The financial crisis – risk transfer, insurance layers and (no?) reinsurance culture

Michael Fackler

freelance actuary
Munich, Germany
michael_fackler@web.de

Abstract

The financial crisis of 2007 has triggered various debates, ranging from the stability of the banking system to subtle technical issues regarding the Gaussian and other copulas. All these debates are important, and it might be good to start even a further one: Credit derivatives have much in common with reinsurance, e.g. risk transfer via pooling and layering, scarce data, skewed distributions, and a limited number of specialized players in the market. All this leads to a very particular mixture of mathematical/statistical and cultural challenges. Reinsurers have been struggling to cope with these, not always successfully, but they have learnt many lessons during their more-than-one century in business. This has led to certain rules adopted by the reinsurance market and to a certain mindset adopted by the individuals working in the industry. Some of the traditions established in the insurance and reinsurance world could possibly inspire markets like the credit derivatives market.

Keywords

Financial crisis, GFC, reinsurance, risk transfer chain, model risk, heavy tail, skewed distribution, LMX spiral, incentive

1. Introduction

The global financial crisis of 2007-2010 (GFC), apart from the economic consequences, has created a lot of anger – to the general public it seemed that a lot of players in the financial industry were fraudsters or incompetent, or both. That might be true in some cases – but there must be deeper reasons for the crisis: It could be that in today's world we have much more uncertainty than we (want to) believe. If so, the tasks in the financial market are far more difficult than expected, thus even honest and knowledgeable people cannot always avoid big errors.

However, one should not take the easy way out by claiming that the crisis was totally unforeseeable. In particular the critical assessment of credit derivatives could rely (or could have relied) on plenty of past experience from another area – their structure is essentially non-proportional reinsurance and both markets have a lot in common: The risk transfer works via pooling and layering, the data available to evaluate the risks is typically scarce, the probability distribution of losses/defaults is typically skewed, risks sometimes change rapidly over time, etc. Thus the financial instruments aimed to transfer big insurance or credit risks are often very difficult to assess, i.e. their fair price is unclear. As a result, the number of players in the market is rather limited (at least the number of institutions that "rate" the transferred risks).

This situation leads to a very special mixture of mathematical/statistical and cultural challenges. Reinsurers have been struggling to cope with these, not always successfully, but they have learnt many lessons during their more-than-one century in business, and have above all found their way to live with a great deal of uncertainty. Some culture established in the insurance and reinsurance world could in the future possibly inspire markets like the credit derivatives market, however, it is just as important to understand in what way the markets are fundamentally different.

Many aspects of the GFC have been addressed by experts of various areas, covering topics ranging from very sophisticated mathematics (e.g. how to deal properly with copulas, see Donnelly & Embrechts, providing besides a comprehensive list of references) to regulation issues (e.g. how to enhance the stability of the banking system, see Hellwig, Sinn). The aim of this paper, apart from a close look at chains of consecutive risk transfer, is to draw the attention to the everyday situation of the individual practitioner in the financial industry, namely to the interaction of uncertainty and human behaviour: On the one hand it is often extremely difficult to find an adequate mathematical language for complex economic issues and to do reliable statistics about the limited data you have available, which makes decisions during the modelling process very delicate. Based on the (somewhat uncertain) results of these models, together with other (maybe somewhat uncertain) input, situated further in a very dynamical (thus somewhat uncertain) environment, the hard actuarial decisions are then followed by maybe even harder underwriting and management decisions. On the other hand the overall uncertainty makes it very difficult to appraise the quality of the decisions made, i.e. whether they were prudent and responsible or not. As we will see, the resulting lack of control (and self-control) tends to push decision makers somewhat towards the risky side.

Section 2 of this paper explains what risk transfer via credit derivatives and non-proportional reinsurance have in common. Section 3 describes the risk transfer chains that have evolved in both markets and the troubles they have caused, revealing among many common features some fundamental differences between the two markets. Section 4 deals with skewed distributions and scarce data, a combination challenging not just the quants (actuaries and alike) but also the management. Section 5 narrates from practical experience how certain uncertain deals come about and how quants struggle with various temptations they are exposed to at work. Section 6 concludes with lessons learned from the GFC – or maybe partly already some 25 years earlier.

2. Common features in credit derivatives and reinsurance

The common essence of the risk transfers by credit derivatives and reinsurance is: Diversify by pooling risks, reduce the pool's probabilities by layering. In order not to miss the subtle differences between the two markets we describe this well-known procedure in detail.

Pooling: Risks are aggregated together in order to create diversification. It is well known that if the risks are not totally positively correlated the outcome of the pool is less volatile than that of the single risks. The danger of greatly overstating this diversification effect has been discussed in the aftermath of the GFC (see Duffie, Donnelly & Embrechts). However, totally correlated risks are rare, thus pooling is in principle useful.

Layering: The aggregate risk of the pool is split by what is called non-proportional risk transfer. If the loss of the pool is X then the simplest option is to transform the risk into the so-called first risk min(X,d) up to a maximum d and the second risk or excess $(X\text{-}d)^+$. The former is "better" than X in terms of Coefficient of Variation (and other criteria), the latter is much more unbalanced than X but usually has a much lower probability of loss. This kind of risk may be preferred by financially strong market players as the premiums paid for such risks normally yield a higher profit margin. (Furthermore rare loss payments require far lower administration expenses.) Different risk preferences have led to the so-called layers (in banking the French word tranches is more common) where the excess is cut into slices in a way that is explained by the following example from industrial insurance, being very similar to reinsurance.

