Universal life – a boiling frog problem?

By Paul Truyens and Mike McDougall

Presented at the Actuarial Society of South Africa's 2020 Virtual Convention
6–8 October 2020

ABSTRACT
Universal life policies were introduced to South Africa in the 1980s. The policies were designed to give the customer increased flexibility but, due to the leveraged effect of investment returns on the risk charges, policies have become unsustainable during a period of low nominal investment returns. This paper reviews the progress of these policies, mitigating actions that can and have been taken and reviews the development of this market from a professionalism perspective

KEYWORDS
Universal life, sum at risk, investment returns, sustainability, professionalism

CONTACT DETAILS
Paul Truyens; Email: paul.truyens@gmail.com
Mike McDougall; Email: mmcdougall@actuarialsociety.org.za

EXECUTIVE SUMMARY

Universal life styled products dominated the market for whole life and endowment policies during the 1980s and 1990s. Universal life products introduced flexibility and transparency to the market as they replaced the traditional with-profit reversionary bonus style products.

This paper is in three sections. Firstly, the review of the history and evolution of universal life policies as they evolved from being a vehicle for predominately investment policies to dominating the risk product market.

The second and most significant section focuses on the performance of these products and the impact on policyholders. While universal life policies introduced transparency and flexibility, they also transferred more of the risk of investment under-performance from the insurer to the policyholder. The products were introduced during a period of sustained double-digit inflation and nominal yields.

With risk policies being sold in an increasingly price sensitive and highly competitive market, insurers combined aggressive pricing with shortened guarantee periods, the guarantee period being the term during which the insurer could not increase premiums to offset deteriorating investment performance. Any shortfall in performance would be carried forward to after the expiry of the guarantee period, with the premium being increased via the policy or premium review process.

However, with declining inflation, economic and political challenges and the increased exposure to global markets, nominal yields fell to levels where policies were no longer earning the returns required to sustain themselves for the full policy term.

After the completion of the guarantee period and at regular intervals thereafter, insurers were able to review the policies and recommend and, after the premium guarantee period, insist on premium or cover changes to ensure the policies remained sustainable and solvent. In practice few companies conducted these reviews until the investment account started decreasing or went negative. At this point the level of premium increase required to sustain the policy would be substantial. The paper models some situations based on actual returns earned over the period.

The third section questions whether actuaries involved with universal life business fulfilled their professional obligations. This is reviewed against the requirements of the current Code of Professional Conduct for Actuarial Society as well as some of the requirements of Guidance Note 1 of the Faculty of Actuaries and Institute of Actuaries that was the primary source of guidance at that time. Areas of focus are the transfer of risk from insurers to policyholders, the appropriateness of communication and the gap between what customers were sold and what they thought they were buying. These issues are considered from the perspectives of the individual actuaries as well as the profession and the insurance industry.

PART 1 – THE HISTORY OF UNIVERSAL LIFE BUSINESS

While biologists refute the concept that if a frog is put in tepid water which is slowly heated, it will be boiled alive before it jumps out, the analogy is still a valid one for describing situations where gradual change (especially adverse change) is not noticed until it is too late to do

anything about it. We believe that the analogy is appropriate to describe the development over the last 25 years or so of the challenge which the South African life industry is currently faced with in terms of the in-force book of universal life policies.

We first define what is meant by a universal life policy and then we trace its history.

It is policy with a selected sum assured and an arbitrary (within limits) premium and a fixed term, in which an investment (aka accumulation) account is maintained to which premiums and investment returns are credited and from which expense charges and the cost of life cover are debited in a monthly cycle.

At the end of the term of the policy the amount in the investment account is paid out as the maturity benefit. On death the sum assured is paid out. On surrender the investment account less any applicable "surrender charges" is paid out.

The amount that is debited for life cover is calculated as the product of the sum at risk and the risk premium for one month's cover applicable to the life assured's then current age. The sum at risk is the difference between the sum assured and the balance in the investment account.

The American roots

Its history starts in the USA. The following is closely based on an abstract from a Society of Actuaries (SOA) study note published in the SOA magazine.

The product was first developed in the United States in 1977 but was previously envisaged by GR Dinney in his presidential address to the Canadian Institute of Actuaries in 1971 where he warned of trouble on the horizon for the insurance industry due to the higher interest rates prevailing at the time brought about by rising inflation. He described a product called a "Universal Life Plan", being a product that combined term assurance and an accumulation fund.

In 1975 James CH Anderson, president of Tillinghast and Company, a US actuarial consulting firm, defined further "The Universal Life Insurance Policy" in speeches and articles. His basic theme was that social instability and change were making traditional whole life insurance irrelevant to the needs of large segments of the market. Members of an increasingly informed and cynical public were being pressed by inflation and mounting taxes into re-evaluating all their financial commitments, including their whole life insurance programmes. (Endowment assurance had never been sold in the USA due to the harsh tax treatment on surrender and maturity. Whole life policies were relatively lightly taxed or not at all either on death or on surrender.)

Frustrated by whole life insurance's clumsiness as a financial instrument and doubting its ability to keep pace with inflation, many traditional customers were opting for term assurance combined with an investment programme utilising some other savings medium. Thus, life assurance companies were in danger of losing their share of the investment dollar.

The first Universal Life policies were unsuccessful as they combined (packaged) term assurance with deferred annuities (equivalent to our Retirement Annuities) which created tax problems on death.

In 1978 Life Insurance Company of California (Lifecal) combined the two elements into a single contract called Total Life and later Complete Life.

Lifecal's vision was that the product should be targeted at wealthy individuals and to achieve that made a deal with EF Hutton, then one of the largest stock broker firms in the US, to get their thousands of brokers, who were in reality equities salesmen rather than the execution only traders that we are used to in South Africa, to sell the product to their well-off customers. Hence the initial product was designed to allow ad hoc premiums to be injected into the policy at any time with relatively low commission paid on an as and when basis on each premium. This made the charging structure very simple with no surrender penalties.

At that stage EF Hutton purchased Lifecal and renamed it EF Hutton Life

The product was not accepted, however, by the EF Hutton broker force, possibly because the way Universal Life distributed investment returns was via declared crediting rates based on corporate bonds and other fixed interest securities. The unit-linked version (variable universal life) would only be developed and sold much later.

The company had to switch to marketing the product through traditional life assurance agents whose client base could not afford the higher premiums initially targeted and for whom as-and-when commission was unacceptable. The flexibility of premiums which the basic design allowed was retained initially but to justify paying initial commission the concept of target recurring premiums was introduced, as opposed to contractual premiums. Up-front commission was based on these target premiums. But policyholders could pay less than that from time to time provided they made up for it later. Inevitably the actual premiums collected very often did not match the target premiums and eventually contractual premiums were introduced.

The product was successful and many other companies followed suit in offering it to dissatisfied traditional whole life policyholders who were persuaded to surrender their policies and inject the surrender value as a once-off single premium into the policy and continue paying the recurring premiums thereafter. The ability to earn another bout of initial commission made the product very popular with agents. In fact, companies encouraged agents to churn their own policies with new commission incentives to prevent other companies from doing so.

Eventually as par (with profits) dividends (bonuses) on traditional products started reflecting the higher prevailing interest rates the rate of replacement tapered off, but by then fixed premium universal life was included in the industry body's (LIMRA) statistics "traditional permanent insurance" category, as opposed to the "flexible-premium and variable UL" category.

To finally round off the history, EF Hutton Life was sold to Capital Life, which then suffered the second largest insurance company failure in the US in 1991 due to overborrowing. It was taken into administration by the Californian Insurance Commissioner and finally merged into Pacific Mutual Life in 2003.

South African experience

In the early 1980s the predominant types of products being sold were on the one hand conventional reversionary bonus policies and on the other the newer unit-linked ones. There were non-profit policies sold as well but these were mainly short duration term assurances, as long duration term assurance, non-profit endowment and whole life policies were prohibitively expensive due to the guarantees inherent in them.

Reversionary bonus (RB) policies were either endowments, with and without life cover, and whole of life. Retirement annuities were effectively endowments with annuity options at maturity. RB policies were either basic or reinforced. In reinforced policies the basic life cover provided by the policy was augmented by a decreasing sum assured term assurance rider (DTA). The reduction in life cover on the rider matched the projected increase in total death cover on the basic policy due to future bonus declarations. The latest declared bonus rates formed the basis of the projections. If future bonus declarations continued at the latest declared ones the total life cover would remain constant. If bonus rates increased, then the death benefit payable would become more than the initial amount and vice versa.

