Digital DIALOGUE
A SourceMedia Web Capture

Executing Effective Validations

By Sarah Davies | Senior Vice President, Analytics, Research and Product Management, VantageScore Solutions, LLC

One of the key components to successfully utilizing risk management models and decision analytics is an effective validation. When executed properly, model validations verify that models are performing according to expectation within their original design and purpose. Effective validations also confirm that models remain sound, even amidst changing environments, while identifying and measuring the impact of any potential limitations or errant assumptions.

Regulators are also stressing the importance of model validation. In April 2011 the Office of the Comptroller of the Currency (OCC) expanded long-existing guidelines about model validation in response to the lending industry's increasing reliance on models to drive decision making. As part of its guidance, the OCC explicitly recommends that financial services firms utilizing predictive models and decision analytics perform regular validations to gauge model efficacy. The OCC's guidelines apply both to lenders that utilize proprietary models and, importantly, those that use vendor generated models where there might not be the same transparency and understanding in terms of how the model was built.

VantageScore Solutions recently completed its own validation of VantageScore® 2.0, the model's second version, and enhanced its validation execution in accordance the OCC's expanded guidelines. A webinar was then hosted in conjunction with *American Banker* to discuss the results and help risk managers perform their own validations in light of the OCC's guidance on model risk management.

The webinar is available online at www.VantageScore.com/research. A synopsis of the event is presented here.

VANTAGESCORE®

Advertorial Produced by SourceMedia Marketing Solutions Group

ASSESSING RISK … HOLISTICALLY AND ROUTINELY

One of the flagships of the OCC's validation guidelines is the idea of having "effective challenge," which is the notion that lending institutions should treat validations as an independent function outside of model development and even outside of model applications. Such treatment affords cleaner processes that allow the validation team within organizations to operate objectively.

A trend among larger institutions is to acknowledge the seniority of the role responsible for model validation and the importance of that person having access to the chief executive officer and the board. These individuals and teams will routinely examine whether risk-management models and processes are fair, transparent, and appropriately used in the context of a lender's business. Many of these institutions have created roles such as chief model compliance officer or chief risk officer. Most smaller institutions are still trying to work through how to develop and identify the appropriate resources for model compliance.

While there is no prescribed method for analyzing risk, validation teams can follow several generally accepted principles. Materiality is a significant issue as a starting point. What is the organization's risk appetite? In a statistical sense, how much mis-estimation can be tolerated? What's the appetite for financial risk? When does too much risk within the business model really impact P&Ls? And what is the reputational risk?

These are critical questions to address before the validation has begun.

EFFECTIVE VALIDATION

An effective validation can focus on three major themes: (1) conceptual soundness, (2) ongoing monitoring, and (3) outcome analysis.

Validation teams need to be intimately familiar with the architecture of the segmentation, the actual data that was developed and synthesized to feed the model and the initial development's performance. This occurs organically when a lender has developed its own model, but for those using third-party developed models it is incumbent upon lenders to conduct due diligence to understand how each model they use is designed.

Much of the necessary information can be provided by the vendor. Upon receipt, understanding the model's assumptions is important. Are the assumptions intuitively correct? Is there sufficient transparency in terms of why the model is performing the way it does? Veiled answers are unacceptable.

Secondly, the concept of ongoing monitoring can be built into the process. This includes routinely drilling down into the performance of the model ensuring that the model is rank ordering within the right markets and on the right products. Questions to ask include whether there are various consumer or product segments performing in the appropriate ways and if it is possible to create better reporting systems.

Ideally this will include dedicated resources, which underscores the OCC's desire that validations become a priority level function. The goal is to understand how and where a model's potential errors are exposing a business to unexpected risk as well as the impact of those errors.

OUTCOME ANALYSIS

The "meat" of the validation is in its testing , or outcome analysis. The OCC describes effective outcome analysis to include the following:

- **Backtesting**, which analyzes whether the model predicted what actually happened;
- **Benchmarking**, which compares a model's performance with other models in the marketplace;
- **Stress testing**, which looks at how the model performs in high-risk areas; and
- **Sensitivity analysis**, which evaluates the impact of small changes in the input data.

