STOCHASTIC ANALYSIS OF UNIVERSAL LIFE PRODUCTS

Godfrey Perrott

ABSTRACT

Most of the discussion at AFIR Colloquia has focused on asset modelling, and in particular the term structure of interest rates. This paper presents a more practical aspect of stochastic modelling dealing with Universal Life plans as they are commonly sold in the U.S. This involves modelling both the asset and the liability sides. The asset modelling is fairly normal. The liability modelling requires modelling of behaviors that are variously inefficient including interest crediting strategy, premium suspension, policy loan activity and policy surrender. The paper explores how these activities may be modelled.

1. INTRODUCTION

This paper addresses stochastic modelling of Universal Life. This is a practical application of asset modelling theory that has been extensively covered at AFIR Colloquia. The paper focuses on the problems of liability modelling where the policyholder's options are complex and not exercised efficiently on a purely economic basis. The balance of the paper is divided into four main sections:

- Description of Typical UL Products in the U.S.
- Modelling of In Force
- Scenario Creation
- Behavior Models

2. TYPICAL UL PRODUCTS IN THE U.S.

Universal Life products have been common in the United States since around 1980. The earliest and most common policies are general account products. The investments are part of the insurer's general
account (although the Universal Life portfolio may well be a separate segment of that account for management purposes) and the investment results are not reflected directly in the policy mechanics. Instead the policyholder fund (account value) receives interest at an interest crediting rate which is declared from time to time by the company. More recently there have been Variable Universal Life products where the investments are made in a separate account, and the investment results reflected directly in the policy mechanics. This paper does not address these products which generally have less disintermediation risk.

The basic flexible premium Universal Life policy has the following mechanics.

- A premium is paid periodically by the policyholder; the amount of the premium is not fixed but may be subject to constraints such as a maximum and minimum if any premium is paid.

- The premium (frequently less an expense load) is added to the policyholder's account value.

- Interest is credited to the policyholder's account value monthly at the declared interest rate. This declared rate may be changed at any time and is subject to a minimum specified in the policy. Most policies include a policy loan provision which allows the policyholder to borrow against the cash value with no other collateral or recourse. The interest rate credited on the portion of the account value that is loaned is usually 1% to 4% less than the interest rate charged on the policy loan. Some policies credit only the guaranteed interest rate on the first thousand dollars of non-loaned account value.

- Mortality and expense charges are deducted from the policyholder's account value monthly. The mortality rate used is a rate declared by the company from time to time subject to a maximum rate specified in the policy. The expense rate may be a declared rate but is more usually fixed.

- The insurance benefit is either a level amount (subject to tax qualification), or a level amount plus the account value. These are commonly referred to as type 1 and type 2 policies respectively. Under a type 1 (and occasionally type 2) policy the insurance amount may increase at later durations to preserve the insurance to cash value ratio required to qualify the policy as life insurance. (The U.S. has minimum standards for a life insurance policy to qualify for favorable tax treatment.)
• The mortality charge is always calculated by applying the declared mortality rate to the net amount at risk.

• The cash surrender value is equal to the account value less a surrender charge. The surrender charge is normally a dollar amount per thousand dollars of insurance which varies by age and duration and usually does not vanish until a relatively high duration such as 15 or 20.

Policies which deduct an expense load from the premium are referred to as front-end loaded policies. Policies that deduct a surrender charge from the account value are referred to as back-end loaded. Pure front-end loaded policies were quite common in the early 80's. The typical policy today is back-end loaded with or without a modest front-end load. The typical interest rate and mortality guarantees will be the appropriate valuation rates.

Some policies lapse with no value when the account value is insufficient to pay the next mortality and expense charge while others lapse when the cash value is insufficient. Products that guaranty insurance for a certain period (such as 10 years) provided a minimum premium is paid each year are quite common.

Typical policies today include a target premium provision where the target premium is roughly at a whole life premium level. First year commissions are quite high up to the target premium (for example 80%), and much lower in excess of the target premium (for example 3%). Renewal commissions are typically around 3% regardless of size of premium with a higher percentage of target premium common in years 2 to 4.

