ACTUARIAL RESEARCH CLEARING HOUSE
1999 VOL. 1

Portfolio Optimization in Corporate Models

William Babcock Steve Craighead
Nationwide Financial Services

November 1998

Abstract

Design and future maintenance of an asset portfolio backing a new line of business is critical for proper asset and liability management for that business. Most portfolio optimization methods utilize linear or quadratic programming and require the user to specify the asset and liability attributes and cash flows into the program. The programmer must also supply an objective function to allow the program to find the optimal asset mix for the associated liabilities. Two problems with this approach are that the liability cashflows are fairly static and that one must frequently rebalance the portfolio.

An alternative approach would be to develop a corporate model of both the assets and liabilities and incorporate various economic scenarios as input into the model. Asset strategies would be measured against a specific objective function that is calculated by the corporate model. Unfortunately, the majority of maximization algorithms available are very time consuming, and obtaining a reasonable portfolio mix becomes impractical. We overcome this difficulty by using a very rapid optimization method from the chemical engineering profession. We will include an example of the use of this process to determine an optimal portfolio. We will also discuss modifications of the algorithm that is required when the shorting of assets is not permitted.

Key Words:

Corporate Models, OAVDE, ROE, Portfolio Optimization, Floppy Triangle, Experimental Design, Stochastic Immunization

1 Introduction

Many asset portfolio optimization methodologies employ linear or quadratic programming to obtain a portfolio that best matches or immunizes a generally static set of liability cashflows. These methods are fairly computer inexpensive, and work well when the liability models are not extremely complex.

In the life insurance industry, most liabilities are very dynamic and complex, and proper pricing requires stochastic modeling. Dembo in [5] addresses stochastic portfolio design by using simple linear or quadratic programming on each scenario to find the best asset that matches that scenario's cashflows. Using a weighting system on each scenario, he associates the amount of that specific asset to be held in the overall portfolio. His method tends to avoid the barbell effect, which is observed with other optimization methods by not discarding any of his assets.

All optimization problems require limits or constraints. The insurance industry is very unique, and the following are industry-related constraints that should be included in any insurance portfolio optimization problem.

1. Risk Based Capital. The statutory requirements prescribe additional surplus be established for asset default, pricing risk, interest rate risk and business risk. The amount of surplus depending upon the level of the risk.

2. Reserve requirements. There are several methods to set up appropriate liability reserves, and generally the most cost-effective reserve utilizes company surplus most efficiently.

3. Statutory Limitations of Distributable Earnings. There is a statutory limitation of the amount that a stock insurance company can distribute to their stockholders.

4. Capital Requirements of a Company. Investments of an insurance company should exceed the companies cost of capital; otherwise the company is using its capital inefficiently.

Because of the above the actuary should seriously consider the efficient use of surplus. The analysis of surplus requires accurate asset and liability models that not only produce reliable asset and liability cash flows, but also properly

model reserves, risked based capital and statutory distributable earnings. From these models, the actuary needs to obtain the cash flow streams of the distributable earnings to either calculate the economic value of the stream or to determine the internal rate of return of the stream. This internal rate of return would correspond to the company's return on equity. Additionally, the actuary would use the models to obtain the best product design and asset mix that would have the greatest impact on the bottom line. Later in the paper we will describe a new business model that will maximize the return on equity while at the same time take into consideration the various risks associated with the distributable earning streams.

Taking the above approach creates complex models that require optimization on either the economic value or the return on equity. These sophisticated models can be very expensive models to produce results. The concept of expense is related to the length of time to create, audit and use the computer to process. Also, as the number of scenarios processed increase, the run times of the model also increase. Most optimization methods such as chief descents or Levenberg-Marquardt methods (see [12]) require extensive computer simulations to obtain the proper estimation of the gradients on the non-linear surface. One is not even guaranteed that the solution obtained by these optimization methods will be the best global solution. In fact the best global solution might not actually be the best business solution. This situation corresponds to the physical concept of a stable or unstable equilibrium. The best global solution may give you the highest value or highest return on equity, but could require a constant rebalancing of the portfolio to maintain the position. The solution would not be a stable solution that could easily move away from optimal to sub-optimal quickly. Whereas, the best business solution may be a product or asset mix that may not have the highest return but would give the highest stable return, without frequently rebalancing.

Due to all of these possible constraints, both theoretically and practically, a very good answer in a timely manner would have more value than the untimely best answer.

These constraints of time, expense and stability were reasons that chemical engineers in the 1960's developed a non-linear optimization method called the Floppy Triangle (hereafter denoted FT). The oddity of the name will become apparent from the geometric explanation given later in Section 2.

In Section 2 we will give a geometric description of FT. In Section 3 we will describe the process of setting up the initial experiments. In Section 4 we will describe the necessary risk and return metrics and define the optimization

target. In Section 5 we will describe the basics of the business models and the asset universe used. Here we will discuss the risk/return of the various test portfolios and the optimal portfolio obtained by the FT algorithm. In Section 6 we will use an efficient frontier method to display our experimental results. In Section 7 we will discuss our conclusions and further possible research.

