Mat-Øk 3.2 Lecture Notes: Mathematical Modeling and Optimization with Applications in Finance

Soren S. Nielsen

March 17, 1998

Contents

1 Introduction

These notes support the Mat-Øk 3.2 course taught at the Department of Operations Research, University of Copenhagen, Fall 1997. The notes this year are just being written and are certainly in draft form. They are supplemented by [15]. The students in the class are being used as guinea-pigs and deserve thanks!

For up-to-date information on the course, please visit the author's home page at

`http://www.math.ku.dk/~nielsen/`

and follow the course link.

2 The Dedication, or Cash-Flow Matching (CFM) Model.

Companies such as insurance companies often face a liability stream reaching several years into the future, for instance, representing future payouts on insurance products such as life insurance. Usually such future liability streams are stochastic, that is, it is not known today precisely when they will occur, or how big they will be. For a life insurance, this depends on customers' longevity and possibly on options built into the product, such as cancellation rights. Other examples of future liability streams are home owner mortgages (with fixed or variable payments over a typically 15-30 year period), lottery payouts (for instance, the Texas Lottery pays out larger prizes in 20 annual installments), etc.

It is often of interest to determine a portfolio of bonds (obligations) whose cash-flows replicate that of the liability stream. Regulatory committies often require insurance companies to demonstrate solvency, and one way to do this is to determine a "fair market value" of their liabilities by finding a replicating portfolio consisting of default-free bonds, such as Treasuries.This is similar to determining the present value (or "expected present value" in case of stochastic liabilities) of the liabilities using, for instance the zero-coupon yield curve, but is a more realistic measure because actually trading bonds, such as par bonds or even corporate or municipal bonds could also be used (although then credit risk needs to be addressed). Such a market value could also be used to sell the liability stream. This is done by many lotteries that pay out over many years but are not in the business of (or are prohibited from) money management.

2.1 The Deterministic Case

We consider first the deterministic case, where assets as well as liabilities are known in advance. Define:

$T = \{0, ...m\}$**:** The set of time periods, for instance measured in years, from $t = 0$ ("now") to $t = m$, the **horizon**.

$U = \{1, ..., n\}$**:** The *Universe* of assets under consideration for inclusion in the portfolio.

$F_{i,t}$**:** The cash flow arising from asset i at time t.

L_t**:** The liability due in period $t \geq 1$.

The cash flows $F_{i,t}$ can be both positive and negative; by convention, a positive number indicates incoming cash and a negative number outgoing cash. For a bond which is purchased today at price P_i per unit face value, with an annual coupon payment of c_i and a maturity of 4 years we would have $F_{i,\cdot} = (-P_i, c_i, c_i, c_i, 1 + c_i, 0, ..., 0)$.

The decision variables $x_i, i \in U$, denote the amount, in face value, of bond i that should be purchased. The objective of the model is to find the least expensive portfolio whose cash flows at least satisfy the liabilities, and this cost is represented by the variable λ in this model:

CFM-1:

(Cash-Flow Matching Model)

$$
\begin{array}{llll}
\underset{x \in U, \lambda \in \Re}{\text{Minimize}} & \lambda & & \\
\text{Subject to} & \sum_{i \in U} F_{i,0} \cdot x_i + \lambda & \geq 0, & \\
 & \sum_{i \in U} F_{i,t} \cdot x_i & \geq L_t, & t \geq 1 \\
 & x_i & \geq 0 & I \in U
\end{array}
$$

The model works as follows, assuming positive liabilities and cash flows as shown above: Since portfolio cash flows in each time period must at least satisfy the liabilities the decision variables, or some of them, are forced to become positive. This in turn forces the cost, λ, to be positive (since the cash flows at time 0 are negative). By minimizing this cost, the least expensive portfolio is arrived at.

It would seem natural to formulate this model using the objective

$$\text{Minimize} \sum_{i \in U} P_i \cdot x_i,$$

remove the first constraint above, and keep all cash flows positive elsewhere. This is a valid approach, but for instruments where cash flows may be either positive or negative (e.g., investments that are paid over several periods before returning to positive cash flows, or fixed-for-floating swaps), there may not be a natural notion of an up-front price, so the formulation shown is somewhat more general.

Note that the liabilities themselves may be negative, and so may the optimal "cost". For this reason, λ in this model is often called a "lump" sum, to be paid or received up front. These possibilities come into play in the extended models given below.

It is important to note that, even though several time periods are modeled, the **CFM** model is *static* in the sense that a single investment decision is made now, then not changed. In the context of a stochastic future (a redudancy...) one typically employs a truly *dynamic* model that allows for future rebalancing of the portfolio. In Section 2.3 we return to such *Stochastic Programming* models.

2.1.1 Short-term reinvesting and borrowing

The only way for **CFM** to meet a liability in period t is for it to purchase instruments with combined cash flows in that period to meet the liability. If the asset universe does not include an asset with a substantial cash inflow (for instance, a bond that matures) in period t, then instruments that mature later than t must be purchased in excessive amounts so their coupon payments can cover L_t. Even if a bond matures at t, it may be advantageous to purchase a much better priced bond that matures prior to t and simply deposit the money in the bank to meet the period t liability. Similarly, one could borrow, at the bank's rate, against a security which matures sometime in the future (in reality one might instead sell such a security, but the present model does not take such sales into account). Finally, short-term reinvesting and borrowing becomes natural with the integrality (tradeability) constraints we introduce later.

Short-term reinvesting and borrowing can be modeled by adding variables r_t and b_t to represent the amount of cash reinvested from period t to $t+1$, or borrowed in period t from period $t+1$. There is usually no need to model multi-period reinvesting or borrowing, since this is basically equivalent to a series of single-period acts. Assume that future, short-term investments have a period return of ρ_t, and that money can be borrowed at a rate of β_t. The augmented model is then:

CFM-2:

(Cash-Flow Matching with reinvesting and borrowing)

$$
\begin{array}{lll}
\underset{x,\lambda,r,b}{\text{Minimize}} & \lambda & \\
\text{Subject to} & \sum_{i\in U} F_{i,0}\cdot x_i + b_0 + \lambda = r_0, & \\
& \sum_{i\in U} F_{i,t}\cdot x_i + (1+\rho_t)\cdot r_{t-1} + b_t = L_t + r_t + (1+\beta_t)\cdot b_{t-1}, & t \geq 1 \\
& b_m = 0 & \\
& x_i, r_i, b_i \geq 0 & I \in U
\end{array}
$$

In this version cash can be carried between periods at the respective rates. Note that borrowing in the last period is, of course, prohibited. Reinvesting cash in the last period is not prohibited, but such cash is essentially wasted, although the variable r_m can be interpreted as an end-of-horizon surplus. We have changed the inequalities in **CFM**'s cash-flow balancing constraints to equalities since, with the possibility of reinvesting/borrowing there is no need for inequalities.

The model does not usually, despite the increase in number of variables, become much harder to solve, but it is important to realize that we have introduced a risk that was not previously present in the model. This is in the assumption that we will in the future be able to borrow and reinvest at certain rates. If this assumption turns out not to hold, the model portfolio may actually generate a surplus or loss. In reality this is not usually a great concern as long as the reinvested and borrowed amounts are relatively small (which can be enforced easily using bounds), and the assumed rates are reasonable. However, it is customary to use quite conservative rates, such as $\rho_t = 0$ and $\beta_t = 0.15$ or higher. Note that reinvestment and borrowing rates need not be constant through time. They could, for instance, be a function of the planner's expectations of future interest rates.

2.1.2 Tradeability Considerations

Large institutional bond investors are usually interested in purchasing large blocks of individual bonds, in "round numbers" ("even lots") of face value. This is primarily because such blocks are more easily traded than smaller or "odd-lot" holdings, but also for liquidity reasons, to avoid the risk of getting stuck with a small holding of a bond with poor liquidity. In addition, brokers usually require a premium for "odd-lot" trades because they may have to split an "even lot" and incur a risk to sell the remaining odd lot. In short, we may need the model to allow only:

1. *even-lot* purchases, i.e., face-value amounts in multiples of, say, \$100,000, and/or

2. *minimum-lot* purchases, i.e., purchase either none or at least some amount, say \$500,000, of each bond.

Even-lot purchases are naturally modeled using integer variables. If the decision variables x_i denote for instance the number of \$100 bonds to purchase and lots in multiples of \$100,000 are desired, we just add integer variables y_i and constraints

$$x_i = (100,000/100) \cdot y_i$$

to **CFM**. Minimum-lot purchases, for instance \$500,000 or nothing, are handled using binary variables, $b_i \in \{0, 1\}$, and the constraints

$$(500,000/100) \cdot b_i \leq x_i \text{ and } x_i \leq M \cdot b_i,$$

where M is some large number such that, if $b_i = 0$ then $x_i = 0$, but if $b_i = 1$ then $x_i \geq (500,000/100)$ and essentially unconstrained upwards.

When integrality conditions are used with **CFM**, it is important to also allow for reinvesting and possibly borrowing, even at very conservative rates, otherwise the model will be forced to overinvest and discard cash in most periods, which could result in a very sub-optimal solution.

The above extensions both change the model from an LP to a Mixed-Integer Program, MIP, which can be extremely expensive to solve to optimality, even for moderate numbers of discrete variables. There are certain modeling tricks that may help (such as giving y_i an upper bound as low as possible, and keeping M as small as possible, while still not in reality constraining x_i), but in practice it is often necessary to resort to *heuristics* such as rounding: Solve **CFM** in its LP version, then successively round variables to the nearest feasible value until a reasonable solution is found. Of course there is then no guarantee that an optimal solution has been found, but by comparing the value of a solution to the LP solution one can at least put a bound on the best possible improvement and thus estimate whether the solution found is "sufficiently good".

2.1.3 Transactions costs

Associated with every investment is transactions costs, such as brokerage fees and bid-ask spreads. Certain model types (such as the Markowitz model, Section **??**) are notorious for occasionally returning solutions containing small amounts of each of a large number of securities. Such a portfolio is in practice impossible to justify due to the costs of establishing and maintaining it. By explicitly modeling transaction costs, one can avoid this problem.

We distinguish between two types of transactions costs: Fixed and variable. A *fixed* cost is incurred for each instrument that is traded. This is modeled using a binary variable $z_i \in \{0, 1\}, i \in U$ for each instrument, and adding the term

$$F \cdot \sum_{i \in U} z_i$$

to the objective, where F is the fixed cost. These new variables need to be linked to the x_i such that if $x_i > 0$ then $z_i = 1$. A set of constraints

$$x_i \leq M \cdot z_i,$$

where M is a suitably large number, takes care of this. The resulting program is, of course, a mixed-integer program, and can as such be very difficult to solve to optimality.