There are a variety of bells and whistles designed to enhance the persistency or to enhance illustrations. There may be an interest bonus where the interest credited rate is increased by a half to 1% after ten years on policies that have paid at least the target premium in each year. Often this interest credited rate may be tied to some external index such as Treasury rates of a particular duration or some other external index.

The following table shows typical current rates in November 1992. (The 90 day treasury rate is 3%, the 5 year treasury rate is 6% and 1 to 10 year high grade corporates are 6.8%).
Many companies allocate investment income (at least partly) by the portfolio method. This will result in high credited rates after a period of decline in interest rates (such as the recent past). These interest credited rates will continue to fall if the yield curve stays about its current positions.

There are also various forms of fixed premium Universal Life.

The most common is called interest sensitive whole life and consists of a conventional whole life policy with a fixed periodic premium which also has a parallel Universal Life style accumulation mechanism. The cash value under this policy is the larger of the whole life cash value and the Universal Life style cash value.

3. MODELLING OF IN FORCE

One of the difficulties of any insurance modelling project is minimizing the number of cells (to minimize the time required to run the model) while maximizing how well the model represents the actual portfolios. Variables that need to be considered in selecting a model include:

- Distribution Channels and Market Segmentation - Universal Life is sold anywhere from the payroll deduction market (typically low premium, small amount, and economically unsophisticated) to highly sophisticated markets such as corporate owned life insurance (COLI).
and wealthy individuals. If a company is using more than one distribution channel it will almost certainly need to model Universal Life products differently since the behavior will be quite diverse.

- Major Plan, Issue Age and Issue Date - The typical company will have a large number of plans which vary only slightly. As with any insurance block of business the in force will frequently be concentrated in many fewer plans than are available. Small differences in plan can usually be ignored in modelling.

Issue age ranges must be selected similar to conventional modelling; the in force is typically divided into ranges (for example decennial) which are represented by central ages. The central age selection may be different than in conventional modelling. A major concern in modelling traditional business is reproducing reserves. This usually results in selection of a central age close to the mid-point of the range. For UL the policy mechanics are normally modelled directly. This means the account value will be modelled well. The reserve is usually modelled as the account value less an expense allowance which grades off over time. The most important variable to be reproduced by the central age selection is the mortality charge. The mortality rate curve by age is much more convex than the typical reserve factor curve. Thus fitting mortality will usually require a higher central age than would be used for traditional business.

Premium mode, smoker status, sex and death benefit option may also need to be modelled separately.

Issue date also requires modelling. Traditional modelling of non par products takes advantage of the fact that all policies of a particular plan and age cell are the same at a given duration, regardless of the issue year. This is not true for UL because of the mechanics of the UL policy; a 1988 policy in Duration 4 will have different values than a 1989 policy in Duration 4. Hence, issue date has to be modelled similar to issue age or plan.

- Premium level or account value level may require cells to be split since the operation characteristics of low premium cells will be different than the operational characteristics of high premium cells on the same plan of insurance. Low premium cells may operate more like level term whereas high premium cells will operate more like whole life or whole life with an annuity rider. This may be derived either from the premium per $1000 or the account value per $1000 of insurance.
4. SCENARIO CREATION

Effective stochastic modelling of Universal Life requires creation of scenarios similar to other investment products. However, a scenario has to consist of more than a set of yield curves for future periods. Maintenance expenses for Universal Life are higher (relative to premium) than SPDA’s and premium volumes tend to be lower so that the inflation in maintenance expenses become significant to the model.

It is even more difficult to estimate economic variables such as the unemployment rate. I believe this has a greater effect on premium suspension, policy loans and surrenders for lower premium policies than the competitiveness of the interest credited rate.

Figure 1 shows the unemployment rate, CPI inflation, and selected points on the yield curve, for the last 30 years.

It is not adequate to use a constant inflation rate to model expenses. A simple inflation model is the current 90 day Treasury Bill rate less 2.5% to 3%. This will probably perform acceptably when averaged over the long term but will not match the volatility that has been seen. A better model is to project the real rate of return (difference between the one year Treasury rate and the inflation rate) as a normally distributed random variable with a mean of about 2% and a standard deviation of 2.4% (see Figure 2).

The unemployment rate shows virtually no correlation to interest rates ($r = .2$ against inflation - see Figure 1). It can best be modelled by a random walk (since it shows strong continuity).

