MODEL RISK MANAGERS' INTERNATIONAL ASSOCIATION

Tech Report 2021-01, 8 March 2021, Version 1.0, ©2021 MRMIA

IMPACTS OF COVID-19 ON MODEL RISK MANAGEMENT

WORKGROUP MEMBERS

Joseph L. Breeden, Chairman
LiMing Brotcke
Bin Duan
Andy Johnson
Charles Maner
Paul O'Neal

Summary

The COVID-19 pandemic provided both a test of the lending industry's model risk management practices and an education into issues not seen before. Following the SARS outbreak in 2002 in China and Hong Kong, many institutions globally started using a pandemic as a stress testing scenario, but the reality has been more complex than any prior stress test.

MRMIA's panel of experts explored many areas where COVID-19 impacted financial services. This document collects those insights, primarily related to lending. Issues in consumer data and macroeconomic data highlighted weaknesses in data gathering, data monitoring, and modeling, all of which need to be improved in the future. Even so, no model can be expected to anticipate the nature and magnitude of government and lender assistance programs, so part of model monitoring is how best to construct model overlays. In other words, we are far from being able to run models on autopilot through crises such as this. Human domain experts are still required.

Part of that expertise is also needed in operational issues around new loan originations, account management, and collections. As borrowers were (and still are being) put into unprecedented situations, so too must managers try new approaches to assist borrowers, preserve their portfolios, and support the broader economy.

Even when vaccination brings the pandemic under control, we cannot forget the data of 2020. How should the industry use or not use that data for model development and validation? The answer will necessarily be specific to the kind of model being built and possibly the market segment addressed. Although a simple answer might be to exclude such data, newer portfolios will not have that option. Time horizon will also be important, as losses that might have been realized in 2020 may yet appear in 2021 or beyond, depending on the future of the pandemic and future borrower assistance programs.

Table of Contents

1. Introduction to MRMIA

Model Risk Managers' International Association (MRMIA.org) is a non-profit industry advocacy group focused on promoting best practices in model risk management. With a membership primarily from financial institutions, MRMIA seeks to provide an independent voice on model risk management that explicitly takes into account a cost-benefit analysis behind the practices being promoted. Specifically, MRMIA seeks to avoid chasing methods with diminishing returns, but rather to promote those practices with real benefits to model developers, model owners, and those in various risk management functions who would need to oversee the process.

The present document is created by an MRMIA workgroup tasked with gathering best practices from personal experience, industry interviews, and academic publications. This is expected to be a living document that will be revised as new insights and methodologies become available.

2. Introduction to the COVID-19 Impacts Workgroup

The COVID-19 global pandemic brings extraordinary challenges to the use of data and models at financial institutions. Models are pattern recognizers, but they cannot perform as designed in the face of such unprecedented events. With changes in the definition of input variables, massive new government assistance programs, and extensive borrower support from financial institutions, model risks in 2020 have reached new heights and senior management have needed to incorporate a great deal of judgment and intuition into normally automated processes.

This white paper seeks to collect lessons learned from the COVID-19 crisis to date regarding shocks to the input data and how we can better monitor for these in the future, insights on model monitoring, and insights on how operations and specific products were impacted. As the recession enters a more typical recovery phase, attention will turn to how FIs should use data from 2020 in future modeling and model validation.

The workgroup and broader MRMIA community have gathered these lessons and early thoughts about future impacts into this white paper as a way to assist the industry and spur further industry research on these topics.

3. Data Monitoring

For both model development and model validation, the availability of complete, accurate, and timely data is a precondition to success. To build or validate a model without good data is like asking a chef to cook a meal without rice, as the ancient Chinese saying goes, which is impossible in Asian culture. The COVID-19 crisis and the subsequent actions taken by government agencies and the financial services industry to deal with the economic fallout have altered the models' input data in definition, severity, accuracy, and correlations / relationships in ways not before seen.

3.1. Credit bureau data

The US CARES Act made changes to how data was reported to credit bureaus in an attempt to provide some protection to consumer credit ratings.[1] Consumers who received a loan accommodation, including the federally mandated accommodations for GSE-based mortgages and guaranteed student loans or lender provided accommodations, would not see their delinquency status deteriorate for the duration of the accommodation. The UK's FCA extended similar payment holiday protections across an even wider range of loan products in June 2020 and again in November 2020.[2]

Many who accepted accommodations did so as preventative measures and would not have moved into delinquency, so this rule makes sense for them. For others, it has the obvious effect of masking their current ability to service their loans.

Bureau scores as a rank-ordering tool tend to be quite robust through crises. Bureau scores will reflect whether consumers paid down debt or increased utilization through the pandemic due to the many disparate impacts of the crisis. So some reported information will be timely while data on delinquency would not be likely to reflect the consumer's current risk. All of this is to say that users of bureau scores and data will probably find that the information or bureau scores directly can still be used to roughly rank order borrower risk, but with significantly higher uncertainty than normal. Obviously, this means that data monitoring needs to be increased with an eye to when sudden changes in government regulations in

[1] Consumer Financial Protection Bureau, "Statement on Supervisory and Enforcement Practices Regarding the Fair Credit Reporting Act and Regulation V in Light of the CARES Act", April 1, 2020, https://files.consumerfinance.gov/f/documents/cfpb_credit-reporting-policy-statement_cares-act_2020-04.pdf

[2] Financial Conduct Authority, "Finalised Guidance: Consumer credit and coronavirus: updated guidance for firms", Nov 19, 2020. https://www.fca.org.uk/publications/finalised-guidance/consumer-credit-and-coronavirus-updated-guidance-firms

the future can create cliff-edge moments. In addition, subpopulations are likely affected nonuniformly, so the uncertainty in predictive ability can vary widely across subpopulations. This argues strongly for using traditional bureau data in combination with cash flow analysis or other more timely data to identify subpopulations with disparate intrinsic credit risk.

3.2. Macroeconomic data

Loss forecasting models including FASB's CECL, IASB's IFRS 9 loss, and the UK/EU IRB reserving models and stress testing models suffered from a wide range of unexpected shocks and changes in definition. These models typically leverage government-reported macroeconomic series measurements in order to predict near-term losses. Scenarios for those same factors are used to predict longer-term or lifetime losses.

3.2.1. Data Integrity Problems

Employment measures are key to many forecasting models. In March 2020 as part of the CARES Act, the US federal government enacted the Pandemic Unemployment Assistance program (PUA) that boosted unemployment benefits and expanded the definition of those eligible for unemployment to include the self-employed, comparable to the UK Coronavirus Job Retention Scheme (CJRS). Figure 1 shows initial unemployment claims, both regular and PUA claims. The initial spike in April was under the regular program, but as eligibility information spread and systems were updated, PUA claims exceeded regular claims through most of the summer 2020. PUA claims alone were at a higher level than regular claims at the peak of the 2009 recession. January through June of 2009 saw 16 million initial claims. In the comparable period from March 21, 2020 through September 19, 2020, 38.4 million regular initial claims were filed, almost exactly twice the 2009 period, and an additional 18.7 million PUA claims were filed, exceeding the 2009 claims alone.

Models taking initial unemployment insurance claims as an input were thus hit with a 49% increase over what would already have been a dramatic shock. Models use many kinds of transformations of the input macroeconomic factors, so this does not mean that a given model would automatically overpredict by 49%, but it does mean that this change in definition for those eligible for unemployment insurance adds a dramatic bias to the forecast. Further, this change is not currently permanent. The added benefit to self-employed consumers was set to expire in December 2020.

Systems that monitor input data for errors could not flag this as an error. The data alone cannot show that the underlying definition changed. Definitions do change from time to time. In the 2009 recession, US unemployment benefits were extended to 99 weeks, which was well in excess of any previous level. Consequently, employment measures can be logically critical to loss modeling and yet seem poorly correlated. However, the shifts in 2020 raise questions about how the input data should be used.

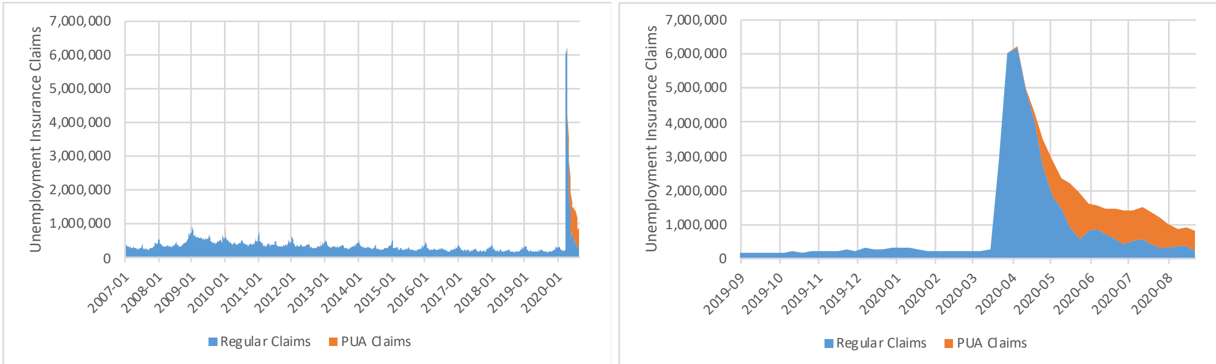

FIGURE 1: CHANGE IN DEFINITION OF THOSE ELIGIBLE TO CLAIM UNEMPLOYMENT BENEFITS. DATA FROM THE US DEPARTMENT OF LABOR.

The US Bureau of Labor Statistics (BLS) conducts surveys of employment status during the first two weeks of every month. Precautions during the coronavirus pandemic required those activities to move from in-person interviews to phone interviews. Further, the response rates were much lower than normal. Although that would naturally cause uncertainty in the estimates, BLS also announced that their field staff had difficulty in assigning the correct status to those temporarily laid-off due to the pandemic.

Figure 2 shows the unemployment rate as reported by BLS. As part of the data release, BLS also provided an estimate of how much the unemployment rate was undercounted because of the misclassification. According to tradition, BLS will not statistically adjust the field count, so the lower number is the official report, but it is seriously wrong.

In March 2020, many economists were predicting an unemployment rate of 20% based upon a detailed analysis of the weekly unemployment insurance claims, when the official number was reported as 14.7%. In fact, adjusted for the undercount, the real number, according to BLS, was 19.2%. However, if an FI's model is loading unemployment rate data from any provider like FRED (fred.stlouisfed.org), the model will obtain 14.7% as the input.

One should note that the undercount is 4.5% in April. That is greater than the total unemployment rate in February 2020. The reporting **error** in April is greater than the reported **value** in February. That would lead any model to significantly underestimate the shock to borrowers.

