The Rational Policyholder

Prepared by: Ákos Gröller
Specialisation in: Actuarial Studies
2012
# Contents

1 Abstract

2 Introduction

2.1 Why we care for policyholder rationality in valuation .............................................................7
2.2 Why is it hard to estimate market dependency in policyholder behaviour? .........................8
2.3 Dynamic policyholder behaviour as a risk ...............................................................................10
2.4 Rational? How come they hold policies? .................................................................................11

3 Revealed preferences

3.1 A call for knowing the client’s taste ........................................................................................13
3.2 Readings in microeconomics ...................................................................................................13
3.3 Calibration in our case .............................................................................................................14

4 Valuation of financial instruments with optionality

4.1 The Black–Scholes model ........................................................................................................16
4.2 Implied volatility .....................................................................................................................17
4.3 Risk neutral valuation ..............................................................................................................17
4.4 Risks for writing options and guarantees .................................................................................20

5 Policies with embedded options

5.1 A simple setup to exercise rational policyholder behaviour ....................................................23
   5.1.1 Guaranteed product based on closed end fund...............................................................23
   5.1.2 Guarantee provided on an open end fund ......................................................................24
   5.1.3 Decisions of the policyholder – reduction to clean strategies .......................................25
   5.1.4 Recycling the BSM formula .............................................................................................26
5.2 Client implied parameters ........................................................................................................27
   5.2.1 Volatility ...........................................................................................................................28
   5.2.2 Interest rate ......................................................................................................................29
   5.2.3 Dividend rate ....................................................................................................................29
   5.2.4 Parallel shift of dividend and interest rate .....................................................................29
   5.2.5 Suggestion .........................................................................................................................29
5.3 Decomposition of client behaviour for modelling ...................................................................30
   5.3.1 The goal: contingency projection ....................................................................................30
   5.3.2 Enumeration of alternatives ............................................................................................33
   5.3.3 Valuation of outcomes .....................................................................................................33
   5.3.4 Comparison of two alternative values .............................................................................34
   5.3.5 Decision and action based on the comparison .................................................................36
   5.3.6 Multiple choice ................................................................................................................38
5.4 Summary of the proposed approach and areas of use ..............................................................39

6 Quantitative example

6.1 Framework for the calculations in the thesis .........................................................................42
6.2 Calibration of implied parameters ............................................................................................42
6.3 Value impact of dynamic behaviour .......................................................................................45
6.4 Sensitivity to technical assumptions .......................................................................................48
6.5 Concluding the calculations ....................................................................................................49
Summary and conclusions

7.1 Achievements..........................................................................................................................50
7.2 Areas for further exploration...................................................................................................50
7.3 Limitations...............................................................................................................................50

A Additional calculations

B Bibliography
Index of tables and figures

Table 6–1: Typical interim and output figures of persistency modelling .............................................. 31
Table 6–2: Independent transition probabilities for persistency modelling ....................................... 31
Figure 6–3: Example for parameterizing the decision function ......................................................... 38
Table 7–1: Introducing the sample policies .......................................................................................... 41
Figure 7–2: Client implied valuation adjustment parameters, base case .......................................... 43
Figure 7–3: Sensitivity chart for implied parameters .......................................................................... 43
Figure 7–4: Client implied parameters for higher σ, 5-y closed fund with Asian call ...................... 45
Figure 7–5: Parameterizing the lapse decision formula – base case on 10 years open fund policy. 45
Figure 7–6: Impact of dynamic behaviour to SP guaranteed UL product .......................................... 47
Table 7–7: Summary of dynamic lapse intensity impact on open end fund product ......................... 48
Figure 7–8: Selection of the comparison function is quite neutral .................................................... 49
1 Abstract

More and more insurance products, especially unit-linked ones, include as a major part of the service some sort of investment guarantee, which inherently provides financial options to the policyholders. Meanwhile competition pushes for narrower and narrower spreads. Hence the prediction of client behaviour in pricing and valuation gains importance. A particular aspect of this, which is regarded as an important stress test – or even a tollgate – is the impact of the policyholder behaving ‘rational’, whatever that may mean.

The thesis discusses this ‘whatever’. As a general principle, we will call for distinction between standard market consistent techniques and valuation of choices from client’s perspective. The reasoning is very brief: an insurance product sold as being market consistently profitable cannot be bought on the same principle. We propose applying the method of revealed preferences, and on a few examples illustrate possible approaches to quantify it.

Before presenting the idea, we briefly introduce and comment on the following concepts that provide the starting points for it, namely:

- revealed preferences;
- the Black–Sholes–Merton formula for option pricing and implied volatilities;
- using risk neutral distributions for market consistent valuation.

The particular solution outlined here is applicable only for products where the utility of the – stochastic – future cash-flow streams can be expressed as a function of some expected monetary value, and a financial option or guarantee is an important factor for the client in assessing this monetary value. Nevertheless I believe the principle, seeking simple and ‘explainable’ mechanics for rationality that is consistent with the – observed or assumed – behaviour of the selected clientele, can be used more extensively than a narrow product range.
2 Introduction

Financial options and guarantees, embedded into or attached to saving policies, are gaining popularity. They present a sound combination of guarantees known appreciated in traditional products and unit linked structures with freedom, flexibility and transparency. This implies that the client can make decisions which may considerably impact the value of the benefits paid and earnings received by the insurance company, decisions that depend on financial market conditions. The question arises whether and how this should be accounted for in actuarial calculations.

2.1 Why we care for policyholder rationality in valuation

Besides the George Mallory argument (“Because it’s there”), we also have formal reasons to do so. The CFO Forum MCEV Principles (CFO Forum, 2009; the acronym stand for Market Consistent Embedded Value) sets out the following criteria to valuation:

“Principle 7: Allowance must be made in the MCEV for the potential impact on future shareholder cash flows of all financial options and guarantees within the in-force covered business. The allowance for the time value of financial options and guarantees must be based on stochastic techniques using methods and assumptions consistent with the underlying embedded value. All projected cash flows should be valued using economic assumptions such that they are valued in line with the price of similar cash flows that are traded in the capital markets.

G7.3 Dynamic policyholder behaviour should, where material, be in the allowance for the time value of financial options and guarantees.

G7.4 The techniques used to calculate the allowance for the time value of financial options and guarantees should incorporate an allowance for stochastic variation in future economic conditions consistent with Principle 15. The economic projection assumptions should be consistent with how the capital markets would value such cash flows and Principles 12, 13 and 14.”

Blending these, calculation for an MCEV disclosure compliant to the principles needs to estimate dynamic decisions of clients within a stochastic framework, so must incorporate rules in the cash-flow projection that drive policyholder decisions meaningfully across a wide range of possible scenarios. Of course, if the impact is material. To judge whether it is material we need to measure the exposure to this risk, more precisely, how exposure to market risk may be affected by considering or disregarding dynamic behaviour of the customers. Similar expectations are set in the QIS5 Technical Specifications (2010), but – of course – in a more technical formulation:

“TP.2.83. Regarding contractual options, the assumptions on policyholder behaviour should be appropriately founded in statistical and empirical evidence, to the extent that it is deemed representative of the future expected behaviour. However, when assessing the experience of policyholders’ behaviour appropriate attention based on expert judgements should be given to the fact that when an option is out

1 We do not state that all unit linked products really bear these three attributes. Yet most of them are marketed with this image and bought for this image. And some actually deliver these features.

2 Note that “In April 2011, the CFO Forum withdrew the intention that the Market Consistent Embedded Value (‘MCEV’) Principles® are the only recognised format of embedded value reporting from 31 December 2011”, see http://www.cfoforum.nl/embedded_value.html. Yet the discussions behind this were not on the main principles, they are commonly understood as good, even necessary, professional practice.

3 Quote from Section V.2. Technical Provisions (hence the acronym TP). Loosely, this is the fair value of liabilities.
of or barely in the money, the behaviour of policyholders should not be considered to be a reliable
indication of likely policyholders' behaviour when the options are heavily in-the-money.

TP.2.84. Appropriate consideration should also be given to an increasing future awareness of policy
options as well as policyholders' possible reactions to a changed financial position of an undertaking. In
general, policyholders' behaviour should not be assumed to be independent of financial markets, a firm's
treatment of customers or publicly available information unless proper evidence to support the
assumption can be observed.

Article 21 TP8 of the SII L2IM (in reference to Art. 78 of 2009/138/EC) also prescribes the
presence of dynamics in the projected incidence rates. Quoting only by concept:

Assumption on policyholder behaviour (wrt. exercising options) shall be based on analysis, covering
‘ITM-ness’, past economic conditions, past management actions and any other relevant factor.
The likelihood (projected execution rate) should not be assumed to be independent of the factors above
unless there is empirical evidence to support such an assumption.

So the modelling decision process is the following:

• Optionally, estimate whether the covered business is materially exposed to risk from dynamic
  behaviour.

• Either on general principle, or because of justified materiality, should dynamic behaviour be
  included in the model?

• If yes, how?

If we can answer the last question somehow, we can obviously use the same approach to assess the
risk. In turn, if we can analyse the risk, we may not need to care for the third. A good answer the
latter is hard to find as we will argue below, whilst an approximation can be enough to analyse the
first. All this suggests to focus on the risk – from after the coming subsection.

2.2 Why is it hard to estimate market dependency in policyholder
behaviour?

There are many, maybe one by one not so striking issues that add up to quite a difficulty.

Timing and relevance

There do exist publications, numerous of them, that present findings on how past experience on
lapses\(^4\) relate to drivers from portfolio and product attributes, company characteristics and
economic figures. Cherry-picking from recent ones, Kiesenbauer (2012) analyses aggregate data
and published their analysis on large portfolios from the German insurance industry. They find
significant relations, but by nature their results are only suitable for aggregate prediction and the
drivers themselves require assumption setting. The article reviews the reference points in the
literature so I skip quoting one sentence from each, though many have even more than a sentence
worth quoting. These tend to address aggregate moves. The magnitude of total response to expect
on economic moves may be assessed on the basis of these reports, often per high level product
categories. But this is less dynamic behaviour, rather reasonable correlation of insurance business
with the overall evolution of the economy.

Two common hypotheses are often considered that aim to explain beyond registering. These can
be used as starting points for the mechanics of a projection model, just the results from different
authors and on different data sets are, say, unambiguous. The emergency fund hypotheses expects
withdrawal of funds, primarily as surrenders, from saving products when adverse economic
environment, usually benchmarked by unemployment rate, necessitates it. Kiesenbauer (2012)

\(^4\) I use the common ‘non-naming’ convention – formally we all distinguish between lapse and surrenders as two
different ways to prematurely terminate an insurance policy. However, as long as we do not need to focus on the
difference between them, we loosely denote the union of these with either word to decrease redundancy in the text.
adds GDP rates. Kim (2006) and Park (2009) extend it with many indicators in actually seeking to predict market total lapse rates. Knoller et al. (2011) also accept personal expenditure, esp. on housing and health care, to qualify a lapse event as emergency fund access and implicitly identify it within their age dependent regression factors.

The interest rate hypothesis plainly assumes that high market interest rates increase the opportunity cost of sustaining older saving plans and trigger lapses for seeking better reinvestment opportunities. However this 'plain' assumption, as reported by Kiesenbauer (2012), is specified in many different ways and with varying success. His study also limits the relevance of it to only a few of the product types reviewed.

Shortly, all these paperwork aid more in getting an insight to what happened in the past, but tangible projection tools do not emerge easily. Furthermore, on this topic increasing customer awareness should be assumed and past data can hardly reveal how that works. The fall of Equitable Life Assurance Society and the eight digit loss of White Mountain Re were driven by not considering risk in policyholder behaviour that had not been not observed before (see Knoller et al., 2011 and references therein).

In aiming to build models calculating MCEV or some similar figure (could be Solvency II Technical Provision and/or the attached risk capital) we should also seek drivers that can be sampled from a market consistent scenario generator, mainly drivers that are traded themselves or link to traded prices. Otherwise the risk part of the link (e.g. augmented equity exposure) can not be consistently channelled into the reporting processes, only the impact on the value. An inconsistent translation of results into parameters of, e.g., a risk neutral valuation framework can easily do more harm than benefit.

Shortly, aggregate analysis of past lapse rates linked to unemployment ratio of the ‘80s is more interesting than useful in predicting dynamic policyholder behaviour in a projection model that has stochastics for only interest and equity.

**Data granularity**

I tend to formulate one of the key problems as policyholders just acting too slowly. Insurance, especially those saving products where the insurer takes the large risk – on investment, longevity, persistency and morbidity, mortality – in the expectation of proportionate margins, is long term business. Especially for the type the products where long term experience data on large volumes could be available, things just don’t happen every second. A traditional endowment policy, where all of long term risk, large amounts and embedded options are present to conduct a nice analysis, merely changes over a policy year. So the series of observations we might have need to be counted in years, and then it is not so long – the time series for market risk indicators to that we would map the outcome are much more frequent. Weekly data is usually considered as infrequent enough to have new information from each item.

Then the individual policies, many as they may be, often do not differ enough to be analyzed separately. If out of a hundred thousand policies we observe three thousand to lapse, we have a very accurate and reliable data point 3%, but that’s just one observation. And as policies of a company are usually to the same or nearly the same economic indicators, splitting the data may also happen not to increase the amount of real information in our hands.

**Data availability**

Recent, detailed and relevant data could come from competitors who launched similar products lately, but somehow this is not common practice to ask for and deliver these. Analysis results may, however, be accessible with a few years delay – as do Knoller et al. (2011) share their findings on Variable Annuity lapses observed at a life insurance company in Japan.
Eling–Kiesenbauer (2011) claim to have worked with maybe the largest portfolio ever used for a published study as they covered 2.5 contracts through 8.9 million policy years from a German insurer. This also indicates that detailed analysis is usually reserved for internal analysis.

Assumption setting

Regardless of difficulties, insurance companies do set up stochastic decrement assumptions for projection and valuation of their business. These may be based on simple heuristics like nominal in-the-moneyness of the option, the company’s internal experience of even fairly complex algorithms. (Knoller et al. (2011), Kling et al. (2011)). A balance between vague and demanding (calculation wise, in data need or in parameterisation) I could hardly find in the literature. An elegant rule I heard of is to make the surrender assumption contingent on the Delta of the option(al benefit) provided by the contract.

The approach that we will suggests is the complete opposite of data intensive – the parameter estimate actually provides information on the product instead of consuming a lot. Parameters to estimate are not many and some plan language meaning can be attributed to them.

2.3 Dynamic policyholder behaviour as a risk

To assess the excess exposure to market risk due to client decisions, the company has to take into account the range of choices offered to the client, project some reasonable pattern of behavior and test how sensitive the final result is to the selection of this modeled behavior. Notwithstanding that even a reasonable estimation of the strategy is a challenge, it’s even tougher to add well established confidence domain and spot the element therein representing the worst case to the company.

A fairly acceptable substitute for a precise worst case is to test the policy assuming the client behaves rational. We may expect that the most valuable pick of the client will be the most costly for the company. This would clearly be the case if

- the valuation principles of the client and the company coincide and
- there are no external sources of gains or losses, so the policy is a zero-sum game between the two entities.\(^5\)

The latter roughly applies if all financial instruments backing up the policy are sufficiently liquid and marked to market, and the liabilities are appropriately covered by reserves.\(^6\) Meanwhile the somewhat loose formulation is, at a closer look, subject to a consistent understanding of what is a gain or a loss, so practically the first item. We will raise concerns on the first bullet right away; however we still accept the rationality of the client as a meaningful stress test on an insurance contract.

As an additional area of use, in the examples we will derive the best estimate assumptions via ‘weakened rationality’, i.e. starting from the rationality stress test setup but applying smoother response thresholds.

---

\(^5\) At least in forward-looking valuation. Gains and losses already realised in the past may have been offset by hedging.

\(^6\) Accounting may prohibit using negative reserves in the books. The main point rather is that the insurer recognizes the respective liabilities in its own valuation. E.g. negative ‘reserve’ incurs a loss when the policy is terminated as it typically cannot be transferred to the beneficiary as a decrease of service.
2.4 Rational? How come they hold policies?

Either if we argue for the rationality test as above, or if we attempt to quantify what the rational decision in a modeled situation is, we get to the question of how policyholders valuate financial instruments to compare their choices before rationally selecting the most valuable one.

