



SESSIONAL MEETING DISCUSSION

Consideration of the proxy modelling validation framework by the proxy modelling working party

[Institute and Faculty of Actuaries, Sessional Webinar, Tuesday 14 November 2023]

This discussion relates to the paper presented by the Proxy Modelling Working Party at the IFoA sessional event held on 14 November 2023.

Moderator (Mr D. J. Harrison, F.I.A.): In 2019, the Prudential Regulation Authority (PRA) led an industry-wide review on proxy models. The Proxy Modelling Working Party has been developing this paper for the past 3 years. The paper takes the PRA feedback into account and explores potential approaches to calibration and validation. It considers the different heavy models used within the industry.

The speakers are Maynard Kuona, a Fellow of the IFoA and a Senior Manager of Life Actuarial Practice at KPMG. Maynard has over 15 years' experience in financial services, covering both industry and consulting experience. Maynard's experience is primarily in helping clients either build or validate economic capital models, including risk calibrations, proxy modelling and independence modelling. Maynard has a particular interest in the application of modern analytical techniques to both new and old actuarial problems.

Matthew Thomson is a Senior Consultant within the Insurance and Financial Services practice at Hymans Robertson. Matthew has 8 years of experience with roles in industry and consulting. Since joining Hymans Robertson in 2019, he has worked on a variety of client assignments, with a particular focus on managing and modelling market risks. He has worked across all types of life insurance products in line 1, line 2 and line 3 roles. Matthew also leads Hymans Robertson's industry-wide benchmarking surveys on the matching adjustment and Limited price indexation (LPI) inflation risk. Both surveys cover all eight of the current UK bulk purchase annuity firms. He has also worked with several smaller UK-based insurers and has been seconded to a large overseas insurer. Before joining Hymans Robertson, Matthew worked in the actuarial reporting team at Standard Life. He is a Fellow of the IFoA and holds the Chartered Enterprise Risk Actuary accreditation.

Mr M. Thomson, F.F.A.: I am going to start with some context around why the Working Party was established and go over some definitions and scene setting.

I believe we are at least the second incarnation of the Proxy Model Working Party. Our work builds, at least in part, on the work carried out by the Working Party's 2014 paper, which took a closer look at some different types of proxy models at a time when their general use was increasing due to the imminent introduction of Solvency II. Our version of the Working Party was established in response to the PRA's June 2019 "Dear CRO" letter, which set out what it viewed as best practice in relation to the use of proxy models. Our objective was to consider the observations raised by the PRA and set out a framework for how businesses could apply this feedback to show that their proxy models were appropriate for use. When the Working Party was established, the area of proxy modelling was one that was clearly at the top of the PRA's priority list.

In preparing their views on best practice, the PRA surveyed a range of UK insurers to understand their existing proxy models and the frameworks and processes surrounding them. The

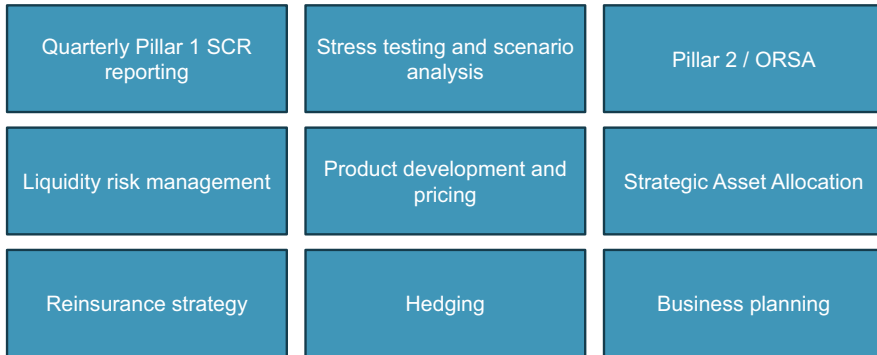


Figure 1. Proxy model uses.

best practice letter noted that no one firm exhibited best practice across all areas of proxy modelling. In preparing their views on best practice, the PRA also noted that proxy modelling was an area where thinking and techniques continue to evolve and also one where an approach that is well suited to one firm may not be the best approach for another. We understand that some firms did begin to take action following the publication of this letter. However, anecdotal evidence suggests that proxy model improvements were de-prioritised by many firms in early 2020, as a result of the Covid-19 pandemic. In the last few years, the regulatory focus has fallen on operational resilience, resolution and recovery planning and then on Solvency II reform, rather than proxy model improvements. However, we do expect the regulators' attention to return to proxy modelling at some point.

