

Acceleration Techniques for Life Cash Flow Projection Based on Many Interest Rates Scenarios – Cash Flow Proxy Functions

Author: Martin Janeček, Tools4F, s.r.o. and Economic University in Prague, 2017

Acknowledgment:

I want to thank very much to my smart colleagues Lucie Fremundová and Dana Fidrmucová who kindly helped me with the development of the approaches described here, especially with their implementation into real calculation tools accompanying by valuable remarks.

Motivation

Majority of actuarial tasks in life insurance is based on projection of future cash flows from insurance contracts (existing contracts in-force as at the *valuation date* or/and future new contracts concluded after the *valuation date*; *valuation date* is the date when the projection starts).

Such tasks are for example:

- Accounting purposes:
 - Liability Adequacy Test (LAT) – i.e. testing of adequacy of the accounting value of life technical provision,
 - IFRS17 – where the technical provision is based on the present value of the future cash flows and also the expected cash flows themselves are crucial in the financial reporting.
- Solvency II:
 - valuation of:
 - best estimate value of life insurance liabilities (BEL),
 - fair value of life insurance liabilities,
 - stress testing of the BEL according to different assumptions of mortalities, lapses, expenses, interest rates, equities, etc. – all these are used for determination of SCR (Solvency Capital Requirement), ORSA (Own Risk and Solvency Assessment), internal risk management and many other applications.
- Other value and risk management purposes:
 - Assets Liability Management (ALM),
 - Embedded Value (EV) calculation,
 - Determination of the Value of New Business (VNB),
 - Profit Testing of new products,
 - Business Plans.

Traditionally, actuaries solve these tasks by running a projection model that calculates a development of the cash flows (and other variables) per each individual policy in scope to future periods based on defined inputs. Such inputs typically are list of contracts (existing and/or including future new business) and

assumptions of the future development (typically probabilities of deaths and lapse, unit costs, level of commissions, interest rates, etc.).

Let us call this detailed modelling approach as the "per-policy model" for the purposes of this paper.

Such calculations (using the per-policy model) are quite extensive. Let us imagine, a mid-sized life insurance company:

- with **500 000** policies in the portfolio;
- for which we want to project next 50 years (usually even more) on the monthly step; therefore, 50 years represents **600** months;
- every month, let us assume a calculation of **500** variables (usually 200 – 1000).

This already represents **150 x 10⁹** (500 000 x 600 x 500) **calculations of variables under one set of assumptions**. We will call such one calculation of the per-policy model (under one set of assumptions) as one "scenario run" or just a "scenario".

Obviously, already this one scenario run takes some time – it strongly depends on powerfulness of the software and hardware used and optimization of the projection model structure. Therefore, usually only a limited number of scenarios can be processed in a reasonable time by this per-policy model.

However, actuaries are more and more asked (by managers, regulators, auditors, etc.) to deliver results based on many more than units of scenarios, often hundreds or even (many) thousands.

Typically, such tasks are for example:

- valuation of Time Value of Financial Options and Guarantees (TVFOG) often included in life contracts, where results (usually the present value of the future cash flows – PVCF) of several hundreds of interest rates scenarios are needed;
- dynamic and stochastic ALM, where inter-relations between assets and liabilities are modelled on monthly (or annual) basis and specific investment strategy is decided dynamically based on the assets cash flow, the cash flow from policies and the actual investment returns.

Due to this quite complex calculation, usually even one dynamic scenario of the per-policy model described above is very time-consuming. And again, different stakeholders can ask for results under many scenarios – e.g. for the purposes of TVFOG valuation, optimization of the investment strategy, empirical distribution of the BEL and many others;

- internal model (in a sense of Solvency II) where again thousands of scenarios should be run.

To obtain results of these tasks, the runtimes of the per-policy model might be unacceptably long (several days, even weeks/months) even if the actuary employs professional software (as Prophet, Sophas, MoSes, ...) and hardware (powerful computers with many cores etc.).

This is why some alternative solution of how to obtain the life cash flows projection results (and other results based on the cash flows projection – typically present value of the cash flows (PVCF)) based on

many scenarios in reasonable time with acceptable deviation from the per-policy model results are to be investigated.

Objective

Within this paper, we will focus on two approaches how to determine (estimate) the future life insurance **cash flows** based on many **interest rates** scenarios in a reasonable runtime.

This is very important for many real life applications. In practice, usually the first task where it is required is the valuation of the TVFOG (required at least by the Solvency II).

We will call these techniques as "Analytical" and "Interpolation" approximations in this paper.

Both approaches have been developed within our team in the Tools4F actuarial consulting company during the years 2015 and 2016. We have also tested them on real life insurance portfolio with very good (for the Interpolation approach even with excellent) results.

Note:

Note that in the real life practice, we might be interested in approximation of several variables. Very often:

- a) average of PVCFs under many interest rates scenarios – to determine $BEL = E\{PVCF\}$,
- b) PVCF under each interest rates scenario individually (especially for the risk management purposes) – to obtain the full PVCF distribution ,
- c) cash flow at any time in future under any interest rates scenario – this is what we are focusing on here in this paper,
- d) other variables projection (premium, expenses, claims, technical provisions, profits, ...) under any interest rates scenario.

In this paper, we will focus on approximation of the future cash flows – ad c) – as:

- if we are able to estimate the future cash flow at any time t under any interest rates scenario, then we are automatically able to calculate the present value these cash flows (i.e. to obtain PVCF (as ad b)) and naturally also to calculate the average of these PVCFs to obtain the estimation of the BEL (as ad a)) and
- for ad d), the analogical approach as for the cash flows estimation could be applied.

Important note!