A plausible technical approach would be to use the market consistent valuation framework which is right at hand anyway. Let us now assume the product delivers positive MCEV(0) to the company, which means the sum of the market values of premiums\(^7\) covers for expenses, commissions, capital charge, value creation and finally the benefits. On day 2 the client usually can decide on holding the policy or surrendering. Assuming little has changed after posting the first or single premium, there are at least the two choices of taking the surrender value or feeding on the company with the earnings sufficient for all the above and getting the respective benefits, so let us compare these two. The surrender value is the premium paid in, less initial loadings, less surrender charge. The benefit less outstanding premiums can be expressed as the benefit, plus premium already paid, minus all premiums. As the company has just calculated, all the premiums are equal to all the loading plus the benefits. Simplifying the comparison by the benefits and the premiums already paid, the two sides become:

- surrender: a loss equal to initial loads and surrender charge; versus
- hold (to maturity): pay all the loading for expenses, commission, capital charge and profit of the company.\(^8\)

If the second choice is better, we probably have one of the following.

- The surrenders are so penalized that the product may be very hard to sell.\(^9\)
- The margin on the product is so low that a normal level of surrender penalty is sufficient to withhold from surrendering. The product has very low value to the company and very likely does not provide major financial options (or they are provided by some highly rated external party on large scale and low price, and just packaged into the insurance product) or else even the risk capital would be hard to finance.
- The client has a fairly different opinion on how she/he will take the options provided by the product, and so even with the same financial methodology the value of benefits is higher and/or the cost of premiums and fees is lower than assessed by the company.\(^10\) The product is mispriced by undervaluation of policyholder options.

None of these is really appealing. But even if we can explain this, there’s no excuse for buying the policy. If the product is correctly priced and the company makes money on that,\(^11\) the client, using the same valuation framework, doesn’t. We can either drop rationality, or find a more sophisticated formulation of it which does not render all clients irrational.

However, first we take a detour and look around how client side valuation is addressed by the theory (and philosophy) of revealed preferences, and how (financial) market valuation is

\(^{7}\) Charges and fees are implicitly appearing in the benefits.

\(^{8}\) A more concise set of choices would include paid-up conversion, deferred surrender and deferred paid-up. For clarity we skip these. Similar logic could be applied in a broader context as well.

\(^{9}\) There may be an initial period with no surrender value and such products are marketable. Then instead of day 2, consider the first time with ‘normal’ surrender penalty.

\(^{10}\) Adverse selection for non-financial risks should not be considered here, though has similar implications. Let us assume a savings product with low enough risk coverage that this is not an issue. Actually the value of standard size non-financial risk covers is in fact typically assessed higher by clients than they are costly to the provider.

\(^{11}\) We can even disregard all but the positive MCEV(0) and financing the MVM on the economic capital.
addressed by risk neutral distributions, Black–Scholes–Merton (BSM) option pricing formula and implied volatilities. Then we will try to get all this together and blend something consumable.
3 Revealed preferences

The mechanics of rational customer behaviour within our scope will be described and illustrated in later chapters. Here we present the idea – the definition of rationality will be given by matching to the revealed preferences of the client. Let’s see how this comes into the picture.

3.1 A call for knowing the client’s taste

The core of being rational is taking the most valuable choice. To fill this with meaning, the decision maker needs

- information with regards to the set of choices and their consequences and
- a set of preferences to valuate (or at least compare) them.

In an actuarial projection we can’t really model any policyholder information that is not a part of the projection itself. So if we aim to model the behavior, we should condition it on inputs that are also available to us. On the other hand, parameters – especially market prices – that the company uses in the projection and are closely related to the consequences of the choices (surrender value, share prices, interest rates) should not be excluded from the modeling of the decisions, as otherwise we would limit the scope of rationality and underestimate its possible impact.

What we have the most limited knowledge about is the preference set of the policyholder. We have just concluded that it should differ from what an insurance company considers market consistent. Just sticking to the words, that market – large scale, low transaction cost, liquid financial market where short and long positions are equally available, instruments can be traded both at fractional and large numbers – is not available to the majority of clients, besides that, dispersing just the effort of trading requires a larger portfolio than one insurance policy. What we may aim at is a valuation somehow consistent with the market of financial instruments truly available to the clientele. Unfortunately we can’t observe prices and build e.g. a risk neutral framework around this market, as, the market being fairly imperfect, we could not use the no arbitrage argumentation anyway. Instead we may – though it must be admitted we have no heavy machinery theory to support that – try to adjust the calibration of our MC valuation to reasonably represent some client behavior. The benefit of this may be that we don’t have to invent something totally new on a few bits of information. What we can be sure in – and what definitely sets the policyholder apart from the company – is buying the policy. Calibrating on this decision is actually taking it as a revealed preference, a common topic in quantitative microeconomics.

3.2 Readings in microeconomics

As observed by Varian (2006), the literature on, or related to, the concept of revealed preference counts rather thousands than hundreds of publications. Very likely, all of them include the basics we present here, but currently we keep on digesting this article, just let loose of some of the mathematics rigor in the presentation.

The definitions are the following:

12 It is of course the model builder’s responsibility to include all material information. Here we will focus on investment parameters where any key factor is either traded or a laid down condition of the policy. Well, we could do scenario testing on unknown, but – in some model we trust in – important parameters that affect the policyholder. This possibility will not be evaluated here. Let us just point out how subjective such approach would necessarily be.
13 No fractional positions on several assets, no short positions on lots of the assets, large transaction delays etc.
Revealed preference

A bundle (aka. bucket) of goods $x^t$, taken by a customer at observation time $t$, is *directly revealed preferred* to any bundle $x$ that was available at the same time for not more than the price of $x^t$, and was not taken. We denote this as $x^t \preceq_D x$.

A bundle $x'$ is *revealed preferred* to $x$, $x' \preceq_R x$, if we can get from $x'$ to $x$ through a series of bundles where each is directly revealed preferred over the next one.\(^{14}\)

**Weak axiom of revealed preference**

As set of observed choices fulfils the WARP if these don’t directly reveal contradicting preferences. Namely, if $x \preceq_D y$ and $y \preceq_D x$ are mutually exclusive.\(^{15}\)

**Strong axiom of revealed preference**

A set of observed choices fulfils the SARP if these don’t reveal contradicting preferences: $x \preceq_R y$ and $y \preceq_R x$ are mutually exclusive. Even shorter, $x \preceq_R x$ never holds.

The power of the notion of revealed preference is fed by several sources:

- SARP must hold if choices are made rationally on the basis of a fixed utility function with positive marginal utility.
- The fundamental theorem of revealed preference states that if SARP holds (and choices are on the budget margin\(^{16}\)), there in fact exists some fixed utility function that would yield the observed choices as the rational ones. For finite observation sets, an example utility function is actually constructed by Afriat (1967).
- Various numerical methods were invented to efficiently check the validity of SARP on large samples. When there are only two goods, WARP is equivalent to SARP and is very easy to check.\(^{17}\)
- Several empirical studies on aggregate demand found SARP to hold in their examples (see Varian (2006) for references). Even though aggregate demand is not a consequence of a single, fixed utility function maximized over the budget set of a single decision maker, it can in these cases be approached as if it was. Shortly, aggregate demand sometimes seems rational.

We must remember that a utility functions based on a revealed preference, though usually linked to observations, are imaginary technical constructs. Just as is any utility function – these all, by nature, only exist ‘indirectly’ via decisions. But this does not cancel their predictive power in various cases. Many criticism on consumer theory claims that it enforces rationality\(^{18}\) on phenomena that are not as simple to be rational in reality. I think that, to the contrary, consumer theory used carefully can enforce us to extend our understanding of rationality, and resist the temptation to give it up at the first hindrance.

### 3.3 Calibration in our case

Clients, very roughly, may consider a particular product a bargain, a fair trade or a bad deal. Counting on the first class we would tend to assume too high persistency as benefits are rated high

---

\(^{14}\) So ‘revealed preferred’ is the transitive closure of ‘directly revealed preferred’.

\(^{15}\) We could go a bit further and introduce the ‘strict’ version of the direct relation where the price of the omitted bucket needs to be strictly lower and define ‘strictly revealed preferred’ through sequences where at least in one step the strict direct relation holds, and get more convenient versions of WARP and SARP – for now, such technical explorations are left to the curious reader.

\(^{16}\) or we add the residual budget as an additional good and SARP still holds for this extended choice set

\(^{17}\) The two-dimensional case is more interesting than it first seems, just put consumption of one good against all the budget not spent on this good – we can quickly traverse to Slutsky’s works in consumer theory.

\(^{18}\) see Wiki4
against costs. Hopefully those in third class get sufficient information that most of them do not get
to buying a policy. Let us hence postulate that the rational client as we will model her is one to
whom the value of future benefits at issue is equal to the value of future premiums at issue. From
now on we shall focus on how we may build up such valuation rules that are also feasible,
apparently reasonable and as unambiguous as possible.

In this scope definition we have hidden even two things important enough to be stated explicitly:

**Philosophy**

We assume rational maximization of a regularly shaped utility function as long as we don’t
have clear evidence of the contrary, i.e. contradicting choices. We know from
microeconomics that this is technically valid assumption.

A good question whether this assumption makes sense, but a question about impossible to
answer in theory. Pragmatically: if we can apply it for good, it does.

Finally we will calibrate to just a few inputs: choice(s) made by the client with regards
to the particular policy, and market prices that compress information on decisions of
lots of clients to just a few numbers. These observations are too few to have even
theoretical chance for such contradictions. So will just try to fit some ‘nice’ utility
functions to the decisions.

**Applicability**

As we measure client value by twisting and bending financial valuation techniques, we
limit ourselves in terms of products and policyholders: the utility function of a selected
client on a selected product should be well expressible through a monetary value, and the
function that yields the monetary value is scalable and bound to a few additional regularity
constraints. With somewhat loose formulation, the more popular maximization target
\( \mathbb{E}(u(\bullet)) \) needs to be replaced by some \( u(\mathbb{E}^*(\bullet)) \).

For instance, this scalability does not hold if a non-guaranteed investment policy is bought
by client to ensure a sustained saving plan to reach some target amount at the end of the
term, say the tuition fee for the child – additional upwards potential has lower marginal
utility value in such a case, meanwhile some of the utility is attributed to the penalty-driven
motivation for not quitting prematurely.

Products explicitly designed around an investment guarantee, and clients buying them,
seem to be much better candidates for this approach. From now on we stick to such
constructs.

This is what we will put in practice, but only after we’re done with financial mathematics.

---

19 Strict equality is a bit prudent, but is a clear condition to calibrate on. It is probably a more realistic rule that the client
perceives the policy to provide a small gain in benefits over the premiums to be spent, but if this really is a small gain
(‘is’ here means how we assume the hypothetical client feels about the policy… never mind), the impact from
disregarding it in the fitting should also be small to the end results.

Besides, we could have added ‘expected value of …’ – for easier reading we will often skip this word, value of future
cash-flows in a risky environment always is some sort of expected value.

20 It should be: scalable (multiple mapped to multiple), continuous (similar to similar), additive (sum to sum – a very
extensive condition as cash-flows themselves are stochastic processes), positive (positive to positive). These are strong
constraints, e.g. scalability says that saturation effect can not appear in the monetary value, while with additivity,
representation of risk aversion is restricted to generic parameters (this will be clearer after the text on risk neutral
measure, 4.3).
4 Valuation of financial instruments with optionality

In this chapter we look at how to value financial instruments that:

- have embedded options and guarantees, and
- do not have an efficient and liquid direct market.

Our primary interest in the topic is insurance contracts with built-in derivatives.

Note that I will a few times take position in scientific debates concerning models of mathematical finance in general, and the Black–Scholes model in particular. Thus the chapter partly presents a personal view – yet one that is blended with consent for many arguments floating around. These debates and opposing views can be looked up starting from the list of references – their review is out of scope here.

I have combined and reformulated inputs from many sources, but from all of them Baxter–Rennie (1996) should be nominated as the most heavily used one.

4.1 The Black–Scholes model

The basic version of the model considers a European put or call option on a stock or index (called underlying instrument), setting out many strict simplifying presumptions. The main result is a closed formula for the value of these options at any time prior to maturity, and closed formulae on the ‘Greeks’ (first order partial derivative of value of position with respect to time and market parameters, and second order derivative with respect to price of underlying). The basic building blocks for getting there are:

- Stock price assumed to follow geometric Brownian motion (with constant drift and volatility).
- Risk free interest rate is known and constant.
- Time, prices and trading units are infinitely divisible.
- Liquid, efficient and arbitrage-free market exists for the underlying stock and cash, both for short and long positions.
- In this situation, a self-financing portfolio (or rather: investment strategy) of cash and stock can be constructed and continuously maintained so that its payoff at the maturity of the option equals the payoff of the option itself. This is called the self-financing replicating portfolio (or strategy) for the option. The strategy itself is derived as the solution to a partial differential equation (PDE).
- In lack of arbitrage and transaction costs, the option is worth just what the replicating portfolio is worth. The latter is built up of traded assets, its value is thus known – then so is the value of the option.

The formulae themselves will be presented in 5.2 where we will work with them the most directly.

Within this framework, the parameters other than volatility of the underlying are simply observable, and independent from even the existence of stock markets. In particular, interest rate(s) can be taken from prices of fixed income securities or even better, interest rate swaps.

---

21 Later we will deal with some of these in detail. See Wiki1 for the list and commentary.
22 The original article rather constructs a ‘neutral’ self-financing portfolio from constant 1 unit of the option and varying volume of stock and cash. Neutral means its value progresses independently of the movements in stock price. The two are conceptually equivalent.
23 In even more formal wording, the portfolio that corresponds to the current state of the replicating dynamic investment strategy. Always keep in mind that the replicating portfolio itself is a multidimensional stochastic process.
24 OK, future dividend rates, that come up in the most commonly used extension, are not fixed. It’s only a somewhat informal ‘deal’ and common expectation between shareholders and managers of quoted companies that normally keeps dividend ratios relatively flat, especially for indices with large samples.
These markets being fairly close to theoretical efficiency (trading margins are much lower than for stock exchange) and very liquid, for analysis of equity options the interest rate(s) and its (their) possible movements are either external inputs, or preceding stock price movements in the calculation logic.\textsuperscript{25} Hence the BSM formula can be understood as one that translates volatility to option prices.

### 4.2 Implied volatility

Naturally – and of course also in the BSM formula – higher volatility increases the value of a standard put or call option at any time before maturity. Besides monotonicity, BSM also yields a continuous mapping. This gives rise to the idea of turning it around – instead of assuming to know the volatility and telling the prices, rather reading off market prices of put and call options and attributing to them the volatility parameter that generates their actual price in the formula. This use of the formula has the following benefits and consequences.

- It translates option prices to volatility figures. The latter are much easier to compare across time, underlying instruments, currencies.
  - Note that the transition between the price and volatility parameters is fairly similar to how utility values and in nature trading rates were translated to nominal prices as money was 'introduced'. Whatever one thinks about the Black–Scholes model as a valuation tool, no true economist should underestimate its capacity to express market information in a compact and commonly understood manner.
  - The gap between historical and implied volatility (expressed as difference or proportion) may also be used in comparisons, as it is even more 'scale-free'.
- Valuation models for more complex instruments (e.g. Monte Carlo simulation methods) can be calibrated to the observed market prices of less complex, more liquid instruments through the implied volatility.
  - In the loop price (observed) → volatility (implied) → price (modelled), quite some of the bias due to deficiencies and limitations of the Black–Scholes model may happen to net out. Of course, for this the valuation model must align in its methodology with the way the volatilities were implied.\textsuperscript{26}
- Implied volatilities do not describe any expected distribution, they just explain prices.

In a perfect world (perfect in the sense of the Black–Scholes model assumptions), implied volatilities would not depend on the duration, relative strike price or whether a put or call option is taken. The last independency roughly holds because of the put-call parity (there is some noise due to transaction costs and different liquidity), but duration and in-the-moneyness notably influence the implied volatility. The ‘implied volatility surface’ (charting by these two dimensions) tends to be smooth and flat enough so that we don’t throw away the entire idea, but uneven enough to alarm us about the limitations. See Wiki\textsuperscript{6} for illustration and commentary..

### 4.3 Risk neutral valuation

The idea of using implied volatility to mark financial instruments to a model that is marked to market can be further generalised. A short dive may help in assessing the nature and applicability

\textsuperscript{25} Let us skip the rigid generic formalisation here.
\textsuperscript{26} The contrary is a very common error – partly because reliable data on implied volatility is not available for a multitude of alternative models and parameterisations, just consider the dividend rate that has to be estimated. E.g. the 'live' version of the numerical example brought up later in the thesis also suffered from this: the implied volatilities (probably) came from a fixed interest model and the dividend rate estimates used there are unknown, yet the calculation used the implied vols in a joint IR–EQ model. Dividend rates did not matter as they were recapitalised.
of such calculations. We are interested as our quantitative examples happen to rely on risk neutral scenario set. Yet in the particular section we are not interested in the rigorous notations, quotations, definitions and all the details that one get lost in. (There are no other ways how risk neutral measure can be a subsection – though see Wiki5 as minimalistic and still useful treatment. Another concise, clear and not too long article explaining the ins and outs is Giesiger (2010)). We stick to the common names that can be resolved starting from Wiki2 Where precision is necessary to start research on the topic, we make sure to use the common names and notations even if we don’t explain them at all.