In our paper, we define a proxy model as any model developed to replicate or approximate the output of a more complex model. We will refer to the more complex models as full or heavy models. You may hear the term "Lite Model" used to describe a proxy model.

There are two other definitions that are important to understand. Calibration is the process through which the proxy model parameters are determined, such that the proxy model output is acceptably close to that of the heavy model, for a given set of inputs. Validation is the process of testing that your calibrated proxy model replicates the heavy model to a desired level of accuracy. In the paper, we go into a lot of detail on different calibration approaches, but we won't go into them here. Instead, we will focus on validation and the different ways in which one can validate a proxy model and provide comfort to key stakeholders that the model is fit for purpose.

Proxy models have been used by insurers for many years. For example, you might not want to run heavy actuarial cash flow models to understand the impact of a small change in an input assumption. The proxy model can be run more quickly and easily and will give you a reasonable idea of the types of impacts that you might see. Proxy modelling really took off during the introduction of Solvency II. Many firms have internal models that require multiple runs of heavy actuarial models under a variety of different stresses and scenarios to calculate the full Solvency II balance sheet. The time taken to do these runs can be huge, even before you think about the human time needed to produce, understand, communicate and validate the results. For this reason, many firms developed proxy models for use in quarterly reporting processes. These models can give results to a reasonably high degree of accuracy, and they can produce results much more easily or quickly than the heavy models, which is beneficial given the relatively short quarterly reporting cycle.

Additionally, under Solvency II, in order to get internal models approved, insurers have to demonstrate that their models are fit for purpose and used across the business for a range of purposes, as shown in Figure 1.

In terms of validation, we need to be able to demonstrate that the proxy model produces acceptably accurate results in several different scenarios, and we have to demonstrate this to a number of different stakeholders. For example, the first thing that many people will think about when they hear the term “proxy modelling” is a model that helps you calculate your solvency capital requirement (SCR). The proxy model needs to be capable of replicating the heavy model results for the 1 in 200 all-risk event that materialises over a 1-year time horizon. However, if I work in a team responsible for liquidity risk, I am likely to be interested in much shorter time periods, potentially only a few weeks or even a few days. Given this, I might be more interested in specific, fast-moving risks, such as market risks or a mass lapse event, than some of the longer-tailed risks that contribute to the SCR. Additionally, I might be more interested in less extreme percentiles than 1 in 200, maybe something like 1 in 20 or 1 in 40.

From this example, you can see how a proxy model designed only to be a good fit for calculating SCR might be less useful for other purposes. It is therefore important that we are clear about what our models are used for and that they are calibrated and validated accordingly. Model limitations should also be well documented and communicated so that you do not use a model that has been designed for one purpose, to inform decisions in an area where it was never intended to be used. Maynard (Kuona) will now go through some validation approaches in a bit more detail.

Mr M. Kuona, F.I.A.: I am going to discuss validating proxy models, what the objectives are when we are validating them and how the various tests that we have set out in the paper help towards achieving those objectives.

When validating a proxy model, we are seeking to obtain evidence or assurance that the model outputs will appropriately reflect the heavy model and that the results are reliable for the use to which we are putting the model. Proxy models, much like other models, will invariably have limitations. The proxy models remain useful as long as those limitations do not become an obstacle to assessing the risks that the proxy models are used to measure. It is therefore important that the validation allows us to get good insights into the likely impact of those limitations on the model uses. In an ideal world, we would be able to obtain a guarantee that the results produced by proxy models are within a target low tolerance of the heavy model results, either in relative or absolute terms. This would be easy to communicate to model users, including non-practitioners. However, obtaining such a guarantee is almost certainly infeasible. What we have instead is typically a suite of tests that we can apply to proxy models, each of which provides us with different perspectives and intuitions on the performance of the proxy models in relation to the heavy models.

Earlier, we discussed model uses being a key consideration when designing and operating proxy models. To add to the model uses, we also need to think of the direct and indirect users of the proxy models. We should consider what insights they would need from the model validation to provide them with comfort that the model’s results are appropriate and not misleading.

