

Improving operational risk management with external loss data

Operational Risk & Regulation convened a panel, sponsored by **IBM**, to discuss how external data elements can assist risk managers and professionals in their operational risk programmes, particularly for informing scenarios and assessments, and also providing tail events for models

Can you provide a general overview of how you use external loss data?

Günther Helbok, UniCredit Bank Austria: We use external data for various purposes in our various legal entities. Specifically, we use external data heavily for modelling, for benchmarking and for performing and informing operational risk scenarios. But we use them both in a statistical sense and in the other sense of informing scenarios via, for example, Operational Riskdata eXchange Association (ORX) news stories. We also have requirements at the level of individual countries where legal entities are encouraged to contribute to local databases. So we have a local database, for example, in Italy and Hungary where country-specific data is shared, whereas group-wise data is shared via ORX. We encourage the use of those databases at the local level.

Jonathan Howitt, Professional Risk Managers' International Association (PRMIA): I've always found external data really helpful in prompting discussion with business management. Historically, business management may not always welcome the risk function; there may be a certain resistance because they feel that you might be getting at them. But if you can show them something that has happened at a peer group firm, or you can show them a loss that, in theory, they're exposed to, you can have a healthy discussion with them, because you can say 'if you don't think it would happen here, why not? Why are our controls better? Why is our management better?'

And there's now quite a body of external loss data that we can tap into. Ten years ago I'd have said there wasn't, but you now have a history of maybe 20, 30 or even 40 relevant events for each business area in a large bank or financial firm. You're looking at large tail losses, i.e. unexpected losses, which are, hopefully, very infrequent in your own firm, but there are enough observations of them in the market to be able to scale them and model them a little bit. And, because they're all categorised by industry, by risk type and by control failing, and so on, you are able to model the data as well and populate areas of your own loss curves that you didn't have any internal information on.

Do you need different details about a loss event to use it as a qualitative example rather than as a quantitative data point?

Jonathan Howitt: I'm not really in favour of these databases where all the

The Panel

Günther Helbok, Head of operational risk and reputational risk, credit risk validation and Basel compliance, UniCredit Bank Austria
Jonathan Howitt, Chair, Professional Risk Managers' International Association (PRMIA) Education Committee
Laura Polak, Head, Risk analytics, consulting engineering and content, IBM

facts have been scrubbed out. They are of very limited use to me. Anybody who tries to scale that data doesn't even know what they're scaling unless, behind the scenes, they really know what institution it related to. I think there is enough information out there that has been in the public domain that has then been classified and referenced properly, in what I call a qualitative database, and you can rely on the loss amounts. So I don't see a problem now in using a qualitative database for quantitative purposes as well. And, frankly, if I'm going to use a loss to ask 'if that happened in my institution, how might it affect us?', I'm going to have a much better shot at doing that if I have a full picture of the company that it happened to in the market place – because I would know their comparative size, I would know their type of business, and I would probably have a view of the strength of their management, their systems and their controls. So I'm a buyer of qualitative databases, rather than these offshore-held, data-swept quantitative databases that, frankly, lack any narrative.

Günther Helbok: I would agree, to the extent that we use them for informing and validating our scenario analysis. I think the narrative, the detailed story and the reasons for an event to happen are very important. We challenge the business to come up with relevant scenarios to really touch the profile of the risk, which would actually hit our institution. When I look at benchmarking, I am not so much interested in the very large, new scenario-type losses; there I'm interested in more low to medium losses, in order to really see what is driving the volume of my losses in a certain region or in a certain bank – say,

in private banking. So there, I actually see a value in databases that go down far more than what I would see in new stories, looking at losses of between €30,000 and €50,000, or €400,000 and 500,000, for example. And it's in this range that I see big value in benchmarking and doing statistical analysis. Of course, we do have the third use, which is regulator-driven, for the modelling where we will also need anonymised loss data. But I also see a need, especially for benchmarking, for the quantitative statistical data.

Laura Polak, IBM: I agree. I think there are needs for both. We don't deal in anonymised data; all of our data is either quantitative or qualitative, and all known. For the purpose of scenario analysis, I think that is absolutely essential. I actually believe more in the kind of very rich, qualitative analysis that tells people the control breakdowns, gives an idea of the culture, the environment and the factors that were involved in that loss. I think that is critical to having a better understanding – and also, when you're bringing it in front of your peers or the various business departments, to informing them and getting them to think really clearly about whether the event could happen to them. In terms of the anonymised data, from the quantitative modelling piece there is value to it. The one question I have always had is whether or not a certain amount of cherry-picking goes on. I've always wondered whether or not institutions are actually holding back certain events that they would not be able to report anonymously because of the loss amount or timing. That would be unfortunate, because it would take away from the value in the data.