As future bonus rates were not guaranteed and the term assurance, although guaranteed, was reducing over time, this was the cheapest available form of life cover at the time.

The prime determinants of the premium rate were the basic premium assumptions, the latest declared bonus rates and the DTA assumptions. Even though bonus rates had been increasing steadily due to the improvement in investment returns accompanying inflation, these increases were not anticipated in determining the rundown rate in the decreasing term assurance rider. Doing so would have led to an even cheaper policy.

Whether this was due to some unspoken agreement between the large RB life offices or just due to in-built conservatism is unsure. Many, many years ago rate books used to illustrate future death benefits on RB policies using increasing bonus rates on the basis that bonus rates had been increasing in the past, but this practice subsequently stopped, perhaps after periods of declining bonus rates.

The reinforced whole life RB policy, with its maximum commission scale of 85% of first year premium and a third of that in the second year, was a very attractive proposition to both policyholder and sales agent whenever maximum life cover was sought for the lowest premium. Longer term changes in the need for life cover versus savings could be handled either by surrendering the policy at the time of cash needs or, if anticipated at an earlier stage, by converting the whole life policy to an endowment maturing in the near future, for a constant premium.

RB endowment policies were sold primarily for a combination of death cover and savings and the typical premium rate was approximately 1/n of the initial sum assured where n was the term. The reinforced version of the endowment was not that popular because if cheap life cover was the aim then a reinforced whole life policy was more suitable.

Unit-linked policies had become more and more popular with the attraction of unit trusts amid a near continuous bull run in equities, bought about by the constantly high inflation rate.

One of the most stable actuarial variables at the time was inflation at about 15%.

As unit-linked policies were essentially pure accumulation products with the death benefit being the account balance, life cover had to be incorporated by building in a decreasing term assurance element. The rate of decrease matched the "guaranteed" rate of growth in the death pay-out of the accumulation account. At that time unit-linked policies typically guaranteed a return of 4% per annum, not each year but on average over the life of the policy. The invested premiums accumulated at 4% was referred to as the guaranteed maturity value and the decreasing term assurance rundown was determined to produce a total death benefit equal to the guaranteed maturity value. There were "cheaper" versions also sold where the DTA rundown assumed a 6% growth in the unit-linked portion.

The term assurance premium was level and was deducted from the total premium. The balance of the premium was used to pay the policy fee and premium based expense charge and to fund the unit-linked account. There was no interaction between the DTA and the investment account.

As the premium for such an endowment policy more or less matched a non-reinforced RB endowment policy unit-linked products were sold essentially as savings products with some death benefit, with the marketing emphasis being on the superior growth prospects of unit-linked contracts versus RB one. For cheap life cover reinforced RB whole life reigned supreme.

Until Liberty Life launched their "whole life" version of Universal Life in about 1986.

But they were not the first to launch universal life (UL). (For purpose of clarity universal life products will be referred to as "UL" while "unit linked" refers to unit-linked policies.).

The development and marketing of the early versions of UL products in South Africa were driven by the assumed attractiveness of the flexibility in the product.

For a given recurring premium any level of life cover could be chosen between the statutory minimum of 8 times the annual premium and a maximum. And once the policy was in force the life cover level could be changed without any change in premium if it was less than the maximum allowed. It was envisaged that as a policyholder's needs for life cover versus savings changed the policy could be amended very simply as compared to the complicated changes that were required to achieve the same in current RB or unit-linked policies.

Since the basic mechanism allowed the sum assured to be changed very simply in any one month, reductions in sum assured were simple to implement. A reduction in sum assured would reduce the risk premiums going forwards and result in higher maturity and surrender values in the future.

Increases in sum assured could also be accommodated, subject to underwriting to prevent anti-selection and subject to the overall maximum sum assured limit for the given premium.

That maximum sum assured was set by means of doing simple projections using risk premium rates and investment growth rates that were assumed to be safe, in the sense that under most foreseeable scenarios, the investment account would not become negative at any stage.

This ability to change the use of the premiums from life cover to savings and back at

any stage and very simply, within one contract, was seen by the development actuaries and marketing departments, to be the essence of the new product.

The names given to these products bear that out. The Adaptor from Southern Life, MasterPlan from Anglo American Life (AAL), Flexidowment from Old Mutual, The One Plan from Sanlam. Product launches and advertising campaigns emphasised this flexibility.

Later versions of the product launched by Southern Life after its merger with AAL were named MasterAdaptor and Adapta Life.

Even the original US names reflected this: Universal Life, Total Life, Complete Life. You did not need any other policy.

Investment options for the investment account were taken straight over from the unit-linked range – market linked or the so-called smooth bonus option.

This contrasts with the US development where unit-linked versions were only sold later, in the form of variable UL.

To cater for the need to mitigate against the effects of the high inflation that was prevalent, options were included to increase the premiums either at fixed rates or at inflation as measured by the CPI. This was already standard with unit-linked policies at the time. The difference was that whereas with unit-linked policies the increase in the sum assured was determined by the premium rates for that policy type and current age of the life assured, for UL products, as there were no fixed premium rates per se, another method had to be devised.

Allowing the sum assured to increase at the same rate as the premium was not feasible as with risk premiums increasing with age this would lead to the investment account going negative long before the maturity date, other than for very low sums assured. So rough rules of thumb were developed by doing various projections. In most cases the sum assured was allowed to increase by about 60% of the premium increase rate. This was the average rate at which the sum assured increased in unit-linked endowments.

The actual combination of premium increase rates and sum assured increase rates was a choice that the policyholder could make. He could choose not to increase the life cover at all, which meant that more and more of the premium would go into savings. In most cases the maximum percentage was chosen as there was usually a need for the life cover to increase as well.

But this new product could not provide the cheap whole of life cover that was offered by the re-inforced RB whole life policies. The maximum sum assured was based on conservative projections, so the maximum sum assured was close to what basic RB endowments and unit-linked endowments offered. Hence it did not make a massive impact on the market initially.

This all changed when Liberty Life launched its version in 1985. It was called Universal Lifestyle.

Liberty had never developed RB products or even smooth bonus products, and even after it absorbed the South African book of Prudential Life, it closed that book to new business. In consequence it was never considered an option by brokers when cheap whole life cover was needed.

However, it saw in UL an opportunity to change that. And the kernel of the boiling frog problem was in the innovation that Liberty introduced.

As we saw from the analysis of the re-inforced RB whole life product, its relative cheapness derived from the fact that a large part of the life cover at later durations would be provided by the bonuses accruing at the assumed bonus rates. The expensive components of the guaranteed life cover in the form of the basic sum assured and the supplementary DTA became less and less important at the later durations where the probability of death became more significant.

Since future bonus rates are not guaranteed the total life cover is only guaranteed to the extent of the basic sum assured, already declared reversionary bonuses and the decreasing term assurance add-on.

If investment performance deteriorates or mortality or both, bonuses will reduce leading to a smooth reduction in the total life cover provided by the contract. The life office does not take all the risk. The policyholder bears some of it as well.

Liberty Life took this sharing of risk to a much greater level in its UL design.

Basically, it did three things differently from the other UL designs in the market.

Firstly, it extended the maturity date to the life assured's 90th birthday. As many lives assured were not expected to survive beyond age 90 that effectively made what was an endowment into a whole of life contract. If lives assured did attain the age of 90 and did not want to take the proceeds the policy could be extended beyond that with the benefit on death being the investment account without a sum assured minimum.

Secondly, it used aggressive projection assumptions to determine the maximum sum assured, so aggressive that the minimum premium for a given sum assured was significantly cheaper than for reinforced RB whole life contracts available in the market.

Thirdly, it transferred nearly all the risk onto the policyholder by only guaranteeing the sum assured for a limited period. At the end of the guaranteed period the policy would be reviewed and the premium could be increased by whatever amount the company deemed necessary to ensure that the policy remained sound throughout its lifetime. If the policyholder was unwilling to pay the increase in premium the sum assured was reduced to achieve the same effect. And the sum assured was only guaranteed for a further limited period and subject to further review then.

This transfer of risk was not only in terms of the investment component but also in terms of the mortality component. The monthly risk premium rates were not guaranteed. So, if the mortality experience of the company deteriorated the premiums could be increased not only to offset any poor actual and future assumed investment returns, but also poor actual and future assumed mortality rates.