Variable, or *proportional*, transactions costs, on the other hand, are easily modeled without any penalty in model complexity or difficulty. A variable transaction cost is a per-unit cost, and this cost can simply be added to the cost of the instrument (or, more precisely, subtracted from $F_{i,0}$ in the **CFM** models). They arise, for instance, in modeling bid-ask spreads, in addition to brokers' commissions.

See also Mulvey, [16].

2.1.4 Rebalancing an existing portfolio

It is perhaps not the most usual situation for a portfolio manager to have to construct a portfolio from scratch. More likely there's an existing portfolio (for instance last year's optimal portfolio) that needs to be re-optimized, or *rebalanced*. Of course one could just sell the old portfolio, then buy the new, optimal portfolio, but that would incur probihitive transactions costs, and would seem unnecessary if the changes in position of most instruments were small. It is therefore necessary to model transaction costs, especially the fixed ones, explicitly to encourage the model to limit turnover. Using additional parameters, x_i^0 to represent the existing portfolio, this can (as in Section 2.1.3) be done using binary variables; the precise model is left as an exercise to the reader (the model should take both fixed and variable costs into account).

In some situations certain holdings may be frozen, i.e., not to be changed. This can happen when the portfolio manager, for whatever reason, needs to keep her position in certain instruments. One can either "fix" the related variables at the current holdings or totally remove such holdings from the model by removing them from the universe U and modify the liabilities by subtracting the instruments' cash-flows. This is one example where the resulting liabilities may become negative.

2.1.5 Time-mismatched cash flows

We have implicitly assumed that all incoming cash flows in a given time period was available to meet the liabilities in that period. This may not always be the case. If the time periods are years, and the liabilities are due on the first day of the year, then some of the cash flow from assets such as bonds would not be available yet. One solution is to augment the set of time periods T to include all cash flow and liability dates, but this leads to a larger model, and is usually not necessary. An alternative is to assume that money can be borrowed against certain future cash flows (still within the same time period), and that incoming cash that arrives before it is needed will be invested until needed. In this case the cash flows can be adjusted in time value. For instance, if a bond coupon payment c_i arrives 6 months after it is need to cover a liability, and the annual rate of borrowing against it is β, then the value $c_i \cdot (1+\beta)^{-1/2}$ (or $c_i \cdot e^{-\beta/2}$) should be used instead of c_i in the model.

2.1.6 End-of-horizon effects

Many types of liabilities do not have a specific horizon beyone which they have been paid out. For instance, the life insurance company which insured Methusalem would need a longer than usual horizon to accurately model their liabilities. Similarly, some assets, such as stocks, do not have definite horizons. However, the **CFM** models need a finite horizon. The usual way to handle this problem in practice is to calculate a final liability (and a final cash flow for the assets) at the model horizon, m, equal to the present value (at time m) of the remaining, outstanding liabilities or cash flows.

This again introduces a risk to the model because an assumption needs to be made regarding future interest rates, or the future, present value of the outstanding cash flows. By using conservative estimates and a long horizon one can safeguard against surprises, but there is an inherent trade-off between model accuracy and complexity.

2.1.7 Diversification considerations

Many institutional investors have limitations on the allowable exposure to risky investments, such as corporate or municipal bonds, mortgage-derivatives, etc. Such limits are usually expressed as a maximum percentage of the portfolio value that may be invested in certain classes of investment vehicles. Such diversification, or limitation, constraints are easily modeled.

Assume that an investor wishes to limit her exposure to the bonds in some set $S \subset U$ to a fraction, p, of total value, and assume that the instruments in question have positive cash flows once they have been purchased at a price P_i. The appropriate constraint can then be written as follows:

$$\sum_{i \in S} P_i \cdot x_i \leq p \cdot \sum_{i \in U} P_i \cdot x_i.$$

Naturally, there can be several such constraints covering different subsets of the asset universe. If a limitation expressed in face value amounts, rather than value is desired, the P_is above can just be omitted.

2.1.8 Net present values and duals

To the optimal solution of an optimization program such as an LP is associated so-called *dual variables.* Dual variables complement the model (or *primal*) variables, but are associated with the model's constraints, including the bounds on variables. Expressed informally, the dual value corresponding to a constraint

$$\textit{linear expression} \left\{ \begin{array}{c} \leq \\ \geq \end{array} \right\} \textit{constant}$$

tells us *the change in objective value per unit change of the constraint's right-hand side*, up to a point. If the constraint is tightened, the objective either stays the same (if the dual is zero) or gets worse (i.e., larger in a minimization program, smaller in a maximization program), and vice-versa if the constraint is loosened. For equality constraints the rules are slightly more complicated.

The dual variables of the cash-flow constraints of **CFM** are interesting because they represent the *present value* of money at the time with which the constraint is associated. It is easy to see that this is correct: If the right-hand side of such a constraint (i.e., the liability) were increased by one unit, then the objective would increase by the additional cost needed today to cover the additional future liability.

This notion of present value is not the usual one, obtained by discounting according to a zero-coupon yield curve. Rather, it is determined implicitly through the prices and cash flows of the assets in U. It is interesting to note that there is no explicit notion of "yield curve" in **CFM**, yet the model automatically "bootstraps" one.

2.2 A variation: Maximizing Horizon Return

CFM ignores what happens to the final surplus cash, r_m, and there is in fact no incentive for the model to carry any surplus cash forward to the end, considering the possibility of "wasting" money by simultaneously reinvesting (at a low rate) and borrowing (at a high rate). We therefore consider two changes to the model:

1. the objective of the model is now to *maximize the final cash position*, subject to meeting the liabilities,

2. Instead of minimizing an initial lump sum, it is assumed that there is a *budget* for investing.

Hence, we now model an investor with a specific sum to invest, with liabilities (possibly 0) to satisfy, who wants to maximize investment performance. The focus is now on investing rather than dedication.

MAX:

Maximizing horizon position

$$
\begin{array}{llll}
\underset{x,h,r,b}{\text{Maximize}} & h & & \\
\text{Subject to} & \sum_{i\in U} F_{i,0}\cdot x_i + b_0 + B & = r_0, & \\
& \sum_{i\in U} F_{i,t}\cdot x_i + (1+\rho_t)\cdot r_{t-1} + b_t & = L_t + r_t + (1+\beta_t)\cdot b_{t-1}, & t\geq 1 \\
& b_m & = 0 & \\
& r_m & = h & \\
& x_i, r_i, b_i & \geq 0 & I \in U
\end{array}
$$

Here, B is the investor's budget, and h, the *horizon position*, is the objective, explicitly set equal to the final amount available. The liabilities can be interpreted as the investor's desired consumption along the way. Note that the interpretation of the dual prices is different from **CFM** (what?).

2.3 The Stochastic Case

Unfortunately, it is not always realistic in practice that the future data, even in a relatively simple model such as the Dedication/CFM or **MAX** model, are known with certainty. In fact, large financial institutions (banks, insurance companies) often make a business precisely in taking on stochastic liabilities, thus freeing cheir customers from assuming risk. An example is annuities: The insured (annuitant) is relieved of the risk of outliving his funds, but the insurance company now has a stochastic liability (of course, the company does not assume this risk for free. For each individual annuity contract, it *expects* to gain).

We start with the **MAX** model with reinvesting and borrowing, but with stochastic cash-flows, including liabilities. The model can be written:

Maximizing horizon position under uncertainty

$$
\begin{array}{lll}
\underset{x,h,r,b}{\text{Maximize}} & \tilde{h} & \\
\text{Subject to} & \sum_{i\in U} F_{i,0}\cdot x_i + b_0 + B & = r_0, \\
& \sum_{i\in U} \tilde{F}_{i,t}\cdot x_i + (1+\rho_t)\cdot \tilde{r}_{t-1} + \tilde{b}_t & = \tilde{L}_t + \tilde{r}_t + (1+\beta_t)\cdot \tilde{b}_{t-1}, \\
& \tilde{b}_m & = 0 \\
& \tilde{r}_m & = \tilde{h} \\
& x_i, \tilde{r}_i, \tilde{b}_i & \geq 0
\end{array}
$$

where the ˜ indicates stochastic parameters. For simplicity we have dropped the index set indications $t \geq 1$ and $i \in U, t \in T, s \in S$; they correspond to the previous models. Note that the investment variables, x_i, are not stochastic, since they need to be determined up-front. We have assumed that first-period data are known with certainty, but there's nothing inherently important about this. On the other hand, one now has the opportunity to also let ρ and β be stochastic.

As this model stands, it is too general for our use, and it is not even clear what it means to "Maximize $\tilde{h}$". If the stochastic parameters (we avoid the term "random variables" to avoid confusion with the optimization variables) have general, continuous distributions, the problem is analytically intractable, except possibly if the dimensions involved are very small. The practical way to proceed is to assume that all distributions involved are *discrete*. This is justified in that, in principle, any distribution can be approximated arbitrarily closely by a discrete distribution. Actually, it is customary to assume that the random data are given by a set of *scenarios*, where a scenario $s \in S$ is given by a collection of (joint) realizations of the random data; that is; we are given data

$$(F_{i,t}^s, L_t^s, h)_{t \in T; i \in U}$$

for each scenario $s \in S$, and then have to determine optimal x, h, r^s and b^s variables. We also assume that scenario probabilities, p_s, are given, although these are not yet used in our model. Note that the variables r^s, b^s and h are stochastic in the sense that they are scenario-dependent, but they are still outputs from the model, not inputs. They are example of *stochastic optimization variables*, which we'll encounter again shortly.

With a change of notation to emphasize the scenario-based nature of the model, we have:

S-MAX:

Maximizing expected horizon return

$$\begin{array}{lll}
\underset{x,h,r,b}{\text{Maximize}} & \sum_{s \in S} p_s \cdot h^s & \\
\text{Subject to} & \sum\limits_{i \in U} F_{i,0} \cdot x_i + b_0 + B & = r_0, \\
& \sum\limits_{i \in U} F_{i,t}^s \cdot x_i + (1+\rho_t) \cdot r_{t-1}^s + b_t^s & = L_t^s + r_t^s + (1+\beta_t) \cdot b_{t-1}^s, \\
& b_m^s & = 0 \\
& r_m^s & = h^s \\
& x_i, r_i^s, b_i^s & \geq 0
\end{array}$$

where we have made precise also the objective: Maximize *expected horizon return*. This is only one possible objective, corresponding to a risk-neutral investor; we will see other possibilities later (max worst case, max expected utility etc.).

For a practical application of a similar model see Adamidou et al. [1]. For a related approach, *Scenario Immunization*, see Dembo [8].

2.4 Scenario Generation

It is clear that the set of scenarios chosen for **S-MAX** is extremely important for the optimal solution returned by the optimization. The scenario set needs to be:

Comprehensive: It should capture all aspects, both extreme and "normal" instances, of the underlying distributions,

Consistent: It must capture *correlations* among the stochastic data well.