The conflict that the concept of a risk neutral (RN) measure resolves is this.

- We really like to use the expected present value operator to combine risky outflows of a financial instrument (or a plan to run a portfolio of them) and get the value in the end.
- We can not do it on the distributions (measures) that we believe the respective random variables (stochastic processes) follow in reality. Well, even worse, we can easily do it but the consequences would be disastrous if we would try to trade on the values we find. Many examples from straight to complex are provided in the introductory chapters of Baxter–Rennie (1996) on why this would fail.27

Because when we really want to, or even more, need to trade financial instruments, it is not our beliefs that drive the prices but supply and demand. Where there is enough liquidity to keep our transaction costs and liquidity risk low, our beliefs still add a tiny little part of the information pool driving the prices, but it will never be as much as two tiny little parts.

Passing this for a second we present the causes of addiction to expected value calculus.

- Prices properly calculated as expected values on a known distribution can never contradict. If we calculate the value of assets A and B via the expected value operator, the value of asset A+B will be the same of the values.
- It allows for convenient simulation.
  - Take a financial instrument X that we want to price now, on the condition that we can evaluate it on all possible outcomes that we want to allow for in our modelled world.
  - Assume we know a (the) probability measure Q that delivers the market prices of future cash flows, and assume that X has finite expected value over the set. Add some further regularity constraints that we don’t even attempt to explain in a thesis.
  - Because of the (...) conditions, expected value of X will be a continuous function of Q (!) – if we have a series of distributions Q(i) that converge in probability to Q, the expected value sequence converges in distribution to the expected value by $E_Q(X) \xrightarrow{d} E_Q(X)$. The right hand side being a constant this implies convergence in probability.
  - Convergence in probability to Q is something that empirical distributions will deliver to us and even the asymptotics are known. If we draw samples s of a statistical process that simulates Q, calculate X for each case and take the average, what we get is exactly $\frac{1}{n}X(s_i) = E_Q(X)$ and so it converges to $E_Q(X)$.
- The simulations can be back tested by checking that if we calculate the value of liquid traded assets, the process actually delivers a statistically accurate estimate.

27 Including the somewhat suprisingly thorough presentation of the authors' cultural shock on seeing that expected value as such was not yet banned from high society, together with the deservedly infamous law of large numbers. Then few dozen pages later they are deep in normal distributions and expected values. Whatever, the examples are well explained and the book is more than worth reading.
Under certain conditions (which in particular apply to the Black-Scholes model), the way we do the back testing can be turned abstract to define the risk measure $Q$. Namely, assume we have a set of reference assets and prices for them that fulfill the following.

- The prices are consistent: there is no arbitrage possible (neither immediate or as a risk free strategy).
- They are complete: the evolution of the prices of these assets in the future provides sufficient information to unambiguously know the payout of all $X$ we care for (in other words, the risks priced by the observed market span the risks we are looking at).
- Smoothness, boundedness and various technical items.

Then there exist a unique stochastic measure $Q$ which can price exactly the assets spanned by our reference market (does not measure anything else than what follows from our inputs), and in fact prices them properly. In addition, for discrete models and for many known continuous stochastic processes, there exist actual constructions for deriving $Q$. When it exists, $Q$ is called the risk neutral distribution. The name itself reflects that the value of a stochastic cash flow is really priced by $Q$ by just adding up the probability weighted outcomes. A one in a thousand chance (chance measured by $Q$) to loose a thousand euros is not any worse than a single euro fixed payout, and a (by $Q$) one in a million chance to win a million euros is not worth one euro and a cent. This is not how human value and risk assessment team up, hence the distinctive name.

There is a formal name for $Q$ not knowing more things that can happen than do the prices it is constructed from. The supply of $Q$ is the same as that of the realistic measure $P$ we start from on the respective stochastic variables, the one that needs to be nicely dressed for $Q$ to exist. $Q$ can not be zero on any event where $P$ is not and vice versa. With all the regularity conditions around, if $P$ is an Itô-process then so is $Q$.

**Some words of warning**

Many authors criticise the continuous dynamic replication to be an improper pricing technique for its practical impossibility, and restrictive and unrealistic conditions of use. It may appear as if using risk neutral valuation did cut the Gordian Knot. However, in the true essence of the concept there is absolutely no difference between the two. The risk neutral expected value process is the value of a replicating strategy, and solving the differential equations that define the replicating strategy is in fact the construction of a risk neutral measure.

There are yet differences in practical use and appearance.

- The replicating strategy needs to be solved for any instrument – any payout cashflow – we price. The risk neutral measure, once derived, can ‘deliver’ the price of any compatible instrument as expected value of cash-flows.
  - Replicating strategy may be easy to solve or approximate on some specific cases. Besides, if we can really solve the PDE, the end value is accurate (relative to the, as we have mentioned, impossible conditions). But specific calculations already performed do not help as we pick a new cash-flow pattern to be priced.
  - Once we have the RN measure, we may either calculate the accurate expected value, or use the Monte Carlo method. In the latter we draw samples from the risk neutral distribution, evaluate the cash-flow along each particular realization sample and take the observed average as an unbiased estimator of the exact expected value. This is very convenient for practical calculations and is in fact the definitive approach how risk neutral valuation is carried out. Quite often even the sample – the scenario set – is pre-generated for use with all the various instruments we will price. A notable risk due to the
comfort of not having to solve PDE's all the time is that we can get carried away in choosing the instruments to price.

Instruments that have a fancy payout shape may need a very high number of sample scenarios to price – the pre-generated set is inapplicable then. This threat can be avoided by checking the confidence interval of the average as an estimator.

Instruments whose value depends on risks that were not included in the calibration of the RN measure (in the generation of the scenario set) can not be priced at all, but the practical steps may be possible to carry out and one may misinterpret the respective results as the price. This risk can only be avoided by thinking carefully, or – if available – occasionally back testing the prices using more sophisticated scenario sets (ones that were generated from more complex models and represent a broader range of observed prices).

• If the person establishing the RN distribution is different from the one using it, the latter can work without having to understand at all the fitting process and its assumptions. This is very efficient for splitting the work and implementing operationally robust processes, but clear guidance must be provided on the proper usage of the distribution / scenarios.

• When pricing a selected instrument with replicating strategy, we declare reliance on more liquid, better traded reference instruments on the underlying risk. For risk neutral valuation we calibrate on instruments with less ambiguous, more observable prices that depend on the same risk factors. This, of course, is not a difference, but feels different. And for the RN way we can more easily forget about these quality criteria and fit to just any price we find.

• A replicating strategy does not tempt the user to read off quantiles – e.g. to assess VaR – as does a RN distribution. But the latter is just as meaningless – more precisely, must only be taken as a qualitative, not a quantitative measure of risk. The risk neutral measure is a technical tool forced to provide expected values that match actual prices. Other statistical attributes are sacrificed within the process and can only have meaning by mere accident. We can make nice graphs from them to understand how our instrument works, but as long as we don’t take the average present value per the date of the calibration, the numbers themselves must be handled with extreme care.

So be careful ahead – our example model will calculate MCEV on a fixed size scenario set drawn from a pre-calibrated risk neutral distribution. Shortly, it is a risk neutral Monte Carlo (RNMC for internal use) model. And we will even look at quantiles – just as indication I promise, and for Black-Scholes

4.4 Risks for writing options and guarantees

The price of options and financial guarantees is higher than the expected payout of them – using the words we introduced lately, implied volatilities are higher than expected ones. The consequent gap – in expectation – is the risk premium of the provider. In the perfect world of the Black-Scholes model, such premia would not exist – the provider could execute a perfect hedge program. In reality, providers still execute hedge programs but they are imperfect, and this is the source of the risk that deserves a premium.

Below we nominate the most important risk types for a hedge strategy not being perfect, for three reasons.

---

28 E.g. the Numerix application used by many financial institutions avoids this risk by dynamically drawing new and new samples up to the time that the prices convincingly stabilise, 'converge'. Meanwhile this technique makes auditing tougher, as the very same figures can not be reproduced again. Run times are also less predictable. Numerix is primarily a pricing tool, not a reporting tool. See www.numerix.com.
• These identify where the technical assumptions fall short of tracking real life (and not the other way round). Whenever using textbook option pricing formulae, going through the list and assessing for each item how much it might bias our end results can help in understanding the applicability and accuracy of our results.

• The risk mitigation techniques applicable for the different risks are also different. These range from plainly buffering with own funds (solvency capital) to sophisticated transfers to hedging with more fundamental options from other providers or writing offsetting positions. Financial providers vary in their appetite and capacity to take and possibly neutralize each of the risk types – e.g. an insurance company may decide to take the equity risk and manage dynamic delta hedge to absorb it, but purchase cover for volatility risk from an investment bank.

When comparing retail products and the corresponding financial operations, an important aspect is who finally bears the various risks, and how are the risk premiums paid for.

In the classification we follow ING (2008) and Wiki1.

**Interest rate risk, equity risk**

Direct (linear) risk from the value of the liability depending on the risk free rate (yield curve) and the price(s) of the underlying instrument(s). The primary purpose of a hedge program is usually to eliminate these. The insurer may decide to run them all or in part. For interest rate exposure, the respective risk premium is zero or negligible and the exposure is not, so it is typical to hedge all or most of it if accounting tracks market value (fair value).  

If accounting does not capture market value fully (e.g. expected future fees and charges of standard unit link products are not capitalized), hedging is usually not applied. For equity, the insurer could decide to take more of it to earn some risk premium, but as it is not core business, the usual approach is to hedge at least the accounted exposure.

Avoidance of these risks is how far a dynamic delta+rho hedge can go. For these risk, overhedging is counterproductive, for others just costly.

**Basis risk or gap risk, liquidity risk**

Risk from the impossibility of continuous dynamic hedging due to discrete time, discrete prices and transaction costs.

Liquidity risk on equity prices can be treated by Gamma hedging, but is costly (transfers most of the risk premium), imposes credit risk and if the option written is more complex than the option(s) bought for the hedge, a Gamma hedge fails on large moves.

A Gamma hedge strategy is hard to update in times of large movements, so it only covers the ‘normal flow’.

We can add here operational risk, when we plainly fail to execute what we know should be done and is possible.

Note that ‘illiquidity risk’ often heard lately around Solvency II discussions is fairly similar just at a different scale. Liquidity risk here is loosely the lack of partners to quickly perform small adjustments to our portfolio at low transaction costs and on *current prices* (i.e. there might be a temporary disequilibrium in supply and demand but no or not enough transactions have happened yet to set a different price). We fear that the prices might change against our favor, but ‘the market’ has not yet confirmed or rejected this.

Illiquidity risk there is the lack of partners to buy out large quantities of assets, usually with correlated risks, or not at any reasonable price. Smaller quantities might even turn at prices that we might as well consider reasonable, We are just sitting on our pile of assets but no one comes over to take it all. We

---

29 A reason for partial hedge can be high exposure to parameter risk, e.g. interest rate hedge on premiums receivable in the future is very sensitive to persistency assumptions.
are quite sure that the value of our asset is, even with adjustments e.g. for issuer credit risk, is apart from the currently available market data, but ‘the market’ does not seem to confirm this.

**Volatility risk**

The risk of the *implied* volatilities changing over time.\(^{30}\) Can be hedged with options – of course, at notable price.

Within this thesis this will not investigate volatility hedging, so this risk will go wherever basis risk goes.

**Tracking risk**

Risk from an option being written on some instrument (e.g. a stock index), but the covered investment and/or the hedge positions being in proxy instruments (e.g. an index tracking fund). Usually arises when a perfect match is plainly not possible or incurs excess transaction costs, so ultimate offsetting is not possible. Based on how other risks are treated, may be transferred to the hedge provider or retained.

**Tail risk**

Underestimation of extreme moves due to using distributions with quickly vanishing tails (e.g. and most typically, normal or log-normal). Can be hedged with out-of-the-money options.

**Credit risk**

Risk incurred by possible default of counterparties involved when trying to offset the other risks via purchase of options and derivatives. Though there are techniques to partially disperse credit risk to multiple providers and to ones with higher ratings, credit risk is ultimately unchangeable by nature.

**Parameter risk**

Risk from misestimation or improper application of parameters of the stochastic processes behind the derivative instrument to be priced, absence of a material factor in the pricing etc. – e.g., negligence of dynamic policyholder behaviour that has material impact.

Can be addressed by historical back-testing, comparison of alternative pricing approaches, simulations. By nature, there is no perfect technical way to exclude it (any validation may itself miss important factors).

---

\(^{30}\) Note that the question ‘What is the risk in historical / observed / expected volatilities changing?’ seems interesting, but is, unfortunately, meaningless. We will not enter the details as that could be embedding an additional thesis here.
5 Policies with embedded options

We will start our investigations on using all the machinery from the previous chapters in a simple setup. The suggested client-side valuation approach will be introduced with reference to this example. This way neither is the theory floating in the air without practical application and illustration, nor is the problem raised in the example waiting through long pages for a resolution. However, we will add a few remarks on items (e.g., future premiums) that will only come up later in practice, as we get to more complex layouts.

5.1 A simple setup to exercise rational policyholder behaviour

Let us constrain ourselves to a unit linked policy with a fixed nominal maturity guarantee and an upside potential linked to the performance of one selected equity or index. Be it single premium to exclude the paid-up conversion from the list of available choices. The death benefit should be the account value plus nothing or a very low extra, so we don’t have to precisely assess the value of the death cover and the respective charge from the client’s point of view. The reason for this is that risk insurance is a different market where also mortality expectations (including e.g. the effect of underwriting) are ‘traded’ and we would have too many parameters to the valuation. The analysis will remain relevant for savings policies where the mortality cover is either minor or the respective charges and benefits may be separated as an embedded ‘fair trade’. Similarly we exclude any other risk components.

There are two basic and fairly different ways of packaging such a guarantee into a unit linked policy – using closed end fund or open and fund.

5.1.1 Guaranteed product based on closed end fund

The main characteristics of these policy types are the following.

- The guarantee embedded in the fund value, typically in the form of a structured note. For simplicity let us assume the note is issued by an external provider – this is the typical solution.
- In essence, all policies start as the note is issued – in operation, there is some subscription period before. Multiple funds can not be accessed by the client, because entry is only possible when the fund is started. Regular premiums are also not possible, at least not in any meaningful way. The guarantee is the same to all clients.
- The fund value at all times represents the market value of the maturity benefit without decrements and on surrender or death this market value exits the fund, giving the basis of the benefit. Surrender charges can be low. At maturity the fund value by definition equals the guaranteed payout.
  - The maturity benefit can be replicated as a zero coupon bond plus a call option, the dynamic hedge portfolio is maintained by the provider. Decrements need to be disregarded in the hedge target.
  - From the point of view of the insurer, the asset portfolio behind the reserves and the hedging strategy are simple – the only asset is the note, and the proportional part needs to be sold back to the provider on decrements as the reserves are released.
- Management fee is typically embedded into the structured note as cash outflows that is earned by the insurer.

---

31 Regular premiums could in theory be directed to a sequence of closed end funds, but either the guarantee levels could not be declared at inception, or the product would become overly complex and expensive. Shortly, not viable.
• The market risk of the insurer is very low\textsuperscript{32} – it’s practically just the fee income varying due to market-dependent persistency of the policies, and even that can be removed if the fee is fixed with the provider and swapped with an implicit front-end load in the fund pricing. The market risk of the provider is moderate in the sense that he assembles the structure note with consent for his own risk appetite and reliance on dynamic hedging capacities. The provider need not count with decrements. Consequently, the option can be purchased ‘cheaply’.

• Tracking risk and basis risk, if applicable, are borne by the provider. The insurer runs notable issuer credit risk.

• The margins are also small, because any competitor can set up a similar product with – the effort required to find a provider and set up operation is not more than for any UL product.

• All in all the product is just a split-up and repackaging of a wholesale financial product, the note, to a portfolio of retail products, the insurance policies.

To avoid repeating ‘guaranteed unit linked product built on a closed fund’, we will use the shorthands closed fund product or (structured) note.