We consider that the validation framework, or process, needs to have the users of the outputs in mind. It should provide them with the appropriate evidence, insights and intuitions into the workings of the proxy model to allow them to make appropriate decisions on how they use the outputs. Indeed, if those models are not performing adequately, it should also provide them with enough insight to adjust the outputs. For example, they may have to calculate capital overlays to address model weaknesses. One of the key constraints in a proxy model validation exercise is the number of heavy model runs that one can realistically perform. Therefore, the evidence that we can gather is generally somewhat limited. Practitioners are invariably operating under a limited run budget, and we need to ensure that we can provide satisfaction to all model users from that limited run budget.

Moving on to the validation test discussed in the paper, which I won’t go through in detail, we have thought about the eleven validation tests that the PRA outlined in their paper and which are commonly used in industry. We have grouped them into three categories.

Those in the first category primarily test the goodness of fit of the calibrations using in-sample metrics. Tests in the second category provide validation and feedback loops and are generally out-of-sample tests. Lastly, tests in the third category help support the sign-off process, and these could be both in- and out-of-sample tests.

In the first category, these tests, which are largely based on in-sample goodness of fit, can provide diagnostics that allow prompt detection of potentially significant issues with the proxy models ahead of their use. This category is very useful to practitioners working at the coalface, as it were, as they provide indicators for possible model limitations that may require resolving ahead of full out-of-sample validations. In some scenarios, these may be the key tests that are available to test the goodness of fit and appropriateness of proxy models. Knowing how to interpret the output of these tests, as well as having a predefined set of interventions in the event of test failure, should form a core part of a validation framework for the proxy model.

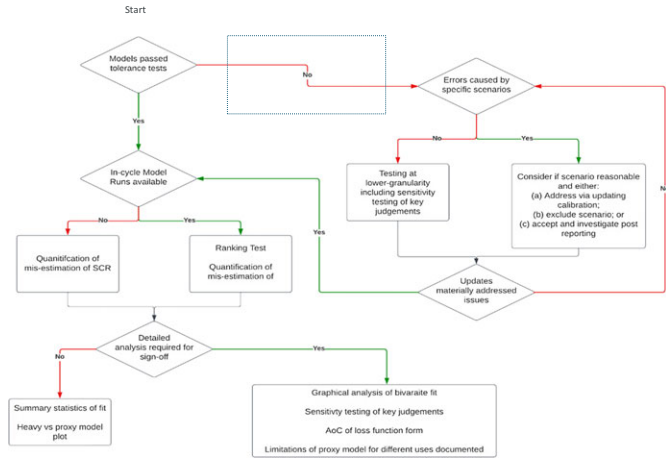
One of the limitations of using in-sample tests is that there will be some circumstances where increasing the goodness of fit in-sample can result in a reduction in out-of-sample goodness of fit (over-fitting). In addition, some in-sample tests may generally not be available or could be meaningless for certain calibration approaches. A good example of this is a precise interpolation, where effectively the residuals of the proxy model fitting are zero, so there are no fitting errors to analyse. In these circumstances, there will need to be more reliance on the second category of tests. These focus more on the out-of-sample performance of the model. As part of the production process, second or third lines of defence may seek evidence that the proxy models generalise and perform well out-of-sample. If not, they may want to be able to assess the likely impact on the SCR or any other relevant metrics. For reasons of independence, it is likely that these tests, including the specific scenarios that are used to produce the relevant metrics and the overall framework governing this testing, are defined by the second line. The choice of the validation scenario is key to this. While out-of-sample tests are viewed as a gold standard in terms of demonstrating the appropriateness of the model, the results are only as good as the scenario selected to form these tests. Performing these tests in cycle is generally a challenge, and it may be necessary to rely on out-of-sample testing performed off-cycle. In other words, testing based on proxy modelling exercises on a previous balance sheet date demonstrates that the design of the proxy model is appropriate and likely to produce results that are aligned with the heavy model. This would usually be coupled with putting the proxy models to the test, which would be discussed later.

Finally, those who have overall responsibility for the models may want to have a good understanding of their performance. They may require tests that are understood more intuitively before they are comfortable signing off the proxy models. These tests may be less statistically rigorous than those in other categories. However, they are still a valuable component of a validation framework that provides appropriate evidence for all model users. Graphical tests may well be something that a line 1 practitioner desires as part of their toolset when assessing the reasonableness of a proxy model.

