the **pipeline**



ISSUE NO. 175 | March 2020 HOW TO AVOID BEING TRIGGERED BY MODELS

By Tom Parrent

Introduction

Borrower behavior constantly evolves in response to changes in economic conditions, new product introductions by lenders, and development of new government policies. Ensuring that Andrew Davidson & Co., Inc.'s (AD&Co) models remain good predictors of consumer behavior requires constant testing and, when appropriate, timely modifications. Performance trigger reports help us identify model drift and developing trends in prepayments, delinquencies and defaults. Properly adjusted models deliver better forecasts, valuations and sensitivities for client use in portfolio management, relative value analysis and hedging.

AD&Co's modelers rigorously test all models during development using standard statistical techniques such as hypothesis testing, in and out of sample model runs and out of time model analysis. Our Model Risk Management (MRM) team validates all new models and significant updates to existing models.

In addition to validating models, MRM reviews model performance on a monthly basis to ensure all models continue to perform in line with expectations as economic conditions and consumer behavior evolve.

Monthly trigger reports provide a quick overview of model performance while also giving guidance on particular areas of developing concern. Catching developing trends early allows us to adjust models through tunings, re-estimation or development of new functional forms that may include additional explanatory variables.

Trigger reports help segment variation between model projections and actual results into noise that averages out over time and signal that identifies changes in conditions and behaviors that require closer investigation. Such segmentation focuses our efforts on adjusting models only when trends clearly indicate a compelling need.

Continuous performance monitoring helps answer questions such as:

- Are the models performing in line with expectations?
- Are coefficients directionally correct and intuitive?
- Are updated goodness of fit measures reasonably in line with those calculated during model estimation?

- Does persistent drift or bias indicate a need for tuning, re-estimation with updated data, or inclusion of new explanatory variables?
- Do emerging behavioral patterns tell us anything about causation vs correlation?

Developing parsimonious models is critical to our ability to project future consumer behavior. We want our models to include all significant variables but exclude variables that may lead to better statistical fits but result in overfitting. Overfit models perform poorly when future conditions do not closely mimic conditions in place during the model estimation period.

Our trigger reports deliberately address the need to identify early adverse trends while avoiding overfitting by using multiple approaches to highlight deviations that represent true signal instead of background noise. Monthly observations tend to be quite noisy, so we use consistent behavior over multiple consecutive months to trigger watches and warnings.

Currently, we have defined watches as patterns that persist for three consecutive months and warnings as patterns that persist for six consecutive months.

We monitor four different types of triggers:

1. Moving Average: Are residuals deteriorating over time?

The moving average trigger compares the 12 month moving average value of residuals to the three month average. The same sign three month average exceeding the 12 month average indicates a model is performing worse in the short term than it has in the long term.

2. Trend: Are residuals consistently becoming more positive or negative?

We define a trend as residuals having a consistent sign and the absolute value of those residuals increasing over time. Model results are thus consistently deteriorating without interim correction. Trending residuals may indicate a fundamental shift in behavior not contemplated in the model.

3. Bias: Is the model persistently overpredicting or underpredicting actuals?

Consistently positive or negative residuals indicate a model may need to be sped up or slowed down even if the absolute value of residuals is not increasing. Deciding on whether to tune away the model/actual difference usually depends upon the level of bias.

4. Magnitude: Are model misses large enough to matter?

Magnitude triggers use the standard deviation of residuals calculated over 12 months. An observation of recent residuals that is large relative to the standard deviation indicates the possibility of recent and significant model slippage. We currently define a Magnitude Warning if any residual in the past three months exceeds three standard deviations. A Magnitude Watch occurs if any residual in the past six months exceeds three standard deviations. Note that Magnitude Triggers are an exception to our general approach of looking for three or six consecutive months of signal.

Monitoring models to ensure that the balance between parsimonious models that avoid overfitting and acceptable performance against actuals is maintained, MRM generally looks for simultaneous trigger warnings

on at least two of the trigger types. For instance, a persistent bias may not indicate a need for model adjustment if the magnitude of residuals is very small. On the other hand, large or persistent trigger violations in any single category are investigated to get ahead of significant model deviation from actuals.