5.1.2 Guarantee provided on an open end fund

This is quite a different animal as seen below.

• The insurer connects the product to a fund tracking the index.

• The fund value does not include anything from the guarantee. At maturity the fund value does not cover the benefit if the guarantee strikes.

• The guarantee is administered as an extra component, its payout is financed from additional unit charges over the regular management fees. Reserves must be set for this component as the fees are prepaid.
  ◦ The maturity benefit is replicated as a long equity position (the fund value) plus a complex put option. The put can be purchased from third party or hedged dynamically using various instruments.
  ◦ Except for constraint from accounting rules (e.g. floor at zero), the value of this put within the invested assets can be matched with the reserve of the guarantee component.
  ◦ Benefits on decrements can only be based on the fund value. The release of the guarantee reserve – which, if policy level flooring does not apply or we disregard it to rather consider an unconstrained fair value, can be negative and trigger loss as released – is accounted against company P&L.\textsuperscript{33} In the meantime the unrealised gains or losses on the corresponding hedge assets are realised.\textsuperscript{34}

• Policies can start at any time. Regular premiums are possible. The guarantee has different content for each policy.
  ◦ Furthermore, the insurer may plug multiple fund to the product and can allow switches and redirections of (regular) premium.\textsuperscript{35}

\textsuperscript{32} In the subscription period, i.e. after declaring the guarantee rules and agreeing with the provider on a minimal booking amount, but before the premiums are received and the note is purchased, there can be substantial exposure on interest rate risk. Losses arise if the value of the guarantee decreases when the company already committed to buy it, but clients did not yet commit to take it up – the insurer may and up with the devaluated note partly on general account. We will not deal with this risk here. Some pre-hedging can be used to limit it, actually.

\textsuperscript{33} No immediate cash-flow impact, but future cash-flows will be different from expected.

\textsuperscript{34} This is a cash-flow item with no net impact on P&L.

\textsuperscript{35} Some care need to be taken because when funds with very different volatilities are put together, a flat guarantee charge across funds is hard to maintain, meanwhile for back end operation and client communication a fund-dependent guarantee charge is ... simply dead.
• Management fee can be charged on the fund just as for non-guaranteed fund.
  ◦ However, the fee level should be uniform over the range of accessible funds – if not, it’ll be about impossible to set all policy conditions consistently, let alone fairly and clearly.
• The insurer bears substantial market risk.
  ◦ Even if using put options, a truly dynamic hedge strategy needs to be maintained, because decrements need to be taken into account in the hedge target.
  ◦ Tracking risk is on the insurer. Basis risk can mostly be taken over by the provider if hedging with puts. Credit risk is limited if hedging with non-optional instruments but increases if purchasing options.
  ◦ Given all these complications in any case, the typical approach is to maintain a – possibly outsourced – dynamic hedging plan that relies primarily on non-optional instruments.
  ◦ Expected gains on the gap between implied and realised volatility are gathered by the insurer.
• Margins can be higher because of the subtle operation, especially the maintenance of the hedge portfolio and the guarantee reserves. The most inconvenient boundary here is that all margins taken out increase the chance that the guarantee ends up in-the-money, so kick back on the net guarantee fee.
• The product combines a complex set of wholesale financial instruments, and its own investment capacities, to turn them into a retail insurance policy.

These products we will call open fund product or just VA, partly as variable annuities are important prototypes of insurance products with attached financial guarantees managed on general account, partly as it’s industry practice to use this name even if annuity benefit as such is absent from the design.

For both designs the market value at issue of the benefits could be calculated with the Black–Scholes formula if there were no decrements. With some considerations, also if decrements are independent of the market movements. But results change if decrements and market movements interfere.

5.1.3 Decisions of the policyholder – reduction to clean strategies

Let us first concentrate on the inception of the policy. The customer chooses from buying or not buying the policy. Let us constrain the meaning of the latter to holding the policy to maturity (when assessing the value at issue – the realized strategy later on may suggest surrendering) or a predefined non-dynamic surrender plan.\textsuperscript{36} For a reasonably built product the payout cash-flow stream of such a non-dynamic plan is the weighted average of clean strategies of definitely surrendering at some given time, except that for partial surrenders the fixed charges (if any) may decrease the CF elements.\textsuperscript{37} We are assuming that the monetary value function of the client is linear and positive monotonous, so the maximum can only be taken at an extremal point of the set of strategies, thus a clean strategy.\textsuperscript{38}

A clean strategy of surrendering at an interim date is assumed to be calibrated by the company to provide positive market value on the balance of premiums and benefits – possibly deteriorated by expenses etc. – and provides no explicit access to the guarantee (maybe implicit through the MV within the AV of the closed fund). The benefit is based on investing the premium, rolling forward

\textsuperscript{36} Just to clarify: under such a plan we mean in the most complex case predefined probabilities on all the future decisions between retention, surrender and maybe partial surrender.
\textsuperscript{37} Here we implicitly treated a 50% partial surrender decision as being halfways between retention and surrender.
\textsuperscript{38} This is a standard theorem in linear programming, or can be derived easily from the Krein–Milman theorem (Wiki3).
on market value, applying charges and fees and finally a surrender charge. Such construct, even if preferred by a client, may be accessed through a non-guaranteed UL product which is probably cheaper. Hence the option of not buying should be better. So from the selected set of choices we should only deal with ‘hold to maturity’ and ‘do not buy’.39

We still have the mortality which might be considered by the client. Here comes our assumption that the risk of and the cover for death (also taking into account that the maturity guarantee is lost on death) is negligible or a fair trade.40 Now we have a client who, when deciding whether or not to become a policyholder, only compares the premium and the no-decrement value of the maturity benefit.

From our definition (decision) on rationality, for the modeling of future behavior we will start from the assumption of these two evaluated as equal by the client. We are now seeking for a valuation rule that enables this.

5.1.4 Recycling the BSM formula

A quantification of such a valuation rule that can be carried forward to project projected behavior is altering one or more market-related parameters of BSM on the value of the guaranteed benefit so that it becomes equal to the premium. Again for simplicity we will assume the interest rate and share price dynamics are uncorrelated, so we don’t have to adjust the formula for correlation and volatility of the interest rate. Leaving the spot and strike price, duration, participation and time intact, we can modify the volatility, interest rate41 and dividend rate.

This technique as applied to the volatility resembles the way how implied volatilities are derived from market quotations on options. The heart of the concept is, in fact, the same. However the backgrounds are remarkably different as we will point out, so let us summarize all the possibilities. But before that it is the highest time to present the formulae on the value of an option and the Greeks. There are several forms and notations of it so first of all we enumerate the inputs that we will use.

<table>
<thead>
<tr>
<th>Notation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T, t \in [0, T], \tau = T-t$</td>
<td>Total duration for that the derivative was originally contracted, time elapsed within the duration, time left until settlement date.</td>
</tr>
<tr>
<td>$S = S_t$</td>
<td>The price of the underlying at the respective data – we usually skip the index.</td>
</tr>
<tr>
<td>$K$</td>
<td>Strike price of the derivative. A forward agreement (short equity) pays $K - S_t$ at maturity, a plain vanilla call option pays $\max\left{0, S_t - K\right}$ and a plain vanilla put pays $\max\left{0, K - S_t\right}$ .</td>
</tr>
</tbody>
</table>

39 The inclusion of a dynamic strategy in the upfront valuation of the client – which could make the picture complete – cannot fit in the current scope. Very briefly, it’s more important for the company that such a strategy can be executed when possessing the contract than the possibility that it can be planned for when entering the contract.

40 Probably for a typical savings instrument its impact is lower than other unavoidable errors in the overall calculation.

41 Logarithmic zero rate for the outstanding term.
### Policies with embedded options

#### Notation Description

<table>
<thead>
<tr>
<th>Notation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r = r_t = r_{f,T-t}$</td>
<td>Risk free interest rate (continuous, logarithmic) in the currency that K and our portfolio value is denominated in, for the remaining period. It is the earning on cash – or the opportunity cost of not investing directly into cash. The indexed versions are primarily for stochastic interest models, but we will not add here the rigor in the notation and the comments to deal with those properly.</td>
</tr>
</tbody>
</table>
| $q$ | Usually referred to as foreign interest rate when S is originally traded in FX, thus K as a fixed amount in our currency is a floating amount in the currency of S. Moreover it can represent:  
  • The (continuous) dividend rate that gradually flows out of S, thus devaluing calls and appreciating puts – this is how we use it when calculating the value of the call option embedded into a closed fund.  
  • The can be the annual management charge decreasing the value of the policy account in standard UL product. This is how we use it when evaluating the put option attached to an open fund guarantee UL.  
In all cases it is an earning on directly owning the underlying or an opportunity cost penalizing the contrary. |
| $F = S_t \cdot e^{(r-q)(T-t)}$ | Forward price of S when contracted at $t$ for settlement date $T$. |
| $DF = DF_t = e^{-r \cdot \tau}$ | Discount factor for the remaining period (in local currency). |
| $\sigma$ | Volatility (annualised, of the logarithmic yield, with stochastic interests we would need to further adjust the logarithmic returns by the risk free rate and take the volatility of the residuals). |

Having these notations clarified, the rest is unambiguous and all formulae we use later on (e.g. vega of a plain vanilla option) can be decrypted easily. All are presented at Wiki1, with some comments and explanations.

#### 5.2 Client implied parameters

As these parameters are connected to the yield and risk of the replicating portfolio, we may remotely attach some interpretation, or rather, impression on what difference between the standard financial and client accessible market such changes may represent.

We also attempt to collect pros and cons of the alternatives in terms of applying an adjusted BSM formula to evaluate future cash flows to the client from a saving insurance contract he owns. The convenience in so limiting the scope of application is that we can stick to imagining how various positive payout patterns (linked to equity performance, cash returns or nominally fixed) change their value under the parameter change. We can skip short positions that could easily confuse us.

It is also important that the client is not fully and directly owning shares, even if the account value represents some units of ownership in a fund ultimately comprised of equities. Immediate withdrawal may be impossible or penalized, and to the time of a possible future withdrawal not the total return will be received by (accounted to) the policy. We should evaluate assets behind
the account value through some benefit that they can be released to, like immediate surrender or take up at maturity.

5.2.1 Volatility

The volatility only affects the price of options written on the index, and very importantly the impact has the same sign on call and put options. Upscaling the volatility captures the fact that the average client has limited access to any type of option, i.e. at higher prices (possibly only embedded into products like the one we consider). The effect is the strongest around the ATM position\footnote{Not in the simple nominal sense, but meaning that the strike price is close to the forward price.} and vanishes for far ITM or OTM, but for those cases the rational decision is quite straightforward anyway. The level of adjustment may be comparable between different guaranteed products as it has the relatively clear effect which higher risk aversion would also incur. The client in accepting an upscale of $\sigma$ in the quote of the insurer would act similarly to the insurer buying options at the implied volatilities for hedging the portfolio, though knowing that the way they are manufactured in the backgrounds creates margin to the provider. But the channel sustained from this margin provides retail access to the product.

Meanwhile the volatility adjustment disregards that access to practically any type of saving – equity, even bank account – is subject to considerable charges. Consequently the volatility-adjusted rational client tends to surrender any policy with low or no guarantee. Therefore the adjustment has different effect on European options versus Asian and other path dependent options where the effective market risk of the benefit is more concentrated to the beginning of the option term, whilst charges might rather be distributed along the term uniformly. Looking at the formula of vega:

$$
\nu = Se^{-\tau} \phi(d_+) \sqrt{\tau} = Se^{-\tau} \sqrt{\tau} \Phi(t) (S_T + \frac{1}{2} \sigma^2 \tau \approx K)
$$

$$
= Ke^{-\tau} \phi(d_-) \sqrt{\tau} = Ke^{-\tau} \sqrt{\tau} \Phi(t) (S_T - \frac{1}{2} \sigma^2 \tau \approx K),
$$

we see that the change we commit to the value of options matters most for roughly ATM options. Far ITM and far OTM option values are not affected – they behave either like a fixed income security or like long or short underlying. Approaching the maturity date, the effect vanishes by the square root of which is proportionate to how the value of still interesting options moves. So we can expect the necessary level of volatility adjustment to be consistent along different policy terms.

Let us point out in the meantime, that such client implied volatility is not so well established as the financial market implied one. For the latter, all the other parameters are already traded in large volumes and with low margins, so when implying anything by option prices it can only be the volatility. Impossible as it is, we can at least imagine a continuous, perfect delta hedge to replicate the price. And still the implied volatility shows high variance by the strike price (especially on short and middle terms) so it is not always as unambiguous as we would like it to be. This ambiguity escalates for the client implied one. Especially if we try to use client implied volatility as driver of rational behaviour across products on different risk, or as benchmark across such products, it will be hard to define consistency.

Rather multiplicative than additive adjustment on $\sigma$ could be considered to use common factors across products with different levels of equity risk covered by guarantees. For the thesis we will not have that many alternative products, especially not products so accurately priced, that additive vs. multiplicative would make a difference.
5.2.2 Interest rate

A decrease of interest rate parameter renders any nominally defined future benefit to be more valuable to the client. This adjustment may account for the extra charges on bank account and any fixed income securities available on the retail market, or tax benefit against holding cash or noninsurance investments. For a pure financial valuation of future outcomes, investment units in a money market fund subject to 70bps annual management fee can not compete with directly buying T-bills. It may compete with deposits at retail banks that also offer a lower rate of return that T-bills. If we see that money market funds are taken by the clientele we focus on, we should assume that for whatever reasons, the better FI security offerings are outside their investment radar and the reference rate they use are lower than the reference rates we see.

With recurring payments the effect of appreciating future benefits is partly offset by higher present value of premiums. Insurance premiums, esp. investment premiums are paid in advance, the net balance is still higher perceived customer value.

For plain equity(-linked) positions this adjustment has no impact. Spot prices (equities in the account value that can immediately be cashed in, e.g. accumulated top-ups) do not change as those prices are factual. Value of indirect equity ownership gains on lower discounting rate from cash-in to current, but in the logic of the Black-Scholes framework this is perfectly phased out by having to expect lower average rates of return. Forward equity prices decrease for the same reason. Call options become cheaper and puts more expensive so if we split the benefits to replicating parts the changes might be confusing and unclear.

We may think of a lower interest rate primarily as a decrease of expected return which affects all non-guaranteed instruments available to the client. The impact decays with outstanding term.

5.2.3 Dividend rate

For equity funds capitalizing the dividend gain this has no impact on the value. Lower dividend rate increases the forward equity prices and has no impact on fixed income. Decreasing the dividend yield increases the price of the call and the forward, and decreases the price of the put.

We may think of lower dividend rate as a relative loss on the equity (linked) elements of our portfolio. The same effect also hits the writer of a call-like option ($\Delta > 0$) as the writer of the option is exposed to the same loss in the replicating portfolio – this loss, of course, is transferred to the client as higher price. The important principle of our approach is that the clients perceives this price increase, or eventually the price that incorporates the respective overcharge as compared to the market consistent price, as fair.

The impact decays with outstanding term.

5.2.4 Parallel shift of dividend and interest rate

For all instruments without optionality this just decreases the expected return, but forward equity prices are adjusted back. Looking at the Black–Scholes formula, both $N_t = \Phi_{0,1}(d_t)$ remain unchanged and in $C = e^{-rT}(FN_1 - KN_2)$ and $P = e^{-rT}(F\bar{N}_2 - K\bar{N}_1)$ all elements except the external discount factor stay as they are. This shows in formula that the expected return on an option is also decreasing – which increases the current price as the future benefit is fixed – and nothing else changes. So this is really an adjustment capturing higher charges on any investment, but with no particular extra loading for options.

5.2.5 Suggestion

Some mixture of volatility upscale and parallel downward shift of the two rates seems to be able to capture both phenomena faced by the customer: access to any investment is subject to charges,
especially ones with financial options. Calibration is discreentional as long as we align two parameters on one condition. We can overcome this at least in two different ways:

- Using different products and/or policy terms for a parallel calibration. It should however be remembered that such calibration may never be more valid than the tariffs are marketable, which can a problem if we align the parameters used for one product / duration to other products / durations.

- Calibration first the shift of the rates to charge levels of comparable / competing non-guaranteed products, possibly including the offering of mutual funds and retail bank. Then only the volatility needs to be fitted to our design.

This approach is very close in nature to how implied volatility for (efficiently traded) financial instruments is derived after fixing interest rates to the FI securities market prices.

