British Actuarial Journal, Vol. 21, part 2, pp. 364–380.
doi:10.1017/S1357321716000027

Expert judgement

Abstract of the Edinburgh Discussion

[Institute and Faculty of Actuaries, Sessional Research Event, Edinburgh, 8 June 2015]

This abstract relates to the following paper: M. Ashcroft, R. Austin, K. Barnes, D. MacDonald, S. Makin, S. Morgan, R. Taylor and P. Scolley. Expert Judgement. *British Actuarial Journal*. doi: 10.1017/S1357321715000239.

The Chairman (Mr J. R. Crispin, F.F.A.): Welcome to the discussion on the expert judgement paper by the Solvency and Capital Management Working Party.

For those of us who have been heavily involved in the Solvency II internal model approval process (IMAP) and all that it entails, expert judgement is a topic close to our hearts. It is clearly a topic close to the prudential regulation authority's (PRA's) heart as well.

It is one of the challenging aspects of Solvency II. It underpins all aspects of the modelling and the documentation. Even the simplest risks, like equity risk, have expert judgement underpinning them. This paper sets out views on what expert judgement is and a process for how you can go about documenting it. It is a useful addition to the actuarial literature.

Four members of the working party are going to give a short presentation on the paper. Michael Ashcroft, who has been chairing the working party, which he joined in 2009, is currently at Scottish Widows, where he is Head of Investment and Capital Analysis.

Roger Austin, who joined the working party in 2007, is currently a partner in Austin Professional Resourcing, which he founded in 2006. Stephen Makin joined the working party more recently in 2013. He is a consulting life actuary partner at Hymans Robertson. Peter Scolley joined the working party in 2012. He has just completed his MBA at Manchester Business School and has started a role at Standard Life. He is going to update part of the paper.

Mr M. R. Ashcroft, F.F.A. (co-author presenting the paper): The agenda for today is for me to cover the background and the framework. My colleagues will cover the key elements of the process, touching on the worked example that we use within the paper, and then validation.

Starting with the "who", the "why", and the "what" of this topic, and how the working party arrived at looking at this subject. The "who" come from the wide membership involved in pensions consulting, from independent actuarial companies who work with general insurance and life insurance companies, and from some large life companies in the UK. We also have someone from the PRA, who has given us great insight into the expectations or directions from the regulator and from European insurance and occupational pensions authority (EIOPA).

The group over the past 4–5 years has focussed on some of the practical areas around solvency and capital management, focussing primarily on Solvency II.

About 4 or 5 years ago we produced a paper for staple inn actuarial society (SIAS) on the implementation challenges of the internal model. We were proud that it was submitted to EIOPA in order to feed into their thoughts on how to implement the model approval process.

We have also conducted talks on other areas such as transitional arrangements and materiality.

Why expert judgement? It is not new as being a member of the institute & faculty of actuaries (IFoA) is evidence of expertise within certain areas of the field. We have had for a long time standards and regulations about being able to demonstrate this sort of judgement: the technical actuarial standards (TAS) standards and the approved person regime, and so on.

However, the bar has been raised significantly by the requirements under Solvency II. The same message comes from the regulator. Some of the early letters of feedback on the IMAP commented that expert judgement standards were not being met by many firms. We saw this as a difficult issue with which companies were struggling.

The standards of documentation, process, and governance that are required by Solvency II are a significant step up from the historical standards that we have tended to meet.

One issue here is the "model answer". We are not saying this is a "one size fits all" type of process. Businesses will have to shape it to their particular need, depending on the materiality or the complexity of the issue at hand. Proportionality has to be a cornerstone for how we set this out.

Given that Solvency II is the big driver for this, it is important to highlight where expert judgement is placed in the regulations.

There are no specific references to expert judgement within the Level I text.

Level II has a specific reference in Article 2 about setting a level of expertise that companies have to demonstrate. This is primarily around the assumptions that feed into technical provisions, SCR, and so on.

It is in the Level III text that expert judgement is referred to throughout some of the key sections: materiality, governance, communication and uncertainty, documentation, and validation. There is a strong theme throughout the Level III guidance of a minimum standard and a minimum set of expectations, for expert judgement.

When we started this work many of the companies we talked to perceived expert judgement to be related to assumptions and capital. That is not the case. There have been clarifications from the PRA and others over the past few years. This is an internal model and standard model firm requirement and it applies to the whole balance sheet. It is not a small piece of work and cannot be pigeonholed to one part of the business.

In actuarial work there are many areas where we are required to make judgements. Not all of these are what we term "expert" judgements. This is all to do with materiality and uncertainty – a variety of judgements that can have very different implications for the balance sheet.

https://doi.org/10.1017/S1357321716000027 Published online by Cambridge University Press

Also, defining expert judgement is in itself a judgement. One of the key stages we start out with is to decide whether or not something falls into the definition of "expert judgement". That is dependent on the circumstances of the particular judgement in hand.

For example, for a life company trying to set a mortality basis, there is a variety of judgements and decisions that need to be made. If I find I have someone in the data that was born in 1760, what do I do with that? I know that it is wrong and I need to fix it. There are many ways of doing that. There is no "expert judgement" within that. That fits into a decision, a judgement.

