

scenarios = {1,.., 324 }: set of scenarios.

links (t intime) = {(s1 in scenarios, s2 in scenarios)}: is a set of ordered pairs indexed over the set "time", where each element of the pair belongs to the set "scenarios". This set is necessary to define the SP model in split-variable form (see Brandimarte (2002) for an example of a split-variable formulation and its AMPL format).

Data

Transaction cost:

TransCost: transaction cost of acquiring / selling FWD. Theoretically there is no charge for entering a forward agreement though the bank could charge for selling back the outstanding forward contract (closing out a forward position). Transaction cost can also reflect an ask-bid spread. In this study we set the value of transaction costs at 1% of the contract value. Also, at this level of transaction costs any potential arbitrage opportunities embedded into the scenario tree (resulting in spurious profits from speculating with FWD) would probably disappear.

Exchange rates:

Xrate (a1 in assets, t in time, s in scenarios): a "a1" type of USD:GBP exchange rates at time t under scenario s. For example, t = 0 refers to currently observed market rates whereas t > 0 refers to future exchange rates under some scenario s.

Cash Flows:

Revenue (t in time, s in scenarios): expected USD revenues (accounts receivables) at time t under scenario s.

Initial data:

FwdPrev(a1 in assets | $a1 \notin \{Spot, 12Fwd\}$*)*: number of FWD of a1 type brought forward from the previous quarter (from 3 months ago) with maturity month at least 1 quarter into the future from the current time period.

Probabilities:

Prob: probability of a scenario. Since all scenarios are equiprobable in the created tree $\Pr ob = \frac{1}{|scenarios|} = \frac{1}{324}$.

UpperLimitOnFwd:

UpperLimitOnFwd: treasury set upper limit on the proportion of the net cash flows to be offset by taking a position in FWD.

<u>Decision variables</u>

FwdHold (a1 in assets | $a1 \notin \{Spot\}$*, t in time, s in scenarios)*: total amount of FWD of a1 type held at time t after rebalancing under scenario s. When t = 0 *FwdHold* refers to the first stage variable, i.e. a decision is taken at the current time period. When t > 0 *FwdHold* refers to the second or higher stage variable with t = 3, 6, 9 corresponding to the second, third and fourth stages respectively.

FwdBuy (a1 in assets | $a1 \notin \{Spot\}$*, t in time, s in scenarios)*: amount of FWD of a1 type purchased at time t after rebalancing under scenario s. When t = 0 *FwdBuy* refers to the first stage variable, i.e. a decision is taken at the current time period. When t > 0 *FwdBuy* refers to the second or higher stage variable with t = 3, 6, 9 corresponding to the second, third and fourth stages respectively.

FwdSell (a1 in assets | $a1 \notin \{Spot\}$*, t in time, s in scenarios)*: amount of FWD of a1 type sold (settled) at time t after rebalancing under scenario s. When t = 0 *FwdSell* refers to the first stage variable, i.e. a decision is taken at the current time period. When t > 0 *FwdSell* refers to the second or higher stage variable with t = 3, 6, 9 corresponding to the second, third and fourth stages respectively.

Reporting (substitution) variables

Cost (s in scenarios): variable representing the yearly costs resulted from trading FWD under scenario s.

ExpYearlyCost: variable representing expected yearly transaction costs resulted from trading FWD.

YearlyRev (s in scenarios): variable representing the yearly GBP-converted inflows, which consist of accounts receivable and gains / losses made holding FWD under scenario s.

ExpYearlyRev: variable representing expected yearly GBP-converted inflows, which consist of accounts receivable and gains / losses made holding FWD.

Variance: variable representing variance of *YearlyRev*.

Objective Function

The objective function represents a trade-off between volatility of future yearly GBP-converted inflows from both accounts receivable and trading with FWD and expected yearly transaction costs from trading with FWD:

$$\text{Minimise } \left(\sqrt{Variance} + ExpYearlyCost\right) \tag{5.3}$$

The decision model in AMPL format is formulated in Figure 5-4.

```
### SETS ###

set assets ordered;

set time ordered;

set scenarios := 1..324 ordered;

set links {t in time: t <= 9} within {1..324, 1..324};

### PARAMETERS ###
```

```
param Prob ;

param Xrate {assets, time, scenarios};

param Revenue {time, scenarios};

param TransCost;

param FwdPrev {a1 in assets: a1 <> '12Fwd'};

param UpperLimitOnFwd;

### VARIABLES      ###

var FwdHold {a1 in assets, time, scenarios: a1 <> 'Spot'};

var FwdSell {a1 in assets, time, scenarios: a1 <> 'Spot'};

var FwdBuy {a1 in assets, time, scenarios: a1 <> 'Spot'};

var Cost {s in scenarios} = TransCost * (

        + (FwdBuy ['12Fwd', 0, s] + FwdSell ['12Fwd', 0, s]

        +  FwdBuy ['9Fwd', 0, s] + FwdSell ['9Fwd', 0, s]

        +  FwdBuy ['6Fwd', 0, s] + FwdSell ['6Fwd', 0, s]

        +  FwdBuy ['3Fwd', 0, s] + FwdSell ['3Fwd', 0, s]) / Xrate['Spot', 0, s]


        + (FwdBuy ['9Fwd', 3, s] + FwdSell ['9Fwd', 3, s]

        +  FwdBuy ['6Fwd', 3, s] + FwdSell ['6Fwd', 3, s]

        +  FwdBuy ['3Fwd', 3, s] + FwdSell ['3Fwd', 3, s]) / Xrate['Spot', 3, s]


        + (FwdBuy ['6Fwd', 6, s] + FwdSell ['6Fwd', 6, s]

        +  FwdBuy ['3Fwd', 6, s] + FwdSell ['3Fwd', 6, s]) / Xrate['Spot', 6, s]


        + (FwdBuy ['3Fwd', 9, s] + FwdSell ['3Fwd', 9, s]) / Xrate['Spot', 9, s]  );

var ExpYearlyCost =    sum {s in scenarios} Prob * Cost[s] ;

var YearlyRev {s in scenarios} =

          Revenue[3, s] / Xrate['Spot', 3, s]

        + Revenue[6, s] / Xrate['Spot', 6, s]
```

```
        + Revenue[9, s] / Xrate['Spot', 9, s]
        + Revenue[12, s] / Xrate['Spot', 12, s]

        + FwdHold['12Fwd', 0, s] * (1 / Xrate['12Fwd', 0, s] - 1 / Xrate['Spot', 12, s])
        + FwdHold['9Fwd', 3, s] * (1 / Xrate['9Fwd', 3, s] - 1 / Xrate['Spot', 12, s])
        + FwdHold['6Fwd', 6, s] * (1 / Xrate['6Fwd', 6, s] - 1 / Xrate['Spot', 12, s])
        + FwdHold['3Fwd', 9, s] * (1 / Xrate['3Fwd', 9, s] - 1 / Xrate['Spot', 12, s])

        + FwdHold['9Fwd', 0, s] * (1 / Xrate['9Fwd', 0, s] - 1 / Xrate['Spot', 9, s])
        + FwdHold['6Fwd', 3, s] * (1 / Xrate['6Fwd', 3, s] - 1 / Xrate['Spot', 9, s])
        + FwdHold['3Fwd', 6, s] * (1 / Xrate['3Fwd', 6, s] - 1 / Xrate['Spot', 9, s])

        + FwdHold['6Fwd', 0, s] * (1 / Xrate['6Fwd', 0, s] - 1 / Xrate['Spot', 6, s])
        + FwdHold['3Fwd', 3, s] * (1 / Xrate['3Fwd', 3, s] - 1 / Xrate['Spot', 6, s])

        + FwdHold['3Fwd', 0, s] * (1 / Xrate['3Fwd', 0, s] - 1 / Xrate['Spot', 3, s]) ;
var ExpYearlyRev =      sum {s in scenarios} Prob * YearlyRev[s] ;
var Variance = sum{s in scenarios} (((YearlyRev[s] - ExpYearlyRev)^2) * Prob);
### OBJECTIVES      ###
minimize Volatility: sqrt(Variance) + ExpYearlyCost;
### CONSTRAINTS     ###
subject to
UpperHedgeLimit3 {s in scenarios}: FwdHold['3Fwd', 0, s] <= UpperLimitOnFwd *
Revenue[3, s];
UpperHedgeLimit6 {s in scenarios}:
        (FwdHold['6Fwd', 0, s] + FwdHold['3Fwd', 3, s]) <= UpperLimitOnFwd *
Revenue[6, s];
UpperHedgeLimit9 {s in scenarios}:
        FwdHold['9Fwd', 0, s] + FwdHold['6Fwd', 3, s] + FwdHold['3Fwd', 6, s] <=
```

```
        UpperLimitOnFwd * Revenue[9, s];
UpperHedgeLimit12 {s in scenarios}:
        FwdHold['12Fwd', 0, s] + FwdHold['9Fwd', 3, s] + FwdHold['6Fwd', 6, s]
        + FwdHold['3Fwd', 9, s] <= UpperLimitOnFwd * Revenue[12, s];
BalanceConstr0_3Months {s in scenarios}:
        FwdHold ['3Fwd', 0, s] = FwdPrev ['3Fwd']
        + FwdBuy ['3Fwd', 0, s] - FwdSell ['3Fwd', 0, s];
BalanceConstr0_6Months {s in scenarios}:
        FwdHold ['6Fwd', 0, s] = FwdPrev ['6Fwd']
        + FwdBuy ['6Fwd', 0, s] - FwdSell ['6Fwd', 0, s];
BalanceConstr3_6Months {s in scenarios}:
        FwdHold ['3Fwd', 3, s] = FwdBuy ['3Fwd', 3, s] - FwdSell ['3Fwd', 3, s];
BalanceConstr0_9Months {s in scenarios}:
        FwdHold ['9Fwd', 0, s] = FwdPrev ['9Fwd']
        + FwdBuy ['9Fwd', 0, s] - FwdSell ['9Fwd', 0, s];
BalanceConstr3_9Months {s in scenarios}:
        FwdHold ['6Fwd', 3, s] = FwdBuy ['6Fwd', 3, s] - FwdSell ['6Fwd', 3, s];
BalanceConstr6_9Months {s in scenarios}:
        FwdHold ['3Fwd', 6, s] = FwdBuy ['3Fwd', 6, s] - FwdSell ['3Fwd', 6, s];
BalanceConstr0_12Months {s in scenarios}:
        FwdHold ['12Fwd', 0, s] = FwdBuy ['12Fwd', 0, s] - FwdSell ['12Fwd', 0, s];
BalanceConstr3_12Months {s in scenarios}:
        FwdHold ['9Fwd', 3, s] = FwdBuy ['9Fwd', 3, s] - FwdSell ['9Fwd', 3, s];
BalanceConstr6_12Months {s in scenarios}:
        FwdHold ['6Fwd', 6, s] = FwdBuy ['6Fwd', 6, s] - FwdSell ['6Fwd', 6, s];
BalanceConstr9_12Months {s in scenarios}:
        FwdHold ['3Fwd', 9, s] = FwdBuy ['3Fwd', 9, s] - FwdSell ['3Fwd', 9, s];
NonAnticipBuy {a1 in assets, t in time, (s1, s2) in links[t]: t <= 9 and a1 <> 'Spot'}:
```

```
        FwdBuy [a1, t, s1] = FwdBuy [a1, t, s2];
NonAnticipSell {a1 in assets, t in time, (s1, s2) in links[t]: t <= 9 and a1 <> 'Spot'}:
        FwdSell [a1, t, s1] = FwdSell [a1, t, s2];
NonAnticipHold {a1 in assets, t in time, (s1, s2) in links[t]: t <= 9 and a1 <> 'Spot'}:
        FwdHold [a1, t, s1] = FwdHold [a1, t, s2];
LowerBound1 {a1 in assets, t in time, s in scenarios: a1 <> 'Spot'}: 0 <= FwdHold [a1, t, s];
LowerBound2 {a1 in assets, t in time, s in scenarios: a1 <> 'Spot'}:  0 <= FwdBuy [a1, t, s];
LowerBound3 {a1 in assets, t in time, s in scenarios: a1 <> 'Spot'}:  0 <= FwdSell [a1, t, s];
```

Figure 5-4 The four-stage SP model in AMPL format

In what follows we will give an explanation of each of the constraints in the model.