Note that the only type of interest rates that affects the future cash flows themselves are the rates that are the basics for the profit sharing decisions. The (cumulative) profit sharing amount is usually paid to the beneficiary at the policy termination (lapse, death, maturity, etc.). These rates follow the company profit sharing principles. Very often, and this is what we expect here in this paper, they depend on the market interest rates development. Typically, the profit sharing rate is based on the investment performance of the company assets portfolio (in accounting or market value view) or might be defined as the future risk free interest rate (potentially plus some spread) for a given maturity, etc.

There are other types of interest rates used in the life insurance calculations.

For example:

- for discounting of future cash flows (usually the risk free rates should be applied) or
- for projection of future accounting profits – investment returns (following accounting principles) should be projected.

However, these interest rates do not affect the cash flow itself.

Therefore, if not specified otherwise, the interest rates in this paper from now on, will mean the **interest rates that are credited to the policyholders'** funds as profit share plus a guarantee (if the guarantee is included in the policy definition).

Notation

We will use the following notation in this paper.

<i>valuation date</i>	represents the date from which the projection starts
t	the time in time unit (in practice usually 1 month) from the <i>valuation date</i> ; $t=0$ at the <i>valuation date</i>
n	total number of months projected, $t = 1, 2, \dots, n$
$\mathbf{i}^{(j)}$	interest rates scenario j , vector of interest rates $\mathbf{i}^{(j)} = (i_1^{(j)}, i_2^{(j)}, \dots, i_n^{(j)})$
$i_t^{(j)}$	the interest rate (of the scenario j at the time t) that represents the total interest rate that is credited to the policyholder (profit share + potential guarantee)
${}_r I_t^{(j)}$	${}_r I_t^{(j)} = (1 + i_{r+1}^{(j)}) \cdot (1 + i_{r+2}^{(j)}) \cdot \dots \cdot (1 + i_t^{(j)})$ $r \in \mathbb{N}$ and $r < t$
$I_t^{(j)}$	$I_t^{(j)} = {}_0 I_t^{(j)}$.
s	total number of interest rates scenarios considered, $j = 1, 2, \dots, s$
$CF_t^{(m),(j)}$	cash flow of the policy m at the time t under interest rates scenario j
m	the policy number, $m = 1, 2, \dots, M$
M	total number of policies within the portfolio
$CF_t^{ptf,(j)}$	cash flow of a portfolio of policies at the time t under interest rates scenario j $CF_t^{ptf,(j)} = \sum_{m \text{ (all policies)}} CF_t^{(m),(j)}$

$f_t^{(m),(j)}$	value of the fund of 1 policy m , at the time t under a given scenario j , if no decrements (deaths, lapses, etc.) are assumed.
$F_t^{(m),(j)}$	value of the fund of 1 policy m , at the time t under a given scenario j , taking into account the decrements.
$F_t^{ptf,(j)}$	value of the fund of the whole portfolio of policies, at the time t under interest rates scenario j , taking into account the decrements. $F_t^{ptf,(j)} = \sum_{m \text{ (all policies)}} F_t^{(m),(j)}$
$l_t^{(m)}$	number of policies in force at the time t , for a policy m
l_t^{ptf}	number of policies in force at the time t , for the whole portfolio; $l_t^{ptf} = \sum_{m \text{ (all policies)}} l_t^{(m)}$

There also might be some additional notations defined if necessary for the specific purpose in the following text.

Analytical approach

Usually, to obtain $CF_t^{ptf,(j)}$, the cash flow for the whole portfolio, a per-policy model calculates cash flows $CF_t^{(m),(j)}$ per each individual policy m and then sums it up (according to each projection month t).

It is,

$$CF_t^{ptf,(j)} = \sum_{m \text{ (all policies)}} CF_t^{(m),(j)} \quad (1)$$

Idea and Objective

The Analytical approach aims to find the analytical formula of $CF_t^{ptf,(j)}$ where:

- the only variable inputs are the interest rates;
More specifically, for a determination of the $CF_t^{ptf,(j)}$ at the time t , the interest rates $(i_1^{(j)}, i_2^{(j)}, \dots, i_t^{(j)})$ – i.e. until the time t – are the inputs. They all are known at the time t .
- other parts of the formula are coefficients that are not dependent on interest rates.

It is, to find the formula for $CF_t^{ptf,(j)}$ in the following form:

$$CF_t^{ptf,(j)} = CFfix_t^{ptf} + \sum_q [coef_t^{(q)} \cdot fn_t^{(q)}(i_1^{(j)}, i_2^{(j)}, \dots, i_t^{(j)})] \quad (2)$$

where

$$CFfix_t^{ptf} = \sum_{m \text{ (all policies)}} CFfix_t^{(m)} \text{ and}$$

$CFfix_t^{(m)}$ represents the value of the part of the cash flow at the time t for the policy m , that does not depend on the interest rates assumption; $CFfix_t^{(m)} \in R$;

$coef_t^{(q)}$ is a coefficient ($coef_t^{(q)} \in R$) relevant to the time t determined for the whole portfolio.

Note that both, $CFfix_t^{(m)}$ as well as $coef_t^{(q)}$ can be determined from results of one run of the per-policy model (no matter what interest rates scenario is assumed).

$fn_t^{(q)}$ is a function of interest rates at the time t ; same function for every policy at the time t and

q identifies the pair of related $coef_t^{(q)}$ and $fn_t^{(q)}$, $q = 1, 2, \dots$

If we are able to find such an analytical formula for $CF_t^{ptf,(j)}$, then we can calculate it out of the per-policy model (for any time t and scenario j). This may cause the calculation of the cash flows projection faster than when the per-policy model is used.

Example:

To be more illustrative, let us look closer at a specific example.

Let us consider:

- universal life type of policy where in case of death the insurance company pays
 - fixed sum assured (SA) plus
 - the value of the fund at the time of death including the cumulative profit share ($f_t^{(m),(j)}$).

Let us investigate the formula for the cash flow of the whole policy portfolio at time $t = 2$ (i.e. second projection month) – $CF_2^{ptf,(j)}$.