Comprehensiveness is important because extreme scenarios must obviously be present in the set so the model can "see" them and take them into account. On the other hand, the model should not be blindsided by extreme events (which are usually very low-probability events), so "normal" cases need to be present as well. This may not be too serious for **S-MAX**, where a low-probability event may not influence the objective too much, but it could have a very large effect on the portfolio generated by **S-CFM**.

Consistency is equally important. In real life events are very often correlated. In Finance, some instruments pay out precisely when other instruments do not (puts and calls). If a risk-averse investor can find negatively correlated assets, or assets which are positively correlated with the liabilities, they should be included in the model, so that their cash flows reflect this correlation. This means that the F_I^s, t are consistent by not being uniformly "all high" or "all low". The scenario set must also be consistent in the time dimension. Some financial instruments pay out a quite predictable amount, but the *timing* of payouts are unpredictable. Examples include CMOs or life-insurance. Hence, low initial payouts should correspond to higher, later payouts for the same instrument under a given scenario.

We might want to add as a third requirement to the asset set that it *should be small*! The model above, with a reasonable time horizon and a reasonable number of assets is already big. If on top the size of the model is multiplied by $\mid S \mid$, it quickly becomes unsolvable. For this reason, *Scenario Generation*, or the way in which scenarios are produced to satisfy all three requirements, is in practice a major component of designing a stochastic program (or is simply ignored!). We give below some examples how scenario generation in Financial planning has been approached. Other examples from real life can be found in, e.g., Cariño, Myers and Ziemba [27], and Mulvey..*Scenario Generation for the Towers Perrin ref.*.

Scenario-Generation in Fixed-Income Planning. Very often one can identify a single or a few underlying, stochastic processes, whose behavior can be modeled and which determines the stochastic parameters. An important example is financial planning involving fixed-income instruments, where the cash-flows of the instruments are (primarily) driven by the future development of interest rates.

The crucial point is that future cash flows must be unique functions of future interest rates, which is the case for Treasury Bonds, and to a large extend to many other fixed-income instruments. This is one way to proceed:

1. One starts with an interest model, either a one-factor like Vasicek which models short rates (see Hull [11], 17.4), or a multi-factor like Nielsen-Ronn [23], which models two factors, short and long rates. The model should be calibrated so it is consistent with expectations of future events. This can be done by matching the model to historical observations (such as the volatility and correlations of the factors), or by matching the model to observable prices of traded instruments (mostly options), or by the decision maker simpley postulating a set of interest rate scenarios (s)he feels covers the possibilities.

2. The interest rate model, if not already in (discrete) scenario form, is then used to generate scenarios of future interest rates. The scenarios need to indicate the whole term structure at each future time point in the model. In the case of a popular model such as Vasicek, this step could involve simulation.

3. Finally, given these scenarios, the fixed-income instruments' cash flows are calculated. This may often be done only approximately. For instance, with corporate bonds or mortgage-securities, interest rates are only one determing factor. The company may go bankrupt for un-economic reasons, or the home-owner get divorced and sell the house.

Examples where this approach was taken, and where the liability side was also stochastically dependent on interest rates is Nielsen-Zenios [25], upon which the Winvest case is built upon.

The New-York 7 The so-called "NY-7" is a set of interest rate scenarios which is used by the industry regulators in New York. It is a set of 7 interest rate scenarios, and each financial intermediary in the state of New York each year needs to demonstrate to regulators that they, with their present asset and liability portfolios, would stay solvent if interest rates over some specified future horizon behaved according to a NY-7 scenario. Since many assets and liabilities are not governed by interest rates alone, this is a very inexact science, but it actually forces financial institutions to look hard at their asset-liability matches and validate (internally and externally) their approaches in calculating future in- and out-flows.

The NY-7 only considers parallel shifts to the term structure. Starting with today's term structure, they are:

1. Rates stay the same,
2. Rates increase by 50bp per year for 10 years,
3. Rates increase by 100bp per year for 5 years, and then drop by 100bp per year for 5 years,
4. Rates increase by 300pb, then stay the same,
5. Rates decrease by 50bp per year for 10 years,
6. Rates decrease by 100bp per year for 5 years, and then increase by 100bp per year for 5 years,
7. Rates decrease by 300pb, then stay the same,

A "bp", or basis-point, is 1/100 of a percentage point, so an increase of 50bp of an interest rate of, say, 5.5% would take it to 6.0%. All rates are truncated at 0%. The scenarios are equally likely.

The NY-7 has a time horizon of 10 years, and includes both high, immediate shifts, long, slow shifts, and reversing shifts. Since the changes are quite dramatic, it is often argued (and hoped) that a model which hedges well against the NY-7 will be well-hedged against virtually any real change in rates, possibly with the exception of non-parallel shifts. Many institutions are content with using just the NY-7, possibly adding a few more scenarios, for planning purposes. While the NY-7 certainly contains dramatic (extreme) events, it can be argued that it probably leaves companies over-hedged.

(Under networks cite [18]; perhaps [17]; under stochastic networks cite [20]). Under diversification/fixed income [22].

3 The Two-Stage Stochastic Program

The models in the previous sections were *static*, meaning that the initial decision (portfolio) is never modified. We now introduce *dynamic* models, where future changes, *recourse*, to the initial decision are allowed. This leads naturally to *Two- and Multi-stage Stochastic Programs.*

3.1 Modeling future rebalancing: Dynamic models

Although the stochastic cash-flow matching problem and the **S-MAX** model for maximizing the expected horizon return are better approximations to an uncertain world than their deterministic counterparts, they have a serious limitation: They model a strict *buy-and-hold* situation, where the investor buys a portfolio, then never changes it. There is, of course, nothing to keep the investor from periodically reviewing the portfolio, but then we might as well model this possibility explicitly.

We first consider a simplified model where the decision maker can rebalance once in the future, after one time period. To do this, we introduce, in addition to the initial, or *first-stage*, decision variables x_i, a set of *second-stage* variable, y_i^s, where y_i^s is the holdings of security $i \in U$ in time periods $t = 1, ..., m$, under scenario $s \in S$. Note that the second-stage variables are scenario-dependent. This is natural, otherwise we would force the same second-stage portfolio no matter what was learned about the economy (or cash flows) during the first period.

The **S-MAX** model is modified as follows:

2S-SP:
A Two-Stage Stochastic Program

$$
\begin{array}{lll}
\underset{x,y,h,r,b}{\text{Maximize}} & \sum_{s \in S} p_s \cdot h^s & \\
\text{Subject to} & \sum_{i \in U} F_{i,0} \cdot x_i + b_0 + B & = r_0, \\
& \sum_{i \in U} F_{i,t}^s \cdot y_i^s + (1 + \rho_t) \cdot r_{t-1}^s + b_t^s & = L_t^s + r_t^s + (1 + \beta_t) \cdot b_{t-1}^s, \\
& \sum_{i \in U} P_i^s x_i & = \sum_{i \in U} P_i^s \cdot y_i^s \\
& b_m^s & = 0 \\
& r_m^s & = h^s \\
& x_i, y_i^s, r_i^s, b_i^s & \geq 0
\end{array}
$$

There are two new things in this model: First, the x and y variable now explicitly relate to different time periods, or *stages* in the decision process. The first stage is at time $t = 0$ where the x_i are determined, and the second stage is at time $t = 1$ where the y_i^s are determined. Second, we need constraints that link the second-stage to the first-stage variables, since the portfolio cannot change value (possibly except for rebalancing transaction costs, which are ignored here). This is expressed by the third constraint above, stating that the total value, at time 1, of the existing portfolio, $sum_{i \in U} P_i^s x_i$, must equal the total value of the new portfolio, $\sum_{i \in U} P_i^s \cdot y_i^s$, and where P_i^s is the price of instrument i at time 1 under scenario s.

We have now introduced new, stochastic parameters to the model, namely the future prices. This means that we need a way to estimate these prices. A reasonable way to do this is very problem-dependent. A possibility in a fixed-income world is to let the price equal the present value (at time 1) of the future cash flows, depending upon scenario, but this requires some assumption (or knowledge) of future interest rates. If the model is generated from an interest-rate model, such knowledge may be available. If the first-stage variables are restricted to short-maturity bonds, then the price-uncertainty is smaller, and the dynamics of the model come from the price changes in the longer, second-stage universe. There seems to be no general way to answer, or get around, this question.

Exercise: How could trading costs be incorporated into the above model?

3.2 A Formalization of the Stochastic Program

The two-stage stochastic program is a problem class of such importance that it deserves to be viewed in a more general and formal framework than that of financial planning. What follows is a general introduction to SPs, without any specific application in mind. The exposition here is from Nielsen and Zenios [26].

The two-stage stochastic linear programming model addresses the following situation:

> A decision is made at the present time facing future uncertainties. At a future time the uncertainties are resolved, and a recourse action is taken.

The uncertainties of the model are represented by stochastic parameters, that is parameters whose actual values are not known until a future period, but their probability distributions are known *a priory.* The decision to be made at present is called the *first-stage* decision, and the (future) recourse decision is called the *second-stage* decision, which is contingent upon the first-stage decision and on the observed realization of the uncertain parameters.

In algebraic notation, the problem can be formulated as follows:

[**SLP**]

$$\underset{x \in \Re^{n_1}}{\text{Minimize}} \quad c^T x + \mathcal{Q}(x) \tag{1}$$

$$\text{Subject to} \quad Ax = b, \tag{2}$$

$$0 \leq x \leq u. \tag{3}$$

where

$$\mathcal{Q}(x) = E\{Q(\mathbf{d}, \mathbf{h}, \mathbf{T}, \mathbf{v}, \mathbf{W} \mid x)\},$$

E is the expectation operator defined on some probability space $(\Omega, \mathcal{F}, P)$, and

$$Q(d, h, T, v, W \mid x) = \quad \underset{y \in \Re^{n_2}}{\text{Minimize}} \quad d^T y \tag{4}$$

$$\text{Subject to} \quad Wy = h - Tx, \tag{5}$$

$$0 \leq \mathrm{y} \leq v. \tag{6}$$

Bold letters are used to designate stochastic quantities, and the corresponding roman letters designate instances of the stochastic quantities. If the second minimization problem is infeasible, we take its value to be $+\infty$. In this paper we assume that the problem has *relatively complete recourse*, i.e., problem (4) – (6) has a feasible solution for any value of the first-stage variable x which satisfies (2)- -(3).

In this formulation, n_1 and n_2 are the number of first-stage and second-stage decision variables, respectively. There are m_1 first-stage constraints (2), and m_2 second-stage constraints (5). All other matrices and vectors have conformable dimensions. Uncertainty of the second-stage problem is represented by the stochastic quantities **d, h, T, v** and **W**.