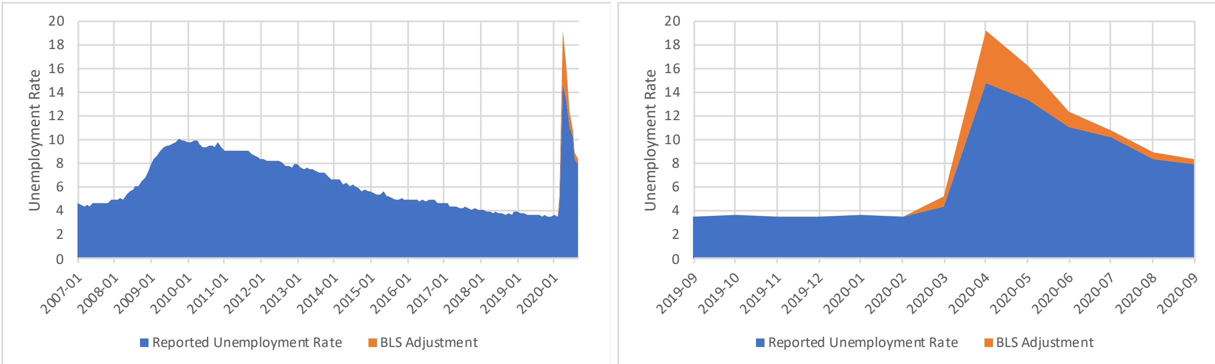

FIGURE 2: BLS REPORTING OF UNEMPLOYMENT RATE UNDERCOUNT.

In both cases shown here, to maintain consistency with the development data, adjustments should be made to the data being fed into the models. Initial claims should include only regular claims. Unemployment rate should be adjusted up to include the estimated undercount. The problem is that both of those adjustments are not readily reported. Additional work is required by analysts or systems engineers to make these changes in a time of crisis when resources are stretched thin. In addition, the only way to be aware of the issues is to diligently read news from those government agencies reporting the data. This resource may not exist at most lenders. Clearly the financial services industry could benefit from an alert service regarding changes to essential model inputs and perhaps even a standardized data feed that adjusts the historic data for those changes as much as possible.

This information also has important implications for error estimates when testing models. Accuracy tests are conducted assuming the input data is correct. Error is measured relative to that data. That is appropriate, but model users need to understand that those error estimates are not complete. They do not include reporting errors and biases in the data itself.

None of the above speaks to what is usually meant by data monitoring. In light of the abruptness and severity of the onset of the COVID-19 crisis, data monitoring is even more important than previously considered. Is the input data going into production models within the range of the training data? If not, how extreme are the observed values? This could be quoted in terms of standard deviations relative to the training data.

However, most of the models that broke, meaning that they could not produce any forecast, did so because of the transformations applied to the input data. Transformation that may have been the best fit in-sample could generate undefined values in light of the COVID-19 crisis data. In general, diminishing returns transformations are advisable for extreme conditions, such as logarithmic, sigmoid or square root transformations that scale less than linearly.

Regardless of the transformations used, the best way to do data monitoring is to measure how extreme the inputs are after transformation. This may not be possible for models that embed the transformations in a networked or multidimensional framework. Still, whenever possible, measuring the standard deviation of the transformed data is most informative.

3.2.2. Unexpected Dynamics

Even when the data is correct, the structure through the recession and correlations between variables has shifted in unexpected ways. This poses a challenge for models that embed such assumptions.

House prices are an important example. In the previous recession, house prices fell between 2007 and 2009. At the same time, months of supply rose as the market was flooded with excess inventory. In 2020, US house prices were flat from May through July but then started to rise again. These rising prices directly relate to a drop in housing supply. Those with steady employment continue to purchase homes, but many who would normally seek to sell their homes do not want to go through that process in the midst of a pandemic.

In this example, supply and demand work, but they do not correlate to economic growth as seen in previous recessions.

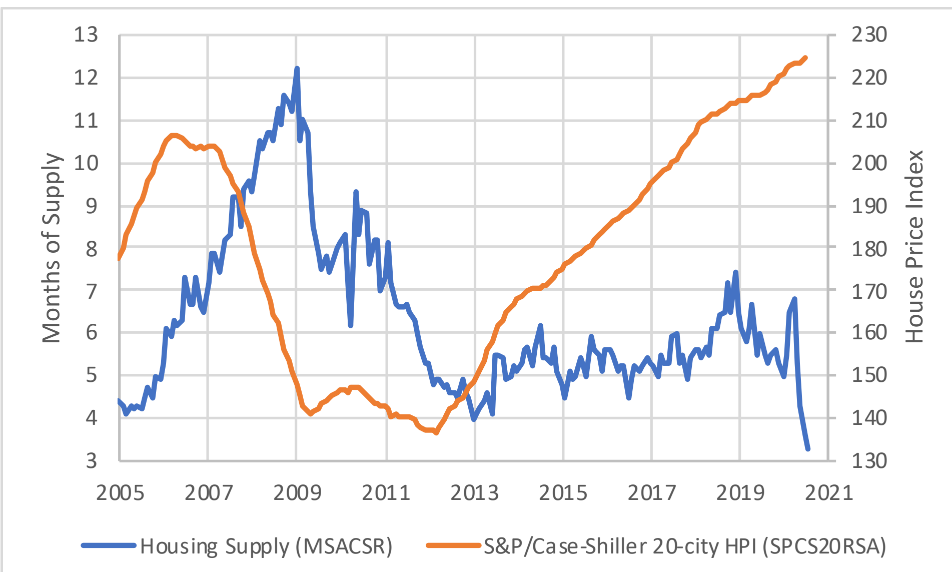

FIGURE 3: HOUSING SUPPLY AND HOUSE PRICES.

Used car prices are important to the auto loan market. Here also, significant and unusual oscillations have occurred. Initially, prices fell on used vehicles as both auctions and dealerships closed. At the same time, auto manufacturers had their plants shuttered and supply lines disrupted, leading to a reduction in new vehicles entering the market. Once the lockdown eased and dealerships re-opened, there was an increased demand for used vehicles due to their lower price and to the lack of new vehicle supply. This has put upward pressure on used vehicle prices, which translates to higher recoveries for lenders selling repossessed cars at auction.

This type of phenomena has been seen before with hurricanes: the storm causes economic damage, but also creates demand for new vehicles to replace those lost to flooding. Supply is temporarily constricted; when consumers purchase a new vehicle, they often add a used one (their trade-in) to the dealer's inventory. This cycle is broken during a hurricane since so many vehicles are flood losses and cannot be resold. There are lots of ways the pandemic is *not* like a hurricane, but it may be the closest event present in lender's historical performance

Multiple indices are now projecting higher used vehicle prices for 2020 than 2019, and continued increases into 2021.

3.3 Use of Alternate Data

Through the rapidly changing early months of the COVID-19 crisis, traditional data sources too often were slow and flawed to effectively guide business decision-making. Although traditional sources, most often from government agencies, are considered reliable, they are insufficient if managers know that the world has already changed by the time the official numbers are released. Similarly, for individual consumer data, credit bureau data continues to be valuable, but cannot offer a complete picture of a consumer's situation in (near) real-time. This will undoubtedly cause some reconsideration of what data sources should be used in models.

3.3.1. Scoring inputs

The primary lesson from the crisis relative to underwriting must be the important of understanding cash flow in real time. The vast majority of financial institutions use underwriting scores that look at credit bureau attributes, not cash flow. As a result of the Covid-19 crisis, the credit bureaus have started provided some bureau data at an accelerated rate, although still weeks behind real-time. Employment verification and self-reported income have always been part of underwriting, but they do not provide a real borrower cash flow perspective. In recent years, many fintech companies had already discovered the value of incorporating cash flow analysis into either the underwriting process or the credit scores directly. The COVID-19 crisis brought this to the forefront of the entire industry.

The US is following a website scraping approach to obtain cash flow data, with customer permission. In the European Union, the UK, and Australia, the 'Open Banking' initiative will be providing a more elegant solution. This API effectively allows consenting consumers to share their transaction account information as owners of their own data.

Cash flow analysis is nothing new in commercial lending. Even there, where defaults are still commonly viewed as a condition that occurs when debt exceeds assets, the value of cash flow modeling is clear. For consumers, model developers are increasingly seeing that loan defaults occur more from a cash flow failure than a debt-to-asset calculation.

If cash flow analysis is established at loan application as part of the underwriting and maintenance process, it can also be available for post-origination analysis of cash flow suspensions, volatility, or a change in source, such as with government assistance. Such shifts in cash flow can feed into behavior scoring models to assess risk, but they can also trigger specific account management actions to support consumers through a crisis, such as proactive discussions of loan accommodations.

As valuable as cash flow modeling has been proven to be, other alternate data sources may be less reliable in a crisis. Some lending innovators have been looking at behavioral patterns in alternate data sources. Famous examples include using the operating system of the device used for the application, the time and date of application, purchase transaction patterns, and many other more esoteric fields. Although the stories have not yet come out, one must assume that those inputs failed rather spectacularly through the crisis. With a huge fraction of the population suddenly working from home, home schooling, and other dramatic shifts, all of the inputs just mentioned would experience a sudden regime shift with no corresponding outcomes on which to revise the models. Well after the crisis, we will learn which kinds of alternate inputs were valuable and robust through the crisis and which were not.

3.3.2. Portfolio forecasting inputs

Obtaining timely macroeconomic data has always been challenging, but the economic slides into the 2001 global recession and 2009 Global Financial Crisis were slow enough to make the wait for updates manageable. By contrast, the COVID-19 crisis was so rapid that even monthly unemployment numbers were out of date by the time they were released. GDP numbers were so slow as to be of mostly academic interest rather than of practical value. Most amusing was the US Index of Leading Economic Indicators (FRED:USSLIND) was suspended completely after its February 2020 value was reported.

Stress test models can be run on hypotheticals, but to understand the likelihood of a specific scenario, information much closer to real time is required. Many US portfolio managers turned to credit card transaction data from sources like Opportunity Insights[3] and Womply[4]. These sources and others can provide daily data on consumer spending by industry category, the fraction of businesses that are open by industry category, and unemployment by income tier. Free data can be obtained with as little as a two-week delay.

Clearly FIs need more rapid and more segmented data than what has been provided traditionally. Unfortunately, many of these sources have not been available with enough

---

[3] https://tracktherecovery.org
[4] https://www.womply.com/blog/data-dashboard-how-coronavirus-covid-19-is-impacting-local-business-revenue-across-the-u-s/

history to verify their reporting accuracy and reliability. Still, something must change. We believe that the most advanced FIs should be allowed to explore using such alternate economic data sources as a way to estimate the inputs to stress test models that would normally take slowly reported government data. This is likely to be an area where vendors will step in to provide estimated values.