At the time $t = 2$, we know the interest rates of the specific scenario for the first two projected months – it is $(i_1^{(j)}, i_2^{(j)})$. Remember that both rates are the interest rates that are credited to the policyholder (profit share + guarantee (if exists)) – as mentioned above (*Important note* on the page 3).

It is,

$$CF_2^{ptf,(j)} = \sum_{m \text{ (all policies)}} \left[l_2^{(m)} \cdot P_2^{(m)} - l_2^{(m)} \cdot E_2^{(m)} - d_2^{(m)} \cdot (SA^{(m)} + f_2^{(m),(j)}) - w_2^{(m)} \cdot f_2^{(m),(j)} - mat_2^{(m)} \cdot f_2^{(m),(j)} \right] \tag{3}$$

Where (on top of the notation defined above in the part Notation):

$d_2^{(m)}$ number of policies considered to die within the second month, of the policy m ;
 $w_2^{(m)}$ number of policies considered to lapse within the second month, of the policy m ;
 $mat_2^{(m)}$ number of policies that are at maturity in the second month, of the policy m ;
 $P_2^{(m)}$ premium of 1 policy related to the second month, of the policy m ;
 $E_2^{(m)}$ expenses of 1 policy related to the second month, of the policy m ;
 $SA^{(m)}$ sum assured related to the policy m .

Let us now continue adjusting the formula (3) in order to separate interest rate sensitive and non-sensitive parts:

$$CF_2^{ptf,(j)} = \sum_{m \text{ (all policies)}} \left[l_2^{(m)} \cdot P_2^{(m)} - l_2^{(m)} \cdot E_2^{(m)} - d_2^{(m)} \cdot SA^{(m)} \right] + \sum_{m \text{ (all policies)}} \left[(d_2^{(m)} + w_2^{(m)} + mat_2^{(m)}) \cdot f_2^{(m),(j)} \right] \quad (4)$$

The first sum of the formula (4) is usually not dependent on the specific interest rates scenario – let us denote it $CFfix_2$.

It is,

$$CFfix_2^{ptf} = \sum_{m \text{ (all policies)}} \left[l_2^{(m)} \cdot P_2^{(m)} - l_2^{(m)} \cdot E_2^{(m)} - d_2^{(m)} \cdot SA^{(m)} \right]. \quad (5)$$

Let us now look to the second sum of the formula (4) and remember that the fund value at specific time t (here $t = 2$) could be determined as:

$$f_2^{(m),(j)} = (f_0^{(m),(j)} + pr_0^{saving,(m)}) \cdot (1 + i_1^{(j)}) \cdot (1 + i_2^{(j)}) + pr_1^{saving,(m)} \cdot (1 + i_2^{(j)}) \quad (6)$$

where

$pr_0^{saving,(m)}$ is the saving part of the premium that is paid by policy m at the time $t = 0$ and analogically $pr_1^{saving,m}$ relates to the time $t = 1$.

$f_0^{(m),(j)}$ is the starting fund value (i.e. the fund value at the valuation date) of 1 policy m . This naturally does not depend on the interest rates development in future and thus $f_0^{(m),(j)} = f_0^{(m)}$.

Then, we can write the formula (4) in the following form:

$$CF_2^{ptf,(j)} = CFfix_2^{ptf} + \sum_{m \text{ (all policies)}} \left[(d_2^{(m)} + w_2^{(m)} + mat_2^{(m)}) \cdot (f_0^{(m)} + pr_0^{saving,(m)}) \cdot (1 + i_1^{(j)}) \cdot (1 + i_2^{(j)}) \right] + \sum_{m \text{ (all policies)}} \left[(d_2^{(m)} + w_2^{(m)} + mat_2^{(m)}) \cdot pr_1^{saving,(m)} \cdot (1 + i_2^{(j)}) \right] \quad (7)$$

Let us now denote:

- $coef_2^{(1)} = \sum_{m \text{ (all policies)}} \left[(d_2^{(m)} + w_2^{(m)} + mat_2^{(m)}) \cdot (f_0^{(m)} + pr_0^{saving,(m)}) \right]$

- $coef_2^{(2)} = \sum_m (\text{all policies}) [(d_2^{(m)} + w_2^{(m)} + mat_2^{(m)}) \cdot pr_1^{saving,(m)}]$
- $fn_2^{(1),(j)} = (1 + i_1^{(j)}) \cdot (1 + i_2^{(j)})$
- $fn_2^{(2),(j)} = (1 + i_2^{(j)})$

and remember that:

- $fn_2^{(1),(j)}$ and $fn_2^{(2),(j)}$ do not depend on the specific policy, i.e. are the same for all the policies and thus can be extracted from the sum and
- $coef_2^{(1)}$ and $coef_2^{(2)}$ do not depend on interest rates and can be obtained from one run of the per-policy model (no matter of what interest rates scenario is assumed).

We therefore finally obtain the formula (4) in the form:

$$CF_2^{ptf,(j)} = CFfix_2^{ptf} + coef_2^{(1)} \cdot fn_2^{(1),(j)} + coef_2^{(2)} \cdot fn_2^{(2),(j)} =$$

$$= CFfix_2^{ptf} + \sum_{q=1}^2 coef_2^{(q)} \cdot fn_2^{(q),(j)} \quad (8)$$

that already is the desired form as in the formula (2).