However, there is a debate within the industry about where we set the level of the improvement factor. That, by definition, fits in the "expert judgement" process.

There is a consensus about mortality risk factors. But things change over time. Particular lives in different companies may be different. This is dependent on the complexity of the business and the complexity of the lives that are covered. There is a grey area. Part of that process is deciding whether the judgement is "expert" or not.

Assume you have now decided something falls under "expert judgement". What does that mean? What do you have to do differently? There are additional standards that you now need to meet that we consider Solvency II sets for companies. This rigour appears in a variety of different areas. I will focus on the approach to follow, demonstrating that you have a robust process, which is consistent, appropriate, and proportionate.

A decision made without appropriate documentation is something you cannot easily justify. Documentation is a critical part of expert judgement, probably more so than for many of the other areas of the Solvency II requirements.

We see validation as a strong way of challenging any decision that is made throughout, and making sure you are comfortable with the process. The validation requires some extra work.

Finally, you make a decision based on the information at hand, but the environment and information can change. So making sure you have a robust process to monitor and manage the judgements that you have made and ensure that they remain appropriate for the balance sheet and for the decision-making, is also critical. That was the "how".

I now want to touch on the "what". Expert judgement applies to the creation of the balance sheet, using the standard formula or an internal model. In the paper, we break it down to three main areas: methodology, assumptions (including parameters), and approximations. This should not be considered to be an exhaustive list. They can appear in a variety of ways: modelling, application of data, or even aggregation. These are judgements for many companies and many companies do not appear to be documenting them as highly as they should.

Our paper sets out these areas in more detail and some particular examples of expert judgement.

Roger [Austin] will later take us through the expert judgement process we are proposing. I start with the concept of the formation of judgement. What are the key elements? Identifying potential useful data sources is an important part of the process. Those data sources are used and applied by the experts, and then both the data and those experts' views feed into the decision-making.

https://doi.org/10.1017/S1357321716000027 Published online by Cambridge University Press

This gives the foundation of the judgement process. How robust your decision-making process will be at the end depends on the materiality of the decision, the level of judgement required, the level of in-house expertise and capability, and the firm's decision-making and governance.

Proportionality is important. You need to make sure that, given the materiality of the decision, the scale of your judgement process is appropriate. In many ways the expert judgement process is just an extension of what many companies should already have in place in terms of their assumption-setting processes. However, it is trying to input something more robust, more size scalable, and that brings consistency across the work that they do.

There are areas where natural bias can come in, for example, where the experts are also the decision-makers.

The expert judgement process sits within a wider framework. Much of the work we did at the start focussed on the framework in which the process is embedded.

A strong governance framework is important. A good, clear, expert judgement policy is relatively critical. Quite a few companies have a stand-alone expert judgement policy. Some, however, embed it into other policies such as materiality. Certainly, the two are closely linked. Areas that we see the policy needs to cover include scope and definition, the process and documentation required, having regard to materiality and proportionality policy. Consistency, and how that is maintained, and appropriate validation and challenge are other stages in the process.

More generally, in managing an overall strong process and tracking it over time, a variety of tools can be used. The paper discusses some of the practical tools that can support an expert judgement framework. The one we have seen used most successfully in companies is the expert judgement register. This is a database of judgements, which allows you to track over time the particular judgements that were made and all the background information that you need to manage them. This will typically be the expert judgement description, materiality, scope, and application limitation. Understanding when it should be applied is vital to understanding whether it fits and can be used for other decisions. Names of experts and their qualifications would be included to record who advised in helping to reach decisions. Governance process, links to documentation evidence, relevant dates, and review triggers are all critical too. Understanding when you need to look at the judgement again, given its materiality and application, is crucial.

Having touched on the framework, Roger Austin will cover the process.

Mr R. Austin, F.F.A. (co-author presenting the paper): A process, at a high level, is familiar to most of us, as effectively it is a form of the actuarial control cycle.

The elements of the proposed process can be grouped into five key stages. The first stage is the preliminary assessment of the judgement. In this stage, you test the judgement against your definition of expert judgement. If it meets the definition, then you follow the expert judgement process. If it does not, it is probably appropriate to follow a less rigorous process.

The second key stage is defining the problem. It has quite a few steps and will cover aspects such as articulating what the problem is, identifying potential experts, and putting together a brief for those experts.

https://doi.org/10.1017/S1357321716000027 Published online by Cambridge University Press

The third key stage is eliciting experts' views. This will involve obtaining experts' views but also analysing their views and clarifying those views, where necessary.

The fourth key stage is the important decision-making stage. It is at this stage that the decision-makers will be reviewing the sources of information, and also the experts' views and subjecting them to appropriate scrutiny and challenge in order to make an informed decision.

The fifth key stage is ongoing monitoring. There are two aspects to this. There is the scheduled review of the decision, perhaps on an annual basis; but also there is the monitoring of triggers where a non-scheduled review might be appropriate. That completes the loop.

Before coming on to our worked example, we have put together a few concepts that are useful in the context of expert judgement. These are the "plausible range", "uncertainty total impact", and "regions of expert judgement". All judgements should lie within a plausible range, which we mean judgements within that range would be considered plausible by a number of experts.

