Key Excerpts from FCIC Staff Interview of Roger Stein

Roger Stein is a Group Managing Director at Moody's in charge of the Quantitative Research and Analytics Group. Beginning in 2005, Mr. Stein led the group at Moody's that developed a new model for rating subprime RMBS called "M3 Subprime".

On Wednesday, May 23, 2010, Ryan Bubb and Vic Cunicelli of the FCIC interviewed Mr. Stein of Moody's at the New York office of Moody's outside counsel, Sullivan and Cromwell.

The key finding in the interview was that Moody's calibrated the M3 Subprime model so that the levels of losses that it predicted matched the views of Moody's mortgage analysts. This was done by adjusting the parameters of the model that had been initially estimated based on historical data.

To generate the credit enhancement level needed for a bond to be rated Aaa, the losses predicted by the model were adjusted further. When a pool of loans was run through the model, the model first generated a set of 1,250 economic scenarios; each scenario predicted the losses on the mortgage pool (using the calibrated parameters described above). To calculate the credit enhancement level needed for a bond issued by the pool to be rated Aaa, Moody's then took the *highest* pool loss from among those 1,250 scenarios and multiplied it by a factor greater than two to generate a "stress" scenario. This factor was determined during a calibration exercise performed during the model's development so that the resulting losses in this worst scenario generally matched Moody's analysts' expectations for how high a loss to consider for some set of representative mortgage pools.

Below are the key passages from the transcript. The full transcript for the interview is forthcoming.

Mr. Bubb: You talked about calibrating the model so it was consistent with experts on subprime pools in the RMBS group. A natural way to do that I would think would be to calibrate it to the prior tool. But that wasn't the approach?

Mr. Stein: No.

Mr. Bubb: Why? Do you have a sense?

Mr. Stein: No, I don't actually know why that tool wasn't one of the tools that people wished to use for calibration. But I do know that we calibrated to a set of pools that a number of committees looked at carefully and felt were representative and for which the partners of ours within the RMBS team felt they had a good understanding.

Mr. Bubb: When you say calibrate, I think I understand what it means, but will you explain what it means?

Mr. Stein: Sure, it may be good to take a step back and just talk about models more generally. So I've been building models for many, many years, and almost every model I can think of is used, let me say it more precisely, I cannot think of a model in current use in finance that doesn't get calibrated in some way. Unless three conditions hold. I think those conditions need to be that the data that one uses to build the model is plentiful, descriptive, covers maybe several economic cycles. That one believes that one's model is in fact complete, and functionally correct, a complete description of the entire problem. And thirdly, that one believes the world will not change in any way materially from whatever the data and model development period was. I don't think that we believed any of those three things to be true. So, in such a setting, it's necessary to calibrate the model to something else. And one could choose many different things. Some people have advocated using for example CDS spreads, which turned out to be very, very tight during the period that we were developing the model. There are academic models that one could use. I suppose one could look at the estimates of analysts on the Street. We have our own analysts who have expert views, so we chose to calibrate the model to those views.Mr. Bubb: So what does that mean, to calibrate a model to the expert views? I mean, describe that in more detail what that means.

Mr. Stein: So, the model output that comes out is, let's say there's an average loss in the pool, call it the expected loss. If the loss comes out much lower than the analysts feel is appropriate, then what needs to be done is to change some of the historical parameters that are based on historical data to increase the risk such that the losses come out to higher levels that are consistent with what the analysts and our experts believe the true risk of those pools are. That was particularly true in the tail.

Mr. Bubb: What were the historical parameters that were used for calibration?

Mr. Stein: There were many. So a number of things were done. The prepayment baselines and sensitivity to home prices, to home price increases, was dramatically reduced. There were additional adjustments for different types of loans that made those loans riskier in the model than they would have been historically in the data. There were adjustments to the actual losses for a variety of different types of things, such as I/O provisions and so forth, documentation type was made riskier, a bunch of things like that. You can imagine a whole suite of these types of things. I don't recall all of them though.

Mr. Bubb: So I just want to make sure I understand this. So you've got a bunch of data which we're going to get to eventually, actually just to talk about what the data was, that provided a set of parameters, for example, you mentioned one of them was the prepayment sensitivity to increases in housing prices. The historical data would tell you that based on that period from which the data came, on average, at this sort of sensitivity, but experts in maybe the RMBS group, experts on subprime, thought that those sensitivities were wrong, that they were insufficiently, the historical data reflected insufficient sensitivity...

Mr. Stein: More generally, they felt that the losses the model was producing were not consistent with the risks in the pools. They were too low.

Mr. Bubb: I see. So this was not based on expert views on this intermediate parameter but rather views on the...