This very innovative move sounded the death knell for reversionary bonus contracts throughout the industry.

Brokers and advisors rushed to Liberty Life to sell this incredibly attractive product that still paid maximum whole life commission. The benefits of the underlying unit-linked investments needed no pushing, as Liberty Life was seen as providing good investment

returns. They had never offered smooth bonus products and Lifestyle did not offer a smooth bonus investment option either.

Nobody, other than actuaries in the competing life offices, questioned how life cover could suddenly have become so much cheaper. Doubts about the limited guaranteed period were glossed over. The initial assumptions were not so aggressive that benefit illustrations, which at the time assumed 12% and 15% (net of management fees) growth rates, indicated any probability of the premiums being insufficient to maintain the life cover through to maturity.

In no time at all the other life offices copied them.

And then the only game in town was to outcompete each other on price.

Each subsequent product was cheaper than the existing ones. This was achieved by increasing the assumed investment returns and by shortening the guaranteed periods.

Eventually guaranteed periods were as short as five years.

For these later aggressively priced versions the benefit illustrations did indicate that at the lower 12% growth rate the investment fund would be depleted before maturity but this did not have much impact on their sales.

More conservatively priced versions continued to be marketed and sold as well.

While some companies did not introduce new names for the whole life versions, others did. Southern Life called its standard version Adaptalife and its cheaper version Economy Life. Others added a descriptor such as Old Mutual's Maximum Cover Flexi.

The other aspect was that the inherent flexibility in the product was no longer a marketing feature. As the investment account took a long time to get to a respectable level the savings aspect of the product did not feature at all. For that need, lower sum assured shorter term endowment versions were sold.

Then one company took the concept of risk shifting to another level.

While investment return assumptions became ever more optimistic the premium was always sufficient to maintain the life cover until the end of the contract on those assumptions, even though that was only guaranteed for a shorter period. This company jettisoned even that. It merely made sure the premiums were sufficient for the guaranteed period. That meant that it was a certainty that, unless investment returns exceeded all expectations, the premium would have to increase to keep the life cover in force. In effect it developed a non-guaranteed renewable term assurance contract that paid whole life commission. The other companies were also forced to add this variant to their product range.

These were the beginnings of the Boiling Frog problem.

PART 2 – UNIVERSAL LIFE PERFORMANCE

The water in the pot stayed tepid during the rest of the 1980s and into the 1990s. Inflation remained in double digits from 1984 to 1995 except for one year. But from 1991 there was a steady if jagged decline in inflation from a high of 16% to a low of 3% in 1999.

Interest rates followed suit. The 20-year government bond rate declined from 17.6% in 1989 to 7.1% in 2005.

Equities, while still reflecting good, if volatile, performance eventually followed suit as well.

The impact of these declines on the investment performance of the portfolios backing UL policies is reflected by the decline of the 5-year moving average of the average yields declared on Old Mutual's and Sanlam's smooth bonus funds from 19% in 1990 to 10% in 2002.

The water was slowly heating up.

The impact of the reduction in yields on the desirability of UL as a product began to be felt quickly.

The main reason for this was one of the essential elements of the product design.

Because the amount of life cover that is "purchased" every month is the difference between the fixed sum assured and the balance in the investment account, there is an inherent gearing in the interaction between the investment returns and the cost of purchasing the life cover. In good times if the investment return is greater than that assumed in the initial premium rate setting the investment account increases more quickly than assumed. This meant that the amount of life cover that is purchased is less than assumed. This then means a higher balance in the investment account that earns the higher interest return.

This virtuous circle was the reason that maturity values in benefit illustrations for endowment versions of the UL so far exceeded those of similar premium and sum assured unit-linked policies. In a unit-linked policy, with its fixed rundown pattern DTA, better investment returns increased the total life cover but only improved the investment account to the extent of the better investment return. There was no gearing effect in unit-linked contracts.

But in bad times this gearing results in a vicious circle. The lower investment returns cause the amount of life cover to be purchased to increase thereby reducing the investment account by more than under the original assumptions. Maturity values were going to end up much less than was illustrated at the point of sale. In the extreme the investment account could become depleted before maturity.

The following graph – Figure 1 – illustrates this sensitivity.

The policy, which is assumed to commence in 1995, is priced at 10% growth. Actual investment returns until 2019 are used and then assumed growth rates of 10% and 12% are assumed. The policy requires 12% per annum future returns to meet its original objective. At 10% per annum it bombs out at age 86.

This was not the case for traditional RB or unit-linked policies.

Maturity values would be less but not that much less as there was no gearing effect. There was no possibility of the policy maturing for nothing.

The realisation of the possibility of this downwards spiral led Old Mutual to design its first Flexidowment policy not as a pure true blood UL but as a hybrid.

While it adopted the monthly deduction of the cost of life cover at the then current age, the amount of life cover purchased was based on a fixed run-down pattern DTA akin to that used in unit-linked endowments. However, this meant that it did not benefit from the virtuous circle effect inherent in true UL which became obvious in benefit illustrations. At the higher projected investment returns their maturity values were very much lower than the

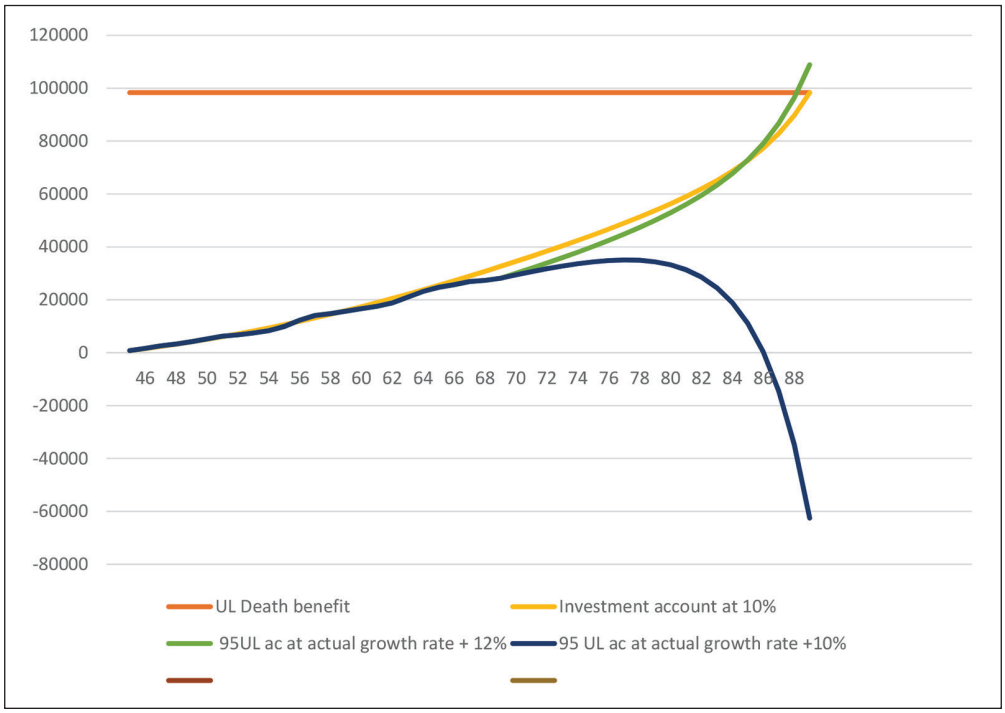

Figure 1 UL sensitivity to returns

competitors. They eventually modified the design and offered both versions, pointing out in their marketing literature the risks of the new version.

As it became clear that lower investment returns had become a permanent feature of the economic landscape, so the universal life concept itself started receiving a bad press. This was unjustified. Permanently lower investment returns should have led to re-pricing at lower assumed growth rates. Existing policies would not be able to deliver the illustrated maturity values, but the changing economic and financial environment could be blamed for that.

Instead in 2000 a new player on the block decided to utilise this disillusion to, again, completely revolutionise the market. The new mantra became "Do not mix savings and life cover in one policy".

With the launch of the new lapse supported whole of life policy, that did not build up an investment account or provide a surrender value, Discovery Life stole a march on the industry by bringing a new whole of life policy to the market that was even cheaper than the UL versions. It effectively killed off the Universal Life policy. But that is a story for another time.

