

ensemble~~x~~

How Broken is Your Model?

White Paper

Shawn Budde and Mike Armstrong
May 2020

Introduction

Many of us have managed lending businesses through recessions, including the Great Recession, but managing through a Pancession¹ is a first for everyone. In numerous conversations with lenders and peers, we are frequently asked the same question:

“Is my underwriting model broken now?”

Almost certainly the answer is, “Yes.” All underwriting models are suffering from some model deterioration due to the Pancession.

In this paper, we share how lenders can use a “Flash Assessment” to determine *how* broken their model is. Is it “add some oil and don’t rev it too high” broken, or is it “have a seat and call a tow truck” broken? If lenders can diagnose how and where their model is broken, then they can make the necessary repairs and adjustments. Lenders who understand and react to these shifts can emerge from the Pancession stronger.

How Recessions Affect Credit Models

In a “typical” recession, models continue to rank order risk, but overall levels of risk increase across model scores. The traditional playbook is to pull back on credit by gradually raising cutoff scores. The end result is volumes contract while risk levels are managed to an acceptable level.

In the Great Recession, we saw that some geographies demonstrated outsized levels of credit risk. The mortgage crisis hit the “sand states” the hardest (AZ, CA, FL, NV). Many lenders paused all lending in the sand states. More informed lenders realized that their models still worked but that they needed to adjust their policy by supplementing their models with geographic overlays. We’ve seen lenders use a similar playbook for more localized crises (e.g. 2017 Harvey Flood in Houston).

With the Pancession, the additional concern is that underwriting models are broken because they are no longer able to rank order effectively. The current unemployment shock means that many Americans, who otherwise routinely

¹ We have not seen this label used elsewhere, but we feel it is the best description of the current pandemic driven economic crisis.

made their loan payments, will face increased economic stress the longer the shutdown persists. Models built on pre-Pancession data have no visibility into how Pancession-affected Americans will perform. Therefore, all current models are broken to some degree.

Defining the Components of Model Deterioration

Most lenders will review model performance on a routine basis to monitor the deterioration and determine when to build a new model. The natural “time and distance” between model development and in-production usage will deteriorate rank ordering. We refer to this as “**Decay**,” which is a natural part of the model lifecycle. Our assumption is that in normal circumstances the model would continue to decay at a regular, predictable rate. There are many tools for assessing decay. Therefore, we will focus on other causes of deterioration directly related to the Pancession.

Defining Shift

It’s possible that the model can rank order a scored population as well as it did before a recession, but that the credit environment has gotten worse. In that event, lenders can simply raise their approval threshold to yield the same credit performance – albeit at lower volumes – that they were able to get previously. We are using the term “**Shift**” to describe the change in performance that is independent of the model’s ability to rank risk.

It is important to measure shift for a variety of reasons, including knowing how much you need to throttle back, calculating loan loss reserves, FP&A forecasting, capacity planning for collections, and forecasting losses to manage capital markets. Accurate assessment of shift also allows you to throttle up when risk improves.

Defining Distortion

While it’s possible that a model could rank order risk just as well during a recession, it is unlikely. It is more likely that an economic shock will deteriorate rank ordering by fundamentally altering previously stable relationships between variables and outcomes. We are using the term “**Distortion**” to describe the incremental and potentially temporary deterioration in rank ordering caused by an economic shock. We summarize these terms in Figure 1.

