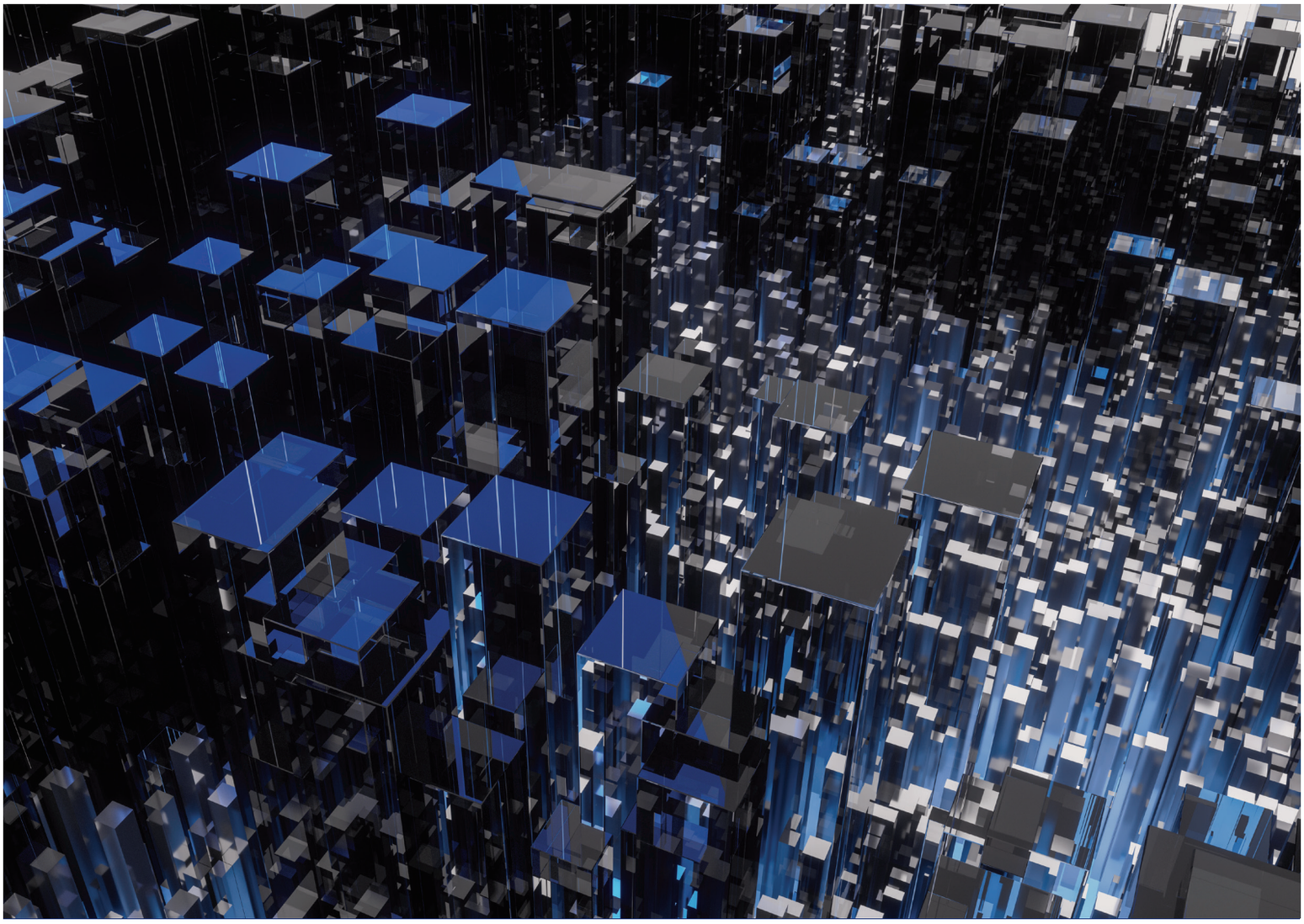
Risk publications

Stress-testing

Special report 2020

Risk.net



Risk.net

Sponsored by

Moody's Analytics

MSCI

2020 stressing you out?

Staying ahead of the crisis curve

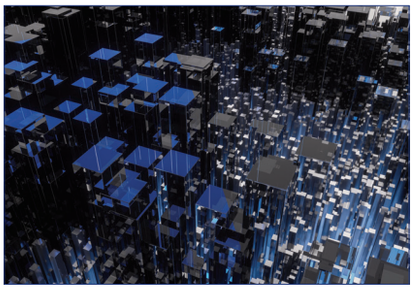
With non-financial risks, such as geo-political, climate change, technological and unforeseen health crises like the COVID-19 pandemic, coming at us with unprecedented velocity, having a robust stress testing framework is more important than ever to understand and address risks early.

At MSCI we help clients seeking to address regulatory requirements for stress testing, but also enable you to 'imagine the unthinkable' with stress testing solutions built to define scenarios early on in a crisis so you can stay on top of risks as new information becomes available.

MSCI's extensive library of historical and predictive stress tests can help with business planning by giving you a clear view of your potential structural and tail risks. Our stress tests are also fully customizable by adding shock parameters, correlation assumptions and other measures.

Learn more at msci.com/riskmanagement

MSCI



Tom Osborn, **Editor, Risk Management**
tom.osborn@infopro-digital.com

Joanna Edwards, **Global Director, Commercial Sales**
joanna.edwards@infopro-digital.com

Stuart Willes, **Commercial Editorial Manager**
stuart.willes@infopro-digital.com

Alex Hurrell, **Senior Commercial Subeditor**
alex.hurrell@infopro-digital.com

Antony Chambers, **Publisher, *Risk.net***
antony.chambers@infopro-digital.com

Philip Harding
Global Head of Commercial Content
philip.harding@infopro-digital.com

Esha Khanna, **Commercial Manager, *Risk.net***
esha.khanna@infopro-digital.com

Rob Alexander, **Sales, *Risk.net***
robert.alexander@infopro-digital.com

David Pagliaro, **Group Managing Director**
david.pagliaro@infopro-digital.com

Ben Wood, **Managing Director, *Risk.net***
ben.wood@infopro-digital.com

Ben Cornish, **Senior Production Executive**
ben.cornish@infopro-digital.com

Infopro Digital (UK)
133 Houndsditch
London EC3A 7BX
Tel: +44 (0)20 7316 9000

Infopro Digital (US)
55 Broad Street, 22nd Floor
New York, NY 10004-2501
Tel: +1 646 736 1888

Infopro Digital (Asia-Pacific)
Unit 1801-05, Shui On Centre
6–8 Harbour Road
Wanchai, Hong Kong
Tel: +852 3411 4888

Cover image
Jian Fan/Getty

Published by Infopro Digital

Why the US election fallout was not a surprise to banks

There is anecdote in Michael Lewis's bestseller, *The Big Short*, that makes the case for stress-testing as a backstop to risk modelling more effectively than any regulatory dictum could.

Lewis is describing Morgan Stanley's massive bets on subprime mortgage bonds in the run-up to the US housing market collapse of 2007, which – after lucrative early successes – were being run from a specially formed unit led by bond trader Howie Hubler.