Assume a large property is to be insured against fire; losses could be as high as 50 (say million Euro). The owners are financially strong enough to cope with losses up to 1, so they need a capacity of 49 in excess (xs) of that, shortly written: a cover 49 xs 1. They find an insurer willing to write the risk but only being able to bear a maximum amount of 4, i.e. they get an offer for an XL (= excess of loss cover) 4 xs 1, which means that they must find further insurance on top of 5. Another insurer preferring layers having a quite low probability of loss steps in, agreeing on a 10 xs 5 layer but not more. To insure the whole risk a further cover of 35 in excess of 15 is needed, which could be placed with a specialized insurer offering large capacities but only accepting layers having an extremely low probability of loss.
In this way the risk is split into four parts: The first risk, usually retained (in the case of industrial insurance it could alternatively be placed as a so-called primary cover); a low "bottom" layer, being much more unbalanced that the risk itself but not on the extreme side; an intermediate layer; and finally an extremely unbalanced "top" layer. All four partners participate according to their risk preferences. (In practice the single layers could be proportionately split among various insurers, with some of them writing shares of several layers.)

In reinsurance (see Swiss Re) this basic layer structure is applied to a large variety of situations, according to how the loss is defined:

- Single loss from a single risk in the insured portfolio: Per Risk XL
- Accumulation of single losses caused by a catastrophic event (e.g. hurricane, earthquake): CatXL (Accumulation XL)
- Aggregate of all losses in a year (in a line of business): Stop Loss

In order to make layered (re)insurance more easily comparable to credit derivatives let us interpret these layers just as financial instruments. (Notice that in this paper we only treat traditional reinsurance – Financial reinsurance can have quite different cash flows.)

- The insurers of the top layer receive a premium every year and in most years they will not have to pay anything. Due to the low probability of loss the premium is only a small percentage of the cover. Thus they get a steady return, being however small compared to the cover (= capital at risk). In the very rare (say once in 100 years) case of a loss, in particular if it is a total loss, they will have to pay a high multiple of the annual yield. This cash-flow has much in common with investment grade bonds, which have a low annual return (being usually

defined as spread over the risk-free interest rate) compared to the invested capital, and in the extremely rare case of default the loss is as large as the return accumulated over decades.
- The bottom layer receives a premium every year, being a higher percentage of the cover due to the higher probability of loss. There might still be loss-free years, however, losses are not exceptional events. If we compare this to bonds, the most similar one would be a junk bond, having very high returns compared to the invested capital but on the other hand a considerable default risk.
- Intermediate layers are somewhere in between, in the same way as bonds having an intermediate security.

Now we explain credit derivatives (see e.g. McNeil et al., Duffie, Hellwig), namely the simplest case of an ABS (Asset Backed Security). This is created out of a pool of mortgages (or loans, etc.). What corresponds to the insured loss is the difference between the money (principal plus interest) that contractually has to be paid back by the borrowers and what is actually paid back, being less in case some of them default. The cash flow is different from insurance in that here the whole money at risk is handed over in advance and then gradually paid back during the credit period. We will see soon the importance of this difference.
The layering works as follows: The aggregate amount expected to be received from the borrowers is sliced into tranches, which are given different priorities for the incoming cash-flow. Three layers are common but there can be more. The bottom layer is called Equity tranche, it has the highest (nominal) return and the highest probability of default. The intermediate layer is called Mezzanine tranche (adopting a term for an intermediate storey). The top layer is called Senior tranche, it has the lowest return but the (estimated) probability of loss is such low that it gets an investment grade rating. The incoming cash flow from the borrowers at first goes wholly to the Senior tranche unless this has collected the whole amount due, then it is redirected to the Mezzanine tranche. Once this has received what it deserved the remaining payments (if any) go to the Equity tranche. This order of priorities ensures the desired order of probabilities of default.
Formally each layer is a separate bond, assigning to the investors the right to get the payments from the borrowers according to the defined rules.

When this kind of derivatives evolved it was largely expected that the originators would retain the Equity tranche (which would have been an effective incentive for them to carefully select the borrowers, see Hellwig). In the end it turned out that there were investors, e.g. hedge funds, willing to buy such tranches, just as there are investors buying unrated junk bonds.

If we compare ABS tranches to insurance they are closest to Stop Loss reinsurance treaties. The former “insure” the aggregate default from a portfolio of mortgages, the latter insure the aggregate loss of a certain portfolio of insurance policies in a certain period.

For completeness it shall be mentioned that in both areas, insurance and credit derivatives, there are various refinements of the explained layer structure. E.g. reinsurance layers have a variety of features working essentially as loss participations, credit derivatives may split interest from principal, etc. However, these features do not change the peculiar characteristics of tranches of different priority and probability of loss, so we can leave them aside here.

From an actuarial point of view, it is clear that while it should be not too difficult to calculate the net premium for a portfolio of insured risks or the adequate coupon for a pool of mortgages, respectively, by assessing the average frequency and severity of losses/defaults, far more is needed to associate the right portion of this risk premium to the single layers/tranches. This requires the calculation of the distribution function of the losses/defaults, i.e. far more complex models taking into account in particular the dependency of the single risks, which in turn might require far more data. We shall come back to this point several times later on.

3. Risk transfer chains

In order to optimize its risk profile it is straightforward for an insurer, or an investor in credit risks, to think about transferring a part of the taken risk to other parties. In fact this happens in both markets.