While the other life offices scrambled to develop and launch their versions of the new whole life policy, attention drifted away from the consequences of the lower investments returns on existing UL policies. As investment returns settled down to a lower level, the immediate impact was almost unnoticeable.

The water keeps warming up in the pot, but the frog does not notice.

A 10% return was not that bad after all. And the investment account was still increasing due to premiums being credited, so there was not a reduction in the investment account, which might have been noticed.

But each year of lower investment returns had a double impact on the progress of the investment account. The lower investment return had a direct impact on the investment account through the lower crediting rate, but also an indirect one in that the amount of life cover that had to be purchased was higher due to the lower increase in the investment account.

There was a double whammy which caused the vicious circle as described earlier.

But this slow raising of the temperature was not noticed by the policyholder or his advisor. This was partly because policy reviews were not yet being performed on a regular basis for all policies. These only commenced from about 2010. Also, since benefit illustrations on existing policies were still being done on unrealistic assumptions the impact of a slightly lower investment account on the ultimate maturity values was hardly significant.

Shortly afterwards the Life Officers' Association (LOA – now ASISA) banned benefit illustrations for new business. A little while later they were also banned for existing business, which in the author's opinion was a bad decision, equivalent to removing the thermometer from the pot, even if that thermometer wasn't actually reading the correct temperature in any event.

To see why, we need to understand the communication the policyholder and his advisor received at the inception of the policy and annually thereafter.

At the point of sale and on inception the benefit illustration would have shown the initial sum assured and premium, illustrated maturity values for endowments and some future surrender values for whole life versions, at two different growth rates. But crucially no projection of the investment account.

The annual notice includes the current premium and sum assured, and the current balance in the investment account. After the banning of benefit illustrations there were no projected maturity values shown, although this was not consistently applied. Some offices continue to this day to show maturity values on some existing endowment policy types. But crucially not for the whole life versions.

So, a policyholder or his advisor has no idea as to whether the balance in the investment account is keeping up with where it should have been under the original benefit illustration, or not. And because of the initially slow effect of compound interest the difference between the actual investment account and the original projected one would have been barely noticeable in any event. It is only in the later years of significant increases in the risk premium rates that the two investment accounts start to diverge significantly.

If he compared this year's notice with last year's the investment account would have shown an increase. He has no idea how to judge that increase. He shrugs his shoulders and hopes for the best.

In the graph above the shortfall after 20 years is only 3% and even after 25 years is only 14%.

And the water kept warming up but without the thermometer the frog had no idea of the disaster ahead. And neither did the life offices even though their actuarial valuation systems which, after all, project all policies till maturity at realistic growth rates, should have indicated that policies were "bombing out" before maturity. But that bombing out did not seem to have raised any alarms if in fact it was even noticed.

Eventually the water reached boiling point, sooner for the very "cheap" versions of the policy, later for the more conservatively priced ones. Eventually even the life companies started noticing that investment accounts were starting to become depleted, and they realised they had better do something about it and remembered the policy review provision.

They started doing policy reviews on policies that were looking like they might be in trouble soon, (but tellingly only on policies that had passed their premium guaranteed dates) and sent out communication to policyholders, effectively demanding that the policyholder increase his premium by a certain percent to maintain the current level of life cover or to accept a reduction in life cover at the same premium level or a mixture of both, including increasing the rate of premium increases or reducing the rate of sum assured increases.

The increase was determined in most cases to maintain the policy in force for life in the case of whole life versions. The higher premium was only guaranteed until the next review date, usually five years hence.

In the case of the "term assurance" versions, the increase was determined to maintain the policy in force until the next premium guarantee date, effectively realising the renewable term assurance aspect of the design.

The investment assumptions used were the then realistic assumptions used for new premium rate setting or valuation.

Premium reviews were not done for endowment versions where their contract conditions did not allow the companies to increase premiums.

In many cases the increase in premium that was necessary to meet the above objectives would have required premium increases of over 100%, but most companies capped the increases to more affordable amounts of 25% or 30% with smaller additional annual increases thereafter for life. One company limited the increase to 5% per annum.

But in the case of the very aggressively priced term assurance versions, there have been much bigger increases. One of the authors has seen a policy where the new premium was three times the old one.

One can imagine the reactions of policyholders to this first-time notification that there was a problem with their policy, having been paying premiums diligently for 20 to 30 years. Many simply lapsed their policies, while others sought relief from the Ombudsman or the FSB (the predecessor of the FSCA). Appendix 1 reflects the nature of the complaints received by the Ombudsman.

The same vicious circle also affected the endowment versions but in most cases the depletion in the investment account was not sufficient to turn the account negative. However, the ultimate maturity values fell very short of the originally projected illustrative values. The same author has seen one case where the original illustrative value at 15% (which is what

most policyholders focus on) was about three times the sum assured, whereas 30 years later it matured for 20% of the sum assured.

The surprising thing in this case was that the response from the company to the complaining policyholder. "The product was not originally designed as a savings contract but as a low-cost protection policy. Any maturity value was a cherry on the top." It is difficult to reconcile this explanation with the original benefit illustration.

While many endowment versions have already matured and many have surrendered, there are still approximately one million UL policies in force, of which about 300 000 are whole of life versions with limited premium guarantee terms. Many of these have premium and sum assured increases. For many of these premium increases will be necessary in the future.

In some cases, even though policies have already passed their premium guarantee date and the investment accounts have already started reducing, because of the lack of adequate systems, reviews have still not commenced.

The bulk of the problem is still ahead. They are all sitting in the water without realising the steadily increasing temperature. The recent poor performance of equities has just increased the temperature significantly. And the double whammy effect gets worse as lives assured reach the ages beyond 55 when the risk premiums start increasing significantly.

The reluctance of companies to do reviews prior to the premium guarantee date is difficult to comprehend. The argument is that there was no point in doing the review before as the company would not have been able to enforce an increase. What this ignores is that a review is the only effective way the policyholder could have been informed about how his policy was doing. The review was the only effective thermometer available to warn him of the increasing temperature of the water.

While policyholders might have ignored the outcome of the review, which they had the contractual right to do, they would at least have been given an indication of the likely increase with which they would be faced if they did nothing in the meantime. A meaningful policyholder communication protocol would have required that. The intention of the LOA benefit agreement was that. It just had not been amended to cater for the new technical issue introduced with the development of UL. Today TCF does require that.

More importantly they could have immediately taken action to ameliorate the consequences of the lower investment conditions, by exercising any or all the options they had, such as immediately increasing the premium, increasing the rate of premium increase, reducing the sum assured or reducing the rate of sum assured increase. Any one of those would have meant a less severe situation at the end of the premium guarantee period.

In the authors' opinion the review is an inherent aspect of the policy, based on the lack of investment guarantees and the inherent "danger" of the gearing aspect of the contract, that is there not only for the benefit of the life company, but also for the benefit of the policyholder, regardless of the actual legal wording of the review clause.

TCF outcome number 5 requires a policy to "behave as the policyholder would expect it to" to paraphrase the actual outcome. If a policyholder has had no meaningfully

understandable communication whatsoever since inception to alert him to the fact that his investment account has not been growing at the required pace, he would certainly not expect to be suddenly faced 25 to 30 years later with a very large increase. That can surely not be considered to be meeting Outcome 5.

The situation is analogous to a pensions actuary, who determines the initial contribution rate required for a defined benefit arrangement, and then disappears for 25 years without communicating at all with the sponsor in the meantime, and on his first review of the fund enforces a massive increase in the contribution rate because of inadequate investment returns. Had he done the professional thing and communicated the likely outcome of the poor investment performance earlier the sponsor could have managed his affairs better, by for example reducing the benefits or increasing contributions in line with salary negotiations.

Unless something is done soon there will tens of thousands of policyholders facing demands for steep increases in premiums to maintain their current levels of life cover, without any pre-warning.

What should the life offices have done then and what should they do now?

The solution was and still is very simple. They should have done reviews as soon as it became evident that lower investment returns had become the new reality. As mentioned above they should have and should still use the fundamental feature of the original design of the policy – the flexibility and adaptability of the combination of life cover and premium – to allow policyholders to manage the temperature of the water themselves. The original attractiveness of the UL design is still in place.

To achieve that, policyholders should be given enough meaningful information, as soon as possible, to see the effects of maintaining the current and future combination of life cover and premiums and encourage them to make any changes to those that they consider suitable. The policy conditions allow them to do so in any event.