Jonathan Howitt: That would be my strong suspicion as well. I think that, as Günther said, you can use what I would call expected loss data – less than €500,000 – for benchmarking. I don't see why people would be reticent about submitting lots of smaller losses. That's helpful to everybody. But, for the larger unexpected losses, I would need it with a name on it with a storyline. And I am suspicious of these anonymous losses – nobody's going to put a €6 billion loss in their external loss database.

Laura Polak: All data sets have a bias of some kind – but I think there's value in combining them, because they give a broader picture. And I think it does depend on the type of firm that you are; some firms may not get as much value out of those anonymous data points. Having the detailed analysis really does help to challenge the way people think. The other point is that people tend to have a certain understanding of their own world, so it's really important to show them those cases and how they happened so they can get a sense of the things that can come from outside that box.

Is there a temptation to look at the worst scenarios and reassure yourself by saying 'it couldn't happen here'?

Jonathan Howitt: I think, if anything, the opposite. As a risk person, I think there is a temptation to grandstand a loss at another company, and assert that yes, it could happen here. You're right from management's point of view; some of them may disingenuously say it could never happen here, or if it did it would be a much smaller loss. There's a tension because the risk people are probably going to say, 'yes, it could happen here', and management are going to say, 'no, it couldn't'. And that in itself is probably a healthy sort of tension to have in a discussion.

It's very difficult to say what is a relevant external loss. Every bank may have the same business lines in terms of products, but you don't share employees, you don't share systems. You may share market interfaces, but you won't always share clients. You won't necessarily deal with the same deal sizes. And there will be differences in timing and market volatility. There are many, many reasons why a loss that you might experience could be quite irrelevant to someone else.

Günther Helbok: In the central oprisk function, I use the stories to challenge people, and to develop the external loss events into something that can actually happen internally. And this is really what starts and informs discussion, and this is why those external databases are so useful. At the end of the day, it is not so much about whether a loss of so many million euros can happen, it is about what controls we have in place to avoid a similar loss.

The discussion about the amount comes at a very late stage. But it is the informed discussion beforehand that really adds value to the management of operational risks.

Another issue is whether to use one loss database or several – and, if you are using several, are you maybe giving up the guarantee that they're all comparable?

Günther Helbok: At UniCredit we have both local and central requirements for external loss data, because we are required to take part in very local data services, say in Hungary or in Italy, and also in more global databases. So, we are currently active members of the ORX for the capital modelling, also the ORX news service as well as the IBM Algo OpData, and I'm aware of two local consortiums where we employ loss data. The most important currently is the ORX service, which we use in the capital modelling function.

And are there issues with using data from several sources?

Jonathan Howitt: To compute our capital or to populate areas where data was missing, we used the case study material and we did two things with it. First, we verified the correctness of the severity distribution for losses that we were using; and second, we looked at correlations of losses, and also tried to understand to what degree operational risk might be cyclical – in the sense that, if there was a cyclical downturn in the economy, such as in 2008, did we see much higher levels of losses? – which, in fact, we did.

So we were able to use the case study material perfectly well, quantitatively. For modelling purposes, we were only really interested in losses of more than €30 million. Of course, we were an investment firm and therefore we weren't looking at what I call small, expected loss amounts in retail banks. But capital, after all, is about unexpected loss in large numbers. So I was perfectly happy with a case-study-based loss data.

Is that fairly typical of your customers' requirements?

Laura Polak: It is. And we do find that people tend to blend it in a way that Günther mentioned, for the purpose of the bank, and also the way that Jonathan describes it. Some of them use quantitative aspects of the data and some of them just use it to look for trends or correlation in the information for the purpose of refining their operational risk programme.

What kind of refresh rate would you think was appropriate for external loss data?

Jonathan Howitt: For management purposes, monthly. We would take the external losses to a monthly risk committee and occasionally there might not be a relevant loss, but if there were it would be flagged at the risk committee. There was a very strong appetite for anything to do with the regulatory environment and regulatory mood because fines by regulators and sanctions by regulators, even if there was no cost, were very important for management to understand.