3.1. Structure

For the (somewhat diversified) portfolio of a primary insurer there is non-proportional reinsurance available, protecting against fluctuations due to exceptionally high losses (see Swiss Re). A reinsurer writes hundreds or thousands of such layers, creating again a (somewhat diversified) pool. For the protection of this pool the reinsurer can in turn buy a program of layers. Reinsurance for reinsurers is usually called retrocession. (There are specialized companies doing this kind of business, however, some reinsurers act as retrocessionaires, too.) A retrocessionare can in turn look for XL protection for the own (somewhat diversified) book of business, and so on. This results in a risk transfer chain where each step consists of pooling and layering as explained above:

Insured risk ➔ Insurer ➔ Reinsurer ➔ Retrocessionaire ➔ 2nd Retrocessionaire ➔ …

The same occurs in the credit derivatives market (see e.g. Sinn), and here in particular with the Mezzanine tranches. While it is not a problem at all to place top-rated tranches among the many investors preferring (or being restricted to) investment grade bonds, and on the other hand one often finds risk-loving investors buying Equity tranches, there is apparently no substantial demand for intermediate tranches. However, by pooling say some hundred non-investment grade tranches – maybe covering different countries and different kinds of underlying risk: mortgages, loans, credit card, leasing, etc. – and slicing this (somewhat diversified) portfolio into say three layers one can create new bonds, being called CDO (Collateralised Debt Obligation): The new top tranche should have a much lower probability of default than the average Mezzanine tranche, thus if the layer is chosen high enough it should get a top rating; the new bottom tranche is a further product with high risk and high nominal return suitable for another type of investor; and in between there is a new but now much smaller Mezzanine tranche. This tranche, if no one wants to buy it, could again be pooled with other ones, and so on. This results in a risk transfer chain where likewise each step consists of pooling and layering:

Mortgage ➔ ABS ➔ CDO ➔ CDO^2 ➔ CDO^3 ➔ …

In principle this is a good idea. Why not create tailor-made products having exactly the probability of loss that investors like most? However, one should bear in mind that each step requires the modelling of the aggregate loss/default distribution of a portfolio of risks being possibly not totally independent of each other. Even without regarding the details (see McNeil et al., notably being published before the GFC) of such modelling it is clear that any error (due to scarce data or whatever reason) will propagate from one chain link to the next and further along the chain.

However, some non-mathematical issues are also worth being mentioned.

3.2 Wandering defaults

See the following risk transfer chain:

A ➔ B ➔ C ➔ D ➔ E ➔ F ➔ …

What happens if D fails during the period the risk transfer is in force?

Here we have to distinguish the seemingly similar cases of insurance and credit derivatives. It will turn out that the different order of payments creates a fundamental asymmetry between the two markets.

In the case of reinsurance, if a loss hits various chain links and D is unable to pay its part this firstly matters to C: C reinsures B and has to indemnify B no matter whether it has reinsurance from D, whether D pays, or not. In other words, C bears the risk of a default of its reinsurer D – which should automatically boost the selection of financially strong reinsurers. Only if the failure of D causes the insolvency of C the parties further to the left (B, possibly A) are affected by D's bankruptcy. E and the further parties to the right are not affected instead; they will have to indemnify their part of the loss independently of the insolvency of D. In short, if insolvencies spread, they spread to the left.

Consequently, if you are a reinsurer thinking about extending a chain by taking and ceding risks in the described manner you have the incentive to look both left and right, i.e. to carefully select the parties to take risk from and to cede risk to, as bad events can come from either direction: From the left you assume (a part of the) underwriting risk, from the right you assume credit risk. This both-way cautiousness should generally keep weak players out of retrocession chains, ensuring the stability of such chains and keeping them generally rather short as the number of available market players being both professional and financially strong is limited.

Let us now regard credit derivatives chains. Recall in an ABS/CDO deal the whole money at risk is handed over at first before (most probably) being returned. Thus during the creation of the chain every participant pays a huge amount of money to his predecessor (acquisition of the bond), then over the years all this money gradually flows back from left to right starting from the borrowers through the whole chain. If now D defaults the payments to E are interrupted or reduced, which might create a domino effect to F and to the further parties to the right. In short, if defaults spread, they spread to the right.

This completely changes the characteristics of the chain. Like in the reinsurance case each player entering a chain should look left where the risk is taken from. But here you have no incentive at all to look right, i.e. to carefully select the people you are selling CDOs to: As soon as they have paid the bonds the only link to them is that they will receive payments from you, thus their financial situation does not affect you any longer. Summing up, in the CDO market bad events can only come from the left. There is no risk in the cession (except for maybe a reputational one), you can cede to whoever wants to take the risk, be they financially strong, professional, or neither of both.

What makes this situation yet a lot worse is the availability of potential risk takers: In hot market phases when the demand for investments is very high the originators of credit derivatives are able to transfer 100% of the risk they have assumed by selling all tranches completely at favourable prices to investors, be they credit market experts or not. In this situation a market player does not even have any incentive any more to look left, i.e. to carefully evaluate the assumed risks – these will anyway be repackaged and sold within a few days. In fact this largely occurred during the crisis (see Duffie, Hellwig).

The fact that most investors need to base their investment decisions on ratings by recognized rating agencies, together with the time pressure of the booming credit derivatives market, led to situations like the following, being reported by practitioners: While the originators would have been able to dedicate considerable time to the pooling and layering of credit risks, employing complex modelling tools, which was followed by the assessment on behalf of more than one rating agency in order to confirm the aspired rating, afterwards the newly designed bonds were marketed by brokers who gave potential investors not more than ten minutes (!) to decide whether to take or to leave the deal. In the end investors willing to buy this kind of product had no chance but to rely completely on the assigned ratings, i.e. the assessment of the investment was outsourced to the rating agencies without double-checking it by thorough in-house analysis. (It goes without saying that ten minutes are not even enough time to do the non-mathematical part, i.e. to read the documentation about structure and content of the preceding chain links.)