Rather than having a hard and fast policy on how many tripped triggers warrant a model update, we take a holistic view of performance monitoring and evaluate all results in the context of changes in market conditions and policies. For instance, as COVID related forbearance became a reality in April 2020, we knew that model tunings were warranted regardless of any particular trigger status.

Generally speaking, when a model trips a watch, MRM looks into the aberration. A warning results in a broader conversation with the modeling team including discussion of possible mitigation actions.

CDR, CRR and CPR are the primary metrics monitored but we also investigate significant deviations at the transition level when appropriate.

In addition to observing performance at a point in time, we also look at trigger trends over time. As shown in Figure 1, we monitor triggers over a six month period in each report. Our reports are also flexible enough to allow reporting over any historical time period. We use this capability when evaluating new models under development by observing residual patterns in different historical environments. This particular report uses special post-COVID tunings released in January 2021.

Resid CPR

UPB

2,073,343,678,416

Loan_Type

FHLMC_30YR

Trigger Type

Mov Avg

Trend

Magni

Rias

Figure 1. LDM 2.2 FHLMC 30 Year Trigger Report – January 2021 Tunings

Agency Pool IDX: 245 Trigger Date: 01DEC2020 Performance Metric: Resid_CPR

Cohort: Loan_Type Minimum Cohort Size: 50,000,000,000 Watch Months: 3

Warning Months: 6

Short Moving Average: 3 Months Long Moving Average: 12 Months

Magnitude Watch: At least one 2 StdDev in 6 Months Magnitude Warning: At least one 2 StdDev in 3 Months



Bias Resid_CPR	Analysis_Date_SAS							Magnitude Regid CBR	Analysis_Date_SAS						
	07-2020	08-2020	09-2020	10-2020	11-2020	12-2020		Resid_CFR	07-2020	08-2020	09-2020	10-2020	11-2020	12-2020	
100								100							
FHLMC_30YR		•	•	•	•	•		FHLMC_30YR	•	•	•	•	•	•	

Figure 2 displays the information for the same securities but using a pre-COVID set of tunings. The improvement is immediately obvious as the new tunings eliminated the persistent moving average and bias warnings.

Figure 2. LDM 2.2 FHLMC 30 Year Trigger Report – Pre-COVID Tunings



The trigger reports also contain time series charts of the data behind the triggers. Figure 3 displays the Post-COVID tunings on the top with Actual/Model results on the left and trigger measures of moving averages and residuals on the right side. The bottom two charts show the Pre-COVID tunings over the same time period. The charts clearly demonstrate that the retuning of the model eliminated the persistent bias present in the Pre-COVID tunings as well resulting in improvements in the six month window of moving average residuals.

Figure 3. Trigger Charts – January 2021 and Pre-COVID Tuning Sets



January 2021 Tunings

Pre-COVID Tunings



Overall model fit at a given point in time tells us how a model is performing on average. However, not every portfolio represents the market average and developing trends often occur at the sub-portfolio level. For instance, overall model performance might look good, but the model may overpredict prepayments on high coupons and underpredict prepayments for low coupons. To guard against good average model performance hiding problematic performance at a more granular level, our trigger reports include analysis of all significant model variables. Table 1 lists some of the granular factors we monitor on a regular basis.

Table 1. Factor Level Trigger Reports

Product Type	House Price Appreciation						
Loan Term	Geographic Location						
Origination Vintage	Loan Purpose						
Credit Score	Occupancy Type						
Loan to Value	Loan Age						
Loan Size	Spread at Origination						

Figure 4 displays trigger results for Credit Score. This demonstrates how a holistic approach produces a better feel for model performance than simple rules. Note that several Warning triggers are tripped for the Credit Score sub-cohorts. Exhibit 6 shows that, while the results do appropriately fire warnings, the size of residual CPR is quite small relative to the level of actual CPR, thus requiring no additional action.

11-2020 12-2020

. .

. .