VantageScore Solutions performs these test as a matter of routine and provided here is a roadmap that risk managers can use to shape their validation procedures. Keep in mind that implicit in the OCC's guidelines are recommendations to drill down to understand the details behind the outcomes. An unexpected result should trigger analysis of what has happened and whether the model is still performing within the organization's tolerance.

BACKTESTING

Numerous metrics may be employed for backtesting, which might surprise some risk managers that are accustomed to the industry mainstay: the Kolmogorov-Smirnov (KS) statistic. A KS measurement looks at the cumulative percentages of good consumers and bad consumers (current borrowers and defaulting) and identifies the point at which there's the greatest separation of those two distributions.

However, it's also important to go beyond an examination of a single point on a credit score model's range so that the validation affirms there is proper separation throughout the distribution. Supplementing the KS measurement is the trade-off curve, or ROC curve. It measures the quality of the distribution. Within this tool, risk managers examine the "C-statistic," whereby a score of 1 equates to perfection and .5 equates to complete randomness.

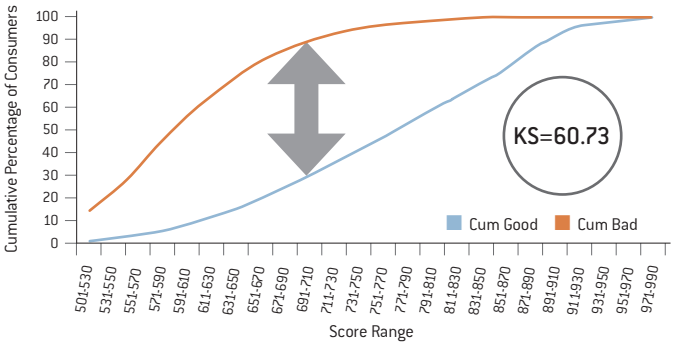
When used in conjunction with one-another, these back tests provide tremendous confidence in how a credit score model is performing, both in terms of identifying goods and bads in an absolute fashion but also along the entire distribution.

The graphs in Figure 1 below demonstrate how these two measurements were utilized when a validation was performed on the VantageScore 2.0 model.

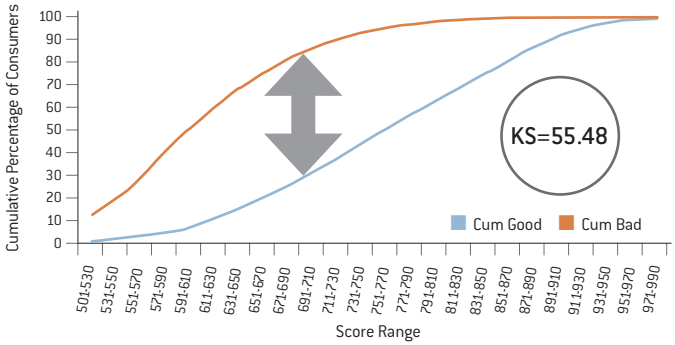
> "KS values show that the model is doing an excellent job of pulling the 'goods' to the top of the distribution and the 'bads to the bottom'."