2 Floppy Triangle

C. D. Hendrix [1] in his introduction of the FT algorithm to the chemical engineers at Union Carbide, says:

> Few people can think in terms of three or more independent variables acting simultaneously. Those who can are usually quashed by the surrounding "two-dimensional thinkers". As a result, one rarely finds projects in which more than two variables were investigated before a report was written. Hence the origin of the sequences: vary one or two variables, write a report. Then vary another variable (perhaps two), write another report. Each report discusses the effects of the variable upon selected responses rather than directing the effort to the objectives: Find the best combination of the variables.
>
> Variations on this theme include factorial and fractional factorial experiments. These methods are inadequate for systems of more than five or six potent variables. They will reveal effects of the variables upon each response.
>
> The matter is further complicated by the fact that variables may interact. That is, the effect of a variable upon a given response depends upon the levels of one or more other variables in the system. It then follows that the optimum level of any variable may depend upon the level of several other variables, and indeed that it may be quite meaningless to speak of the optimum of any single variable. It is clear that the simultaneous optimum of all such variables is most meaningful.

[1] This section is derived in part from C. D. Hendrix's paper "Empirical Optimization in Research and Development". This paper is a very old internal working paper for the Union Carbide Corporation. We will be glad to provide a copy of the original upon request.

It is well, even necessary, to know the effects of each and every variable in a system. But would it not be better to find the near-optimum combination of all variables simultaneously, then explore the effects of variables near the optimum?

He goes on to say:

This method begins with $k+1$ trials in k dimensions (k independent variables), those $k+1$ trials being arranged in the form of a regular simplex. The $k+1$ outcomes are ranked from best to worst. Here a full ranking is not necessary. Only the worst outcomes are of immediate interest. The worst trial is then rejected. A trivial calculation indicates a trial to replace the one rejected. The procedure is repeated sequentially, maintaining an inventory of $k+1$ trials at every stage.

A variation on this scheme was developed by J. S. Bodenschatz at Union Carbide. In this variation, the m worst trials are rejected and replaced by m new trials. The choice of m is at the discretion of the experimenter. However, some choices of m are better than others.

There are several advantages of the FT method. The question of "when to move" does not arise. Once the original set of $k+1$ trials is completed, a move is made every m trials. The time intervals between moves are thereby abbreviated, thus supporting a high level of interest...

The question of "where to move" is settled by the trivial calculation rule:

Twice the average of the best, minus the worst.

We will demonstrate the use of this rule in the following subsection. Hendrix goes on to say:

The FT method is not without shortcomings. The principle difficulty is that the method may lead in the wrong direction (or fail completely) if the gradient/error ratio is too low. A "false optimum" can arise if an outcome is fortuitously "good". These shortcomings have not greatly detracted from the advantages of FT. The simplicity of the FT methods requires little or no training in statistics or optimization methodology.

Next, we will use a simple geometric presentation that demonstrates the simplicity of the algorithm (as well as reveal why FT is called the "Floppy Triangle").

2.1 A FT Example in Two Variables

Consider a portfolio selection process with three separate assets to be purchased. Call the percentages of the assets purchased $X_1, X_2, 1 - X_1 - X_2$. The objective is to increase the rate of return. We will outline how to improve the combinations of X_1 and X_2. We will begin with three separate asset allocations arranged as a triangle in the two variables. Denote the three as points A, B, and C. See Figure 1.

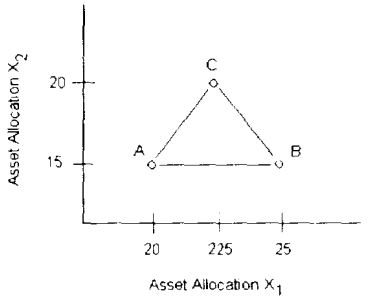

Figure 1:

Trial	X_1	X_2	$1 - X_1 - X_2$	Rate of Return
A	20%	15%	65%	7.4%
B	25%	15%	60%	7.6%
C	22.5%	20%	57.5%	8.2%

Table 1: Initial Portfolio Allocations

Example values of the rate of return are contained in Table 1. Note that point "A" is the worst of the three trials. We will discard "A" and obtain a new trial at "D". (Note how triangle ABC "flops" over into triangle BCD). See Figure 2.

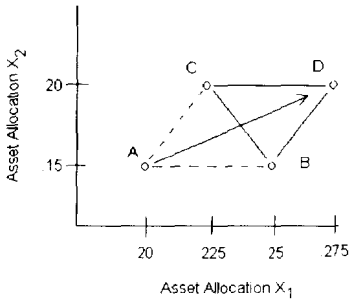

Figure 2:

Trial	X_1	X_2	$1 - X_1 - X_2$	Rate of Return
B	25%	15%	65%	7.6%
C	22.5%	20%	57.5%	8.2%
D	27.5%	20%	52.5%	8.5%

Table 2: Second Portfolio Allocation

Processing the allocation we obtain the result of the "D" allocation in Table 2. Note that because of the discarding of "A", the inventory of trials remains constant at three.

Now examine the current inventory of trials and select from these the worst point "B". See Figure 3.

Trial	X_1	X_2	$1 - X_1 - X_2$	Rate of Return
C	22.5%	20%	65%	8.2%
D	27.5%	20%	52.5%	8.5%
E	25%	25%	50%	8.9%

Table 3: Third Portfolio Allocation

As before, reflect the worst point "B" to a new trial point "E". Discard B from further consideration. Process asset allocation "E". See Figure 4 and Table 3.

Continuing the process, one will approach an optimal answer.