We will denote with $\eta$ the additive modifier to be applied to the interest and dividend rates in parallel (a negative number), with $\omega$ the additive modifier on the volatility (a positive number). The two together, irrespective of the underlying ambiguity, we will call the client implied volatility shift and the client implied rate shift. When benchmarking products, maybe the names design implied volatility / rate shift express the concept even better.

In our example we will model the client behavior based on such adjusted Black–Scholes valuation, with a narrow but smoothened threshold range around ATM where a surrender probability is chosen. Let us see how it is built up and calibrated.

### 5.3 Decomposition of client behaviour for modelling

It is of course not enough to introduce a concept for rational financial valuation by the customer to enter rational behaviour into our actuarial projection and valuation model. We need to translate our idea to a set of rules that end up as inputs to contingency modelling in the projection of cash flows and accounting figures. To achieve this, we will expand the notion ‘behaviour’ to the following sequence:

- enumeration of alternatives,
- (financial) valuation of consequent outcomes,
- comparison of acquired values, and
- actions taken based on the comparison results.

As typical for actuarial calculation, these actions will be probabilistic. Let us now see what this all means.

#### 5.3.1 The goal: contingency projection

Briefly, we want to get to a set of figures like the one below. Of course, the figures should depend on the projected fund prices and yield curves – this linkage is where rational behaviour enters.

---

43 If we miscalculate the client implied parameters for a product because it’s own tariff is out of bounds, it’s not big deal – if there will be no sales anyway, the policyholder behaviour risk calculated for the imaginary portfolio is not a risk for the business.
### Policies with embedded options

<table>
<thead>
<tr>
<th>Time from projection start (years)</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Current state probabilities</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>In force, premium paying</td>
<td>.95525</td>
<td>.92639</td>
<td>.88816</td>
<td>.84094</td>
<td>.79744</td>
<td></td>
</tr>
<tr>
<td>In force, premium cancelled</td>
<td>.92639</td>
<td>.88816</td>
<td>.84094</td>
<td>.79744</td>
<td>.74471</td>
<td></td>
</tr>
<tr>
<td>Policy was surrendered</td>
<td>.03631</td>
<td>.05675</td>
<td>.07906</td>
<td>.11134</td>
<td>.14063</td>
<td></td>
</tr>
<tr>
<td>Policyholder has died</td>
<td>.00844</td>
<td>.01687</td>
<td>.02533</td>
<td>.03377</td>
<td>.04217</td>
<td></td>
</tr>
<tr>
<td><strong>Transitions in current year</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deaths from PP</td>
<td>.00844</td>
<td>.00843</td>
<td>.00843</td>
<td>.00835</td>
<td>.00824</td>
<td></td>
</tr>
<tr>
<td>Deaths from PU</td>
<td>.00843</td>
<td>.00843</td>
<td>.00843</td>
<td>.00835</td>
<td>.00824</td>
<td></td>
</tr>
<tr>
<td>Surrenders from PP</td>
<td>.03631</td>
<td>.02044</td>
<td>.02222</td>
<td>.03185</td>
<td>.02873</td>
<td></td>
</tr>
<tr>
<td>Surrender from PU</td>
<td>.00845</td>
<td>.00846</td>
<td>.00846</td>
<td>.00838</td>
<td>.00825</td>
<td></td>
</tr>
<tr>
<td><strong>No-decrement modifications in year</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average partial surrender ratio (PP)</td>
<td>.27 %</td>
<td>.292 %</td>
<td>.529 %</td>
<td>.740 %</td>
<td>.958 %</td>
<td></td>
</tr>
</tbody>
</table>

Table 5–1: Typical interim and output figures of persistency modelling

We model in discrete time steps, so what we need is the independent\(^{44}\) state transition probabilities and the partial surrender rate (withdrawal rate) for each period. We will use annualised rates and we will condition them on the state at the start of the period (premiums due at that time already received and invested).

Considering all possible states of a regular premium policy, the rates and ratios we need are the following.

**Notation**

- \(P_{PP|PP}^*, P_{IF|IF}^*, P_{Hold}^*\) Holding the policy and continuing to pay premiums due (if applicable).
- \(P_{SU|PP}^*, P_{SU|IF}^*, P_{surr}^*\) Surrendering an in force, premium paying policy.
- \(P_{PU|PP}^*, P_{pu}^*\) Converting a premium paying policy to paidup status.
- \(P_{PU|PU}^*\) Retaining a paidup policy.
- \(P_{SU|PU}^*\) Surrendering a paidup policy.

**Independent transition probabilities for persistency modelling**

- \(P_{PSU|PP}^*, P_{psurr}^*\) Partial surrender ratio (withdrawal ratio) for a premium paying policy. For single premium policies we blend this into the full surrenders.

- \(P_{PSU|PU}^*\) Partial surrender ratio (withdrawal ratio) for a paidup policy – this we will not use, partial surrenders can be blended into surrenders for paidup policies.

\(^{44}\) Independent here means that the input is, e.g., the probability of surrendering from premium paying state with all other possible events filtered out from the basis. The actual probability that a premium paying policy will be surrendered by the end of the period is less, as the policyholder may die before surrendering.
However, stochastic binary decrements are a good theoretical benchmark to what we rather do, blending of policy states. Blending has no adverse impact when applied to states after death and surrender, or any other termination – an already terminated policy implies no cash-flows or accounting figures, so its characteristics need not be kept track of in the model. Consequently, differences between policies ending up in this state do not matter any more, modelling just the total number of discontinued policies (or rather the number of in force one) is sufficient.

For non-guaranteed unit-linked products also the partial surrenders are OK to model on average – one policy withdrawing 100 units at a given date and another one not withdrawing any yields the same total cash-flows onwards as both withdrawing 50. Paid-up are always a bit fuzzier, because policies converted in different policy years can be quite apart in major characteristics, so it is sometimes hard to find a good 'weighted average' to represent the total cash-flows of them. Policies earning different returns on their fund values can also be blended quite well, because the charges and benefits are functions of just the totals.

For guaranteed products with dynamic behaviour assumed, blending is a material simplification and decreases the cost of the guarantees. This is because a blended state is always closer to the best estimate path than the individual paths it is combined from, so less volatile, so concludes to less convex payout patterns as a function of market prices.

For an example take two policies A and B that start as and remain identical by up to the fifth anniversary. Withdrawals are only allowed at anniversaries 3, 5, and 7 – just now client B needs money and withdraws 100 units out of the total 150, client A keeps all invested. There are another five years to maturity. Assume in the next year the fund price goes up and the yield curve drops. Policy A is out of the money (OTM) for the guarantee as the earning on past premiums is far above the guaranteed return. Client A will surrender her policy so stop paying the guarantee fee. Client B did not benefit a lot from the fund price increase, but he now has a guarantee on the future premiums that compares nicely to the decreased risk free rate. He will retain the policy and with good chances cash in the maturity benefit.

In turn, the blended policy earns less than A and more than B on the equity, so the inclination to withdraw all the money and realize the gain is balanced out by the increasing value of the forward guarantee as interest rates drop.

Both setups deliver the average 33% withdrawal rate at the fifth anniversary, which might even be our average experience from the past. But A and B will on their own will surrender one policy in the next year, while the blended policies remain.

This, of course, is an extreme textbook example. But the heart of it is the same on smaller scale – individually behaving policies span a larger set of states, thus have a higher chance to make decisions between significantly different alternatives, so altogether gain more on being smart.

We can not give up using blended states, but should keep in mind that we downscale the rationality risk in doing so. Hopefully not by magnitudes, though.

By now the calculation scheme is:

• We have some assumptions and parameters, and the projected state of the policy at the beginning of the step. The state is actually a combination from various paths but we accept to live with this.
• Then a miracle appears.
• We have the annualised, independent probabilities and ratios for the transitions and partial surrender. We use these through the step to project the changed state probabilities. We filter out the exits from the state probabilities and blend in the arrivals. Then we can proceed to the next step in the same manner.

We will elaborate on the second step in the next headings.
5.3.2 Enumeration of alternatives

Using the arguments from section 5.1, we only consider the extremes. Thus at inception, we put ‘buy’ against ‘do not buy’. Within ‘buy’ we assume an intention to keep the policy up to maturity, in case of a regular premium policy also to pay all future premiums.\(^{46}\)

Throughout the lifetime of a single premium contract, we allow to choose from immediate full surrender or retention until maturity. With regular premium, this extends with immediate paidup conversion, and withdrawal of the current fund value but continuing payment of premiums.\(^{47}\) The immediate surrender is available separately for policies paying premium and policies already paid up.

Note that each pair of initial state and available alternative now corresponds to rate or ratio listed in Table 5–2, the ones we aim to calculate. A transition probability is the – probabilistic – decision between the current state and the modified state. The alternative of keeping everything unchanged maps to the probability to stay – this we need not calculate as it the residual after the movements. Finally the withdrawal ratio, obviously, teams up with the withdrawal option.

5.3.3 Valuation of outcomes

We use the client-side adjustment parameters to derive a modified market value for the cash flows corresponding to each possibility. The option value is calculated with Black–Scholes formula using the parameter adjustments. Where future premiums are present, the underlying is a bit tricky to calculate. But is possible – see appendix A.1. Very similar calculation is part of the spreadsheet model with that we evaluate the structured note behind the closed fund product if it includes an Asian option.

Still a comfortable flexibility remains in where we set the ‘0’ mark. This can sometimes help to build up plausible comparison functions.

It is always hard to reasonably quantify how numbers possibly close to zero, or having different sign, compare to each other. In most contexts, the difference between –1 and +1 is more than between 100 and 110, so arithmetic difference is not descriptive. Besides, the output range is dependent on policy size, so it requires further transformation to apply some standardized decision function on the outcome.

Ratio is even worse if inputs are not safely away from 0, preferably in the positive direction.

We will occasionally rely on this flexibility to shift comparison inputs to a safe domain, or just to bend the comparison output a bit without altering the comparison function.

This may seem a too subjective, too pragmatic element of the calculation chain. I argue this ambiguity we can spot in how we can think about the client’s utility assessment, without targeting the analysis for the technical approach outlined in the thesis. Imagine a regular premium policy and the decision whether to pay on or to surrender. The straight comparison is between the surrender value for the latter, and future benefits less premiums due for the former. But we can as well compare surrender value plus future premiums unlocked from the plan on one side and the maturity benefit at stake on the other. If the client already planned her future household budget with the insurance premiums ‘spent’, deducted from the regular income just as monthly utility costs, maybe the second comparison is better. If our customer has irregular earnings pattern and has to take care each month to reserve for the premiums due, his views would rather be represented by the first version. And this difference is not because we want to play around with the valuation function in our projection model ☺.

---

\(^{46}\) A determination to stop premium payment at a preset time before the end of the payment term could be added here. However, proper product design – in particular, paid-up conversion penalty – should ensure that for a customer willing to buy the policy, such a strategy is inferior to complete payment.

\(^{47}\) The latter is a technical extreme of partial surrender that is for sure excluded by product design, but as a theoretical option may help to assess the motivation for partial surrender within the available range.
5.3.4 Comparison of two alternative values

Let us first outline a comparison of two possibilities. A purely rational choice would pick a single alternative with full intensity, of course the one with the higher value. However it seems better, at least for analysis but I think also for setting our actual expectations, to allow for some 'laziness' here on the customer's side, appearing as smoothness in the modelling.

Let us introduce some standards for comparison functions $\pi(x, y) \quad x, y \in \mathbb{R}$. They should be as follows

- Normalized: $-1 \leq \pi(x, y) \leq +1$ and $\forall c \in (-1, 1) \exists x, y : \pi(x, y) = c$.
- Symmetric: $\pi(y, x) = -\pi(x, y)$, this implies that it is centered: $\pi(x, x) = 0$.
- Strictly monotonous: $y < y' \Rightarrow \pi(x, y) \leq \pi(x, y')$ and equality is only allowed if $\pi(x, y) = \pm 1$
  
  With symmetry this implies that $\pi(x, y) \left\{ \begin{array}{ll} > & 0 \Leftrightarrow x \left\{ \begin{array}{ll} \geq & y \end{array} \right. \\ < & \end{array} \right.$

- Continuous.

Shortly, if $\pi(x, y)$ is +1 then $x$ is absolutely better than $y$; if it is −1 then completely inferior to it, when 0 they are identical. In between the function behaves nicely and does not throw tapioca pudding at innocent passers-by.

This way various comparison functions can be plugged into the calculation process somewhat consistently, and we can experiment with their impact on the end results. Of course, only if we can invent multiple comparison functions that we like. First it seems that for us a comparison function is made out of thin air, because its inputs are artificial: the monetary valuation function we insisted on introducing, and it feeds an artificial calculation afterwards: the decision function (see in the next heading) that can never be observed on its own. In fact, monotonous transformations of the $[0, +1]$ interval can be carried out freely – but this, as we will see, will apply equally to what we will call a decision function so the two flexibilities collapse to one, which will finally represent the sensitivity of the decisions on the input values compared.

Meanwhile the shape of the function, precisely, the shape of the level curves – for any given comparison distance, the set of $(x, y)$ pairs yielding that value in the comparison function – have some microeconomic interpretation.

To illustrate this we need some functions. First a technical one (rather notation than function), it is just a cap and a floor at ±1.

$$[x] = \begin{cases} x, & -1 \leq x \leq +1 \\ -1 & x < -1 \\ +1 & +1 < x \end{cases}$$

An easy way to get to a regular comparison function is to pick a utility function $u$ with positive marginal utility. Then $\pi_u(x, y) = \left\lfloor u(x) - u(y) \right\rfloor$ clearly is a good comparison function.

Let us now select some $x$ that is preferred over zero by a little, say $\pi_u(x, 0) = \varepsilon$. Then select further elements in a row so that $\pi_u(x', x) = \varepsilon$, $\pi_u(x'', x') = \varepsilon$, ... – the respective solutions can be read from the lever curves of the comparison function... This way we can (in theory) construct a sequence that has constant utility gaps between each step. Gathering a lots of such sequences and using some continuity argument ends up in reconstructing the utility function from the

---

*The client will probably not act immediately on any slight difference – partly because of the excessive effort to be able to do that, partly because we declared earlier that some small benefit from owning the policy we factor out when fitting the client implied parameters, so there will be a narrow threshold zone at least. In the model, extremely narrow tolerance range in the response renders outputs too volatile and increases the exposure to parameter misestimation risk.*
comparison function through these level curves. Still in theory, if there exists at all a utility function compatible with the level curves, it is a well defined, just somewhat hidden attribute of the comparison function.

If we were a perfectionist theoretic dedicated to the concept of revealed preferences, we would consequently limit ourselves to comparison functions that are derived from utility functions. If there is a utility function hidden in the comparison, it seems elegant to declare it up front and not packaging it into a bivariate function. If there is none, the comparison function can not at all result from the client evaluating a utility function and making a probabilistic decision where the decision is purely a function of how the utilities compare. In a very abstract sense, an infinite number of the decisions made on the basis of such a comparison function would violate a very abstract probabilistic generalization of the SARP. Well, the reader may not see this as a tragical event, and then the author humbly joins this opinion, but still a slight inclination grows to prefer utility based comparison functions over others.

We aim to invent a few different comparison functions just to test whether the particular choice can make a difference, or from the alternatives easily popping in our mind we can just choose one freely. Given that comparison functions and utility functions are so good friends, we will use some utility functions on monetary value (of a stochastic cash flow pattern, evaluated via the adjusted BSM formula) to generate the desired diversity of comparison rules.

So hereunder are a few that we will check in our numerical examples. The parameter $T$, where used, serves to make the function scalable – in our model we will set it to the single premium and it will ensure that comparing double values will not result in double the lapse driver if it is just a consequence of projecting a double size policy. The contrary would introduce a substantial noise factor in our model that would be driven by policy size. Such projection is not desired. Besides, it seems fair to me to assume that a policyholder entering a contract at a face amount of 10 million HUF has her nominal thresholds (on amount at stake for considering lapsation) higher than someone contracting for 2 millions.

Having spent so many words on the level curves I plot those as well – I must admit, rather as decoration. Well, they images may help a bit in calibrating the thresholds parameter so that the range spanned by the comparison functions over the stochastics in the projection will be similar across functions. Note that a monotonous transformation on the output of the comparison function does not change the curves, so the corresponding utility function also remains the same, but the way it drives the decision function obviously changes.