- ***UpperHedgeLimit constraints***

Constraints comprising this group put an upper limit on cumulative FWD exposure for a maturity date as a proportion of accounts receivable (USD revenues) expected at that date. For example, if the maturity date is t = 9 months from the current (t = 0) time period, FWD exposure for t = 9 will be: FwdHold[9Fwd, 0, s] + FwdHold[6Fwd, 3, s] + FwdHold[3Fwd, 6, s], which represents cumulative outstanding position in FWD established from t = 0 to t = 6 and maturing at t = 9.

- ***Balance constraints***

These constraints make sure that after rebalancing positions in FWD the amount of FWD held for a specific maturity, say t, is equal the amount of FWD inherited from 3 months ago plus amount of FWD bought less amount of FWD sold with the same maturity t.

- ***Non-anticipativity constraints***

These constraints make sure that if at certain time, t, some scenarios pass through the same tree node then the decisions taken at that time, t, under those scenarios will be identical. These constraints prevent the model from "seeing" the future.

- ***LowerBound constraints***

These constraints represent lower bounds on variables related to FWD trading. They require all such variables are non-negative hence precluding short-sales.

5.4 Analysis of the Results

5.4.1 Dynamic Data Model

The role of the historical market data, the organisational data and their interaction with the decision model is illustrated in Figure 5-5. The experimental set up requires that we dynamically:

(i) Use market data in order to revalue the forward positions, a well-known "mark to market" procedure.

(ii) We also record the decisions made in the current step of the model as an input of the starting position of the next "roll" of the model.

Whereas in futures currency contracts there is an external requirement for "marking to market", for forward positions there is no such obligation. As an "internal good practice procedure", however, we have introduced this in our FWD decision model so that we are able to compute the "moneyness" of the current positions to give some indication of ongoing performances.

Thus for each time movement the model database is updated with the most currently available forward rates and spot rates. By accessing our current forward commitments along with their current marked to market forward rate from the model database, we adjust our forward rates to be the same as the current month forward rates. The process of "marking to market" of our currently held FWD involves realigning the contracts by one time period as well as determining the financial losses or gains made on our forward positions. Similarly, we close out the opening income stream using a combination of currently maturing FWD and the current spot rate. All these cash transactions, namely the marking to market of FWD and conversion of the current income revenue are recorded in our financial database.

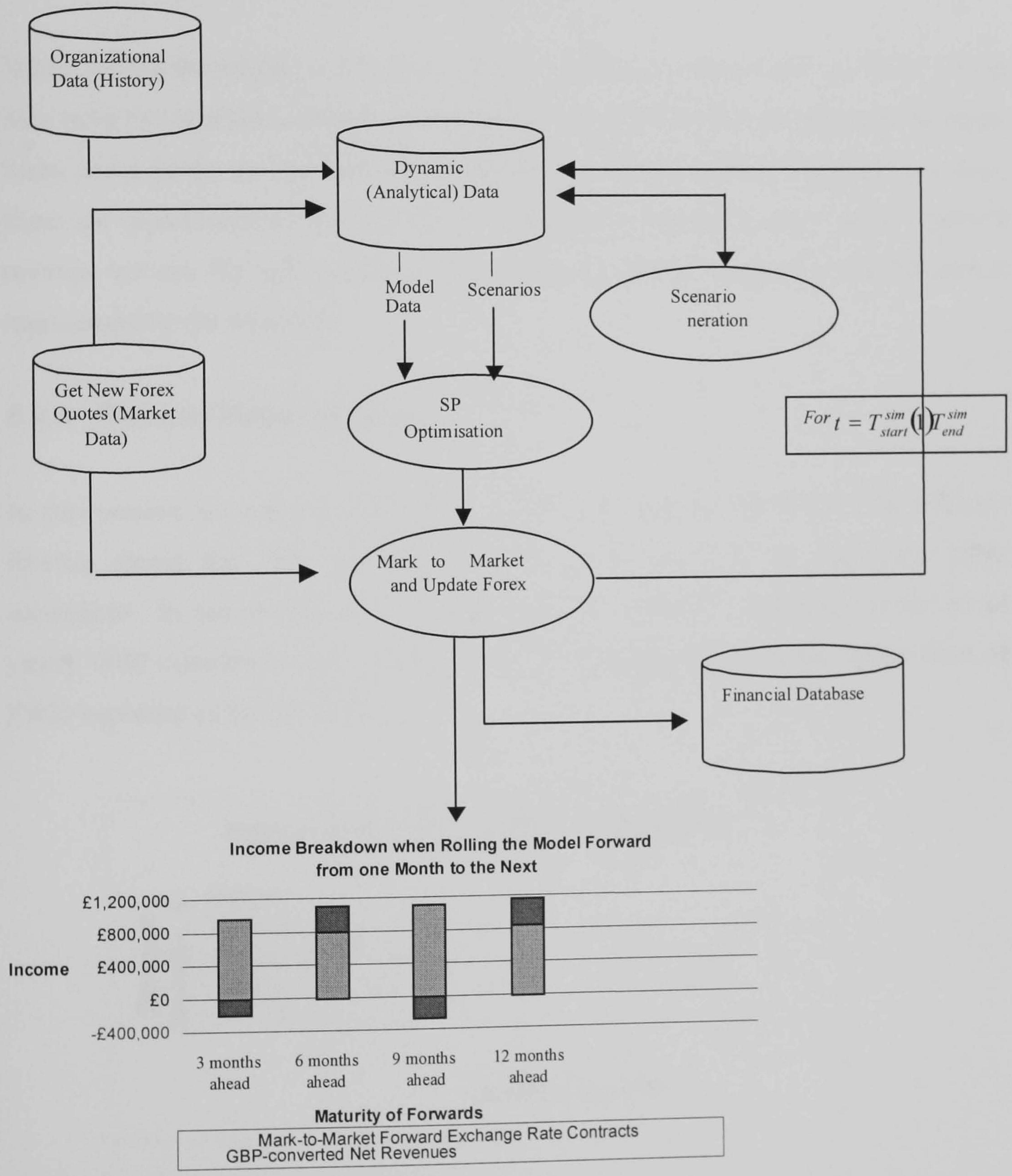

Figure 5-5 Rolling the model forward and "Mark to Market" process.

5.4.2 *The Rolling Decision Model*

The decision model uses data sets, which are updated every quarter. The scenario generator uses the historical spot rates, historical forward rates and revenue data having stepped through by one quarter t = t + 1. Thus the scenario generator creates a completely new set of scenarios looking ahead over a time horizon of T = 4 quarters.

When rolling the model forward we mark to market positions held in FWD, which may have two outcomes. Firstly, in realigning our FWD to the current rates we either make some profit on our currently held FWD or our speculation has led to a loss, these are represented by the red bars. Secondly, in processing the current quarter's revenue we use the spot rate thus the income revenue is marked to market and is represented by the blue bars.

5.4.3 *Risk and Return Analysis*

In this section we estimate and draw an efficient frontier for MNF. The efficient frontier shows the best possible return for a certain level of risk given other constraints. In the model risk is represented by volatility (standard deviation) of yearly GBP-converted cash inflows, which is directly related to the upper limit of FWD exposure as shown in Figure 5-6.

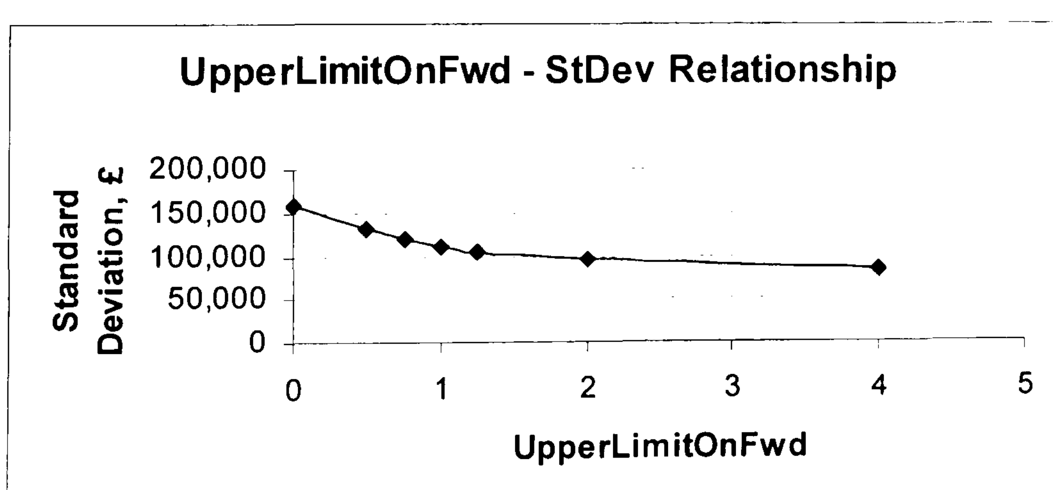

Figure 5-6 Relationship between UpperLimitOnFwd coefficient and Volatility (Standard Deviation) of GBP-coverted Cash Inflows.

The smaller this limit is the less flexibility MNF has and the smaller FWD position it can take. At the extreme case if the upper limit of FWD is zero then MNF is not allowed to trade in FWD at all hence all future accounts receivable will be converted at spot exchange rates. This increases the risk since there is no asset MNF could use to offset the change in both future USD cash flows and in future exchange rates (there is no hedge in this case). The equivalent of return is expected yearly GBP-converted cash inflow.

Any distinct value of UpperLimitOnFwd represents a particular FWD trading strategy. Some of the FWD trading strategies optimised by SP model are tabulated in Table 1.

Table 1 Optimal risk and return pairs for some FWD trading strategies produced by SP model

Strategy Type (UpperLimitOnFwd)	*ExpYearlyCost, £*	*StDev, £*	*Volatility: StDev + ExpYearlyCost, £*	*ExpYearlyRev, £*
4	33,147	82,258	115,405	2,530,850
2	27,216	94,527	121,743	2,532,020
1,25	21,674	105,479	127,152	2,534,580
1.00	18,649	111,719	130,368	2,535,530
0.75	15,164	119,818	134,982	2,536,730
0.5	11,144	129,988	141,133	2,538,290
0	0	159,291	159,291	2,541,680

These various trading strategies represent an efficient frontier and can be visualised by a graph as shown in Figure 5-7.