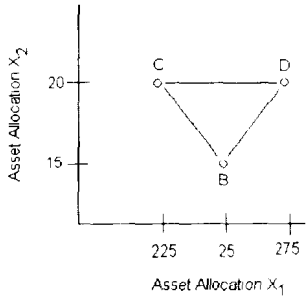

Figure 3:

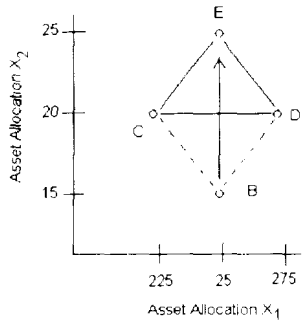

Figure 4:

To extend the FT to more than two variables requires one to use the following rule:

Twice the average of the best points, minus the worst point.

There are some difficulties associated with using FT and that of proper portfolio allocation. If the portfolio manager is allowed to short from the assets in the asset universe, the standard FT algorithm will be sufficient for designing the optimal portfolio. However, if shorting is not allowed the FT algorithm may create unreasonable asset allocations. A simple three asset allocation example will demonstrate this problem. In Figure 5, there

are three different asset allocations, A, B, and C. However, suppose that experiment "B" is the worst experiment in the inventory. The next step in the FT algorithm will cause the situation in Figure 6 to arise.

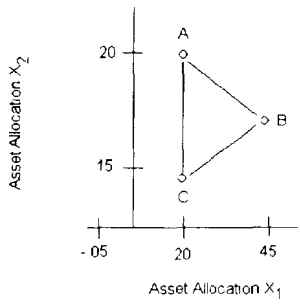

Figure 5:

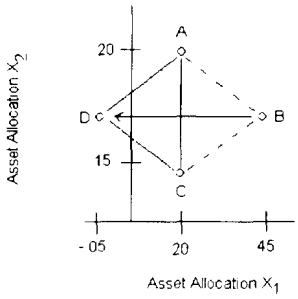

Figure 6:

We developed two methods to prevent the asset allocation from leaving the allowable allocation space. The first method is to use the standard FT algorithm, and if a specific asset's allocation goes negative, set that specific allocation to zero and adjust all of the other asset's allocations pro rata by the sum of all of the remaining positive allocations. Mathematically, assume $\{X_1, X_2, \ldots, X_n\}$ is the asset universe, and $\{a_1, a_2, \ldots, a_n\}$ is the asset

allocation such that $\sum_{i=1}^{n} a_i = 1$. If there exists an a_j such that $a_j < 0$, the new allocation will be $\{a_1, a_2, \ldots, a_{j-1}, 0, a_{j+1}, \ldots, a_n\} \frac{1}{\sum_{i=1}^{j-1} a_i + \sum_{i=j+1}^{n} a_i}$. This method effectively reduces the asset universe by the single asset with the negative asset allocation. This technique actually speeds up the optimization because the corporate model increases in speed as the asset universe is reduced. However, because of the reduction of the asset universe, portfolio allocations are created that replicate the problem of barbell portfolios as discussed in Dembo [5]. He discusses that most asset allocations obtained through the use of linear or quadratic programming lead to a small asset universe of one or two assets unless certain restrictions (e.g. position limits) are placed on the optimization algorithm. This also appears to occur when using this type of modification of the FT algorithm. (Note: We have observed that when shorting is allowed this problem does not occur with the FT. Here, the optimal portfolio mix leads to a mixture of all of the assets in the universe.)

The second method prevents the reduction of the asset universe and allows an asset allocation of all of the assets. The allocation may become very small for some of the assets, but position limits do not have to be forced upon the FT algorithm modification. Effectively the rule:

Twice the average of the best points, minus the worst point.

is replaced with

The square of geometric mean of the best points, divided by the corresponding coordinate value of the worst point.

This method effectively replaces the new experiment design by that of logarithms to prevent the possible occurence of a negative a_i. However, the allocations obtained purely by this method violate the condition of $\sum_{i=1}^{n} a_i = 1$. By adding the pro rata approach as in the first method, where the a_i are rescaled by the sum of the allocations, this problem is removed. This, of course, transforms the pure geometric interpretation of the FT into that of a FT algorithm with scaling. However, the algorithm is still effective in designing the subsequent experiments. This modified algorithm is the method that we used in our example in Section 6.

3 Initial Experimental Design

One disadvantage of the use of the FT algorithm is the need to develop the initial experiments. Neither of us are experts in the various forms of experimental design so we simply used the structure that Hendrix laid out in his paper [7]. His methodologies were to use certain types of experiment designs called Plackett-Burman plans or near-saturated or saturated factorials. Hendrix gives the following discussion and rules:

> As a rule, a large excess of trial points (beyond $k+1$ in k variables) will decrease the rate of progression. In spite of this, we have found it convenient to use near-saturated (rather than saturated) factorials and Plackett-Burman plans in lieu of simplexes in high dimensions. Hence we recommend the following:
>
> 1. if $k = 2$, use simplex.
> 2. if $k = 3$, use simplex or saturated factorial (same thing).
> 3. if $4 \leq k \leq 7$, use factional factorial in 8 experiments, or a simplex as seems appropriate.
> 4. if $8 \leq k \leq 11$, use Plackett-Burman plan in 12 experiments.
> 5. if $12 \leq k \leq 15$, use fractional factorial in 16 experiments.
> 6. if $k \geq 16$, use higher Plackett-Burman plans (or consider using supersaturated two-level plans.)
>
> Samples of each such plan are included here.
>
> 1. Simplex Plans (1),(2)
>
> In two variables, a simplex(triangle) is just this:

Variables			*Variables*	
A	*B*		*A*	*B*
−	−	, for example:	6	20
+	−		8	20
0	+		7	26

> where "0" implies average of *all above*.
> In three variables, a simplex(tetrahedron) is this:

Variables				*Variables*		
A	*B*	*C*		*A*	*B*	*C*
−	−	−	, for example:	6	20	37
+	−	−		8	20	37
0	+	−		7	23	37
0	0	+		7	21	41

Or, in three variables, a saturated factorial (a rotated simplex) is this:

Variables				*Variables*		
A	*B*	*C*		*A*	*B*	*C*
−	−	+	, for example:	6	20	41
+	−	−		8	20	37
−	+	−		6	23	37
+	+	+		8	23	41

2. Fractional Factorials (3), (5)

The initial pattern for four to seven variables can be developed by assigning "high" and "low" levels to the "+" and "-" signs in the following table.

	Variables						
	A	*B*	*C*	*D*	*E*	*F*	*G*
1	−	−	−	−	+	+	+
2	−	−	+	+	+	−	−
3	−	+	−	+	−	+	−
4	−	+	+	−	−	−	+
5	+	−	−	+	−	−	+
6	+	−	+	−	−	+	−
7	+	+	−	−	+	−	−
8	+	+	+	+	+	+	+

This initial pattern for 8-15 variables is developed from the following table in the same manner.

	Variables													
	A	*B*	*C*	*D*	*E*	*F*	*G*	*H*	*I*	*J*	*K*	*L*	*M*	*N*	*O*
1	−	−	−	−	+	−	−	−	−	+	+	+	+	+	+
2	−	−	−	+	−	−	+	+	+	+	+	−	+	−	−
3	−	−	+	−	−	+	−	+	+	+	−	+	−	+	−
4	−	−	+	+	+	+	+	−	−	+	−	−	−	−	+
5	−	+	−	−	−	+	+	+	−	−	+	+	−	−	+
6	−	+	−	+	+	+	−	−	+	−	+	−	−	+	−
7	−	+	+	−	+	−	+	−	+	−	−	+	+	−	−
8	−	+	+	+	−	−	−	+	−	−	−	−	+	+	+
9	+	−	−	−	−	+	+	−	+	−	−	−	+	+	+
10	+	−	−	+	+	+	−	+	−	−	−	+	+	−	−
11	+	−	+	−	+	−	+	+	−	−	+	−	−	+	−
12	+	−	+	+	−	−	−	−	+	−	+	+	−	−	+
13	+	+	−	−	+	−	−	+	+	+	−	−	−	−	+
14	+	+	−	+	−	−	+	−	−	+	−	+	−	+	−
15	+	+	+	−	−	+	−	−	−	+	+	−	+	−	−
16	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+

3. Plackett-Burman Plans (4),(6)

A Plackett-Burman plan in 12 experiments is outlined in the following table.

	Variables										
	A	*B*	*C*	*D*	*E*	*F*	*G*	*H*	*I*	*J*	*K*
1	+	+	−	+	+	+	−	−	−	+	−
2	−	+	+	−	+	+	+	−	−	−	+
3	+	−	+	+	−	+	+	+	−	−	−
4	−	+	−	+	+	−	+	+	+	−	−
5	−	−	+	−	+	+	−	+	+	+	−
6	−	−	−	+	−	+	+	−	+	+	+
7	+	−	−	−	+	−	+	+	−	+	+
8	+	+	−	−	−	+	−	+	+	−	+
9	+	+	+	−	−	−	+	−	+	+	−
10	−	+	+	+	−	−	−	+	−	+	+
11	+	−	+	+	+	−	−	−	+	−	+
12	−	−	−	−	−	−	−	−	−	−	−

4. Supersaturated Plans (6)

If there is serious question about the potency of some variable, there is evidence that a supersaturated plan (more variables than experiments) is appropriate. Two of many such two-level plans are shown here. The first is for up to 16 variables in 12 experiments, the second is for up to 24 variables in 12 experiments.

	Variables															
	A	*B*	*C*	*D*	*E*	*F*	*G*	*H*	*I*	*J*	*K*	*L*	*M*	*N*	*O*	*P*
1	+	+	+	+	+	+	+	+	+	+	+	−	−	−	−	−
2	+	+	+	+	+	+	+	−	+	−	−	−	−	−	−	−
3	+	+	+	+	+	−	−	+	−	−	+	+	+	−	+	+
4	+	+	+	−	−	−	+	−	−	+	−	−	+	+	+	+
5	+	+	−	−	−	+	−	−	+	−	+	+	+	−	+	−
6	+	−	−	−	+	−	−	+	−	+	+	+	+	+	−	+
7	−	−	−	+	−	−	+	−	+	+	+	+	−	+	+	+
8	−	−	+	−	−	+	−	+	+	+	−	+	−	+	−	+
9	−	+	−	−	+	−	+	+	+	−	−	+	+	+	+	−
10	+	−	−	+	−	+	+	+	−	−	−	−	+	+	−	−
11	−	−	+	−	+	+	+	−	−	−	+	−	−	−	+	+
12	−	+	−	+	+	+	−	−	−	+	−	−	−	−	−	−