The change in unemployment is difficult to reproduce. However, representing it by a random variable with normal distribution, 0.0 mean and 0.3 standard deviation is a reasonable fit, and much better than doing nothing.

Interest scenarios may be generated by any of the acceptable methods. It is normally sufficient to use a quarterly period or even an annual period since the Universal Life policy is not as volatile as SPDA’s. There is still debate as to whether the best method is binomial or trinomial lattices, continuous generation or a representative set of yield curves together with transition matrices in a Markov chain model.
Fig. 1 Unemployment, Inflation, Treasury Rates
Fig. 2 Year Treasury vs. Inflation
5. Behavior Models

The most interesting and challenging area of modelling Universal Life is modelling company and policyholder behavior. This involves options that we know are not exercised efficiently from a purely economic viewpoint. Models are needed for:

- Market rates (what interest rates competitors are crediting)
- Interest crediting strategy
- Lapse surrender rate
- Policy loan rate
- Premium suspension rate

**Market Rates**

Average credited rate is usually modelled as the greater of the five year Treasury Note rate minus some spread and the 90 day Treasury Bill rate minus a different spread.

Figure 3 shows the relationship between a theoretical market rate calculated as described above and the actual average credited rate for the period 1984 to 1991. It is generally known that, during the early part of the interest rate decline, companies were allowing their spreads to reduce on universal life out of fear that to do otherwise would cause excess lapses. Later in the decline companies realized that they had to get their spreads back, and the rates resumed a more normal pattern. The best model we could obtain was to use spreads of 100 basis points off the five year Note and -100 basis points off the 90-day Bill adjusted by a lag term which depends on the number of quarters consecutive reduction in this index. The lag term is:

<table>
<thead>
<tr>
<th>Number of Consecutive Quarters of Reduction</th>
<th>Addition to Market Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>50</td>
</tr>
<tr>
<td>4</td>
<td>150</td>
</tr>
<tr>
<td>6</td>
<td>200</td>
</tr>
<tr>
<td>8</td>
<td>250</td>
</tr>
</tbody>
</table>
Fig. 3 Average UL Interest Rates

- Average UL Rate
- Market Rate
- Market with Lag
Interest Crediting Strategy

Interest crediting strategy involves a trade-off between the spread desired on the product and the competitive marketplace. Theoretically, the correct credited rate is that which will maximize the present value of future profits on a block of business. In practice, not only is this not done, but many companies cannot articulate their strategy for setting credited rates.

One approach to interest crediting strategy might involve two spreads: $S_1$ the spread assumed when pricing the product and $S_2$ the marginal spread required for a breakeven statutory position. $S_2$ will take account of the following components:

- Maintenance expenses,
- Excess of increase in reserve over increase in account value,
- Expected gain on surrenders, and
- Expected gain on mortality.

(Acquisition expense is not recognized in $S_2$ because it is a sunk cost).

For a purely front-end loaded product there will be no expected gain on surrenders or cost of reserve increase in excess of account value increase. For a back-loaded product (whether or not it has a front-end load) the statutory reserve will normally be between the cash value and the account value with all three converging as the surrender charge wears off. Thus, the reserve increase is higher than the increase in account value (resulting in a statutory charge each year) and a gain is realized on surrender as long as there is a surrender charge.

Once these spreads have been determined the interest crediting strategy can be easily formulated. If:

$I_i$ is the current earning rate for this block of business,

$I_m$ the current market rate, and

$I_c$ the credited rate then:

a. $I_c = I_i - S_1$ if $I_i - S_1 > I_m$

b. $I_c = I_m$ if $I_i - S_2 > I_m > I_i - S_1$

c. $I_c = I_i - S_2$ if $I_m > I_i - S_2$

Some companies will lower $S_2$ below the marginal breakeven point in the belief that this avoids the costs of excess lapse. This may be a valid strategy for SPDAs but it is not clear if it is as valid for Universal Life.
Total Lapse Rate

The total lapse rate must be determined by the model and then allocated between premium suspensions, policy loans and actual policy lapse or surrender. Almost all interest sensitive testing of universal life to date has used lapse rate formulas that were originally developed for SPDAs. A typical formula might be:

\[ \text{total lapse rate} = w + f(I_m - I_c - a)^2 \]

where:

- \(w\) is the basic lapse rate independent of any interest sensitive effect,
- \(I_m\) and \(I_c\) are the market and crediting rates expressed as percentages,
- \(a\) is an adjustment factor equal to \(1/4\) of the surrender charge expressed as a percentage of the account value but not more than \(I_m - I_c\) and not less than 0,
- \(f\) is a volatility factor (typically between 1 and 4),

and the term in parentheses is squared as an absolute value and retains its original sign.