Hubler's positions in collateralised debt obligations were stress-tested for losses on subprime pools reaching 6% – the highest in recent history – and still showed a healthy profit. Morgan Stanley's then-chief risk officer, Tom Daula, asked what would happen if losses hit 10%. Lewis documents the reaction of Hubler and his traders: "It was more than a little weird," said one of them. "There was a lot of angst about it. It was sort of viewed as: 'These folks don't know what they're talking about. If losses go to 10%, there will be, like, a million homeless people.'" (Losses in the pools Hubler's group had bet on would eventually reach 40%.)

As a senior Morgan Stanley executive outside Hubler's group put it: "They didn't want to show you the results. They kept saying: 'That state of the world can't happen.'"

Fast-forward to today, and traders seem more willing to ask 'What if?'

As Lewis's bemused executive had it, plenty of people had thought the current state of the world – a contested US presidential election in the middle of a pandemic that has killed millions and wrecked the global economy – couldn't happen, either. Traders could have been forgiven for dismissing the scenario of an election result – which the incumbent has blamed on millions of fraudulent votes – looking set to remain unclear for days, if not weeks.

But, clearly, many did hypothesise such a situation – which might go some way to explaining why market reaction in the immediate aftermath has been muted to the point of being "eerily quiet", as one trader has described it.

"This scenario was not out of the realm of possibilities. People were considering it, and would have priced in something that looked like this," says a senior risk manager at a US bank.

"They might've assigned a lower probability to it, but it's not taking them by complete surprise."

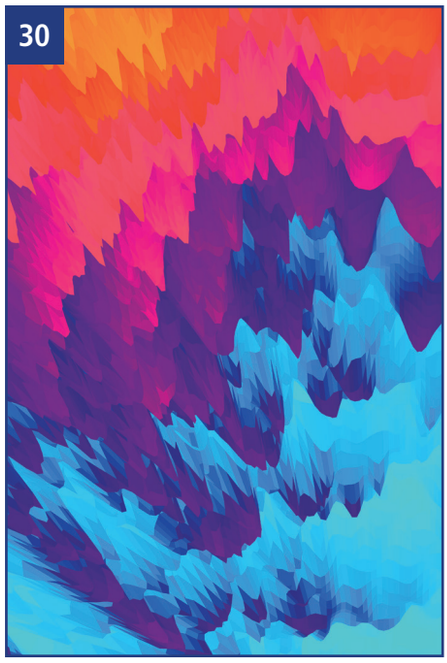
At its crudest, stress-testing forces firms to look at their risk drivers, scratch their heads, and talk to their traders and portfolio managers about whether they'd still be comfortable with their exposures if a given situation came to pass.

As the risk manager sees it: "After 2016, people learned not to take the most obvious outcome as a given, and to plan for chaos. Every firm should have planned and strategised around these outcomes. Banks will have a playbook they can follow for what is likely beyond this point."

Of course, like any risk measure, stress tests can be gamed – and, as in the unfortunate case of Morgan Stanley's credit teams, their outputs ignored. But, even if a given outcome is unexpected, banks shouldn't be unprepared.

Tom Osborn
Editor, Risk Management, *Risk.net*

Contents



Features

Sponsored feature

Sponsored Q&A

MOODY'S ANALYTICS

Stress-testing under Covid-19

Stress-testing is a challenging exercise to regularly assess a bank's level of risk or capital adequacy. Olivier Brucker, Sunayana Mehra and Ed Young of Moody's Analytics explore an approach that can address this, proposing an alternative methodology that can forecast and stress-test entire balance sheets and profit-and-loss statements for US banks

Banks need to build numerous models and try to separate bank-specific decisions from macroeconomic effects. Projecting the balance sheet and income statement under the most likely scenarios is no easy task – and banks cannot rely solely on internal performance data.

When stress-testing needs arise from *ad hoc* situations such as the Covid-19 pandemic, the challenge is even more complex as the process lies outside of 'business as usual'. Timelines are compressed and validated models might not be adapted for the specific stress.

The Moody's Analytics paper *Stress testing under Covid-19* explores an approach that addresses these issues, proposing an alternative simple, coherent methodology that allows Moody's Analytics to forecast and stress-test the entire balance sheet and profit-and-loss statement consistently for all of the approximately 6,000 banks in the US. This methodology can be used as a primary approach for banks without the means to produce such stress-testing exercises, or as a challenger or benchmark to validate the results of a set of primary models.

The Moody's Analytics approach is also useful for strategic planning as it allows banks to compare their balance sheet and income statements with those of their peers and the industry, and to explore potential mergers and acquisitions.

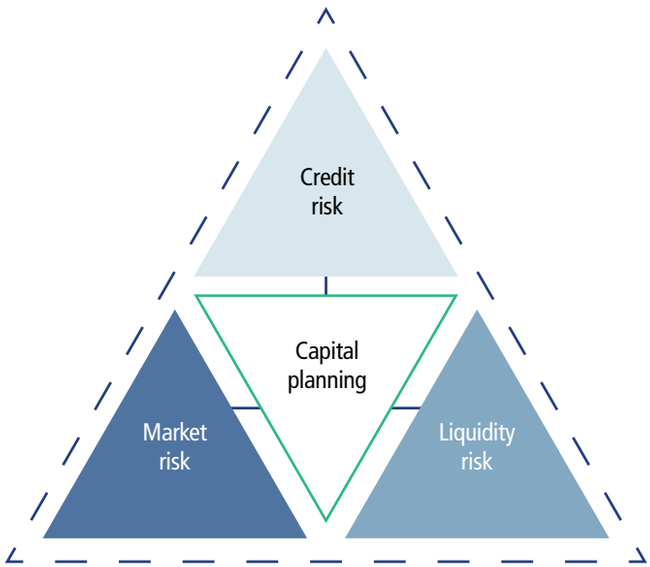
Moody's Analytics has always been on the front line of quantitative analysis to support loss forecasting and forecasting the other line items of the balance sheet and income statement. A few years ago, at the peak of the Comprehensive Capital Analysis and Review exercise, Moody's Analytics helped customers develop models and execute simulations to produce regulatory reports, such as the US Federal Reserve's FR-Y14. We have harnessed this experience in Capital Risk Analyzer, a capital planning and stress-testing solution that offers a cloud-based platform to execute champion and challenger forecasting models, and off-the-shelf models to automate stress-testing.