Figure 1

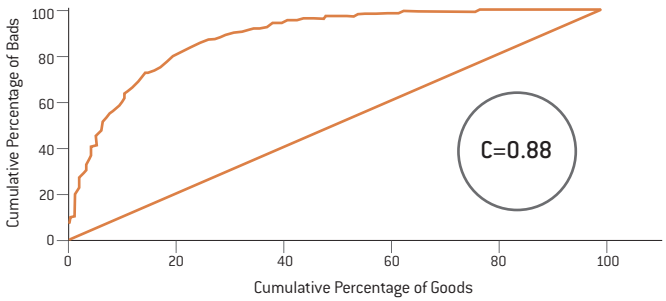
Existing Accounts



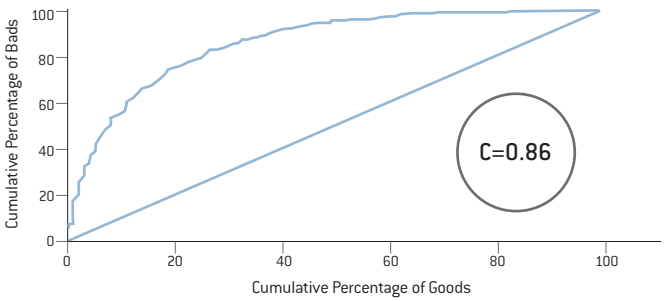
New Accounts



Existing Accounts



New Accounts



As demonstrated, the KS values show that the model is doing an excellent job of pulling the "goods" to the top of the distribution and the "bads" to the bottom, which allows lenders to more effectively manage risk, and the trade-off curve demonstrates the model's separation across the entire distribution.

Across all industries, the model is performing similarly, as demonstrated by the KS results right, Figure 2.

Figure 2

VantageScore 2.0 KS Values

Industry	New Accounts	Existing Accounts
Overall	55.45	60.73
Bankcard	63.99	63.02
Finance	40.89	49.44
Revolving	58.54	63.06
Installment	49.54	49.75
Auto	46.92	51.30
Retail	51.62	63.17
Real Estate	55.80	59.68
Department Store	52.26	64.95
Credit Union	45.90	58.54

BENCHMARKING

The next question to address is whether the model provides the absolute best results available. This is where benchmarking may be employed. In the credit scoring industry, benchmarking determines how a model performs against other competitive models, both in-house, proprietary models and those provided by third parties.

For example, VantageScore 2.0 is measured against the best credit score models from each of the three largest credit reporting companies (CRCs). As the chart in Figure 3 below shows, for mortgage originations, there is exceptionally strong performance, with VantageScore 2.0 outperforming the CRC models in a range from eight percent to 12 percent. The average range of outperformance is three percent to four percent across the board for most of the key industries.

By and large, when lenders are comparing their models to others, it's important to conduct enough benchmarking to feel confident they're getting the best possible results from the best available tools.

Figure 3
New Accounts: 2009-2011 (KS Values)

Industry	VantageScore 2.0	CRC 1	CRC 2	CRC 3
Overall	55.5	54.4	54.0	54.1
Bankcard	64.0	63.3	63.3	63.1
Finance	40.9	39.9	38.5	39.4
Revolving	58.5	56.9	56.8	56.9
Installment	49.5	48.9	48.8	49.1
Auto	46.9	43.9	43.5	43.9
Retail	51.6	49.1	48.6	48.6
Real Estate	55.8	52.0	49.8	51.5
Department Store	52.3	49.3	48.6	48.4
Credit Union	45.9	43.1	43.0	43.3

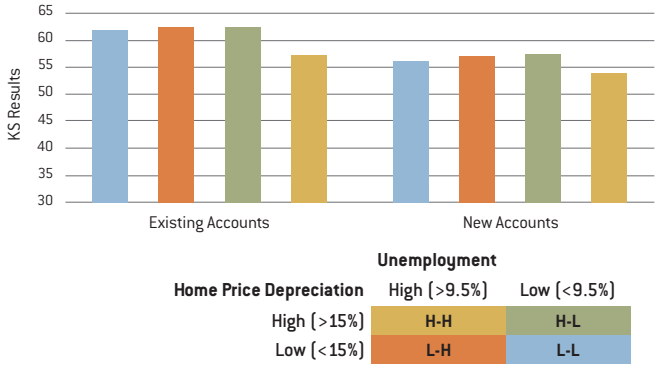
VantageScore Percent Improvement

Min	Max
2	3
1	1
3	6
3	3
1	2
7	8
5	6
8	12
6	8
6	7

STRESS TESTING

Ensuring a model doesn't break down in the most volatile conditions is very important as well, especially as the lending industry emerges from just such an environment. As part of the validation exercise for VantageScore 2.0, the model was exposed to stress testing. Demonstrated below in Figure 4, the validation examined performance in geographic regions of the country that experienced varying degrees of unemployment and home price depreciation.