Aside from timeliness, one of the biggest advantages in alternate economic data sources would be in segmentation. Today's stress test models implicitly assume that the impact of macroeconomic stresses on borrowers will be in line with previous recessions. However, the COVID-19 crisis impacts are much more concentrated among lower income service workers. Economy-wide averages of the impacts to consumers will miss effects that are very important to lenders with large exposures to specific segments. Forecasting and stress testing with the FRB DFAST variables is not enough to see this disparate impact and therefore limits the value of models using those variables in a crisis such as this.

We need to look beyond the DFAST variables to create models that will be helpful in crisis management. Of course, building models requires historic data, so these segmented variables would be most useful if they can be extrapolated back through the 2009 recession.

4. Operations

COVID-19 is now the pandemic that many had both hoped would not occur and feared would. While many firms were prepared from an operational risk perspective, preparedness from the perspective of individual retail credit impacts was likely not clear, further complicating the financial industry's preparedness for it. Additionally, while never a given, the advent of profound government financial support has assisted with softening the sharp impact that mandatory quarantining has had on numerous industries, namely restaurants, travel, and entertainment.

The COVID-19 crisis impacted almost every aspect of FI operations. For the scope of this work, we want to consider the impacts to new loan underwriting, forbearance, and collections.

4.1. Underwriting and New Loan Pricing

Loan pricing, while predominantly (and obviously) driven by market demand, should take into account both cost of funds (CoF) and expected credit losses. Credit losses in a somewhat benign economic environment, while not trivial, doesn't necessarily require complicated methods to estimate. However, when faced with a changing (and daunting) economic environment, or event, disparately affected industries, changing consumer risk appetites, etc., pricing loans fairly for both the lender and borrower can be tricky.

To date, the Federal Reserve has been keenly interested in banks' capital plans as they relate to the COVID-19 impact. To wit, on September 17, 2020, the Federal Reserve (The Federal Reserve, 2020) has required banks to perform "an additional round of stress tests … due to the continued uncertainty from the COVID event."[5]

Generally, when (retail) loan applications are submitted, they are decisioned by generic credit attributes, typically supplied by the applicant's credit bureau file (in the US and UK, that is TransUnion, Equifax or Experian). Those attributes are combined, sometimes with attributes from the loan application, depending on the particular financial institution. However, intentionally not included are macroeconomic factors, specifically economic forecasts, which would incur concerns about fair lending and general compliance risks. That is, an applicant during benign times might be considered an acceptable credit risk. However, purely due to the anticipated economic outlook conditioned on the applicant's particular credit attributes, that applicant would be declined. This situation would not be one a financial institution's Compliance Risk function would be likely to support/endorse nor would regulatory bodies likely permit.

Instead, financial firms should consider how to link their traditional underwriting/application decisioning models to the firm's credit risk appetite, which does include macroeconomic considerations. Often those models are the expected loss (EL) developed and employed in the name of Comprehensive Capital Analysis and Review (CCAR) and Current Expected Credit Loss (CECL) efforts. In fact, the CECL approach of using economic scenarios during the reasonable and supportable (R&S) period and then relaxing onto long-run averages would align well with using those loss estimates for pricing. This is true for term loans and true if the models are better quality than the minimum standard set for CECL. For lines of credit, the CECL rules about payment allocation and not reserving for future spend do not align with the kind of financial simulations required for pricing.

One challenge with using economic scenarios during the COVID-19 crisis is the split that has occurred between those with jobs and those without. Those who still have jobs during the crisis might be at no more risk of losing their jobs than during normal economic conditions, so once employment and cash flow are confirmed, the economic scenario used for pricing could be quite different than what is used to predict the risk in the existing portfolio where employment and cash flow are not reconfirmed.

Arguably, the firm's risk appetite and underwriting decisioning models should be linked. However, that linkage should be manifest in the use of the underwriting models, not so much in the models themselves, so as to avoid uncomfortable discussions with compliance risk and regulatory parties.

---

[5] The Federal Reserve. "Federal Reserve Board releases hypothetical scenarios for second round of bank stress tests", September 17, 2020, https://www.federalreserve.gov/newsevents/pressreleases/bcreg20200917a.htm.

In this era of a pandemic, COVID-19, firms are endeavoring to quantify their expected losses. However, that is only for what is already booked. Dynamically linking the loss forecasts (models) – stressed due to the pandemic -- to the underwriting criteria would not only streamline underwriting criteria, but would also allow a firm to incorporate risk-based pricing into their loans thereby getting properly compensated for the likely increase (or, could be, decrease) in expected credit losses attributed to impending macroeconomic conditions.

4.2. Borrower Accommodations

Although the COVID-19 crisis is a global pandemic, some countries have experienced much greater economic disruption. The US and UK are notable examples for which to study the response of lenders to these extreme impacts.

4.2.1. US Forbearance Programs

As COVID-19 spread across the US and as the subsequent shutdown of much of the US economy occurred, the unemployment rate shot to its historical high on record since the Great Depression. In a nation with significant numbers of its people living paycheck to paycheck, the first challenge for those who lost their jobs was how they were going to pay their bills and, most importantly, how they were going to stay in their homes by paying their mortgages. With millions of people losing their jobs and without money to pay their mortgages thus facing mortgage default and eviction, the US Congress passed the CARES Act, and the GSEs, Fannie and Freddie, have installed mortgage forbearance policies to prevent foreclosure for people who are impacted by the pandemic. Similar actions have been taken by banks across the financial industry to deal with the inability to pay the debt obligations to help those impacted by COVID19.

If a borrower takes the forbearance option, his or her loan will not be recorded as delinquent during the forbearance period even though there is no payment and cash flow to their lenders during that period, up to the maximum allowable time permitted by the forbearance policy. For mortgages, all forbearance requests were approved. In other product categories, forbearance programs were designed by lenders. Some of those programs mirrored the automatic approval for mortgage forbearance. Other lenders required proof of hardship in order to qualify. The laxness of the forbearance criteria will have many downstream impacts on loan performance and portfolio losses.

There was potential for immediate higher losses. In addition to financial stress on consumers, many states suspended all repossession activity. Repossession is the single greatest loss mitigation tool available to auto lenders. Repo generally occurs before a loan faces mandatory charge-off, and a lender is able to record the loss after post-repo auction proceeds are applied. With repossession suspended, there was a prospect for a higher number of loss events occurring with greater severity

These losses have not materialized in 2020 due to government support of borrowers and due to widely available payment extension plans offered by lenders. Both of these factors are still in place, though most lenders' pandemic extension programs have been scaled back or terminated by October 2020. These programs have helped borrowers who were distressed during the lockdown, but returned to stability afterwards.

4.2.2. UK Payment Holidays

In March, as the pandemic proliferated, the Financial Conduct Authority (FCA – one of the UK's regulatory bodies for financial services) issued guidance to mortgage lenders regarding the provision of a Payment Holiday (PH) to customers who are, "*experiencing or reasonably expects to experience payment difficulties as a result of circumstances relating to coronavirus…*" Guidance was issued to extend PHs of up to three months initially, extended in June to six months, effectively on a 'self-certified' basis. At the same time, the Prudential Regulation Authority (PRA – the arm of UK regulation focused upon prudential oversight) issued complementary guidance as to how PHs should be treated in capital models, and issued considerations for impairment modeling – though, of course, this must also be undertaken in discussions with External Auditors.

In addition to mortgage lenders, instruction was also issued on the same basis to firms offering wider credit facilities, credit cards, unsecured fixed term loans, current account overdrafts, high street finance, rent to landlords (who, in turn may have their own Buy To Let related lending exposures), housing associations, etc.

The extension of a PH facility was just one of several macro-prudential measures emanating from the UK government, others including the 'furlough' scheme to support employers and their employees, grants and loans to SMEs, etc. Unprecedented times require unprecedented measures. Though there are examples of payment moratoria issued through government statute as a result of, for example, natural disaster – this was the first time such an instruction had been issued in the UK and extended to an entire nation.

4.2.2.1. How are Payment Holidays Treated?

Essentially, in the UK, models are 'blinkered' to PHs per sé. Guidance issued by the PRA stated,

"*Our expectation is that eligibility for, and use of, the UK Government's policy on the extension of payment holidays should not automatically, other things being equal, result in the loans involved being moved into [IFRS 9] Stage 2 or Stage 3*[6] *for the purposes of calculating ECL or trigger a default under the EU Capital Requirements Regulation (CRR). This expectation extends to similar schemes to respond to the adverse economic impact of the virus.*" (Dear CEO Letter, Bank of England, March 2020)

Similar guidance was issued to ensure that customers' credit reference files were not adversely affected by PHs, effectively *freezing* credit files whilst PH arrangements are in place. Rather than being centrally administered through a PH 'flag' submitted to the CRAs,

---

[6] Stage 2: Significant Increase in Credit Risk, relative to origination, evident. Stage 3: Defaulted

individual lenders were charged with ensuring that increasing arrears positions were not submitted, given the PH 'hold' status over the 3/6-month period. Additionally, these have not been reflected as arrangements – something which a measure of forbearance might ordinarily be recorded as.

Whilst this protected consumers' credit files while the PHs were in place, it also begs a number of questions in the short, medium and longer term; what will be the impact when the big freeze starts to defrost? Who is likely to take a PH holiday both now and as the pandemic unfolds, i.e. will the demographic change? What happens when government mandated schemes draw to a close?

Figures published in the UK suggest that roughly one in five UK mortgagors requested a PH, and (perhaps, surprisingly) a slightly lower proportion furthermore requested a PH on their range of unsecured exposures. These exposures are significant drivers of credit file based variables used in models. Examples of this might include arrears status, credit line utilization, full payer or revolver, turnover – all potentially used within internal models, and within widely used bureau scores (like FICO, the UK credit reference agencies offer a number of 'generic' scores) – often used in their own right or as continuous or classed variables within models.

4.2.3. Impact Upon Models in Production

Unfortunately, as a modeling community, the lack of precedence (and data) for loan accommodation programs (forbearance, payment holidays, or others) presents a significant challenge. Not only do we need to forecast likely outcomes, or a range of potentially plausible outcomes, but we also have a spectrum of stakeholders with differing motives, expectations and acceptable landing pads to assuage.

Capital model suites under the Internal Ratings Based (IRB) approach and CCAR, and impairment model suites under IFRS9 and CECL (Expected Credit Losses) typically include credit reference data and / or bureau scores, as do operational decisioning models which, themselves, also often feed into capital and impairment models. In the absence of explicit loan accommodation *flags* against individual exposures at the credit reference agency, some of the questions we should ask ourselves as modelers and lenders are:

- How well do we know our customers and underlying portfolio credit quality?
- What 'type' of customer takes an accommodation – immediately and further down the line?
- What is the rationale for taking it?
- What are the potential impacts of credit file *freezing* in an already uncertain environment?