As mentioned above they can reduce the current level of life cover to any level, even zero. They can reduce the rate at which life cover increases. They can do both. They can increase the level of premiums at a higher rate than they are currently doing.

Small increases now are usually more manageable than a very large increase in the future. Because of the gearing effect a smaller percentage reduction in sum assured is needed to "save" the policy than the necessary percentage increase in premium. While "actuarially" equivalent smaller percentage reductions in sum assured will be more acceptable to policyholders.

Do this immediately and annually. Do not wait until the next premium guarantee date. A review is not the same as a premium increase.

Re-introduce annual benefit illustrations. These must be on very realistic assumptions, say currently 8% and 3%. Remember these are after tax and management fees.

Illustrate future maturity values for endowments or at what stage the investment account becomes negative. For whole life versions indicate the level of premium increase that is likely at the next premium guarantee date.

Tell the policyholder that those results are not cast in stone. Illustrate a few very basic alterations that the policyholder could choose to make, such as to stop life cover increases or reduce life cover by say 20% and show the impact of those on future values.

Make clear that their advisors can show them more options and give the advisors the necessary programs or apps to do so.

Then implement an e-mail and/or SMS and/or call centre "call back" programme, first to the advisors and then to the policyholders themselves using call centre authorised advisors to "force" them to make a conscious decision, even if that is to do nothing".

Where absolutely no contact can be made, then there is nothing that can be done other than wait until the next premium guarantee date and then increase the premium accordingly. Should the policyholder then make contact, offer him alternatives.

This should also be done for endowments. It is a mistake to assume that the default choice for policyholders who chose a reasonable balance between life cover and savings at the outset is to maintain that life cover at the expense of the savings objective. The fact that the life company may not be able to enforce a sum assured reduction or premium increase does not mean that a review is meaningless. Most policyholders probably do not realise that they can influence the outcome themselves. They just need to be told that they can.

Only then can you claim to be the customer's best friend.

Save the frog!

What have life companies done?

The following is based on information provided by the four large life companies, that have the most UL policies on their books.

It is clear that they are all taking the issue with UL very seriously. The philosophy on premium reviews is not uniform but they are all attempting to address it constructively. One office has by and large implemented all the suggestions made above.

In only one case are reviews being done before the premium guarantee date, and as mentioned above, one company has not started the reviews yet on a significant book of policies.

In all cases there is a limit to the increase in premium of some sort or other.

In one case, premium increases are limited to 5% per annum for life and the shortfall between the actual premium and the theoretical premium is subsidised by the company.

In another case the initial increase is limited to 25% but thereafter the premium increases by a further 5% each year. The policyholder in this case is offered an alternative of age-related premium increases each year capped at 14% in any one year.

In another case there is an initial premium increase of 14% with further annual increases of 6%.

In most cases, unless future assumptions change for the worse the premium is guaranteed not to increase beyond the amounts mentioned.

In another case where the policy was essentially a term assurance contract, the premium increase cap is only until the next premium guarantee date, and the increase in premium is only sufficient to maintain the investment account until the next premium guarantee date.

In some cases, if the investment account is projected to deplete only after age 85 the premium is not increased at all.

In one case, if a policy is projected to run out of funds within five years, the office does not increase the premium on the existing policy but replaces the policy with a new style whole of life policy.

In another case while the premium is increased on the existing policy the increased premium is limited to the premium that would be payable on a new style whole of life.

For policies that have the existing premium and/or life cover increasing, the options take these into account, for example, increasing the premium by an additional 5% per annum on top of the existing premium increase rate.

The extent of the premium increase required is a function of three basic factors. The pricing assumptions of the original policies, the timing of the review, and the actual historic investment performance.

The more aggressive the original pricing, the poorer the historic investment performance, and the later the review is undertaken, the more severe the theoretical increase is.

In some cases the original pricing basis required 12% or even 15% future investment returns. Investment performance has varied by as much as 2% per annum on average for 25 years between different portfolios, and in some cases reviews have not been done yet on lives assured that are 75 years old on policies taken out in 1999, more than 10 years after the first scheduled review and a few years after the premium guarantee period.

The worse these factors are the more "generous" is the subsidy inherent in the cap on premium increases.

In terms of communication, all the offices appear to be doing a reasonable job of contacting the policyholders and their advisors and explaining the various outcomes and options as best they can. But understandably they have been receiving a lot of resistance.

PART 3 – PROFESSIONALISM CONSIDERATIONS

How much has unacceptable or unprofessional conduct by the actuarial profession contributed to the heating of the water and the frog's apparent ignorance of or indifference to its fate?

The actuarial profession has been built on a foundation of trustworthiness, competence and relevance. During 2012 this was codified as the Code of Professional Conduct[1] (the Code). All work and contributions to broader society from members of the Actuarial Society should be judged against the standards of this Code.

This section of the paper reviews the requirements of the Code, then evaluates actuarial work on universal life business against this Code.

Requirements for professional conduct

The Code is a principles-based approach to self-regulation which makes it both less specific and more enduring. It does not highlight specific actions or conduct that are seen to be

1 Code of Professional Conduct published by the Actuarial Society of South Africa in 2012

consistent with the expected standards of professional conduct but lays down broader principles. This leaves it open to interpretation but does allow it to remain applicable as actuarial practice evolves.

The essence of the Code is set out in the opening paragraph that summarises Professional Promise which states:

> Members are expected to render quality service to their clients through
> a. The application of specialist and up to date actuarial knowledge and expertise,
> b. The demonstration of ethical behaviour, especially when doing actuarial work, and
> c. The member's accountability to the Society for professional oversight.[2]

In providing further explanation in the following paragraph the Code highlights the need for actuaries to only perform functions for which they are competent and appropriately experienced but to do so in a manner that is ethically sound and supports the public good.

> Members are expected to conduct themselves
>
> honestly, with integrity, competence and due care, and in a manner that fulfills the profession's responsibility to the public.[3] Members need to ensure that they "shall do nothing that brings the actuarial profession into disrepute.[4]

In discussing knowledge and expertise, the Code does not restrict its focus to technical competence and experience but also refers to communication.

> Effective and appropriate communication is an essential part of all actuarial services. The form and content of any actuarial communication must meet the need of the particular circumstances, taking into account the knowledge and understanding of any potential users and the relationship of the actuary to such users.[5]

Paragraphs 9 to 22 focus on values and ethical behaviour. This section covers a wide range of issues including whistleblowing and conflicts of interest that are not directly relevant to this subject. However, the following two are core considerations:

> — A member must act honestly, with integrity, competence and due care, and in a manner that fulfils the profession's responsibility to the public and upholds the reputation of the actuarial profession.[6]
> — In fulfilling assignments, members must consider the likely implications of their recommendations for all parties that are likely to be materially affected, and also draw the attention of their clients to such implications.[7]

2 Code of Professional Conduct, paragraph 1
3 Code of Professional Conduct, paragraph 2b
4 Code of Professional Conduct, paragraph 2c
5 Code of Professional Conduct, paragraph 8
6 Code of Professional Conduct, paragraph 9
7 Code of Professional Conduct, paragraph 11

The final eight paragraphs consider the actuary's responsibility regarding professional accountability. These paragraphs consider both the accountability of ASSA to providing education and professional development, as well as guidance and standards, and the accountability of members to act in the public interest.

Acting in the public interest is open to a wide range of interpretations. The Code provides some bounds to this in paragraph 24, where it states:

> Members are encouraged to consider the public interest when rendering actuarial services, but provided that members meet the requirements of the applicable Law, the Constitution of the Society and any applicable Standards of Practice and the Code, they will be deemed by the Actuarial Society to have met the expectations of the profession with respect to the public interest.

In short these requirements require members of the Society to ensure they are competent, appropriately transparent, that they fulfil their obligations to the public as well as their clients, and to create awareness of the implications of their recommendations to all parties that are likely to be materially affected.

Universal life – areas for consideration from a professionalism perspective

When considering universal life products and the market for these products, professionalism issues arise in the following areas:

— Transferring of risk from the insurer to the customer
— Value sold to the customer versus the value bought by the customer
— Benefit illustrations and transparency

These will be reviewed against the following considerations from the Code:

— Appropriateness of communication for all users of the service.
— Competence and due care in a manner that fulfils the profession's responsibility to the public and upholds the reputation of the actuarial profession.
— Implications of recommendations for all parties that are likely to be materially affected and to also draw attention of their clients to such implications

For this paper it is assumed that the actions of insurers reflect the actions of actuaries working for or consulting to insurers. The professionalism focus is limited to the communication and public interest requirements of the Code and not the technical competence of any individuals involved.