Practical process:

In practice, the calculation process is as follows:

- 1) Derive the formulas for the functions $fn_t^{(q)}$ and $CFfix_t^{ptf}$ and $coef_t^{(q)}$. This is what is necessary to do just once for each of the insurance product until the per-policy model of the product is changed.
- 2) Every valuation date (usually every month) then:
 - a. Run the per-policy model once (with whatever interest rates scenario) and derive the values of the coefficients - $CFfix_t^{ptf}$ and $coef_t^{(q)}$ and save them.
 - b. Now, follow this cyclical calculation:

FOR $j = 1$ to s

FOR $t = 1$ to n

i. Take the interest rates $(i_1^{(j)}, i_2^{(j)}, \dots, i_t^{(j)})$ from the given interest rates scenario j (these are the inputs)

ii. Calculate $CF_t^{ptf,(j)}$ by the formula (8). It is,

$$CF_t^{ptf,(j)} = CFfix_t^{ptf} + \sum_q [coef_t^{(q)} \cdot fn_t^{(q)}(i_1^{(j)}, i_2^{(j)}, \dots, i_t^{(j)})]$$

Result of this cycle is the projection of $CF_t^{ptf,(j)}$ from $t = 1$ to the final future ($t = n$) for one specific interest rates scenario j .

Results of this cycle is the $CF_t^{ptf,(j)}$ projection for s interest rates scenarios.

Practical issues:

Although this approach could be seen as quite simple and clear with no deviations between the $CF_t^{ptf,(j)}$ from the per-policy model and from the approximation, quite often it is not easy to determine the analytical formula $CF_t^{ptf,(j)}$ exactly.

Typical issues for example are:

- products where the death benefit is defined as a maximum of the specific sum assured (SA) and the current (at the time t) value of the client's fund ($f_t^{(m),(j)}$). It is, the death benefit equals to $\max(SA, f_t^{(m),(j)})$. The complication here is that to estimate future death benefit part of the cash flow, we need to sum this max values per each individual policy. But for some policies $f_t^{(m),(j)} < SA$, for some the opposite holds $f_t^{(m),(j)} \geq SA$. At the same time, naturally, $f_t^{(m),(j)}$ is very much dependent on interest rates;
- function $fn_t^{(q),(j)}$ might be quite complex if (what is usually the case in practice) the profit share value is added to the fund value once per 12 months (usually at the end of the year);
- to derive the values of the parameters $CFfix_t^{ptf}$ and $coef_t^{(q)}$ the per-policy model often has to be extended heavily and also, number of parameters $CFfix_t^{ptf}$ and $coef_t^{(q)}$ might be huge (also tens of thousands). This is why the first run of the per-policy model from which all the parameters are derived might be quite time-consuming;
- the first time preparation is usually extremely demanding.
 - To succeed, it is required to:
 - understand all the details of the modeled life products;
 - understand the per-policy model formulas in details;
 - be advanced in mathematics to be able to derive the correct formula in the desired form;
 - be extremely patient – the formulas usually include many indices, sub-parts, members, etc. All (100%) of them must be correct, even one mistake in one index makes all the results incorrect;
 - have experience in life modelling and be smart to solve the practical issues;
 - be advanced and very careful in programming the approximation formulas – again, no mistake is allowed.

High-level results

Naturally, the results of this Analytical approach very much depends on specific company situation and cannot be generalized correctly.

Here, we comment briefly just basic results from our case study that was a mid-sized life insurance company, using Prophet tool, having large variety of life product types (with many insured persons, riders, funds, etc.).

Cash flow fit

The quality of the cash flow fit, i.e. the difference between the $CF_t^{ptf,(j)}$ from the approximation and the per-policy model very much depends on the product complexity. There are products (like in the example on the page 6 above) where the cash flow could be calculated exactly with no difference. However, for some products simplifications and partial alternative solutions must be developed.

In our case study, we managed to obtain the difference in vast majority of interest rates scenarios far below 1% every year, even for complex products and even if the tested interest rates scenarios were extreme (from -40% to +60%).

Present value of the cash flows fit

PVCF differences (per-policy model vs. approximation) in each of the individual scenarios were up to 0.5% ($\pm 0.5\%$).

Average of the PVCFs (BEL estimation) from the approximation were very much close to the value from the per-policy model (usually, the differences were below 0.1%).

Acceleration results

In our case, the calculation runtime decreased from 5 to 40 times (for 1 000 to 10 000 scenarios). More for higher number of scenarios as the most time consuming operation is to calculate the parameters of the Analytical approach that are based on one run of the per-policy model, but the model that is heavily extended from the standard one to obtain all the detailed results necessary for the parameters determination.

We also implemented the analytical formula calculations in MS Excel. The speed would be definitely be much higher if more proper tool would be used (e.g. R, Python, MatLab).

Conclusion to the Analytical approach

We see the Analytical approach as quite powerful with the following features:

- replicates very well the per-policy model cash flows and their present values;
- runtime might be significantly shorter if proper calculation tool (R, Python, Matlab, ...) is used;
- however, the practical difficulty is that the first time preparation (determination of the formulas for all the parameters and extending the existing per-policy model) is usually extremely demanding.

Interpolation approach

This approach again (same as for the Analytical approach) aims on finding an estimation of cash flow at specific time t under given interest rates scenario j for given portfolio of policies (e.g. one product) – $CF_t^{ptf,(j)}$.

Interpolation approach ideas:

The main ideas of this approach are the following:

- 1) Imagine now that we already have a set of s interest rates scenarios $\mathbf{i}^{(1)}, \mathbf{i}^{(2)}, \dots, \mathbf{i}^{(s)}$ for which we intend to calculate/estimate $CF_t^{ptf,(j)}$.

$(\mathbf{i}^{(j)} = (i_1^{(j)}, i_2^{(j)}, \dots, i_n^{(j)}))$ has the same interpretation as defined in the part Notation)

- 2) Let us build a new (fictive) scenario of interest rates $\mathbf{i}^{max} = (i_1^{max}, i_2^{max}, \dots, i_n^{max})$ that is determined for every time t as a maximum of $i_t^{(j)}$ from all the scenarios j .

It is, for every t , $i_t^{max} = \max_{\forall j} i_t^{(j)}$.

Under this fictive \mathbf{i}^{max} scenario, the projected future profit share credited to clients at each and every time t (based on i_t^{max}), will obviously be higher or equal to i_t^j under any scenario j (within our set of scenarios).