	Variables																					
	A	*B*	*C*	*D*	*E*	*F*	*G*	*H*	*I*	*J*	*K*	*L*	*M*	*N*	*O*	*P*	*Q*	*R*	*S*	*T*	*U*	*V*	*W*	*X*
1	+	−	−	+	+	+	+	−	+	+	−	−	−	+	−	−	−	−	−	−	−	−	+	+
2	−	−	−	−	+	+	−	+	−	−	+	−	+	−	−	+	−	−	−	−	+	−	−	−
3	−	−	+	+	+	−	+	+	−	+	−	−	−	−	−	−	+	+	+	−	−	+	−	−
4	−	+	−	−	−	+	+	−	+	−	−	+	−	−	−	+	−	+	+	+	−	−	−	−
5	+	+	+	−	+	+	−	+	+	−	−	+	+	+	−	+	+	−	+	−	+	−	−	−
6	+	+	+	+	+	−	+	−	−	−	+	−	+	−	+	−	−	+	−	+	−	−	+	−
7	+	−	−	−	−	+	−	−	−	+	+	+	−	−	+	−	+	−	−	−	+	+	−	−
8	−	+	−	+	+	−	−	−	−	+	+	+	+	+	+	−	−	−	+	+	−	+	−	+
9	+	+	−	−	−	−	−	+	−	+	−	−	−	+	+	+	+	+	−	−	+	−	+	+
10	+	−	+	−	−	−	+	+	+	−	+	−	−	−	−	−	−	−	+	+	+	+	+	+
11	−	−	+	+	−	+	+	+	+	+	+	+	+	+	+	+	+	+	−	+	+	+	+	+
12	−	+	+	+	−	−	−	−	+	−	−	+	+	+	+	+	+	+	+	+	+	+	+	+

The actual construction of an initial table of experiments is illustrated using the Plackett-Burham plan in 12 experiments.

1. Assign variable names to the columns.
2. Consider Column "A". Select a low level and a high level of the first variable, "A". Assign the low level of "A" wherever

a (-) is shown in Column "A". Assign the high level of "A" wherever a (+) is shown in Column "A". *CAUTION: It is wise to be conservative at this point. Do not space the levels of the variables too widely.*

3. Repeat (2) for each variable under consideration.

NOTE: It is not necessary that each experiment be executed at precisely the prescribed conditions. The achieved conditions should be reported and used to progress to the next experiment(s).

In our example in Section 6, the raw asset allocations (before division by the sum of the raw allocations), were 80% for variables calling for "+" and 20% for the ones calling for "-". If we used 0% for "-", the FT algorithm would exclude that asset and reduce the asset universe, as discussed in Section 2.1.

We are exploring other experimental design methods including low discrepancy sequences [3, 6, 9, 11, 13], Latin hypercubes, or the new merger of the two–Latin supercube sampling [10].

4 Profit Metrics and Risk Measures–OAVDE, Expected ROE

In the first subsection, we will discuss various "bottom-line" profit measures, such as option-adjusted value of distributable earnings (OAVDE) or expected return of capital (ROE). These stochastic profit measures allow the actuary to measure the anticipated profits of the company by incorporating the effects of the embedded options in both the assets and liabilities.

In the second subsection, we will discuss various risk measures that quantify the risk in a stochastic pricing environment. Some of these are percentile estimators, and modifications of the standard deviation.

4.1 Profit Metrics[2]

Let us adopt the following notation:

[2]The following is based in part on Russ Osborn's article *Key Profit and Risk Measures: Definitions*

Let M be the number of projection months for the corporate model.

Let m be the projected month index. $m = 1 \ldots M$.

Let G_m be the net statutory gain at the end of the projection month m.

Let RS_m be the required risk surplus (also known as target surplus) at the end of projection month m.

Let $\Delta RS_m = RS_m - RS_{m-1}$ be the increase in required surplus in month m.

Let $NIRS_m$ denote the net (after-tax) investment income on required surplus in month m.

Let $DE_m = G_m - \Delta RS_m + NIRS_m$ be the distributable earnings at the end of month m.

Let N be the number of stochastic scenarios processed.

Let s denote the stochastic scenario index. $s = 1 \ldots N$.

Let p_s be the probability assigned to scenario s.

Let $DE_{s,m} = DE_m$ be the resultant distributable earnings for scenario s.

Let $\tau_{s,m}^{(2)}$ denote the gross short-term Treasury rate for scenario s in month m. This rate is in bond equivalent yield format(BEY).

Let $\tau_{s,m}$ denote the gross short-term Treasury rate for scenario s in month m. This rate is an annual percentage rate format (APR).

Let $r_{s,m}$ be the after-tax short-term Treasury rate for scenario s in month m. This is an effective rate.

Let TR denote the corporate tax rate.

Let ROE denote the expected return on equity. This is also known as the expected return on total capital (ROTC) or option-adjusted yield.

Let OAS represent the option-adjusted spread over risk-free Treasury rates.

Let $OAVDE$ denote the Option-Adjusted Value of Distributable Earnings.

Let Y be the target ROE (also known as the target ROTC).

Let S denote the target OAS.

$OAVDE$ represents the expected present value of future profit, where the expectation is taken over a probability space of stochastic model scenarios.