Figure 1 - Summary of Terminology



How to Assess Your Model

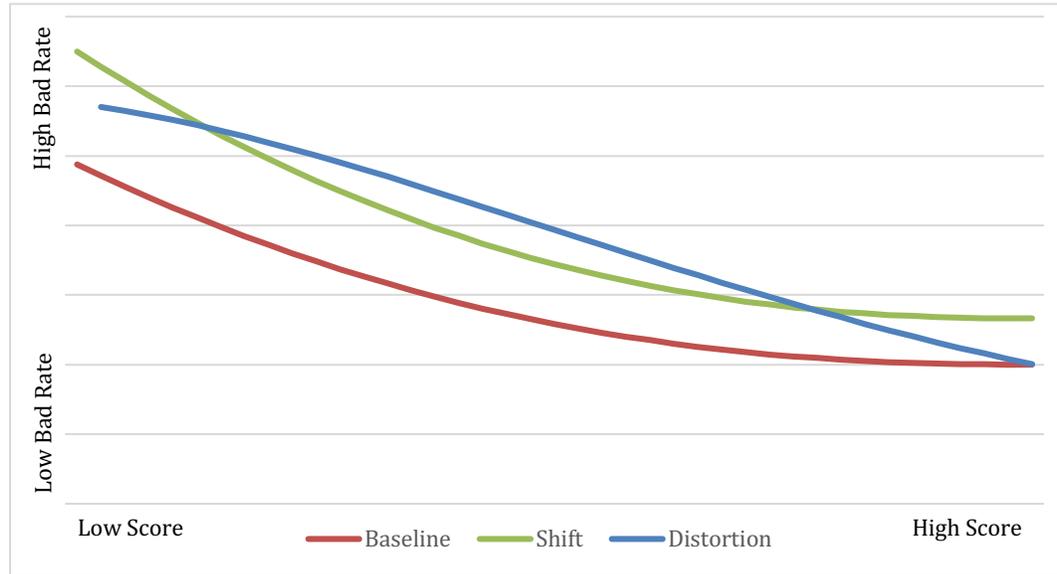
In a normal recession, models will experience accelerated deterioration from both shift and distortion. We've shown what each of those might look like in Figure 2. The baseline shows how the model was performing prior to the recession; the shifted line shows the results of uniformly higher risk with the same model ranking power; and the distortion line shows the impact of comparable risk with a loss of model power.

Figure 2 shows shift and distortion in isolation. In reality, it is very unlikely that only one of the two effects will occur. Despite that, in recessions lenders tend to react as if only shift is occurring. To better understand and adapt to deterioration we want to measure shift and distortion independently.

Measuring the Model

In normal circumstances, we would assess a model using the same objective function, observation period, and universe. For example, a typical model might be built to predict bad rate (any 90+ DPD) over eighteen months, on a through-the-door population. These are not normal circumstances, and we certainly don't have eighteen months to let a portfolio age to adapt our strategies. We want an instrument that can give rapid empirical feedback from Pancection performance data to ground our intuition. We are calling this instrument a "Flash Assessment."

Figure 2 - Bad Rate by Cause of Deterioration



To meet these requirements, we have tightened up the timeframe and objective function to be accounts that miss a payment in a given month. We have also, necessarily, altered the universe from a through-the-door population to the existing portfolio. These changes will allow us to see how the portfolio risk is shifting and distorting every month. We have summarized the metrics in Figure 3.

Figure 3 - Model Assessment Frameworks

	Traditional	Flash Assessment
Timeframe	Modeled Timeframe (e.g. 18 months)	Each Month on File
Universe	Approved accounts plus universe test and/or reject inference	Booked accounts
Good	Never 30+ DPD	Made a payment this month
Bad	Ever 90+ DPD	Missed a payment this month ²

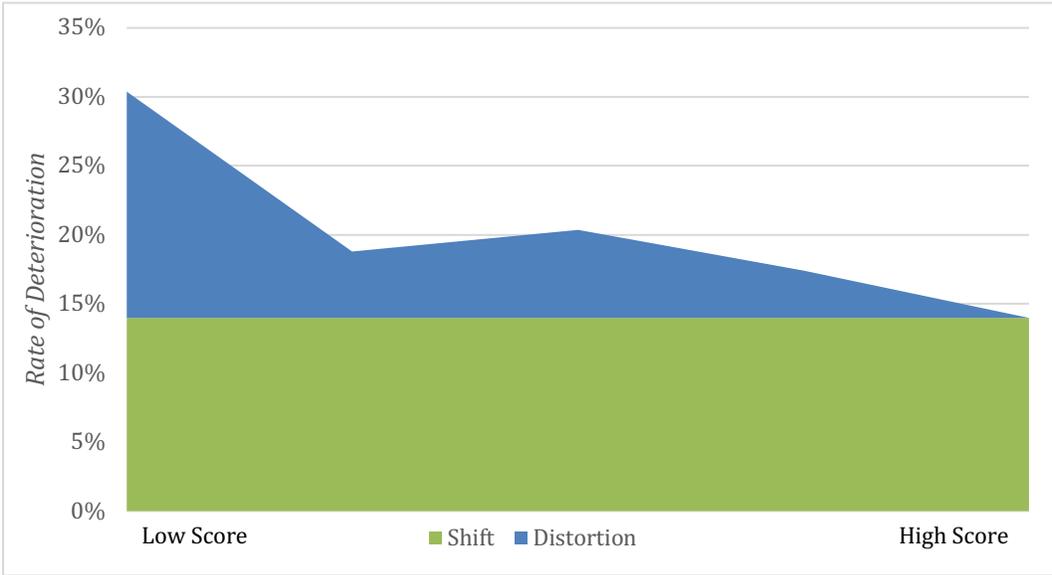
² We have seen analysis that suggests that modified loans are as little as 30% of the risk of comparable early delinquent loans. However, this tool is designed to identify borrowers experiencing stress. As such, we are inclined to consider modified loans 70% bad.