There is some overlap between these categories, and all model users may find tests from any of the categories to be useful for their needs. However, we believe it is worth thinking about including tests from each category in an overall framework to ensure that the framework is maximally useful for all model users.

This brings us to the validation framework. We have set out in our paper the outlines of a proxy model validation framework that can form the basis of a dynamic and responsive proxy modelling process. The approach to validating proxy models should provide assurance to all users that the proxy models are fit for their intended purpose. This should consider the different users of the model, the materiality of the business being modelled, the heavy model it is designed to replicate and, in particular, any inherent limitations. Additionally, errors on lower percentiles may be more material relative to the effect or risk being measured than errors in the SCR percentile. At times, we can find that the errors do not necessarily scale with the size of the effect that we are measuring.

- Should reflect model uses (e.g. capital, liquidity, stress and scenario testing)
- Should reflect different users' requirements
- Should include information on tests from all categories



Illustrative decision-making process

Figure 2. Validation framework.

Therefore, if you have a smaller change, some of the results can look significantly different from the model in context.

There are additional considerations when firms calibrate their proxy models out of cycle and roll them forward to the calculation date. These are discussed in the paper. We believe that a defined framework for validation should be outlined within a firm’s suite of model documentation. This should justify the specific tests chosen, their relation to the metrics that the proxy models are used for and the interpretation of those tests, so the model users are able to decide how each of the tests will affect their use of the model.

The chart we have outlined in Figure 2 provides a high-level framework for informing what validation tests are appropriate.

We have considered a number of tests that we believe should be part of an overall framework. These include the results of out-of-sample testing, bias tests and testing the independence homoscedasticity and normality of errors. There may be circumstances in which some of these tests are rendered unnecessary by other tests. For example, if it can be established that all the proxy model errors are below the materiality threshold, we may be inclined to de-scope or put less emphasis on other tests. This should be something that the validation framework makes clear.

As previously intimated, the choice of validation scenarios is crucial. If the scenarios are not chosen appropriately, then the results of the validation can be less useful and potentially misleading. When you have a fairly low number of validation scenarios then the choice of the individual scenarios becomes quite important. If you can produce a large number of validation scenarios, then maybe you can place more reliance on the statistical approaches to assessing the errors and less focus on individual errors and residuals. In this situation, the process by which the validation scenarios are generated becomes much more important. Matthew (Thomson) will now go through some of the additional considerations regarding the roll-forward of the proxy models.

Mr Thomson: So far, we have focused on the types of validation that you can do to show that your proxy model gives a reasonable output for a given set of inputs at a particular point in time. However, the effort required to recalibrate a proxy model means that they tend not to be continuously recalibrated. Proxy model calibrations can be grouped into three types. The first approach is where the proxy model form and parameters are calibrated when the model is used; for example, within the quarterly reporting cycle. The second approach is where the proxy model

form is fitted out of cycle, but the parameters are calibrated when the model is used. The final approach is where both the proxy model form and parameters are calibrated out of cycle.

The first approach, where everything is calibrated when the model is used, is obviously preferable from an accuracy and relevance perspective. However, it might not be achievable for many firms, depending on their modelling capabilities or the types of business held. For example, this approach might be more difficult for stochastic business. The second approach, where the model form is fitted out of cycle and the parameters are calibrated in cycle, might be more proportional as it relies on a smaller number of full model runs to calibrate the coefficients of the polynomial. We consider these two approaches to be most aligned with the PRA's best practice letter. The letter noted that firms showing best-observed practice recalibrate and validate their proxy models in full each quarter.

However, we do recognise that some firms' limited modelling capabilities, and the complexity of their books, mean that they have to fully recalibrate both the model form and parameters out-of-cycle. They then apply adjustments to allow for known changes in economic or business conditions since the calibration date. This process is known as rolling forward your proxy model, or simply as roll-forward.