Stress testing under Covid-19 discusses one of the approaches in this solution based on a top-down model, the Call Report Forecast. This model uses publicly available historical data through the Federal Deposit Insurance Corporation report to derive forecasts for industry-level aggregates based on asset sizes and geographies. The paper describes how this approach can be useful for producing stress-testing results with unexpected scenarios, such as the Covid-19 pandemic: first, by examining the methodology used to produce industry-level forecasts, then looking at Covid-19-based scenarios before finally combining them to quickly produce bank-level forecasts under these scenarios. ■

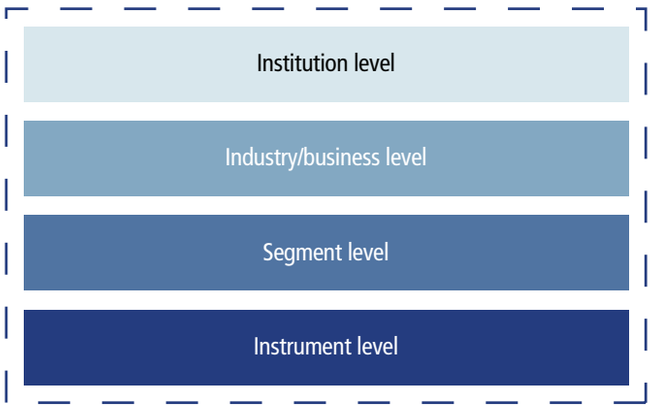
To learn more

Read the full paper, *Stress testing under Covid-19* at *https://bit.ly/34j5m0r*

1 The new stage of planning – capital planning
Processes must be inter-risk and multidimensional



The Covid-19 crisis has accentuated the need for solutions that can incorporate the impact of a wide range of potential risks. Harmonising risk measurement across risk silos is required to answer questions the current crisis is raising



Second round of stress tests to push banks' limits

Stress scenarios recently released by the US Federal Reserve Board to gauge resilience against simulated economic shocks project unemployment to peak at 12.5% in the worst-case scenario. By Louie Woodall

A fresh set of stress scenarios released by the US Federal Reserve Board will gauge banks' resilience against a simulated economic shock tougher in some respects than the one it conjured up seven months earlier.[1]

The newly minted severely adverse and alternative severely adverse scenarios, rolled out so top lenders can reassess their capital adequacy in light of the ongoing economic disruption inflicted by the Covid-19 pandemic, project peak unemployment rates of 12.5% and 11%, respectively. This is higher than the 10% rate used in the severely adverse scenario published in February for the regular Dodd-Frank Act Stress Tests. However, the increase from the starting level is smaller than in previous years because US unemployment is already at a historically elevated level, and is expected to be around 9.5% as of the third quarter of this year.

Real GDP is projected to fall to -5.9% year-on-year under the severely adverse scenario and to -9.1% under the alternative scenario in the first quarter of the simulation. The February scenario projected a -9.9% trough in the first quarter of the scenario, Q2 2020. Real GDP actually contracted -31.7% that quarter as the pandemic took hold.

The Dow Jones Industrial Average is projected to fall 43% from its Q2 2020 level under the new severely adverse scenario, and 42% under the alternative scenario. The February scenario projected a 50% point drop from a Q4 2019 starting point. The Dow Jones fell 21% in Q1 this year.

The Fed said the severity of the alternative scenario is on a par with that of the W- and U-shaped scenarios featured in its Covid-19 sensitivity analysis in June 2020.

Fed scenario amendment

In the wake of the Covid-19 crisis, the Fed amended its annual stress-testing programme to incorporate a sensitivity analysis, which was used to appraise banks' capital adequacy under the unprecedented economic conditions created by the pandemic.

The central bank issued this latest set of scenarios so banks could reassess their capital needs before resubmitting their plans to make capital distributions to the agency in Q4.

Why it matters

Two things stand out from the Fed's latest scenarios. First are the unemployment assumptions. Typically, the central bank projects unemployment to gradually rise to a peak and then drift lower over an extended period. The reality of the Covid-19 recession shows that the unemployment rate can move very rapidly quarter to quarter. This lesson appears to have been adopted, at least in part, by the Fed. The alternative downside scenario takes five quarters to hit its peak unemployment rate. In comparison, the February severely adverse scenario took seven.

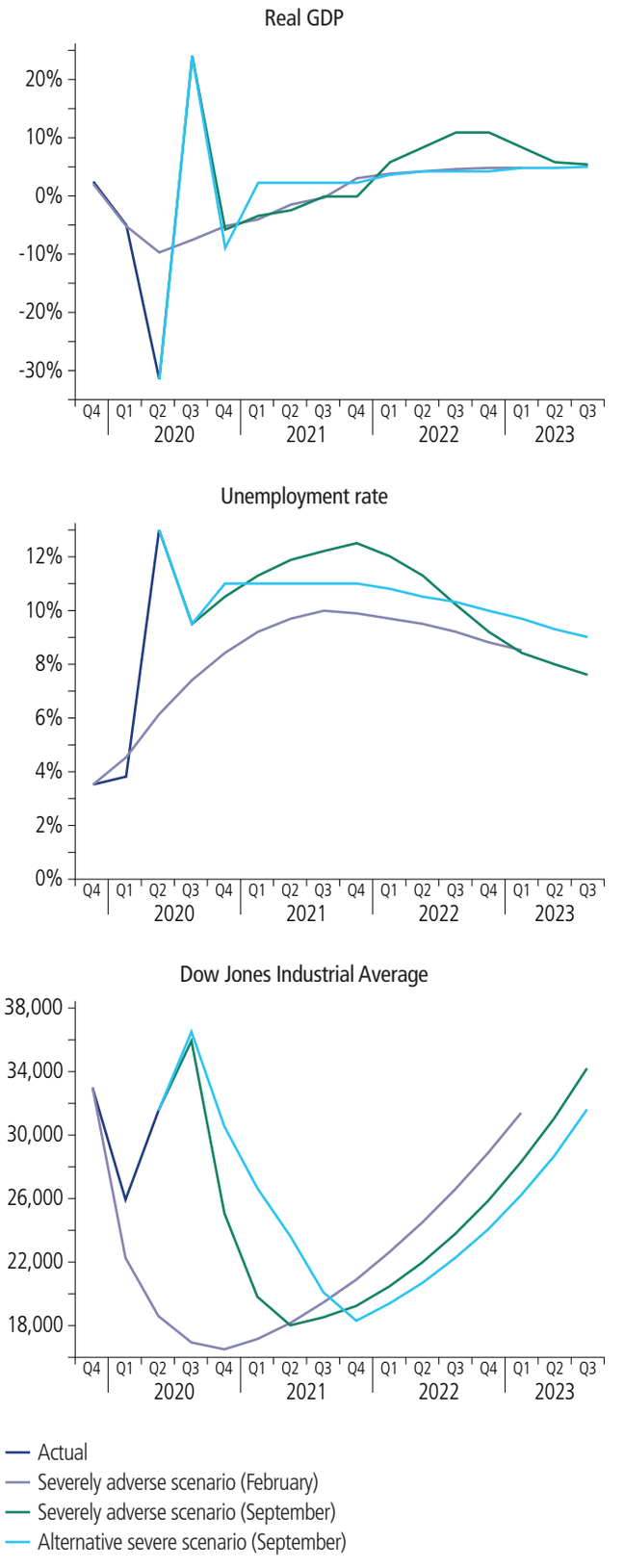
Still, between March and June this year the unemployment rate jumped 9.2%. The absence of a similar upward jolt suggests the Fed doesn't think unemployment could climb by such a large amount at its already elevated level.