The underlying issues – transferring of risk from the insurer to the customer

Before the advent of universal life, insurance companies carried the downside mortality and investment risk with policyholders benefiting through the usually opaque and complex bonus rules if they had with-profit policies. However, where policies under-performed to such an extent that bonuses could not be declared, the insurer incurred the additional cost. While the

universal life structure offered more direct and transparent participation, the pure universal life structure offered no downside protection[8] in the event of market under-performance. The leveraged impact of investment under-performance on the sum at risk and mortality charges further increased the risk to the customer. Under the previous reversionary bonus product structure, insurers could remove non-vested bonuses but not erode the basic sum assured.

With the advent of universal life business, policy structures became more transparent with policy holders being credited with the actual investment returns[9] and incurring the then current mortality and other charges on the policy. Should investment markets under-perform, the retrospective accumulation of reserves (accumulated funds) could lead to negative fund values. In the absence of premium guarantees, this could be passed to the customer in the form of premium increases.

The underlying issues – value sold to the customer versus value bought by the customer

In short, did the customers buy what they thought they were buying and were the products fairly represented by intermediaries and other advisors.

This relates to whether customers would have reasonably understood what they were purchasing and whether actuaries can be held professionally accountable for any misunderstanding.

Policies sold for needs ranging from estate planning and provision for young families to business assurance were marketed as whole life policies. These ranged from aggressively priced products with short guarantee periods to more conservatively priced ones with higher premiums and longer guarantee periods.

The guarantee period represents the period during which the insurer may not increase premiums or reduce cover regardless of the performance during the period. Any performance shortfall is effectively capitalised for the customer to make up the shortfall on completion of the guarantee period.

As the policies were all marketed as whole life policies, they would have attracted maximum commission though few had guarantee periods close to the full commission term.

In policy documentation some companies committed to conducting regular reviews of the policy after completion of the guarantee period. Many companies conducted no such reviews while the policies remained solvent and the premiums continued to cover the risk charges.

The underlying issues – benefit illustrations and transparency

The LOA[10] projections were primarily focused on investment policies with the objective of limiting benefit projections during a time of very high nominal yields. The focus of the projections was on maturity and surrender values for endowment and pure endowment policies.

8 Some smoothed bonus portfolios offered guaranteed minimum bonuses with less direct market participation
9 Smoothed if invested in a smooth bonus portfolio
10 Life Office Association – one of the organisations that merged to form ASISA

As nominal yields reduced, the gap between the projected yields at policy inception and the realised maturity and surrender values widened, leading to a loss of credibility of the projections and the eventual demise of the LOA agreement.

The benefit illustrations were of limited value to high cover policies as they gave no indication as to whether the policy values were sufficient to sustain the level of life cover. As yields reduced and pricing of risk policies became more aggressively competitive, customers were usually unaware that their current premiums would be insufficient to sustain their policies and were given no indication of the extent of premium increases that might be required in the future.

There was no disclosure requirement for the yield required to sustain the current premium on the policy or the premium increases at different durations and on the expiry of the guarantee period should different yields be obtained over the term of the policy.

Had such projections been provided, premium or cover adjustments following a review could have been communicated to policyholders as they would have complied with the policyholder's reasonable benefit expectations.

The essence of the case

Actuaries held key roles in the development and management of universal life in South African life assurers and held a wide range of specialist, management, leadership and non-executive roles within the insurers. The Appointed Actuary, in accordance with what was then Guidance Note 1 (GN1),[11] had wide ranging oversight responsibilities covering everything from the adequacy of reserving to the communication to customers.

The case against actuaries would contend that:

1. Communication with customers focused on the potential increased returns without consideration for the consequences of investment under-performance to the customers.
2. No clear explanation of the transfer of risk from the insurer to the insured.
3. The term guarantee would be misleading to individual customers who would view a guarantee in the same way as a manufacturer's guarantee on material goods – if it under-performs, the manufacturer repairs it for their own account.
4. The drive towards more aggressive pricing of risk policies with shortening guarantee periods focused on the initial premiums and not the potential increase in premiums in later years.
5. More aggressively priced policies with shorter guarantee periods were inappropriately labelled as whole life policies to maximise commission without regard for the impact on the customer.
6. Statutory actuaries were either insufficiently aware of the severity of the potential impact of low yields on policyholders or took insufficient steps to raise awareness of this sensitivity.

11 Guidance from both the Institute and Actuaries and the Faculty of Actuaries (which were separate entities at that time and all South African actuaries belonged to one or both bodies). Version 3 released July 1992

Aggravating factors

At the time universal life policies were introduced the market consisted of conventional with- and without profit policies as well as unit-linked savings policies. Non-profit policies were fully guaranteed with the premiums, sums assured and maturity values established at inception. With-profit had guaranteed minimum values but the opportunity to benefit from superior performance through an opaque combination of reversionary and terminal bonuses (usually declared as a percentage of the sum assured). While the underlying product structures were opaque and the bonus formulae complex, customers had a clear understanding of what their premiums were for the duration of the policy and the minimum (with profit) or guaranteed (without profit) benefits would be. Marketing for universal life was largely driven around the allure of flexibility, the apparent attractiveness of the bonuses/ returns which were no longer expressed as a percentage of the sum assured but as a percentage of the accumulated premiums less expense and risk charges. The removal of minimum full policy term premium and cover guarantees and the overall transfer of risk to the customer did not receive the same exposure.

The knowledge gap between actuaries representing the insurer, customers and intermediaries was more significant than it is now as information was less available with the intermediaries and customers relying almost exclusively on information provided by the insurers.

Although policy conditions spoke of regular reviews of the policy, in practice few, if any, companies established systems and processes to implement these systematically on their policies on a regular basis. The net effect of this is that under-performance of the policy is not regularly tracked. Therefore, remedial action in the form of increased premiums or reduced cover is not recommended until the policy ceases to be viable.

As the universal life structure was used to provide aggressively priced whole life cover policies, policyholders were not made aware of the impact of low investment returns on their policies. It is unclear whether this was due to insufficient awareness of this risk to policyholders or to insufficient communication with real and potential policyholders.

Two key aggravating factors can be summarised in two words – guarantees and commission.

— **Guarantees** in that to make initial premiums lower, insurers offered shorter guarantees. However, this was a misuse of most rational customers' understanding of the term 'guarantee'. Instead of the guarantee ensuring that the insurer would make good any performance shortfall during the guarantee period, it meant that the insurer would defer the requirement for the customer to make good the shortfall – thus increasing future payments from the customer.

— **Commission** in that aggressively priced policies with short guarantee terms continued to be labelled as whole life so that full term commission could be paid. Actuaries would have been aware that, unless there was a sustained generation-long bull market, the premium would not sustain the policy for the full term yet still acquiesced to the payment of whole life term commission.

Mitigating factors

Mitigation factors concern primarily the broader market and the tools and technology available to actuaries at the time.

The market was, as now, highly competitive with a small number of large companies dominating the market. Distribution of these policies was almost exclusively by independent brokers and employed agents. Maximum commission scales for brokers were regulated and companies needed to show equivalence of reward for their employed agents. Maximum commission was expressed as a percentage of premium multiplied by the term of the policy with a ceiling that set the maximum commission term to twenty-seven years. It was widely understood in the industry that the market for risk cover was price and commission term driven. With-risk cover being regarded as a commodity, these conflicting market forces encouraged the sale of more aggressively priced products that paid full whole life commission. It is unlikely that any single company, even one of the bigger ones, acting in isolation could have changed the market dynamics. While this might be a mitigating factor for actuaries working for individual companies, it can be argued that it is an aggravating factor for the profession – and particularly the Life Insurance Committee and Council.

Over much of the time the country had endured a prolonged period of high nominal yields and high inflation. At the start of the universal life era, it is possible that actuaries would have thought that by restricting nominal yield projections to 12% or 15% was conservative.

The hard- and software to do the risk analyses and projections that are now routinely used were not available to actuaries. Computers were slow with limited capacity for multiple iterations on multiple scenarios.

Actions taken by some companies to limit the consequences for individual customers and transfer the cost of under-performance in part to the companies is evidence of possible awareness of professionally questionable past practices as well as the commercial and reputational risks of the full cost being passed on the customers.