- What information is available to us, in order to make informed risk-based decisions as the situation unfolds?

These are non-trivial questions in the current environment. They pose particular issues for the majority of lenders who do not have transaction level visibility into their customers' checking / current accounts and are therefore reliant upon credit reference file data as a 'proxy' for this. Solutions (typically a hybrid of automated and manual) implemented since the onset of the COVID-19 crisis -- in terms of collecting customer information around employment status, income and expenditure, and future prospects – have needed to be developed and rolled out at speed. Content is therefore rudimentary and largely unverified.

In addition, although the credit bureaus / credit reference agencies have enhanced their products and service in response to the COVID-19 crisis, an inherent time delay still exists in data being updated. Typically, though more timely solutions are being rolled out, an event in (say) March would be unlikely to be visible on a credit file until May. Contrast this limitation in the current environment to real-time sight of consumer turnover on a checking / current account. Add further to this that economic outlooks are changing more frequently than perhaps ever seen before, and any range of consensus has yet to converge.

4.3. Collections

Following application decisioning, or post loan approval, is loan servicing, also referred to as collections. Similar to application models during benign economic scenarios, the majority of clients, solidly underwritten, pay per their contractual obligation. For clients who do miss payment, or are late, are typically targeted for collection treatment. In financial parlance, the collection "treatment" usually deals with (1) who to call vs. not call, (2) when to call relative to their days late, both before or after a payment due date, (3) when to call to optimize agent-client contact likelihood and (4) how many times to call before leaving a message. In some jurisdictions, there are limits to how often a client can be called or pestered, per se, by a financial collector.

In a pandemic situation, such as the current situation with COVID-19, financial institutions have made numerous client loan accommodations, some initiated by clients and some sponsored by government support (which provide financial guarantees/backing to financial institutions that make them). Those accommodations typically take the form of deferred payments with either a pseudo term extension with interest accruing throughout the remaining life or a re-amortization such that payments and interest are re-apportioned through the remaining life of the original term.

Regardless, the core "problem" during the current pandemic is how to allocate collection personnel and resources to maximize returning clients to good standing. While there are myriad vendor-based collections models and systems, there is no analogous pandemic situation from which to model, at least in modern times. (The 1918 flu pandemic would arguably be similar but obtaining consumer payment data from those times is neither practical nor likely possible.) In the absence of hard data against which to model, and in turn leverage to determine who to (not) call, would be to take a qualitative tack.

That is, to help differentiate who may be more likely to pay vs. not pay, financial institutions may consider such dimensions as:

1. expected loss
2. pre-COVID-19 payment performance
3. deposit changes at the lender's bank (if any)
4. payment coverage vs. deposit amount at the lender's bank (if any)

These dimensions would likely be factors either in or considered for a collections "roll" model, which would then be used to inform the order of a collector's calling queue, combined with other information resulting from known successful strategies (e.g., when to call relative to loan payment due date or loan accommodation expiry date).

Once a financial institution determines what data elements they both need and have readily available, they could then qualitatively combine them, e.g., 10%, 15%, etc., with their respective values, and sum those weights to yield a qualitative score, so to speak.

This suggested approach would not only provide a rational approach for banks to help with collection efforts but would also provide a solid basis for data collection on which a "real" model could be developed. Again, it's not that financial institutions don't have data or models for collection efforts, but rather they, or even the financial industry, does not have historical loan accommodation data, due to a pandemic, from which to estimate a model.

It would be remiss, however, to neglect to mention one important consideration here. That is, while this model would be fine for prioritizing collection calls, it would not be prudent to serve as the basis for decisioning action for a particular client. Doing so could risk incurring compliance risk issues due to decisioning the client's credit relationship, e.g., forcing a charge-off or renewing a loan accommodation, specifically based on the qualitative score.

5. Model Monitoring

Model performance ongoing monitoring is a standard practice at large banks as an effective way to evaluate the necessity for model adjustment, redevelopment, or replacement. Appropriately selected metrics and thresholds determine probable model break points up front that enable timely capture of model degradation. A number of well-known changes attributing to model performance breaches include, but are not limited to, model inputs, assumptions, products, exposures, activities, clients, or market and environment conditions.

Regardless of the sources, the impact of those changes is enlarged forecasting errors and hence under- or over-prediction.

The COVID-19 pandemic, and the subsequent recession caused widespread model degradation. A great deal of effort has been put in place by bankers and regulators to mitigate the unprecedented model deterioration, among which are identification of specific root causes, and quantification of the impact at the individual model level as well as the aggregate model.

5.1. Sources of Model Deterioration

Model risk occurs during model development and use, demonstrated by model errors before being exemplified in other forms.[7] Fundamental errors can take place in any stage of a model build from target design, data sampling, theory application, variable selection, model estimation, diagnostic testing to implementation. Presence of model errors take various forms. A series of weak or failed hypothesis tests may suggest multicollinearity, autocorrelation, selection bias, overfitting, etc. If not properly corrected, those errors will cause the model to produce incorrect or less accurate estimates. Models with such errors are prone to faster deterioration and even breakdown. Application of erroneous estimates to business decision making not only elevates model risk but, through their supporting role to the business, could also increase other risk such as credit, market, interest rate, liquidity, and so on. Model risk stemming from fundamental errors can be mitigated during the development via sample selection, data manipulation, testing alternative model specification, or applying mechanical and technical solutions, though it can never be eliminated. Understanding and quantifying model risk helps improve model performance and increase model longevity. The magnitude of such model inherent risk is a common factor of performance deterioration and determines the speed and size of breaches.

A regression model is developed to construct relationships between target and explanatory variables using realized events. Put differently, we analyze past correlations in order to predict the future. The degree of difference between future and past correlations is the primary reason most models degrade over time and must eventually be replaced. Properly identified metrics and thresholds create the ability to measure the difference between the current data and historical observations, and the subsequent timely capture of material deterioration.

Rapid advancement of machine learning algorithms in the recent years has caused the financial industry to begin adoption of alternative modeling methods that allow nonconventional pattern recognition and in-production live adjustment. Rapid growth in data accumulation and computing resources have fueled this expansion. Although such progress has the potential to allow models to identify new, nonlinear patterns, it does not change the

[7] Cox, D. Role of Models in Statistical Analysis, *Statistical Science*, 1990, 5:2, pp. 169-174. file:///H:/My%20Research/Role_of_Models_in_Statistical_Analysis.pdf, last access on 8/5/2020.

fact that future prediction continues to largely rely on historical experiences, nor does it significantly mitigate model risk. Therefore, performance monitoring becomes more important in the new era of model development.

Model deterioration resulting from economic shocks does not necessarily mean the conceptual soundness of a model is in doubt but does call for immediate attention to confirm the model remains viable. Large banks tend to conduct sensitivity analyses to quantify the magnitude of forecasting variance, and then leverage the analyses to form model overlays and other adjustments including in-model adjustments that directly change the value or range of the explanatory variables. The most basic sensitivity analysis typically implies estimating changes in model output by shocking individual or combined independent variables including macroeconomic indicators and portfolio risk drivers. Model developers can also perform a more sophisticated sensitivity analysis to quantify business risk exposure change due to portfolio mix change driven by the overall outlook of market and economic conditions. This exercise is considered to be an assumption reevaluation. For example, model developers can re-estimate the portfolio Probability of Default (PD) by increasing the percentage of high-risk applicants if anticipating higher losses from newly originated accounts during the pandemic. Similarly, a sensitivity analysis can be designed to help quantify loss magnitude changes by assuming higher default rates for the same underlying risk score bands.

It is generally recognized that the underlying economic principles are not fundamentally broken. Rather, their impacts based upon past empirical analyses are heavily distorted, diluted, or delayed in the presence of massive government and bank intervention to keep the economy stable. In the recent Covid-19 response, federal and state stimulus checks and generous application of forbearance programs by lenders created sustainable cash flow for obligors and hence disguise, to a certain extent, the magnitude of financial stress on borrowers. Among several attempts to address the loss of model accuracy and reliability, the most frequently sought-after approach is the application of overlays and adjustments to correct the widening forecasting errors.

It is worth noting that some shocks introduced to econometric models are short lived and self-correcting. Careless overlay application can further worsen the model performance and increase model risk. Performance outcome evaluation becomes more critical during the epidemic. Instead of reacting to the breaches, the model developer and user should first assess if existing metrics and thresholds, and their monitoring and reporting frequency, need to be modified to prevent overreaction to severe but transitory shocks to the model. For example, model developers and model users may need to incorporate ultra-short-term metrics or change thresholds in order to manifest the impact of short-term shocks to the model estimates. Changing the monitoring frequency will not mitigate model deterioration but may offer timelier detection of model misbehavior. For example, if the state level unemployment rate is a model input and has been assessed on a quarterly basis, a month-over-month measure can be added to provide granular information on unemployment rate each within a quarter. Additionally, model performance with and without adjustments should be compared as well to further evaluate the overall effectiveness of model overlays and overrides.

5.2. Key Components of Ongoing Monitoring

Effective model ongoing monitoring, analysis, and reporting are critical to identifying, controlling, and managing risk. A comprehensive and effective ongoing monitoring program includes evaluation of estimate accuracy, input and output stability, and model robustness.

The role of statistical models can be described as substantive, empirical, and indirect, even though combinations of those roles are probable in a specific application. The majority of models developed and used by the financial industry are probability based, and therefore fall into the empirical category. Accuracy measurement therefore is critical for models with a predictive nature.

In lending, accuracy can take two forms, absolute and relative. For a pool of loans, we can ask whether the risk distribution is well predicted through time or if the relative ordering of the accounts is effectively predicted.

For absolute accuracy, the smaller the variance between model estimates and the actual observations, the more accurate the model. Accuracy can be gauged by several well-known metrics such as Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Deviation (MAD)/ Mean Absolute Percentage Error (MAPE)/Mean Absolute Percentage Deviation (MAPD) etc. MAE is the most natural measure of the average error magnitude while the RMSE is an unambiguous measure of the average error magnitude.

For relative accuracy, discrimination statistics such as AUC, Gini, and K-S can be monitored for degradation through time. Notably, risk ranking has been found historically to be more robust through economic cycles than the absolute forecast accuracy.

Comparisons of risk scores to odds of default are also commonly monitored for degradation through economic cycles. However, these are often misinterpreted. Shifts in the slope of a score-to-odds line do not indicate a model failure in itself. Rather, only if the relationship becomes so flat that accounts can no longer be ranked to within the confidence interval of the linear relationship can the model no longer be used for risk-based pricing without a recalibration.