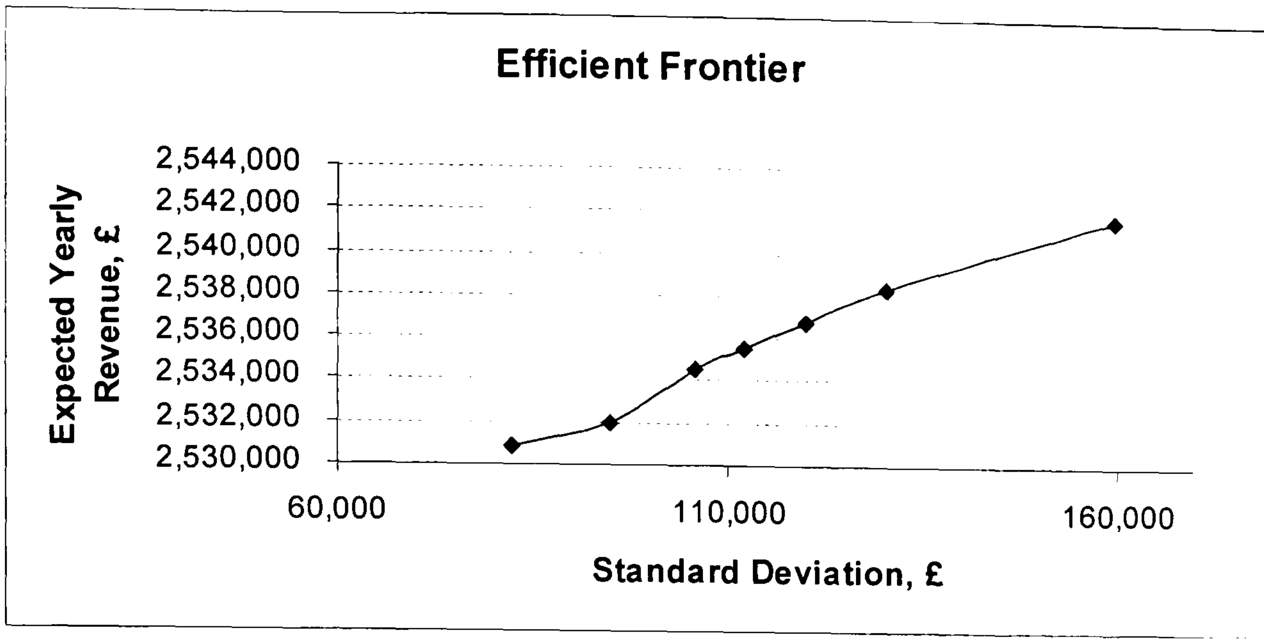

Figure 5-7 Volatility – Revenue Efficient Frontier when SP Model is Applied

As can be seen from Figure 5-7 the tighter upper limit of FWD trades results in higher GBP-converted yearly revenues volatility hence higher risk. If MNF's base strategy is not to use any FWD and convert all USD revenues at spot exchange rates then this case represents a "risky" strategy, since MNF does not use hedging at all. By increasing the UpperLimitOnFwd coefficient the volatility of the future yearly revenues declines quite rapidly. For example, by raising UpperLimitOnFwd from zero (base strategy) to 2 standard deviation drops from 159,291 to 94,527 (approx. 41%). Of course these drops in volatility are accompanied by drops in Expected Yearly Revenues but the relative revenue drops are immaterial compared with those for standard deviations. For example, by raising UpperLimitOnFwd from zero (base strategy) to 2 expected yearly revenue drops from 2,541,680 to 2,532,020 (approx. 0.38%).

Also, since MNF uses FWD to hedge not only FX risk but also some change in USD accounts receivables, excessively high values of UpperLimitOnFwd can be deemed risky and therefore other company internal regulations and policies with regard to FWD exposure are necessary. By using UpperLimitOnFwd coefficient MNF can control such an exposure.

In order to evaluate the effect of incorporation of random future USD accounts receivable in 3 months from the initial time period in the scenario tree we analyse the volatility – revenue relationship of strategies when MNF uses forecasts (expected values) of future USD revenues instead of a scenario tree. This strategy, which is quite often employed in "real" world, "assumes" that the mean value of the USD account receivables will realise. Since both the expected USD cash inflows in 3 months from the current time period and USD cash inflows in 6, 9 and 12 months into the future equal $1,000,000, the only risk MNF can hedge is FX risk (USD cash inflow risk is eliminated by replacing scenarios of future inflows by expected values), which is done by purchasing 3, 6, 9 and 12 months FWD to sell $1,000,000. Table 2 shows various risk-return (standard deviation - revenue in our context) combinations when the strategy (amount of FWD for 3, 6, 9 and 12 months ahead) is a multiple of the expected USD revenues (e.g. $1,000,000 of revenues after the first quarter).

Table 2 Risk and return pairs for some FWD trading strategies based on expected revenues (when SP model is not applied)

Strategy Type, $	*ExpYearlyCost, £*	*StDev, £*	*Volatility: StDev + ExpYearlyCost, £*	*ExpYearlyRev, £*
4,000,000	99,385	279,466	378,851	2,513,260
2,000,000	49,693	126,162	175,855	2,527,470
1,500,000	37,269	107,539	144,809	2,531,020
1,250,000	31,058	105,203	136,261	2,532,800
1,000,000	24,846	108,121	132,967	2,534,570
500,000	12,423	127,644	140,067	2,538,120
250,000	6,212	142,385	148,597	2,539,900
0	0	159,291	159291	2,541,680

At one extreme, MNF may not hedge its expected USD revenues at all and hence no FWD are purchased. Since this is obviously very limiting strategy, its objective of minimising volatility of the cash inflow stream is limited as well, which is reflected by higher standard deviation, £159,291. If MNF increases its FWD exposure from 0 to about $1,250,000, it may achieve lower volatility (accompanied by lower expected revenues) since it has more flexibility to use FWD to offset some of the USD revenues volatility. However, if gone to the other extreme, such as purchasing $4,000,000 worth of FWD, MNF's strategy would become sub-optimal since it results in higher volatility and lower expected revenues.

In Figure 5-8 we compare the two approaches: 1) choosing FWD strategy based on application of SP model and 2) choosing FWD strategy based on forecasts (expected values) of future USD revenues.

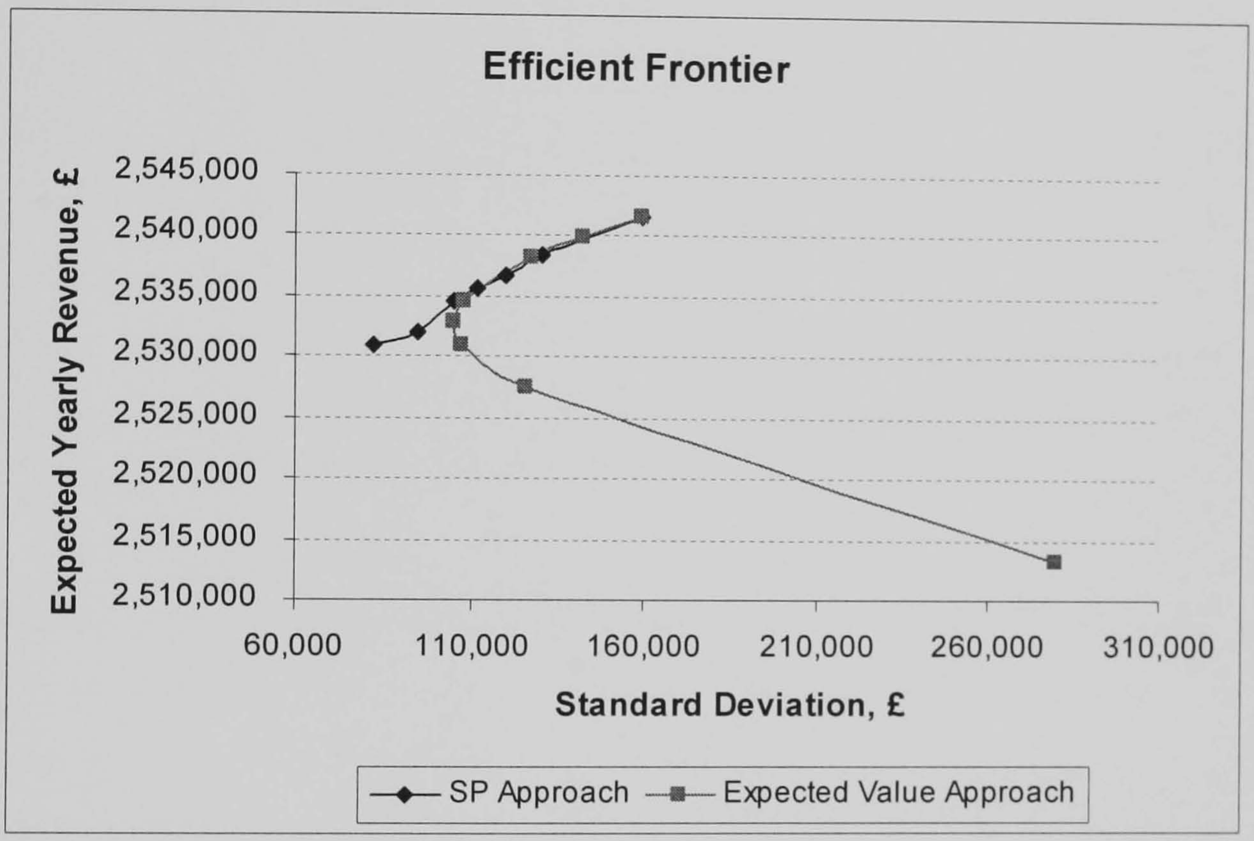

Figure 5-8 Volatility vs. Revenue Relationship for SP Model and Expected Value Approaches

As expected the expected value approach is sub-optimal to the SP approach, though coming very close to it when the amount of FWD exposure for each of the four future quarters is less than $1,000,000. These results demonstrate that if a company's policy does not allow maintaining FWD exposure higher than expected revenue stream then both SP and expected value approach are approximately equivalent. As long as MNF's FWD exposure is below its expected revenues then the optimal FWD strategy lies on the efficient frontier such as shown in Figure 5-8. If MNF's policies allow holding FWD exposures higher than the expected accounts receivables then the SP approach is superior and the optimal FWD strategy will lie on its efficient frontier.

Chapter 6. Conclusions

Summary

In this thesis we illustrate a complete process of financial planning under uncertainty. We begin with an overview of types of SP decision models and stochastic measures used to assess the quality of problem solutions. We then outline the whole process of building a scenario tree for an SP decision model. We split this process into three major steps: (1) calibration of *data processes* underlying the dynamics of random parameters; (2) *sampling*, which represents approximation of a continuous distribution of random parameters and (3) scenario *tree construction* necessary for multi-stage decisions and for adjusting the tree structure to certain predetermined characteristics. We provide a theoretical review of the techniques involved in each of the steps. Then we describe some of the most important tree properties, namely, the arbitrage-free condition, and show techniques that can be used to build a tree, which satisfies such properties.

A case study in FX hedging is presented in order to illustrate the implementation of all the steps of the problem modelling process. The application of matching statistical moments of a probability distribution to generate a multiperiod scenario tree for our problem is described in detail. A four-stage SP decision model is formulated using the random parameter values. This model computes currency / cash flows hedging strategies, which provide rolling decisions on the size and timing of the currency forward market exposure. We compute an efficient frontier from which an investor can choose an optimal strategy according to his risk and return preferences. The flexibility of the SP model allows an investor to analyse various risk-return trading strategies. The model decisions are investigated by making comparisons with decisions based purely on forecast (expected) future cash flows. The investigation shows that there is a considerable improvement to the "spot only" strategy and provides insight into how these decisions are made.

Contributions

The contributions of the thesis are summarised below. (i) Review of the complete process of problem modelling and decision-making under uncertainty, and description of alternative methods used at each step of the process. (ii) The FX forward scenario trees are derived using arbitrage-free pricing strategy and is in line with modern principles of finance. (iii) Use of the SP model and forward contracts as a tool for hedging decisions is novel. (iv) In particular smoothing of the effects in exchange rates and smoothing of account receivables are examples of innovative modelling approach for FX management.