Since we've changed our method for evaluating the model, we need to establish a new baseline. We will refer to the period leading up to the Pancession (performance through February 2020) as "Pre" and the period after the Pancession start (performance from April 2020 onward) as "Post."³ We establish our new baseline by measuring natural model decay during the Pre period with our new performance metrics. We look at model score band performance by vintage over every month on file since the model was implemented, making adjustments when needed to account for major shifts in credit policy or universe. This allows us to set a baseline expectation for the model performance immediately before the Pancession begins.

Measuring Shift and Distortion

Every Post month we will look at the bad rate performance for that month relative to the Pre baseline. By indexing the bad rate for that month against the baseline, we can get an assessment of how the model is performing through the Pancession. We calculate shift by measuring the minimum deterioration experienced across every score band. Distortion is then the incremental deterioration within a score band. We show an example in Figure 4.

Figure 4 - Deterioration by Cause



³ We will essentially ignore the brackish performance of March 2020.

In this case there has been an increase across all ranges, a "shift" of 14%. There has also been significant distortion at the bottom of the score ranges. Measuring the area of each segment tells us the relative impact of shift and distortion. In this case, we can say that about 70% of the increase in risk is due to shift with the remainder due to distortion. It is normal for distortion to move across score ranges like a wave crashing on the shore – it appears first in the higher risk populations and moves across with decreasing impact to the lower risk populations.

Using Shift and Distortion

Having measured shift and distortion, the natural question is “What do we do with it?” That, of course, depends on what the data is showing.

- If the model is **performing within the normal bounds** of natural decay (i.e. there is minimal shift or distortion) then you can keep using your model with confidence.
- If you begin to **observe shift**, then your model is continuing to rank order as intended and you should adjust cutoff scores to account for the higher risk.
- The most likely outcome is that you will observe **changes in shift and distortion** over the course of the Pancession. This can't be solved just with changes to score cutoffs. It requires more detailed analysis to identify and remediate the causes of distortion.

Continually tracking the shape and magnitude of shift and distortion throughout the Pancession allows lenders to leverage the latest information to mitigate risks and maximize opportunities.

Conclusion

Good lending is a blend of intuition based on experience and data-driven decision making. In the early days of the Pancession, lending tilted almost exclusively towards intuition. Some lenders facing liquidity issues suspended new credit extensions entirely. For lenders with stronger balance sheets, suspending all originations is unnecessary.



At Ensemblex, we are working to find ways to bring data back into balance with intuition. The Flash Assessment is a tool that enables lenders to use Pancession data to perform real-time adjustments to their model. Lenders who use these kinds of new, creative techniques will emerge as market leaders while others remain frozen on the sidelines.

About Ensemblex

Ensemblex is a boutique professional services firm focused on the lending industry. We are a team of strategic, data-driven credit experts who worked at companies like Capital One, CitiBank, and ZestFinance. Ensemblex offers analytically advanced consulting services and machine learning solutions across the full lending life cycle. We have decades of experience working with companies ranging from leading banks to fintechs across multiple asset classes, including credit card, auto, and personal loans. Learn more at www.ensemblex.com.

You can reach us at shawn@ensemblex.com or mike@ensemblex.com.