When experts are providing their central view, they should also provide a lower and an upper bound for that view at a particular level of confidence. That effectively gives the plausible range. If number of experts are involved on a particular process, their plausible ranges can be brought together into an overall plausible range. That plausible range gives an indication of the uncertainty around that particular judgement.

The plausible range can be used to estimate the sensitivity of the output metric of interest across that plausible range, effectively giving an impact range.

If we sum the impact ranges across all expert judgements, we get an indication of the sensitivity of the particular output metric of interest resulting from the uncertainty around inherent expert judgements in the model. We call this concept the "uncertainty total impact".

The last concept is that of "regions of expert judgement". Most firms will be keen to try to reduce the uncertainty total impact. One way of doing so would be to allocate resources and budgets to particular regions of judgement, and firms might find that efficient. By regions we mean our estimates of longevity, expenses, and so on.

For certain regions we might potentially go further than that into areas of judgement; so within longevity it might be the areas of base mortality table and improvement factors.

For certain regions of judgement it might be efficient to think of the various expert judgements within that region together as we go through the expert judgement process. That can improve efficiency and potentially consistency.

We have a worked example in our paper and I want to go through that now.

To set the scene, we have a new insurance company. It intends to write only bulk purchase annuity business, and needs to establish its mortality assumptions.

The first key stage of the overview of the process is the preliminary assessment of the judgement. The first part of that is to identify the judgement required. In our worked example it is mortality improvements.

https://doi.org/10.1017/S1357321716000027 Published online by Cambridge University Press

We then need to assess whether it meets the definition of an expert judgement, which as Michael Ashcroft said earlier, is one of the tricky areas. There can be quite a few judgements lying in the grey area as to whether they meet the definition or not.

In our particular example, clearly mortality improvements are going to be a key risk to such a company. Therefore, they do fall within the expert judgement process.

Peter Scolley is to continue by going into more detail at the "defining the problem" stage.

Mr P. Scolley, F.I.A. (co-author presenting the paper): Properly defining the problem brings significant advantages to the efficiency and effectiveness of the expert judgement process. For our annuity company we have split out base mortality rates and improvement rates. For the remainder of the example we will just cover annual improvement rates.

It is important to define correctly the terminology used. This enables consistent understanding between users and over time. For our example, we have defined our longevity improvement as the annual improvement rate, by calendar year and age.

The next stage is to articulate why the judgement is required (in this case it is for assumptions), and what metrics will be of interest. Here, it is important to bear in mind the eventual use of these metrics.

As well as setting out the process for the judgement and the terminology, we find it useful to state a high-level understanding of the firm's exposure. In this case it is relatively straightforward. In other cases it might be more complex, particularly where there are interactions between assets and liabilities.

It is important to define the level of granularity required in the judgement. This can lead to conflicts between different users, as some users may require more detailed assumptions and judgements than others.

After the judgement has been articulated, it is important to set out the reason why a judgement is required. In this case it is because a new product is being sold. However, it might be an existing judgement that is being reviewed. This might be as a result of a trigger requiring the assumption to be reviewed.

At this stage we have established what is required and why. The next stage is to set out an understanding of the current situation around the judgement.

The first part of this is to understand what has been done previously and identify the drivers to change. There is an argument that you should not look at what has been done previously as it might unduly bias the judgement this time. Nevertheless, there are advantages to understanding what has been done previously and knowledge that has been gained by the company. In our example, the choice is simplified by the lack of a previous judgement.

Earlier, Roger [Austin] introduced the concept of a plausible range. This is the next stage of the process. For example, this would start off by considering possible ways to model future mortality improvements. Here we have used the continuous mortality investigation (CMI) model. Again, this is to demonstrate the process as the CMI model is based on a number of judgements. Many of these have default parameters in the model, but the long-term improvement rate is left to the user; and it is the long-term improvement rate that we will focus on as the judgement to be made.

https://doi.org/10.1017/S1357321716000027 Published online by Cambridge University Press

The next stage is to come up with the plausible range. In this case we have shown an example of historical data and benchmarking being used by the expert to construct the initial plausible range. Purely, for example, in this case we have come up with a range of $1\frac{1}{2}\%$ and $2\frac{1}{2}\%$ for the 25^{th} and 75^{th} percentiles, respectively.

Now we have the plausible range, the next stage is to assess the impact of the plausible range. Obviously, a range of metrics could be used. In this case we are using the best estimate of liabilities, and looking at the difference between the 75^{th} and 25^{th} percentiles, which is £53.4 million.

This does allow a quantitative assessment of the impact of the plausible range, which may be helpful for assessing materiality and prioritising judgements for further attention.

The next stage is to consider what information or analysis could be used to reduce the plausible range, and gain greater comfort over the estimate. We have given a few examples of possible analysis in this situation, as it is important to think widely about what can be done even if it requires a significant amount of effort and time.

So far we have considered clarifying what the judgement is, and what can be done to gain additional insight. The final stage of defining the problem is to decide what actions are going to be taken. Practically, this needs to balance the time available, which unfortunately is not unlimited, the expertise available, which again is not unlimited, and also the materiality of the judgement. It is important to remember that regardless of how much energy or effort is put into a judgement, it may never reduce the uncertainty in the judgement to 0.

Once the decision about what actions to take has been made, it is important to set out the overview of the need for the judgement and identify the roles each individual will play.