3.3 Cycles

As we have seen, risk transfer chains can be problematic on their own – but the really bad variant is as follows. The parties C and F could be … the same. This can happen to all players ceding and writing risks of the same kind. If they always read all the documentation about the risks to be assumed such cycles should be avoided, but it is clear that when having to check a long series of pools and layers some details can be missed and so inadvertently you become your own reinsurer…

Old reinsurance practitioners occasionally tell stories like this:

- You have to pay a huge loss: You get a 50 million Dollar cash call from one of your clients. You check your retrocession treaties. To your great relief a retro layer covers the part of the loss exceeding 5 million. Serenely you send a 45 million Dollar cash call to your retrocessionaire.
- However, to balance your portfolio you have written a bit of retrocession business. Some weeks later your retrocedent sends you a 38 million Dollar cash call, referring to the above loss event.
- If the retro business you have written is protected by your retrocession treaty you can prepare your next cash call and then wait: Sooner or later you will get a further cash call.

It is truly a challenge to calculate the final share of that loss event that you, and all the others involved, will have to pay.

What makes things even worse is the fact that chains are not just chains – every risk transfer step is a pool connecting hundred or more market players. It is merely a kind of spider's web linking everyone to everyone through chains and cycles. If further some of the people in this web go bankrupt things start getting really complex…

This is not just a funny story about some minor market players. As was even reported in the mass media, a considerable part of the GFC was exactly this: an all-embracing web of CDOs and other credit derivatives trapping investors world-wide, with defaults, downgrades, and liquidity problems infecting other players, and so on. (It must be noted that beyond layered risk transfer there were further critical links between players, e.g. commitments of liquidity assistance, see Hellwig.) To the general public this phenomenon may have seemed unheard of, but it was not the first big event of its kind: Reinsurers still remember well what has happened to the Lloyds of London market in the 1980s – it is called the LMX (London Market Excess) Spiral. Focusing on what matters for our topic the story was as follows (see O'Neill et al.):

In that period Lloyds in order to enlarge the number of players in their market ("Names") radically eased the access conditions, and many people became new members. At a certain stage there was a surplus of insurers. To supply the increased demand for business and to take advantage of the resulting soft-market conditions a lot of players began to retrocede a lot of their written business via XL retrocession. This process was strongly pushed by brokers – they got a considerable commission for any placed deal, whether it was a useful risk transfer or not. As long as no large losses occurred, everyone seemed to make money. Then the Piper Alpha oil rig exploded, causing a huge market loss involving many parties. It was followed by some more big claims, driving several Names into insolvency and making clear that the only ones having made profit from this business model had been the brokers.

Lloyds have survived their crisis, however, it was a huge task. They created an appropriate entity, Equitas, to manage run-off and correct allocation of the losses of that period.

3.4 Tail geometry

Coming back to modelling issues we want to draw the attention to a problem arising when layers protect (pools of) layers, being due to the typical geometry of loss size distributions. We explain this "XL on XL" situation with an example from insurance. Say a portfolio produces 10.000 losses per year. How many of these will exceed 1 million Euro? In practice it is often only a tiny part, say 10 per year, which is 1000 times less. What is the frequency of losses exceeding 2 million Euro? It could be 0.01 per year, thus 1000 times less than at 1 million. The frequency at 3 million could again be 1000 times lower than this, and so on – in the case of an Exponential distribution we would exactly have this situation. However, in practice in the million Dollar range one often observes much heavier tails than Exponential, sometimes leading to surprisingly low decreases of the loss frequency. It could be that 80% of all losses exceeding 2 million also exceed 3 million and most of these exceed 4 million, etc.

Assume there is a 4 xs 2 reinsurance layer protecting this portfolio, which is pooled together with 99 alike 4 xs 2 layers protecting other businesses. The portfolio consisting of these 100 layers shall be protected by a retro cover 200 xs 200. How much of the aggregate premium of the 100 layers has to be paid for the XL protection of the pool? (All premiums are intended to be fair net premiums.) The two extreme cases are the following:

- If the 4 xs 2 layers are stochastically independent an aggregated loss of more than 200 million Euro is extremely unlikely – to exceed this amount more than 50 independent layers need to suffer a loss at the same time. Thus the risk premium of the retro layer would only be a negligible part of the pool's aggregate premium.
- If the 4 xs 2 layers have identical loss distributions and if these are further totally correlated then the layer covering the pool suffers a loss if and only if each single 4 xs 2 layer suffers a loss greater than 2. In other words, if we artificially split each 4 xs 2 into two layers 2 xs 2 and 2 xs 4, the retro layer covers exactly the losses of the hundred 2 xs 4 layers, thus its risk premium is the sum of the premiums of these artificial layers.
  If in addition all the underlying businesses are very heavy tailed the 2 xs 4 layers will not be much cheaper than their 2 xs 2 counterparts, thus require not much less than half of the premium of the original 4 xs 2 layers. Hence the adequate premium for the 200 xs 200 retrocession can be in the range of 50% of the total pool premium – an astonishingly high figure, due to the missing diversification effect in this XL on XL case.

These are extremes, in practice we are mostly somewhere in between. However, it must be noted that although total dependency is extremely unlikely cases of strong dependency can produce similar results. Many people when seeing a layer xs 200 million Euro think it must anyway be extremely unlikely to suffer a loss. The above example illustrates instead that if heavy tails and strong dependency coincide the probability of such events may be hundreds of times higher than expected.

