Metrics Matter

By Douglas Gordon

Lenders are greatly increasing their use of Automated Valuation Models to confirm property value and to streamline the lending process, by using the AVM as a property value screen in home-equity lending. As the market has grown, so has the number of commercially available AVMs. Each AVM has different performance characteristics, geographic coverage, and confidence scores. For the most powerful risk management, lenders must understand the statistical performance characteristics of the AVM they use. That also means understanding Forecast Standard Deviation. The benefit of FSD is that it statistically links the confidence scores of an AVM to the AVM's forecasts of property value. Just how important is FSD, and how do you determine if your AVM supplies it accurately?

FSD is the best measure of expected accuracy of an AVM value in actual use. FSD is an AVM value's expected (forecasted) proportional standard deviation around actual subsequent sales price for the given property value estimate. Sometimes FSD is called sigma. But some AVMs may use the generic term sigma to refer to other measures of standard deviation, such as the standard deviation of coefficient estimates. These measures are not forecast standard deviation and are not as reliable a measure of expected performance of the AVM value.

FSD measures how accurately the particular AVM value estimates value for the specified property.

FSD measures how accurately the particular AVM value estimates value for the specified property.

The lower the forecast standard deviation, the closer the AVM value will be to the actual sales price. FSDs are generally reported as proportional standard deviations rather than dollar standard deviations. Percentage error is often used to approximate the proportional error used in the calculations, and we do so in this article. Using proportional standard deviations rather than dollar standard deviations, the forecast accuracy for a $100,000 house can be more easily compared with the accuracy for a $300,000 house. For forecast errors (proportional difference between AVM value and sale price), about 68% of sample observations will fall within +1 standard deviation of the estimate. For example, if FSD for property estimates are 10%, about 68% of actual sale prices will fall within +10% of the AVM values.

AVMs typically report a confidence score associated with their value estimates. A confidence score is a measure providing information about the accuracy of an AVM's value. Some AVMs, such as Freddie Mac's Home Value Explorer service, use FSD to determine the reported confidence score. Using HVE as an example, a high confidence score includes FSDs of 13% or less, while a medium score includes FSDs between 13% and 20%, and a low confidence score includes FSDs greater than 20%. Some AVMs do not tie their confidence scores to FSD, instead basing scores on unreliable factors such as number of local properties used in the model's estimate, neighborhood range of values, or some other measure that does not correspond precisely to expected AVM performance against sale

(A,B,C,D,E, etc.) or numbers (1,2,3,...,100) that are harder to compare. Such a confusion of scores and lack of connection to statistical performance in actual use forces lenders to guess at their risk management.

For lenders using confidence scores rather than FSD directly, it is important to verify that an AVM value's confidence score corresponds to an FSD or range of FSDs. Lenders are concerned with the expected accuracy of the AVM in actual use. When some AVM vendors describe a "standard deviation" associated with their model, they may mean that it is a measure of the fit of the model coefficients when the model is created, not the expected accuracy in actual use. Lenders should verify that a reported model standard deviation is truly an FSD and that the reported FSDs match standard deviations in their tests -- that the FSDs are reported accurately by the model.

The best way to validate confidence score or FSD is by performing an out-of-sample test of property valuations.

For an out-of-sample test, the model should be applied to property sale transactions that are still in process or so recent that they do not appear in public records (for example, loans closed in the last 30 days). When testing an AVM a lender should validate not only that the model is accurate and the differences from the sale price are small, but also that the FSD and confidence score are reliable.

The best way to validate confidence score or FSD is by performing an out-of-sample test of property valuations.

Here are questions to ask your vendor, or to verify with a test:

1. Do confidence scores correspond to FSDs? How do they correspond? Do tests validate the relationship?

2. Are about 68% of the AVM values tested within 1 standard deviation (reported as the model's FSD) of the actual price? For example, if FSD is 10%, does an interval of +10% cover around 68% of the deviations of estimate from sale price? Similarly, does an interval of +20% cover 95% of the deviations?

If the lender is unable to test against property sales and must test against appraisal values, the lender should expect a less strong correspondence of FSD to actual difference from appraisal value. In addition, the model value estimates will typically average a bit below appraisal values. A lender should verify that the AVM provider regularly validates its confidence scores and FSDs using out-of-sample tests with recent property sale prices, to ensure that the measures correspond to customer experience. Ideally, the lender will perform its own tests to validate confidence scores and FSDs.

A reliable forecast standard deviation (or a confidence score based directly on FSD) enables a lender to develop an integrated credit and collateral policy.

A reliable forecast standard deviation (or a confidence score based directly on FSD) enables a lender to develop an integrated credit and collateral policy.

The FSD allows a lender to calculate probability of loss due to collateral deficiency and the expected size of any loss. Thus, the lender can determine risk and evaluate costs and benefits of different property valuation programs and policies. Because the FSD of individual property value estimates is supplied by the AVM provider, out-of-sample validation by the lender is important. If an AVM does not provide FSD, then a lender's risk management is more difficult.

Here is a way that a lender can test forecast standard deviation. Ideally, each AVM value

supplied will have an FSD. Using the standard deviation supplied with the AVM value, the lender can calculate intervals around the AVM value estimates. Two intervals that are useful to test are 1 standard deviation (a 68% interval), and 1.28 standard deviations (an 80% interval).

For example, calculations to create a test based on a 1 standard deviation interval can be done as follows. If the standard deviation is 10% for one read, the lender creates a range +10% from the AVM value. Recall that percentage standard deviation is an approximation for proportional standard deviation, so this test is approximate. If the standard deviation is 15% for another value, the lender creates a range +15% from that AVM value. Next, the actual sale prices are compared with the AVM values and the differences are calculated. The lender counts the number of times the sale price fell within the 1-standard-deviation range created around each AVM value, then divides by the total number of AVM values, and multiply by 100. This will give the percentage of the time that the AVM value comes within 1 standard deviation of the sale price. The lender then compares [to see] that this percentage is at least 68% to determine the reliability of the model's FSD. The advantage of this technique is that multiple FSDs can be combined to get a single measure of actual performance versus expected performance. Similarly, a 1.28 standard deviation interval should contain about 80% of the differences from sale price. A lender can also test performance in a 90% interval, by creating a range 1.65 standard deviations around the AVM's value estimate.

If the lender is using confidence scores that are based on FSD, the same test can be performed for the scores. Using HVE as an example, all the high confidence score estimates from the AVM can be combined and the differences from the test values calculated. The lender can verify whether at least 68% of the test values fall within +13% of the high confidence values. All the medium scores can be combined and the performance compared with a +20% deviation from the test values, and similarly for the low confidence scores.

Lenders are using AVMs more and more often to streamline their home-equity mortgage lending.

Lenders are using AVMs more and more often to streamline their home-equity mortgage lending.

As AVM usage grows, so does the importance of understanding the performance characteristics of the AVM used. Lenders can use the forecast standard deviation provided by the AVM in risk management calculations. But, to gain the greatest value from their models they must validate the FSD as well as the AVM values, to ensure that they understand the risk from using the AVM in their processes. Confidence scores that are built up of ranges of FSDs can give lenders similar risk management capabilities. Without FSD or a confidence score based on FSD lenders are forced to guess at risk management.

Douglas Gordon is director of collateral modeling at Freddie Mac.