A model, or equation, is “stable” if it can be applied to different time periods of data without significant loss in its prediction accuracy. A stability measure evaluates if the distribution and value range of inputs and outputs have shifted over time such as between the production data and the development sample, or between different production data. The

Population Stability Index (PSI)[8] is commonly used to evaluate model stability and is a useful tool to detect population shift due to business strategy change or shocks to the macroeconomy drivers. Coefficient re-estimation using production data is another common approach to assess model stability, which is particularly useful for models requiring regular refitting with updated input data.

Model robustness analysis is also important. It ensures developers' conclusions hold under different assumptions. Put differently, robustness is the model's ability to perform effectively under internal and external disturbances. Robustness can be viewed as a model resilience or sensitivity given dramatic changes to model inputs. Model robustness assessment is typically performed as part of the diagnostic testing during the model development. Some statistical tests should be performed again periodically during the model's ongoing use to measure changes in model robustness.

6. Post-Model Adjustments and Model Overlays

We know from experience – or we are advised by those who purport to know – that *all downturns are different*. In terms of root cause, this is likely true. Pandemics are not a new phenomenon – but they are largely unprecedented, in terms of having a major impact upon a globalized banking system. Regretfully, the scale, the driver(s) of the crisis, the macroeconomic and governmental intervention measures, and the potential end position, make the COVID-19 crisis potentially unique in this respect.

Despite these, even when those factors are considered, the log-normal shape of emerging defaults and losses in a downturn / crisis tends to be a familiar one – no matter what the underlying drivers were, or what was done to manage it. A sharp ascent to the peak (into downturn) and a prolonged descent (through recovery) to more normative levels, is the familiar trend. Attempts to 'flatten the peak' in the UK, Europe and the US have certainly had some early success in consumer finance, residential mortgages and commercial lending, but the question to ask is: are all of these measures just delaying events – are we 'kicking the can down the road'?

In reality, we do not yet know the endgame of the pandemic and associated economics. Perhaps *endgame* is not even the appropriate expectation. If our complex, creative, innovative brains cannot accurately conceptualize the outcome and/or the timeline over which this will play out *generally* -- and *specifically* for those who have taken accommodations and those who have not but could need assistance at any time in the near term future -- what chance do our *simplified models of life* with their limited inputs stand?

[8] Karakoulas, G. (2004). "Empirical validation of retail credit-scoring models." *The RMA Journal*, 87(1), 56-60.

They have been trained on the past, with no prior experience of such a pandemic and no mass accommodation data or outcome. Models assume relationships hold and the future will, ceteris peribus ("*other things being equal*"), be the like the past. They will need to be retrained – potentially redeveloped – when the data exists to do so and when existing relationships have been confirmed or new relationships have formed. Given this, there is limited merit in frantic kneejerk attempts to repair models when we as a banking community do not really *know* what 'fixed' will look like.

In the meantime, we should watch and learn with the utmost diligence, maintaining prudence in our capital and impairment levels through scenario and sensitivity[9] stressed testing of our models, and evolving our risk appetite ranges of acceptability to potentially manage a 'new norm'.

Before continuing with this train of thought, all is not lost for us modelers, and COVID-19 is not a reason to discard well established model theory. In these days of uncertainty, we should still have absolute confidence that our existing models are the best quantitative tools to establish a range of plausible outputs which inform wider discussions, based upon a number of expectations we can draw upon with a perfectly reasonable degree of confidence:

Discrimination:	Most lenders have a number of years' evidence in retail lending from internal capital models that, in a downturn, model discrimination tends to hold up well and (for example) our Gini's tend to improve – or certainly hold fast. Volumes of bads fluctuate through time, but the fundamental drivers of such events tend to be the same.
Accuracy:	Model calibration is expected to slide over time; however, if we assume that discrimination holds, we can adjust for accuracy in our forecasts through scenario analysis and sensitivities.
Stability:	Like accuracy, stability is likely to be impacted at both population and characteristic levels. Again, scenarios on sensitivities can be run to emulate this phenomenon.

Using these as our theoretical basis, we can use both the models we have and the multi-dimensionality of our creative thought processes to create a dynamic hybrid (of models and minds) to derive forecasted capital / impairment outputs. Regularly refreshed data emerging from enhanced sources of information are becoming available:

- additional, though perhaps rudimentary, **internal** data captured during the accommodation customer contact

[9] When referring to *sensitivities*, reference is made to univariate changes – often referred to a 'turning the dial up/down'. *Scenarios* typically refers to multivariate changes to inputs, such as a revised economic forecast.

- enhanced credit bureau feeds, **external** data made available in a timelier fashion (though still not current account 'real time')
- additional **macroeconomic** data produced more often; how has government intervention impacted upon the relationships between unemployment, GDP, default rates, etc.

These can inform a discussion across a range of the plausible impacts and the temporary overlay of Post Model Adjustments (PMAs). Let senior management draw upon their experience and subject matter expertise of their portfolios to set the expectation and be accountable for the level of overlay. Good governance, evidencing sound thought and discussion based upon knowledge of their business in the context of the pandemic's effects – and the level of accommodation activity observed during and following the period of moratorium – is the key for our regulators and auditors. This is not a time for statistical overload – better to be roughly right than precisely wrong. Rather, it is a time for risk-based pragmatism. In the meantime, we test, we learn, and we regroup with our model suites to prepare for the future.

Returning to the 'new norm' train of thought . . . as more data emerges, we can assess whether discrimination is indeed holding up with greater certainty, or if it is time to rebuild. Additionally, as the outcome emerges, accuracy measures may become more credible and prudence can be dialed out – the 'jaws' of credible ranges can be narrowed as we hone our predictions.

6.1. Account-level Considerations

While COVID-19 and its economic fallout is enormous, the impact on borrower payment ability and behavior are not equal. For those whose jobs are impacted in the pandemic, or for those who have amassed enough savings or assets in the past, the lender should expect their payment behavior won't be changed. However, for those who lost jobs and income sources in the pandemic and, at the same time, don't have enough savings or assets to cover the economic fallout of lost jobs, the lender would expect payment delays or default. Thus, the default or loss forecast model should be adjusted accordingly to incorporate the legislative and policy changes during the pandemic.

Most lenders have used data and models for many years, and we have evidence / experience of model performance across economic cycles. The historical relationship between cyclical trends and risk appetite is relatively well established. The fundamentals have not (yet) changed; drivers of downturn change, but outcomes do not. To attempt to quantify, even heuristically, what the future performance of an account will be following an accommodation, several questions should be considered.

6.1.1. Borrower Attributes with Accommodations

As mentioned above, not all borrowers are actually impacted by COVID-19 economically, thus who will actually take up the accommodation option from their lenders is an important factor to impact on the model performance during the pandemic.

Bureau and internal scores pre-pandemic suggest that those who took an accommodation initially are generally higher risk: more unsecured exposures, higher utilized credit lines, and more credit active. However, immediate post-accommodation performance has been good and volumes of those extending from three to six months, or recent adopters of PHs have been lower than anticipated.

That said, the accommodation extension / recent adopter sub-population are more likely to be those who have genuine short-term affordability issues – as opposed to early adopters who appear to include those taking an accommodation as an insurance measure. Early indications suggest we should prepare for a higher failure rate within this sub-population.

Over the accommodation period, unsecured exposures appear to have reduced initially, indicating perhaps that some customers have used mortgage accommodations to de-leverage on their credit cards, etc. However, in the UK this is returning to pre-pandemic levels and early evidence suggests that those taking accommodations are starting to exceed pre-accommodation unsecured exposure levels.

One approximation for measuring the risk of borrowers in accommodation is to take the fraction of borrowers who accept accommodation over all borrowers in the same sub-population. This can be one of the metrics used to adjust payment related behavioral models. This rate could be initially estimated based on the historical economic and behavioral data, then consistently updated and adjusted by the real time accommodation acceptance data from the loan servicing system.

Also useful is to differentiate between early adopters, late adopters and those who have extended their accommodations. It is likely their motivations and risk levels differ.

Beyond this, both accommodation and non-accommodation sub-populations place particular focus upon pockets with higher unsecured exposures or trends of accelerated unsecured balance increase / utilization limits. This helps to establish early contact strategies.

6.1.2. Utilization during Accommodation

Not everyone who accepts an accommodation will actually miss the scheduled payment, and it is important to use limited data to estimate the percentage of borrowers who have actually missed payments after taking up the forbearance option. Once a borrower has actually missed a payment after taking the accommodation, it is reasonable to assume that the borrower will miss all the payments in the full term of the accommodation period, and the loan payment behavior should be assumed to roll straight to default or charge-off.

6.1.3. Exiting Accommodation

The Federal Financial Institutions Examination Council (FFIEC) provided advice to FIs on how to work with borrowers as they near the end of their accommodation periods.

There are three likely loan performance outcomes: 1) those who would catch up on the payment requirements during the accommodation period or after the end of the accommodation, 2) those who would ask, be qualified, and successfully entered a loan modification program with the lender, and 3) those who could not catch up on the payments or enter a loan modification and eventually default or charge-off from the loan book.

Among the above three borrower behavior metrics, the accommodation outcome is the most challenging to estimate for the model overlay as it takes time to observe this payment behavioral outcome. Historical data in similar programs, if it exists, could be leveraged, but the only somewhat similar programs in wide use are for deployed military personnel. Such programs are neither widespread enough nor similar enough in risk characteristics to provide an effective comparison.

6.2. Portfolio Performance Overlays

Over the last decade, financial institutions around the world have put significant resources into building stress testing and capital models. Immediately following the Hong Kong SARS Recessions in 2003, pandemics were a common stress testing scenario.[10] The pandemic crisis modeling approach has been to build a model that responds to macroeconomic drivers, translate pandemic impacts into those same key drivers, and run the forecast as usual.

That approach should have worked for the COVID-19 crisis, if not for trillions of dollars in US government assistance and more trillions around the world, and for the accommodation programs from lenders. Assuming an FI's model handled extreme events appropriately, those models would predict performance if not for this massive assistance to consumers. Such forecasts set an upper bound on what could have happened through 2020 and provide some guidance to what could happen in 2021 and 2022 after these assistance programs end.

For questions of portfolio forecasting, FIs will not want to rely purely on the upper bound expectations from the models, nor should they assume that the low loss levels of 2020 will persist. Instead, portfolio-level model overlays are needed to adapt the model forecasts to the realities of the COVID-19 crisis, now and under scenarios for the future.

6.2.1. Designing Overlays

The hope of government assistance and loan accommodation programs was to sustain the public until the pandemic was under control and the economy could return to normal. If the

[10] Breeden, J.L., "Stress Testing for Pandemics", RMA Journal, April 2020, pp. 34-39.

novel coronavirus could have been eradicated in a few months, that approach would have provided a bridge across the pandemic months to when the economy could restart.