The extremely wide range of possible outcomes in XL on XL deals makes clear that there can be an enormous uncertainty in the model selection and the parameter estimation, which furthermore increases with the length of the chain due to the accumulation of XL on XL situations (see Donnelly & Embrechts).

Although model risk is being discussed more and more it seems that it is still somewhat popular to take any bad surprise for a random event ("very bad luck"): During the GFC some bankers claimed they would have experienced "several 25-standard deviation moves in a row" – an event as rare as winning the lottery in more than 10 countries the same day (see Dowd et al.). However, what they experienced was certainly a case of model uncertainty. In other words: The applied models were deadly wrong.

3.5 Workarounds

Reinsurers have been facing XL on XL situations and risk transfer chains for some time. In fact they have established some rules and traditions to reduce the likelihood of certain very bad situations:

- To begin with, 100% risk transfer is close to impossible in reinsurance (apart from special cases where the reinsurer is in control of the underwriting process). The philosophy is: If a company does not want to retain even a tiny part of a certain business they must be very afraid of something – whatever it might be, don't cover it.
- Reinsurers traditionally calculate the adequate premium for the risks they assume themselves, i.e. they hardly rely on premium ratings from external parties (apart from adopting some external expertise about issues like earthquake probabilities). They would typically be involved in the structuring of the reinsurance (determination of the retention level, etc.) or check at least whether the proposed structure makes sense.
- They routinely conduct accumulation control in order to limit their exposure to several named very large single risks and accumulation events, including traditionally natural disasters but nowadays also e.g. terrorism scenarios. That means that if a company notices they are about to exceed the given limit for say hurricanes in Florida they would stop writing Property reinsurance in the area – or look for an adequate transfer of a part of the risk.
- Reinsurance treaties traditionally exclude excess business, i.e. the reinsured policies must not be insurance layers. (Policies with small deductibles are admitted, though.) If excess business is not excluded total transparency is required: The insurer has to provide a bordereau, i.e. a list of the individual layer policies with all information needed to do a sound XL on XL rating.
- To avoid cycles reinsurance treaties mostly exclude reinsurance business, retrocession treaties may exclude retro business.

Of course there are situations where total transparency is unacceptable for the ceding party. Retrocession tends to be more "anonymous" than reinsurance. However, in the important field of Accumulation XLs for natural catastrophes there is a fair workaround: One can refer to figures from geophysical simulation models being known and accepted throughout the industry, which are able to quantify the exposure per geographical area without having to disclose the single risks.

And then there are situations where lack of transparency is knowingly accepted, maybe between parties having a long-term relationship including the commitment to let the other party recoup somehow in the years following an unexpected big loss. Alternatively one can find retrocessionaires of the more risk-loving side, asking extremely few questions, being aware that what they do is essentially gambling, and playing the game. It is possible to work which such people, however, one should know the rules. As practitioners put it: "You can be certain that this market will triple premiums as soon as it suffers a loss."

Apart from the exaggerations of the LMX spiral (and maybe some minor comebacks during soft-market periods) the typical risk transfer chain in reinsurance has not more than four to five chain links. People in the industry seem to be well aware that longer chains have such a high model risk that their added value is questionable.
In the credit derivatives market, though, there are rumours about of some far longer chains (see Sinn), having been created to cater for the huge demand for investment grade bonds in the years preceding the crisis.

Some of the measures developed by the reinsurance industry to deal with risk transfer chains might be not too difficult to adapt to markets like the credit derivatives market, in particular those preventing cycles and the transparency standards. A mandatory retention of a part of the assumed risks has been recommended by many experts (see e.g. Hellwig). What would arguably be much harder to address is the lack of incentive in credit derivatives chains to carefully select whom to cede risk to. Maybe it is worth thinking about whether it would be possible to implement the certainly very useful pooling/layering system in a more insurance-like fashion.

4. Skewed distributions and scarce data

Whether very heavy-tailed or not, distributions in insurance are typically skewed towards the "bad" end, and the same is true for mortgages/loans: Gains are rather limited, losses are potentially huge. To assess such distributions one typically needs more data than for a risk following a nice bell curve, and often you have less data available than you would feel comfortable with. The decisions to be made in such situations are not only a statistical problem, but also a cultural one, as can be seen from the following (fictitious but maybe not unrealistic) examples.

The interpretation problem

Imagine you are a quant (an actuary or alike) having to look at a business having a skewed loss distribution. Very simplifying skewed means that if you observe such a risk for say fifty years you see a very bad year, some more years are a bit negative, but most of the years yield very nice results.
A typical loss history to work with, however, would be the past ten years. (If older data exists you can often not use it as the former data base structure is incompatible to today's or because it is felt that the world has changed too much since then to rely on that data.)

Two situations are possible:

Case A: There is **no very bad year** in your data.

This loss history looks stable (maybe even almost symmetrically distributed). Modelling should not be too difficult.
You wonder whether the data is representative. Of course it is! Everyone agrees about that: The people having originated the business, the broker trying to place the risk transfer with your company, and your colleagues being keen on creating business volume.

Case B: There is **a very bad year** in your data.

This loss history is very unfavourable for modelling purposes.
Can this data be representative? Of course not! That bad year was without doubt a very exceptional event, having a return period of certainly more than 200 years. Everyone agrees about that: The people having originated the business, the broker trying to place the risk transfer with your company, and your colleagues being keen on creating business volume.

It is clear that doing a good job in such a situation does not only challenge the mathematical skills of a quant.