To find the appropriate short-term Treasury rates, we convert the nominal rates to annual effective rates and adjust for quarterly taxes as follows:

$$\tau_{s,m} = \left(1 + \frac{\tau_{s,m}^{(2)}}{2}\right)^{\frac{1}{2}} - 1, \text{ and}$$

$$r_{s,m} = \left\{1 + (1 - TR)\left[(1 + \tau_{s,m})^{\frac{1}{4}} - 1\right]\right\}^{4} - 1.$$

The formulas for $OAVDE$, ROE, and OAS will exhibit the following general form:

$$PresentValue(\bar{i}) = \sum_{s=1}^{N} p_s \left\{ \sum_{m=1}^{M} \frac{DE_{s,m}}{\left(\prod_{j=1}^{m}(1 + i_{s,j})^{\frac{1}{12}}\right)} \right\}, \tag{1}$$

where $\bar{i} = \{i_{s,m}\}$ is the given array of monthly interest rates by scenario by projection month. These rates are determined according to the purpose at hand, as described below.

Assuming the present value of future distributable earnings is never negative, we can compute profit measures as shown in Table 4. However, since some scenarios may have trailing negative distributable earnings, it would be improper to discount such future earnings at the assumed discount rates above. Instead, we need to discount future negative earnings in the same

To Compute:	Set $i_{s,m}$ equal to:	Solve Eq 1 for:	Assuming:
$OAVDE(Y)$	Y	$OAVDE$	$i_{s,m} = Y$ is constant
$OAVDE(S)$	$(\tau_{s,m} + S)$	$OAVDE$	S is constant
ROE	ROE	ROE	$PresentValue(i) = 0$
OAS	$(\tau_{s,m} + OAS)$	OAS	$PresentValue(i) = 0$

Table 4: Profit Measures

way we would discount future benefits to compute a reserve by using after-tax investment earnings rates. Therefore, we cannot write the present value computation as a simple summation. Rather, we must state the calculation as an algorithm that starts at time M and discounts backward month-by-month, iteratively discounting each month's value to the preceding month at the appropriate interest rate:

$$
\begin{aligned}
PresentValue(s, M) &= DE_{s,M} \\
PresentValue(s, m-1) &= DE_{s,m-1} + \frac{PresentValue(s,m)}{(1+\phi_{s,m})^{\frac{1}{12}}}, \text{for } m = M \ldots 1, \\
\text{where } \phi_{s,m} &= \begin{cases} i_{s,m} & \text{when } PresentValue(s,m) \geq 0, \\ r_{s,m} & \text{when } PresentValue(s,m) < 0. \end{cases} \\
PresentValue(i) &= \sum_{s=1}^{N} p_s PresentValue(s, 0)
\end{aligned}
$$

We use realistic random interest rates for our scenario set, and so we assign $p_s = \frac{1}{N}$ for all s.

The above discounting algorithm is discussed in Becker [1]. A less technical overview that explains the motivation behind this type of profits model is Becker [2].

4.2 Risk Measures[3]

The measurement of risk requires the portfolio manager to quantify the possible dispersion of results from the expected. The portfolio manager will

[3]The following is based in part on Russ Osborn's article *Key Profit and Risk Measures: Definitions*

use the risk metrics in one of two ways. The first is that they will produce a risk/return trade off graph (efficient frontier) where the various potential returns will be graphed against the level of risks measured. See Section 6 for our example of this graph. The other use is to use the metric in an optimization scheme to maximize a return metric of the portfolio while reducing the risk metric. This will design the portfolio that will best fit the situation being studied. This is the approach that we will take in this paper where we are using the FT algorithm as the maximization scheme. However, we will deviate from the common portfolio approach of finding the best asset portfolio that matches a liability cash flow. Here we will maximize the stockholder return metric while reducing the overall risk metric.

The various risk measures that we will discuss besides the sample standard deviations are partial sample standard deviation, percentiles and "comfort levels".

The formula of the partial sample standard deviation (denoted PARSTD) is:

$$PARSTD = \sqrt{\frac{\sum_{i=1}^{N} Min(X_i - \bar{X}, 0)^2}{n-1}} \tag{2}$$

The justification of the $Min(X_i - \bar{X}, 0)$ term is to make sure that the metric measures the dispersion of results associated with downside risk. The common standard deviation, when used to measure dispersion of results, includes values both above and below the mean. $PARSTD$ however only emphasizes the contribution to the dispersion due to the lower "tail" results. When trying to maximize return while reducing risk, the standard deviation is not the best risk metric. If one tries to maximize the return and reduce risk, the portfolio manager will be discarding potential upside profit if he or she uses standard deviation as the measurement of risk. PARSTD is a better indicator of the downside risk, where the measurement of risk by PARSTD does not include any potential upside profits. (Note: If a portfolio manager were trying to match a fund index exactly (e.g., S&P 500 Large Cap), the standard deviation would be a correct measure of the risk).

If $0 < p < 1$, then the $(100p)^{th}$ percentile of the probability distribution of a continuous variable X is a value ξ_p for which $Pr(X < \xi_p) = p$.

Suppose we would like to conservatively estimate a given percentile level of a distribution using data from a random sample, such that we have a certain level of confidence that we are not overstating the value of that percentile. In mathematical terms, we want to find $X_{p,c}$ such that $Pr(X_{p,c} \leq \xi_p) = c$

where c is the level of desired confidence, ξ_p is the true $(100p)^{th}$ percentile of the distribution and $X_{p,c}$ is termed the $[1-(100p)]\%$ "comfort level". For example, the 80% comfort level at a 98% confidence gives us a conservative value for the 20^{th} percentile of the distribution, with only a 2% probability that the true 20^{th} percentile is actually higher than the stated 80% comfort level. For $N = 50$, the 80% comfort level is given by the 5^{th} order statistic. See Hogg & Craig [8].