In this paper we consider the case where the stochastic quantities have a discrete and finite joint distribution, represented by the *scenario set* $\Omega = \{1, 2, 3, \ldots, S\}$. This set can be an exhaustive enumeration of the possible realizations of the stochastic parameters, or it can be obtained by sampling from the joint distribution, see, e.g., Dantzig and Infanger [1991]. In this case, we have

$$\mathcal{Q}(x) = \sum_{s=1}^{S} p_s Q(d_s, h_s, T_s, v_s, W_s \mid x), \tag{7}$$

where p_s is the probability of realization of scenario s,

$$p_s = P\{(\mathbf{d}, \mathbf{h}, \mathbf{T}, \mathbf{v}, \mathbf{W}) = (d_s, h_s, T_s, v_s, W_s)\}, \text{ for all } s \in \Omega. \tag{8}$$

It is assumed that $p_s > 0$ for all $s \in \Omega$ and that $\sum_{s=1}^{S} p_s = 1$.

Assuming that the uncertainties are represented by a finite scenario set, the stochastic program [SLP] can be reformulated (Wets [1974]) as the following *deterministic equivalent* program:

[**DELP**]

$$\underset{x \in \Re^{n_1}, y_s \in \Re^{n_2}}{\text{Minimize}} \quad c^T x + \sum_{s=1}^{S} p_s d_s^T y_s \tag{9}$$

$$\begin{aligned}
\text{Subject to} \quad & Ax &=& \quad b, & (10)\\
& T_s x + W_s y_s &=& \quad h_s, \text{ for all } s \in \Omega, & (11)\\
& 0 \le x &\le& \quad u, & (12)\\
& 0 \le y_s &\le& \quad v_s, \text{ for all } s \in \Omega. & (13)
\end{aligned}$$

This large-scale, linear problem consists of $n_1 + S \cdot n_2$ variables and $m_1 + S \cdot m_2$ equality constraints. The constraint equations can be written in matrix form as the following dual block-angular system:

$$\begin{pmatrix} A & & & & \\ T_1 & W_1 & & & \\ T_2 & & W_2 & & \\ \vdots & & & \ddots & \\ T_S & & & & W_S \end{pmatrix} \cdot \begin{pmatrix} x \\ y_1 \\ y_2 \\ \vdots \\ y_S \end{pmatrix} = \begin{pmatrix} b \\ h_1 \\ h_2 \\ \vdots \\ h_S \end{pmatrix} \tag{14}$$

It is now clear that, if a vector $\hat{x}$ is given, satisfying (2)–(3) for the first-stage variables, the system for the second-stage variables can be written as:

$$\begin{pmatrix} W_1 & & & \\ & W_2 & & \\ & & \ddots & \\ & & & W_S \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_S \end{pmatrix} = \begin{pmatrix} h_1 \\ h_2 \\ \vdots \\ h_S \end{pmatrix} - \hat{x} \cdot \begin{pmatrix} T_1 \\ T_2 \\ \vdots \\ T_S \end{pmatrix} \tag{15}$$

This system can be decomposed into S scenario blocks and solved in paralle l. This observation is the basis for the frequent use of *decomposition algorithms* in solving large-scale SPs. For a short introduction to Benders decomposition in this context, see Nielsen and Zenios [26].

Exercise: Show in detail how the various matrices and vectors in the formulation in this section relates to the formulation **2S-SP** in Section 3.1.

3.3 A Network Formulation

We now return to financial planning model very similar to **2S-SP** from Section 3.1. The purpose of this section is to show the power of networks in visualizing an optimization problem, and the almost mechanical way in which an algebraic formulation of the problem can be arrived at, once the graphical network is correctly understood.

Definition: A *Network* is a directed graph... (fill in later).

The problem addressed in the following is very similar to the Winvest case. It is a two-stage stochastic network, where the first-stage decision is an investment, and the second-stage decision is a subsequent rebalancing. There is one major difference, however, between the following formulation and **2S-SP**: The decision variables represent the *actual value of holdings*, not the *face value.* So a holding of 5 bonds, each with a face value of \$100 and a price of \$92,is represented by some $x^s_{pi} = 460$, not \$500. It turns out this change in decision variables is very convenient for network formulations.

The paper by Nielsen and Zenios [25], from which the material in this section is adapted, funds an insurance company's exposure from selling an annuity product called a *Single Premium Deferred Annuity*, or SPDA.

The portfolio is to be constructed from a universe $\mathcal{U}$ of financial instruments. The model is dynamic, and all events, such as asset trading or coupon payments, occur at discrete time points, $t = 0, ..., T$. An initial portfolio is constructed at time $\tau_0 = 0$, and is subsequently rebalanced at time points $\tau_1 < \tau_2 < ... < \tau_Y = T$. During the periods between rebalancing points (where *period* p is the time span $\tau_{p-1} \leq t < \tau_p$) the portfolio composition remains unchanged, except that cash-flows are reinvested at the short rate. Transaction costs are included, and limited borrowing is allowed.

Uncertainty is modeled using a set of interest rate scenarios, generated according to a suitable term structure model. Lapse behavior for SPDA annuitants is driven by the short-term rates under each scenario, leading to a stochastic liability stream. On the asset side, the prices of assets in future time periods, as well as the short-term borrowing and lending rates, are also interest rate driven. Their estimation is described in Section 4.2 in [24]. The objective of the stochastic model is to construct a portfolio whose cash-flows match the liabilities in each time period under all scenarios, while having a risk-return profile consistent with a prescribed level of risk tolerance.

In its basic form the model is a two-stage, stochastic program with recourse. The *first-stage* decision is the construction of the initial portfolio. After a realization of interest rates has been observed, the portfolio is rebalanced. The construction of the rebalanced portfolio constitutes the *second-stage* decision. Of course, rebalancing decisions are contingent upon the realized scenario and the composition of the initial portfolio.

For a given interest rate scenario, the model has a network structure, as shown in Figure 1 for two assets (Mulvey and Vladimirou [21]). Columns of nodes correspond to different time points, while rows of nodes correspond to different assets. The bottom row of nodes corresponds to cash. Horizontal arcs — between nodes corresponding to the same instrument $i \in \mathcal{U}$ — model the holding of instrument i in the portfolio. Arcs which link nodes in the bottom row model short-term cash reinvestment and borrowing. The vertical arcs which link cash and instrument nodes model changes in the position of each instrument, that is, rebalancing. The initial infusion of cash into the model consists of the SPDA premium and possibly insurer equity, and is used for constructing the initial portfolio.

Triangles on arcs designate arc multipliers, i.e., that the value (cash or holding) entering the arc is changed by some proportion (multiplier) before leaving the arc. For arcs modeling sales of assets, this multipler represents a transaction cost on sales. For reinvestment and borrowing, the multiplier represents short-term reinvestment and borrowing rates. Multipliers on the arcs which represent the presence of instruments (holding arcs) model the yield of the instrument for each time period. Of course, these multipliers are dependent on the specific scenario, $\mathcal{S} = \{1, ..., S\}$. The arcs in bold type for investments at time 0, i.e., construction of the initial portfolio, are *first-stage* decisions, and must be the same regardless of the scenario. The remaining arcs (later period holdings and rebalancings) represent *second-stage* decisions which are allowed to depend on the scenario realized. This distinction between first- and second-stage decisions is a key feature of two-stage, stochastic models, where immediate decisions cannot depend on, as yet, unrealized data, but future decisions can.

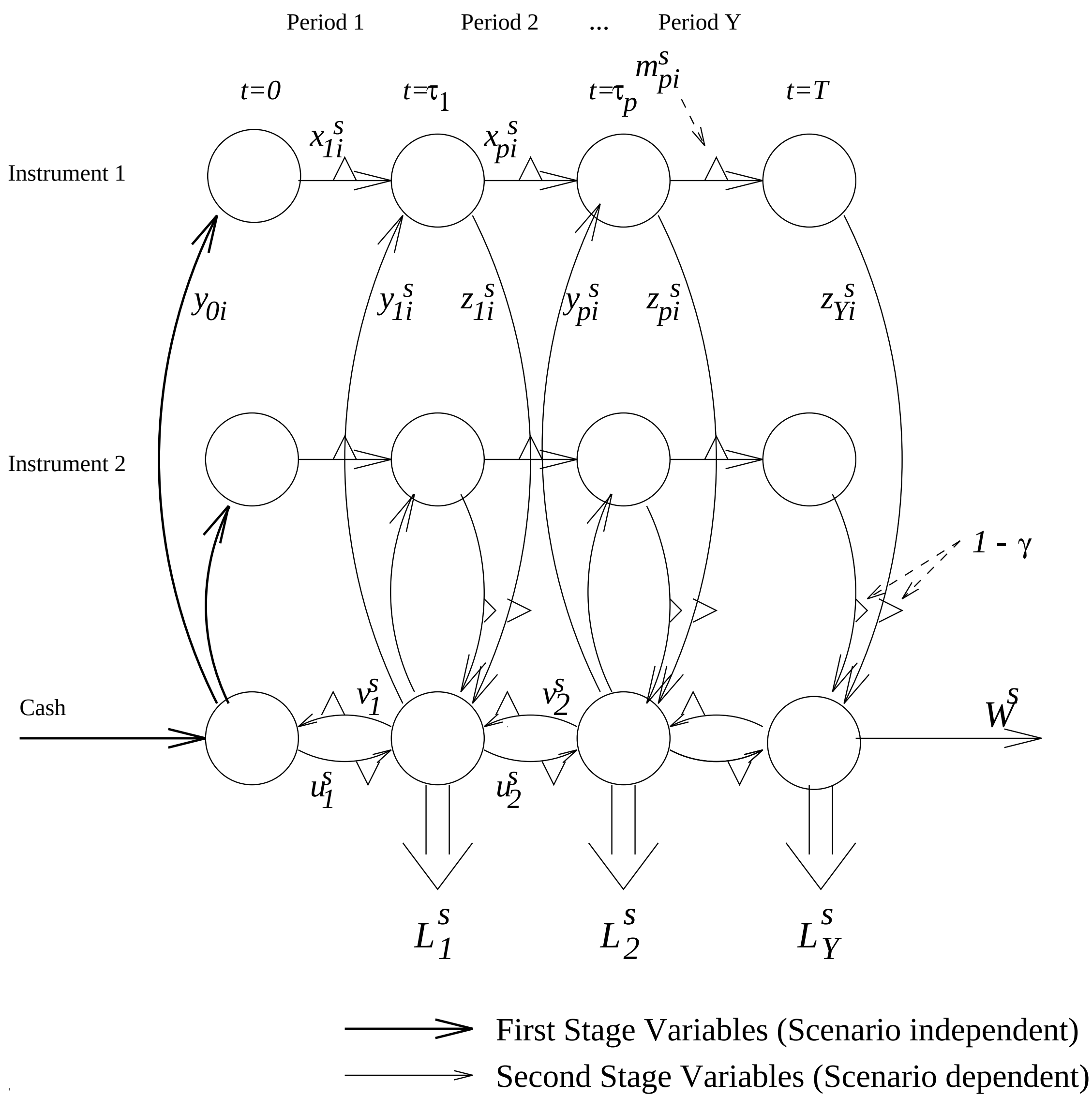

Figure 1: Network model underlying the two-stage, stochastic SPDA model. This figure includes two instruments, and depicts a 4-period model. Stochastic quantities are denoted by a superscript, s.