Speaking a bit more mathematically, if the loss distribution for a single year is skewed, so is it for periods of about 10 years. Of course the latter will be less skewed, however, it will still be far from a symmetric bell curve. That means that most of the typical statistics you come across when assessing such "skewed" risks will show a result being better than the long-term average. If in the rare cases when the loss history reveals such a bad year you ignore it (as an outlier) or assign an excessive return period to it your rating will be too cheap on average.

This problem is in principle well known and as a consequence people work with parametric models for loss distributions, being able to account for issues like skewness. However, in the practice of non-proportional risk transfer things are not that easy. Most statistical methods have been developed for large amounts of data. They can be applied to small data sets, too, but unexpected things may happen, see the following example.

The calibration problem

Imagine again you only have data from the past ten years available but need to calculate the 100 year event, i.e. the 1% quantile (maybe in order to rate a high layer or for a VaR calculation).

You start your powerful statistics software, providing a bunch of distributional models and being able to select the best fit according to standard statistical methods.

Tool output: Best fit is **Weibull**, VaR = **100** (say million Euro).

The day after you get updated data (slight changes due to run-off). You repeat the calculation.

Tool output: Best fit is **Lognormal**, VaR = **50**.

Later that day you get a new data update. (Someone found a typing error.) You rerun the calculation.

Tool output: Best fit is a **Pareto** variant, VaR = **150**.

What do you do then? If this result is not just the pricing of a minor deal but an important figure for e.g. solvency purposes, what do you tell management? They probably don't want to hear stories about best fit procedures yielding surprisingly unstable results.

The numerical effect of this example might be unrealistically high, however, it describes the uncertainty inherent in the modelling of skewed distributions with limited data. Sometimes such high deviations of the final result due to small variations of the data input even occur if one restricts the analysis to a single distribution model having few parameters. If you take the time for some sensitivity testing you might even come across counterintuitive results, e.g. while you thought you were passing to a more conservative approach your price becomes cheaper instead. And notice this may occur with just one-dimensional fits, being far simpler than the complex copula models (and alike) needed to rate risk transfer via pooling and layering in the credit derivatives and retrocession markets. It is clear that in such complex situations even 30 or 50 years of representative data cannot remove that uncertainty – model selection and parameter estimation will be highly sensitive anyway.

What makes things even worse is the dynamical environment both in the reinsurance and the credit market. Earthquake activity may be constant over long time periods, but other natural hazards, social and economic conditions are not. Thus many risks change quite rapidly over time. If you do not want to be restricted to statistical data of the very recent past you have to incorporate the changing environment somehow into the models, which makes them even more complex, requiring in turn further and more detailed data…

Whatever methodology quants adopt, when data is not abundant they will have to make some hard decisions. But they are not the only ones.

The management problem

Companies need to be protected against very bad years. That is the classical field where managers come across skewed distributions – if they don't do any risk management the economic result of the company will have a distribution being skewed towards the bad end. Essentially the management have the same problem as the quant when having to do a rating based on 10 years of data. The recent past in most cases has not seen any bad year, the remote past seems not to tell anything as "those were completely different times". Of course some plausible bad scenarios have been developed but no one can tell for sure whether these are 50-year events or far less probable.

One thing is clear: If the management, while the company faces cost pressure, have to choose between two strategies to reduce the risk of the enterprise, then the less effective strategy must appear very appealing – it will be cheaper (at least in 49 out of 50 years). In such a situation it is very hard to stay risk-averse and to enforce the safe option.

There is no point in playing managers and quants against each other. The uncertainties they have to deal with make it extremely difficult for both professions to do a good job. However, this is no excuse for negligence.

Reinsurers have developed a bit of culture to address the problems arising from skewed distributions and scarce data. Their advantage to capital markets is that a lot of their business relationships last for decades and involve several reinsurance treaties, being revised and renewed every year.
This firstly gives them the chance to negotiate a premium level that both parties regard as a fair long-term average even though they might not agree about every single deal in every single year. In a way they are able to smooth several uncertain premium calculations over time and lines of business.
Such long-lasting ties secondly lead to data series far beyond the past 10 years, which enable reinsurers to learn – slowly, however, steadily – over time which of their assumptions came true and which did not.
Apart from being this a base for continuous improvement of rating methods it might also have let reinsurers develop some intuition about which stories to believe and which not. So, if a huge loss has occurred in the past 10 years and the ceding company or the broker assures: "That will happen never again in this portfolio", they will remember cases when they had heard exactly that and a similar loss would occur just three years later. Such experience possibly creates the right portion of scepticism needed to avoid some of the very bad deals.

Generally, long business relationships help reinsurers to keep in mind the remote past. This is not only a matter of data availability – tradition helps, too. Reinsurance became a big market due to its decisive role in various rare events, starting with the 1906 San Francisco Earthquake (see e.g. Swiss Re). It is natural for the industry to look back very far and this attitude helps both the actuaries and the risk managers to be listened when they warn against bad events.

The calibration problem is arguably the most difficult to address. The word calibration is often used to describe a rather automatic and standardized statistical inference process. Such procedures make certainly sense if tons of data have to be processed, say for a weather forecast. However, the above example makes clear that the results of a statistical procedures are basically random outcomes, in case of scarce data very volatile ones. Powerful statistical tools cannot solve this problem. Reinsurers in many areas of their work have adopted what is proposed in various publications about Extreme Value Theory (see Embrechts et al., McNeil et al.): Inference of heavy tails is basically a step-by-step procedure requiring many intermediate decisions, employing successively various methods from deep exploratory analysis to parametric models (though using maybe non-standard estimators having good small-sample properties, see e.g. Brazauskas & Kleinfeld), and one should always test the sensitivity of the model assumptions and plausibilise the final results. It might be more difficult to implement such time-consuming procedures in the hectic placement of capital market deals, however, in the long run slow and careful calculations should be the cheaper option.