Future research

The work presented in this thesis could be extended in a number of ways, which are summarised below. (i) The model could include more volatile currencies, which makes the model more flexible and accessible by users of new currencies. (ii) Interest rates in different countries could be included in the scenario tree and the decision model in order to incorporate the "time value" concept into the optimal decision-making. (iii) The objective function could be extended in order to allow for a risk and return trade off in different investment decisions. (iv) Other hedging instruments such as currency options could be included, which would enhance the characteristics of the portfolio compositions. (v) And finally, other alternative scenario generation methods, for example, parsimonious vector autoregressive model of lag one could be analysed.

References

Abdullah, F., and Wingender, J., "*Multinational Financial Management: Foreign Exchange Exposure and International Cash Management*," Journal of the Midwest Finance Association, 1987.

Adler, M., and B.Dumas. "*Exposure to Currency Risk: Definition and Measurement*." Financial Management, 13 (Summer 1984), 41-50.

Barraquand, J. "*Numerical Evaluation of High Dimensional Multivariate European Securities*", Paris Research Laboratory, March 1993.

Beltratti, A., Consiglio, A., and S. Zenios "*Scenario Modelling for the Management of International Bond Portfolios*", Annals of Operations Research, 85 (1999), 227-247.

Beltratti, A., Laurant, A., and S. Zenios "*Scenario Modelling for Selective Hedging Strategies*", Journal of Economic Dynamics and Control, 28 (2004), 955-974.

Berkelaar, A.B., Hoek, H, and A. Lucas "*Arbitrage and sampling uncertainty in financial stochastic programming models*", Econometric Institute Report 147 (1999), Erasmus University Rotterdam, Econometric Institute.

Birge, J., and F. Louveaux "*Introduction to Stochastic Programming*", Springer Series in Operations Research, 1997.

Birge, J., and J. Mulvey "*Stochastic programming*", Mathematical Programming for Industrial Engineers (1996), Dekker, New York, 543-574.

Boender, G. "*A hybrid simulation / optimization scenario model for asset / liability management*". European Journal of Operational Research, 99 (1997), 126-135.

Boender, G., van Aalst, P., and F. Heemskerk "*Modelling and management of assets and liabilities of pension plans in the Netherlands*", in Ziemba & Mulvey (1998), 561-580.

Boyle, P., Broadie, M., and P. Glasserman "*Monte Carlo methods for security pricing*", Journal of Economic Dynamics and Control, 21 (1997), 1267-1321.

Brandimarte, P. "*Numerical Methods in Finance: a MATLAB-based Introduction*", Wiley series in probability and statistics, 2002.

Britto, R. "*The Simultaneous Determination of Spot and Futures Prices in a Simple Model with Production Risk*" Quarterly Journal of Economics, 99 (May 1984), 351-365.

Canestrelli, E., and S. Giove "*Scenarios identification for financial time series*", Current Topics in Quantitative Finance (1999), Physica-Verlag, 25-36.

Carino, D., Kent, T., Myers, D., Stacy, C., Sylvanus, M., Turner, A., Watanabe, K., and W. Ziemba, "*The Russel-Yasuda Kasai Model: An Asset/Liability Model for a Japanese Insurance Company Using Multistage Stochastic Programming*", Interfaces, 24 (Jan.-Feb. 1994), 29-49.

Christofides, A., Tanyi, B., Christofides, S., Whobrey, D., and N. Christofides "*The optimal discretization of probability density functions*", Computational Statistics and Data Analysis, 31 (1999), 475-486.

Chvatal, V. "*Linear Programming*", Freeman, New York / San Francisco, CA, 1980.

Clemen, R. "*Making Hard Decisions: An Introduction to Decision Analysis*", PWS-Kent Publishing Co., Boston, 1991.

Cornuejols, G., and R. Tutuncu "*Optimization Methods in Finance*", Course Material, Summer 2005, Carnegie Mellon University, Pittsburgh, PA 15213 USA; web link: http://integer.tepper.cmu.edu/webpub/finance/OptFin05.pdf.

Darius, D., Ilhan, A., Mulvey, J., Simsek, K. and R. Sircar "*Trend-following Hedge Funds and Multi-period Asset Allocation*", (2002), Department of Operations Research & Financial Engineering, Princeton University E-Quad, Princeton NJ 08544; web link: http://www.princeton.edu/~sircar/Public/ARTICLES/dimss.pdf.

Dempster, M., and A., Thorlacius "*Stochastic simulation of international economic variables and asset returns: The Falcon Asset Model*", Proceedings of the 8th international AFIR colloquium (1998), Institute of Actuaries, London, 29-45.

Dempster, M., Germano, M., Medova, E., and M. Villaverde "*Global asset liability management*", British Actuarial Journal, 9 (2003), 137-216.

Dominguez-Ballesteros, B. "*The modelling and analysis of mathematical programming problems: tools and applications*", PhD thesis (2001), Department of Mathematics and Statistics, Brunel University.

Dumas, B. "*Short- and Long-term Hedging for the Corporation*", Discussion Paper No. 1083 (1994), Centre for Economic Policy Research, www.cepr.org/pubs/dps/DP1083.asp.

Dupacova, J., Consigli, G., and S.W. Wallace "*Scenarios for multistage stochastic programs*", Annals of Operations Research, 100 (2000), 25-53.

Eaker, M. R., and D. Grant. "*Optimal Hedging of Uncertain and Long-term Foreign Exchange Exposure.*" Journal of Banking and Finance, 9 (June 1985), 221-231.

Ederington, L. "*The Hedging Performance of the New Futures Markets.*" The Journal of Finance, Vol. 34 (Mar., 1979), 157-170.

Eun, C., and B. Resnick "*International diversification of investment portfolios: U.S. and Japanese perspectives*", Management Science, 40-1 (1994), 140-161.

Fleishman, A. "A method for simulating nonnormal distributions", Psychometrica, 43 (1978), 521-532.

Fourer, R., Gay, D., and B. Kernighan "*AMPL: A Modeling Language for Mathematical Programming*", Brooks/Cole – Thomson Learning, 2003.

Golub, B., Holmer, M., McKendall, R., Pohlman, L., and S. Zenios, "*A Stochastic Programming Model for Money Management.*" European Journal of Operational Research, 85 (1995), 282-296.

Gondzio, J., Kouwenberg, R., and T. Vorst. "*Hedging options under transaction costs and stochastic volatility*", Journal of Economic Dynamics and Control, 27 (2003), 1045-1068.

Grimmett, G., and D. Welsh "*Probability: An Introduction*", Oxford University Press (2000).

Gulpinar, N., Rustem, B., and R. Settrgen "*Simulation and optimisation approaches to scenario tree generation*", Journal of Economic Dynamics & Control, 28 (2004), 1291-1315.

Halling, M., Popa, C., and O. Randl "*Stochastic Optimiztion, Tree Structure and Portfolio Choice*", Working Paper of the University of Vienna, May 2005.

Hansen, P., and B. Jaumard "*Cluster analysis and mathematical programming*", Mathematical Programming, 79 (1997), 191-215.

Harrison, M., and D. Kreps "*Martingales and arbitrage in multiperiod securities markets*", Journal of Economic Theory, 20 (1979), 381-408.

Herzel, S. "*Arbitrage opportunities on derivatives: A linear programming approach*". Technical report (2000), Department of Economics, University of Perugia.

Hirschleifer, D. "*Risk, Futures Pricing, and the Organization of Production in Commodity Markets.*" Journal of Political Economy, 96 (Dec. 1988), 1206-1220.

Hochreiter, R., and G. Ch. Pflug "*Scenario Tree Generation as a Multidimensional Facility Location Problem*", Technical Report AURORA TR 2002-33. University of Vienna, 2002.

Hoyland, K., Kaut, M., and S. Wallace "*A Heuristic for Moment-matching Scenario Generation*", Computational Optimization and Applications, 24 (2-3) (2003), 169-185.

Hoyland, K., and S.W. Wallace "*Generating scenario trees for multistage problems*", Management Science, 47 (2) (2001), 295-307.

Hull, J., and A. White "*The Pricing of Options on Assets with Stochastic Volatilities*", Journal of Finance, 42 (1987), 218-300.

Infanger, G. "*Planning under uncertainty*", Boyd & Fraser Publishing Company, 1994.

Infanger, G. "*Dynamic asset allocation strategies using a stochastic dynamic programming approach*", to appear in *Handbook of Asset and Liability Management*, S. Zenios and W.T. Ziemba (Eds.), Elsevier, Amsterdam, 2006.

Ingersoll, J.E. Jr. "*Theory of financial decision making*", 1987, Studies in Financial Economics, Rowan & Littlefield, New Jersey.

Jamshidian, F., and Y. Zhu "*Scenario simulation: theory and methodology*", Finance and Stochastics Journal, Springer Verlag, 1 (1997), 43-67.

Jorion, P. "*Mean-variance analysis of currency overlays*", Financial Analyst Journal, May/June (1994), 48-56.

JP Morgan, "*RiskMetrics*", Technical Document, 4 ed., 1996.

Kaut, M., and S. Wallace "*Evaluation of scenario-generation methods for stochastic programming*", Working Paper (2003).

Keefer, D. "*Certainty equivalents for three-point discrete distribution approximations*", Management Science, 40 (1994), 760-773.

Keefer, D., and S. Bodily "*3-Point approximations for continuous random variables*", Management Science, 29 (1983), 595-609.

Keefer, D., and W. Verdini "*Comparison of Simple Approximations in Estimating Means and Variances of Beta Distributions: Detailed Results*". Technical Report DIS 90/91-1, Department of Decision and Information Systems, College of Business, Arizona State University, 1990.

Kerkvliet, J., and M.H.Moffett "*The Hedging of an Uncertain Future Foreign Currency Cash Flow.*" The Journal of Financial and Quantitative Analysis, Vol. 26 (Dec. 1991), 565-578.

Klaassen, P., Shapiro J.F., and D.E. Spitz "*Sequential Decision Models for Selecting Currency Options.*" Technical Report IFSRC No. 133-90, Massachusetts Institute of Technology, International Financial Services Research Centre, July 1990.

Klaassen, P. "*Discretized reality and spurious profits in stochastic programming models for asset / liability management*", EJOR 101 (1997), 374-392.

Klaassen, P. "*Financial asset-pricing theory and stochastic programming models for asset-liability management: A synthesis*", Management Science, 44 (1998), 31-48.

Koskosidis, Y., and A. Duarte "*A scenario-based approach to active asset allocation*", Journal of Portfolio Management, Winter (1997), 74-85.

Kouwenberg, R. "*Scenario generation and stochastic programming models for asset liability management*", European Journal of Operational Research, 134 (2001), 279-292.

Kouwenberg, R., and S. Zenios "*Stochastic programming models for asset liability management*", Working paper 01-01, "Hermes" European Centre of Excellence on

Computational Finance and Economics, School of Economics and Management, Univercity of Cyprus, Nicosia, 2001.

Kouwenberg, R., and T. Vorst *"Dynamic portfolio insurance: a stochastic programming approach"*, Working Paper (1998).

Kusy, M.I., and W.T. Ziemba *"A Bank Asset and Liability Management Model."* Operations Research", 34 (3) (1986), 356-376.

Kwok, C. *"Hedging Foreign Exchange Exposures: Independent vs. Integrative Approaches."* Journal of International Business Studies, 18 (Summer 1987), 33-51.

Loretan, M. *"Generating market risk scenarios using principal components analysis: Methodological and practical considerations"*, In *The Measurement of Aggregate Market Risk, CGFS Publications No. 7*, pages 23-60. Bank for International Settlements, November 1997. Available at http://www.bis.org/publ/ecsc07.htm.

Luenberger, D., *"Investment Science"*, Oxford University Press, 1998.

Lurie, P., and M. Goldberg "An approximate method for sampling correlated random variables from partially-specified distributions", Management Science, 44 (2) (1998), 203-218.