The scaled absolute difference pairs up with linear utility. The plot shows the bent version as that is nicer.

$$
\pi_{\text{abs-crop},T}(x, y) = \left\lfloor \frac{x - y}{|T|} \right\rfloor \quad T \neq 0
$$

$$
\pi_{\text{abs-bend},T}(x, y) = \frac{2}{\pi} \arctan \left( \frac{x - y}{|T|} \right) \quad T \neq 0
$$

The square root difference represents a utility function that assigns minus infinity to values below the threshold (negative infinity meaning that the actor pays ‘any’ cost to avoid such an outcome), then the marginal utility by a square root law (marginal utility halved when input quadrupled).
and for our current investigations irrelevant that a parallel change to lapses impacts the portfolio
some static or even irrational part in the decision functions.

Even if they cause financial losses of clients on the respective policies. So we should still allow for
design might cap the partial surrender rate, urgent need for money can trigger some surrenders
However, these may depend on several additional factors not linked to the market status: product
volume of lapsation. First, that can as well be achieved by adjusting the static model. It is known
and famous utility functions on money

5.3.5 Decision and action based on the comparison

In this heading we have to live with the fact of not having extensive and significant experience
data to evidence the modelled behavioural patterns. In the summary and conclusion section we
will touch upon how and for what our results can still be useful. Here we just describe what we
can find for doing it.

The goal is straight – based on the results of the comparisons, we have to assign probabilities and
rates to the actions that the modelled client can perform.

However, these may depend on several additional factors not linked to the market status: product
design might cap the partial surrender rate, urgent need for money can trigger some surrenders
even if they cause financial losses of clients on the respective policies. So we should still allow for
some static or even irrational part in the decision functions.

The latter is a clear example to an important generic phenomenon. Just looking at the insurance contract,
client behaviour seems irrational, and it is fairly common to criticise the rationality postulate on such
observations. We could argue that it is not irrational, just a change in preferences – still a bit of an own
goal, working with SARP or even WARP on a changing set of preferences is uneasy at least.

An argument ‘compliant’ with establishing our work on revealed preferences is spotting that from our
point of view it is the budget line that has changed. Still makes our life harder as we try to use rational
choice as base principle and calibrate the mechanics accordingly – we need to prepare for noise. But such
events does not prove that the technical assumption of rationality is irrational.

To compare with results of static modelling it is desired not to dramatically change the overall
volume of lapsation. First, that can as well be achieved by adjusting the static model. It is known
and for our current investigations irrelevant that a parallel change to lapses impacts the portfolio

\[
\pi_{\text{sqrt-crop}, r}(x, y) = \left[ u_{\text{sqrt}, r}(x) - u_{\text{sqrt}, r}(x) \right]
\]

\[
\pi_{\text{sqrt-bend}, r}(x, y) = \frac{2}{\pi} \tan^{-1} \left( u_{\text{sqrt}, r}(x) - u_{\text{sqrt}, r}(x) \right)
\]

\[
u_{\text{sqrt}, r}(x) = \begin{cases}
-\infty & x < -T \\
\sqrt{1 + \frac{x}{T}} - 1 & 0 < T, T \leq x \\
\sqrt{-1 + \frac{x}{|T|}} - 1 & T < 0, |T| \geq x
\end{cases}
\]

The norm-adjusted difference does not correspond to any utility function, but with the power
parameter is set to 2, it delivers a smooth transition when both positive and negative values need
to be compared.

\[
\pi_{p, r}(x, y) = \begin{cases}
\frac{x - y}{(|x|^p + |y|^p + |T|^p)^{1/p}} & p \geq 1 \\
\frac{x - y}{(|x| + |y| + |T|)} & p = 0
\end{cases}
\]
value. Second, the needless fixed effect makes it even harder to assess how much truly the dynamics matter.

Then because we know that we do not know what to really expect, we should do these.

- Set up smooth relationship so we don’t introduce excess noise for no particular reason.
- Parameterize the behaviour in a flexible way so comparisons can be made.
- Attempt to select parameters that, to the extent possible, we can name in human language.

Then we have the chance to expect something when we change some and we can perform reasonability checks on the response of the model.

Unfortunately these principles are insufficient to unambiguously define the response function – especially that the input to it, the comparison function we invented as fed by the valuation rule just recently introduced, will not be crystal clear. So let us choose functions of the following form and move on.

\[
\Gamma_p(x) = \begin{cases} 
  y_L & x \leq x_L \\
  y_R & x \geq x_R \\
  y_M & x = x_M \\
  y_R + (y_L - y_R) \left( \frac{x_R - x}{x_R - x_L} \right)^\gamma & x_L \leq x \leq x_R 
\end{cases}
\]

This family of functions has the following parameters.

**Left pinpoint**

\((x_L, y_L)\) sets the maximum response and the threshold value at and below that this maximum response is activated. The height is left for professional judgement and experimenting, the width can be calibrated to the effective input range produced by the model through valuation and comparison along stochastic scenarios.

**Right pinpoint**

\((x_R, y_R)\) sets the threshold and amount similarly for minimum response. The level, at least as I believe, can be set reasonably on the basis of observed minimum lapses across a large variety of products, as a frequency estimate on inevitable external factors that can trigger any policy to lapse or surrender. The threshold is rather floating, but in the typical parameterizations that I expect its exact value barely impacts even the curve, not to mention the end results.

**Middle pinpoint**

The curve is expected to pass through \((x_M, y_M)\). The level can be set as a static assumption would be set for such a product or a similar one. Professional judgement may be used to set the location, but we will see in the examples that the principles we listed will bind the choice (which is actually good).

**Exponent**

\(\gamma\) is a consequence of the other parameters and the prescribed functional form.
The idea was taken from gamma correction used in image and video processing. In less than a nutshell, it (re)balances the contrast within a finite input-output box.49

5.3.6 Multiple choice

In the abstract generic case we would need to condition Markov chain state transition matrices on the values of possible outcomes accessible from each of the current states. Of course in true multistate modelling these values would not depend on just the target state, but as well the state that we look at it from, actually on the path how we got there, and maybe also on the next year GDP of Hungary. It is by itself a challenge to ensure that the transition rates over the targets from a selected source add up to 100% if these rates are individually dependent on the respective values and comparisons. Shortly, hereunder we skip this generalised path and stay within a particular example.

A regular premium VA product (open fund UL with externally attached guarantee), as most regular premium UL, offers surrender, partial surrender and paidup conversion as alternatives to keeping up the initial status, in force premium paying. Just what we summarized in Table 5–2.

Reviewing the offering of the product, all the respective independent transition probabilities qualify for dynamic behaviour.

- The policy provides the standard UL investment option – the client can pay in money and withdraw it if she is satisfied with the accumulated earnings.
- An equity guarantee is added to offset the past and future losses.
- An interest rate guarantee is implied by providing future equity guarantees on the original rates.

Consequently:

---

49 http://en.wikipedia.org/wiki/Gamma_correction
Both surrenders and partial surrenders can capitalize on the first feature. Partial surrenders are more appropriate if the interest rate guarantee is valuable, otherwise the full surrender.

Paid-up focus on the equity guarantee to remedy losses already suffered and give up the interest rate guarantee.

Sustained premium payment combines all three and is likely to be the best choice in ‘neutral’ positions.

Partial surrender and surrender on policies already in paidup status are nearly the same (“esp. for modelling”).

Thinking through the possible combinations we find that if one of the cancellation options is notably more valuable than some other or keeping up premium payments, then also the other decrement alternatives are lagging behind. When there really is something happening it is really just one, of course when limiting our thoughts to reasonable client actions of (non)persistency.

Well, PU→SU or PU→PSU come together with PP→SU and PP→PSU but these do not interfere. This suggests that we can restrict ourselves to the binary decisions on the following:

- Should a PP policy…
  - perform full surrender or rather stay?
  - convert to paidup and stay in force or rather keep paying premium?
  - withdraw (lots of) units via partial surrender or rather leave the fund value as is?

- Should a PU policy…
  - surrender or part surrender (a notable part of all its units) or rather keep the investment in the fund?

Projection of these decisions yields the independent rates of Table 5–2 and we can feed the contingency calculations. Well, one could probably imagine more complex product design where such decomposition of the decisions to a set of binary choices would be too simplifying. Yet it was a quite simple and generic observation that when some decrement option is highly favorable to the client instead of keeping the contract, it is probably favorable to any other option as well (or the policy will completely disappear soon). Then a fair estimate on that single large movement should matter more for the value then subtle joint rules considering all possible interactions.

So the startup strategy to dynamic modelling of multiple choice decrement decisions should be to decompose the set of states to pairs and model binary decisions.

### 5.4 Summary of the proposed approach and areas of use

We suggest client implied valuation adjustment parameters to be calculated for and attached to guaranteed UL product designs. These may be used in the following areas.

**Benchmarking and comparison**

Just like market implied volatilities are much easier to compare across times, instruments and countries, we can use the client implied rates for understanding products better. The way we have defined the parameters concludes that they may heavily depend on product characteristics. Meanwhile we should avoid to think that there are as many completely contradicting groups of rational as many products we launch. The client implied parameters reflect some special (opportunity) cost that our targeted clientele should be willing to take when purchasing the product. So it is a compact presentation of how good (or how bad) our offer is.
Pricing and valuation for financial reporting
Here we project future liabilities to value them from the point of view of shareholders. The target audience for the information in the results are shareholder and managers. Outputs can be used to drive strategic decisions like buying vs. selling equities of the insurance company at the stock market, introducing or not a new product, reviewing exiting tariffs and charges, choosing an appropriate hedge program.

Reserving
We project the balance of future non-funded benefits vs. premiums, fees and charges covering for those benefits, expressed in present value terms. (For products and reserving schemes where we would like to account for dynamic behaviour, ‘present value’ will actually be fair value.) If the balance is positive (company is liable for more payout than what its related incomes are worth), reserves must be set up to ensure that the company can meet its liabilities. Based on regulation and the company’s policies, negative reserves may also be allowed (at least at the level of policies or blocks of business – a negative amount for total reserves of the company are unlikely to be accepted in any country).

Hedging
A block of business is evaluated along various projections or rather various sets of projections. This yields estimates on exposures to various market risk factors. By purchasing assets with opposing exposures, the market risk of the total (fair value) balance sheet can be decreased.

When a hedge program is in place, the calculation results directly drive the trade transactions to execute this program.

We provide numerical examples to the first two areas.
We illustrate the presented methodology on single premium UL products with a guarantee. We will use six sample policies and slight modifications of them through the analysis. The values and explanations of major attributes are as follows.

<table>
<thead>
<tr>
<th>Policy term</th>
<th>Fund type</th>
<th>Index</th>
<th>Man fee</th>
<th>Guarantee fee</th>
<th>Rollup (strike)</th>
<th>Participation</th>
<th>Par rate</th>
<th>MCEV(0)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10</td>
<td>open</td>
<td>12.5 %</td>
<td>1.75 %</td>
<td>1.04 %</td>
<td>122%</td>
<td>--</td>
<td>6.1%</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>open</td>
<td>12.5 %</td>
<td>1.75 %</td>
<td>1.44 %</td>
<td>103%</td>
<td>--</td>
<td>6.0%</td>
</tr>
<tr>
<td>3</td>
<td>10</td>
<td>closed-1</td>
<td>12.5 %</td>
<td>1.75 %</td>
<td>--</td>
<td>125%</td>
<td>133%</td>
<td>6.1%</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>closed-1</td>
<td>12.5 %</td>
<td>1.75 %</td>
<td>--</td>
<td>106%</td>
<td>110%</td>
<td>6.0%</td>
</tr>
<tr>
<td>5</td>
<td>10</td>
<td>closed-10</td>
<td>12.5 %</td>
<td>1.75 %</td>
<td>--</td>
<td>125%</td>
<td>238%</td>
<td>6.1%</td>
</tr>
<tr>
<td>6</td>
<td>5</td>
<td>closed-5</td>
<td>12.5 %</td>
<td>1.75 %</td>
<td>--</td>
<td>106%</td>
<td>176%</td>
<td>6.0%</td>
</tr>
</tbody>
</table>

Table 6–1: Introducing the sample policies

Policy term

Obviously, the policy term in years.

Fund type

Distinguishes between products based on an open end fund plus attached put option vs. closed end fund linked to a structured note comprising of fixed income securities plus call option. The numeric suffix indicates the number of averaging points for the call, one for European call and multiple for Asian.

Man fee

Annualised management fee, either over the initial single premium (contracted in advance with the provider of the note so market risk in fee income is minimized) or over account value.

Guarantee fee

The open fund product charges a dedicated fee in excess of the ‘usual’ management fee. In reality, such split is rather the optics of presenting the product to the client. In our model, it is directly linked to the ‘fair’ fee of the attached put option with a 125% multiplier.

Rollup (strike)

The guaranteed maturity benefit over the single premium, so the cumulated and not the annualised figure.

For a closed end product, this refers directly to the underlying (though, based on the design, dividends may be taken out of the total return). For an open end product the guarantee applies to the value (price) after the deduction of the value based fees, so actually represent a strike price higher by (1 + annual fee) on the power of the policy term.

Participation

Payout delta of the call option if it strikes - the $\pi$ parameter in the payout of the note:

$$V_T = K + \pi \cdot (X_T - K)_+ = \begin{cases} K & (X_T \leq K) \\ K + \pi \cdot (X_T - K) & (X_T \geq K) \end{cases}$$

Beware that the underlying of an Asian option has far lower expected value and volatility than a European, hence the higher participation – the nondecremented opening value of the two options is equal to the cent.
Par coupon rate

Indication of the interest rate environment assumed for the date of issue.

MCEV(0)

Is expressed as a percentage of single premium – negatives appear because of the distortions in the assumption set.\(^5\)

6.1 Framework for the calculations in the thesis

We use an Excel model targeted at pricing and analysis of dynamic behaviour in single premium guaranteed UL products, based on either closed or open fund. It is running on outdated and slightly distorted assumptions to avoid confidentiality problems. This also means that the absolute value of the results is not always meaningful. Relative changes within each product setup are relevant as these only rely on the mechanics of the calculations.

The model itself is not considered to be an inherent part of the thesis, partly for confidentiality (is based on a pricing model I have built at ING Insurance, Hungary), partly because some performance optimization techniques used have negative side effects on the stability of the Excel application that I would not love to share with the reader’s computer. On request, the spreadsheet can be made available.

To project dynamic lapse behaviour, we follow this scheme.

• For each projection step (calendar quarter), we calculate the value of surrendering as the immediate cash value, and the value of keeping the policy as the adjusted BSM applied on the current state and assuming no decrements until maturity.

• The resulting values are passed through one of the decision function seen at 5.3.5 to get the ITM measure for the guarantee.

• The assumptions contain pinpoints for standard / floor / cap lapse and the respective \(\Gamma\) function is used to derive the annualised lapse rates from the ITM level.

• These lapse rates are picked up by the model and all else happens as if we had it from a deterministic assumption.

We will review how the existence of the dynamic link as such and how alternative parameterizations impact the projection results. We will also assess if the behaviour really is rational 'ex post', so do the modelled clients benefit from that we alter their lapse pattern based on the ITM of the guarantee.

6.2 Calibration of implied parameters

First we fit the client valuation parameters in a way that the client value of the payouts equals the single premium.

We fit in three ways – rate only (affecting the interest rate and alternative rate parameters of the BSM), volatility only (please guess) and mixed that starts with a 60/40 mix of the previous two and scales it slightly to get the policy purchase to par. The first chart shows pairs of interest adjustment (a negative value) top-down, scaled to the left axis and volatility adjustment bottom-up, scaled to the right axis.

\(^{5}\) Bit of a cheating – this MCEV(0) was already calculated with the base case dynamic behaviour that we just introduce.
Within similar products the results are fairly consistent. The open end products require more adjustment for the volatility approach and less for the interest rate approach – we will return to this a few paragraphs below. The overall image is pleasant – the pure rate adjustment competes with regular bank and mutual offers (in the retail sector) and a tax benefit may apply as an extra. It’s harder to relate the volatility spread parameter – based on ING (2008), similar proportions are normal for the implied vs. historical volatilities, so we do not seem to be out of bounds.

Then we check how these parameters change if we stress the policy design or change the economic environment.

For each policy, the second column represents the base case. To the left we have just increased the annual management fee by 50 basis points. To the right, once the volatility was shifted up from 12.5% to 17.5%, once the yield curve was shifted down by 100 basis points. For these economic
stresses also the rollup rate, participation rate and guarantee charge are adjusted automatically by the model.

As expected, the implied parameters detect the excess charge and signal that rationally buying the altered product requires higher assumptions on the policyholder tolerance level. The net impact of yield curve change and repricing is minimal – which is good, the implied parameters identify that the product has 'not changed'. The response to volatility driven repricing – or, rather, application of the existing product design to a more volatile underlying – shows controversial figures at the first sight.