Three examples will make this clearer (\(s\) is the surrender charge):

<table>
<thead>
<tr>
<th></th>
<th>w</th>
<th>f</th>
<th>I_m</th>
<th>I_c</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10%</td>
<td>1.5</td>
<td>8%</td>
<td>8%</td>
</tr>
<tr>
<td>2</td>
<td>10%</td>
<td>1.5</td>
<td>8%</td>
<td>5%</td>
</tr>
<tr>
<td>3</td>
<td>10%</td>
<td>1.5</td>
<td>0%</td>
<td>1.5%</td>
</tr>
</tbody>
</table>

Lapse rate 8.3% 10% 13.4%

This may be a reasonable model when applied to annuity products such as SPDAs, but it is not very satisfactory applied to UL. First of all the surrender charge percentage tends to be large enough that the lapse rate stays at the base rate for a relatively wide range of values of \(I_c\) compared to \(I_m\). Secondly, lapse activity on UL is driven by other variables than simply the interest rate environment. For the low premium or low account value forms of universal life lapse rate is somewhat correlated to unemployment. This correlation is probably also true in higher premium forms although not as pronounced.

Universal life policies are much harder to change than annuities if the policyholder wishes to retain the product but change to a different insurer to get a more attractive interest rate. Most life insurance is purchased (among other reasons) for its tax deferral. In the U.S., that tax deferral can be preserved if the policy is moved from one insurer to another, but the procedure is quite complicated and typically takes a
month or two to complete even after the prospective insured has been accepted by the new insurer. Both of these changes can be reflected by using a lapse rate with a lower volatility (for example $f = 1$) and an additional term dependent on the unemployment rate (for example $u - 6\%$ not less than 0 as an additive term).

The allocation of the total lapse rate between premium suspension rate, policy loan rate and actual lapse or surrender of the policy is difficult. Premium suspension rate in this description does not include planned premium cessation of policies sold on a vanish basis where a relatively small number of premiums (such as 6 or 7) are sufficient to fund the policy benefits based on the credited interest rate and charged mortality rate. All three of these rates are forms of policy lapse (the company takes in less money than expected or pays out money). However, they have different incidences since (for example) surrenders usually involve surrender changes whereas premium suspension does not. Premium suspension may be restricted by minimum premium requirements. Policy loans and surrenders may have different tax consequences.

On low premium and small size policies (such as payroll deduction) I believe that little policy loan activity will take place and that premium payment will either continue for the full amount or cease and that relatively small account values will be built up. Accordingly, it seems reasonable to model this type of business assuming no premium suspension, no policy loan activity, and that all lapses are actual surrenders.

For higher premium and larger amount policies all three activities will take place. The study of the particular company records may indicate a premium payment pattern that can be assumed. Policy loans have now largely been structured so that the insurance company does not have a negative spread risk. Thus, policy loans will be driven by cash flow needs which may realistically be coupled to the unemployment portion of the lapse rate. Policy surrender rates become the balancing item.

6. Conclusion

Modelling universal life policies (and by extension, other interest sensitive policies such as participating whole life in the United States) requires consideration of more variables than just the term structure of interest rates, and these variables are not easy to model directly. However, the modelling process can be improved significantly by reflecting
inflation and unemployment as separate random variables. Unemployment is not correlated to the yield curve. Inflation can be modelled by assuring the real rate of return is a random variable.

Econometricians would argue that the correct way to model any economic process is start out by projecting changes in GNP. If a stochastic GNP model were constructed, it should be reasonably easy to connect the unemployment rate and the inflation rate to the change in GNP. However, I believe it would be very difficult to connect the term structure of interest rates to the economic variables and problems similar to those discussed in this paper would occur from the opposite direction. Actuaries should talk to economists but we probably should continue to explore our own approaches.