Markowitz, H. *"Portfolio Selection: Efficient Diversification of Investment"*, 1959, Yale University Press, New Haven, CT.

Maurer, R., and S. Valiani *"Forward versus options for hedging the currency exposure risk: an application to international portfolio selection"*, School of Businss and Economics, Goethe-University Frankfurt/M, Working paper 109, June 2003.

McKay, M., Conover, W., and R. Beckman *"A comparison of three methods for selecting input variables in the analysis of output from a computer code"*, Technometrics, 21 (1979), 239-245.

McNamee, P., and J. Celona "*Decision Analysis for the Professional with Supertree*". The Scientific Press, Redwood City, CA, 1987.

Merton, R. "*Continuous-time finance*", Blackwell Publishers, 1990.

Miller, A.C., and T.R. Rice "*Discrete approximations of probability distributions*" Management Science, 29 (1983), 352-362.

Mulvey, J., and A. Thorlacius "*The Towers Perrin global capital market scenario generation system*", In Ziemba & Mulvey (1998), 286-314.

Mulvey, J., and H. Vladimirou "*Stochastic network optimisation models for investment planning*", Annals of Operational Research, 20 (1989), 187-217.

Naik, V. "*Finite State Securities Markets Models and Arbitrage*", in Handbooks in Operations Research and Management Science, Vol.9 (1995) Finance, R. Jarrow et al. Eds., Amsterdam: North Holland.

Perold, A., and E. Schulman "*The free lunch in currency hedging: Implications for investment policy and performance standards*", Financial Analysts Journal, 44 (1988), 45-50.

Pflug G. "*Scenario tree generation for multiperiod financial optimization by optimal discretization*" Math. Programming, Ser. B 89 (2001), 251-257.

Pliska, S.R. "*Introduction to Mathematical Finance: Discrete Time Models*", 1997, Blackwell.

Poojari C., Lucas C., and G. Mitra. "*A Decision Model for Natural Oil Buying Policy under Uncertainty*" in: Proceding of Industrial Mat. Conference. Editors: M. Joshi and A. Pani, Springer Verlag (2004).

Rachev, S. "*Probability metrics and the stability of stochastic models*", 1991, J. Wiley and Sons, New York.

Rolfo, J. "*Optimal Hedging under Price and Quantity Uncertainty: The Case of a Cocoa Producer.*" Journal of Political Economy, 88 (Feb. 1980), 100-116.

Rubinstein, M. "*Implied Binomial Trees*", Journal of Finance, 49 (1994), 771-818.

Shapiro, A.C. "*Currency Risk and Relative Price Risk.*" Journal of Financial and Quantitative Analysis, 19 (Dec. 1984), 365-373.

Sharda, R., and Musser K. "*Financial Futures Hedging via Goal Programming.*" Management Science, (August 1986), 933-947.

Sharda, R., and Wingender, J. "*Multiobjective Approach to Hedging with Foreign Exchange Futures.*" Advances in Mathematical Programming and Financial Planning, Vol. 3 (1993), 193-209.

Sims, C.A. "*Macro-economics and reality*", Econometrica, 48 (1980), 1-48.

Smith, J. "*Moment methods for decision analysis*", Management Science, 39 (1993), 340-358.

Stigliz, J.E. "*Futures Markets and Risk: A General Equilibrium Approach.*" In Futures Markets: Modelling Managing, and Monitoring Futures Trading. Manfred Streit, ed. Oxford: Basil Blackwell (1983), 75-106.

Swanson, P.E., and Caples, S.C. "*Hedging Foreign Exchange Risk Using Forward Foreign Exchange Markets: An Extension.*" Journal of International Business Studies. 18 (Spring 1987), 75-82.

Thorlacius, A. "*Arbitrage in asset modelling for integrated risk management*", Swiss Re Investors, 1998. web link: www.actuaries.org/AFIR/colloquia/Tromsoe/Thorlacius.pdf.

Taylor, F. "*Mastering Foreign Exchange and Currency Options: A Practical Guide to the New Marketplace*" 2003. Financial Times Prentice Hall (2-nd edition).

Topaloglou, N., Vladimirou, H., and S. Zenios "*CVaR models with selective hedging for international asset allocation*", Journal of Banking and Finance, 26(7) (2002), 1535-1561.

Topaloglou, N., Vladimirou, H., and S. Zenios "*Dynamic stochastic programming models for international portfolio management*", Technical report, "Hermes" European Centre of Excellence on Computational Finance and Economics, School of Economics and Management, Univercity of Cyprus, Nicosia, 2004(a).

Topaloglou, N., Vladimirou, H., and S. Zenios "*Risk management for international investment portfolios using forward contracts and options*", Technical report, "Hermes" European Centre of Excellence on Computational Finance and Economics, School of Economics and Management, Univercity of Cyprus, Nicosia, 2004(b).

Uryasev, S., and R. Rockafellar "*Optimization of conditional value-at-risk*", Research report 99-4, University of Florida, Dept. of Industrial and Systems Engineering.

Vale, C. and V. Maurelli "Simulating multivariate nonnormal distributions", Psychometrica, 48 (3) (1983), 465-471.

Vitoriano, B., Cerisola, S., and A. Ramos "*Generating scenario trees for hydro inflows*", Working Paper (2001).

Volosov, K., Mitra, G., Spagnolo, F., and C. Lucas "*Treasury Management Model with Foreign Exchange Exposure", Computational Optimization and Applications, 32 (2005), 179-207.*

Wingender J., and Sharda R. "*A Goal Programming Approach for Hedging a Portfolio with Financial Futures: an Empirical Test*" Advances in Mathematical Programming and Financial Planning, Vol. 4 (1995), 223-249.

Wu, J. and Sen S. "*A Stochastic Programming Model for Currency Option Hedging.*" Annals of Operations Research (2000), 100, 227-250.

Zellner, A. "*An efficient method of estimating seemingly unrelated regression models*", Journal of the American statistical Association, 57 (1962), 348-368.

Zenios S., and M. Shtilman "*Constructing Optimal Samples from a Binomial Lattice*", Journal of Information & Optimization Sciences, 14 (1993), No. 2, 125-147.

Appendices

Appendix A Discrete approximations of probability distributions

The following discussion as well as the notation used is taken from Miller & Rice (1983).

Notation

$\{x\}$: denotes the probability density function (for a continuous random variable) or the probability mass function (for a discrete random variable) of the random variable x;

$\langle x \rangle = \bar{x} = \int_{-\infty}^{\infty} x\{x\}dx$: is the expected value of x; (A.1)

$\{x\}^{\leq} = \int_{-\infty}^{x} \{x\}dx$: is the cumulative distribution of x; (A.2)

$\langle g(x) \rangle = \int_{-\infty}^{\infty} g(x)\{x\}dx$ is the expected value of the function $g(x)$. (A.3)

Measuring the Accuracy of Discrete Approximations

The accuracy of approximation of a continuous probability distribution by a discrete distribution is measured by the extent to which statistical moment of the discrete distribution match those of the original continuous distribution. To show this, let us assume we can approximate expected value of an arbitrary function $g(x)$ as follows:

$$\int_{-\infty}^{\infty} g(x)\{x\}dx \cong \sum_{i=1}^{N} p_i g(x_i), \qquad \text{(A.4)}$$

where the probability distribution $\{x\}$ is represented by a set of discrete values x_i and their respective probabilities p_i.

Assuming that the function $g(x)$ can be approximated as a polynomial the equation (A.4) can be rewritten as:

$$\int_{-\infty}^{\infty}\left(a_0 + a_1 x + a_2 x^2 + \cdots\right)\{x\}dx \cong \sum_{i=1}^{N} p_i\left(a_0 + a_1 x_i + a_2 x_i^2 + \cdots\right). \qquad \text{(A.5)}$$

Equation (A.5) can be rewritten in terms of original and approximate moments:

$$a_0 + a_1\langle x\rangle + a_2\langle x^2\rangle + \cdots \cong a_0\sum_{i=1}^{N} p_i + a_1\sum_{i=1}^{N}(p_i x_i) + a_2\sum_{i=1}^{N}(p_i x_i^2) + \cdots. \qquad \text{(A.6)}$$

The equation (A.6) is satisfied for all coefficients $(a_0, a_1, a_2, \ldots)$, provided the moments of the approximation equal the moments of the original distribution such that:

$$\langle x^k\rangle = \sum_{i=1}^{N} p_i x_i^k \qquad \text{for } k = 0,1,2,\ldots \qquad \text{(A.7)}$$

Thus equation (A.7) is the criterion for the accuracy of the approximation of a continuous probability distribution to be a discrete distribution.

The Accuracy of Typical Approximations

It is common to approximate continuous probability distribution by dividing the range of values of a cumulative distribution into mutually exclusive intervals. Then each interval in the partition of the range of the cumulative distribution will have a corresponding interval in the corresponding partition of the domain of the cumulative distribution. Each interval of the domain can be approximated by its mean or median. Figure A-1 illustrates the partition of the domain and range of the cumulative probability distribution into four intervals with corresponding four mean values of each interval.

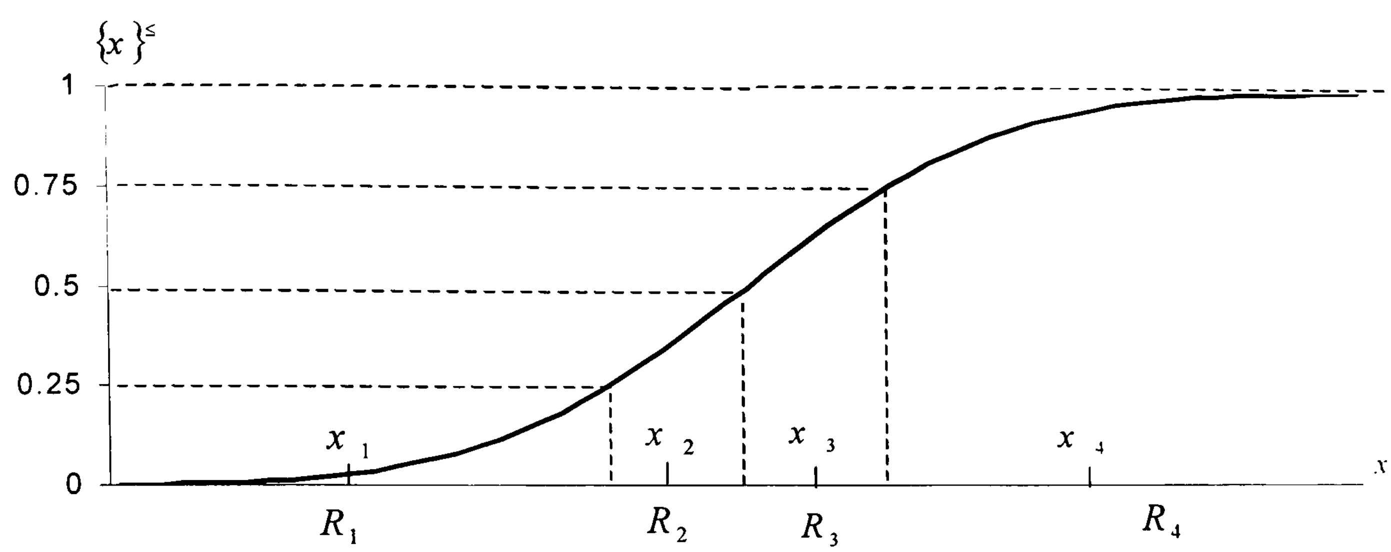

Figure A-1 Approximation of the Probability Distribution by the Mean Values of Equally Likely Intervals.

The probability of each mean value x_i is assumed to be the probability of the true values to fall within the interval R_i.