A clear brief is important at this stage, as it helps define what the experts are being expected to do and we propose that this brief covers the following information:

- Articulation of what the expert judgement relates to and why it is required.
- Definitions of terminology.
- Documentation of what was done previously.
- Initial estimate of the plausible range.
- Potential information sources.
- Practicalities such as areas of concern for decision-makers.

This brief may require more than one iteration as experts challenge the brief and it is further refined.

This concludes the "defining the problem" stage of the process.

Stephen Makin will now cover the rest of the process.

Mr S. J. Makin, F.F.A. (co-author presenting the paper): Having set out a clear brief for the experts, the next stage of the process is about getting hold of their views.

In the paper, we highlight four possible approaches, which are not exhaustive, nor are they mutually exclusive. You could ask for written responses to the brief, interview each of the experts individually,

https://doi.org/10.1017/S1357321716000027 Published online by Cambridge University Press

or together, to test interactions and challenge from all of them at the same time, which could be done without decision-makers present or, alternatively, with the decision-makers present.

You might also envisage some sort of iterative approach, starting with asking for the experts' initial responses in writing and then following up with interviews, perhaps, to clarify any areas of uncertainty or to ask any supplementary questions that arise.

In the paper, we have a concept of an "elicitation manager". One of the roles of the elicitation manager is to take responsibility for creating the brief, and to identify and reach out to the experts. At this stage of the process, the elicitation manager has an important role, which is to collate all the information, to highlight all the areas of agreement and disagreement, and then to see whether a further round of elicitation is necessary.

Whatever happens, the end result is that we get the views of the experts. In our worked example we have three such experts whose learned views are that the long-term improvement rate parameters in the CMI model that Peter [Scolley] referred to are 1.75%, 2%, and 2.5%/annum. For the avoidance of doubt, these are just hypothetical numbers for the purposes of our worked example.

To go along with all of this, there will be supporting rationale from each of the experts that will ultimately feed into the decision-making process, which is our next stage.

An important part of the decision-making process is for the decision-makers themselves to scrutinise and challenge what has already been done, especially if the decision-makers have not been involved in the process so far. Scrutiny and challenge should not relate only to the proposals put forward for longevity improvement assumptions, but involve being mindful of the bigger picture view.

For example, you would want to ensure consistency with mortality improvement assumptions – longevity and mortality improvement assumptions ought to be consistent – and you would also want to consider consistency, perhaps, with longevity stresses, to help ensure consistency between technical provisions and capital requirements.

Decision-makers should also be mindful of not introducing bias, which might happen if there is a particularly strong or dominant individual in a group of decision-makers. Peter [Scolley] mentioned earlier a different source of bias called "groupthink" or "herding".

It is obviously hard to eliminate bias entirely, but there are some practical techniques that can be used, some of which we highlight in the paper.

It is important that the decision-makers are clear in setting out their thought processes. What was it that led them to reach the decision that was ultimately taken? What information was used and why was it used? What information was important, and what less so? What were the plausible ranges, and what was the ultimate decision?

It is important that the thought processes of the decision-makers are all captured within the expert judgement register.

The final responsibility of the decision-makers at this particular stage is to loop back with the individual experts themselves; that is, to communicate the final decision, the plausible range, the overall rationale,

https://doi.org/10.1017/S1357321716000027 Published online by Cambridge University Press

and so on. To be clear, we do not see this as a polite nicety. It is an important part of the process. Under Solvency II, it is explicitly mentioned and mandated in the Level III guidelines. However, it is common sense that, having gone through this process, you should go back and close the loop off with the experts.

Our experience is that it is not always something that is done with relish but we cannot overstate the importance of going back and giving the experts the right of reply to make sure that their views have not been misrepresented or misunderstood in any way. Once the decision has been made, it is important to make sure that there is a robust system in place to monitor that decision. Roger [Austin] mentioned earlier the two main strands to the review that we advocate. The first is a fairly straightforward scheduled review in advance of the expert judgement, passing what you might call its "use by" date. So, for example, we might make sure we have reviewed the longevity trend assumption by a year from today. This is fairly standard and self-explanatory.

The second and most useful, insightful and valuable strand, would be some sort of trigger-based monitoring, the idea being to ask, "has something happened that would cause the judgement to lapse out of currency or usefulness?"

The expert judgement register is something we feel is quite important in this particular part of the process. It captures many of the points that have gone into helping formulate decisions so it might give you ideas as to what some useful triggers would be. An example would be if you come across material data that would give you cause to go back and re-examine the judgement that was reached. Even more pragmatically, as another example, you might be on the receiving end of some formal guidance from the regulator, which would certainly cause you to go back and challenge what has been done before. You do not necessarily have to agree with what the regulator has said but still have to address the formal guidance.

Validation is an important part of the process. Solvency II internal models require independent validation and the consideration of a minimum suite of tools. As Michael [Ashcroft] said, the need for this validation is definitely not the preserve of Solvency II internal models, and it is definitely not the preserve of Solvency II more widely. It is quite the opposite. All actuarial models use expert judgement – to a greater or lesser degree – and that judgement should be validated.