For market risks, shifts and scalars can be applied to proxy model risk distributions. For example, this might be to allow for changes in interest rates or movements in equity markets since the model was last calibrated. For stresses that are applied in absolute terms, such as an increase in interest rates, it might be most appropriate to apply a shift to the proxy model risk distribution. For example, if the yield curve had increased by 20 basis points since the model was last calibrated, you might shift the whole risk distribution for interest rates by 20 basis points. However, for stresses that are applied in relative terms, such as a percentage increase in equity values, it might be more appropriate to scale the risk distribution. Market risk adjustments tend to be straightforward, and we would expect most firms to be doing something to adjust their proxy models for observable changes since the last calibration date. Non-market risks tend to be more difficult, and we generally would not expect such adjustments to be made. This is mainly because demographic assumptions do not change very frequently and when they do (for example following an annual assumptions review), we would expect firms to be recalibrating their proxy models anyway. We would, however, expect firms to adjust their proxy models for the impact of new business and for the runoff of existing business since the calibration date. The way in which these adjustments will be made will depend on the business profile of the firm and the types and volumes of business in question.

That covers what you might want to do in terms of rolling forward your proxy model. However, where firms are using these sorts of adjustments, it is vital that there is sufficient validation in place to provide assurance that the model fit remains appropriate. The way in which you validate your adjustments will depend on the approach taken to the roll-forward, but there are a few principles that we think should be applied in the validation of all roll-forward approaches.

The first is that validation should be performed in cycle and prior to reporting. This will provide assurance to the senior stakeholders who are reviewing and signing off on the model results.

The second principle is that additional validations should take place outside of the reporting cycle. This will allow the roll-forward approach to be continuously refined. For example, a quick ongoing test of your roll-forward methodology is to roll your previously calibrated proxy model up to the date at which you fully recalibrate. You can then compare the output of the roll-forward proxy model with the output of the newly recalibrated model. We would expect firms to be able to account for any material differences here and have triggers for highlighting potential limitations in the roll-forward approach.

The third principle is that a trigger framework should be developed to identify when the roll-forward is not appropriate. This framework should set out the size or type of event that would trigger a full model recalibration, that is, where the movement observed is so material that the

roll-forward adjustments may no longer be appropriate. In our paper, we set out some suggestions for developing such a framework. For example, the roll-forward trigger framework should specify tolerances for all material individual risks that are included in the scope of the roll-forward, and the framework should be designed such that it is appropriate for each of these individual risks. Typical framework indicators that could be used to set triggers include movements in the 10-year swap rate for interest rate risk or movements in the FTSE 100 Index for equity risk. The trigger framework should define clear pass/fail criteria for each test. The framework should also consider changes in the firm's asset holdings over the period as well as changes in external economic variables. Finally, the trigger framework should set out how to apply any true-ups required to the proxy model. We would expect this trigger framework to be reviewed regularly, probably annually and more frequently if appropriate, for example in response to a change in risk appetite.

Material changes to the heavy model will also lead to discrepancies between the heavy model and the rolled forward proxy model. We would recommend that, where possible, any model developments are planned to coincide with proxy model recalibration dates. However, if this is not possible, the impact of changes to the heavy model should be analysed and understood so that the impact on proxy models can be estimated. Adjustments should then be made to the proxy model accordingly.

The final matter our paper touches on is the roll-forward of interactions and diversification within proxy models. We are not aware of any firms that do this currently, potentially because of how difficult or how judgement-based it would be. As with demographic risks, we would not expect risk dependency assumptions to change materially between model calibrations, but there may still be some interactions that could change. For example, the relationship between interest rate risk and lapse risk on types of business will vary in different interest rate environments. The roll-forward of interactions and risk dependency assumptions might be something that we could see firms try to allow for in the future, as focus returns to proxy modelling capabilities and modelling capabilities continue to develop.

Now we move to the question-and-answer session.

Mr P. Scolley, F.I.A.: I have three main comments on the paper. First, Section 5.2 of the paper discusses fitting, and it immediately jumps into polynomials. I felt this was a bit premature. I feel that the first stage of calibrating any proxy model should be to understand the business that we are trying to model. We need to understand what the key risks are and what the key interactions are, and we can then decide what an appropriate choice of loss function form might be. In many cases, that might be a polynomial because they do tend to fit business quite well, but they do not work in every situation. For instance, at the end of the paper, in the Least Squares Monte Carlo example, it notes value and an equity call option, which gives a poor fit generally because of the polynomials fitted. There are other situations that may be better fitted by interpolation rather than by polynomial.