Answering the case

Considering specific points against actuaries:

1. *Communication with customers focused on the potential increased returns without consideration for the consequences of investment under-performance to the customers.*

   The move to universal life focused on flexibility and the ability to share in superior performance more directly. This focused on the greed of customers without due regard to the potential consequences of under-performance. In mitigation it can be argued, that after sustained nominal bull markets, a sustained period of single digit performance was not anticipated or tested for. This mitigation would not have applied as readily towards the end of the universal life generation as nominal yields fell and there was no focus on warning customers of the risks of downside performance.
2. *No clear explanation of the transfer of risk from the insurer to the insured.*

   This was not emphasised in marketing material to customers as the focus remained on

the upside potential and not the increased risk. As insurance policies with life cover are primarily a vehicle for mitigating risk, this can be seen to be a significant oversight.

3. *The term 'guarantee' would be misleading to individual customers who would view a guarantee in the same way as a manufacturer's guarantee on material goods – if it underperforms, the manufacturer repairs it for their own account*.
   Here there were clear deviations in understanding with insurers understanding guarantees to be the period during which there will be no mandatory premium increases thus giving comfort to customers with short-term cover needs. However, for customers looking for stability over a longer period this was not understood. Even policy conditions specifying possible premium increases following future reviews were not fully clear that this would include the accumulated shortfall during the guarantee period.
4. *The drive towards more aggressive pricing of risk policies with shortening guarantee periods focused on the initial premiums and not the potential increase in premiums in later years*.
   This flows on from the guarantee point above and refers specifically to the period when sustained high nominal yields were no longer routinely expected. This was an opportunity for the industry to see that allowing continued low prices in the expectation of high nominal investment yields was not sustainable. However, with life assurance risk cover seen as a price sensitive commodity, pressure for retaining low prices led to the shortening of guarantee periods while retaining the whole life policy veneer. This presented an opportunity, that appears to have been missed, for actuaries to clearly explain why the continued low prices were not sustainable for whole life policies. However, it would have been difficult for any company acting alone to go against the industry by changing its pricing and commission practices, when access to the customer was controlled by the intermediary.
5. *More aggressively priced policies with shorter guarantee periods were inappropriately labelled as whole life policies to maximise commission without regard for the impact on the customer*.
   In many respects this is at the core of the matter. The pressure to retain commission rates led to shortening guarantee terms while retaining the whole life labelling. Without this commission pressure, products could have been marketed as term assurance policies with guaranteed insurability benefits. These would have paid lower commission (thus reducing premiums) and customers would have expected age-related premium increases on each policy renewal. However, one company operating in isolation would have been unlikely to have implemented this successfully as the product would rarely have been presented as an option to customers.

   It can be argued that by being either complicit in, or actively leading, plans to drive down premiums to levels that were unlikely to be sustainable for a full life term while retaining the whole life labelling, actuaries fell short of the requirements of the Code discussed above. There was clearly not appropriate communication for all users, the impact on all those materially affected does not appear to have been highlighted, and it does not appear that due care was taken to ensure the long term reputation and standing of the profession.

As this was a systemic and broader market conduct issue, it can be argued that attention should be focused less on individual actuaries in various roles in insurers and more on the leadership of the profession – especially the Life Insurance Committee at the time – who may have been in a position to influence the profession and the industry.

6. *Statutory actuaries were either insufficiently aware of the severity of the potential impact of low yields on policyholders or took insufficient steps to raise awareness of this sensitivity.* While sensitivity analyses have been done to show the impact of variations in experience on shareholders and the solvency of the insurer, little regard appears to have been paid to the impact on customers of lower yields and increases in risk charges.

   In modelling the effect on shareholders, explicit and implicit mitigating actions would show reduced benefits and increased charges for policyholders. No similar work appears to have been done or publicly reported showing the potential impact on policyholders – especially the risks that highly reinforced whole life policies will not remain solvent.

   While there may not have been an explicit reporting requirement for actuaries to conduct such analyses from the perspective of the policyholder, actuaries bound by the Code and GN1 have a professional duty to see the impact of factors on all parties impacted by the actuary's work (the Code) and that policyholder reasonable benefit expectations are met (GN1).

In terms of GN1, "It is incumbent on all Appointed Actuaries to ensure, so far as it is within their authority, that the long term business of the company is operated on sound financial lines and with regard to its policyholders' reasonable expectations."[12] This simple statement summarises the competence and trustworthiness expected from the profession – that we would ensure businesses remain sound and customer value is preserved in line with reasonable expectations. However, with the transfer of risk from insurers to customers and aggressive use of "guarantee periods", there was a clear shift in balance between the customer and the company. The emphasis changed from ensuring reasonable policyholder expectations were met to ensuring the financial stability of the insurer. Against this standard it can be argued that the Appointed Actuaries fell short of the standard of having regard for customer expectations while those with responsibility for product development actively worked to redefine the product offering away from what would have been seen as reasonable customer expectations.

12 Guidance Note 1, Faculty of Actuaries and Institute of Actuaries, V3, 1992, paragraph 1.1

CONCLUSION

There are few things as easy as judging the apparent failings of another era using the standards and knowledge of the current era. However, despite the presentation of facts and arguments above, that is not the objective of this section of the paper.

The first objective is to create awareness of the risk of actuaries of being caught in the hype of the raging commercial battle without considering the long-term implications for the profession, the industry and the ultimate customers.

The second objective is to raise awareness of the risks of a herd mentality where most market participants follow a course of action that actuaries, as professionals working in the industry, should be able to see potential adverse consequences and work to create awareness. This, however, is not a risk confined to the narrow constraints of the industry practices described in this paper, but a more universal risk to the profession – especially in times of rapidly changing market conditions and industry practices. The importance of questioning the established practices and assumptions from the perspective of all parties impacted by the actuarial work cannot be over emphasised.

With the establishment of the Market Conduct Committee, the formalisation of the Code of Professional Conduct and the leadership of the Professional Matters Boards, individual actuaries are now better positioned to create awareness of inherent risks that may have future negative consequences for the profession and the customers who are impacted by their work.

As we return to our frog in the simmering water, actuaries have contributed to both the management of the temperature and the soporific mood music. Steps taken to reduce the impact for individual customers represent the first steps taken by members of the profession to lower the temperature of the water and provide an escape option for the frog.

APPENDIX

Issue Number 44 – August 2020

Ombuzz

PREMIUM REVIEWS

Policyholders, often nearing or in retirement, are having to face steep premium increases (or a drastic drop in their cover amount) as a result of premium reviews to their universal life policies, giving rise to an increasing number of complaints to our office. When we investigated recently we identified at least 68 complaints received since 2018 about this issue.

Universal life policies

Universal life policies are policies which were popular in South Africa from the mid 1980s to early 2000s, where the product design is characterised by both risk and savings elements in a policy.

The savings element operates by means of an investment/accumulation account for each policy in a similar way as a bank account operates. Premiums, after the deduction of premium charges, are credited to the account and administration charges and the cost of cover are deducted on a monthly basis. The balance earns growth in the underlying investment portfolio, credited to the account on a monthly basis.

The cost of life cover increases annually as the life assured grows older and is based on the amount at risk, which is the difference between the cover amount and the balance in the investment account. The idea was that in the earlier years of the policy, when the cost of life cover was less than the nett premium, the investment account would grow, decreasing the amount at risk, and that this growth in the earlier years will subsidise the increasing cost of life cover as the policyholder ages. The extent to which this materialises depends on the actual investment returns achieved relative to the assumptions made at inception.

For a given premium, the policyholder could elect the level of risk cover ranging from minimum risk cover, where virtually the entire premium is applied to the savings element, to maximum risk cover, where the risk component is as large as possible, with a very small savings component. In order for the policy to maintain the level of initial cover, the investment account would need to grow at an assumed rate. The higher the level of cover for the given premium, the more optimistic the assumed rate was, with maximum cover commonly assuming a growth rate of 15% per annum, the maximum illustrative projection rate allowed by the then Life Offices Association ("LOA"). If the actual investment returns exceeded the assumed projection rate resulting in the balance in the investment account exceeding the cover amount (break-out point), no more cover would need to be purchased and the investment account growth would be accelerated. If the investment growth rate was lower than projected, the converse could happen, with the investment account becoming depleted, that is, reaching break-down point.