Then there are the stock market assumptions. The Fed is holding fast to a simulation where the Dow drops roughly by half over the course of its simulations. But even at the height of the coronavirus crisis, the index dropped little more than 20% before shooting right up again. This suggests the Fed isn't letting the past performance of equities dictate its stress-testing programme. In fact, it hints that these tests may be even harder on banks than the recent crisis. ■

Previously published on Risk.net

[1] *US Federal Reserve Board (September 2020),* Supervisory scenarios for the resubmission of capital plans in the fourth quarter of 2020, *https://bit.ly/30C8LX4*

1 Key domestic variables used in the Fed's supervisory stress scenarios

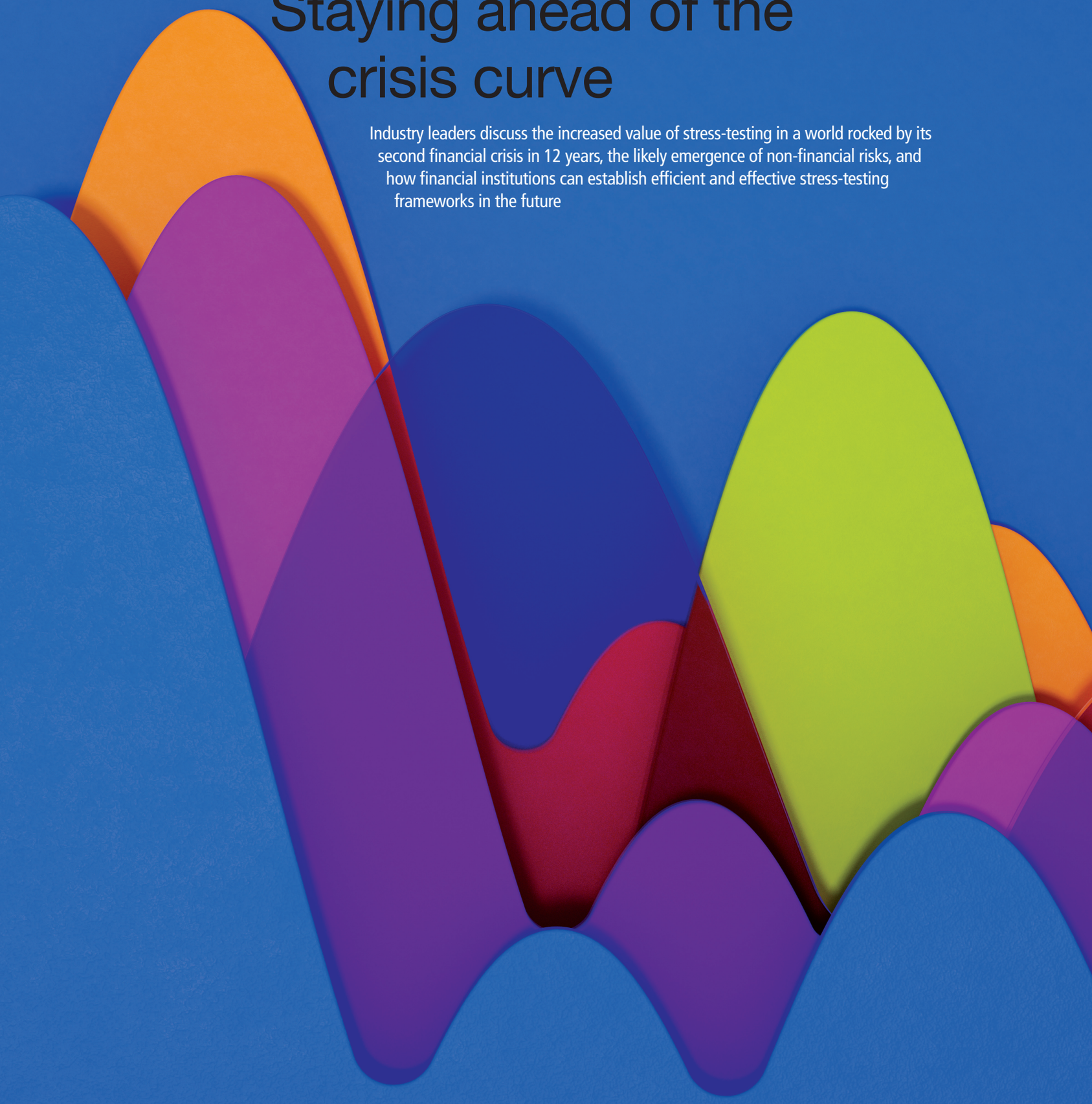


Source: US Federal Reserve Board

Thinking the unthinkable

Staying ahead of the crisis curve

Industry leaders discuss the increased value of stress-testing in a world rocked by its second financial crisis in 12 years, the likely emergence of non-financial risks, and how financial institutions can establish efficient and effective stress-testing frameworks in the future



MSCI

Thomas Verbraken
Executive Director, MSCI Research
www.msci.com

How has the crisis brought about by the Covid-19 pandemic highlighted the value of stress-testing for financial services firms?

Thomas Verbraken, MSCI: It is precisely during periods such as the recent Covid-19 crisis – when historical analysis has its limits – that stress-testing proves to be most useful because of its forward-looking nature and the ability to inject expert knowledge into the analysis. That makes stress-testing a great tool for modelling unprecedented scenarios.

Having a stress-testing framework in place may leave investors more prepared for market turmoil. Of course, it is unlikely that an exact pandemic-like stress test would have featured on the list of scenarios run by investors prior to Covid-19. But running other crisis scenarios – even if not capturing the exact nature of the pandemic – may help get a sense of tail risk and the vulnerabilities and exposures of portfolios.

Furthermore, meaningful stress-test scenarios are often characterised by economically sound narratives, expressed as a set of shocks on major market and macroeconomic variables. A solid understanding of how macroeconomic variables are linked to financial markets is crucial, especially in cases such as the Covid-19 crisis, from which an unprecedented economic shock originated. Understanding those relationships can also be used in a reverse stress-testing exercise and help answer such questions as: which levels of economic growth, real rates and equity risk premia could explain an observed market downturn? In particular, during periods of high uncertainty, such analysis can provide an anchor for a better understanding of what is happening.

Finally, it is not only useful to focus on drawdown-type scenarios. If Covid-19 taught us one lesson, it is that it may be equally important to think about possible recovery scenarios during market turmoil so the impact of potential tactical shifts can be understood. For example, the swift 'V-shaped' recovery seen in the past months might have been an unpleasant surprise if de-risking near the bottom meant partly missing out on the rebound, ending up with underperforming the market and certain investors able to ride out the crisis.

Laurent Birade, Moody's Analytics: The stress-testing framework established after the 2007–08 financial crisis prepared firms to understand the impact of new stress events and to have a well-rehearsed process for determining potential impacts on capital and firm viability. Because of those improvements to stress-testing, many banks were already maintaining substantially higher capital and thus have been better positioned for the economic shocks caused by Covid-19.

The ongoing Covid-19 crisis reminds us that having a decision-ready, stress-testing framework – and being more prepared for the next crisis – is of value well beyond the cost of implementing that framework.

Vivian Chan, Nomura Asset Management: Stress-testing is very useful for financial firms, but everyone has their own approach. Regulatory tests are necessary but highly prescriptive by nature, and are therefore less useful for commercial purposes. Internal stress tests allow you to adapt the scenarios to better reflect your business – Nomura, for example, looks at whether the bond or equities markets will move against us.