We will in the following see some consequences of uncertain calculations and their very random results, being partly far too expensive, partly far too cheap. Notice that this problem is not restricted to fancy businesses and over-complex risk transfer deals. Even in very traditional and professional markets there will occur situations where the data is far from sufficient to fairly rate a deal. E.g. new insurance lines come into existence, new insurers, too, sometimes even new economies or states. They need (re)insurance, loans and other financial services, just as business segments with excellent and reliable data.

5. Decisions in face of uncertainty and incentives

It is easy to make decisions, and to stand up for them, if you have plenty of information. In case of a lot of uncertainty it is much harder. Say you rate a deal and data is such scarce that you need to make an assumption (= to judgmentally choose a model / a parameter). Say you have two scenarios, both looking plausible, being not even much different from each other – and the more conservative one leads to a three times higher premium. Data cannot tell you which assumption is more realistic (and colleagues cannot either). In this situation there are pressures of various kind pushing you towards the cheaper option.

Here is an example, as practitioners in reinsurance tell it (but it could possibly have occurred likewise in other areas, too):

The deal placing game

If a “bad data” deal comes across some players in the market give it a chance and have a closer look at it. And then it is always the same:

- With so many hard decisions to make, time is never sufficient.
- People get hectic, lots of discussions and meetings.
- The frequency of the broker’s telephone calls is steadily increasing.
- At a certain point a market player (unknowingly and inadvertently) offers a share in the deal at an arguably too cheap price.
- The broker is happy, assigns that company the lead in the deal (say they take a 30% share) and attempts to place the remaining 70% at the indicated price.
- At this point other players, who were thinking about charging much more, hope that the leader know what they are doing, and take a minor share in the deal.

This is basically herding behaviour – someone being supposedly knowledgeable leads the way, and others follow. At the same time it is a variant of the well-known winner’s curse.

The hazard is obvious: If many deals are placed this way, the written business will have on average a too low price, and the market as a whole will loose money. But it is so hard to say no – after all the offered price is within the range of plausible assumptions. Furthermore, if you go for the cheaper pricing option no one will be able to prove it was a wrong calculation. And finally the fact that others in the market accept the deal is – a kind of rating classifying the deal as acceptable. Even if you are generally sceptical about external ratings, in such a situation they are difficult to argue against.

What you experience here is a combination of double peer pressure (by the people around you being keen on business and by the players having already accepted the deal) and lack of control of your work (due to the huge uncertainty). This is a triple incentive to stay a bit on the cheap side.

The appraisals of other market players have a particularly strong weight in markets like reinsurance, consisting of a quite limited number of active market players, which essentially “know” each other. (Of course there is a large number of less known minor players, however, these typically write following shares of deals after all terms and conditions are agreed on, thus do not influence the structuring/pricing process.) So if the proposed premium comes from a renowned company, how dare you say it is inadequate? “They must know that kind of business in that country.”

While the credit derivatives market has many more players than reinsurance, at least during the credit derivatives boom nearly all of them adopted a passive role (see Hellwig, Duffie): The assessment of the probability of default of the structured bonds was done by the extremely few big rating agencies (being supported on the way by the originators), and most risk takers would never question their ratings. How can you challenge the assessment of a big company being officially in charge of the rating of these bonds? Recall that data is neither sufficient to confirm their judgment nor to contest it.

Various variants of herding behaviour could be observed in the credit derivatives market. Investors followed rating agencies, which in turn followed … mainly a certain way to apply the Gaussian Copula model ("Li model", see Duffie, Donnelly & Embrechts), which for being easy and efficient to use became the most widespread method to assess default correlations, despite some inherent flaws (see McNeil et al., Whitehouse, both notably being published before the GFC), leading among other things to frequent underestimation of upper tail probabilities.
(Could it in contrast be that the somewhat low probabilities the model assigned to high tranches after all boosted its popularity?)

Sticking to what appears to be a market standard is tempting, particularly in uncertain situations. How many debates can you spare if everyone agrees about the methodology – you can justify what you are doing by simply saying: "It's state of the art." From a risk management perspective, though, this behaviour constitutes a concentration risk for the market as a whole.

A further kind of herding behaviour was investors' overall view of Senior CDO tranches. At a certain point it had become that much doctrine that these were the most attractive business opportunity of the decade that even sceptical investors could hardly stay away from them (see Hellwig).

Coming for a moment back to the above deal placing game, there seem to occur occasionally really bad variants, as practitioners report: E.g. it is suggested that the proposed premium was quoted, or at least accepted, by a certain renowned market player, while this is not true. In this or similar ways various reinsurers (or maybe rating agencies) happen to be played against each other – or, as an extreme case, various colleagues of the same department. People just shop around until they get the signatures they need…

All in all, there are various incentives to abstain from too expensive calculations:

- Human beings tend to go with the crowd. If you accept the deal everyone around you is happy (at least for the short term).
- An assessment without calculating many variants and asking many questions saves a lot of time. The more time you dedicate to a single deal the more other work will be waiting for you afterwards.
- Admitting uncertainty is uncool. If you try to discuss the case with colleagues, this might harm your reputation – further they are too busy to dedicate much time to your problem.
- It is even hard to admit to yourself that you are very uncertain, i.e. to acknowledge that all the skills you have acquired and all these hours of hard work are insufficient to carry out a certain task with the desired degree of professionalism. Humans feel much more comfortable when they believe they are in control of the situation, so in difficult moments they are somewhat vulnerable to deluding themselves.
- If you take the uncertainty into account instead and choose a more conservative rating variant, this might affect the business result assigned to your department for variable salary purposes. In other words, a rather expensive calculation could affect your bonus and/or – yet worse – the bonus of your boss / your colleagues.
- Due to the huge uncertainty it will be close to impossible to classify your calculation as wrong. This kind of deals generally has such a low probability of loss/default (high layers) that the outcomes in the near future will most probably be positive or at most slightly negative. In the rare case the deal produces a catastrophic result you will get away with just saying this was an extreme random result ("bad luck") and moreover totally unforeseeable.

It is very difficult to handle the problem that a bit too cheap calculations tend to create (for the short term) less trouble and less workload in the company. To assess the quality of the work of the quants one has to look at a lot of calculations together in order to smooth the random effects inherent in the single deals. In insurance this is called re-pricing – written deals are analysed later on to see whether the ratings were on average too cheap or too expensive compared to the actual loss experience. This is a complicated and tedious exercise, which might be easier to implement in the underwriting cycle of a business like reinsurance, having essentially one pricing season per year, than in the capital market placing deals throughout the year with a lot of time-pressure. However, if the rating of difficult business is not given sufficient time the variety of pressures illustrated above will push the rating results towards the cheap side.

Finally a word about a special case of gambling comprising a special temptation: It was reported that during the GFC some players made profit from the failure of others, in particular through speculation with Credit Default Swaps (CDS). These derivatives were originally designed to protect lenders and other stakeholders against the default of a certain institution, but can be traded on the capital market. Thus someone having no stake in the so-called reference entity of the CDS could buy the derivative – and later on be induced to accelerate the default of that entity, say by spreading certain rumours in the financial market, etc. (see O'Neill et al., Sinn).
This problem has been known to the insurance world for a very long time. E.g. someone could take out life insurance referring to the life of his neighbour. If this healthy neighbour then dies surprisingly early, was it a coincidence – or was it arsenic? To prevent the moral hazard inherent in such situations a principle was introduced centuries ago: To be indemnified by an insurance policy one has to prove an **insurable interest** – one must actually have suffered a loss.
It might be an idea to think about whether some areas of the capital market, e.g. CDS arrangements, could adopt this principle.

6. Conclusion

The comparison of the GFC with challenges met earlier by the (re)insurance industry might identify various approaches to enhance the credit derivatives market. At last this paper wants to focus once more on the situation of the individual practitioners in the financial industry – quants, managers, and others. Most of us want to do a good job and at the same time would like to be comfortable at work, despite the hard decisions we have to make. What can we learn from the crisis?

Probably we must develop a better feeling for randomness. Although we have learned how modern statistics deal with extremes and dependencies, it could be that our intuition is still somewhat stuck in the old times when the so-called "Normal" distribution was considered the "normal" case and fluctuations could always be "diversified away" thanks to (quite) independent risks.

When having to model about scarce data we should insist on getting sufficient time to quantify at least the impact of the critical assumptions we have to make.

We need better data and should make the effort to collect, structure, and maintain potentially useful information.

We must learn from history. Although the world around us is in continuous change there are quite some still useful lessons, learnt by our predecessors in the past centuries. It would be a pity if we repeated all those errors over and over again.

Rather than creating illusions about manageability, we should acknowledge that the modern world is far more uncertain than we used to, and would still like to, believe. Maybe in certain moments we are not able to handle more than 10% of the overall uncertainty and are nevertheless doing an excellent job – like a mariner successfully steering a small vessel through a storm.

Finally, we should learn to live happily in spite of all these uncertainties around.

References

Brazauskas V, Kleefeld A (2009) Robust and efficient fitting of the generalized Pareto distribution with actuarial applications in view. Insurance: Mathematics and Economics 45, 424-435.

Donnelly C, Embrechts P (2010) The devil is in the tails: actuarial mathematics and the subprime mortgage crisis. ASTIN Bulletin 40(1), 1-33.

Dowd K, Cotter J, Humphrey C, Woods M (2008) How unlucky is 25-sigma? Open Access publications from University College Dublin, http://ideas.repec.org/s/ner/ucddub.html

Duffie D (2008) Innovations in credit risk transfer: Implications for financial stability. BIS Working Papers, Monetary and Economic Department. Number 255.

Embrechts P, Klüppelberg C, Mikosch T (2008) Modelling extremal events for insurance and finance. Springer, Berlin Heidelberg.

Hellwig M F (2009) Systemic risk in the financial sector: an analysis of the subprime-mortgage crisis. De Economist, 157(2), 129-207.

McNeil A J, Frey R, Embrechts P (2005) Quantitative risk management. Concepts, techniques, tools. Princeton University Press, Princeton NJ.

O'Neill W C, Sharma N R, Carolan M T, Charchafchi Z (2009) Coping the CDS crisis: lessons learned from the LMX Spiral. Journal of Reinsurance, Vol. 16, No. 2, 1-34.

Sinn H-W (2010) Casino capitalism. How the financial crisis came about and what needs to be done now. Oxford University Press, Oxford.

Swiss Re (2002) An introduction to reinsurance. Expertise publication, www.swissre.com

Whitehouse M (2005) Slices of risk: How a formula ignited market that burned some big investors. Wall Street Journal, September 12 2005.