3.3.1 Model Definition

We now describe the components of the stochastic model. The model is described in three parts, dealing with security holdings, cash position, and the model's objective. The complete notation and algebraic formulation of the model is given in Appendix A.

Security Holdings: The variables x^s_{pi} are used to represent the holding (in dollar value) of instrument i during period p under scenario s. Purchases and sales of securities are represented by the variables y^s_{pi} and z^s_{pi}, respectively, where the index p refers to transactions which take place at the end of period p.

The constraints which define the holdings of securities during each time period under each scenario can be written as:

$$x^s_{1i} = y_{0i}, \quad \text{for all } s \in \mathcal{S}, i \in \mathcal{U}. \tag{16}$$

$$x^s_{(p+1)i} = m^s_{pi}x^s_{pi} - z^s_{pi} + y^s_{pi}, \quad \text{for all } s \in \mathcal{S}, i \in \mathcal{U} \text{ and } p \in 1, ..., Y-1. \tag{17}$$

Equation (16) states that the initial holdings equal the initial investments. The initial investment, y_{0i}, must not depend on the scenario to be realized, and hence does not have the scenario superscript. These variables are all measured in cash value rather than face value. Equation (17) defines the changes in the cash value of holdings due to yields, sales and purchases of instruments. The multipliers m^s_{pi} represent the yield during period p, of instrument i, under scenario s. We assume that the final portfolio must be liquidated, and this is ensured by the following constraint, which states that sales after the last period must equal holdings after that period:

$$z^s_{Yi} = m^s_{Yi}x^s_{Yi}, \text{ for all } s \in \mathcal{S}, \text{ and } i \in \mathcal{U}. \tag{18}$$

Cash Position Accounting: Next come definitions of cash positions in each time period. The initial amount of cash available (premium and equity) is denoted by C. Cash is used for purchases of securities and the payment of liabilities, and is generated by sales. During each period, excess cash is invested at the short rate, and borrowing is allowed. We use u^s_p to designate short-term cash investments during each time period, and for each scenario, and v^s_p to denote short-term borrowing.

The constraint for the first time period states that the initial amount of cash available, C, plus any first-period loan, v^s_1, must equal the amount invested in instruments, y_{0i} or held in cash during the first period, u_1:

$$\sum_{i \in \mathcal{U}} y_{0i} + u^s_1 - v^s_1 = C, \text{ for all } s \in \mathcal{S}. \tag{19}$$

In general (although not for the first time period), short-term borrowing and cash investment are scenario-dependent, hence the superscript s. The corresponding cash-flow balance constraints for the intermediate time periods are:

$$\sum_{i\in\mathcal{U}}((1-\gamma)z^s_{pi} - y^s_{pi}) - u^s_{p+1} + (1+r^s_p)u^s_p \; - (1+r^s_p+\delta)v^s_p + v^s_{p+1} = L^s_p, \quad (20)$$

$$\text{for all } s\in\mathcal{S}, p=1,...,Y-1,$$

where γ is a transaction cost (charged as a fraction of the amount of the transaction) when selling instruments and L^s_p is the liability due at the end of period p under scenario s. For simplicity, transaction costs on purchases are not used. Rather, γ should account for total costs. The cash-flow constraint for the last time period is:

$$\sum_{i\in\mathcal{U}}(1-\gamma)z^s_{Yi} + (1+r^s_Y)u^s_Y - (1+r_Y+\delta)v^s_Y = L^s_Y + W^s \text{ for all } s\in\mathcal{S}. \quad (21)$$

It differs from (20) in not allowing purchases, and in defining the final wealth under each scenario, W^s, as the surplus cash after the last liability, L^s_Y, has been paid.

Objective Function: The objective of the model is to maximize the expected value of a measure of final wealth across scenarios. We maximize the *Expected Utility of Return on Equity.* Return on Equity, ROE, under scenario s is defined as $r^s = W^s/E$, where E is equity. Utility is measured using the family of *iso-elastic utility functions* (Ingersoll [1987] [13]):

$$U_\alpha(r) = \begin{cases} \frac{1}{1-\alpha}(r^{1-\alpha}-1) & \text{for } \alpha\neq 1, \\ \log(r) & \text{for } \alpha = 1, \end{cases} \quad (22)$$

where $\alpha \geq 0$ is a risk-aversion parameter. Higher values of α implies more risk-aversion, i.e., less tolerance for risk on the investors' part. The value $\alpha = 0$ results in a linear utility function which corresponds to a risk-neutral attitude, and $\alpha = 1$ in a logarithmic utility function which corresponds to a moderate level of risk-aversion and which is known as the *growth-optimal strategy.* The properties of logarithmic utility functions for portfolio selection are discussed in McLean, Ziemba and Blazenko [15]. Assuming that all scenarios are equally probable, the objective function of the model is defined by

$$\text{Maximize Expected Utility } = u_\alpha \doteq \frac{1}{\mid\mathcal{S}\mid}\sum_{s\in\mathcal{S}} U_\alpha(W^s/E), \quad (23)$$

where u_α is the expected utility of return on equity.

Utility measures in themselves are not meaningful, except for ranking uncertain outcomes. Hence, we employ the more meaningful *certainty-equivalent return on equity*, CEROE. The certainty-equivalent has the same utility as the expected utility of the investment. The investor with the prescribed risk-attitude is indifferent between receiving the (deterministic) certainty-equivalent return and the (stochastic) portfolio return. We use the CEROE value of different portfolios when comparing alternative portfolio management strategies. The Certainty-Equivalent is defined by

$$\text{CEROE} = U_{\alpha}^{-1}(u_{\alpha}), \tag{24}$$

and is used in comparing investors' preferences among different investments. A complete description of the model is given in .

3.3.2 Extension to a Multistage Model

The two-stage model described above allows rebalancing decisions at times $t > 0$ to depend on future prices and returns, i.e., on data not yet known at time t. This implies that the decision maker has "perfect foresight" in making rebalancing decisions after the first time period. This is of course unrealistic. To overcome this problem, the two-stage model is extended to a *multistage model* in which decisions at time $t > 0$ do not depend on the specific sequence of events which will be realized during later time periods, but depend only on events observed prior to time t.

Multistage models are a much better representation of reality than two-stage models. They have the same data requirements as two-stage models, but are substantially more complex in structure, and can be significantly harder to solve. Similarly to the two-stage models, they are based on scenarios of future interest rates. In the two-stage case these scenarios are independent of each other but in the multistage case, they are grouped together such that certain scenarios are indistinguishable from each other up to a certain time point (See Figure 5 in [24]), where 8 of the 16 interest rate scenarios are indistinguishable up to time 6, groups of 4 are indistinguishable up to time 12, etc.). The lack of foresight mandates that decisions made under such indistinguishable scenarios, up to the time point where they differ, should be the same under each scenario. This leads to the requirement that some variables should be identical across scenarios. We return to multistage models in Section 4.

4 Multistage Stochastic Programs

As alluded to in Section 3.3.2, the two-stage stochastic program can be extended in a natural fashion to a multistage program. This program models the situation where there are several opportunities for adjusting ones position, such as rebalancing a portfolio. Each such opportunity is again called a *stage*. The following is a formal introduction to the problem (from Nielsen and Zenios [24]).

We now define the linear, multistage stochastic problem with generalized network recourse, followed by a deterministic equivalent formulation. Transposition of a vector x is denoted x^T, and the inner product $x^T y$ is written xy when context makes the meaning clear. Bold letters are used to denote stochastic quantities, and the corresponding roman letters designate instances of these quantities.

4.1 Formulation of the T-stage Stochastic Program

A T-stage stochastic programming problem can be formulated as follows (Birge [2]):

$$
\begin{array}{lll}
[\mathbf{MS}] & \min_{x_1} \left\{ c_1 x_1 + \mathrm{E}_{\xi_2} \left[\min_{x_2} \left(\mathbf{c}_2 x_2 + \mathrm{E}_{\xi_3|\xi_2} \left(\min_{x_3} \mathbf{c}_3 x_3 + \cdots + \mathrm{E}_{\xi_T|\xi_2,\dots,\xi_{T-1}} \min_{x_T} \mathbf{c}_T x_T \right) \right) \right] \right\} & \\
\text{s.t.} & A_1 x_1 = b_1, & \\
& \mathbf{B}_2 x_1 + \mathbf{A}_2 x_2 = \mathbf{b}_2, & \\
& \quad \mathbf{B}_3 x_2 + \mathbf{A}_3 x_3 = \mathbf{b}_3, & \\
& \qquad \ddots \qquad \vdots & \\
& \qquad \mathbf{B}_T x_{T-1} + \mathbf{A}_T x_T = \mathbf{b}_T, & \\
& 0 \le x_t \le u_t, \quad \text{for } t = 1, \dots, T, &
\end{array}
$$

where

$$\xi_t = (\mathbf{A}_t, \mathbf{B}_t, \mathbf{b}_t, \mathbf{c}_t) \quad \text{for } t = 2, \dots, T$$

are random variables, i.e., $\mathcal{F}_t$-measurable functions $\xi_\mathbf{t} : \Omega_t \mapsto \Re^{M_t}$ on some probability spaces $(\Omega_t, \mathcal{F}_t, P_t)$.

The decision variables $x_t \in \Re^{n_t}$, for $t = 2, \dots, T$, are stochastic variables measurable on the σ-field generated by $\xi_\mathbf{t}$. The notation E_ξ denotes mathematical expectation with respect to ξ, and $\mathrm{E}_{\xi_i|\xi_j}$ similarly denotes conditional expectation.