The Covid-19 pandemic is a fascinating scenario. When we were reforecasting post-Covid-19, and using the pandemic as the baseline, we expected to see a fall in assets under management, as well as bonds and equites markets, but we didn't. Instead, they dipped briefly and then rose. It's important we have the stress tests, particularly for capital adequacy, but Covid-19 proved the scenarios and economy are very difficult to predict, with little consistency in its impacts across industries and regions. Covid-19 isn't the only reason the value of stress-testing has increased. Most companies stress-test a global economic downturn, and there are also reverse stress tests as part of firms' risk and control self-assessment processes, where it is considered what could 'break' the organisation.

Model risk and stress-testing specialist at an international banking group: There's value in thinking about what might happen in stress tests, and Covid-19 doesn't change this. However, the elephant in the room is whether current stress-test figures are remotely credible.

Models have proven far more accurate than even the best human intuition, but the results are now being challenged by senior managers and others.

When stress tests are dominated by models, you need confidence in your model results. With both a Covid-19 pandemic-induced recession materialising extremely quickly – much faster than previous recessions – and with results way outside of the typical calibration range, you have to question whether the models are delivering credible results.

The biggest challenge is knowing the limitations of the models and the level of human overrides required, but it's uncharted territory. It is not like previous recessions where industries behaved in a reasonably similar way. There are so many things outside the usual range that we do not have the economic or historic data to model with confidence what might happen.

Even if it was a typical recession, it's atypical in terms of the impact. Some industries are doing very well while others are experiencing extreme difficulties. Government intervention is also far more pronounced in this case, so some of the data is distorted as you cannot see some of the defaults. For example, how do you assess the future stresses in the personal sector when there are few defaults with people being furloughed or benefiting from temporary payment holidays?

Models can only predict what they are calibrated for. When you go outside that range, you need management adjustments to reflect the perceived limitations and weaknesses in the models. The question is how to perform the overrides in a structured, rational way. For example, an International Financial Reporting Standard 9 submission might incorporate relationship managers' views on individual companies' performance alongside the model results. This helps reflect the idiosyncratic effect of this recession where some industries are booming and others have been shut down.

A recent regulatory review of banks' internal stress-testing programmes found deficiencies such as siloed operations, overly lenient scenarios and a paucity of *ad hoc* stress-testing capability. Are these criticisms fair? What are the key tenets of a robust stress-testing framework?

Vivian Chan: I think this is mostly fair. Some larger banks have their own scenario expansion teams and undertake a diligent governance process to understand the outlook and macroeconomic behaviour for each market. For the Internal Capital Adequacy Assessment Process, for example, the UK Financial Conduct Authority has guidelines on how banks should stress the budget or business plan. It's not just that it's siloed, but that each bank will have a very prescriptive view. There are no publicly published scenarios each bank should be using, so it is difficult for anybody trying to compare or assess whether banks could withstand a certain economic impact. One firm could decide a particular market will drop 5% while another thinks it will be nearer to 10%.

Laurent Birade: Although financial institutions are relatively better prepared than they were before the previous financial crisis, their work is not complete. The evolving Covid-19 crisis has revealed some weaknesses in the regulatory framework that are typically run once a year and supported by many governance activities. The deficiencies we're seeing today have only come to light due to the rapid evolution of this year's unforeseen events. *Ad hoc* but timely stress-testing must supplement the regulatory framework to give banks the ability to run any scenario, at any time, for any set of assumptions. A robust stress-testing framework must incorporate:

- A wide range of risk evaluation (accounts for emerging risks)
- Consistent loss estimation
- Scenario analyses (quantitative and qualitative, plus outer-bounds estimates)
- Recent data for more informed decision-making (data from the past few days, not months).

Moody's Analytics provides rapid-fire scenario analysis capabilities to clients through its award-winning Capital Risk Analyzer platform, which can be used as a complement to a robust Comprehensive Capital Analysis and Review/Dodd-Frank Act Stress Tests framework.

MOODY'S ANALYTICS

Laurent Birade, Senior Director
Risk and Accounting Solutions
www.moodysanalytics.com

Model risk and stress-testing specialist at an international banking group

The recent crisis has brought non-financial risks sharply into focus. To what extent will climate, geopolitical, technological and pandemic risks climb the stress-testing agenda?

Laurent Birade: Before the previous financial crisis, home price declines were not considered a serious risk. Similarly, Covid-19 now moves pandemic risks into the spotlight. Climate, geopolitical and technological risks follow closely behind, now that these elements are surfacing more frequently – particularly climate risks.

The speed and depth of the disruption caused by Covid-19 is unprecedented. How can risk managers become more adept at 'thinking the unthinkable' and creating effective future scenarios? Scenario design is a crucial part of risk managers' jobs – thinking outside the box and looking at risk in the following paradigms:

- Known
- Known unknowns
- Unknown unknowns.

Thomas Verbraken: Even before the Covid-19 crisis we had noticed non-financial scenarios becoming increasingly important to investors, for example the Brexit referendum, the US-China trade war and the US elections. Their concern is not only about the turbulence that may occur around these events themselves, but also the potential longer-lasting effects they may have through their macroeconomic impact. At MSCI we are spending a lot of time and effort on understanding the links between macroeconomic shocks and their consequences for financial markets. It is a complicated problem because of the different timescale in which macroeconomic changes and financial market shocks play out, and because financial markets may react to perceived changes in long-term macroeconomic expectations that ultimately may not materialise. Having a credible model to connect macroeconomic shocks to financial markets has proven very helpful for creating meaningful scenarios about geopolitical events.

The second major trend we see is around climate change. Scenario analysis lends itself well to this topic for various reasons. First, there is no real precedent for climate change, so the analysis must be forward-looking in nature and stress-testing allows for that. Second, there is a large degree of uncertainty around future outcomes, not only in the evolution of the climate itself but also because of the offsetting nature of two types of climate risk for financial portfolios: transition and physical risk. The former is caused by the efforts undertaken to slow down climate change and can potentially impact carbon-intense companies negatively while creating opportunities for renewable energy and other climate-friendly technological innovations. Physical risk, on the other hand, is the cost of inaction, such as increased wildfires, floods, rising sea levels, and so on. Each climate scenario represents a trade-off between these two risks, whereby the balance between transition and physical risk varies. Being able to run a range of plausible climate scenarios (for example, from 1.5° to 3° Celsius global warming scenarios) and assessing their possible impact on financial portfolios is an increasingly important exercise for many investors.

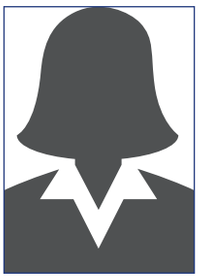
Model risk and stress-testing specialist: Stress-testing is dominated by the needs of regulators to ensure banks are well capitalised to survive extreme stress. However, the level of detail and associated cost of stress-testing is increasing for firms. As a consequence, there is less appetite to work beyond the regulatory requirements.

Stress-testing processes need to become much more efficient before a wider range of scenarios can be considered. Climate change and global warming will certainly be among these scenarios, but there are many more that could be valuable to banks – for example, wars or conflicts that disrupt the global supply chain.

This risk classification will enable risk managers to think more actively about what may trigger the next crisis.