Policyholders could choose a guaranteed cover term, which guaranteed the cover for the selected term at the set premium (level or increasing). The guaranteed cover term is the period the policy can sustain the cost of providing the level of cover. The longer the guaranteed cover term, the higher the premium is set when the policy is purchased, so the guaranteed cover term was mostly shorter than the term of the policy, e.g. 10 or 15 years.

To mitigate the possible loss of cover after the guaranteed cover term, universal life policies commonly made provision for policy reviews, in terms of which the insurer would monitor the performance of the policies, and advise the policyholder should it appear that the investment account balance and future premiums would be insufficient to maintain the cover beyond the guaranteed cover term. However, the review clauses differ from insurer to insurer and are often vague as to what constitutes the review, when the review will be performed and what steps would follow. It appears that insurers did not necessarily institute policy reviews for a very long time, some state they did but decided to take no action, and in some cases subsidised the cost of cover for a while before implementing premium increases (see Case 3).

From the time these policies were sold, the inflation rate has declined considerably over the years, leading to lower investment returns than assumed. In addition, there were prolonged adverse economic conditions, especially after 2008, which compounded the problem of poor investment returns. The lower the investment account balance, the greater the amount at risk and the amount of the cover charge. This is further exacerbated by the increase in the cover charge with ageing, and the snowballing effect leads to a rapid depletion of the investment account. Insurers were jolted into action and policyholders were taken by surprise, especially by the level of increased premium required, leading many to question the insurer's right to increase premiums and the timeliness of the review. Needless to say, the maximum cover policies require the steepest increases, as can be seen from the following example:

Case 1

1. The Maximum Cover whole life policy commenced on 1 January 1998. The policyholder was 38 years old at the time.
2. The life cover (sum assured) was R1 000 157 for a fixed premium of R316 p.m., which included the cost of a rider benefit of R31,50 p.m.
3. The policyholder selected a guaranteed cover term of 16 years, which therefore expired on 1 January 2014.
4. In September 2019 the insurer wrote to the policyholder to inform him that his policy had been reviewed and that the current premium was insufficient to sustain the life cover. It was projected that the investment account would be depleted by 2024. As the insurer had closed its universal life product range to new business, including an increase in premiums, it offered the policyholder another pure risk product that could provide the same cover.

5. The premium for the new policy, with the same amount of life cover of R1 000 157 and a new guaranteed term of 10 years, increased from R316 to R1 231 p.m., an increase of almost 400%. The insurer explained that the increased premium under the universal life policy would have been similar as the premium would have been based on the insured's current age, and without the benefit of a substantial balance in the investment account.

The following case illustrates a variation of the universal life policy, where the policy is designed so that the premium would only be sufficient to sustain the level of life cover until the end of the guaranteed cover term.

Case 2

1. The whole life policy forms part of a capital protection package, together with a compulsory life annuity purchased with the proceeds from a pension fund at retirement age of 62 years.
2. The policy commenced on 1 October 2004, with life cover equal to the annuity purchase amount of R277 730.
3. The premium of R300.52 p.m. is deducted from the gross annuity of R2 156,67, leaving the annuitant with an income of R1 856.15 p.m.
4. The guaranteed cover term is 16 years, thus ending on 1 October 2020.
5. The annuitant has been informed by the insurer that in order to retain the level of cover on the whole life policy, the premium would have to increase to R588.22 as from 1 October 2020, meaning his net income would drop to R1 568.45. Alternatively, the life cover would drop to R137 110. Other alternatives are also provided, such as annual increases to the premium or annual decreases to the sum assured, as well as a further option of cancellation of the whole life policy in exchange for a joint and survivorship annuity.
6. The insurer explained the reason for the increase as follows: "The contract was initially priced for a specific term resulting in the initial premium being lower in the specific guaranteed term (i.e. 16 years) and will be renewed for the next specific term without underwriting being applied but incorporating other factors, such as the life insured's age etc. The higher increase in premium is purely because of the change in the risk due to higher mortality factors caused by the change in the life insured's age. A review is necessary to accommodate the increasing cost of the risk cover on the Whole Life policy. Without this premium review mechanism, the premiums would have been much higher at the start date."

The following is a synopsis of a case that demonstrates a drastic effect in Rand terms of the premium review.

Case 3
Date of commencement – 1994
Life Cover at commencement – R300 000; Lifeline benefit – R200 000
Premium – R125 per month
Contractual Cover increase – 15% per annum
Contractual Premium increase – 15% per annum
Premium review provision – every 5 years
First actual premium review – February 2019
Cover prior to review – R9 875 642
Premium prior to review – R4 114
Investment fund at review date – a negative balance of R770 168
Premium after review – R25 206
Next premium review – 2024
Commission payable on the review increase – R123 389 in the first year and R41 129 in the second year

ONE OF THE ALTERNATIVE OPTIONS
Premium – remains at R4 114
Cover – reduces from R9 875 642 to R1 612 032
Continuing premium and cover increases at 15%.

NOTE
What is unusual about this case is that both the premium and the cover amount increases at 15%.

The insurer subsidised the premiums for many years prior to the first premium review.

The following case study demonstrates the effect of the premiums being priced for term only, requiring a regular premium increase.

Case 4
The complaint involves three similar whole life policies taken on the life of the same person at the same time with the same insurer – treble the trouble.

At the time the policies commenced in July 2000, the life assured was 54 years old.

All three policies had a guarantee period of 10 years and provided for scheduled premium increases of 10% per annum with concomitant cover increases of 6% per annum.

At the end of the initial guarantee period in 2010, the policies were reviewed and the option to retain the cover and increase the premium with a new 5-year guarantee period was taken up on all the policies. The other option was to reduce the cover and retain the premium. Note that with both options, the scheduled annual increases in premium and cover continued.

At the expiry of the guarantee period in 2015 the policies were again reviewed, and two similar options as before were given, and again the option of increasing the premium with a new guarantee period of five years was taken up on all the policies.

The prospect of the July 2020 premium review, which later materialised, led to the complaint.

The effect of these reviews on the respective policies is demonstrated in the following table. Only two policies are illustrated as the details for the third policy are identical to those of the second policy.

POLICY	**1**	**2**
Initial Premium	R580.31	R2 320.18
Initial Life Cover	R1 000 000	R4 200 000
1st Premium Review	**2010**	**2010**
Cover prior to review	R1 689 477	R7 095 808
Premium prior to review	R1 368.31	R5 470.85
Cover after review	R1 790 845	R7 521 556
Premium after review	R2 334.08	R9 547.05
2nd Premium Review	**2015**	**2015**
Cover prior to review	R2 260 898	R9 495 789
Premium prior to review	R3 417.33	R13 977.84
Cover after review	R2 396 551	R10 065 536
Premium after review	R5 079.75	R21 061.07
3rd Premium Review	**2020**	**2020**
Cover prior to review	R3 025 588	R12 707 506
Premium prior to review	R7 437.27	R30 835.52
Cover after review	R3 207 123	R13 469 956
Premium after review	R10 424.56	R46 337.34
Premium had there been no reviews	**R3 549.04**	**R14 189.98**

Option 2		
Cover after review	R2 570 297	R10 702 077
Premium after review	R8 181	R13 469 956
Note: The "cover after review" amounts all reflect the scheduled annual cover increase only, and the "premium after review" amounts include the scheduled annual premium increase and the increase as a result of the review.		

Discussion

Universal life policies essentially transferred the investment risk, and in some instances the mortality risk, to the policyholder. Whether policyholders, or even the financial advisors selling the products, understood this risk, or contemplated the possibility of the premium being insufficient to maintain the life cover, is questionable.

Complainants are of the view that insurers prejudiced them by not informing them earlier that premiums were not sufficient to cover the cost of the cover. These complainants argue that this deprived them of the opportunity at an earlier stage to consider whether to carry on with the policy or to terminate it. Regular reviews would also have prepared them for gradual, reasonable increases rather than the drastic increases now required.

This newsletter focused on the issue of premium reviews only. There are other issues giving rise to complaints about universal life policies, such as low or no maturity values as at the date of maturity, despite regular premium payments.

The problem with universal life policies is not unique to South Africa. It is also not a problem that this office is able to easily resolve by granting relief to individual complainants. These are issues which have a wider impact, and they involve the actuarial viability of policies. We have thus raised awareness of the issues with the Financial Sector Conduct Authority ("FSCA") and National Treasury and to question whether an approach at industry level is required.