Unfortunately, COVID-19 could not be contained, and a vaccine is not likely to be widely available globally until well into 2021. Consequently, these assistance programs cannot create a perfect bridge. When the US government assistance ended in August, the economy had moved past its worst point, but was not yet back to pre-crisis levels. More government assistance may yet be extended, but economic recovery is broadly acknowledged to be a slow process from here.

TABLE 1: DIAGRAM OF CATEGORIES OF BORROWERS RECEIVING FORBEARANCE.

Received Accommodation	During Assistance Period	After Assistance Period
1. Yes	1.1. Those who were not at risk	1.1.1. Cure from accommodation and stay in good standing
		1.1.2. Lose their jobs later and are default risks
	1.2. Those whose risk was deferred	1.2.1. Cure from accommodation by finding new jobs
		1.2.2. Default after assistance is used up
2. No	2.1. Those who were not at risk	2.1.1. Continued good standing
		2.1.2. Lose their jobs and are default risks
	2.2. Those with non-crisis risk	2.2.1. Continue to be a default risk

To translate programs such as forbearance into a portfolio overlay, we can consider the categories as shown in Table 1. If we consider US mortgage as an example, Figure 4 shows that a peak of 12% of all GSE-backed loans[11] went into forbearance.

Approval for forbearance was automatic for GSE-backed loans. Therefore, we know that some of those who accepted forbearance will fall into category 1.1, which could be called Strategic Forbearance. They had no specific risk at the time they accepted the accommodation, but they were concerned that macroeconomic conditions could create future risks. For products besides GSE-backed mortgages, forbearance policies were up to the lenders. Those that required proof of hardship would have lower forbearance rates, so the overall forbearance rate will correlate to the strategic forbearance rate, the ratio of 1.1 to 1.2.

[11] Fannie Mae and Freddie Mac are examples of US Government Sponsored Entities (GSEs) that purchase home loans.

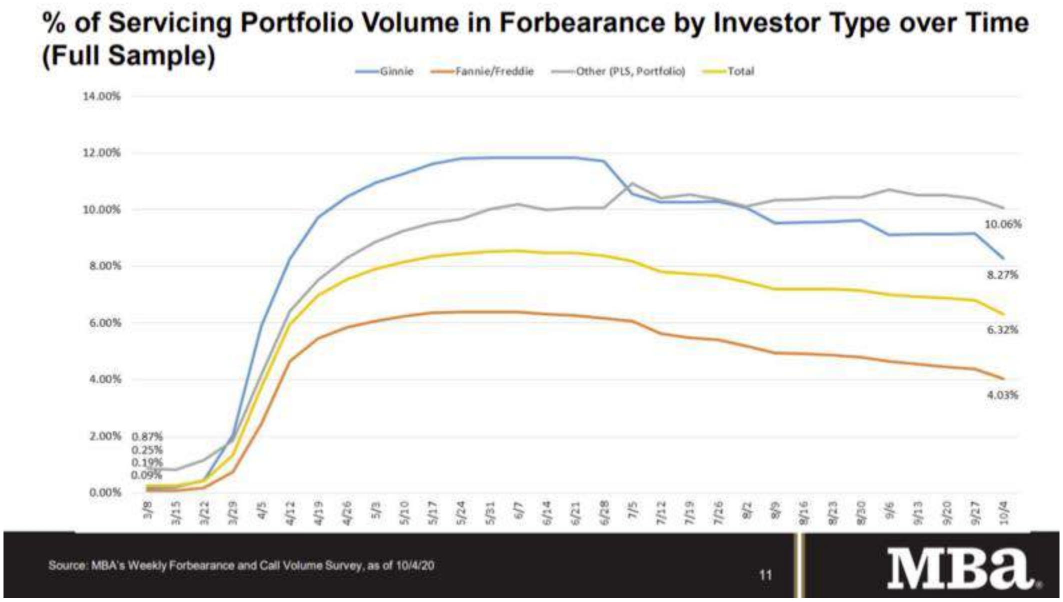

FIGURE 4: FORBEARANCE RATES FOR US MORTGAGE.

Of the strategic forbearance group, if the economy continues to improve, the majority can be expected to cure as in 1.1.1. In fact, some who accepted forbearance never stopped making payments and could be considered an automatic cure.

Some who were not at specific risk at the time of accepting the forbearance could yet lose their jobs and default later. The ratio between 1.1.1 and 1.1.2 is largely a function of future economic conditions.

Of those who were at risk at the time of accepting forbearance, category 1.2, some will return to work while others may eventually default. The ratio of 1.2.1 to 1.2.2 is again largely a function of the economy.

Of those who do not go into forbearance, for GSE-backed mortgages this would either be because they did not feel at risk or were unaware of the program, category 2.1. For other products, they could have been denied forbearance, because no hardship was shown. Some who do not go into forbearance nevertheless may be at risk because of pre-crisis difficulties. Most who do not go into forbearance and show no specific risk are assumed to stay in good standing in the future, category 2.1.1, although some may yet default due to future job loss. Thus, the ratio of 2.1.1 to 2.1.2 is again a reflection of future economic conditions.

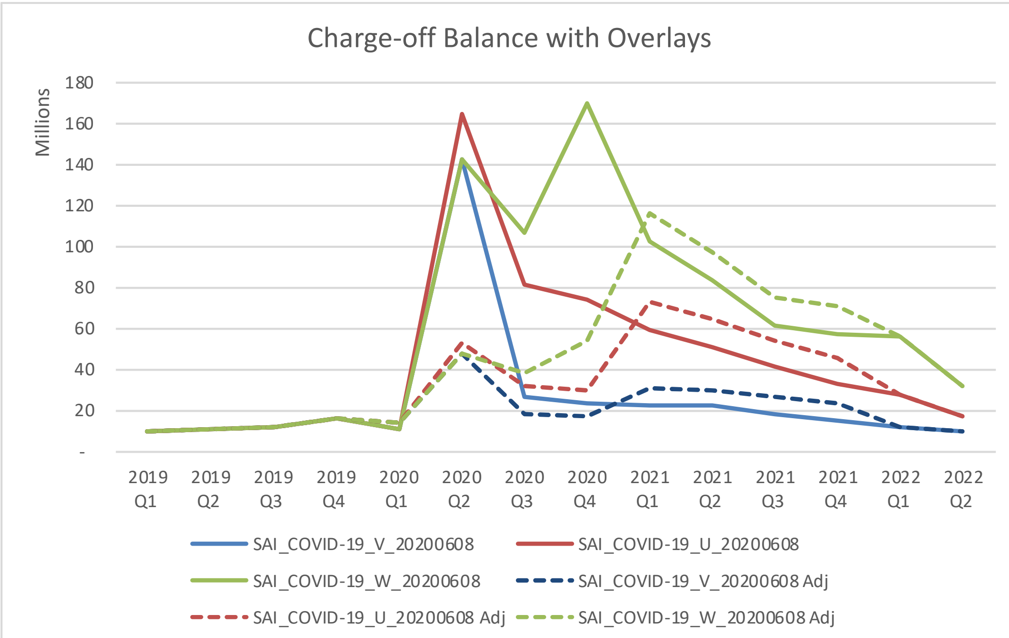

FIGURE 5: HYPOTHETICAL NET CHARGE-OFF BALANCE WITH AND WITHOUT THE DESIGNED OVERLAY FOR ALLIANT FCU.

To create an overlay, we need to create estimates of these proportions and an idea of when those losses would occur. For example, Figure 5 shows a loss forecast for the COVID-19 recession where the solid lines are the original forecast under V, U, and W-shaped macroeconomic scenarios. The dashed lines show a result with overlays.

The overlay assumes those who default without forbearance, category 2.2.1 above, is 25% of the original loss forecast. The cure from forbearance, category 1.2.2 is 75% of those at risk, category 1.2. Those who do not cure are assumed to default over a 12-month period after forbearance. Also, there are future defaults due to economic conditions, categories 1.1.2 and 2.1.2, as predicted from the stress test models. All of this looks only at forbearance and economic conditions, which are influenced heavily by future government assistance.

6.3. Self-Adjusting Models

Self-adjusting models are critical for rapid pattern detection in fraud, AML, etc. It is a means of rapidly learning new patterns on the assumption that those patterns will persist for a useful time period. In the case of accommodations during COVID-19, by the time enough data is available to learn the new pattern, the accommodations could end and the model will be wrong the other way. In either case, it becomes difficult to design overlays when the user cannot probe exactly how the model is adjusting itself.

In this case, information about accommodations was likely missing from the initial training set. Although ML methods can be very good at adaptation, models in production are unlikely to have the ability to incorporate new input variables with no prior history. Humans are good at incorporating new sources of information, because we have a meta-model of how the world works and can make educated guesses about the impact on a forecast. The current generation of AI/ML models lacks this.

Therefore, models that employ self-adjustment capabilities are actually likely to be wrong longer through the current crisis and are more difficult to design overlays for.

6.4. Risk of Overlays

Using overlays is often seen as a failure of the underlying statistical model and negative overlays (reducing loss reserves) have historically been reviled by auditors. Still, these are extraordinary times. Rather than lament the inability of models to predict what has never before been experienced, we need to recognize the valuable role of management and model monitoring in adjusting our expectations. That the overlays will be negative in 2020 is an indication that government assistance and lender portfolio management are having beneficial effects for borrowers.

If overlays become completely untethered from models, a different risk arises. Although the expert opinion of seasoned professionals can be quite good, having no model for guidance can lead to estimates driven by wishing. Even with the extreme conditions of 2020, a process grounded in available data should be followed in establishing forecasts. Understanding model risks means understanding how and when they can be leveraged even through this crisis.

7. Modeling after COVID-19

By the time the COVID-19 Crisis becomes the COVID-19 economic recovery, most of the preceding sections will have become lessons learned and will provide best practices on how to better monitor data, monitor models, and manage models through the next crisis. However, as soon as 2021, developers and validators will have to decide what to do with the data from 2020. With its extremes and anomalous outcomes, how can we utilize this data in modeling and model risk assessment?

7.1. Data Strategy for COVID-19 and Beyond

While the previous discussion on overlays can guide model adjustment to count for the impact of COVID 19 pandemic in the short term, a different kind of adjustment is required to support future development. If FIs want to use data from the COVID-19 crisis in their modeling and validation, they will need to carefully consider how the outcome should be measured. Many portfolios do not have data back through a previous recession and will have no choice but to include 2020.

Once a borrower takes an accommodation, even though no payments are made, there is no increase in delinquency, default, or foreclosure. Any model using these loan performance measures as the model outcome will not capture the real economic impact on borrower payment behavior by the actual economic fallout of the COVID 19 pandemic. Instead, using these data in model development will distort the normal relationship between the macroeconomic drivers and the loan performance behaviors.