Vivian Chan: Some of these elements are already considered, such as the impact of elections or trade wars within certain scenarios, or how a pandemic would trigger financial risk. But I think climate and technological risk will need to feature more. More firms are committing to reporting on climate risk. Environmental, social and governance is also a factor that could be added into stress-testing scenarios as it could have a major impact in terms of credit risk. For example, in current circumstances, we've seen some retail firms exposed as having workforces operating outside Covid-19 guidelines to maintain production.

Technology has changed the way the world operates but again it's difficult to predict the behaviour. I've seen reports that more people are dabbling in share dealing as a way to supplement reduced income under Covid-19, and chasing futuristic tech stocks such as Tesla, so there's potential for a bubble to emerge if prices are artificially inflated.



Vivian Chan
Executive Director, Head of Finance
Nomura Asset Management
www.nomura.com

The speed and depth of disruption caused by Covid-19 was unprecedented. How can risk managers become more adept at 'thinking the unthinkable' and creating effective future scenarios?

Vivian Chan: It's very hard. Most banks are using severe adverse scenarios but, with Covid-19, the impact on such areas as unemployment is extremely hard to predict from country to country. What could be worse than a third world war? A data war perhaps? Banks are using reverse stress tests of the most severe scenarios such as a US, China or European Union crash. But every bank could do something completely different.

Thomas Verbraken: We have learned through conversations with our clients that risk management becomes more useful when you can actively involve more stakeholders in the exercise. Portfolio managers, senior executives and board members can all provide meaningful input into stress-test scenarios. Drawing on more diversified expertise may not only improve the scenarios themselves but could also make the stress-testing exercise more impactful, because stakeholders other than risk managers will have bought into the analysis early on.

We have observed it is useful to have a framework that allows investors to express stress-test narratives quantitatively at a high level, with a relatively limited set of shocks on major market and macroeconomic variables (for example, MSCI USA, the 10-year Treasury yield and US GDP growth). Once the narrative is specified, the risk system should be able to take care of the rest and propagate such views to granular portfolios in a robust way. However, most effort – in particular from the other stakeholders – should go into the definition of economically sound scenario narratives at a high level.

Investors with the risk management processes and systems in place to do such stress-testing exercises efficiently might benefit on two fronts. Not only can they draw on much more expertise to 'imagine the unthinkable' during calmer times but also – and perhaps more importantly, once a crisis starts to unfold – they will be able to react quickly by defining stress-test scenarios early on in the crisis and stay on top of things by continuously refining the analysis as more information flows in.

Model risk and stress-testing specialist: I remember a quote about it being impossible to convince someone of a fact if their livelihood depends on not believing it. Similarly, in stress-testing, if a scenario suggests a profitable business strategy could cause large losses in the future, it will be difficult to obtain engagement from senior management if their rewards are tied to the strategy.

Thinking about the different responses to the Covid-19 pandemic, you realise people don't always act logically. Nor are all the facts readily available to make well-informed decisions when a scenario such as Covid-19 is unfolding. History is not necessarily a good guide. Even as the pandemic was unfolding, I would not have believed the response would be a lockdown of the economy. The response to Covid-19 appears far more draconian than the response to the Spanish flu, which was much more lethal, with most deaths occuring in the healthiest adults in the prime of their lives. Remember, when using historic scenarios, peoples' values and behaviour were different in the past. Hence, you need to consider how attitudes have changed to assess the impact of future potential scenarios.

Stress-testing places huge demands on data, systems and resources, but too often the resulting insights are left on the table. How can firms derive more meaningful action from their endeavours?

Laurent Birade: Having appropriate stress-testing based on needs is the right starting point. The regulatory stress test emphasises compliance with a strictly governed process that must meet qualifications of regulatory reviews. However, a more nimble top-down stress-testing process should accompany that infrastructure to ensure results can be produced in a timely manner, and that those results are congruent with the regulatory stress test. Incorporating the process as part of a regular strategic plan is only possible if your process is nimble enough to accommodate expert judgement analysis that enables managers to conduct rapid expert scenario analysis, to complement any quantitative approach.

Model risk and stress-testing specialist: The real question is, what is the purpose of stress-testing? If you're saying it's just a numeric exercise, and a tick-box exercise based on the maths, then I would question the value of the stress-testing.

I always remind myself that 'failing to plan is planning to fail'. Similarly, the objective of stress-testing should be to stimulate debate so that management can make quicker and better decisions when a scenario is crystallising.

Vivian Chan: Banks in particular spend time spent trying to pull a coherent dataset together. You need people who understand the finance and requirements of the stress test, as well as the technology and the data language. For example, I had to hire a specialist team that could enter the data into a system that uses Python, and then process it so we could run the necessary analytics. You often find gaps in the data, and measurement criteria change frequently. You put so much effort into data cleaning, creating the perfect system and looking at the requirements of a particular scenario that it is easy to lose focus on the commercial value.

Considering the operational challenges caused by the Covid-19 pandemic, how can firms drive efficiencies in their stress-testing procedures in the future?

Laurent Birade: By streamlining the stress-testing process and ensuring you can run the right scenario on the right process framework. This means adopting a bottom-up regulatory stress test with all the governance bells and whistles, complemented by a top-down approach that offers directionally correct answers in a timely manner for rapid decision-making. The regulatory stress test provides a robust governance process that, while required, may not accommodate daily scenario requests in times of economic upheaval.

Vivian Chan: One of the most time-consuming aspects is working out what data you need and applying the correct filters to enhance your core datasets. The challenge lies in identifying the dataset you need to apply the correct drivers to perform the stress test. In the scenarios, depending on the product, I might be looking at a bear market in bonds or equities, for example, and whether we have enough data to look at underlying assets in bond markets in particular countries.

It's a process of evolution, deciding what data you might need on a consistent basis and ensuring it is readily available and fit for purpose. The systems are evolving, but currently too much is Excel spreadsheet-driven and manually built. I'd like to be able to pull the data into something like Tableau to give a better visualisation of the impacts, and make it more commercialised to help inform business strategy. Technology has meant people have been able to transition seamlessly to working from home so, from that perspective, the challenge hasn't changed. ■

>> The panellists' responses to our questionnaire are in a personal capacity, and the views expressed herein do not necessarily reflect or represent the views of their employing institutions

Inside the Fed's secret liquidity stress tests

Bank lobby groups and Randal Quarles, vice-chair for supervision at the US Federal Reserve Board, train their sights on horizontal exams that can shape banks' risk appetites. By Sharon Thiruchelvam

On March 24, at the height of the Covid-19 market crisis, the US Federal Reserve Board quietly shelved confidential liquidity reviews and stress tests for the largest banks as part of a raft of measures intended to ease operational burdens.

Horizontal liquidity stress tests were among the planned exams deemed "non-critical", *Risk.net* learned, and ultimately deferred until June 15.

The suspension unnerved some former regulators. Tim Clark, former deputy director of regulation and supervision at the Fed until 2017 and key architect of US liquidity stress-testing, says a hiatus in exam work will impair the Fed's oversight of systemic risk. "It's as important now as it's ever going to be," he says.

The liquidity stress tests are among a range of measures designed to keep banks operating during crisis events. But after the Fed released a $3 trillion tidal wave of cash into the financial system over three months to June, its unfinished liquidity regime has come under renewed scrutiny.