5 Description of the Business model

We utilized the FT algorithm on a single premium life insurance liability. Our asset universe consisted of noncallable corporate "A" rated bonds with various maturities ranging from one year to thirty years. We employed stochastic pricing on a statutory basis for 20 years and assumed that the policy credited interest rate would be determined at each policy anniversary. Also, we assumed that this credited interest rate is based on an asset portfolio net earned rate less a spread. Additionally, we assumed that policyholder lapses would only be the result of disintermediation, and lapses would occur when the competitor rate, (specifically the five-year Treasury plus a spread), exceeded the credited rate by a threshold. The pricing model purchased negative assets when cash was needed. This serves the same economic purpose as selling assets, except that statutory interest maintenance reserve accounting is avoided and no taxable event occurs.

The initial portfolio strategy consisted of a proportion of various corporate "A" rated noncallable bonds. Any reinvestments and disinvestments in the projection used the same initial investment strategy. The initial portfolio strategies are first found by the initial experimental design, then they are determined by the FT algorithm.

6 Efficient Frontier Results

As we discussed in Section 4, many profit and risk metrics could be used for optimization. When applying the FT algorithm we set the objective function to maximize the return on equity (ROE) and to minimize the PARSTD on the distribution of distributable earnings. To accomplish this, we had to rescale the PARSTD value in such a way to maximize the objective function

when PARSTD is minimized. This was accomplished by the following:

$$RS\_PARSTD = \frac{(PARSTD - MAXPARSTD)}{(MINPARSTD - MAXPARSTD)} \tag{3}$$

The above formula limits RS_PARSTD between zero and one, and it is maximizes RS_PARSTD as PARSTD is minimized. (Note: Here MAXPARSTD and MINPARSTD are initially estimated by the highest and lowest PARSTD values obtained from the initial series of experiments. These values are then "grossed up" to make sure that PARSTD does not go outside of the bounds MINPARSTD and MAXPARSTD.)

In a similar fashion, we had to rescale the ROE values obtained from the initial experiments. The formula used is:

$$RS\_ROE = \frac{(ROE - MINROE)}{(MAXROE - MINROE)} \tag{4}$$

Unlike PARSTD, ROE will be maximized if RS_ROE is maximized. This formula also assures that the value of RS_ROE will between zero and one. We determined MINROE and MAXROE in a similar fashion as MINPARSTD and MAXPARSTD.

We used the following objective function, which placed twice the emphasis on ROE.

$$Y = \left[(1 + RS\_ROE)^2)(1 + RS\_PARSTD)\right]^{\frac{1}{3}} - 1 \tag{5}$$

This objective function was designed in the same fashion as recommended by Hendrix [7].

We were somewhat surprised when the FT optimization determined that the best static investment strategy was effectively a barbell strategy. (We actually could have used all of the assets, however the sum of the two assets in the barbell covered over 99.9% of the allocation.) We did additional optimization experiments with other metrics, which confirmed the initial results.

Figure 7 shows the performance of a three-year and a ten-year barbell versus various bullet (or ladder) bond strategies.[4] As you can see, the point

[4]Unlike other barbell type results that increase risk with a barbell strategy, the static barbell strategy in this business model actually reduced the risk.

corresponding to the 3/10 barbell is above the efficient frontier determined by the ladder portfolios.[5] Note, in order to reduce experimental volatility, each maturity consisted of an equally allocated ladder. For example, a five-year corporate bond consisted of an equal weighting of four-year, five-year, and six-year maturity bonds. Similarly, the three-year/ten-year barbell, consisted of two-year, three-year, four-year, nine-year, ten-year and eleven-year bonds.

The explanation for the selection of the barbell portfolio include:

1. The differences of the corporate yield curve at various maturities.

2. We are assuming the initial asset strategy is used for all investments and disinvestments. For example, when the initial strategy is a three-year/ten-year barbell, the asset portfolio at the start of year four would consist of 25% three-year, 25% ten-year and 50% of a seven-year bond, which lengthens duration.

3. The nature of the interest rate generator. The generator produces yield curves from a realistic perspective, and the embedded risk premium within the generator would lead to a bias for longer assets.

4. Other interactions of assumptions in the business model.

5. The possibility that the model inefficiencies were optimized.

We found that the FT algorithm to be very effective in the selection of the optimal static asset allocation for our specific new business model. In fact, in our initial experiments, we used only fifty interest rate scenarios and obtained the barbell strategy. When we expanded the study to 250 scenarios, we observed that the barbell strategy still outperformed the various ladder strategies.

Our emphasis in the use of the FT algorithm was a preliminary foray into attempting to find optimal portfolio mixes that would maximize the stockholder's return. The use of our objective function 5 was in effect a utility function that we placed upon the business model to evaluate the profit position. We will continue our research in the optimal choice of assets using risk neutral pricing on the distributable earnings.

[5]This is not exactly an efficient frontier, since it is in the fourth quadrant, and it uses PARSTD instead of the standard deviation. These types of graphs occur when a product line underperforms from the target ROE.