4.2 Scenarios and the Scenario Tree

The sequential nature of the decision process is apparent from this formulation. At each stage of the decision process, a conditional expectation is to be minimized. We consider in this paper the case where each ξ_t can assume only *finitely* many values. Then a scenario tree, as in Figure 2, can be used to represent the way in which the stochastic variables $\xi_\mathbf{t}$ evolve: The root of the tree corresponds to the immediately observable, deterministic data, A_1, B_1, b_1 and c_1. The nodes

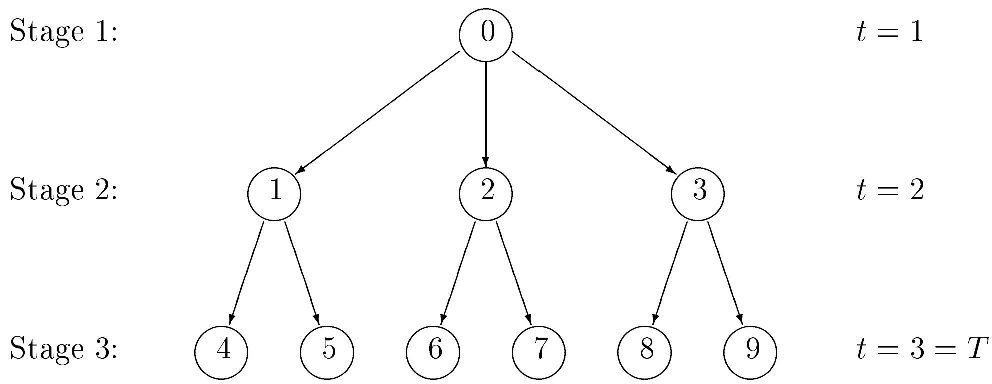

Figure 2: Scenario tree for a 3-stage program ($T = 3$) having 6 scenarios. In this example, ξ_2 has three possible realizations. ξ_3 has two possible realizations for each realization of ξ_2.

of the tree at level $t \geq 2$ corresponds to possible realizations of $\boldsymbol{\xi_t}$. These set of possible realizations (or, equivalently, their probabilities P_t) at level t are, in general, dependent on the preceding observations of $\boldsymbol{\xi_2}, ..., \boldsymbol{\xi_{t-1}}$.

The nodes in the scenario tree are also associated with the sequential decision process, x_t, so that each node at level t corresponds to a decision x_t which must be determined at time t, and which could depend on $\boldsymbol{\xi_2}, ..., \boldsymbol{\xi_t}$ as well as $x_1, ..., x_{t-1}$. This process must, of course, be adapted to $\boldsymbol{\xi_t}$ in the sense that x_t cannot depend on the specific future events, $\boldsymbol{\xi_{t+1}}, ..., \boldsymbol{\xi_T}$ which are not yet realized.

In the formulation [**MS**], the constraints on x_t are influenced directly only by x_{t-1}, not by $x_{t-2}, ..., x_1$. This structure is not a loss of generality (by suitable choice of the state space), and is not imposed by our algorithm, but is used because it conforms more closely to the scenario tree, where quantities associated with a node at level t depends directly only on quantities associated with the predecessor node of the tree. The ancestor node is for a node v represented algebraically by the *ancestor mapping,* $a(v)$.

We need to introduce the notion of "scenario", and its relationship to the scenario tree. A *scenario* is defined as a possible realization of the stochastic variables $\xi_2, ..., \xi_T$. Hence, the *set of scenarios*, $\Omega = \{1, ..., S\}$, is in a one-to-one correspondence with the set of leaves of the scenario tree, and we associate scenario $s \in \Omega$ with the sth leaf of the scenario tree. With this correspondence, we use the notation $a(s), s \in \Omega$, to denote the ancestor of the sth leaf node. The probability of scenario s is denoted by $p^s > 0$.

4.3 The Split-Variable Formulation

We introduce now the *split-variable formulation* of [**MS**], see, e.g., Rockafellar and Wets [28] or Ruszczynski [29]. The split-variable formulation associates with each scenario $s \in \Omega$ and each stage $t = 1, ..., T$ a set of decision variables, $x_t^s \in \Re^{n_t}$, which replace the stochastic variables x_t used in [**MS**]. For each scenario $s \in \Omega$, define the (deterministic) *scenario subproblem:*

$$
[\mathbf{SUB(s)}] \qquad \begin{array}{llcl}
\min\limits_{x_t^s, t=1,\ldots,T} & c_1^s x_1^s + c_2^s x_2^s + \cdots + c_T^s x_T^s & & \\
\text{s.t.} & A_1 x_1^s & = & b_1, \\
& B_2^s x_1^s + A_2^s x_2^s & = & b_2^s, \\
& \qquad B_3^s x_2^s + A_3^s x_3^s & = & b_3^s, \\
& \qquad\qquad \ddots & & \vdots \\
& \qquad\qquad\qquad B_T^s x_{T-1}^s + A_T^s x_T^s & = & b_T^s, \\
& 0 \le x_t^s \le u_t, \quad \text{for } t = 1, ..., T, & &
\end{array} \tag{25}
$$

where superscript s on the data, A_t^s, B_t^s, b_t^s and c_t^s, denote the unique realization of ξ_t associated with scenario s.

The scenario subproblems are obviously independent of each other. In order to obtain a formulation equivalent to [**MS**], we need to impose further constraints linking together the scenario subproblems. The *non-anticipativity constraints* impose the logical requirement that decisions up to stage t must coincide for those scenarios which have common root-to-leaf paths up to level t in the scenario tree. Letting a^{T-t} denote the $(T-t)$th power of the mapping a, the non-anticipativity constraints can be written:

$$
x_t^s = x_t^{s+1}, \quad \text{for } t = 1, ..., T \; \text{ and } s \in \Omega : a^{T-t}(s) = a^{T-t}(s+1). \tag{26}
$$

The complete set of scenario subproblems, together with the non-anticipativity constraints, define a program equivalent to [**MS**]. The split-variable formulation is:

$$
[\mathbf{SPLIT}] \quad \begin{array}{llcl}
\min\limits_{x_t^s, t=1,\ldots,T} & \sum\limits_{s \in \Omega} p^s \left(c_1^s x_1^s + c_2^s x_2^s + \cdots + c_T^s x_T^s \right) & & \\
\text{s.t.} & A_1 x_1^s & = & b_1, \; s \in \Omega \\
& B_2^s x_1^s + A_2^s x_2^s & = & b_2^s, \; s \in \Omega \\
& \qquad B_3^s x_2^s + A_3^s x_3^s & = & b_3^s, \; s \in \Omega \\
& \qquad\qquad \ddots & & \vdots \\
& \qquad\qquad\qquad B_T^s x_{T-1}^s + A_T^s x_T^s & = & b_T^s, \; s \in \Omega
\end{array} \tag{27}
$$

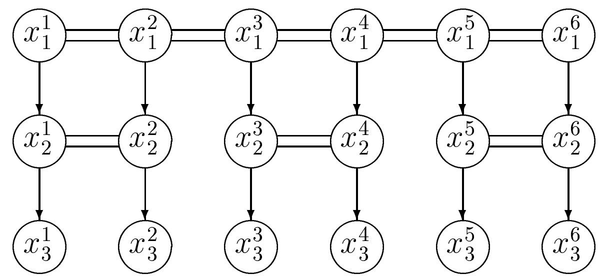

Figure 3: Illustration of the split-variable formulation corresponding to the scenario tree of Figure 1. Variables associated with each scenario/stage combination are shown. Horizontal double-lines indicate non-anticipativity (equality) constraints among these variables.

$$0 \le x_t^s \le u_t, \quad \text{for } t = 1, ..., T, \text{ and } s \in \Omega$$
$$x_t^s = x_t^{s+1}, \quad \text{for } t = 1, ..., T, \quad \text{and } s \in \Omega : a^{T-t}(s) = a^{T-t}(s+1).$$

The split-variable formulation is attractive for any algorithmic framework where the non-anticipativity constraints can be temporarily ignored while the S independent subproblems are solved, possibly simultaneously and in parallel. This is the case with the progressive hedging algorithm of Rockafellar and Wets [28] as well as with the augmented Lagrangian decomposition method of Ruszczynski [30], the diagonal quadratic approximation algorithm of Mulvey and Ruszczynski [19] and the proximal minimization algorithm we study here.

(((For more discussion of the structure and formulation of multistage models, see Gassmann [10] or Nielsen and Zenios [24]. For a state-of-the art overview see Birge [3]. Infanger (1994) [12] is a monograph on the use of importance sampling within Benders Decomposition.)))

5 Immunization: Present Value, Duration, Convexity

The concepts of "Duration" and "Convexity" have a long history in asset-liability management, dating back to Macaulay [14], who in 1938 proposed duration as a measure of the sensitivity of a bond portfolio's value to changes in interest rates.

We consider again a company (such as an insurance company) which has a future liability stream, and which wants to fund these by a bond (or some other fixed-income) portfolio. We look for an asset portfolio which has sufficient value to fund the liabilities, but we relax the requirement of cash-flow matching. Instead, we require that the *sensitivity to changes in interest rates* is equal on the asset and liability sides. That is, we *immunize* against interest rate changes.

Many insurance (and other) companies got into trouble in the late 1970's and early- to mid-eighties where interest rates (in the U.S) rose. Consider the following example. An insurance company in 1985 sold a GIC (Guaranteed Investment Contract), worth \$1,000,000, maturing in 1992 at a rate of 5%. That is, in return for the investment, the insurance company promises to pay $\$1,000,000 \cdot 1.05^7 = \$1,407,100$ in 1992, 7 years later. To ensure this liability, the insurance company (in this simple example) purchased the highest-yielding zero-coupon they could find, a 30-year, yielding 7%. The face value of this investment was $1,000,000 \cdot 1.07^{30} = 7,612,300$.

In 1992, the general level of interest rates had risen by 2%, so the (now 23-year) zero now had a yield of 9%, and hence the company's investment was worth $7,612,300 \cdot 1.09^{-23} = 1,048,800$, — 358,300 less than the liability! A very modest change in interest rates lead to a large shortfall, or asset-liability mismatch. The problem here was that the assets and liabilities were *duration-mismatched*, i.e., reacted very differently to changes in interest rates. This problem is precisely what immunization attacks.

5.1 Present Value and Dollar Duration.

The key formulas we will be using tie together the future cash flows of a bond with its present value, or price. It's important to note that when talking about duration, an underlying assumption is that the yield curve, whatever its shape, only moves in "parallel" up or down. More advanced techniques (bucket or factor immunization) address the general case of yield curve shape changes. By "yield curve" we refer to the annualized yield of a zero-coupon, default-free (e.g., Treasury) bond with given maturity t, called y_t. A simplifying assumption would be a *flat term structure*, where all y_t are identical; we will assume a possibly non-flat term structure.

Let F_{it} be the cash flow (coupon and/or principal) of a bond i at time t, and let T be the set of time points where cashflows occur, measured from time 0 ("now"). Then the present value of the bond's cash flows is given by

$$P_i = \sum_{t \in T} F_{it}(1 + y_t)^{-t}. \tag{28}$$

Ideally, this value is also the price of the bond (although small deviations in richness/cheapness are common) - otherwise there would be arbitrage opportunities between the par and zero coupon bonds.

Notice that cash flows that arrive at different times are discounted by different factors. The *yield to maturity* of a bond is the "average" yield — constant through time — which would equate the right-hand-side of the expression above to the bond's price, i.e. the yield r_i which solves:

$$P_i = \sum_{t \in T} F_{it}(1 + r_i)^{-t} \tag{29}$$

with P_i given in (28), or being the observed market price.