• For the open end design, the product becomes more costly by all means to the client. Looking at the background figures, the guarantee fees have multiplied themselves and even MCEV(0) increased notably. This indicates that the ×125% safety factor over the necessary option fee is excessive, and the total fee drag will make it hard for the fund value to pass the strike price. As volatility increases we run into a vicious circle – the more we charge the less account value remains to pass the barriers and the option gets even more expensive.

    What we have seen in the base case, the open end fund rather performing better on rate adjustment and poor on pure volatility adjustment also indicates that the actual payout to expect is rather fixed income than equity. More precisely, its payout picks less from large ups than that of the open end fund, so more ‘hypothetical’ volatility needs to be granted to balance the costs with the benefits. On the other hand it is cheaper as a secure, low upwards potential saving vehicle, because when its costs are translated to fixed income (fixed outgo…) we only get some 2% p.a.

However, with a high volatility underlying dynamic behaviour can more and more easily skim the upside, and the exposure of the insurer to economic and insurance shocks is also larger as guarantee value to reserve fund hedge for puts larger amounts to own books. The true guarantee price\(^51\) for the insurer goes up in fact and a guarantee fee without safety margin would not be a good idea either.

All in all, that design just does not fit to the high volatility. Maybe the rollup can be cut further – the rollup rate calibration function is fairly draft in our model – but low guarantees do not sell well. Personally, I vote for providing this type of products to clients who can live with funds having single digit σ.

• For the closed end design, the pure fee adjustment parameter drops and the volatility shift parameter increases, the total impact being about neutral. The chart below shows how the curve of possible pairs of implied parameters changed with the volatility jump. For the particular product design, the parameters of the mixed adjustment are fairly stable against this change in the risk of the underlying.

\(^{51}\) The implicit guarantee fee pricing formulae of the model do assume some dynamic adjustments in the decrements, but are not based on stochastic evaluation.
It seems our imaginary policyholders gathered the product knowledge to make dynamic decisions.

6.3 Value impact of dynamic behaviour

We calculate MCEV at issue for the selected policies under four different lapse assumptions.

Static

Deterministic best estimate assumption is used regardless of the current value of the guarantee.

Dynamic, basic

We use a moderate dynamic pattern based on a parameterized \( \Gamma \) decision function. In this parameter set the level of the base pinpoint equals best estimate, and is positioned to 5% ITM. The cap pinpoints drive the maximum lapse from 5% to 20% at an ITM of –60%. The floor lapse is constant 3% at 40% ITM.

Dynamic, sharp

The driving range of the standard function is proportionally downscaled to a fifth of the original size.
Dynamic, steep
Here we rather extend the output range.

Figure 6–5b: Steep curve for lapse decision, 10 years open fund policy

Dynamic, rational
This setup allows for just a very little smoothness between 100% and 0% lapse, changing around 0% ITM.

Figure 6–5c: Rational curve for lapse decision, 10 years open fund policy

Comparison function is set to what we will describe as default in the next subsection.
Below we gather and plot a few figures to summarize the runs.
Figure 6–6: Impact of dynamic behaviour to SP guaranteed UL product

The green and red bars in the middle show the probabilities of maturity vs. cumulated surrenders. Solid are at or below strike at maturity so receive (or would have received) the guaranteed sum, hollow take upside. The ratio of total solid vs. total hollow is granted, and the ratio of total green vs. total we aim not to change a lot by dynamic projection of behaviour (see 5.3.5).

So whatever the clients may be smart in can only be pumping the coloured liquid between the green and red containers. For VA the goal is clear, the clients want green liquid and red gas – the first is cashing in the put guarantee, the latter is walking away with the upside and not paying fee for the OTM option any more.

For the notes there is not a lot to do, exits also happen on market rates. Anyway interesting to see whether something moves.

The blue diagonals to the left are the correlations between ultimate S/K and probability of staying in force for the whole term. Large negative correlation is smart for the VA and interesting but neutral for the notes. The orange or pale yellow sticks to the right show MC value at inception as a proportion of single premium, measured along the upper value axis.

The chart shows that both the clients and had some success. We could keep the overall lapse level close to the original – let alone the fully rational response but there we could not have had such ambitions –, while the policyholder could squeeze out some extra cash from the VA products. Not much has happened to the closed versions except some applause from the clientele – the rationals kept nearly all of them. The latter simply indicates that the charge structure is fairly constant over

52 Deaths we hide within surrenders (could call it terminated) – they are static and salso mall compared to surrender and maturity, so this won’t mislead us.
time, the cash value is on market prices less a small surrender, so who once decided to join can safely stay.

Looking at the numeric indicators, the profit killer clearly is the immediate and intensive response to both ITM and OTM. The rational clients in our projection do not exit with large amounts at a time, but when they exit they all do. They are gone if the index price is up at maturity, and they are also gone if it ever was high in the meantime. Most columns indicate this, and especially that the impact on total surrenders has different direction for the 5-year and 10-year product. They are all in line with how the behaviour was set up, and in fact the portfolio value got hit badly, though the average annualised lapse increase across all the stochastic scenarios is a mere 110-115% shock, but in reality only one scenario will happen, and that 110-115% will be only one picked from the zeros and ten times upcales that have made up the average in the projection. This is clearly indicated by the standard error of the conditional expected probability of hold to maturity with respect to the economic scenario as condition. This figure also strikes out for the 5-years product where the unconditional probability of hold to maturity actually increases as clients go rational.

<table>
<thead>
<tr>
<th>lapse</th>
<th>MCEV(0)</th>
<th>MV @ 0</th>
<th>MV @ 0</th>
<th>MA</th>
<th>SU</th>
<th>MA</th>
<th>SU</th>
<th>SU</th>
<th>(S&lt;T/K, IF)</th>
<th>P(MA)</th>
</tr>
</thead>
<tbody>
<tr>
<td>static</td>
<td>5</td>
<td>0.63% ± 2.0%</td>
<td>14.5%</td>
<td>69.9%</td>
<td>32.1%</td>
<td>47.5%</td>
<td>12.1%</td>
<td>8.2%</td>
<td>-49.2%</td>
<td>0.0%</td>
</tr>
<tr>
<td>dynamic</td>
<td>5</td>
<td>0.53% ± 2.1%</td>
<td>15.2%</td>
<td>69.2%</td>
<td>32.5%</td>
<td>46.9%</td>
<td>12.8%</td>
<td>7.9%</td>
<td>-51.1%</td>
<td>1.3%</td>
</tr>
<tr>
<td>sharp</td>
<td>5</td>
<td>0.37% ± 2.2%</td>
<td>12.1%</td>
<td>72.1%</td>
<td>33.9%</td>
<td>48.1%</td>
<td>11.6%</td>
<td>6.4%</td>
<td>-54.6%</td>
<td>2.4%</td>
</tr>
<tr>
<td>steep</td>
<td>5</td>
<td>0.46% ± 2.1%</td>
<td>16.9%</td>
<td>67.7%</td>
<td>32.6%</td>
<td>45.3%</td>
<td>14.4%</td>
<td>7.7%</td>
<td>-53.2%</td>
<td>3.2%</td>
</tr>
<tr>
<td>rational</td>
<td>5</td>
<td>-0.08% ± 2.5%</td>
<td>11.1%</td>
<td>82.1%</td>
<td>37.8%</td>
<td>52.5%</td>
<td>7.2%</td>
<td>2.5%</td>
<td>-46.2%</td>
<td>5.9%</td>
</tr>
<tr>
<td>static</td>
<td>10</td>
<td>3.71% ± 2.8%</td>
<td>24.8%</td>
<td>43.3%</td>
<td>26.8%</td>
<td>33.2%</td>
<td>22.1%</td>
<td>17.8%</td>
<td>-65.4%</td>
<td>0.0%</td>
</tr>
<tr>
<td>dynamic</td>
<td>10</td>
<td>3.23% ± 2.7%</td>
<td>25.1%</td>
<td>43.5%</td>
<td>27.8%</td>
<td>31.4%</td>
<td>23.9%</td>
<td>16.9%</td>
<td>-69.5%</td>
<td>3.8%</td>
</tr>
<tr>
<td>sharp</td>
<td>10</td>
<td>2.90% ± 2.7%</td>
<td>19.7%</td>
<td>47.9%</td>
<td>30.1%</td>
<td>30.4%</td>
<td>24.9%</td>
<td>14.5%</td>
<td>-73.8%</td>
<td>8.3%</td>
</tr>
<tr>
<td>steep</td>
<td>10</td>
<td>2.80% ± 2.6%</td>
<td>27.0%</td>
<td>42.4%</td>
<td>27.8%</td>
<td>27.3%</td>
<td>28.0%</td>
<td>16.9%</td>
<td>-72.6%</td>
<td>8.4%</td>
</tr>
<tr>
<td>rational</td>
<td>10</td>
<td>0.70% ± 3.1%</td>
<td>2.9%</td>
<td>60.9%</td>
<td>34.0%</td>
<td>20.0%</td>
<td>36.7%</td>
<td>9.3%</td>
<td>-65.7%</td>
<td>31.8%</td>
</tr>
</tbody>
</table>

Table 6-7: Summary of dynamic lapse intensity impact on open end fund product

Worth spotting in the pile of figures above that our naked market risk exposure increased only by a little, where it did at all. Whether the hedgeability of this risk got worse is a more interesting question – probably yes, but this model can not quantify it.

Before concluding, we check the stability of our results.

### 6.4 Sensitivity to technical assumptions

The comparison of the comparison functions is still to come, though we don’t expect surprises here. The data confirms this assessment – below we produce a chart similar to that in the previous heading, but make it more compact because there are more rows there but less things to see.

A +100bps sensitivity to the η, or a -5% shift to ω also have just a small impact to the decrements, though in the mixed setup these practically annulate the respective adjustment terms. Lapses slightly increase, but as not in correlation with the evolution of the value of the option, the portfolio value is not affected. This response being weak is attributable to the surrender penalties and the response function not being aggressive. As planned, we aimed to keep the overall level and a steep function would also have gone high and make this goal hard to reach.
6.5 Concluding the calculations

The closed end fund products are out of the game as expected, we now focus on the open fund with externally attached guarantee.

• Yes, dynamic surrender is confirmed to be a risk for these products.
  - Half percent loss in profit margin for a 5 years SP product is not negligible, but tolerable (especially if some hedge is set up so the actual outcome is not half percent subject to ten percent noise). And it only strikes with the extremist non-surrender behaviour.
  - Three percent off on ten years is a lot for single premium, even if coming from the worst case client. Close to a percent of the initial premium can be lost on more moderate patterns. Meanwhile the expected loss up front is not a striking amount if the ‘basic’ pattern we have set up is not stupidly low. (I hope not.) Less than half percent expected loss (relative to a pricing result probably on deterministic decrements) is not too bad.

• It’s the swift movements that cause the problem, the ‘steep’ pattern on a double output range hurts much less.
  - This is bad news – it increases the convexity of the liability and makes hedges harder, more costly and augment the risk of overhedging.
  - Let us not forget, the problem is not even the clients changing their mind over time, as may apply to usual shocks we count with. We have derived these figures with clients modelled to ‘like’ the product, they have not lost their implied adjustment parameters and were not waivered of surrender charges, just took the money when it made sense, just as it made sense a few years earlier to enter the contract.

If we assume that our real clients made an educated decision in purchasing such policies from us, they thus understood and assessed the nature of the benefits offered by it, the same level of education assumed on their further behaviour is enough to choose wisely. (We just turned that decision upside down to drive their actions onwards in the model.) It is a very tough assumption that sale of these products happens between two knowledgeable parties, but it’s also a goal that insurers must seek to achieve or approach.

Furthermore the financial awareness increases dramatically and even arbitrage can arise if we consider secondary markets of guaranteed investment policies. This does appear in real life, see Scanlon (2010) for an example.
7 Summary and conclusions

7.1 Achievements

In the thesis we introduce the idea of client implied volatility and interest rate shift to practically parameterize an approach to modelling rational policyholder behaviour. We presented the parenting concepts, revealed preference and implied volatility, arguing that our extension is essentially compatible with their own logics.

To my best knowledge, this interpretation of modelling policyholder rationality is unique in actually defining what the word means to us in the particular setup. This does not render it more accurate, but allows for principle based assumption setting in an area where data based approach is hard to perform. In addition, these attributes provide a compact measure on the overall cost level of a product that can be used for comparisons.

We enumerated the modelling steps needed to actually project dynamic behaviour on the basis of it, and finally reported on an example implementation.

7.2 Areas for further exploration

In the thesis we only used the estimated parameters for design benchmarking and value assessment of single premium unit linked products with a guarantee. The product scope can be extended with regular premium.

In terms of application areas, risk management may deserve even more attention than valuation, especially that policyholder rationality is often considered as a worst case scenario.

7.3 Limitations

What we suggest cannot be applied, or at least not in a reliable way, to policies that are not saving oriented. A well defined guarantee, one that materially contributes to the contract’s value, is necessary.
Appendices
A Additional calculations

A.1 Approximate value of an ‘inverse Asian’ option

For the applicability scope it is important to declare that a Black-Scholes formula can be used on special inputs to calculate the option value embedded in a regular premium guaranteed UL. Besides: a very similar calculation routine was used in our example model to deal with Asian call embedded into the structured note.

The description of the calculation is attached because it did not seem to be available from the literature.

A European put or call option is when the payout function depends solely on the price of the some selected instrument at maturity, denote as \(X(T)\). We can formulate it as investing an initial amount \(N(0)\) into the instrument at price \(X(0)\), rolling it forward to time \(T\) and applying the payout function on \(N(0) \times X(T)/X(0)\). The option is driven by the performance of the selected asset in the period \([0, T]\), \(X(T)/X(0)\).

For an Asian option, the underlying input to the payout function is more complex: we invest \(N(0)\) an amount \(N(0)\) into \(X\) at time 0, but then measure the price at multiple interim time points, e.g. \(X(i \cdot T/k)\) for \(i=1\ldots k\), and apply the payout function on \(Y = X(0)/k \cdot \Sigma X(i \cdot T/k)/X(0)\). The option is driven by the performance of \(X\) over a set of gradually extending periods, each starting at 0.

In a regular premium UL policy with a guarantee, the Asian option logic is turned around: investments occur regularly from 0 to (and not including) \(T\), and the underlying of the option is \(Y = AP \cdot \Sigma X(T)/X(i \cdot T/k)\). (For simplicity in the notation we keep the annual premium constant.) In a Black–Scholes model, the theoretical distribution of \(Y\) is correlated convolution of several log-normals – not log-normal but fairly close to it in shape. So with some fitting we can use the generic version of the Black–Scholes–Merton formula to value such options.

The original formula is assuming the spot rate has no stochastic behaviour. In section 6.2 of Baxter–Rennie (1996), the formula on a European call under stochastic spot rate shows that the original formula still remains valid as long as the stochastic noise of the logarithmic equity index and the stochastic noise in the continuous spot rate is uncorrelated and the joint distribution is log-normal, and we use the zero rate of the respective duration when it is used for discounting to the pricing date. Intuitively this applies also to the calculation below on the approximated value of the Asian call, though a rigorous proof will not be given. So we will assume an interest rate structure which is varying by time, but only as it shifts forward on the yield curve known at start and has no stochastic noise in it.

Let us assume an Asian call on some equity index starting at \(S(0)\) with volatility \(\sigma\), the duration of the covered period be 0 to \(T\) with \(N\) averaging points. We introduce the following notations:

\[\Delta T = \frac{T}{N}\]  
Time step between averaging points.

\[t_i = i \cdot \Delta T \quad (i = 1 \ldots N)\]  
Averaging points for the option

\[r_{u,v|t} \quad \text{and} \quad r_{i,j|t}\]  
Logarithmic (forward) zero rate from time \(u\) to time \(v\) or from \(t(i)\) to \(t(j)\), calculated from the term structure at \(t\).

---

53 Precisely, this formula describes the underlying of an arithmetic Asian option. A geometric Asian option pays on the \(k\)-th root of the product of these prices.

54 A simple simulation program verified that the relative error of the closed formula approximation we derive here is not more than 1‰ with realistic parameters on duration, premium frequency and volatility of \(X\).