Therefore:

=> the future benefits payments (that include the cumulative profit share) will be the highest of all from the given set of scenarios and

=> the relevant cash flow, let us denote it $CF_t^{ptf,i^{max}}$ (if defined as incomes minus outcomes as usual), creates the lower boundary of all cash flows from all the scenarios within our set. It is, $CF_t^{ptf,i^{max}} \leq CF_t^{ptf,i^{(j)}}$ for any t and any j .

- 3) And vice versa. If we build a fictive scenario of interest rates $\mathbf{i}^{min} = (i_1^{min}, i_2^{min}, \dots, i_n^{min})$ that is determined for the time t as a minimum of $i_t^{(j)}$ from all the scenarios j ; it is, for every t ,

$i_t^{min} = \min_{\forall j} i_t^{(j)}$, then

=> the future benefit payments (that include the cumulative profit share) will be the lowest of all from the given set of scenarios and

=> the relevant cash flow, let us denote it $CF_t^{ptf,i^{min}}$, creates the upper boundary of all cash flows from all the scenarios within our set. It is, $CF_t^{ptf,i^{min}} \leq CF_t^{ptf,i^{(j)}}$ for any t and any j .

4) Therefore, for every interest rate scenario j , $\mathbf{i}^{(j)} = (i_1^{(j)}, i_2^{(j)}, \dots, i_n^{(j)})$ from the given set of scenarios and every time t , the relevant $CF_t^{ptf,(j)}$ will be positioned somewhere between $CF_t^{ptf,i^{max}}$ and $CF_t^{ptf,i^{min}}$.

It is, $CF_t^{ptf,i^{max}} \leq CF_t^{ptf,(j)} \leq CF_t^{ptf,i^{min}}$ for every t .

5) However, the range between $CF_t^{ptf,i^{max}}$ and $CF_t^{ptf,i^{min}}$ could be quite wide. Therefore, better specification of where to find $CF_t^{ptf,(j)}$ within this range is necessary.

6) Let us build a grid of z (also fictive) interest rates scenarios that is defined in the following way.

For every time t split the interval between i_t^{min} and i_t^{max} into $z-1$ equidistant intervals. Let us identify the specific grid scenario g_k where $k = 1, 2, \dots, z$.

Then for every time t the difference $i_t^{g_k} - i_t^{g_{k+1}}$ is constant for all k .

(Note that $i_t^{g_1} = i_t^{min}$ and $i_t^{g_z} = i_t^{max}$.)

Note: There might be other split of the interval between i_t^{min} and i_t^{max} , e.g. setting them according to quantile values (e.g. 10th, 20th, ..., 90th) for every time t . However, based on the experimental results, both options give similar results and thus the split approach is not crucial.

7) Then calculate the cash flows projection (cash flow for every time t) within the per-policy model for each of the grid scenarios of interest rates - CF_t^{ptf,g_k} . It means, the per-policy model will run z scenarios.

8) For every interest rates scenario j ($j = 1, 2, \dots, s$) and every time t we would like to find the $CF_t^{ptf,(j)}$ between the values of the cash flow at the time t of the "closest" grid scenarios g_k and g_{k+1} .

However, the following questions remain:

a) How to find, for each time t individually, the *closest* g_k and g_{k+1} ?

b) If we find them, how to interpolate between their cash flows to estimate properly the $CF_t^{ptf,(j)}$?
It is, what proportion of CF_t^{ptf,g_k} and $CF_t^{ptf,g_{k+1}}$ to take?

9) To solve these questions, let us find an alternative variable that satisfies both of the following two conditions:

a) It is highly correlated (under interest rates scenarios) to $CF_t^{ptf,(j)}$ (condition 1)

b) It is possible to be calculated easily and quickly out of the per-policy model. (condition 2)

10) If we find such a variable – let us denote it $V_t^{(j)}$ for now – then:

a) calculate the $V_t^{g_k}$ for every grid scenarios ($k = 1, 2, \dots, z$) and every time t (out of the per-policy model);

b) for every scenario j and for every time t :

i. calculate $V_t^{(j)}$ and find the grid scenarios g_k and g_{k+1} where $V_t^{g_k} \leq V_t^{(j)} \leq V_t^{g_{k+1}}$;

ii. see where the $V_t^{(j)}$ ranks between the V_t^{gk} and V_t^{gk+1} .

It is, determine the ratio $p = \frac{V_t^{gk+1} - V_t^{(j)}}{V_t^{gk+1} - V_t^{gk}}$;

iii. determine the estimation of the $CF_t^{ptf,(j)}$ as being between the grid values $CF_t^{ptf,gk}$ and $CF_t^{ptf,gk+1}$ using the same interpolation as for the value of $V_t^{(j)}$ ranking between V_t^{gk} and V_t^{gk+1} .

It is, $CF_t^{ptf,(j)} = p \cdot CF_t^{ptf,gk} + (1 - p) \cdot CF_t^{ptf,gk+1}$.

11) What could be the variable $V_t^{(j)}$?

- Let us consider the value of the clients' fund to be paid at the time t for the given scenario j for the considered portfolio of policies – let us denote it $F_t^{paid,(j)}$.
- For this variable, we expect high correlation with the $CF_t^{ptf,(j)}$ as the fund value paid is usually a significant part of the cash flow that is paid at the time t and is naturally dependent on the interest rates. This satisfies the condition 1 as defined above on the page 12.

12) $F_t^{paid,(j)}$ could be determined by the following calculation:

Let us define and remind (some of the definitions have already been set above in the part Notation):

$$f_t^{(m),(j)} = \frac{F_t^{(m),(j)}}{l_t^{(m)}}$$

gr guaranteed interest rate (theoretically could depend on time t , but usually it does not).
It equals to 0 for the products where minimum interest rate is not guaranteed.

$$f_t^{(m),gr} = \frac{F_t^{(m),gr}}{l_t^{(m)}} \text{.....fund value of 1 policy } m \text{ at the time } t \text{ with no decrements, for a scenario}$$

where all the future interest rates i_t (for every t) equals to the rate gr .