Günther Helbok: You need to have certain key information very quickly to do ad-hoc analysis: On the other hand, the data will be used in a more pragmatic way, during the yearly scenario analysis, for example. But certain data points will be looked at once a year and analysed once a year. So, in our case, besides doing ad-hoc analysis for key events that we see outside of the bank, we would update our scenarios on a yearly basis. And, for modelling purposes, we have a half-yearly refresh rate for our external loss data.

Laura Polak: In terms of what we see of people using data points for modelling, it is more in line with what Günther is talking about. It tends to be more of an annual review. Some may look at it twice a year, with the final review done on an annual basis. And, typically, for the more qualitative data, they may review the data on a monthly basis and decide which cases or events are usable, but they may actually only do the work on a quarterly basis or every six months, with a more detailed analysis on an annual basis. And it depends on the type of business.

Is the rate speeding up?

Jonathan Howitt: That must be true, mustn't it? In the monthly risk pack we put a page of relevant external losses, and people were very interested, especially from the compliance and regulatory angle. There's enormous political and regulatory risk out there, a lot of which is operational risk. The pitfalls have become greater and greater. We've seen fines for mis-selling, money laundering, market abuse and the Libor fines. These represent a whole new realm of cost, compared to what it would have been pre-crash, and a totally different political mood. And the business wants to be on top of that. Half a year or a year might be good for a regulatory report or for modelling, but it's not good for real-time engagement with a business. So we always found a monthly process was the best way to interact.

Laura Polak: Many of my clients have monthly risk committees that they work with and provide with information from a variety of sources, not just external data or internal data. And there's quite a bit more regulatory oversight since the crash and this is something that people want to keep on top of.

How has that increased regulatory pressure – which we're seeing in a lot of areas – translated into different requirements for the use of external loss data?

Günther Helbok: UniCredit has seen clear regulatory guidance, especially on the model, regulators are saying what data to use and how. This is something fairly new, and I think several banks are seeing this pressure more and more.

The regulators are starting to 'prescribe' the use of external data in modelling, and are starting to check individual losses and have a say in classifying them, and in the way they are interpreted in the external databases.

Laura Polak: There's a lot more rigour around what gets included and why it gets included, and they want more detail and justification over the data points.

What is this increased rigour meant to achieve?

Günther Helbok: The increase in rigour is meant to make sure that I use external data properly and objectively. If we had developed the model 10 years ago, I could have used data from all around the world to fit, say, an asset management firm. Now the regulator would clearly look at where I am doing business and would ask whether the data fitted my risk profile. This level of scrutiny has significantly increased over the last few years.

Laura Polak: I think this goes back to what Jonathan was saying about how the data has improved in the last five to 10 years, and how there is so much more information available. So that has put the burden of proof onto the banks – since there is so much more rich data available, the regulators can ask for much more detail.

Günther Helbok: Also, regulators are now coming to us and simply asking: 'we have seen a high number of severe losses in this area in other banks in your country; how can you explain your difference?' So you immediately get into the details of explaining your data, and also explaining the external data points. So you need to develop a very thorough understanding. At the same time, the databases are growing continuously, with more years of data and also by reaching out to more and more banks.

Laura Polak: And with the proliferation of the internet and other sources, there's so much more information – regulatory, legal, and so on – available online.

Jonathan Howitt: When we've dealt with regulators on this issue, we actually haven't tried to scale individual losses. We've simply said that these are relevant losses, and we've made our own internal estimate of how much we might be exposed to losing, and that is informed by a list of relevant losses under that particular type of scenario. But we did use external data quantitatively for verifying the severity distribution and also some of the correlation assumptions that we were making in our model, and that was well-received by the regulator at the time.

I can understand that, for banks, there is a lot of data in what I call the expected loss zone – €500,000 and thereabouts – and there will be a temptation to bias more towards the expected loss. But I would say that the capital level is driven by unexpected amounts, and I would have – these days in banks anyway – a very high *de minimis* threshold for unexpected loss modelling – I wouldn't even start until I had a €10 million loss. At that point, it becomes meaningful for capital.

But that's my personal opinion. I suspect those are the debates that regulators are having with the banks now, on where the division is between

expected and unexpected, and what is relevant for capital versus what is relevant for good day-to-day management.