Judgement is difficult to validate; it is an amorphous and perhaps slippery concept at times. However, a proportionate application of the key features of the process that we have set out – the logical structure with clearly-documented thought processes – should aid that validation process. Specifically, this structure enables each stage of the process to be subjected to an appropriate degree of scrutiny and challenge.

The validation tools, mentioned earlier, required by Solvency II are stress and scenario testing, benchmarking, peer review, simplified models, and so on. We expand on these, so I encourage you to look at the paper to see which of those might be useful for your own particular situation.

One that I will expand upon, a concept in Solvency II, has come to be known as "profit and loss attribution". One of the key areas for expert judgement for any modelling exercise is which factors to model and which ones not to model. If a model successfully explains all of the profits and losses that have arisen over the year it might – although, this is no guarantee – validate the choice of risk factors.

On the flip side, if there was a large amount of unexplained profit and loss over the year, then naturally that gives you good cause to go away and re-examine whether or not the risk drivers you have chosen to model are the right ones. That would form an important part of ongoing validation.

https://doi.org/10.1017/S1357321716000027 Published online by Cambridge University Press

Expert judgement is inherent in all models and will always be required. We are not suggesting that we shy away from it, but rather that we should embrace it. While Solvency II is increasing the focus on expert judgement, it is definitely not, as far as we are concerned, the preserve of, or indeed only a requirement for, Solvency II.

The framework of having an expert judgement policy, a sensible governance structure, a robust but proportionate process tailored to the firm's needs, all backed up by good documentation, validation and ongoing monitoring, is crucial.

The Chairman: The opener is Peter Heffernan, who is an Associate Director of Pricewaterhouse-Coopers (PwC) based in Edinburgh. He has been with PwC for over 3 years focussing on audit and enterprise risk management work. Prior to that he worked at Lloyds Banking Group for 14 years, where he was Head of Financial Risk.

Mr P. J. Heffernan, F.I.A. (opening the discussion): This is an interesting paper, as it sets out a process for expert judgement.

The paper sets out in detail what is in reality a straightforward thought process. You find an expert, ask him what he thinks, write it down and then you move on. The paper talks at length about a detailed process, and takes an almost mathematical, logical approach to the concept of expert advice.

The starting point is proportionality or materiality. There is no point in doing all this if it is not going to make a material impact. That sounds like stating the obvious, but I come across many instances where a lot of work is done over something that is not material. The starting point is to ask "Does this matter?" Then you can move on through the process.

The paper talks about a "plausible range" for this: the idea of trying to quantify upper and lower percentages.

I spend a fair amount of time working with accountants. They perceive an actuarial black box – that it is all witchcraft and that we are all just making it up.

Given that, anything that we can do as a profession to narrow down the uncertainty, or increase certainty or transparency, is definitely good.

There is a note of caution in that it is possible that the reader may think that following the process leads to a more certain answer than it actually does, or even lead to the right answer. That is not what the paper intends, but my concern is that that is what people could think.

As a profession, it is important that we present a nuanced view. By definition, you are applying expert judgement because there is no right answer, because it is difficult or it is complicated. We need to persuade our audience that we are talking about uncertainty. There is a danger that we present too certain a view of the world and that people seek certainty when there is none.

We need to bear this in mind when we are doing our work. The value we add is portraying the fact that these assumptions, these judgements, are uncertain, and even after having gone through this process there is still uncertainty.

https://doi.org/10.1017/S1357321716000027 Published online by Cambridge University Press

There is a good book on this subject by Dan Gardner. It is called *Future Babble: Why Expert Predictions Fail and Why We Believe them Anyway*. The theme of the book is that you ask experts for their advice. Invariably they are wrong, but you still listen to them. You go back and ask them again, even though they have been proven to be wrong in the past. There is a psychological bent to this.

The paper talks about whether you start with updating previous judgements or with a clean slate. That is maybe contentious and I do not know the right answer. It is good that the paper mentions behavioural financial aspects like anchoring and groupthink and that may be a topic for a further paper.

On groupthink, for example, you set out some assumptions and typically the board wants to appear in the pack, and the PRA are comfortable when the assumption is in the pack. All that happens is that the pack just creeps up or down. It requires strength of character and confidence to set an assumption that is out of line with the pack. It is important that groupthink is considered when you talk about expert judgement, and it is certainly an area about which an independent reviewer needs to think. It is important that we are all conscious of our biases.

I liked the worked example in the paper. It brought the process to life. The question I have is whether or not this could resonate with an executive committee or some sort of board. If we said that this is the process that we are going to undertake, how would the committee receive it? To the layperson, to the non-actuarial mind, is there a way that we could communicate this detailed process that would make it easier for the layperson to accept?

Maybe a next step would be to test the process with a friendly executive committee or board. If you ran this process across an organisation and then quantified the uncertainty and put it in front of the members of a management committee, how would they react to it? If you placed some numbers in front of them about how uncertain this is, would they react well or would they be slightly surprised? I suspect the latter.

I feel strongly that it is right that, as the paper says, the purpose of validation is to obtain an independent viewpoint that ensures the judgements are reasonable, or not unreasonable, and not to obtain a second or third opinion. It is right that you make the judgement once and get it validated rather than ask another expert and just receive a different opinion, which may be equally valid.