My second point is the need to select suitable calibration ranges. I think that is a key judgement for anyone who is doing proxy modelling. It is not really touched upon in the paper, but if you fit too narrow a range, for instance, just the range between 1 and 200 up and 1 and 200 down, then your calibration might not capture the full range of modelling that your internal model produces. Equally, if you fit it on too wide a range, particularly in multiple dimensions, then you end up fitting a lot of variables that never get simulated, even under extreme roll-forwards. Therefore, you waste a lot of modelling and can over-fit the model.

My third point is in Section 6.3.3, which is the quantification of misestimation of the SCR. I felt this was a very good section, and I think it is quite important to highlight. Regardless of the work you do to fit proxy models, errors inevitably occur. As you mentioned, the roll-forward is quite key for a lot of firms. Firms always have the practical challenge of getting from their proxy model to a more accurate SCR number for reporting purposes.

There are a few other points I would like to make. One of the approaches was to use a 1 in 200 smooth scenario. The paper does touch on this, but effectively it relies on the assumption that the average scenario will produce an average error. I think that is quite a bold assumption. When making adjustments, this is better than no assumption at all, but I think it can perform quite poorly in practice. A better approach would be to use an average of unsmoothed scenarios and multiple scenarios if they can be modelled and the firms have the modelling capacity to do so. This is likely to produce a more stable adjustment.

There were comments mentioning a trailing error. I struggled to see why that would be appropriate. It would be good to know if firms do use it successfully and the circumstances in which it is appropriate.

Finally, the paper does mention using multiple smooth scenarios, but it does not go into detail about how those would be produced. There is a reference to the Murphy article in *The Actuary*. It is not clear whether the authors are using the technique that is briefly alluded to in the Murphy paper or whether other approaches have been used.

Mr Kuona: The first point was on fitting polynomials. We did focus on polynomials as a technique. It is widely used within the industry, and it is probably the most common technique used in the UK. I am aware that, especially in Europe, different techniques have been adopted, like replicating portfolios and other techniques that might be better adapted to the assets or liabilities being modelled. We focused on polynomials because we wanted to reflect what is largely the practice in the UK. I do agree that, if I were starting with a blank sheet of paper, I would want to assess whether there are better ways of doing this. In practice, what happened in the past, or at least my reading of it, was that a lot of the platforms that firms implemented the proxy models on initially only supported polynomials or largely supported polynomials. That is the way the industry has developed. If you were starting again, I would expect firms to consider alternative approaches.

The second point was around selecting the range of the calibration. It is briefly touched on in the paper, but I do agree that it is one of the more difficult areas. The wider the range, the worse the fit overall, but it may be more stable from period to period. That is one of the judgements that people have to reckon with when they are fitting proxy models.

On the third point, the quantification of SCR misestimation is not an easy thing to do. One difference from when firms started adopting proxy models is that it is probably more likely they can run 1,000 or 2,000 scenarios. Maybe that will become a little bit easier going forward. I am aware that some firms are looking at different techniques, such as using Graphics Processing Units, which may be able to run lots more scenarios than would have been practical when firms were starting out on this journey.

In terms of the average and smoothed scenarios, I agree with the point that an SCR adjustment or correction based on looking at a single smoothed scenario is not likely to perform as well as averaging the errors over multiple scenarios. It goes to my earlier point that it is probably more useful to have more scenarios if you are able to, because you can assess how the proxy model works in general rather than focusing on individual scenarios. It can be quite misleading to look at individual scenarios. From some of the work that we did in the working party, you could see that, whichever way you try to fit the models, you could still have individual scenarios that were some way out. If you avoided those scenarios in your validation, the proxy model could look better or worse than it is in practice.

In terms of multiple smooth scenarios, I do agree that it is not easy to pick those multiple smooth scenarios. One of the ways I have seen firms think about this is by considering different windows of the SCR scenario. They could use those somewhat out of cycle, for example, by taking last year's ones, or last quarter's ones, and maybe trying to construct some smooth scenarios out of that.

Mr Thomson: In relation to the trailing error point, I have not seen that done.

Mr Kuona: I have not seen it used in practice.

Mr Thomson: When we completed the paper, our key conclusion was that model use needs to be a key driver: what you are using the model for, the specific types of business, is fundamental.