We will be interested in protecting the value of a (bond) portfolio against changes in interest rates. As mentioned, we assume that interest rates only change "in parallel", i.e., that the whole yield curve shifts up or down by the same amount. How does a bond's value change in response to such changes?

For simplicity we write the bond's price as a function of its yield-to-maturity (YTM), and then ask how sensivite the price is to changes in YTM. This is equivalent to assuming a flat term structure with constant discount rate, r_i, and considering parallel shifts up or down. A natural measure of price-sensitivity is *dollar duration*, which is just the derivative of the price with respect to interest rates:

$$D_i^{dol} = \frac{dP_i}{dr_i} = \sum_{t \in T} -t \cdot F_{it}(1 + r_i)^{-(t+1)}. \tag{30}$$

Notice that dollar duration is *negative*: If interest rates *rise*, present value *drops*. Zenios et al [6] use k_i for dollar duration.

5.2 The Immunization Model

Let P_L be the present value of a future liability stream (calculated using the same expression as for P_i). We would like to construct a portfolio of bonds (from a bond universe $\mathcal{U}$) such that the portfolio's present value matches that of the liability stream:

$$\sum_{i \in \mathcal{U}} P_i x_i = P_L,$$

where x_i is the face value holdings of bond i. In addition, we would like to immunize the portfolio against *small, parallel shifts* in interest rates — in other words, make sure that for small shifts, the value of the portfolio remains nearly equal to that of the liabilities. Dollar duration matching achieves precisely this:

$$\sum_{i \in \mathcal{U}} D_i^{dol} x = k_L,$$

The classical immunization model (e.g., Zenios et al [6]) is an LP with the above two constraints, plus non-negativity conditions on the x_i. It is tempting to seek a cheapest portfolio which satisfies this, but that is generally meaningless, since the price of the portfolio, by the first constraint, is already fixed to the present value of the liability. To the extend that the present value and the price of the obtained portfolio differ, the model is (in a fair market) just picking mispriced bonds.

Another possibility is to seek to maximize the portfolio yield, which can be approximated as:

$$\text{Maximize } -\sum_{i\in\mathcal{U}} D_i^{dol} \cdot r_i \cdot x_i.$$

This expression is explained in [6], but it is useful to think about it as maximizing the yield of the portfolio (given by the bond YTMs, r_i) weighted by the duration of the bond (D_i^{dol}), i.e., how long the bond's yield affects the portfolio yield (see below for the connection between dollar duration and "average time" till we receive the bond cash flow). This objective is subject to the same objection as minimizing the portfolio price, since low price (possibly due to mis-pricing) is of course associated with appearant higher yield.

Immun-1:

(Simple Immunization Model)

$$\begin{array}{llll}
\underset{x\in U}{\text{Maximize}} & -\sum_{i\in U} D_i^{dol} \cdot r_i \cdot x_i & & \\
\text{Subject to} & \sum_{i\in U} P_i \cdot x_i & = P_L & \\
& \sum_{i\in U} D_i^{dol} \cdot x_i & = D_L^{dol} & \\
& x_i & \geq 0 & i \in U
\end{array}$$

5.3 Macaulay and Modified Duration

Dollar duration is not the only measure of price sensitivity used. The traditional, and more intuitive, measure of "duration" is Macaulay Duration (D_i^{Mac}).

Dollar duration is an absolute measure: If portfolios A and B have identically timed cashflows, but A's cash flows are twice as big as B's, then A has twice the dollar duration of B (and both are negative). It is often preferred to used a *relative* duration measure, where the timing and relative magnitude of the cash flows is what matter, not their scaling. This introduces *Modified Duration*,

$$D_i^{mod} = -D_i^{dol}/P_i = \frac{1}{P_i} \cdot \sum_{t\in T} t \cdot F_{it}(1+r_i)^{-(t+1)} \tag{31}$$

Note that portfolios A and B have the same modified duration. Modified duration is easy to use in estimating price changes when interest rates change. Consider (from Tuckman) a bond with $P_i = \$100$, a yield $r_i = 8\%$ and a modified duration of $D_i^{mod} = 5$. To estimate the change in price when rates increase by, say, 10bp, just multiply the rate change by modified duration: $\Delta P_i/P_i \approx 5 \cdot 0.001$ which equals 0.5%. Hence the price falls by (approximately) 0.5% to \$99.50.

Modified duration is measured in years, but unfornately, a single payment occurring at time t does not have modified duration t. *Macaulay-duration* has that property. It is defined by multiplying modified duration by $1 + r_i$:

$$D_i^{Mac} = (1 + r_i) \cdot D_i^{mod} = \frac{1}{P_i} \cdot \sum_{t \in T} -t \cdot F_{it}(1 + r_i)^{-t)} \tag{32}$$

Macaulay-duration can be viewed as a weighed average of the times to arrival of cash flows, where the weights $\frac{1}{P_i} F_{it}(1+r_i)^{-t}$ are proportional to the magnitude of the cash flow, and sum to 1 (why?). In estimating price changes using Macaulay-duration, one multiplies D_i^{Mac} not by the absolute interest rate change, but by the relative change (in $1+r_i$). The bond given above has $D_i^{Mac} = (1+r_i) \cdot D_i^{mod} = 5.4$ years, so if rates increase by 10bp from 8%, the price change is approximately $0.1\%/(1 + 8\%) \cdot D_i^{Mac} = \$0.50$ again. Macaulay-duration is a more intuitive concept than modified duration, but for our purposes any of the three measures are equally good in hedging interest-rate uncertainty.

5.4 Critique of the simple Immunization Model

The immunization model **Immun-1** has only two constraints, apart from non-negativity. This means that in general, only two bonds will be chosen (only two x_i will be positive). Typically the optimal portfolio is a "bar-bell" portfolio consisting of a very long bond (because they generally have the highest yield, hence contribute the most to the objective), and a very short bond (to match the liabilities' duration). This portfolio is unfortunate because it is *maximally exposed to shape risk*, i.e., the risk of a non-parallel term structure change. This is because the very long bond is extremely sensitive to changes in the long rate, but the short bond is insensitive. In addition, such a "bar-bell" portfolio may satisfy the two constraints *today*, but as soon as the short-maturity bond matures (say, within one year), assets and liabilities are suddenly terribly duration mismatched. A related problem is that the portfolio's duration changes rapidly when iterest rates change – that is, the derivative of the duration is large, so if rates change, the asset/liability sides rapidly become duration-mismatched.

5.5 Convexity

Duration matching protects against small, parallel shifts in interest rates, so that the asset and liability sides stay approximately balanced. Mathematically, we are providing a first-order fit between the present-value functions of the assets and liabilities. But more is needed to avoid "bar-bell"-type solutions. This is where *convexity* enters.

Convexity (a similar concept used historically is "dispersion", see [9]) is defined as

$$Q_i = \sum_{t \in T} t(t+1) F_{it} (1+r_i)^{-(t+2)},$$

and is the second derivative, $d^2 P_i / dr_i^2$, of present value with respect to yields. It turns out that the "bar-bell" portfolio has a very high convexity (known as "positive convexity", according to Fabozzi because it's something good!) By adding to the duration-matched portfolio the requirement that its convexity should be as small as possible, we obtain a portfolio with cashflows that are more centralized in time. But the convexity should not be *too* small: Intuitively because if the first bond cashflows arrive after the first liability we have to borrow, but there's a better, mathematical reason having to do with larger shifts in yields: By requiring that the asset convexity is no less than the liability convexity, we assure that if interest rates change, not only will the asset and liability values change roughly equally, but to the extend they differ, the assets will be worth *more*! (Think of two curves that first-order match at some point, but assets have a higher curvature than liabilities. At both sides of the point, assets have the higher value). In other words, we require the *net convexity* (of assets minus liabilities) to be positive: This is what we mean by "positive net convexity" (and this is why it is "good"!): If rates do change, the asset side comes out ahead of the liability side.

The above discussion then leads to the model which minimizes asset convexity, subject to 3 conditions: (1) present value match, (2) duration match, and (3) positive net convexity.

Immun-2:

(Immunization Model with Convexity)

$$\begin{array}{lllll}
\underset{x \in U}{\text{Minimize}} & \sum_{i \in U} Q_i \cdot x_i & & \\
\text{Subject to} & \sum_{i \in U} P_i \cdot x_i & = P_L & \\
& \sum_{i \in U} D_i^{dol} \cdot x_i & = D_L^{dol} & \\
& \sum_{i \in U} Q_i \cdot x_i & \geq Q_L & \\
& x_i & \geq 0 & i \in U
\end{array}$$

Be aware that this whole discussion centers on *parallel shifts* and resulting *smooth changes* in present values. Many derivative fixed-income instruments such as options, mortgage-backed securities, etc., have non-smooth response functions, which may make a local immunization unreliable. The current state-of-the art response to this problem is the use of interest-rate scenarios leading to stochastic models, or the use of formal term-structure models which, through pricing and cash-flow models, can be used for hedging and immunization, for instance in a multi-stage, stochastic program.

References: See Zenios, Ch. 1, 3.1–3.2 [6] for a complete statement of the immunization and factor models. Chapter 31 in [9], by P.E. Christensen and F.J. Fabozzi contains a good discussion of immunization, with examples. Tuckman, Part 3, [31] contains a more rigorous exposition, including Key Duration Matching (Part 2 outlines the modern approach of using a term structure model to rate changes). A more introductory approach is given in Bodie, Kane and Markus [1998], Ch. 16, [4] and similar texts.

Exercises: (1) Find the Macaulay-duration of a T-year, zero-coupon bond. (2) Show that the Macaulay-duration of a perpetuity (i.e., a bond paying only coupon payments, forever) with coupon c, using a constant discount rate r, is $(1+r)/r$.

5.6 Factor Immunization

The simple immunization model **Immun-1** can be extented to directly address other types of term structure changes than parallel shifts. The underlying concept is to model each type of change as a *factor*, or a vector that specifies how interest rates change. For instance, the first factor may be

$$a_{1,t} = (1,1,1,1,...,1)^T \in \Re^{30}$$

which, when added in some (positive or negative) multiple to the term structure corresponds to a parallel up- or down-shift; a second factor might be

$$a_{2,t} = (-15,-24,-13,-12,...,-1,0,1,...,+13,+14)^T \in \Re^{30},$$

which indicates a twist in the structure, and a third might be

$$a_{3,t} = (0,0.5,1,...,4,4.5,4,3.5,...,0.5,0)^T \in \Re^{30},$$

which is a kind of curvature change.

