How The False Positive Myth Encourages Collateral Damage

Relying solely on false positives when gauging a risk score can lead to serious problems.

By Dr. Mark Fleming

The concept of “collateral damage” is defined as damage inflicted on unintended targets by virtue of their proximity or relationship to the intended target. While this term is most often associated with battlefield conflicts, not the mortgage loan industry, in the war to identify fraud and performance issues prior to funding a loan, it appears that collateral damage in the form of “false positives” may be just what is occurring. A false positive is a loan flagged as high-risk when in reality it is low-risk.

While loan officers complain that they get too many false positives and the vast majority of loans that get flagged with a high-risk score will perform and perform well, new evidence suggests that ranking and then contemplating the risk of some loan files relative to other files can help streamline decision-making with respect to false positives. Let’s take a closer look at how this is possible.

Measuring mortgage risk involves understanding and stratifying risk that is derived from a variety of factors. The obvious factors that come to mind are the riskiness of the borrower, the terms or structure of the loan, the quality and sustainability of the collateral, and the riskiness of the third-party mortgage contract participants (broker, appraiser, loan officer, closing agent, et al). Scoring the risk of all of these factors together allows one to rank order loans from highest to lowest risk as a function of these risk factors. A risk score ultimately indicates the likelihood that a loan will become delinquent and experience loss severity that is higher than a loan with a lower risk score.

For example, loans with the highest FICO scores and lowest loan-to-values (LTVs) sometimes default, and many of the loans with the lowest FICO scores perform. The odds that these loans will perform are very different. Mortgage lenders use the tools at their disposal (risk scoring) to identify the odds that a loan will perform based on the risk characteristics that can be observed. The better the tools used to stratify mortgage risk, the better the odds of correctly identifying loans that perform from those that don’t.

For any risk score, there is an associated likelihood that loans with that score will perform. The better the tool generating the risk score, the closer the predicted odds are to the observed actual odds for that score. In other words, if a particular risk score identifies that three in ten loans with that score will experience serious delinquency, then we would expect that the actual rate of seriously delinquent loans with that score is something close to three-in-ten. Ideally, the higher the odds for a given score, the lower the false positives.

Playing the numbers

Again, an example may illustrate this well. If the odds that a loan will default are three in four, then the false positive rate is one in four. If the odds are two in four that a default will occur, then the false positive rate is also two in four. It is expected that the risk score should be higher for the group of loans with a three-in-four odds of defaulting versus the risk scores for those loans.

Dr. Mark Fleming is director of economics for First American Core Logic Inc., a Santa Ana, Calif.-based provider of residential mortgage risk management and fraud protection technology and services to the U.S. mortgage banking industry. He can be reached at mfleming@corelogic.com.

Copyright © 2007 Zackin Publications Inc. All Rights Reserved.
with two-in-four odds.

Therefore, risk scoring will always generate false positives, and the odds associated with the risk score indicate the false positive rate. But this assumes the risk score is treated as a binary indicator of an adverse event, which was never the intent. The real question is whether the risk scores are more efficient (generate fewer false positives or have increasing odds ratios) as they indicate higher risk.

How efficient a risk score is will be, in part, dependent on how the lender chooses to define performance. Each lender has its own basis for the definition of a bad outcome, such as a 90-day delinquency, first payment default, early payment default, REO, etc. Under this pretext, a false positive is a loan that is not meeting the chosen definition of a bad outcome but has been scored with a risk score above some threshold.

But remember, the risk score is not stating definitively that when a score is above some benchmark value the loan will not perform. Instead, it is saying the odds are higher that it will not perform.

There is a way to structure a test so that the odds can be measured against the risk scores to determine the overall efficiency of the risk model. The faster the odds increase as the risk score rises, the more efficient the model is and the lower false positives will be for any given score cohort. This is accomplished by taking a group of properties for which there is known loan performance data and retroactively scoring them.

In other words, score them based upon information that would have been known at the time the loan was being originated. The act of retrospectively scoring the loans is done to replicate how the score would be used in the production environment when the decision to fund the loan is being undertaken. If the risk score is predictive of the measure of adverse loan performance decided upon by the lender, the odds that a non-performing loan is identified will be higher the higher the risk score. The odds are calculated for each score value by counting all the “bad” loans relative to the total of “goods and bads.”

**Riskier business**

For example, if a low-risk score had 100 loans of which ten were bad and 90 were good, the odds ratio would be 10% (10/100). A high-risk score may also have 100 loans of which 30 were bad and 70 were good. In this case, the odds ratio is 30% (30/100). This clearly shows that the score is indicating an increasing likelihood of the bad event, but it is never 100% certain.

While the risk score is objectively measuring the riskiness of the loan, the lender’s opinion of false positives will be based on its individual determination of what constitutes a bad loan and where it decides to draw the line in the range of risk score values.

Consider again the high and low FICO loans discussed above. It is a well-established principle in mortgage lending that the odds the high FICO loan will not perform are lower than the odds the low FICO loan will perform. So if the odds are calculated on a pool of loans as described above for different levels of FICO scores, false positive rates can be constructed for any FICO group. The findings from this exercise demonstrate that FICO generates false positives like any mortgage risk model, but we often forget that it is a risk model with these characteristics. Yet loan officers never complain about FICO scores generating too many false positives!

Because the risk scores can stratify or rank order the odds that one loan will experience adverse performance relative to another, they can be used to dictate the necessary level of underwriting due diligence. In fact, the utilization of risk models has been one of the primary factors driving the growth and efficiency of the mortgage lending industry over the last ten years. Before the wide-scale adoption of risk models, every loan was underwritten manually. The use of risk scores in automated underwriting engines has enabled the industry to more quickly and efficiently underwrite many loans.

Risk scoring within automated underwriting helps distinguish those loans that should be underwritten minimally because of the odds of adverse loan performance based on the risk scores from the loans that need extensive manual underwriting because the odds of adverse loan performance are much higher. The refinance booms of recent years have been made possible in some part due to the advent of these underwriting efficiencies. Furthermore, the use of risk models has facilitated the liquidity of the mortgage market as participants in the market are able to more objectively measure, rank order, and stratify risk, which in turn speeds up the rate at which portfolios of loans are securitized.

The myth that only a small percentage of those loans flagged as high-risk go bad is just that, a myth. Loans flagged as high-risk are more likely to have adverse performance than those that are low-risk. False positives are red herrings when gauging a risk score.

All risk tools produce false positives - the litmus test for such a tool should be its ability to rank order risk on a relative basis.