It is interesting that the paper says the process relies on prior beliefs and time resource constraints, which naturally introduce biases. That would say to me that you need strong validation. But, we need to be aware that the validation process is often at the back end and more time-constrained, so similar biases can occur when you are validating judgements. Again, I have no particular observation other than the fact that we need to be self-aware and be aware of the biases that a situation will introduce.

Finally, section 6.5 talks about cultural considerations: the ideas of keeping it fresh by rotating the experts, and being humble and recognising your limitations.

The cultural considerations are more interesting than the process, although I do recognise that for other people the process itself will be more valuable. Again, maybe that is a topic for a further paper.

https://doi.org/10.1017/S1357321716000027 Published online by Cambridge University Press

The paper is thought-provoking and detailed with a good mix of theory and practice.

Mr A. J. Clarkson, F.F.A.: As Michael [Ashcroft] said at the beginning, expert judgement is clearly receiving more focus under Solvency II. It is a concept that has been the cornerstone of actuarial work for many years. All actuarial judgement could, in a sense, be viewed as expert judgement, although at times we should consider whether it is expert judgement, or whether it is merely judgement. Often actuaries think of themselves as experts in areas where maybe they are not experts.

There is clearly merit in properly thinking through the way in which expert judgement is sought and used along the lines set out by the authors. However, it is important, as the authors point out, not to lose sight of proportionality and the need to ensure that we do not create a bureaucratic monster. In my experience, the greater the number of boxes that need to be ticked to demonstrate compliance with a particular process or policy, the greater the risk that a tick-box mentality, rather than brain-engagement mentality, will be applied.

If you add up the number of judgements underlying an internal model under Solvency II for an insurer with broad and varied business, the answer would be high, as is evidenced by section 2 of the paper. I am not sure how practical it is to apply the approach suggested to all judgements.

The overall objective of any expert judgement framework is to ensure that, first, all those involved in the setting of assumptions based on expert judgement have a clear understanding of the context in which those assumptions are made and applied; second, that there is transparency relating to the uncertainty of assumptions and the resulting outcomes.

The framework should aim to achieve this in the most efficient way possible, taking proportionality into account. It is important to apply a consistent approach; but equally important not to over-engineer that approach. There are limited resources within a company and across the industry, and these need to be applied where they can add most value. In that context, as the authors have pointed out, a link between the expert judgement framework and a firm's materiality framework is very important.

The authors suggest in paragraph 3.5.11.3 that the framework should equally apply to decision-makers. I am not convinced. Ultimately, the board are the decision-makers. The authors may mean the people proposing the method to the board. In which case, an alternative model is to distinguish between judgements made by those developing and proposing the methodology and judgements from other people that they have relied on in developing that methodology. The person proposing the methodology should make it clear in their documentation when they have relied on the judgement of others. They should not have to follow the same formality in terms of forming their own judgement as long as the rationale for that judgement is clearly set out. When I am independently validating a model, I find that this is an effective approach to use.

In terms of a plausible range for the expert judgement, it is important for the users to understand that there is a range of plausible judgements and to understand the sensitivity of the overall capital requirements to the breadth of this range. The concept of plausible range is discussed when there is discussion about the internal model with the board. However, the challenge in determining a plausible range should not be under-estimated and in itself involves significant judgement. Whilst it might be clear if a particular assumption lies outside a plausible range, it is much less obvious where the boundary of that plausible range lies and where you cross that boundary.

https://doi.org/10.1017/S1357321716000027 Published online by Cambridge University Press

In terms of the use of an "uncertainty reduction budget", in an area where there is inherent uncertainty, more focus on a particular judgement does not necessarily reduce the level of uncertainty. The authors note this in paragraph 3.2.4.3. The authors use the concept of an uncertainty reduction budget where there is more uncertainty and the overall result is sensitive to the judgement. Equally, I find these useful factors in judging how frequently validation should be carried out within a validation framework.

Finally, the authors suggest in paragraph 3.5.7.1 that engaging additional experts, analysing new information sources and considering alternative methods may help reduce the plausible range. These are all important considerations when developing methodology. However, in my experience, they can be just as likely to increase the plausible range as reduce it.

Mr Ashcroft (responding): You probably picked out some of the critical points we are trying to get across here in terms of proportionality. We do not want a 300-strong expert judgement team. More people would help but that is not the aim of this.

One of the most interesting points for me that you raised was around that rationale for the judgement process. In many ways, that is the mini version of this expert judgement process we have gone through. One of the things we have seen least is documentation of the thought process that has been gone through as to why a particular expert judgement has been chosen as the appropriate one. Quite often it is a proposal from the expert that sets it out. The minutes will say it was discussed and agreed. But that discussion, in terms of the challenge, the insight, and the potential issues that were flagged and then closed off as part of the discussion, in some ways are the most valuable part of the decision-making process. All we are trying to say here is that recording and measuring is a critical part.

I agree with your point that there is a hierarchy here. By definition, you start at the lowest level and build judgements on top of that. Quite often, what goes to the board is that top level. Giving sight of how the judgements have been arrived at is quite critical.

Mr Austin (also responding): You touched on a valid point. There are certain judgements that are quite stubborn. It does not matter how much resource and budget you invest in them, they are still unknowable, so it is worth bringing that out.

