

Model Validation Need Not Be a Blood Sport

By Tony Hughes

September 13, 2019

My team did a big validation project for a financial institution a few years ago. We were actually the backup, external validators called in to resolve a disagreement between the model build team and the internal validation group.

The dispute was rather interesting.

The build team was stacked with high-caliber analysts who used a clever mix of relevant academic papers and their own guile to come up with some very interesting research.

In the meeting we held before kicking off the project, I told my team to have the courage to pass the model unless it was obviously flawed. You can always find something to quibble with, something you would have done differently, and human nature makes it far too easy to pull out the “reject” stamp. I wanted my team to be better than this.

The challenger model we built was certainly different from the internal model (otherwise, what's the point?) – but not necessarily much better when it came down to brass tacks. We could have argued that ours was better, but so could they, so we decided to pass the model.

The internal validation team ran the usual set of diagnostics on the initial model, and were able to reject a few null hypotheses. They issued it a failing grade on this basis. Our report highlighted the more novel aspects of the problem at hand and the reasons why many of the conventional tests were not relevant.

Ultimately, I think this process was a narrow win for model risk governance. But it was expensive, time consuming and morale sapping for all the analysts involved. I've worked on both sides of the validation divide and can tell you that these interactions are never pleasant.

Flaws of the ‘Fox-and-Hounds’ Approach

If you are the model builder, you always feel like the validator is nit-picking – finding fault with all the small decisions you face when building any model. You feel that you could win a fair debate with the validator, but that you will never get a chance to rebut their arguments in front of a disinterested judge.

As the builder, you are the fox being chased by a relentless hound, and the best you can hope for is for your model to survive.

As a validator – a role I have only filled in an external capacity – the psychology is perhaps more fraught. You are required to find problems with the model, but what if there aren't any? Will the validation exercise still be viewed as serious if no shortcomings are found? When new validators are engaged, it should be stressed that, if the model is sound, a passing grade accompanied by a positive validation report is a perfectly reasonable outcome.

It's worse if the model you are assessing is truly terrible. A negative report can be difficult for the client to swallow and the validators can be cast as villains, their report disregarded or even derided.

If the internal model builders are new or not held in high regard, the negative validation report may put their jobs in jeopardy. I'm sure some validation hounds love to tear a weak fox to shreds, but I am not one of them. I hate to fail models, but have done it when a passing grade would be impossible to defend in front of the people who are ultimately paying the bills.

So, is there a better way?

The Competing Models Strategy

Instead of commissioning one team to build a new model, imagine for a moment that we engage two. These teams would be independent of each other but would have the same mandate: to build the best possible model for the task. On day one, both teams will believe that they are building the champion model.

When the builds are complete, the team that better achieves the task would take the role of the champion and write the core model documentation. The team that loses would then write the validation report, using their vanquished (but hopefully worthy) model as the primary challenger. The debate between the two teams, part of the process used to identify the winner, would form the guts of the final validation document.

The psychology built into this structure is much more constructive than the traditional "fox-and-hounds" model. Before either team is even aware of their respective roles, they will push each other to build better models.

After the roles are allocated, the validation team will have direct recent experience of the challenges faced by the build team, and thus should be more sympathetic. The build team, meanwhile, will not fear the validators' wrath – if they could do better, they would have done so back when the game was afoot.

I think that the traditional validation model is built on a somewhat shaky premise. It asks whether a given model is valid, but all models, by definition, are ultimately invalid when their premises are tested in reality. A better approach involves seeking the best possible invalid model and casting the net wider to consider two – or 10 – different challengers. You then validate the entire process, rather than a particular specification.

Such an approach provides statistical benefits, and also a more relaxed and focused group of analysts throughout the organization.

Tony Hughes is a managing director of economic research and credit analytics at Moody's Analytics. His work over the past 15 years has spanned the world of financial risk modeling, from corporate and retail exposures to deposits and revenues. He has also engaged in forecasting of asset prices and general macroeconomic analysis.