1. Introduction

1.1 This short note shows how a survival model can be used to identify the major risk factors affecting persistency, i.e. the tendency for policies to either lapse or be made paid-up. Such a model can be used to either identify the highest-risk business for managing customer retention, or else in pricing to ensure new business is taken on profitably.

2. Model for persistency

2.1 We construct a model for the lapse intensity for life $i$, $\mu_i$, as follows:

$$\mu_i = \exp(\alpha_i + \beta_i x_i + \gamma_i r_i + \delta_i t) \quad (1)$$

2.2 where $x_i$ is the exact age of life $i$, $r_i$ is the duration since policy start and $t$ is calendar time. Each one of $\alpha_i$, $\beta_i$, $\gamma_i$ and $\delta_i$ is constructed from risk factors, i.e.

$$\alpha_i = \sum_{j=1}^{m} \alpha_j z_{ij}$$

2.3 where $\alpha_j$ is the effect of risk factor $j$ and $z_{ij}$ is an indicator variable taking the value 1 if life $i$ has risk factor $j$, and zero otherwise. Similar definitions apply to $\beta_i$, $\gamma_i$ and $\delta_i$, where $\beta_i$ is the age-related change in lapse intensity, $\gamma_i$ is the duration-related change and $\delta_i$ is the time-related change.

2.4 With an appropriate regression model, it is possible to identify a wide variety of statistically significant risk factors for persistency, including:

- Age
- Gender
- Lifestyle
- Duration
- Region
- Fund size
- Premium
- Distributor
- Agent

2.5 Using this kind of regression model, it is possible to account for the important sources of risk, and therefore any changes in business mix. Note that interactions between risk factors are possible: for example, $\delta_i$ is the time trend in lapse intensity, but this can vary by risk factor, e.g. separate trends for males and females.
3. Implementing the model in Longevitas

3.1 Longevitas fits a wide variety of models, including the model in Equation (1). Simply upload the data as normal and fit a survival model with the Gompertz* setting.

3.2 When fitting a model in Longevitas, all parameters which include the term \( \text{Age} \) are the components of \( \beta_i \). Similarly, all parameters with the term \( \text{Duration} \) are components of \( \gamma_i \) and all parameters with the term \( \text{Time} \) are components of \( \delta_i \). Note that Longevitas treats \( \text{Duration} \) and \( \text{Time} \) exactly like \( \text{Age} \). For example, the age-related component may be distinct for males and females. Within Longevitas, it is possible to fit distinct gender components for \( \text{Duration} \) and \( \text{Time} \) as well (if the data supports this).

4. Optimisation

4.1 A particularly useful feature of Longevitas is the factor optimiser. Simply put, this takes a variable with a large number of values and reduces it to a small number of homogeneous groups. This is particularly useful where a variable might have a large number of different values. Longevitas will work through the potentially thousands of model combinations and return with the optimal risk groupings.

4.2 The advantage for persistency is that you can use Longevitas to find out which distributors or agents bring you the poorest-persistency business (all other things being equal). Even if you have thousands of distributors, Longevitas would be able to separate them into — for example — three groups according to whether the business they introduce has low, normal or high lapse rates. Such information can be used when considering the terms on which you will do business with each class of distributors.

5. Business management

5.1 Once a model for persistency has been created, you can download the portfolio data tagged according to its risk factors. This includes optimisations across complicated factors such as distributor: if you have created a new risk classification based on distributor or agent, every record can be tagged with its persistency risk according to this new factor.

5.2 You can also download rate tables for use in business calculations, such as new-business pricing or setting commission terms for a distributor. Alternatively, for an in-force portfolio you could download individual modelled persistency rates for calculating embedded values at the per-policy level. If your model includes duration effects or time trends, then such rate tables would also include projections based on your experience to date.

6. Conclusion

6.1 For insurers with the necessary historical transaction records, it is possible to use survival models to deconstruct past lapse experience into the main risk factors and time trends. With this model, you can either identify the highest-risk in-force business for managing customer retention, or else you can use the model in pricing to ensure new business is taken on profitably.

6.2 Visit www.longevitas.co.uk for more information on the application of survival models to life insurance.