Mr A. C. Martin, F.F.A.: I have nothing in my business related to Solvency II but the paper goes beyond Solvency II and it is most welcome.

I see many parallels with risk management, but just as risk registers should not just be tick boxes or criticised with hindsight, they do have a place in the future. In particular, risk registers have been scrutinised after the financial crisis and I hope the expert processes, assumptions, monitoring, and judgements will not only be assessed with hindsight in future. I am sure that HBOS, RBS, the Greek Government, and FIFA have made some good decisions and avoided the risks in the past – at least, at some time!

My first general observation centres on what I call the fourth dimension, and that is time. Like risk registers and processes, expert judgement, especially from the perspective of data monitoring, should focus even more on time. We will be judged by the future. The military analogy is quite good, and that is simply that there is no point in being perfectly prepared for the last war. The only guarantee is change.

In financial services this is particularly important with changing technology, increased competition and regulation. The paper recognises time changes in data manipulation, process overview, plausible ranges, and uncertainty risk budgeting. I would merely suggest some further extrapolation and recognition that uncertainty risk budgets might have to increase rather than decrease over time. The mortality example that was given is very good. The database for longevity improvements does cover decades, but judging the appropriate time period may depend on the answer you want. Personally, I place much greater emphasis on more recent stellar medical advances than anything before, for example, penicillin was discovered or we had an national health service (NHS).

My second main point is that all judgements at an individual company level, and indeed from a regulator's perspective, need to consider the much wider environment, particularly the financial environment. In 2010, I did not have a plausible range for interest rates with negative signs besides the real interest rates. Negative real interest rates on government stocks are still called gilts, but they should perhaps now be more appropriately called gifts. Equally, I would not like to predict the level of real and nominal interest rates for the next decade, which is when it really matters. I am, however, sure they will affect pension scheme funding and life assurance company solvency – not just rates going up or down, but staying lower for longer. Therefore the expert judgement frameworks, processes, and challenges should include much more allowance for wider influences, even societal changes and the inevitability of change. Competition is a good example. We now live in a global environment.

Another aspect I would suggest should receive more attention is perception. I believe an expert judgement could be theoretically, statistically, and intellectually sound, and indeed sound in every other way, but it could fall down because it missed public, political, or regulatory perceptions. This is not an invitation to include psychology, marketing, consumer testing, or public relations in every judgement, but it would be dangerous to ignore them. You can think of some reputational issues that can change quickly. The example that I would put forward now is annuity rates. Amazingly, the life insurance industry was blamed by the government – with alleged vested interests and lack of openness and communications, whereas it was all because of open market options not being exercised.

Finally, I would suggest that the expert judgement framework should place even greater emphasis on the culture of the organisation. It is covered in section 6.5, but an open and transparent CEO or indeed organisation, and even a regulator, that encourages constructive challenge, is more likely to make good judgements. This applies to judgements from decision-makers and regulators. In the public sector regulation, there is a significant issue in the area of freedom of information. FOI is the acronym. Those considerations should not cloud the judgements. But, sadly, the word "FOI-able" is in the public sector dictionary, and I do not believe it is adding to the efficiency of the world.

Mr P. G. Telford, F.I.A.: I agree with what Alastair Clarkson said about the plausible range concept. That is valuable, particularly in communicating with boards and similar end users, because it tells the non-expert which uncertainties matter and which ones therefore they might want to work on reducing.

When judgements collide, then your judgements get tested. This happens quite regularly in merger and acquisitions work, when you talk to the regulator, or an investment analyst. If you have a good process and can explain quickly and credibly why your judgement is what it is, you are in a stronger position.

https://doi.org/10.1017/S1357321716000027 Published online by Cambridge University Press

Keeping sight of the pre-judgement range of view is useful. I come from a firm that values consensus. We work hard to arrive at an internal consensus. But we must not forget that we started from different positions because we need continually to challenge ourselves as to whether or not we have reached the right consensus.

I was not convinced by the uncertainty total impact concept for a couple of reasons. First, there seem to be many different ways you can define it. Do you aggregate with or without interactions and diversifications or do you just add up? Second, and more practically, there is a danger that the number is so large for a complex model it returns us to the realm of actuarial magic from which Peter Heffernan was so keen to avoid.

The follow-on concept of uncertainty reduction budget will most likely be driven by commercial and risk appetite considerations, where our firms are not in the business of reducing uncertainty for the sake of it. We want to do that because with less uncertainty either we will be more confident to pursue a certain line of strategy or we will need to put aside fewer of our scarce resources against a particular risk.

The question of to where uncertainty reduction effort is directed will solve itself, and be different for each firm.

Mr Makin (responding): I like the analogy of judgements colliding. That is in line with our thinking when talking about going back and reviewing your expert judgement if you are on the receiving end of some formal guidance from the regulator. This is something that many in the industry are seeing right now in the context of longevity stress assumptions, where the regulator has a much stronger view than many. A difference of views is really an argument with the regulator as to whose expert judgement is better. This is a hard argument to win, but success – whatever that means – should be easier if you have a process of the type we have set out underlying the assumptions that were proposed.