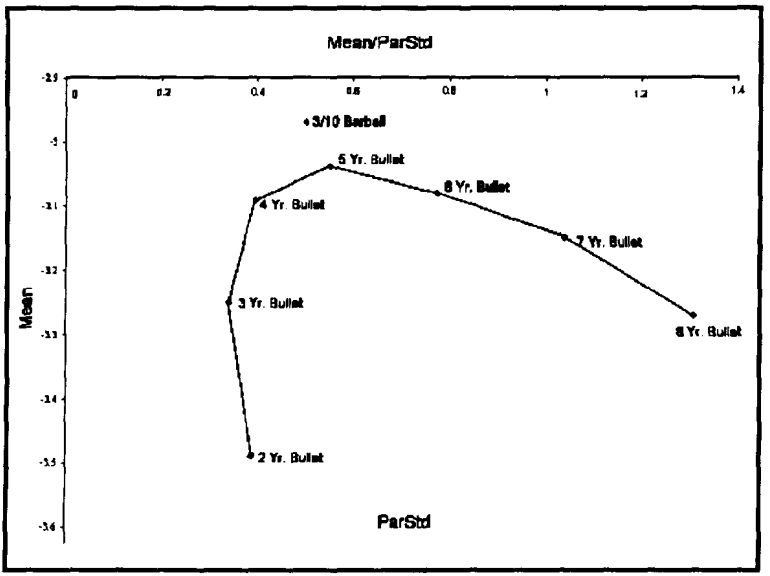

Figure 7:

7 Conclusions and Further Research

We were very pleased with the ease of use and rapid convergence of the FT algorithm. This method would be a wonderful addition to any actuary's optimization toolbox.

One of our next steps will be to employ a dynamic asset strategy throughout the projection period. We have examined various papers that are being used in the generalization of statistical regression. See [4, 14]. These methods discuss a method to graphically obtain the smallest dimensional model that does an adequate explanation of the data. We plan to use some of these methods to determine the optimal dynamic asset strategy to support a line of business.

The usage of Hendrix's experimental design is a very reasonable approach for most problems. However, due to the unique aspects of portfolio optimization, we are exploring alternative methods to construct the initial experimental design. One method we are considering is the use of low discrepancy sequences in the design phase.

We have found that the FT algorithm is a very effective optimization

tool, and we believe that this tool should become a standard within the actuarial community. We hope that its simplicity and utility will help others to optimize their decision making as well.

Acknowledgements

(Will Babcock) I would like to thank Stephen Hodges and Russ Osborn for their support throughout this project. I would like to especially thank Steve Craighead for his key insights, hard work, and dedication to this project.

(Steve Craighead) I would like to thank various discussions of the FT algorithm with Dr. Hongfei Zhang and Dr. John Marcsik. They have had make several insightful comments on the utility and limitations of the algorithm. I am also grateful to my co-author, Will Babcock, for going out on a limb and trying this strange but beautiful optimization method. I also deeply appreciate his business acumen, without which we would not have been able to produce the results that we did. I would also like to thank Stephen Hodges for allowing us the time to make these studies and the ongoing support that he and Russ Osborn have been in the research phase of this project. We would also like to thank Dave Raymond the project manager with encouraging us to continuing with these experiments, as well as his insight in how to explain the risks backing his line of business. Also I am grateful for Dr. Bostick Wyman's class at OSU that were guinea pigs in the instruction and utilization of the FT algorithm. Finally, I would like to thank my wife Patricia Craighead who had to take care of our five children (Samuel, Michelle, Bradley, Evan and Carl) while writing this paper.

References

[1] D. Becker(1988),"A Generalized Profits Released Model for the Measurement of Returns on Investment for Life Insurance",*TSA*. Vol. XL. pp. 61-114.

[2] D. Becker(1989),"A New Model for Investment Return on Investment",*Reinsurance Reporter*. Winter 1989.

[3] R. E. Caflisch and W. Morokoff(1996), *Valuation of Mortgage Backed Securities Using the Quasi-Monte Carlo Method*,August 8.

[4] R. Cook and S. Weisberg(1993) *An Intoduction to Regression Graphics*. New York: Wiley.

[5] R. S. Dembo(1993), "Scenario Optimization",*Financial Optimization*. edited by S. A. Zenios. Cambridge: Cambridge University Press, Cambridge, UK. pp. 290-308.

[6] J. H. Halton(1960), "On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals", *Numerical Mathematics*. Vol. 2, pp. 84-90.

[7] C. D. Hendrix, *Empirical Optimization in Research and Development*,Union Carbide Corporation.

[8] R. Hogg and A. Craig(1965) *Introduction to Mathematical Statistics, 2nd Edition*. New York: Macmillan Company. Equation 4, p. 180.

[9] H. Niederreiter(1992),*Random Number Generation and Quasi-Monte Carlo Methods*, CBMS-NSF Regional Conference Series in Applied Math., Vol. 63, SIAM, Philadelphia.

[10] A. B. Owen (1998), "Latin Supercube Sampling for Very High-Dimensional Simulations",*ACM Transactions on Modeling and Computer Simulations*. Vol. 8, No. 1, January. pp. 71-102.

[11] S. H. Paskov and J. F. Traub(1995), *Faster valuation of Financial Derivatives*, Journal of Portfolio Management, Fall.

[12] W. H. Press, B. P. Flannery, S. A. Teukolsky and W. T. Vetterling(1989),*Numerical Recipes in Pascal, The Art of Scientific Computing*,Cambridge University Press, New York. pp.574-580.

[13] S. Tezuka(1995),*Uniform Random Numbers: Theory and Practice*., Kluwer Academic Publ., Norwell, MA.

[14] S. Velilla(1998), "Assessing the Number of Linear Components in a General Regression Problem", *JASA*. September, Vol. 93. No. 443. pp. 1088-1098.