• • •

•

•

12-2020

• • •

Figure 4. Credit Score Trigger Report



MovAvg Resid_CPR				Analysis_	Date_SAS			Trans Regist	Analysis_Date_SAS				
		07-2020	08-2020	09-2020	10-2020	11-2020	12-2020	Trend Nesia	07-2020	08-2020	09-2020	10-2020	
Loan_Type	Orig Credit Score							Loan_Type Orig Credit Score					
FHLMC_30YR	620-699	•	•	•	•	•	•	FHLMC_30YR	620-699	•	•	•	•
	700-749	•			•	•	•		700-749	•	•	•	•
	750+	•			•	•	•		750+	•	•	•	•

Bias Resid_CPR				Analysis_	Date_SAS			Managed a Day	Analysis_Date_SAS					
		07-2020	08-2020	09-2020	10-2020	11-2020	12-2020	Magnitude ries	07-2020	08-2020	09-2020	10-2020	11-2020	
Loan_Type	Orig Credit Score							Loan_Type	Orig Credit Score					
FHLMC_30YR	620-699	•	•	•	•	•	•	FHLMC_30YR	620-699	•	•	•	•	•
	700-749	•	•	•	•	•	•		700-749	•	•	•	•	•
	750+	•	•	•	•	•	•		750+	•	•	•	•	•





Figure 6 further confirms the well behaved nature of the model against Credit Score by showing that residual deviations remain well controlled with most residuals at less than one standard deviation.

Figure 6. Residual Control Charts – Credit Score

Loan_Type: FHLMC_30VR WAOCS: 620-699 4 3 2 1 0 Red CR. Devidon 4 -2 -3 4 2020-10-12 2020-07-04 2020-08-23 2020-12-01 Resid CPR Deviation □ +13dDev -23dDev +2 StdDev

Loan_Type: FHLMC_30VR

WAOC5: 700-749





WAOCS: 750+



In addition to monitoring performance against modeled variables, our trigger reports also include variables that are not included in model estimation. As mentioned above, we intentionally build parsimonious models to avoid overfitting. At the same time, we want to make sure that the models do not miss out on essential predictive power of variables not included in the current version of a model.

For instance, LDM 2.2 does not include debt to income (DTI) ratio as an explanatory variable. We know that DTI has an effect on prepayments and defaults. However, inconsistency in historical data has often muddled the signal that DTI might provide to improve model performance. Nonetheless, we monitor the performance of models against DTI to make sure that we are not missing a clear signal coming from recent, cleaner data.

Figure 7 displays the trigger report for DTI on FHLMC fixed rate collateral as of December 2020 using our standard Pre-COVID release of LDM 2.2. It shows some deviation across the DTI range indicating that DTI may now exhibit a strong enough signal to merit inclusion in model estimation. In fact, the next release of loan level LDM will include DTI as an explanatory factor.



Figure 7. DTI Trigger Report – Pre-COVID Tunings



Additional non-modeled factors that we regularly monitor include Property Type, Occupancy Status, Pool Delinquency Percentage and Current Balance. While these variables may seem obvious choices for inclusion in model estimation, we must always maintain the delicate balance between well fit and over fit models. Monitoring model performance against these variables helps us decide when that balance reaches a tipping point in favor of adding a variable to the model.

Monitoring model performance in order to separate signal from noise ultimately involves a blend of art and science, judgment and statistics. Bringing together the entire AD&Co team on a regular basis to review the trigger reports not only represents good model governance but also inevitably leads to discussions ranging from economic conditions to borrower demographics to new model forms. Such creative collaboration drives innovation and results in continual model improvements to better serve our clients.

Collaboration does not end with our internal discussions. We value input from our clients and welcome your questions and suggestions regarding model performance monitoring. Client conversations often spark additional innovation in both our model development and performance monitoring. We look forward to hearing your ideas.

Please contact us at tparrent@ad-co.com or support@ad-co.com with any questions or suggestions.

This publication is believed to be reliable, but its accuracy, completeness, timeliness and suitability for any purpose are not guaranteed. All opinions are subject to change without notice. Nothing in this publication constitutes (1) investment, legal, accounting, tax, or other professional advice or (2) any recommendation or solicitation to purchase, hold, sell, or otherwise deal in any investment. This publication has been prepared for general informational purposes, without consideration of the circumstances or objectives of any particular investor. Any reliance on the contents of this publication is at the reader's sole risk. All investment is subject to numerous risks, known and unknown. Past performance is no guarantee of future results. For investment advice, seek a qualified investment professional. Note: An affiliate of Andrew Davidson & Co., Inc. engages in trading activities in securities that may be the same or similar to those discussed in this publication.