If the data from the pandemic is not to be discarded, it will need to be adjusted. Rather than using actual default as the target, a would-have-been default could be estimated and modeled. Whenever a borrower takes an accommodation, such as a COVID-19 accommodation or other policy initiatives, then misses a payment, it is not a delinquency or default from the policy point of view and needs to be recorded and reported as such for compliance purposes. However, the borrower is actually late in payment from behavior point of view and this behavioral outcome actually has internal or external drivers behind it, e.g., the economic stress caused by a lost job during the pandemic. Therefore, it is important to define both of the outcomes separately.

To estimate these would-have-been defaults, not all accommodations can immediately be recorded as defaults. Those who continued to make payments and those who could have made payments but were simply hoarding cash should be excluded. The latter group requires some approximation. Those who cure from default set a floor for those who could have defaulted initially, but we know that accommodations will prevent some of what would-have-been defaults. That assessment can look at how much the economy improved between the beginning and end of the accommodation period and how much screening was performed to determine who actually needed accommodation. If proof of job loss was required, then nearly all can be assumed as would-have-been defaults. With no screening as with GSE-backed US mortgages, then the change in the economic environment might be the best way to statistically generate what would-have-been defaults.

This approach requires building a data infrastructure to collect and store both policy driven payment outcome data for normal business operations, such as compliance reporting, fee income collection or reporting, and would-have-been defaults data that can be used for strategic analysis and behavioral modeling purposes. This data system should be integrated well for data quality checks and other business reporting and operations. A single assessment of would-have-been defaults should be maintained so that different modeling teams all use the same assumptions and assignments.

While the policy driven data collection component of this dual data strategy needs to be flexible enough to be customized to support different business strategies and policy driven based data definitions, such as COVID-19 pandemic accommodations, the borrower would-have-been default component should be consistent across different strategies and economic cycles to record the actual borrower behavior in different business and economic environments. The consumer behavior data collection and storage system in particular should also be built in an environment and infrastructure with good data architecture and governance structure, and that is readily accessible for business analysis or modeling purposes, including model development and model validation.

Information is power and data is the most valuable asset in the age of machine learning and AI. Therefore, building the right technology infrastructure to capture and store the right data is key for competitive advantage in the banking and lending industry.

7.2. Behavior Scores for Credit Risk

Using data from the COVID-19 crisis will be most problematic for building behavior scores. Even if the outcome is adjusted to predict would-have-been defaults, the normal inputs are scrambled. It will be more difficult to simulate the delinquency patterns, payment patterns, and many of the usual factors. As for external inputs to the models, problems in both timeliness and definitional shifts will be problematic. Even with all the recent popularity of Machine Learning methods, those algorithms cannot substitute for scrambled or missing patterns.

If the data from 2020 is not used in future modeling, the implicit assumption is that consumer behavior post-assistance will return to a pre-crisis normal. For models that primarily use payment patterns and credit bureau inputs, this is a reasonable assumption. The basics of consumer finances still drive default.

The greater risk is in models using alternate data sources. This has already been an identified problem, where using inputs like type of phone or operating system of login device is used to predict behavioral patterns. Aside from possible problems with unintended bias, such models often need to be retrained frequently so as to adapt to sudden shifts in the input factors. However, the patterns of 2020 are almost certainly not going to be predictive of 2021, and the patterns of 2019 and before may not be regained in 2021 either. Models using alternate data inputs may be largely flying blind in 2021 until new patterns emerge.

7.3. Origination Scores for Credit Risk

Origination score creation may not be as problematic as behavior scores. For those who have maintained employment through the crisis, 2020 is proving to be a good year financially. After employment and cash flow screening, lenders are finding credit quality to be better than average. Positive selection is a part of any credit cycle and not a sign of model failing. Previous origination scores appear to be working in 2020 with this positive selection adjustment. That suggests that 2020 data might be usable in the future for origination scoring with default modification for accommodations where they occur.

One area of caution with loan origination models is specialization. Just like in biological extinction events, the most specialized fail first in a sudden crisis. Through 2020, general purpose origination models have performed well enough. Highly specialized models run the risk that the patterns they learned could shift or disappear in a crisis.

Machine Learning models excel at finding pockets of predictability – special situations where a group of consumers behaves in ways that do not follow the overriding linear relationships. Those pockets of predictability are the first to shift or disappear when the situation changes. Lenders are likely to find that the simpler linear methods are more robust through the crisis than highly nonlinear methods, precisely because the simpler methods focus upon principles that do not change in a crisis, e.g. borrowers with a past history of delinquency and high utilization are at greater risk of defaulting – nothing fancy.

The COVID-19 crisis may also reopen debate about when to use through-the-cycle versus point-in-time estimates as discussed in "Risk-Grading Philosophy: Through-the-Cycle versus Point-in-Time". Should loans be priced for the current crisis for an "average" environment, or using a transition as with lifetime loss reserves under IFRS 9 Stage 2 and CECL? If the PIT to TTC transition is employed, which is sensible overall, what near-term environment do we use: the overall economic environment, the environment specific to the borrower in keeping with the K-shaped recovery, or the environment that could be in 2021 if no further government assistance comes?

7.4. Stress Testing

Stress testing models seek, at a minimum, to identify the relationship between changes in the macroeconomic environment to changes in portfolio performance at each point in time. Using data from 2020 will pose a number of challenges.

7.4.1. Data for Modeling

The point made earlier about using what would-have-been defaults as the target outcome will be essential to getting the timing right if 2020 data is going to be used in stress testing models.

Further, the dynamics of the input variables must be carefully considered. Any known data biases that are not corrected will then be assumed by the model to occur again next time. For example, using the BLS-reported unemployment rate will cause the model to incorporate an undercounting bias in times of severe stress. More reasonable would be to assume that BLS will better educate staff so that were this to happen again, the data would be accurate. Therefore, the input unemployment rate should be adjusted up to include the BLS estimates of the undercount, as shown in Figure 2.

For initial US unemployment insurance claims, the program allowing self-employed workers to file for unemployment is due to expire at the end of 2020. Assuming that is not renewed, if initial claims is to be a model input, it will need to exclude self-employed claims, as shown in Figure 1.

The general lesson is that each macroeconomic factor going into a model needs to be carefully considered for these sorts of biases. However, not all shocks should be removed. The large discontinuities in disposable personal income (FRED: DSPIC96) are not reporting errors or definitional changes. These are real impacts to consumers driven by government policies.

Timing of data availability has also rarely been considered when creating stress test models. If the stress test model is only for compliance for programs like the US CCAR, data timeliness seems to be a minor concern. However, this crisis showed that usable stress test models are an essential component of business decision making. Therefore, careful consideration should be given to timeliness.

In cases like the US Index of Leading Economic Indicators (FRED: USSLIND), even if it is provided in the future, this experience suggests it should not be relied upon. Among standard stress testing inputs, commercial real estate prices (FRED: COMREPUSQ159N) can be more than six months out of date, and house prices (FRED: CSUSHPISA) and disposable personal income (FRED: DSPIC96) can be two months out of date. These reporting lags mean the inputs must be predicted from more rapidly provided values from alternate services or those services should be the primary inputs.

For example, Zillow provides house price indices with negligible lags (https://www.zillow.com/research/data/). That data could be used to substitute the slowly reported DFAST HPI, or data like that from Zillow could be used to estimate what the DFAST HPI values will be.

Similarly, multiple services provide commercial real estate indices that could be used to replace or predict the DFAST CRE Index. Such steps should be taken the next time these models are recreated so that they are most useful through the next crisis.

Another key issue is that the scenarios provided for the Dodd-Frank stress testing are not broad enough to cover some issues that have emerged as key. Most stress test models do not segment by income tiers, sticking mostly with geographic segmentation. In the COVID-19 crisis, the greatest divergence has been between income tiers where lower income workers were impacted much more severely than middle and upper income workers[12]. Building models using more segmented macroeconomic data can make the models much more useful operationally, but may pull them away from regulatory compliance since government scenarios do not cover such variables.

7.4.2. Transformations

One of the best arguments for including data from the COVID-19 Crisis is to better understand nonlinearities during extreme events. This would only work if using the would-have-been default measure, but it would allow the stress test models to learn about diminishing impacts during extreme changes. The biggest part of those diminishing impacts is due to government intervention. The 2009 Global Financial Crisis and COVID-19 Crisis show that the more sudden and severe the crisis, the greater the response from

[12] https://tracktherecovery.org

governments. Therefore, sensitivities to changes in macroeconomic factors could reasonably assume something roughly similar in the future.

Capturing this nonlinearity can be accomplished with logarithmic, square root, hyperbolic tangent, sigmoid, and other transformations. We are unlikely to identify a clear best answer, but any can serve to avoid unrealistic forecasts.

If we do not use the data from 2020, the gap in the input data will be significantly worse than the gap in the target variable. Input variables are transformed with time lags and windows of change. The maximum lag considered, plus the maximum window for computing change, indicate how much the gap will be extended. For many models, this could mean that a 12-month gap for 2020 becomes a multi-year gap in the transformed factor. That could be such a significant problem for which we have no choice but to adjust the inputs and target variables as discussed previously and to include 2020 in the training.

7.4.3. Modeling Techniques

Correlations between macroeconomic factors have shifted during the COVID-19 Crisis and may not return to pre-crisis structures. The US Federal Reserve announced long-term changes in how they would manage interest rates, and they repeatedly assured markets that interest rates would be kept low for years to come. Aside from the obvious impacts on stock markets, this also impacts the correlations between variables in stress test models.

Stress testing models based upon linear regression may not explicitly incorporate cross-correlations between input factors, but the selection of inputs is influenced by past correlations between variables. In order to minimize multicolinearity, explanatory variables with the least redundancy are selected. The dramatic breaks in the correlation structure between variables that were observed in the COVID-19 Crisis means that changes need to be considered in the way stress test models are built.

One of the simplest ways to be robust to shifts in the correlation structure is to use ensembles of models with different input variables. Creating a collection of models where the dominant explanatory variable in one model is excluded from the next enforces diversity that enables for a more robust forecast when dramatic shifts occur, such as in 2020.

7.5. Economic capital / Basel II

The COVID-19 recession is a tail event like no other. Capital calculations are explicitly intended to take into account such tail events as pandemics. The challenge is to attach the appropriate severity. Looking at the first-year impact of the pandemic can be misleading if creating an annual loss distribution. Either one must integrate over a longer time period or use the should-have-been defaults adjustment.

As stated before, that approach still assumes that future governments will have comparable responses to future crises. If that assumption is undesirable, a stress test model could actually be used to reforecast 2020 with disposable personal income held to previous recessionary patterns. The forecasted losses could then be used as the data point when estimating a loss distribution.