Mr. Stein: In many cases it was.

Mr. Bubb: Okay.

Mr. Stein: But ultimately, if the analysts' theory about a particular parameter was X, and by setting the parameter to X, the losses were not high enough, then we would make other adjustments as necessary for subprime mortgages during that period. Our analysis suggested that prepayment is a big driver of pool level losses, because if a loan prepays, it can't default. And so if prepayment rates are very high, even very risky borrowers leave the pool early. So by lowering prepayments we make the conditional probability of default, the conditional loss rate of the pool, much higher.

Mr. Bubb: I see, but you increase the tweak here, you increase prepayment sensitivity.

Mr. Stein: Yes. We increase prepayment rates.

Mr. Bubb: Sorry. Is it the levels or the slope?

Mr. Stein: Both. The prepayment baseline was decreased.

Mr. Bubb: Was decreased?

Mr. Stein: Yes, decreased.

Mr. Bubb: I'm sorry. And, what about the slope?

Mr. Stein: The slope was made less sensitive to positive home price appreciation.

Mr. Bubb: I see. And was, forgive me if I just, if you said this and I just forgot, was this tweak because of expert views on that prepayment sensitivity, or expert views on the ultimate answer in terms of expected loss?

Mr. Stein: The calibration of the model was a iterative process.

Mr. Bubb: I bet.

Mr. Stein: With members of my team, the RMBS team, and other groups, so we would meet weekly to discuss different approaches to doing calibration. I don't recall whether this specific tweak came from a modeler or whether it came from an RMBS analyst. But there was a fair mix of both I would say.

Mr. Bubb: And, help me think about this calibration, so, um, one view of what was going on here was that your experts are wrong, that the data say what the data say about ultimate pool losses based on historical experience and that these experts' views in fact are subject to various biases, we make mistakes, experts make mistakes, and that you were over-riding what the data were saying based on, sort of, idiosyncratic views. You know, why have a model at all, would go this view. What do you think?

Mr. Stein: I think several things, one is, I don't think *ex post* it's necessarily feasible to determine whether a specific view on a probability is right or wrong. I think probabilities by construction, unless they're zero, have some probability of happening. With a single draw it's impossible to say whether that probability was correctly assessed or not. With that said, I think that history bears out the, let me say this differently. Had we done as you suggest, had we done as you are recommending, for example, we would have had levels that are much lower than what ultimately were assessed by the analysts. Historical data was not capable of producing losses in the tail, particularly, that analysts felt were appropriate. More importantly, it couldn't even produce losses at mean that were sufficiently risky. I've sometimes wondered if we had had more historical data, whether that would have made a difference, but our recent experiments suggest that even using this last period of data, one still could not produce losses that are sufficiently large in the tail based solely on historical data.

Mr. Bubb: So do I infer from that, that essentially the losses that we experienced over the last several years are unprecedented historically?

Mr. Stein: Certainly the depth of the home price decline on a national level is unprecedented since the change of various banking regulations. At a national level that's true I think.

...

Mr. Stein: ... Maybe shocked is the wrong term, but it's certainly not the loss one would have gotten by looking at historical data. It's a much higher loss than that. Then, after we've done that 1200, I'm sorry, 1250 times, we can then have a distribution in some states of the world, a dip in home prices, I'm sorry, for some pools, a dip in home prices in Year Three might be bad, and for in others a dip in home prices in Year Two might be bad. So whatever that last path is that generated the worst loss, that is then shocked again, and that's pushed out into the tail to make it several times, a couple of times, some number of times worse than what it would have shown up in the simulation by itself.

Mr. Bubb: I'm, forgive me, still a little confused. We've got, I was with you when we had our 1250 draws, each of which had already, each of which represented an expected pool loss for that economic scenario. And we shocked it, based on what you were calling before the shock to the historical simulation. And now, I'm sorry, just explain it again, the second shock is applied to which of those draws?

Mr. Stein: So we take the worst case.

Mr. Bubb: The single worst case?

Mr. Stein: The single worst case under the simulation for that particular pool.

Mr. Bubb: Meaning, of those 1250 draws, this is the worst of 1250 draws.

Mr. Stein: Yes exactly.

Mr. Bubb: And that one we're going to...

Mr. Stein: Shock again.

Mr. Bubb: I see, and that's used how?

Mr. Stein: That then becomes the basis for the triple-A enhancement level.

Mr. Bubb: And it's the case that the triple-A enhancement level is set so that under that scenario so shocked out, losses just come to below the triple-A. Is that basically... This is for the model output. I understand that the committee has to do things.

Mr. Stein: There is, there... The model isn't the same as the rating.