The case study material is going to be biased towards unexpected losses. Some of the insurance loss databases or ORX might have a lot of more expected loss-type information.

What role can automation play in handling external loss data?

Günther Helbok: I think certain rules can be applied to select data. The main concern is that they have to be objective, and they can be reproduced at any point in time. But it is really important that risk management and business management develop the skill of going through news stories and selecting the ones that are most appropriate and most relevant for them.

Jonathan Howitt: I'm not sure automation helps in this game. I prefer case study material because you are able to understand how relevant it is. If you simply scale some data points at arm's length, you lose rather a lot in terms of lessons learnt and in terms of applicability. So I'm not sure how automated this process can be.

Going back 10 or 12 years, there was discussion about having loss data approaches to operational risk, and there was an acknowledgement at the time that it might apply to things like credit card fraud, where you had heaps and heaps of data of a sort of expected loss type, and you were looking at how cyclical that might be.

But it's the big, unexpected stuff that's going to drive your capital, so I am not sure how relevant a loss data approach, even now, would be.

I don't believe that by automating the process you'll get more understanding. You might get a little bit more analysis, but you might miss some of the subtleties and relevance and applicability of the data. There's a little bit more art than science still in this. I'm not saying we shouldn't do the quantitative analysis, but I think if you only rely on that you're missing something.

Laura Polak: I see both sides of it, but I do agree with Jonathan. I think that you can't just rely on the automation; you do have to put a certain amount of review into it and time spent looking at it. I would also warn that it's important to look at other geographies and other similar-sized institutions. Subprime started out initially in the US, but then eventually moved into Europe. So, if you only take your lessons learned from your own backyard, I think there's a real risk that you could miss something that is coming your way. You can't leave that to automation.

How far is it safe to extend your view back in time? Does there come a point where you have to say that the industry has been through so many changes since then that we no longer think this kind of loss is relevant. Do loss data points age out?

Jonathan Howitt: Some definitely do, where systems change, processes change. Others I think are still relevant, but they just might manifest themselves in a different way. I mean, after Barings, people thought that was the last great rogue trade – after this, everybody will get all their reconciliations right. They'll get account ownership right, and we'll never see rogue trades again of this order, of this magnitude. And now Barings

looks positively tiny compared to what we've seen more recently, and we've seen rogue trading in an institutional way as well as solo rogue traders.

So I think there's some relevance to the old data, but not quite in the same way that it might have been relevant at the time.

Laura Polak: If you go back and look at how long Ponzi schemes have been happening, the outcomes are the same, but how you get there varies. So where you have control breakdowns, the type of controls that are broken may change, but the outcomes don't change.

If you watch frauds and how they actually happen, the criminals are constantly a little ahead – so it's all about trying to get into step with them to basically beat them at their game. I think that will always be a part of the business.

Günther Helbok: I think external data will continue to be very important going forward; also to develop a more forward-looking approach towards operational risk. External data has the tendency to give us a picture of the past, but we always need to see what the key risks are that we are facing in the future. And there, regulatory action is moving higher and higher on the list, if we look at the fines that many of the banks are incurring. Mis-selling is another area where it is very important to look back into history. We really need to keep these examples alive in our scenarios, to make sure that we don't make the same mistakes in five or 10 years' time.

Jonathan Howitt: There has been a big focus on cultural and business practice types of problems post-crash. For mis-selling, in particular, the penalties have become higher, and this is something that is now just not tolerated.

I think the regulators are only just starting to tap into market abuse, Libor being one example; anti-money laundering and money-laundering issues are another area in which the fines are just getting bigger and bigger. And I would highlight a couple of newer risks, where we've had a few cases – the exchanges and central counterparties could be the new systemic risk; and high-frequency trading feels to me like a very significant operational and systemic risk waiting to happen. We've seen a few near-bankruptcies from that, and I don't think they will be the last. So we can also use external data to think about how risks might manifest in the future.

Laura Polak: I'm also hearing concerns from a number of clients over some of the more systems-related risks; for example, denial of service attacks because so many companies are becoming a bigger presence on the internet, as well as the use of cloud services, loss of client data and privacy issues. We are seeing more variety in the types of risk emerging and I think external data and, basically, any information that the banks can gather for their operational risk programmes will help them to better equip themselves against these types of losses.

This article provides only a flavour of what was discussed during the webinar. To hear the entire proceedings, visit www.risk.net/2260870