Assume that a number $j \in J$ such factors are given by $a_{j,t}$. We can now build a model which explicitly hedges against these types of term structure changes. Let y_t be the term structure. Then a bond price is given by

$$P_i = \sum_{t \in T} F_{i,t} \cdot (1+y_t)^{-t}. \tag{33}$$

Differentiate w.r.t. each y_t:

$$\frac{dF_i}{dy_t} = -\sum_{t \in T} F_{i,t} \cdot (1+y_t)^{-(t+1)} \cdot dr_t. \tag{34}$$

We can formalize the meaning of the $a_{j,t}$'s by noting that some multiple of them, F_j is added to y_t, hence the change in the term structure y_t is given by:

$$dy_t = \sum_{j \in J} a_{j,t} \cdot dF_j \tag{35}$$

which, substituted into (34), yields an expression for

$$f_{i,j} = \frac{dP_i}{dF_j} = -\sum_{t \in T} t \cdot a_{j,t} \cdot F_{i,t} \cdot (1 + y_t)^{-(t+1)}. \tag{36}$$

This number, $f_{i,j}$, is the *sensitivity of bond i to changes in the term structure, as given by factor j*, or the bond's *factor loading.* If we can build an asset portfolio which has the same sensitivity to all factors as the liability side, we are immunized against changes given by the factors. The model is now straightforward:

Immun-3:
(Factor Immunization Model)

$$\begin{array}{lllll}
\underset{x \in U}{\text{Maximize}} & -\sum_{i \in U} D_i^{dol} \cdot r_i \cdot x_i & & & \\
\text{Subject to} & \sum_{i \in U} P_i \cdot x_i & = P_L & & \\
 & \sum_{i \in U} f_{i,j} \cdot x_i & = f_{L,j}, & j \in J & \\
 & x_i & \geq 0 & i \in U &
\end{array}$$

where we again maximize (an approximation to) the portfolio's yield. Note that there is no constraint matching dollar duration (as in **Immun-1**). This can be achieved by including the first factor, $a_{1,t}$, shown above. This model also comprises *key-duration matching* models which seek to immunize against changes in individual rates, $y_{\tau_1}, y_{\tau_2}, ...$, by having vectors $a_{j,t}$ with just a single non-zero each (at $\tau_1, \tau_2, ...$).

This model is very flexible because the decision maker can include any factor of interest. But how many factors are "enough"? Dahl [5] has comments on this. In a study on US data, as few as 3 factors explained 99.9% of the term structure variation (these were level, steepness, and curvature factors), and in Denmark, 4 factors explained 99.6% of the variation: steepness, curvature, a short-maturity factor (up to 10 years), and a very short-term (4 years) twist-factor. The most significant factors, based on historical data, were found by statistical methods (principal component analysis).

A Algebraic Formulation of Stochastic Network Model

Below is given the complete, algebraic formulation of the stochastic model described in Section 3.3.1, and the notation and variables of the model are defined.

Two-Stage Stochastic Model for Funding SPDA Liabilities:

$$\text{Maximize} \quad \frac{1}{S}\sum_{s\in\mathcal{S}} U_\alpha(W^s/E) \tag{37}$$

$$\text{Subject to} \quad \sum_{i\in\mathcal{U}} y_{0i} + u_1^s - v_1^s = \text{C} \qquad \text{for all } s\in\mathcal{S}, \tag{38}$$

$$\sum_{i\in\mathcal{U}}((1-\gamma)z_{pi}^s - y_{pi}^s) - u_{p+1}^s + (1+r_p^s)u_p^s - (1+r_p^s+\delta)v_p^s + v_{p+1}^s = L_p^s \qquad \text{for all } s\in\mathcal{S},\ p=1,...,Y-1, \tag{39}$$

$$\sum_{i\in\mathcal{U}}(1-\gamma)z_{Yi}^s + (1+r_Y^s)u_Y^s - (1+r_Y+\delta)v_Y^s = L_Y^s + W^s \qquad \text{for all } s\in\mathcal{S}, \tag{40}$$

$$x_{1i}^s = y_{0i} \qquad \text{for all } s\in\mathcal{S},\ i\in\mathcal{U}, \tag{41}$$

$$x_{(p+1)i}^s = m_{pi}^s x_{pi}^s - z_{pi}^s + y_{pi}^s \qquad \text{for all } s\in\mathcal{S},\ i\in\mathcal{U},\ p=1,...,Y-1 \tag{42}$$

$$z_{Yi}^s = m_{Yi}^s x_{Yi}^s \qquad \text{for all } s\in\mathcal{S},\ i\in\mathcal{U}. \tag{43}$$

All variables are non-negative.

Portfolio Construction and Rebalancing Variables:

y_{0i} : Contents of the initial portfolio (in cash value) of instrument i (first-stage variables).

y_{pi}^s : Purchase of instrument i at time τ_p under scenario s.

z_{pi}^s : Sale of instrument i at time τ_p under scenario s.

Auxiliary Variables:

x_{pi}^s : Holdings (in cash value) of instrument i during period p, under scenario s.

u_p^s: Amount of cash invested at the short rate in period p.

v_p^s: Amount of cash borrowed in period p.

W^s: Final wealth under scenario s.

Model Data:

C: The initial cash invested, consisting of the SPDA premium, P, and equity, E.

r_p^s: The short rate during period p under scenario s.

m_{pi}^s: The change in value of a \$1 holdings of instrument i in period p under scenario s.

L_p^s: The liability due at time τ_p under scenario s.

γ: Transaction cost (as a fraction of the transaction).

δ: The spread between borrowing and reinvestment rates.

References

[1] E. Adamidou, Y. Ben-Dov, L. Pendergast, and V. Pica. The optimal portfolio system – targeting horizon total returns under varying interest-rate scenarios. In S.A. Zenios, editor, *Financial Optimization*. Cambridge University Press, Cambridge, U.K., 1992.

[2] J.R. Birge. Decomposition and partitioning methods for multistage stochastic linear programs. *Operations Research*, 33(5):989–1007, 1985.

[3] J.R. Birge. Stochastic programming: Computation and applications. *INFORMS Journal on Computing*, 9(2):111–133, 1997.

[4] Bodie, Kane, and Marcus. *Investments*. Irwin, 1989.

[5] H. Dahl. A flexible approach to interest-rate risk management. In S.A. Zenios, editor, *Financial Optimization*. Cambridge University Press, Cambridge, U.K., 1992.

[6] H. Dahl, A. Meeraus, and S.A. Zenios. Some financial optimization models: I. risk management. In S.A. Zenios, editor, *Financial Optimization*, pages 3–36. Cambridge University Press, 1993.

[7] G.B. Dantzig and G. Infanger. Large-scale stochastic linear programs: Importance sampling and Benders decomposition. Technical report, Operations Research Department, Stanford University, Stanford, CA, 1991. Report SOL 91–4.

[8] R. Dembo. Scenario immunization. In S.A. Zenios, editor, *Financial Optimization*. Cambridge University Press, Cambridge, U.K., 1992.

[9] F.J. Fabozzi and I.M. Pollack, editors. *The Handbook of Fixed Income Securities*. Dow-Jones, Irwin, 1987.

[10] H.I. Gassmann. MSLiP: A computer code for the multistage stochastic linear programming problem. *Math. Prog.*, 47:407–423, 1990.

[11] J.C. Hull. *Options, Futures, and other Derivative Securities*. Prentice Hall, Englewood Cliffs, New Jersey 07632, 1993. Second Edition.

[12] G. Infanger. *Planning under Uncertainty: Solving large-Scale Stochastic Linear Programs*. Scientific Press, Danvers, MA, 1994.

[13] Jr. Ingersoll, J.E. *Theory of Financial Decision Making*. Studies in Financial Economics. Rowman & Littlefield, 1987.

[14] F. Macaulay. Some theoretical problems suggested by the movements of interest rates, bond yields and stock prices in the United States since 1856. Technical report, National Bureau of Economic Research, New York, 1938.

[15] L.C. MacLean, W.T. Ziemba, and G.Blazenko. Growth versus security in dynamic investment analysis. *Management Science*, 38(11):1562–1585, Nov. 1992.

[16] J. Mulvey. Incorporating transaction costs in models for asset allocation. In S.A. Zenios, editor, *Financial Optimization*. Cambridge University Press, Cambridge, U.K., 1992.

[17] J.M. Mulvey. Network models for cash flow analysis. *Journal of Cash Flow Management*, 4(1), 1984.

[18] J.M. Mulvey. Nonlinear network models in finance. *Advances in Mathematical Programming and Financial Planning*, 1987. JAI Press.

[19] J.M. Mulvey and A. Ruszczynski. A diagonal quadratic approximation method for large scale linear programs. *Operations Research Letters*, 12:205–215, 1992.

[20] J.M. Mulvey and H. Vladimirou. Stochastic network optimization models for investment planning. *Annals of Operations Research*, 20:187–217, 1989.

[21] J.M. Mulvey and H. Vladimirou. Stochastic network programming for financial planning problems. *Management Science*, 38(11):1642–1664, 1992.

[22] J.M. Mulvey and S.A. Zenios. Diversifying fixed-income portfolios: modeling dynamic effects. *Financial Analysts Journal*, pages 30–38, January/February 1994.

[23] S.S. Nielsen and E.I Ronn. A two-factor model for the valuation of the T-bond futures contract's embedded options. In F. Fabozzi, editor, *Advances in Fixed-Income Valuation Modeling and Risk Management*, pages 135–152, 1995.

[24] S.S. Nielsen and S.A. Zenios. Solving multistage stochastic network programs on massively parallel computers. *Mathematical Programming*, 73:227–250, 1996.

[25] S.S. Nielsen and S.A. Zenios. A stochastic programming model for funding Single Premium Deferred Annuities. *Mathematical Programming, Series B*, 75:177–200, 1996.

[26] S.S. Nielsen and S.A. Zenios. Scalable parallel Benders decomposition for stochastic linear programming. *Parallel Computing*, 23:1069–1088, 1997.

[27] D.R. Cari no, D.H. Myers, and W.T. Ziemba. Concepts, technical issues, and uses of the russell-yasuda kasai financial planning model. Technical report, Frank Russell Company, Jan. 1997.

[28] R.T. Rockafellar and R.J.-B. Wets. Scenarios and policy aggregation in optimization under uncertainty. *Math. of Operations Research*, 16(1):119–147, February 1991.

[29] A. Ruszczynski. Parallel decomposition of multistage stochastic programming problems. Working paper WP–88–094, IIASA, October. 1988.

[30] A. Ruszczynski. Regularized decomposition and augmented Lagrangian decomposition for angular linear programming problems. *Aspiration Based Decision Support Systems*, 1989. A. Levandowski, A. P. Wierzbicki (Eds). Springer Verlag.

[31] Tuckman. *Fixed Income Securities.* John Wiley & Sons, 1995.

[32] R.J-B. Wets. Stochastic programs with fixed resources: The equivalent deterministic problem. *SIAM Review*, 16:309–339, 1974.