Mr. Bubb: Of course. Yep, yep, yep.

Mr. Stein: And so, there's a credit enhancement level that is for the collateral of the transaction. All of the subprime transactions that I'm aware of had cash flow waterfalls as well. There's a separate analysis that needs to be done on those cash flow waterfalls that is not part of this model. So this was a starting point for the enhancement levels. There was additional analysis with stress testing and things that was then done, using some of this output, to arrive at whatever the final enhancement levels would be for a transaction.

And that would also contemplate various types of legal or structural issues as well as behavioral issues of the servicer, and so forth.

Mr. Bubb: But is in general sort of the right way to think about this, that worst draw, the 1250, after being subject again, this is another shock, that that outcome is modeled in the cash, whatever is being done to model the cash flows to the actual bonds of the deal. And that the principle is, the triple-A ought to not be hit in that scenario.

Mr. Stein: I believe a way to say it would be that the analysts would look at that output, make adjustments to that output, and then do what you have described.

Mr. Bubb: The adjustments would be based on the legal issues...

Mr. Stein: Those adjustments might be [inaudible] to the collateral-based adjustments, so they would typically be based on thoughts about the collateral, and servicer, and these types of things.

Mr. Bubb: So how did you figure out how big to make that shock on the last, the worst draw?

Mr. Stein: Again this was a fairly long process. We made adjustments such that the levels were consistent with what analysts and other experts in the RMBS team and more broadly within Moody's were comfortable sufficiently represented the risks in the tail of the transaction, of the population of transactions. It's important, I think, to note that the historical data provided no guidance in this regard, so you might think of it as a Bayesian type exercise, if that works as a way of thinking about it.

Mr. Bubb: So, we're really using the analysts' current beliefs about the tail events to drive the ultimate assumptions used in the model?

Mr. Stein: That's correct. One of the, it wasn't just the analysts, again, I'm not sure how familiar you are with Moody's processes, but these were not just opinions of analysts, they were long committees that would discuss these issues and so forth for resolving them. And so I believe it was the case that these represented the best views of the organization with respect to those tail losses primarily again because the historical data provided no guidance in this regard.

Mr. Bubb: So, that's interesting. So the model in some sense is the tail, the analyst is the dog, the committee is the long... to use a metaphor.

Mr. Stein: I don't want to use that metaphor. I think we have a model that must contemplate events for which there is no data. It's not clear how else one might calibrate that model, besides using expert judgment.

Mr. Bubb: What was the point of having the model at that point? I mean, another way to go here would be, look, we've got analysts who have debated these things for a long time, they're good Bayesians, they've updated based on the latest greatest events in the housing market, they can just rate deals with their own gestalt assessments of the state of things.

Mr. Stein: I think that there are actually, I mean there are several benefits for the model. One of those benefits was transparency. The goal of the model was to provide market participants, investors, issuers, others in the market, with a better understanding of the factors that drive Moody's ratings. In addition, while, while the committee view that was used for this calibration may reflect a general view, on an individual transaction, that committee's view is not available, typically. And so, for an individual transaction, you need to be able to rank the different transactions that come in to be rated based on the general expectation about losses in general on the market. The expectation about losses in general is what's used for the calibration. For individual analysis, an analyst will take to their committee an individual transaction which may be very different than the general view, the general market. So the purpose of the model is two-fold: to provide transparency and to provide consistency in the ranking of the different transactions.

Mr. Bubb: A related question. So in that view, one could have a transparent model, a model that performs the role that you described, both the ranking and the transparency role, without using any historical data. So you could literally have a set of assumed coefficients that would output risks for a loan that had been tuned so that the ultimate losses of the pools matched expert judgment, but that didn't begin with what one might characterize as a contrivance, of using historical data, and generate, and achieve the same two functions, of transparency and ordinal ranking of pools that you described. What am I missing, or am I missing...?

Mr. Stein: I don't think that this is in any way a contrivance. I believe that to do as you propose would be extraordinarily difficult because of the interactions of the different relationships of prepayment, default, and severity within the model. So I think that the historical data provides tremendous guidance on the shapes, the relationships of different behaviors, but may not provide sufficient guidance on the levels of those behaviors. So the shape of the prepayment curve, for example, is a complicated function of the structure of the loan which is probably difficult to just come up with gestalt I think is the word you used. On the other hand, moving that up or down will affect the levels of prepayments, and so while the historical data may provide good guidance on the relationships and the shapes of things, it is probably given the historical record going to provide optimistic levels of the model.

Mr. Bubb: On that note could you describe a bit the data that was used?

Mr. Stein: Certainly. At some point I'd like to take a break...