8. Validation with COVID-19 Data

A fear expressed by many developers is that even if data from 2020 is not used in model development, a validation against 2020 performance could be requested by validators, examiners, or auditors. The problem is obvious – that 2020 contains structure and impact that cannot be learned from any pre-2020 data, so the models can be dramatically wrong. If all involved agree not to include 2020 data in model development, they must also agree not to include 2020 in validation.

The situation becomes more complicated when 2020 data is adjusted as described above so that it can be used in development. For scoring models where hold-out samples are standard practice, developers should consider having separate 2020 and pre-2020 hold-out samples so that performance in normal times and during the pandemic can be tested. By testing pre-2020, the developers will have an idea how well the model will perform once a normal, non-assisted recovery begins. The 2020 test demonstrates how well the model performs during this and similar crises.

Validating time series models like stress test models is the most complex because including the COVID-19 data in development leaves no similar data for out-of-time testing. In 2021, if a validator asks for the common 12-month out-of-sample test, a pre-2020 model is unlikely to capture the nonlinearity of 2020, so that out-of-sample test would fail. However, 2020 data, properly adjusted as described in Section 7.4, could be quite valuable to improving modeling of future crises. Therefore, at least through 2021, a one-year out-of-sample test (or any comparable period) should be taken from the pre-2020 data. A slice of transformed input data of the needed length should be reserved, either from the 2009 recession, if available, or one of the intervening years. A hold-out of raw macroeconomic data would have to be significantly longer than the needed period in order to allow for lag and window-of-change transformations.

9. Conclusions

As this was written, the COVID-19 pandemic was far from over. We expect a mass vaccination campaign in 2021 will bring the pandemic under control and return lending to more normal operating principles. Until then, we continue to learn and adapt and may revise the observations and expectations identified so far in this document.

Among those many observations, the first were about the input data to our models. At an individual borrower level, bureau data could track how consumer borrowing changed through the crisis, but rules about forbearance and delinquency limit the completeness of this picture. Consequently, credit scores, even where they still rank order overall, will likely have higher uncertainty and vary more widely across subpopulations. Many lenders responded to this uncertainty through an increased reliance on cash flow analysis. This is particularly useful in identifying those who can make payments with government assistance versus those making payments through continued employment.

Throughout the pandemic, reports of macroeconomic data have been headline news, and yet the data integrity was severely impacted by the pandemic and government policy changes. The US Bureau of Labor Statistics admitted that unemployment probably peaked at 19.6% when adjusted for the undercount due to flawed surveys, and yet the officially reported numbers remain well below that. Initial unemployment claims peaked much higher than they would have if not for the additional Pandemic Unemployment Assistance program, highlighting a definitional problem with the high-level statistics. Models make assumptions that the input data, though possibly noisy, remains unbiased and definitionally unchanged. Both of these assumptions failed through 2020 and raise questions about how to monitor not just the data feed to models, but how to monitor definitional and reporting problems in the future.

Operationally, lenders still lent through the pandemic, although with dramatic shifts in product and demographic focus. Borrower cash flow analysis was already gaining traction in the industry for use in underwriting, but the pandemic proved this to be essential for analyzing the current risks in the borrower's situation. Care needs to be taken to avoid disadvantaging a group of borrowers through association with increased risks in their geographic or demographic group, however, portfolio management needs to balance risk assumed across subgroups through adjustments to origination volumes.

Account management also moved in new directions with loan accommodation programs and shifts in collections policies. Initially many lenders offered assistance to any that requested, leading to a large amount of strategic forbearance – borrowers accepting a loan accommodation as a way of hoarding cash to protect against potential future pandemic-related problems. Of course, the cure rate from such programs was quite high. In the US, later rounds of assistance tended to require proof of hardship, which preserved more of the loan payment stream while also causing future cure rates to be weaker. As in many things, quicker was better in the beginning. Parsimonious was preferable with more time to prepare.

Throughout the pandemic, models predicted loan performance given the inputs provided. None of those models we observed had previously anticipated loan accommodations or government assistance that would break the correlation between unemployment and loan default. Therefore, model monitoring was not an exercise in how to reject weakly performing models, but in how to quantify the separation between what would have happened and what was happening. These missing defaults cannot be assumed to be completely prevented. Rather, some fraction of those can be expected to return in future quarters as assistance programs end, given the extent to which the economy has not returned to a pre-crisis state. That analysis is at the heart of designing model overlays.

The industry generated a fair amount of noise about rebuilding models during the pandemic to capture the current reality. Although well-meaning, 2020 was too soon to model the impacts of the pandemic, because we do not yet know the borrowers' performance outcome. Loan accommodations cannot be added to a credit score when the cure rate is unknown. This leads to the natural question of when crisis-era data can be used in modeling, and how.

The data can be considered for use in modeling when the usual loan payment dynamics have returned, i.e. after assistance has ended and the delay defaults have started to occur. That will likely be sometime in 2021. How to use the data is more complex. If we take the default history purely as given, any model dependent upon macroeconomic correlations will be damaged by the 2020 data. Instead, we would either need to adjust the performance data probabilistically for what would have happened if not for assistance and when it would have happened, or we must exclude the crisis-era data entirely from modeling.

For credit scores seeking to rank order risk, analysts will still want to consider the performance that would have happened, but timing is less of a problem. In economic capital calculations and other models of tail risk, this data might be used as-is, incorporating the way governments and lenders respond to severe natural disasters.

Future model validation must come with the same caveats as listed for model development. Models built without 2020 data cannot reasonably be validated against 2020 unless performance adjustments are made. If model risk oversight finds this approach unpalatable, then validation may have to be performed against hold-out time periods prior to the crisis.

In both model development and validation, careful consideration must be given to our post-crisis assumptions. Will loan performance dynamics return to a pre-crisis normal such that we can effectively ignore 2020, or will future loan dynamics shift sufficiently that a new normal will arise and we will need a wave of new model development in 2022? For now, we can only go with the assumption that the old models will work, but model monitoring in late 2021 may tell us that significant changes are coming.

Workgroup Members

Joseph L. Breeden, Chairman

Dr. Breeden has been designing and deploying risk management systems for loan portfolios since 1996. He founded Prescient Models in 2011, which focuses on portfolio and loan-level forecasting solutions for pricing, account management, CCAR, and CECL. He co-founded Deep Future Analytics in 2012 to serve credit unions and community banks. He is also the owner of auctionforecast.com, which predicts the values of fine wines using a proprietary database with over 2.5 million auction prices.

He is member of the board of directors of Upgrade, a San Francisco-based FinTech, an Associate Editor for the Journal of Credit Risk and the Journal of Risk Model Validation, and president of the Model Risk Managers' International Association (mrmia.org).

Dr. Breeden has created models through the 1995 Mexican Peso Crisis, the 1997 Asian Economic Crisis, the 2001 Global Recession, the 2003 Hong Kong SARS Recession, and the 2007-2009 US Mortgage Crisis and Global Financial Crisis. These crises have provided Dr. Breeden with a rare perspective on crisis management and the analytics needs of executives for strategic decision-making.

Dr. Breeden earned a Ph.D. in physics, and has published over 50 academic articles, seven patents, and four books. His most recent books, *Living with CECL: Mortgage Modeling Alternatives* and *Living with CECL: The Modeling Dictionary* were published in 2018.

LiMing Brotcke

Liming Brotcke, PhD leads the Model Validation Group at Ally. Her extensive industry financial modeling and model validation experience is enriched by a deep understanding of regulatory expectations on large bank supervision, stress testing, model risk management, etc. Prior to joining Ally, she worked at the Federal Reserve Bank of Chicago, Citi Group, and Discover Financial Services.

Bin Duan

Bin is currently a Principal at Principal in Quantitative Modeling and Model Risk at a GSE Lender, responsible for model risk governance oversight of all aspects of model risk across the enterprise. He has more than 20 years of experience in model development and model risk management, has personally built many predictive models in credit card and mortgage marketing, credit underwriting, credit loss forecast using different data mining, statistical and machine learning methodologies, and has managed teams of different sizes in these areas in a number of banking and financial services companies.

Before joining the GSE, Bin was the Head of Model Risk Execution, managing the global Qualitative Model Validation team at TD Bank. Before TD, Bin worked for Citi Bank in New York, where he was the Director of the CCAR Model Implementation & Execution, heading a group of senior managers and quantitative analysts responsible for the model implementation, production, and performance tracking of all loss forecast models in Citigroup's global consumer loan portfolios. Before that, Bin spent about 3 years with Discover Financial in the Chicago area as the Director of Model Risk Management.

Bin holds a PhD degree in Quantitative Psychology from Tulane University.

Andy Johnson

Andy has worked in and led modelling teams for over 25 years. He is currently Head of Modelling and Measurement at Leeds Building Society. He has worked in banking for over 30 years, in addition to modelling, fulfilling roles in customer services and Information Technology. His experience covers modelling across the lifecycle of lending, across a range of retail and commercial portfolios.

Andy has an MSc. in Software and Information Systems Design, graduating in 2001, including his thesis "*Using Neural Technology for Mortgage Decisioning*".

Charles Maner

Charles is the Retail Credit Risk Model Development Manager for Truist whose core responsibility includes development and monitoring of retail credit risk models assisting with managing all aspects of the consumer credit risk life cycle.

Charles is a seasoned quantitative model developer/manager having the benefit of over 25 years of banking and risk management experience at Bank of America and heritage BB&T. Prior to his current role, Charles was the Retail and Wholesale Quantitative Analytics Credit Risk Manager for BB&T. He joined BB&T in 2014 to build out and lead the Quantitative Credit Risk Analytics function within the Risk Management Organization (RMO). He previously worked at Bank of America in Charlotte for 20 years where he led the Commercial Risk Rating Model Development function. He began his banking career in Charlotte as a quantitative marketing analyst, then led the Customer Contact Analytics group for three years, later joining the Commercial Risk Rating Group.

Outside of his time in risk management, Charles is actively involved in the community. His spare time is spent with his wife and young daughter. He is active outdoors and enjoys both gardening and bee keeping. He graduated with a BS in Operations Research from the University of North Carolina in Chapel Hill, NC, followed by an MS in Operations Research from the Georgia Institute of Technology in Atlanta, GA.

Paul O'Neal

Paul is a Vice President of Portfolio Analytics for an automotive finance company. His core responsibilities involve managing that company's development and maintenance of their CECL model, as well as other modelling efforts that support financial planning, portfolio forecasting, and profitability analytics. His career in auto finance began in 2007. Since that time, he has focused on many aspects of risk management, including credit scoring, bureau data research, and operations metrics. Paul graduated from the University of Texas at Arlington with an MS in Statistics.