(Note that $f_0^{(m)} = f_0^{(m),gr}$ as this is the starting value of the fund at the $t=0$ (the *valuation date*))

$$\overline{f_t^{gr}} = \frac{F_t^{ptf,gr}}{l_t^{ptf}} \text{.....is the average value of } f_t^{(m),gr}$$

$$\overline{f_t^{(j)}} = \frac{F_t^{ptf,(j)}}{l_t^{ptf}} \text{.....is the average value of } f_t^{(m),(j)}$$

$$F_t^{ptf,(j)} = \sum_{m \text{ (all policies)}} F_t^{(m),(j)}$$

$$F_t^{ptf,gr} = \sum_{m \text{ (all policies)}} F_t^{(m),gr}$$

$$\Delta f_t^{(m),gr} = \frac{f_{t+1}^{(m),gr}}{1+gr} - f_t^{(m),gr}$$

(Note that the $\Delta f_t^{(m),gr}$ represents the saving part of the premium.)

$paid_t^{(m)}$ is the total number of policies that are expected to obtain benefit payments at time t related to the policy m ; usually it is a sum of number of policies expected to die, lapse or mature at time t .

Then the policyholder fund to be paid at the time t for the whole policy portfolio is:

$$\begin{aligned}
F_t^{paid,(j)} &= \\
&= paid_t^{(1)} \cdot \left[\left(f_0^{(1),gr} + \Delta f_0^{(1),gr} \right) \cdot I_t + \Delta f_1^{(1),gr} \cdot {}_1I_t + \Delta f_2^{(1),gr} \cdot {}_2I_t + \dots + \Delta f_{t-1}^{(1),gr} \cdot {}_{t-1}I_t \right] + \\
&+ paid_t^{(2)} \cdot \left[\left(f_0^{(2),gr} + \Delta f_0^{(2),gr} \right) \cdot I_t + \Delta f_1^{(2),gr} \cdot {}_1I_t + \Delta f_2^{(2),gr} \cdot {}_2I_t + \dots + \Delta f_{t-1}^{(2),gr} \cdot {}_{t-1}I_t \right] + \dots + \\
&+ paid_t^{(M)} \cdot \left[\left(f_0^{(M),gr} + \Delta f_0^{(M),gr} \right) \cdot I_t + \Delta f_1^{(M),gr} \cdot {}_1I_t + \Delta f_2^{(M),gr} \cdot {}_2I_t + \dots + \Delta f_{t-1}^{(M),gr} \cdot {}_{t-1}I_t \right] = \\
&= paid_t^{(1)} \cdot \left[f_0^{(1),gr} \cdot I_t + \left(\frac{f_1^{(1),gr}}{1+gr} - f_0^{(1),gr} \right) \cdot I_t + \left(\frac{f_2^{(1),gr}}{1+gr} - f_1^{(1),gr} \right) \cdot {}_1I_t + \left(\frac{f_3^{(1),gr}}{1+gr} - f_2^{(1),gr} \right) \right. \\
&\quad \left. \cdot {}_2I_t + \dots + \left(\frac{f_t^{(1),gr}}{1+gr} - f_{t-1}^{(1),gr} \right) \cdot {}_{t-1}I_t \right] + \\
&+ paid_t^{(2)} \cdot \left[f_0^{(2),gr} \cdot I_t + \left(\frac{f_1^{(2),gr}}{1+gr} - f_0^{(2),gr} \right) \cdot I_t + \left(\frac{f_2^{(2),gr}}{1+gr} - f_1^{(2),gr} \right) \cdot {}_1I_t + \left(\frac{f_3^{(2),gr}}{1+gr} - f_2^{(2),gr} \right) \right. \\
&\quad \left. \cdot {}_2I_t + \dots + \left(\frac{f_t^{(2),gr}}{1+gr} - f_{t-1}^{(2),gr} \right) \cdot {}_{t-1}I_t \right] + \dots + \\
&+ paid_t^{(M)} \cdot \left[f_0^{(M),gr} \cdot I_t + \left(\frac{f_1^{(M),gr}}{1+gr} - f_0^{(M),gr} \right) \cdot I_t + \left(\frac{f_2^{(M),gr}}{1+gr} - f_1^{(M),gr} \right) \cdot {}_1I_t \right. \\
&\quad \left. + \left(\frac{f_3^{(M),gr}}{1+gr} - f_2^{(M),gr} \right) \cdot {}_2I_t + \dots + \left(\frac{f_t^{(M),gr}}{1+gr} - f_{t-1}^{(M),gr} \right) \cdot {}_{t-1}I_t \right] = \\
&= paid_t^{(1)} \cdot \left[I_t \cdot f_1^{(1),gr} \cdot \left(\frac{1}{1+gr} - \frac{1}{1+i_1} \right) + {}_1I_t \cdot f_2^{(1),gr} \cdot \left(\frac{1}{1+gr} - \frac{1}{1+i_2} \right) \right. \\
&\quad \left. + {}_2I_t \cdot f_3^{(1),gr} \cdot \left(\frac{1}{1+gr} - \frac{1}{1+i_3} \right) + \dots + {}_{t-1}I_t \cdot f_t^{(1),gr} \cdot \frac{1}{1+gr} \right] + \\
&+ paid_t^{(2)} \cdot \left[I_t \cdot f_1^{(2),gr} \cdot \left(\frac{1}{1+gr} - \frac{1}{1+i_1} \right) + {}_1I_t \cdot f_2^{(2),gr} \cdot \left(\frac{1}{1+gr} - \frac{1}{1+i_2} \right) \right. \\
&\quad \left. + {}_2I_t \cdot f_3^{(2),gr} \cdot \left(\frac{1}{1+gr} - \frac{1}{1+i_3} \right) + \dots + {}_{t-1}I_t \cdot f_t^{(2),gr} \cdot \frac{1}{1+gr} \right] + \dots + \\
&+ paid_t^{(M)} \cdot \left[I_t \cdot f_1^{(M),gr} \cdot \left(\frac{1}{1+gr} - \frac{1}{1+i_1} \right) + {}_1I_t \cdot f_2^{(M),gr} \cdot \left(\frac{1}{1+gr} - \frac{1}{1+i_2} \right) \right. \\
&\quad \left. + {}_2I_t \cdot f_3^{(M),gr} \cdot \left(\frac{1}{1+gr} - \frac{1}{1+i_3} \right) + \dots + {}_{t-1}I_t \cdot f_t^{(M),gr} \cdot \frac{1}{1+gr} \right]
\end{aligned}$$