The above approximation, using means and probabilities of a set of intervals, will always underestimate all statistical moments (including central moments) of the original distribution except for the first moment (mean value). To show why this happens we divide the range of values of the random variable x, into several intervals: $R_1, R_2, \ldots, R_N$. The means of each such interval and their respective probabilities, which are used to approximate the original distribution, are given as:

$$p_i = \int_{R_i} \{x\} dx \qquad \text{for } i = 1, 2, \ldots, N, \qquad \text{(A.8)}$$

$$x_i = \int_{R_i} x \frac{\{x\}}{p_i} dx \qquad \text{for } i = 1, 2, \ldots, N. \qquad \text{(A.9)}$$

For an arbitrary function $g(x)$:

$$\langle g(x) \rangle = \int_{-\infty}^{\infty} g(x)\{x\}dx = \sum_{i=1}^{N} \int_{R_i} g(x)\{x\}dx = \sum_{i=1}^{N} p_i \int_{R_i} g(x)\frac{\{x\}}{p_i} dx. \qquad \text{(A.10)}$$

If $g(x)$ is a *convex* function:

$$\int_{R_i} g(x)\frac{\{x\}}{p_i}dx > g\left(\int_{R_i} x\frac{\{x\}}{p_i}dx\right) = g(x_i) \qquad \text{(A.11)}$$

Substituting inequality (A.11) into equation (A.10) the following inequality is obtained:

$$\langle g(x)\rangle > \sum_{i=1}^{N} p_i g(x_i) \qquad \text{(A.12)}$$

Let $g(x) = (x-a)^n$ be a convex function, where a is any constant and n is an even positive integer. Then, *even* statistical moments of a discrete approximation of such a function will always *underestimate* the even statistical moments of the original distribution. In particular, for the second moments and second central moments (variance) we have the following inequalities:

$$\langle x^2\rangle > \sum_{i=1}^{N} p_i x_i^2, \qquad \langle (x-\bar{x})^2\rangle > \sum_{i=1}^{N} p_i (x_i - \bar{x})^2 \qquad \text{(A.13)}$$

Using the same logic as used above one can show that for *odd* moments the function $g(x) = (x-a)^n$ is convex for $(x-a) \geq 0$ and hence the moments of the discrete approximation will underestimate the moments of the original distribution. For $(x-a) \leq 0$ the function will be concave and the moments of the discrete approximation will overestimate the moments of the original distribution.

An Approximation Procedure Based on Gaussian Quadrature

There is an alternative method to approximate a distribution function, which is more accurate than the method based on equally spaced intervals discussed above. This method uses gaussian quadrature of numerical approximation of integrals. This method approximates an integral of the product of a function $g(x)$ and a weighting function, for example $\{x\}$, by the sum over a finite number of points x_i of the function $g(x_i)$ evaluated at these points and the weights assigned to these points, say p_i:

$$\int_a^b g(x)\{x\}dx = \sum_{i=1}^{N} g(x_i)p_i \qquad \text{(A.14)}$$

In order to achieve the equality in equation (A.14) we will approximate the function $g(x)$ as polynomial and find such values of x_i and p_i that the approximation (A.14) is satisfied for each term (replacing $g(x)$) of the polynomial:

$$\langle x^k \rangle = \int_{-\infty}^{\infty} x^k \{x\}dx = \sum_{i=1}^{N} x_i^k p_i \qquad \text{for } k = 0, 1, 2, \ldots \qquad \text{(A.15)}$$

In order to match (2N – 1) moments exactly the original distribution function should be approximated by N pairs of x_i and p_i that satisfy the following system of equations:

$$\begin{aligned}
&p_1 + p_2 + p_3 + \cdots + p_N = \langle x^0 \rangle = 1 \\
&p_1x_1 + p_2x_2 + p_3x_3 + \cdots + p_Nx_N = \langle x \rangle \\
&p_1x_1^2 + p_2x_2^2 + p_3x_3^2 + \cdots + p_Nx_N^2 = \langle x^2 \rangle \\
&\vdots \\
&p_1x_1^{2N-1} + p_2x_2^{2N-1} + p_3x_3^{2N-1} + \cdots + p_Nx_N^{2N-1} = \langle x^{2N-1} \rangle
\end{aligned} \qquad \text{(A.15)}$$

It can be shown that if the original moments are the moments of the probability distribution and are finite then solving system (A.15) will yield N distinct and real-valued x_i, which all lie in the interval spanned by the original distribution. All probabilities p_i will be positive.

The problem becomes slightly more complicated when the original distribution is based on subjective estimates such that its forms and statistical moments are not known. Usually the cumulative distribution function can be assessed in a graph in a similar fashion to Figure A-1. In such cases the gaussian quadrature approach is applied in a two-step procedure:

1. Use gaussian quadrature to determine the moments of the subjective continuous distribution

2. Use gaussian quadrature to approximate the original distribution with the moments determined in step 1 by a discrete distribution.

Appendix B   Cubic transformation

Cubic transformation can be expressed as:

$$Y = a + bX + cX^2 + dX^3, \tag{B.1}$$

where Y is a non-normal random variable (or vector) with specified (target) four statistical moments; X is any arbitrary random variable (or vector).

The purpose of cubic transformation is to generate univariate random variable (or a vector of random variables) Y with required 4 statistical moments from realisations of another random variable (or vector) X with known first 12 statistical moments.

The main task of the transformation is to find the transform coefficients: a, b, c and d that satisfy the system of non-linear equations. These non-linear equations equate the moments of Y, $E[Y^i]$ and functions of moments of X, $E[X^i]$:

$$E[Y] = a + bE[X] + cE[X^2] + dE[X^3]$$

$$\begin{aligned} E[Y^2] &= d^2E[X^6] + 2cdE[X^5] + (2bd + c^2)E[X^4] + (2ad + 2bc)E[X^3] \\ &+ (2ac + b^2)E[X^2] + 2abE[X] + a^2 \end{aligned} \tag{B.2}$$

$$\begin{aligned} E[Y^3] &= d^3E[X^9] + 3cd^2E[X^8] + (3bd^2 + 3c^2d)E[X^7] + (3ad^2 + 6bcd + c^3)E[X^6] \\ &+ (6acd + 3b^2d + 3bc^2)E[X^5] + (a(6bd + 3c^2) + 3b^2c)E[X^4] + (3a^2d + 6abc + b^3)E[X^3] \\ &+ (3a^2c + 3ab^2)E[X^2] + 3a^2bE[X] + a^3 \end{aligned}$$

$$\begin{aligned} E[Y^4] &= d^4E[X^{12}] + 4cd^3E[X^{11}] + (4bd^3 + 6c^2d^2)E[X^{10}] + (4ad^3 + 12bcd^2 + 4c^3d)E[X^9] \\ &+ (12acd^2 + 6b^2d^2 + 12bc^2d + c^4)E[X^8] + (a(12bd^2 + 12c^2d) + 12b^2cd + 4bc^3)E[X^7] \\ &+ (6a^2d^2 + a(24bcd + 4c^3) + 4b^3d + 6b^2c^2)E[X^6] + (12a^2cd + a(12b^2d + 12bc^2) + 4b^3c)E[X^5] \\ &+ (a^2(12bd + 6c^2) + 12ab^2c + b^4)E[X^4] + (4a^3d + 12a^2bc + 4ab^3)E[X^3] + (4a^3c + 6a^2b^2)E[X^2] \\ &+ 4a^3bE[X] + a^4 \end{aligned}$$

Appendix C Matrix transformation and its properties

The contents of this appendix follow the lines of the respective appendix from Hoyland *et al.* (2003). Matrix transformation is applied to a vector of random independent variables, X, to obtain another random vector, Y, with required (target) correlations, R, among the elements of the resulting random vector and can be expressed as follows:

$$Y = LX,$$

where L is the lower-triangular matrix obtained from decomposing the correlation matrix such as: $R = LL^T$.

If we make the following assumptions about the random vector X:

1. X is an n-dimensional random vector with independent elements, i.e. any X_i and X_j are independent for $i \neq j$;

2. $E\lfloor X^k \rfloor$ exists for $k = 1, 2, 3, 4$;

3. $E[X] = 0$ and $E\lfloor X^2 \rfloor = 1$,

and if the random vector Y is defined as $Y = LX$, then Y will have the following properties:

1. $E\lfloor Y^k \rfloor$ exists for $k = 1, 2, 3, 4$;

2. $E[Y] = 0$ and $E\lfloor Y^2 \rfloor = 1$;

3. correlation matrix of Y is $R = LL^T$;

4. $E\left[Y_i^3\right]=\sum_{j=1}^{i} L_{ij}^3 E\left[X_j^3\right];$

5. $E\left[Y_i^4\right]-3=\sum_{j=1}^{i} L_{ij}^4\left(E\left[X_j^4\right]-3\right).$

If we assume that correlation matrix $R = LL^T$ is positive-definite and hence the diagonal elements of the lower-triangular matrix L are positive then the moments of X can be expressed as follows:

6. $E\left[\widetilde{X}_i^3\right]=\frac{1}{L_{ii}^3}\left(E\left[\widetilde{Y}_i^3\right]-\sum_{j=1}^{i-1} L_{ij}^3 E\left[\widetilde{X}_j^3\right]\right)$

7. $E\left[\widetilde{X}_i^4\right]-3=\frac{1}{L_{ii}^4}\left(E\left[\widetilde{Y}_i^4\right]-3-\sum_{j=1}^{i-1} L_{ij}^4\left(E\left[\widetilde{X}_j^4\right]-3\right)\right)$

Appendix D Exchange rates and their first differences used in the case study.

Table 3 Quarterly spot and forward exchange rates

Date	Spot	3Fwd	6Fwd	*9Fwd*	12Fwd
01/01/1998	1.6537	1.6464	1.6396	1.625	1.6279
01/04/1998	1.6725	1.6651	1.6579	1.65472	1.6466
01/07/1998	1.6623	1.6536	1.6452	1.63405	1.6293
01/10/1998	1.7053	1.6972	1.6895	1.6789	1.6762
01/01/1999	1.6637	1.6505	1.6489	1.65745	1.6481
01/04/1999	1.6049	1.6039	1.6043	1.60555	1.6067
01/07/1999	1.5767	1.5778	1.58	1.5823	1.5833
01/10/1999	1.6547	1.6522	1.6515	1.6538	1.6475
01/01/2000	1.6133	1.6133	1.6128	1.62285	1.6117
01/04/2000	1.5977	1.5982	1.5991	1.59855	1.6018
01/07/2000	1.5127	1.5138	1.5183	1.5222	1.5239
01/10/2000	1.4663	1.4688	1.4707	1.4754	1.4737
01/01/2001	1.4947	1.4967	1.4977	1.4976	1.4988
01/04/2001	1.4209	1.419	1.4168	1.4154	1.4138
01/07/2001	1.4171	1.4122	1.407	1.4005	1.3978
01/10/2001	1.4782	1.4714	1.4644	1.4582	1.4528
01/01/2002	1.4531	1.4455	1.4382	1.43355	1.4253
01/04/2002	1.4239	1.4164	1.4097	1.4186	1.3994
01/07/2002	1.531	1.5225	1.5135	1.5031	1.4965
01/10/2002	1.5691	1.5607	1.5522	1.54445	1.536
01/01/2003	1.6099	1.5998	1.5901	1.5805	1.5716

Table 4 First differences of quarterly spot and forward exchange rates

Date	Spot_Diff	3Fwd_Diff	6Fwd_Diff	9Fwd_Diff	12Fwd_Diff
01/04/1998	0.0188	0.0187	0.0183	0.02972	0.0187
01/07/1998	-0.0102	-0.0115	-0.0127	-0.02067	-0.0173
01/10/1998	0.043	0.0436	0.0443	0.04485	0.0469
01/01/1999	-0.0416	-0.0467	-0.0406	-0.02145	-0.0281
01/04/1999	-0.0588	-0.0466	-0.0446	-0.0519	-0.0414
01/07/1999	-0.0282	-0.0261	-0.0243	-0.02325	-0.0234
01/10/1999	0.078	0.0744	0.0715	0.0715	0.0642
01/01/2000	-0.0414	-0.0389	-0.0387	-0.03095	-0.0358
01/04/2000	-0.0156	-0.0151	-0.0137	-0.0243	-0.0099
01/07/2000	-0.085	-0.0844	-0.0808	-0.07635	-0.0779
01/10/2000	-0.0464	-0.045	-0.0476	-0.0468	-0.0502
01/01/2001	0.0284	0.0279	0.027	0.0222	0.0251
01/04/2001	-0.0738	-0.0777	-0.0809	-0.0822	-0.085
01/07/2001	-0.0038	-0.0068	-0.0098	-0.0149	-0.016
01/10/2001	0.0611	0.0592	0.0574	0.0577	0.055
01/01/2002	-0.0251	-0.0259	-0.0262	-0.02465	-0.0275
01/04/2002	-0.0292	-0.0291	-0.0285	-0.01495	-0.0259
01/07/2002	0.1071	0.1061	0.1038	0.0845	0.0971
01/10/2002	0.0381	0.0382	0.0387	0.04135	0.0395
01/01/2003	0.0408	0.0391	0.0379	0.03605	0.0356

Appendix E Algebraic formulation of moment matching scenario generating optimisation model.