Figure 4
Stress Test (KS)



The yellow bar represents regions that have both high unemployment and high rates of home price depreciation, which is where accurate predictive performance can be harder to achieve due to volatility. The KS results for VantageScore 2.0 remain strong across the high and low stress regions.

This test can be tailored to a risk manager's own business. For example, instead of geographic regions one might test certain products where the risk may vary.

SENSITIVITY ANALYSIS

Another test useful for an effective validation is determining model sensitivity to differentiating data inputs. For a third-party credit score model, sensitivity analysis can be viewed in two ways:

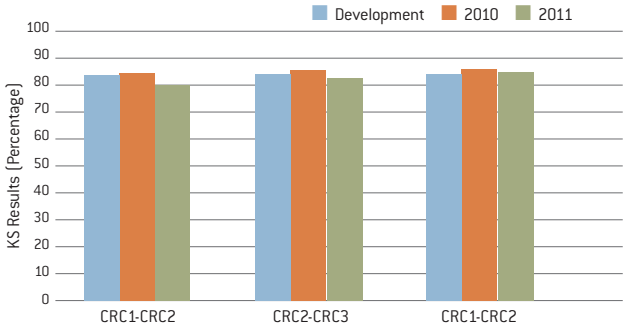
1) The impact from small differences in consumers' credit files at the three CRCs.
2) The impact from changes in consumer payment behaviors that have occurred since a model was developed.

Exposing a third-party consumer credit scoring model to small data variances in consumer credit files among the three CRCs is an important way to measure sensitivity. If the test results in wide variance among scores, then a portfolio may be exposed to hidden risk because the model has potentially become unstable.

The VantageScore model uses a single algorithm with leveled characteristics across all three CRCs, which limits score variance. Model stability across the multiple data sets is proven by KS results, which show consistent accuracy in predicting performance, regardless from of which CRC the data is pulled.

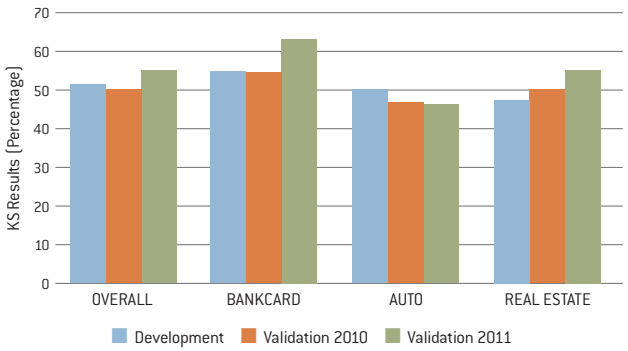
The graph in Figure 5 below shows the percentage of consumers whose score difference is less than 40 points between the CRCs when VantageScore 2.0 is employed. The scores generated are highly predictive and highly accurate across the three CRCs, thus allowing lenders to have the confidence that a consumer will fall under the same risk bracket regardless of the data source.

Figure 5
Sensitivity Analysis: Score Consistency



Sensitivity measurement can also show patterns over time. An interpretation for users of third-party consumer credit scoring models is its ability to remain predictive over a time period where payment behaviors may have changed. The graph in Figure 6 shows VantageScore 2.0's performance improvement over the past two years versus its performance at development. While some deterioration in the auto lending sector occurred, credit card and mortgages have improved 14 percent and 16 percent respectively.