Need to know

- The US Federal Reserve Board's horizontal liquidity exams – known as the Comprehensive Liquidity Analysis (CLAR) and Review – were implicated in last year's repo crisis, which forced the central bank to inject billions in overnight cash to keep markets flowing.
- Very little is known about the stress-testing and review programme, which is subject to the Fed's most stringent confidentiality rules.
- At the peak of the Covid-19 crisis, as the Fed deployed emergency liquidity facilities, the CLAR exams were postponed in a move that unnerved some former regulators who cited risk to the financial system.
- *Risk.net* spoke to senior officials – both former liquidity supervisors and bank treasury officials – to gain a 360-degree view of how the programme works and its future.

"Liquidity is a critical issue," says Clark, who is now an adviser to the public interest group Better Markets. "The good news is we were better prepared for [the Covid crisis]. The less good news is that the Fed still had to take huge actions to keep the markets functioning from a liquidity standpoint."

The injection follows another unprecedented move by the Fed last year to calm jitters in the repo market by unleashing billions in overnight money in response to a sudden spike in borrowing rates.

The episode led to questions from lawmakers in Congress. In a hearing of the House Financial Services Committee on December 4, 2019, the Fed's vice-chair for supervision Randal Quarles highlighted supervisory practices as one of the factors that may have contributed to the overnight repo market volatility.

"They were probably not the decisive contributors, but they were contributors, and I think we need to examine them," Quarles told Congress.

This came after comments from JP Morgan chief executive Jamie Dimon, who said liquidity stress-testing requirements had reduced the amount of cash the bank was willing to commit to the repo market at that time.

Yet the horizontal programme is an object of mystery to all but a few bank insiders and regulators. The tests are subject to souped-up confidentiality laws that limit access to a tight need-to-know list. The laws are so strict that on July 24, the Fed made a legal tweak just to allow firms to disclose information to their staff, auditors, outside legal counsel, or other federal or state banking regulators.[1]

Risk.net has interviewed four former liquidity supervisors and the treasurer of a large US bank with direct experience of the test process. The conversations give previously undisclosed details on the nature of the horizontal stress test and how banks conduct it.

Their insight shows how banks develop the scenarios used to stress liquidity provision, the way that regulators assess these efforts and deliver their verdicts, and the impact of the test on banks' operations.

The Fed did not comment on the horizontal stress tests and reviews.

Randal Quarles, Federal Reserve

All CLAR?

By design, US liquidity supervision can exert a stronger hold on banks' liquidity practices than the Basel liquidity coverage ratio, which acts as a hard backstop. The horizontal exams enable supervisors to stress and assess banks' risk management systems and processes to identify gaps or weaknesses.

The tests are widely known as the Comprehensive Liquidity Analysis and Review (CLAR) – although the US regulator no longer formally uses this term.

"The CLAR body of work is looking to ensure that banks have the right types of flexibilities and contingencies, such as contingency funding planning," says Katheryn Van der Celen, director at the Promontory Financial Group consultancy and former supervisor at the New York Fed.

It is also designed to identify gaming of the Basel prudential ratios, says a former senior Fed liquidity official. The liquidity coverage ratio is based on a set of fixed assumptions about the cashflows generated by certain types of assets and liabilities.

"You need to be able to monitor the activities banks are engaging in that reduce what the liquidity coverage ratio measures as liquidity risk," says the former Fed official.

Introduced in 2012, the exams are termed 'horizontal' because they are conducted across the eight systemically important US banks. The Fed uses the results to establish a broad picture of liquidity risk management at the country's largest lenders. The exams are intended to probe banks' own internal stress-testing efforts, and ensure they meet the required standard.

As the treasurer at one of the systemic US banks puts it: "CLAR is not a thing, it is a process. You should think of CLAR as a collective noun for the Fed's annual programme of review of everything that we do, which manages liquidity and quantifies and measures liquidity risk."

The horizontal exams are staffed by a group of liquidity risk experts, and testing occurs through a year-long cycle that is split into two half-year testing periods.

In a year's cycle, firms can expect to receive four to six exams focusing on specific topics. These quantitative and qualitative tests could range from modelling how banks earn revenues from their high-quality liquid assets in a stress scenario, to prime brokerage management, second line of risk defence, or internal auditing.

Stressed in three ways

The quantitative tests assess how much liquidity risk a firm is exposed to during periods of stress. The tests use three stress scenarios: a market-based scenario, which features a market shock; an idiosyncratic scenario, which stresses individual banks' most significant exposures; and a combined scenario, which features aspects of both.

The Fed examines how banks use internal stress tests to gauge liquidity in a business or activity, and how firms measure their risk profile across each of those three scenarios.

The regulator does not hand banks a set of off-the-shelf scenarios. Rather, banks design their own scenarios within broad guidelines issued by the Fed in Regulation YY enhanced prudential standards, as well as a 2010 policy statement on funding and liquidity risk management.[2]

"The pen, so to speak, is in the hands of the bank to develop the narrative and the details of the scenario as well as the underlying methodology, which ultimately drive the results," says House. "When evaluating individual methodologies the Fed will consider: is this a credible stress test to parse out the key risk and stress them in a meaningfully conservative way?"

> "[The Federal Reserve] might reduce scope to reprioritise one or two issues or topics that are the most relevant in the current environment"
>
> Katheryn Van der Celen, Promontory Financial Group

Typically, in each half-year, banks can expect to receive one quantitative and one qualitative test. But supervisors have the flexibility to tailor the schedule to respond to significant economic or market events and to defer less urgent planned topics – as happened in March.

"Absent those very topical and important things that come up from time to time, they will have their slate of topics that they'll get to more on a cycle-based approach. But if you look at the exam calendar for the next 12 months, you probably have one or two on there that you say these are mission critical, we need to get done," says Mark House, partner at consultancy Chain Bridge and a former Fed programme manager for CLAR.

Within the banks, the horizontal reviews are primarily handled by treasury and risk management staff but can also involve internal audit and, depending on the topic of the review, the leads of specific business functions.

The treasurer agrees: "They are far more focused on our capabilities and inspecting the process that we have, rather than giving us a scenario and telling us to work on it."

The market-based scenario generally contains some of the same macroeconomic factors used in the capital stress tests, says Van der Celen. This enables the capital and liquidity frameworks to be consistent.

The qualitative reviews focus primarily on risk management policies and procedures. Topics might include contingency funding plans, where the Fed would assess the alternative funding lines that banks could fall back on if they lose access to a source of funds.

Alternatively, the Fed might review banks' reporting procedures for monitoring liquidity risks. This could involve analysing how firms source data for reporting liquidity risk metrics and their controls to ensure the data is accurate. The examiners may also review how the reporting is disseminated to management, and whether decision-makers are receiving the information required to make informed decisions.

The overall process is described by the bank treasurer as "very intense", with regular calls and meetings between bank staff and supervisors over a compressed period (see box: *Taking the test*).

Assessing results

Results are typically issued via confidential feedback letters to banks once or twice a year, giving a readout or a grade on their performance in the exam and detail on their progress in meeting the requirements of the regulation.