55 Section 4.5 on quanto options gives more details on a similar argumentation and calculation.
\[ m = \lfloor t / \Delta T \rfloor, \] so
\[ m \Delta T \leq t < (m + 1) \Delta T \]

For any time \( t \) the index of the latest averaging point which is not after \( t \) (or zero if we are before the first one).

\[ \tau = (m + 1) \Delta T - t \]

The remaining duration to the next averaging point at time \( t \).

\[ r_{\tau|t} = r_{t,t+\tau|t} = r_{t,(m+1)\Delta T|t} \]

The logarithmic zero rate to the next averaging point. We will sometimes refer to it as spot rate since rates at time \( t \) for a shorter term will be irrelevant in the calculations.

\[ Q, \quad (Q|t) \]

The risk neutral probability measure for the equity process and its conditional on the filtration at (i.e. information up to) time \( t \). Then \( Q = (Q|0) \).

\[ \sigma \] and \( \delta \)

The volatility and dividend rate of the equity (both annualised, normal in distribution. For the conditionals

\[ K \]

The strike price of the call.

\[ Z_i = \ln(S_{i+1}/S_i), \quad Z_i|t = (Z_i|t) \]

The one-step change in the logarithmic equity price and its conditional on \( t \).

\[ C_u = \ln(S_u/S), \quad C_{u|t} = (C_u|t) \]

The cumulated logarithmic change up to time \( u \) or \( t(i) \). Then \( C_i = \sum_{j=1}^i Z_j, \quad C_{i|t} = \sum_{j=1}^i Z_{j|t} \).

\[ Z_{i|t} \equiv \xi_i \quad \text{for} \quad i \leq m \quad (t_i \leq t) \]

A realization of the random variable \( Z_i \) – the notation stresses that when calculating at time \( t \) and beyond, this is a known constant value.

\[ C_{u|t} \equiv \zeta_u \quad \text{for} \quad u \leq t \]

\[ C_{i|t} \equiv \zeta_i \quad i \leq m \]

A realization of the random variable \( C_u \) or \( C_i \).

Then \( Z_i \sim Q(0) N((r_{t_i,t_{i+1}} - \delta - \sigma^2/2) \cdot \Delta T, \sigma^2 \Delta T) \), the \( Z_i \) series is independent and joint normal in distribution. For the conditionals

\[ Z_{i|t} \sim \begin{cases} 
\xi_i & i \leq m \quad (t_{i+1} \leq t) \\
N((r_{t_i,t_{i+1}} - \delta - \sigma^2/2) \cdot \tau, \sigma^2 \tau) & i = m \quad (t_i < t \leq t_{i+1}) \\
N((r_{t_i,t_{i+1}} - \delta - \sigma^2/2) \cdot \Delta T, \sigma^2 \Delta T) & i > m \quad (t < t_i)
\end{cases} \]

and the series \( Z_{i|t} \) is also independent and joint normal.\(^{56}\) Consequently

\[ C_{i|t} \sim \begin{cases} 
\zeta_i & i \leq m \\
\zeta_i + N((r_{\tau|t} - \delta - \sigma^2/2) \cdot \tau, \sigma^2 \tau) & i = m + 1 \\
\zeta_i + N((r_{t,t_i} - \delta - \sigma^2/2) \cdot (t_i - t), \sigma^2 (t_i - t)) & i > m + 1
\end{cases} \]

Here we used that \( r_{t,t_m+1|t} \Delta T + r_{t_m+1,t_{m+2}|t} \Delta T + \cdots + r_{t_{i-1},t_i|t} \Delta T \), the duration-weighted sum of the forward rates from \( t \) to \( t_i \), is equal to the duration-weighted zero rate \( r_{t,t_i|t} \cdot (t_i - t) \) (all conditioned on \( t \)). The \( C_{i|t} \) series is joint normal and of course the elements are correlated.

\(^{56}\) From now on we will drop the reference to \( Q \) and \( (Q|t) \) where self-explanatory.
For the Asian call, the payout is \((A - K)^+\) for \(A = \frac{1}{N} \sum S_i = \frac{S}{N} \sum \exp\{C_i\}\). Here all \(\exp\{C_i\}\) are log-normal so the sum has approximately log-normal distribution if \(N\) is not very high. We will substitute the distribution of this sum with a log-normal with the same expected value and deviation, and will evaluate the call option by the generic Black-Scholes formula. This states that if \(Z^* \sim N(0, 1)\) and \(F, \bar{\sigma}, K\) are constants and \(\Phi\) the standard normal cumulative distribution function, then

\[
\mathbb{E}\left(\left(F \cdot e^{\sigma Z - \sigma^2/2 - K}\right)^+\right) = F \cdot \Phi\left(\frac{\ln(F/K) + \bar{\sigma}^2/2}{\bar{\sigma}}\right) - K \cdot \Phi\left(\frac{\ln(F/K) - \bar{\sigma}^2/2}{\bar{\sigma}}\right).
\]

So first we have to calculate the mean and standard deviation of \(A\) under the RN distribution. The mean is easy – RN means the equity provides the same expected total yield as the risk free investment so

\[
\mathbb{E}(S_i|t) = \begin{cases} S_i = S_0 \cdot e^{\bar{\gamma}i} & i \leq m \\ S_i \cdot e^{(r_t + i\delta)(t_i - t)} & i \geq m + 1 \end{cases}
\]

Let us introduce \(F_{i|t}\) for this expected value – it is the future price of \(S_i\) at time \(t\) if \(S_i\) is in fact in the future, otherwise it’s just the realised value of \(S_i\) without interest rate rollup. With some calculation

\[
\mathbb{E}\left(S_i^2|t\right) = \begin{cases} F_{i|t}^2 & i \leq m \\ F_{i|t}^2 \cdot e^{\sigma^2(t_i-t)} & i \geq m + 1 \end{cases}
\]

The key for the second case is that \(S_{i|t}\) is log-normal with double drift and double volatility (quadruple \(\sigma^2\)) compared to \(S_{i|t}\). Hence the upscale in \(\mathbb{E}\left(e^{N(\bar{\mu}, \bar{\sigma}^2)}\right)\) over \(\mathbb{E}(N(\bar{\mu}, \bar{\sigma}^2))\) due to the volatility is quadruple in \(S_{i|t}\) compared to \(S_{i|t}\), while the RN balancing part \(-\frac{\sigma^2(t_i-t)}{2}\) in the drift of \(S_{i|t}\) is only doubled, so in \(\mathbb{E}\left(S_i^2|t\right)\) the ‘missing’ \(-\sigma^2(t_i-t)\) increases the result from \(\mathbb{E}\left(S_i^2|t\right)\) to the value shown above.

We now need the covariances between different \(S_{i|t}\) variables. Let us calculate \(\mathbb{E}(S_iS_k|t)\) for \(i < k\). If \(i \leq m\) then \(S_{i|t} = Sc^i\) is known at \(t\) so it can be carried out from the conditional expected value and we get \(\mathbb{E}(S_i|t) \cdot \mathbb{E}(S_k|t)\). For \(i \geq m + 1\) we decompose \(S_{i|t} = S_{i|t}^2 e^{Z_{i+1} + \cdots + Z_{k|t}}\). In this product the factors are independent so

\[
\mathbb{E}(S_iS_k|t) = \mathbb{E}(S_i|t)^2 \cdot \mathbb{E}(e^{C_k-C_i}).
\]

The second is the \(t(i)\) to \(t(k)\) risk free escalation factor because of risk neutrality, and the first we already know. Substituting we have

\[
\mathbb{E}(S_iS_k|t) = F_{i|t}^2 \cdot e^{\sigma^2(t_i-t)} e^{r_{i+k}(t_k-t_i)} = F_{i|t}F_{k|t} \cdot e^{\sigma^2(t_i-t)}.
\]

Summarizing the cases:

\[
\mathbb{E}(S_iS_k|t) = \begin{cases} F_{i|t}F_{k|t} & i \leq m \\ F_{i|t}F_{k|t} \cdot e^{\sigma^2(t_i-t)} & i \geq m + 1 \end{cases}
\]

Putting the pieces together, for \(X = N \cdot A = \sum_{i=1}^{N} S_i\) we have

\[
\begin{align*}
\mathbb{E}(X|t) &= \sum_{i=1}^{m} F_{i|t} \\
\mathbb{E}(X^2|t) &= \sum_{i,k \leq m} F_{i|t}F_{k|t} + \sum_{i,m} F_{i|t}F_{k|t} + \sum_{i,k > m} F_{i|t}F_{k|t} \cdot e^{\sigma^2(u-t)} \\
\mathbb{D}^2(X|t) &= 0 + 0 + \sum_{i,k > m} F_{i|t}F_{k|t} \cdot \left[e^{\sigma^2(\min(t_i,t_k)-t)} - 1\right]
\end{align*}
\]

For a log-normal \(\bar{A}_t \sim LN(\bar{\mu}_t, \bar{\sigma}_t^2)\) with the same mean and deviation as \((A|t) = (X/N|t)\) we use
\[
\sigma_t^2 = \ln \left( 1 + \frac{\mathbb{E}^2(A|t)}{\mathbb{E}^2(A|t)} \right) = \ln \left( 1 + \frac{\sum_{i,k>m} F_{i|t} F_{k|t} \left[ \exp\{\sigma^2 \cdot (\min(t, t_k) - t)\} - 1 \right]}{\sum F_{i|t}^2} \right)
\]

\[
\mu_t = \ln (\mathbb{E}(A|t)) - \frac{1}{2} \sigma^2 = \ln \left( \frac{1}{\mathbb{E}} \sum F_{i|t} \right) - \frac{1}{2} \sigma^2
\]

The value at \(t\) of the Asian call, if we use \(\tilde{A}_t\) instead of \((A|t)\), can be calculated in the following way. Let \(F_t = \exp \left\{ \mu_t + \frac{1}{2} \sigma_t^2 \right\} = \frac{1}{\mathbb{E}} \sum F_{i|t}\) then \(\tilde{A}_t = F_t \cdot e^{\delta_t Z}\) for some standard normal \(Z\). Now by the generic version of the Black–Scholes formula we have

\[
\mathbb{E} \left( (\tilde{A}_t - K)^+ \right) = F_t \cdot \Phi \left( \frac{\ln(F_t/K) + \sigma_t^2/2}{\sigma_t} \right) - K \cdot \Phi \left( \frac{\ln(F_t/K) - \sigma_t^2/2}{\sigma_t} \right)
\]

The value of the option is then

\[
V_t = e^{-r_t,T(t-T)} \cdot \mathbb{E} \left( (\tilde{A}_t - K)^+ \right)
\]

### A.2 Actual implementation to calculate the value

The end formula and how we got there may seem complicated. However with some structuring of the calculations, it can be implemented in a compact and efficient manner.

Looking only at the results, to calculate \(V_t\) for the option we need the following inputs

- \(T, N\) (then we have \(\Delta T = T/N\))
- \(t\) (then we have \(m = \lfloor t/\Delta T \rfloor\))
- the zero rates \(\{r_{i,T|t}\}\) for \(i \geq m\)
- the average of the \(F_{i|t}\) values (already realised components of \(A\))
- the dividend rate \(\delta\) and the volatility \(\sigma\)
- the actual equity price \(S_t\)

**Variables**

\(FSum_i \stackrel{\text{def}}{=} \sum_{k \leq i} F_{k|t} \quad (i = 1..N), \) partial sum of conditional forward values \(F_{i|t}\) values up to index \(i\). \(FSum_m\) is the realised average multiplied by \(m\), further elements to add are \(F_{k|t} = S_t \cdot e^{(t-k) r_{i,(k)|t}} \quad (i = 1..N)\).

By rearranging the sum for the square deviation of \(X\) we get

\[
\mathbb{D}^2(X|t) = 0 + 0 + \sum_{i,k>m} F_{i|t} F_{k|t} \cdot \left[ e^{\sigma^2 (\min(t, t_k) - t)} - 1 \right] = \sum_{i > m} F_{i|t} \left( e^{\sigma^2 (t_i - t)} - 1 \right) \cdot \left[ F_{i|t} + 2 \sum_{k > i} F_{k|t} \right] = \sum_{i > m} F_{i|t} \left( e^{\sigma^2 (t_i - t)} - 1 \right) \cdot \left[ F_{i|t} + 2 \left( F_{\sum}^{(\Sigma)} - F_{i|t}^{(\Sigma)} \right) \right]
\]

This can be calculated with a simple algorithm in linear time, so we have all inputs to the final formula.

If \(t\) is negative, the value of the option is discounted more by the function and the result is a correct market value. If \(t\) is beyond the expiry, the function returns the payout from the realised average, without any time value adjustment.
<table>
<thead>
<tr>
<th>Reference</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Note the existence of the (Draft) Level 2 Implementation Measures that has been circulated for discussion and even many reactions to it are publicly available, but the text itself not. Its future existence and scope is driven by the Directive, so as long as it does not formally emerge, the anticipated text I plainly attribute to the same, compounded bibliography item.</td>
<td></td>
</tr>
<tr>
<td></td>
<td><a href="http://www.econ-pol.unisi.it/~afriat/Econ_PCP.pdf">http://www.econ-pol.unisi.it/~afriat/Econ_PCP.pdf</a></td>
</tr>
<tr>
<td></td>
<td>Actual construction of a utility function, given a finite set of SARP compliant observations.</td>
</tr>
<tr>
<td></td>
<td>A reference on derivatives pricing and risk neutral valuation from binomial models through to multidimensional Brownian motion and Itô calculus. It features substantial rigor in the mathematical presentation, but – or rather: consequently – is not an easy reading. Useful even if one gets lost in the proofs after a while (at least that’s my experience).</td>
</tr>
<tr>
<td></td>
<td>These two are the definitive articles behind the BSM option pricing formula that earned the Nobel prize in economics for the authors in 1997 (Black, deceased by then, was mentioned, but the prize itself can only be granted to the living).</td>
</tr>
<tr>
<td></td>
<td>While the formulae and the mathematical […] have changed a little when compared to current presentations of the topic, interpretation(s), critics, and analysis of the conditions created a diverse set of publications in contemporary economics literature.</td>
</tr>
</tbody>
</table>
CFO Forum, 2009  
*MCEV Principles and Guidance*, Stichting CFO Forum Foundation

http://www.cfoforum.nl/downloads/
CFO_Forum_MCEV_Basis_for_Conclusions.pdf

Principles agreed by the CFO Forum os European Insurares for MCEV calculation and disclosure, guidance for applying the principles, and a supplementary document on the backgrounds. Currently withdrawn to be the definitive rule set for MCEV disclosure, but still a robust reference point for agreed best practices in market-consistent valuation in insurance.

Eling–Kiesbauer, 2011  


Giesiger, 2010  

http://ssrn.com/abstract=1395390

An introduction to risk neutral valuation that focuses on plain language and finishes off by a textual conclusion, not the most extraordinary formula that it could reach in 27 pages – useful.

ING, 2008  
(Szőbeli közlés) ING Investment Management Hungary

Kiesenbauer, 2012  

http://www.soa.org/NAAJ/
http://2012/no-1/naaj-2012-vol16-no1.asp

Kim, 2006  
Kim, C., *Report to the Policyholder Behavior in the Tail Subgroups Project*, working paper, Society of Actuaries


Park, J., *The Study on Dynamic Lapse Ratio Forecasting Methods, working paper*, Korean Insurance Development Institute

http://www.actuary.or.kr/_new/upload/eaac/Session%2013_2_1.pdf

Both studies perform statistical analysis on dynamic lapsation from the aspect of macroeconomic factors, instead of product features and individual client decisions.

Kling et al., 2011  

http://www.actuaris.com/site/pdf/4e8dd713e78a2.pdf


http://www.secondaryinsurancemarketblog.com/weblog/2010/02/


http://www.sims.berkeley.edu/~hal/Papers/2005/revpref.pdf

A brief yet wide-scope review of the revealed preference concept and its presence in economics literature, with emphasis on recent empirical studies. Besides the clear and compact main text, features a very rich bibliography where deeper dives into the topic can start from.

Wiki1  Black–Scholes and Greeks (finance), Wikipedia entries

http://en.wikipedia.org/wiki/Black-Scholes


The sources of the formulae that we have skipped in the text.

Wiki2  Convergence of random variables, Wikipedia entry


Wiki3  Krein–Milman theorem, Wikipedia entry

http://en.wikipedia.org/wiki/Krein%E2%80%93Milman_theorem

Wiki4  Rational pricing, Wikipedia entry

http://en.wikipedia.org/wiki/Rational_pricing

Wiki5  Risk neutral measure, Wikipedia entry


Also consider its links.

Wiki6  Volatility smile, Wikipedia entry

http://en.wikipedia.org/wiki/Volatility_smile

All web links were checked to be accessible in May 2012, so date of download is not indicated individually. These sources were publicly available and appeared to present proper recognition of the authors, involved institutions and copyrights as applicable. Snapshots were saved to the CD attachment as PDF files, per author or company and including all these information to identify the original sources.