Note that:

- for products with the interest rate guarantee, $i_t \geq gr$ for any t and therefore $\left(\frac{1}{1+gr} - \frac{1}{1+i_t} \right) \geq 0$;
- for unit-linked products with no interest rate guarantee this boundaries do not exist.

13) Let us now replace $f_t^{(m),gr}$ with the average value $\overline{f_t^{gr}}$.

It is, we consider that all the policies are the “average” ones. Rewrite the formula for $F_t^{paid,(j)}$ under this (at this moment quite strange) assumption and denote it $\widehat{F}_t^{paid,(j)}$.

It is:

$$\begin{aligned} \widehat{F}_t^{paid,(j)} = & \\ = & paid_t^{(1)} \cdot \left[I_t \cdot \overline{f_1^{gr}} \cdot \left(\frac{1}{1+gr} - \frac{1}{1+i_1} \right) + {}_1I_t \cdot \overline{f_2^{gr}} \cdot \left(\frac{1}{1+gr} - \frac{1}{1+i_2} \right) \right. \\ & \left. + {}_2I_t \cdot \overline{f_3^{gr}} \cdot \left(\frac{1}{1+gr} - \frac{1}{1+i_3} \right) + \dots + {}_{t-1}I_t \cdot \overline{f_t^{gr}} \cdot \frac{1}{1+gr} \right] + \\ + & paid_t^{(2)} \cdot \left[I_t \cdot \overline{f_1^{gr}} \cdot \left(\frac{1}{1+gr} - \frac{1}{1+i_1} \right) + {}_1I_t \cdot \overline{f_2^{gr}} \cdot \left(\frac{1}{1+gr} - \frac{1}{1+i_2} \right) \right. \\ & \left. + {}_2I_t \cdot \overline{f_3^{gr}} \cdot \left(\frac{1}{1+gr} - \frac{1}{1+i_3} \right) + \dots + {}_{t-1}I_t \cdot \overline{f_t^{gr}} \cdot \frac{1}{1+gr} \right] + \dots + \\ + & paid_t^{(M)} \cdot \left[I_t \cdot \overline{f_1^{gr}} \cdot \left(\frac{1}{1+gr} - \frac{1}{1+i_1} \right) + {}_1I_t \cdot \overline{f_2^{gr}} \cdot \left(\frac{1}{1+gr} - \frac{1}{1+i_2} \right) \right. \\ & \left. + {}_2I_t \cdot \overline{f_3^{gr}} \cdot \left(\frac{1}{1+gr} - \frac{1}{1+i_3} \right) + \dots + {}_{t-1}I_t \cdot \overline{f_t^{gr}} \cdot \frac{1}{1+gr} \right] \end{aligned}$$

Therefore, we obtain the final form:

$$\widehat{F}_t^{paid,(j)} = paid_t \cdot \left[I_t \cdot \frac{F_1^{ptf,gr}}{l_1^{ptf}} \cdot \left(\frac{1}{1+gr} - \frac{1}{1+i_1} \right) + {}_1I_t \cdot \frac{F_2^{ptf,gr}}{l_2^{ptf}} \cdot \left(\frac{1}{1+gr} - \frac{1}{1+i_2} \right) + {}_2I_t \cdot \frac{F_3^{ptf,gr}}{l_3^{ptf}} \cdot \left(\frac{1}{1+gr} - \frac{1}{1+i_3} \right) + \dots + {}_{t-1}I_t \cdot \frac{F_t^{ptf,gr}}{l_t} \cdot \frac{1}{1+gr} \right] \quad (9)$$

where

$$paid_t = \sum_m (\text{all policies}) paid_t^{(m)}.$$

Note that all the variables in the formula are known at the time t .

14) Note that (and we are aware of that) the $\widehat{F}_t^{paid,(j)}$ given by the formula (9) does not express the fund value to be paid at the time t for the whole portfolio considered and the scenario j exactly. It is even very likely that it does not fit to the real fund value to be paid at the time t at all.

However, it does not matter much.

Remember what were our conditions that should be fulfilled for the variable $V_t^{(j)}$ – the Condition1 and Condition2 on the page 12.

It seems (based on real case studies) that $\widehat{F}_t^{paid,(j)}$ estimation meets both conditions very well.

15) Alternatively, and more practically with the same result, we can derive the $F_t^{\widehat{paid,(j)}}$ formula following this process:

$$F_t^{\widehat{paid,(j)}} = paid_t \cdot \overline{f_t^{(j)}}$$

$$F_0^{\widehat{paid,(j)}} = 0, \overline{f_0^{(j)}} = \frac{F_0^{ptf,(j)}}{l_0^{ptf}}$$

$$\overline{f_t^{(j)}} = \left(\overline{f_{t-1}^{(j)}} + \frac{\overline{f_t^{gr}}}{1 + gr} - \overline{f_{t-1}^{gr}} \right) \cdot (1 + i_t) = \left(\overline{f_{t-1}^{(j)}} + \frac{F_t^{gr}}{l_t^{ptf} \cdot (1 + gr)} - \frac{F_{t-1}^{gr}}{l_{t-1}^{ptf}} \right) \cdot (1 + i_t)$$

In summary,

$F_0^{\widehat{paid,(j)}} = 0, \overline{f_0^{(j)}} = \frac{F_0^{ptf,(j)}}{l_0^{ptf}}$ $F_t^{\widehat{paid,(j)}} = paid_t \cdot \overline{f_t^{(j)}} = paid_t \cdot \left(\overline{f_{t-1}^{(j)}} + \frac{F_t^{gr}}{l_t \cdot (1 + gr)} - \frac{F_{t-1}^{gr}}{l_{t-1}} \right) \cdot (1 + i_t) \text{ for } t > 0$	(10)
--	------

Where again, all the variables in the formula (10) are known at the time t .