Figure 6
Sensitivity Analysis: Score Stability



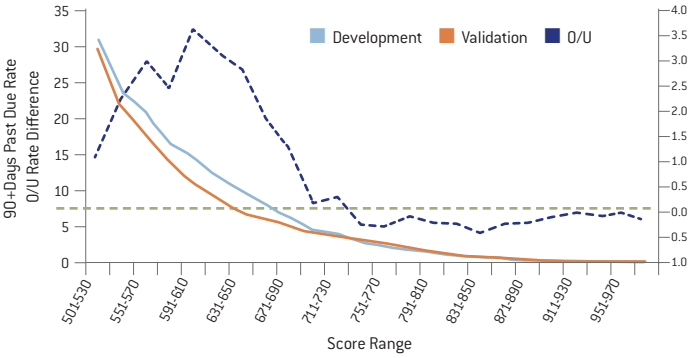
DRILL DOWN

As the VantageScore validation team engaged these tests, key findings included that the model is more accurate overall and particularly in the credit card and mortgage industries. Additional findings reflected statistically insignificant drops in performance. Rank ordering remained effective throughout.

An assessment of the input data reveals there has been a significant shift in the composition of the marketplace, which gives some context to the performance.

The volume of subprime originations in the auto finance industry is increasing. Offsetting this is an overall improvement in default rates. The graph below in Figure 7demonstrates a reduction in default rates in the subprime score band in particular, thus there is little likelihood of a material impact to a portfolio's P&L ratio.

Figure 7
Drill Down: Auto

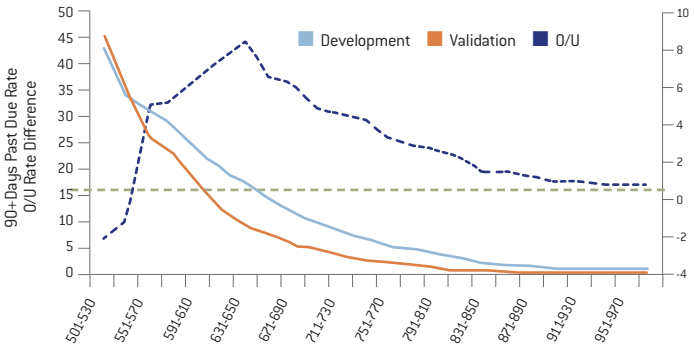


The real estate industry also is experiencing a shift. In the case of the most current timeframe, data shows that there was over a 60 percent reduction in defaults as compared to the original development window. Moreover, the total number of credit inquiries is reduced by about one-third, and there also was one-third fewer opened bank cards with a significant reduction in bank card balances.

Consumers appear to be contracting or reducing their credit footprint, suggesting that they are now more likely living within their credit footprint. Intuitively, consumers are now living within their credit or cash profile and not extending themselves into risky situations. Fundamentally that's driving lower risk and more stable behavior.

Based on the results of the validation and the tightening of credit markets, the chart below suggests it may now be time to expand lending strategies in order to capture more borrowers that have behaved conservatively. The doted-green line exemplifies how the development default rate was higher than the validation's default rate, signaling to lenders an optimal pool of prospective borrowers.

Figure 8
Drill Down: Real Estate



CONTINUED ON BACK

CONCLUSION

Clearly, validations are becoming a much larger priority for the lending community as the OCC's guidelines are taken to heart. It will take some getting used to. Budgets and priorities will be shifted. For small lenders it's a whole new world.

Central to the guidelines is the quest to continually ensure risk models are accurate, fair, and transparent, and that actionable results are reviewed by those with the power to affect change.

Much more on this topic is available in the web seminar online at www.VantageScore.com/research.

The VantageScore model leverages the collective experience of the industry's leading experts on credit data, credit risk modeling and analytics to provide lenders and consumers with a more consistent, highly predictive credit score. Developed as a joint venture among the three largest credit reporting companies (CRCs)—Equifax, Experian and TransUnion, VantageScore marks the first time that the three companies joined forces by combining cutting-edge, patented and patent-pending analytic techniques with an intuitive scoring scale, producing a model that offers more consistency across all three CRCs and has the ability to score more people. ■

VANTAGESCORE®

The entire web seminar, including many more graphs and participant questions, is available online: www.VantageScore.com/research.