SETS:
P set of predecessors
S set of scenarios
A set of assets
L set of links, where $L = \{(a_1, a_2) \ / \ a_1, a_2 \in A, \ a_1 \neq a_2 \ \wedge \ \text{ord}(a_1) < \text{ord}(a_2)\}$

PARAMETERS:
Mean_tar(a):
$YT_a \ \forall a \in A$
Variance_tar(a):
$VT_a \ \ \forall a \in A$
Skewness_tar(a):
$WT_a \ \forall a \in A$
Kurtosis_tar(a):
$KT_a \ \forall a \in A$
Covariance(l):
$ZT_l \ \ \forall l \in L$
Prob_ind:
$PROBIND$
Prob:
$PROB$
mean_tar_ind(a,p):
$YTI_{a,p} \ \ \forall a \in A, \forall p \in P$
var_tar_ind(a,p):
$VTI_{a,p} \ \ \forall a \in A, \forall p \in P$

VARIABLES:
X(a,p,s):
$X_{a,p,s} \ \forall a \in A, \forall p \in P, \forall s \in S$
mean(a):
$$Y_a = \sum_{p \in P} \sum_{s \in S} X_{a,p,s} * PROB \quad \forall a \in A$$
variance(a):
$$V_a = \sum_{p \in P} \sum_{s \in S} \left[\left(X_{a,p,s} - Y_a \right)^2 \right] * PROB \quad \forall a \in A$$
skewness(a):
$$W_a = \sum_{p \in P} \sum_{s \in S} \left[\left(X_{a,p,s} - Y_a \right)^3 \right] * PROB \quad \forall a \in A$$
kurtosis(a):

$$K_a = \sum_{p \in P} \sum_{s \in S} \left[\left(X_{a,p,s} - Y_a \right)^4 \right] * PROB \quad \forall a \in A$$

Covariance(a1,a2):

$$Z_l = \sum_{p \in P} \sum_{s \in S} \left(X_{a_1,p,s} - Y_{a_1} \right) * \left(X_{a_2,p,s} - Y_{a_2} \right) * PROB \quad \forall l \in L, l = (a_1, a_2)$$

mean_ind(a,p):

$$YIND_{a,p} = \sum_{s \in S} X_{a,p,s} * PROBIND$$

variance_ind(a,p):

$$VIND_{a,p} = \sum_{s \in S} \left[\left(X_{a,p,s} - YIND_{a,p} \right)^2 \right] * PROBIND \quad \forall a \in A, \forall p \in P$$

<u>OBJECTIVES:</u>

Diff_MomCov_initial:

$$DMI = \min \sum_{a \in A} \left(Y_a - YT_a \right)^2 + \sum_{a \in A} \left(V_a - VT_a \right)^2 + \sum_{a \in A} \left(W_a - WT_a \right)^2 + \sum_{l \in L} \left(Z_l - ZT_l \right)^2$$

Diff_MomCov_ind:

$$\begin{aligned} DMIND = \min & \sum_{a \in A} \left(Y_a - YT_a \right)^2 + \sum_{a \in A} \left(V_a - VT_a \right)^2 + \sum_{a \in A} \left(W_a - WT_a \right)^2 + \sum_{a \in A} \left(K_a - KT_a \right)^2 \\ & + \sum_{l \in L} \left(Z_l - ZT_l \right)^2 \\ & + \sum_{a \in A} \sum_{p \in P} \left(YIND_{a,p} - YTI_{a,p} \right)^2 + \sum_{a \in A} \sum_{p \in P} \left(VIND_{a,p} - VTI_{a,p} \right)^2 \end{aligned}$$

Diff_MomCov_Final:

$$\begin{aligned} DMF = \min & \sum_{a \in A} \left(Y_a - YT_a \right)^2 + \sum_{a \in A} \left(V_a - VT_a \right)^2 + \sum_{a \in A} \left(W_a - WT_a \right)^2 + \sum_{a \in A} \left(K_a - KT_a \right)^2 \\ & + \sum_{a \in A} \sum_{p \in P} \left(YIND_{a,p} - YTI_{a,p} \right)^2 + \sum_{a \in A} \sum_{p \in P} \left(VIND_{a,p} - VTI_{a,p} \right)^2 \end{aligned}$$

Appendix F Model parameters and sets over time.

- *T = 3 (tree months from the initial time period)*

Predecessors = {1}

Scenarios = {1, 2,…, 6}

Prob = Prob_ind = 1/6

assets	mean_tar	variance_tar	skewness_tar	kurtosis_tar
Spot	1.59980	0.00272	0.00005	0.00002
3Fwd	1.59010	0.00262	0.00005	0.00001
6Fwd	1.58050	0.00253	0.00004	0.00001
9Fwd	1.57160	0.00228	0.00002	0.00001

assets	assets	covariance_tar
Spot	3Fwd	0.002533
Spot	6Fwd	0.002482
Spot	9Fwd	0.002326
3Fwd	6Fwd	0.002445
3Fwd	9Fwd	0.002283
6Fwd	9Fwd	0.002250

- *T = 6 (six months from the initial time period)*

Predecessors = {1, 2,…, 6}

Scenarios = {1, 2, 3}

Prob = 1/18

Prob_ind = 1/3

assets	mean_tar	variance_tar	skewness_tar	kurtosis_tar
Spot	1.59010000	0.00544032	0.00004865	0.00012194
3Fwd	1.58050000	0.00524867	0.00004504	0.00011599
6Fwd	1.57160000	0.00505558	0.00004078	0.00010770

assets	assets	covariance_tar
Spot	3Fwd	0.00253
Spot	6Fwd	0.00248
3Fwd	6Fwd	0.00244

assets	predecessors	mean_tar_ind	variance_tar_ind
Spot	1	1.585866272	0.00272016
Spot	2	1.585032124	0.00272016
Spot	3	1.551702234	0.00272016
Spot	4	1.620916662	0.00272016
Spot	5	1.678790315	0.00272016
Spot	6	1.516384949	0.00272016
3Fwd	1	1.560504083	0.00262434
3Fwd	2	1.599121342	0.00262434
3Fwd	3	1.521190459	0.00262434
3Fwd	4	1.610797786	0.00262434
3Fwd	5	1.663880145	0.00262434
3Fwd	6	1.525609871	0.00262434
6Fwd	1	1.534818064	0.00252779
6Fwd	2	1.580094153	0.00252779
6Fwd	3	1.54076689	0.00252779
6Fwd	4	1.620483437	0.00252779
6Fwd	5	1.642402277	0.00252779
6Fwd	6	1.509149954	0.00252779

- *T = 9 (nine months from the initial time period)*

Predecessors = {1, 2,…, 18}

Scenarios = {1, 2, 3}

Prob = 1/54

Prob_ind = 1/3

assets	mean_tar	variance_tar	skewness_tar	kurtosis_tar
Spot	1.5805	0.00816048	0.00004865	0.00041154
3Fwd	1.5716	0.00787301	0.00004504	0.00039148

assets	assets	covariance_tar
Spot	3Fwd	0.002532752

assets	predecessors	mean_tar_ind	variance_tar_ind
Spot	1	1.580209443	0.00272016
Spot	2	1.611286429	0.00272016
Spot	3	1.490150091	0.00272016
Spot	4	1.546330647	0.00272016
Spot	5	1.668757506	0.00272016
Spot	6	1.582412891	0.00272016
Spot	7	1.45543064	0.00272016
Spot	8	1.527461034	0.00272016
Spot	9	1.580810116	0.00272016
Spot	10	1.576745395	0.00272016
Spot	11	1.572344258	0.00272016
Spot	12	1.683441762	0.00272016
Spot	13	1.604943968	0.00272016
Spot	14	1.730137226	0.00272016
Spot	15	1.656701748	0.00272016
Spot	16	1.592491927	0.00272016
Spot	17	1.46757527	0.00272016
Spot	18	1.516893201	0.00272016
3Fwd	1	1.555450788	0.00262434
3Fwd	2	1.465174831	0.00262434
3Fwd	3	1.583955448	0.00262434
3Fwd	4	1.536287403	0.00262434
3Fwd	5	1.651078113	0.00262434
3Fwd	6	1.553053471	0.00262434
3Fwd	7	1.487991615	0.00262434
3Fwd	8	1.525361556	0.00262434
3Fwd	9	1.609075729	0.00262434
3Fwd	10	1.621947499	0.00262434
3Fwd	11	1.681818663	0.00262434
3Fwd	12	1.55782926	0.00262434
3Fwd	13	1.660098672	0.00262434
3Fwd	14	1.693728849	0.00262434
3Fwd	15	1.573529003	0.00262434
3Fwd	16	1.472659865	0.00262434
3Fwd	17	1.474123339	0.00262434
3Fwd	18	1.580788135	0.00262434

- *T = 12 (twelve months from the initial time period)*

Predecessors = {1, 2,..., 54}

Scenarios = {1, 2, 3}

Prob = 1/162

Prob_ind = 1/3

assets	mean_tar	variance_tar	skewness_tar	kurtosis_tar
Spot	1.5716	0.0108806	0.0000486	0.0019510

assets	predecessors	mean_tar_ind	variance_tar_ind
Spot	1	1.59804109	0.0027202
Spot	2	1.585247438	0.0027202
Spot	3	1.483104274	0.0027202
Spot	4	1.482404735	0.0027202
Spot	5	1.395351433	0.0027202
Spot	6	1.517796991	0.0027202
Spot	7	1.656712357	0.0027202
Spot	8	1.547284199	0.0027202
Spot	9	1.54791388	0.0027202
Spot	10	1.468648991	0.0027202
Spot	11	1.593292789	0.0027202
Spot	12	1.546958397	0.0027202
Spot	13	1.578658513	0.0027202
Spot	14	1.693042714	0.0027202
Spot	15	1.681586004	0.0027202
Spot	16	1.480960358	0.0027202
Spot	17	1.597454874	0.0027202
Spot	18	1.580785311	0.0027202
Spot	19	1.422443756	0.0027202
Spot	20	1.493464507	0.0027202
Spot	21	1.548098229	0.0027202
Spot	22	1.459431709	0.0027202
Spot	23	1.58494535	0.0027202
Spot	24	1.531744154	0.0027202
Spot	25	1.646676833	0.0027202
Spot	26	1.644238276	0.0027202
Spot	27	1.536359458	0.0027202
Spot	28	1.602958244	0.0027202
Spot	29	1.570642102	0.0027202
Spot	30	1.692291195	0.0027202
Spot	31	1.617907524	0.0027202
Spot	32	1.743892611	0.0027202
Spot	33	1.683713019	0.0027202
Spot	34	1.531571066	0.0027202
Spot	35	1.512228789	0.0027202
Spot	36	1.629728716	0.0027202
Spot	37	1.643466363	0.0027202
Spot	38	1.607104944	0.0027202
Spot	39	1.729778867	0.0027202
Spot	40	1.641709062	0.0027202
Spot	41	1.675720037	0.0027202
Spot	42	1.763816252	0.0027202
Spot	43	1.509439668	0.0027202
Spot	44	1.635366597	0.0027202
Spot	45	1.575823685	0.0027202
Spot	46	1.518116826	0.0027202
Spot	47	1.499146004	0.0027202

Spot	48	1.400746469	0.0027202
Spot	49	1.408939942	0.0027202
Spot	50	1.534686411	0.0027202
Spot	51	1.47877347	0.0027202
Spot	52	1.576114005	0.0027202
Spot	53	1.520280185	0.0027202
Spot	54	1.646013865	0.0027202

Appendix G Algebraic formulation of the decision model.