Mr J. Dalmaris, F.I.A.: What standard procedures are typically employed to address tail risks? Have recent economic shifts affected the suitability of proxy models in accurately representing these risks?

Moderator: Have you seen anything in particular as a result of the mini-budget crisis last year, where we were observing market movements that were generally beyond the 1 in 200 percentile for certain risks?

Mr Kuona: In terms of tail risk, it probably goes back to the point around the range that proxy models are fitted for. Provided that that tail risk is an event that is anticipated in the domain of the proxy model, then theoretically you can get it to fit reasonably well. In general, proxy models tend to be validated around the 1 in 200 scenarios. What we saw for interest rates last year was, for many firms, well outside their 1 in 200 risk calibrations. Having spoken to a few companies, it is quite clear that this is an area that companies are looking at now. If the calibration range that companies used was informed by 1 in 200 events, then it is likely that the proxy models would not have captured that event. When you really move outside the range for which a proxy model is calibrated, especially for the more complex proxy models, then it is quite easy for those to not fit very well and to produce results that are almost nonsensical. I think that is one of the challenges trying to balance the accuracy over the range where you expect to be using the proxy model and the individual scenarios where your proxy model might end up being well outside of the calibration range. Those tail risks, like the mini-budget crisis, are probably cases in point where some firms may not have calibrated their models to handle such large stresses and would have needed to address that in some way.

Mr D. Crispin: I have two comments. One was on the tail risks. In my view, you should take this as an opportunity to stop and think about the underlying models. If you have very extreme events happening, you are not faced with a proxy modelling problem, you are faced with your underlying data and you should stop and think, "Is it valid?"

Second, on the averaging of losses to form the capital estimate, what has not been mentioned explicitly is the trade-off between bias and variance. You have a choice, and it is not obvious that one is better than the other. Ultimately, it has to do with how you are presenting and communicating your capital estimate. If you are communicating your errors and your uncertainty, you are free to make that choice.

Mr Kuona: I would agree. Our paper does talk about bias and variance, largely in the context of some of the techniques that we might use to calibrate proxy models. We mentioned the lasso technique for calibrating it, which is much more of an active trade-off that you are making in terms of calibrating the model. I do agree that at times it might be worthwhile having a little more bias to avoid having more variance, that is, to have slightly more reliable results that you know are a bit more wrong on average, rather than trying to have something that is completely accurate but maybe has a lot of variability in the output. That is largely something that you see in machine learning and so on. That is a much more active decision people are making around how they optimise their models.

Moderator: Can you foresee a point where technology has advanced sufficiently that we will not be reliant on proxy models?

Mr Thomson: We conclude in the paper that it will remain part of the risk management toolkit. I am not surprised by the thought that 10 years ago, one perhaps thought they were going to disappear, and they have not done so. Maynard (Kuona) and I were discussing before the session tonight, “When will regulators’ focus come back to these things?” We do genuinely think that some firms will have had plans to improve their modelling capabilities for the past few years. Some have probably followed some of those plans through. However, it seems that for the past few years, and probably for the next few years as well, there are possibly bigger fish to fry in terms of Solvency II reforms and other things that will come alongside that. The technology should be there, but there is a question about where this will rank in the priority list for firms.

Mr Kuona: As the ability to run heavy models more quickly develops, it can be tempting to think there might be a point where those models become the only models. In practice, there will always be a lot of demands on risk management functions and on people trying to understand what is happening within their business. A proxy model is still going to be incredibly useful. You can use it to understand very quickly what is likely to have happened to your business in the last few days. If you consider moving to heavy models to produce your overall SCR, I think the benefit of the additional accuracy that you might get is probably outweighed by the time required. I remember when, around 2012, some firms estimated that trying to run through an SCR calculation with a heavy model might take the equivalent of 45 years. Now, I think you could probably do it in a day with some clever modelling, but the proxy model will still be able to do that in minutes. Why would you want to wait 7 hours when you could have your results in 10 minutes? If you can get that speed in heavy models, you can probably get even more speed from proxy models. I think they will be with us for a while yet. Maybe in 10 years someone can ask me again.

Moderator: That concludes tonight’s presentation. I would like to thank all of our speakers and questioners for their contributions to a very interesting talk.