Practical process:

The real process in practice is the following:

- 1) Determine the fictive grid scenarios i^{gk} , $k = 1, 2, \dots, z$ following the approach as described in the point 6) on the page 12.
Typically, $5 \leq z \leq 20$, usually $z = 10$ showed to us to be the most effective for most of the products and still with a very good fit.
- 2) Calculate $CF_t^{ptf,gk}$ for every grid scenario ($k = 1, 2, \dots, z$) within the per-policy model.
- 3) Calculate $F_t^{\widehat{paid,gk}}$ for every grid scenarios ($k = 1, 2, \dots, z$) and every time t using the formula (9) or (10) – out of the per-policy model.
- 4) Then we have prepared everything for the fast estimation of $CF_t^{ptf,(j)}$ for any interest rates scenario j following this procedure:

FOR $j = 1$ to s	FOR $t = 1$ to n	<ol style="list-style-type: none"> 1) Calculate $F_t^{\widehat{paid,(j)}}$ following the formula (9) or (10) 2) Find the closest values of $F_t^{\widehat{paid,gk}}$ from the grid scenarios; It is, find k where $F_t^{\widehat{paid,gk}} \leq F_t^{\widehat{paid,(j)}} \leq F_t^{\widehat{paid,g_{k+1}}}$. 3) Calculate ratio $p = \frac{F_t^{\widehat{paid,g_{k+1}}} - F_t^{\widehat{paid,(j)}}}{F_t^{\widehat{paid,g_{k+1}}} - F_t^{\widehat{paid,gk}}}$ 	(11)
--------------------	--------------------	--	------

4) Calculate the $CF_t^{ptf,(j)}$ estimation as:

$$CF_t^{ptf,(j)} = p \cdot CF_t^{ptf,g_k} + (1 - p) \cdot CF_t^{ptf,g_{k+1}}.$$

Result of this cycle is the projection of $CF_t^{ptf,(j)}$ from $t = 1$ to the final future ($t=n$) for one specific interest rates scenario j .

Results of this cycle is the $CF_t^{ptf,(j)}$ projection for s scenarios.

Practical note - extrapolation

Note that up to now we have assumed that the set of interest rates scenarios is known from the beginning. Based on them, we create the grid scenarios (including the minimum (first) and maximum (last) scenarios).

In practice, this is usually not the case.

However, a possible solution is:

a) The range of the possible future interest rates could usually be (at least roughly) estimated.

For example based on the last periods' rates or based on some theoretical interest rates scenario generator calibrated to the current market situation, etc.

b) Even, if for some time t and some scenario j the $F_t^{\widehat{paid,(j)}}$ is out of interval $(F_t^{paid,g_1}, F_t^{paid,g_z})$ from the grid scenarios, the proportion p could be calculated by the same (or similar formula) as (11). In this case, it represents not the interpolation but extrapolation and $abs(p) > 1$.

From practical point of view, we recommend not to use just the last two values of F_t^{paid,g_k} from the grid scenarios but more (usually three is fine) and use a regression for the extrapolation. The proxy results then fit with the per-policy model much better.

Results:

Cash flow fit

The quality of the cash flow fit, i.e. the difference between the $CF_t^{ptf,(j)}$ from the approximation and the per-policy model showed similar or even better results comparing to the Analytical approach. It is, that in our case study, the differences on annual cash flows were in vast majority times and scenarios far below 1% each year, even if the tested complex products on extreme interest rates scenarios (from -40% to + 60%).

Present value of the cash flows fit

PVCF differences in each of the individual scenarios were mostly up to 0.2% ($\pm 0.2\%$) and the average of the PVCFs (BEL) from the approximation were very much close to the value from the per-policy model (usually $\leq \pm 0.05\%$). These results we consider as excellent.

Correlations

As mentioned above, we expected that the $F_t^{\widehat{paid},(j)}$ will be highly correlated with the $CF_t^{ptf,(j)}$. The experimental results confirmed that – in vast majority cases the correlation was higher than 95% (very often higher than 99%).

Acceleration results

The calculation runtime (in our case study) decreased considerably:

- for 100 scenarios the Interpolation technique accelerated the runtime 10 times;
- for 1 000 scenarios, it was 60 times;
- for 10 000 scenarios, 130 times.

Conclusion to the Interpolation approach

We see the Interpolation approach as very powerful and effective.

It has the following features:

- replicates the per-policy model cash flows and their present values exceptionally well;
- runtimes are significantly shorter;
- the principles of the Interpolation approach are clear and simple, therefore, its implementation is fast; This is a significant benefit comparing to Analytical approach.
- the approach is “universal” in a sense that it does not depend much on the specific product (and its complexity). Therefore, the analyst does not need to know all the product details (what on the opposite is vital when implementing the Analytical approach).

Conclusion

We introduced two techniques that might be used for estimating the cash flows projection for a given interest rates scenario with a very good fit to the per-policy model results and with significantly shorter runtimes – Analytical and Interpolation approach.

We believe that both techniques are powerful and effective.

Especially the Interpolation approach shows excellent results in all the criteria we focused on:

- excellent fit;
- much shorter runtime;
- the approach is very easy and therefore can be understood and implemented within the real practice in short period of time.