<u>SETS:</u>

A set of assets

T set of time periods

S set of scenarios

L set of links, where $L = \{(s_1, s_2)/s_1, s_2 \in S\}$

<u>PARAMETERS:</u>

Prob:

$PROB$

Xrate(a,t,s):

$RATE_{a,t,s} \quad \forall a \in A,\ \forall t \in T,\ \forall s \in S$

Revenue(t,s):

$REV_{t,s} \quad \forall t \in T,\ \forall s \in S$

TransCost:

TC

FwdPrev(a):

$FP_a \quad \forall a \in A \backslash_{\{12Fwd\}}$

UpperLimitOnFwd:

UL

<u>VARIABLES:</u>

FwdHold(a,t,s):

$YFH_{a,t,s} \quad \forall a \in A \backslash_{\{Spot\}}, \forall t \in T, \forall s \in S$

FwdBuy(a,t,s):

$YFS_{a,t,s} \quad \forall a \in A \backslash_{\{Spot\}}, \forall t \in T, \forall s \in S$

FwdSell(a,t,s):

$YFB_{a,t,s} \quad \forall a \in A \backslash_{\{Spot\}}, \forall t \in T, \forall s \in S$

Cost(s):

$$Z_s = TC * [\ (YFB_{a='12Fwd',t=0,s} + YFS_{a='12Fwd',t=0,s}$$
$$+ YFB_{a='9Fwd',t=0,s} + YFS_{a='9Fwd',t=0,s}$$
$$+ YFB_{a='6Fwd',t=0,s} + YFS_{a='6Fwd',t=0,s}$$
$$+ YFB_{a='3Fwd',t=0,s} + YFS_{a='3Fwd',t=0,s}) / RATE_{a='Spot',t=0,s}$$

$$+ (YFB_{a='9Fwd',t=3,s} + YFS_{a='9Fwd',t=3,s}$$
$$+ YFB_{a='6Fwd',t=3,s} + YFS_{a='6Fwd',t=3,s}$$
$$+ YFB_{a='3Fwd',t=3,s} + YFS_{a='3Fwd',t=3,s}) / RATE_{a='Spot',t=3,s}$$

$$+ (YFB_{a='6Fwd',t=6,s} + YFS_{a='6Fwd',t=6,s}$$
$$+ YFB_{a='3Fwd',t=6,s} + YFS_{a='3Fwd',t=6,s}) / RATE_{a='Spot',t=6,s}$$

$$+ (YFB_{a='3Fwd',t=9,s} + YFS_{a='3Fwd',t=9,s}) / RATE_{a='Spot',t=9,s}\ ] \qquad \forall s \in S$$

ExpYearlyCost:

$$XC = \sum_{s \in S} PROB * Z_s$$

YearlyRev(s):

$$
\begin{aligned}
XR_s = \; & REV_{t=3,s} / RATE_{a='Spot',t=3,s} \\
& + REV_{t=6,s} / RATE_{a='Spot',t=6,s} \\
& + REV_{t=9,s} / RATE_{a='Spot',t=9,s} \\
& + REV_{t=12,s} / RATE_{a='Spot',t=12,s} \\
\\
& + YFH_{a='12Fwd',t=0,s} * \left(\left(1/RATE_{a='12Fwd',t=0,s} \right) - \left(1/RATE_{a='Spot',t=12,s} \right) \right) \\
& + YFH_{a='9Fwd',t=3,s} * \left(\left(1/RATE_{a='9Fwd',t=3,s} \right) - \left(1/RATE_{a='Spot',t=12,s} \right) \right) \\
& + YFH_{a='6Fwd',t=6,s} * \left(\left(1/RATE_{a='6Fwd',t=6,s} \right) - \left(1/RATE_{a='Spot',t=12,s} \right) \right) \\
& + YFH_{a='3Fwd',t=9,s} * \left(\left(1/RATE_{a='3Fwd',t=9,s} \right) - \left(1/RATE_{a='Spot',t=12,s} \right) \right) \\
\\
& + YFH_{a='9Fwd',t=0,s} * \left(\left(1/RATE_{a='9Fwd',t=0,s} \right) - \left(1/RATE_{a='Spot',t=9,s} \right) \right) \\
& + YFH_{a='6Fwd',t=3,s} * \left(\left(1/RATE_{a='6Fwd',t=3,s} \right) - \left(1/RATE_{a='Spot',t=9,s} \right) \right) \\
& + YFH_{a='3Fwd',t=6,s} * \left(\left(1/RATE_{a='3Fwd',t=6,s} \right) - \left(1/RATE_{a='Spot',t=9,s} \right) \right) \\
\\
& + YFH_{a='6Fwd',t=0,s} * \left(\left(1/RATE_{a='6Fwd',t=0,s} \right) - \left(1/RATE_{a='Spot',t=6,s} \right) \right) \\
& + YFH_{a='3Fwd',t=3,s} * \left(\left(1/RATE_{a='3Fwd',t=3,s} \right) - \left(1/RATE_{a='Spot',t=6,s} \right) \right) \\
\\
& + YFH_{a='3Fwd',t=0,s} * \left(\left(1/RATE_{a='3Fwd',t=0,s} \right) - \left(1/RATE_{a='Spot',t=3,s} \right) \right) \quad \forall s \in S
\end{aligned}
$$

ExpYearlyRev:

$$XER = \sum_{s \in S} PROB * XR_s$$

Variance:

$$V = \sum_{s \in S} \left(XR_s - XER \right)^2 * PROB$$

OBJECTIVE:

$$\min \; VOL = \; \min \sqrt{V} + XC$$

CONSTRAINTS:

UpperHedgeLimit3:

$$YFH_{a='3Fwd',t=0,s} \leq UL * REV_{t=3,s} \quad \forall s \in S$$

UpperHedgeLimit6:

$$YFH_{a='6Fwd',t=0,s} + YFH_{a='3Fwd',t=3,s} \leq UL * REV_{t=6,s} \quad \forall s \in S$$

UpperHedgeLimit9:

$$YFH_{a='9Fwd',t=0,s} + YFH_{a='6Fwd',t=3,s} + YFH_{a='3Fwd',t=6,s} \leq UL * REV_{t=9,s} \quad \forall s \in S$$

UpperHedgeLimit12:

$$YFH_{a='12Fwd',t=0,s} + YFH_{a='9Fwd',t=3,s}$$
$$+YFH_{a='6Fwd',t=6,s} + YFH_{a='3Fwd',t=9,s} \leq UL * REV_{t=12,s} \quad \forall s \in S$$

Balance constr. 0-3 months:
$$YFH_{a='3Fwd',t=0,s} = FP_{a='3Fwd'} + YFB_{a='3Fwd',t=0,s} - YFS_{a='3Fwd',t=0,s} \quad \forall s \in S$$

Balance constr.0- 6 months:
$$YFH_{a='6Fwd',t=0,s} = FP_{a='6Fwd'} + YFB_{a='6Fwd',t=0,s} - YFS_{a='6Fwd',t=0,s} \quad \forall s \in S$$

Balance constr.3- 6 months:
$$YFH_{a='3Fwd',t=3,s} = YFB_{a='3Fwd',t=3,s} - YFS_{a='3Fwd',t=3,s} \quad \forall s \in S$$

Balance constr.0- 9 months:
$$YFH_{a='9Fwd',t=0,s} = FP_{a='9Fwd'} + YFB_{a='9Fwd',t=0,s} - YFS_{a='9Fwd',t=0,s} \quad \forall s \in S$$

Balance constr.3- 9 months:
$$YFH_{a='6Fwd',t=3,s} = YFB_{a='6Fwd',t=3,s} - YFS_{a='6Fwd',t=3,s} \quad \forall s \in S$$

Balance constr.6- 9 months:
$$YFH_{a='3Fwd',t=6,s} = YFB_{a='3Fwd',t=6,s} - YFS_{a='3Fwd',t=6,s} \quad \forall s \in S$$

Balance constr.0- 12 months:
$$YFH_{a='12Fwd',t=0,s} = YFB_{a='12Fwd',t=0,s} - YFS_{a='12Fwd',t=0,s} \quad \forall s \in S$$

Balance constr.3- 12 months:
$$YFH_{a='9Fwd',t=3,s} = YFB_{a='9Fwd',t=3,s} - YFS_{a='9Fwd',t=3,s} \quad \forall s \in S$$

Balance constr.6- 12 months:
$$YFH_{a='6Fwd',t=6,s} = YFB_{a='6Fwd',t=6,s} - YFS_{a='6Fwd',t=6,s} \quad \forall s \in S$$

Balance constr.9- 12 months:
$$YFH_{a='3Fwd',t=9,s} = YFB_{a='3Fwd',t=9,s} - YFS_{a='3Fwd',t=9,s} \quad \forall s \in S$$

Non-anticip. buy:
$$YFB_{a,t,s_1} = YFB_{a,t,s_2} \quad \forall a \in A \setminus_{\{'Spot'\}}, \quad \forall t \in T, \quad \forall s_1, s_2 \in S \,/\, (s_1, s_2) \in L_t$$

Non-anticip. sell:
$$YFS_{a,t,s_1} = YFS_{a,t,s_2} \quad \forall a \in A \setminus_{\{'Spot'\}}, \quad \forall t \in T, \quad \forall s_1, s_2 \in S \,/\, (s_1, s_2) \in L_t$$

Non-anticip. hold:
$$YFH_{a,t,s_1} = YFH_{a,t,s_2} \quad \forall a \in A \setminus_{\{'Spot'\}}, \quad \forall t \in T, \quad \forall s_1, s_2 \in S \,/\, (s_1, s_2) \in L_t$$

Lower bound1:
$$YFH_{a,t,s} \geq 0 \quad \forall a \in A \setminus_{\{'Spot'\}}, \forall t \in T, \forall s \in S$$

Lower bound2:
$$YFB_{a,t,s} \geq 0 \quad \forall a \in A \setminus_{\{'Spot'\}}, \forall t \in T, \forall s \in S$$

Lower bound3:
$$YFS_{a,t,s} \geq 0 \quad \forall a \in A \setminus_{\{'Spot'\}}, \forall t \in T, \forall s \in S$$