Mr M. A. Potter, F.I.A.: My question to the working party is whether in their thinking they distinguished between professional judgement and expert judgement. The reason for my question is that there is a guidance note on expert witness work, where an actuary might find themselves giving an opinion on another actuary's work and deciding whether that actuary was right to make that professional judgement. If they did, and if it was within what you might call a plausible range, then all is well.

I could not help but have the impression that getting a panel of experts to do this could lead to an averaging process creating just that effect that you said you did not want, which is the herding behaviour. Is that average a mean, median, or mode?

Mr Ashcroft (responding): I would see that as a subset of expert judgement. We started out at the beginning saying that in many ways the badge of the actuarial profession, like many other professions, demonstrates a certain amount of expertise in specific areas.

We are not just talking about purely actuarial judgements, but judgements that affect the wider insurance industry. Many of those in modern terms are from doctors or from medics looking at the history of cause of death models and so on.

Many of the decisions are insurance decisions that ultimately affect the balance sheet. But sometimes it is not an actuary or finance professional that is making the decision, so we are trying to give

https://doi.org/10.1017/S1357321716000027 Published online by Cambridge University Press

them a better understanding of what they are being asked to do from their own professional background.

Do we end up with a mean, median, or mode? We are trying to avoid that. It is about getting sight of the differences in judgement. If the reality is that there is a consensus, it does not mean that an extreme value does not fall into your range. That should be discussed as well. We see that as part of the challenge process. Why did that person highlight this extreme value or that extreme figure? Is their thinking different? Should you then bring that in?

I accept there is the concern that you will just take the average. That is not what we want. The value here is, after the elicitation of the expertise, demonstrating a clear, logical process to go through in making a decision.

We are not saying that it is easy; we are just trying to put a consistent framework around the process.

The Chairman: Colin Ledlie will close the session. He is currently working independently. Previously, he spent many years at Standard Life, including roles as group Chief Risk Officer and Chief Actuary.

Mr M. C. Ledlie, F.F.A. (closing the discussion): The paper goes into a great deal of detail about a proposed process. It should not be seen as a handbook on how firms should go about managing expert judgement, but rather as a set of guidelines and helpful suggestions.

Most of the suggestions are just plain common sense, reflecting the current expectations for rigour and process that flow from the PRA and the Solvency II regulations. Some areas are perhaps more open to alternative approaches, and firms may find better ways of achieving the objectives than the approaches set out in the paper.

The authors have been keen to stress throughout that proportionality is key. The paper may read at times as a little heavy-handed and process-heavy, but the authors acknowledge that their suggestions need to be implemented in a proportionate manner.

I would summarise the key points of the paper briefly and simply as: establish an expert judgement policy; establish an expert judgement process; establish an expert judgement register; ensure that you have good documentation; ensure that there is appropriate ownership of the process, the register, and the documentation; ensure that you have good communications so that all parties are talking about the same thing, and put in processes to monitor and review what you are doing.

These points to me are the essence of the paper and what firms should be aiming for.

The authors also suggest tools such as the determination of "plausible ranges", the determination of an "uncertainty total impact", and the setting of an "uncertainty reduction budget".

There is scope for firms to decide whether these suggestions are helpful to them or whether there is some alternative approach that will achieve the same objective in a more effective manner.

Looking at some of the comments and areas of discussion that we have had proportionality has come up many times. The authors have highlighted themselves that their suggestions need to be taken in a

https://doi.org/10.1017/S1357321716000027 Published online by Cambridge University Press

proportionate manner. We need to ensure that going through what could be a very detailed process is only done for the most material and key assumptions and not the small ones that have little impact.

The setting of a plausible range is a valid and important step. A number of speakers have echoed that. There is a danger in that potentially the plausible range for the answer, say the future improvement rate, may be seen as defining the potential range of future possible outcomes, which it is not. The range of potential future outcomes for mortality experience could be very different and much wider than the range of likely judgements that our range of experts will make, which probably will start grouping given the range of assumptions used by competitors in the market.

There was talk about the uncertainty reduction budget and whether it is useful – I have some questions about how useful the concept would be in practice – and also some questions about whether it leads to a false mindset.

One of the speakers said doing all the work could lead to more uncertainty in the final outcome. That is true.

The authors sided heavily in the paper, with the use of prior answers to guide the judgement setting. They do highlight that there is a risk of anchoring within that process. Perhaps, they could have been more open in the paper about the different approaches that you could take. You may wish to set your experts out with no knowledge of prior decisions and then, at a later stage in the process, decide how to home in on an answer and take into account what judgements have been made in the past.

I, too, am not entirely convinced about "uncertainty total impact" as a communication tool and would suggest an alternative approach might be a ranking of the various judgements in the same way that firms will do a risk ranking under individual capital assessment (ICA) or Solvency II, an approach which I have found useful in the past.

It is critical to remember that we are modelling the future and that there are limits to the accuracy of any model. We need to make sure that our processes are attuned to the model's inherent limitations.

I like the closing paragraph of the paper where it states by its nature the correct answer is unknown, possibly even unknowable, and to think otherwise is potentially dangerous.

The Chairman: I express thanks to the authors for pulling together a valuable and timely paper, and to thank the opener and the closer for adding their insights and also to all those who participated in the discussion.

https://doi.org/10.1017/S1357321716000027 Published online by Cambridge University Press