If credit risk modelling, with decades of data on default rates and credit losses, is a science, liquidity stress modelling is an art. The process is more qualitative than quantitative, says House, and assessments require an element of judgement.

Firms receive guidance on "the range of the quantity of liquidity they have to hold, whether it's their liquidity coverage ratio stress test or their internal liquidity stress test", says Van der Celen.

Negative feedback is delivered through matters requiring attention (MRAs) or matters requiring immediate attention (MRIAs). For example, a firm may have underweighted a given risk driver in its internal liquidity stress test, casting doubt on the adequacy of its protections against that risk. In this case, the Fed may issue a notice outlining the deficiency with directives to fix it.

How long remediation takes depends on the seriousness of the issues found. If the issue is not significant, the firm may be given more time, House says.

The Fed's advice is not always a diktat, either. Sometimes the feedback is the starting point for a "constructive dialogue" between bank and regulator, says the treasurer. A bank may push back on a particular issue, pointing out certain mitigants and causing the Fed to reconsider.

The feedback can also have far-reaching implications for a bank's business. An exam could identify that certain activities aren't being appropriately captured in stress-testing – in other words, the Fed concludes that the liquidity risk was underestimated by the firm.

The resulting supervisory finding could cause a firm to hold more liquidity against a business, which could ultimately affect the economics of that area activity, depending on how substantial the Fed considered the underestimate to be.

House thinks that standardised measures such as the liquidity coverage ratio establish a baseline cost of conducting many of the affected businesses and activities – for example, the amount of high-quality liquid assets a specific business line is absorbing.

But, he adds, the internal liquidity stress tests can more accurately measure the liquidity risk in a firm's activities. This could lead to the firm determining that a business may be riskier than the liquidity coverage ratio suggests – perhaps too high to keep conducting in its current format.

What's in store?

The Fed announced the resumption of planned examinations on June 15. The former supervisor says they expect Fed staff to be formulating scenarios that better reflect the altered circumstances.

Van der Celen adds that it may translate to a lighter exam-load until the worst of the Covid-19 crisis is over: "They might reduce scope to reprioritise one or two issues or topics that are the most relevant in the current environment and support any shift in the supervisory focus on systemic concerns."

Looking ahead, Quarles has confirmed that the Fed is continuing with a programme of reform aimed at increasing transparency and accountability in bank supervision, some of which will affect horizontal liquidity stress tests.

The first of these, the rule revising confidentiality procedures finalised on July 24, means firms will have scope to more widely share information relating to the tests – though investors will still be off-limits.

Bank lobby groups have long railed against the secretive nature of the liquidity supervision programme, particularly the Fed's methodology and assessment criteria. In a comment letter on the proposed changes to the confidentiality rules in August 2019, the Bank Policy Institute (BPI) proposed the Fed make available general observations arising from exams, and horizontal reviews in particular.

"Releasing this information in an anonymised manner would offer institutions meaningful opportunities to strengthen their own compliance programmes, reduce the potential for compliance violations and enhance risk management practices based on information that would otherwise be inaccessible," the BPI suggested.

But the suggestion was not acted upon in the final amendment to the confidentiality rules.

Also relevant to horizontal reviews are proposed changes to the use of MRAs and MRIAs. The quantity of these requests and the speed at which they must be turned around has been a source of frustration for many banks, says the bank treasurer.

Quarles indicated in January 2020 that he would like to see these withdrawn as tools for enforcing supervisory guidelines.[3] Instead, he wants them to be used only in cases of specific violations of law or regulation by a bank. His proposal echoes lobbying by the American Bankers Association in late 2018, and by BPI chief Greg Baer in testimony to the Senate Banking Committee in April 2019.[4]

Window of opportunity

In the longer term, how liquidity risk is regulated is likely to be a subject of intense debate. Prominent voices have argued that if liquidity regulation makes banks hold excessive amounts of assets against liquidity shocks, then the downside is their lack of participation in financial markets at times of stress.

To this end, Quarles proposed in February that banks should assume they can access the discount window in the liquidity stress-testing process.[5] This would allow them to hold smaller quantities of cash reserves at the Fed, and instead build up holdings of other high-quality liquid assets. Quarles explained: "We could... [allow] firms to rely on the discount window in their [internal stress-testing] as a means of monetising, for example, Treasury securities in their scenarios."

Central to this debate is the extent to which bank liquidity should be resilient to extreme shocks such as Covid-19, or whether regulation should factor in the role of a central bank as liquidity provider of last resort.

Data shows that Fed short-term liquidity lines are quickly repaid, so this is not a bailout in the sense of saving an insolvent bank. Primary credit to banks – essentially, the discount window – peaked at more than $49 billion in the first week of April 2020, and by early August was already back down to just $3.5 billion. Even the strongest advocates of the post-crisis regulatory framework accept the liquidity rules may have had some unintended consequences.

Daniel Tarullo, Fed governor until 2017 and a long-standing critic of subsequent efforts by President Donald Trump's administration to roll back regulation, has talked about the need to move beyond regulations that simply encourage the hoarding of Fed reserves.

Addressing a Brookings Institute event in December 2019 on the role of regulation in the overnight repo market volatility, Tarullo said more attention should be paid to runnable wholesale funding used by non-banks. This could be a better alternative than piling up liquidity regulation on banks. His argument may gather fresh impetus after leveraged hedge funds were identified as one of the factors fuelling a Treasury market liquidity squeeze in March 2020.

"Having every bank completely self-insure its liquidity needs even under severe stress or failure may make sense as a microprudential matter," said Tarullo. "But if every bank must sit on its pool of readily usable liquidity in anticipation of possible failure, the result could in periods of stress be a decidedly suboptimal macroprudential policy that starves an already strained financial system of needed liquidity." ■

Previously published on Risk.net

[1] *US Federal Reserve Board (July 2020),* Federal Reserve Board finalises rule that implements technical, clarifying updates to Freedom of Information Act procedures and changes to rules for the disclosure of confidential supervisory information, *https://bit.ly/33uO1lW*

[2] *P Parkinson, US Federal Reserve Board (March 2010),* Letter to the officer in charge of supervision at each Federal Reserve bank, *https://bit.ly/30w6exG*

[3] *R Quarles, US Federal Reserve Board (January 2020),* Spontaneity and order – Transparency, accountability, and fairness in bank supervision, *https://bit.ly/3iyc1Jt*

[4] *BPI (April 2019),* BPI president and chief executive, Greg Baer, describes opaque subjective and unreviewable examination process in testimony before Senate Banking Committee, *https://bit.ly/33z7VMW*

[5] *R Quarles, US Federal Reserve Board (February 2020),* The economic outlook, monetary policy, and the demand for reserves, *https://bit.ly/34pfVPM*

TAKING THE TEST

The bank treasurer says: "In December, we receive two topics and the regulator tells us they're going to open the exam on January 15. On that date, we receive an incredibly long request for information with 40, 80, 150 line items of data requests.

"For the first two to four weeks, the Fed staff plough through that data, understanding and mapping our processes, and pulling out the line items of our particular capabilities that they want to go deeper on.

"They send their data line requests to us, and we send it back to them. Then they run their statistical sampling tools and they put out 80 or 100 items, and they ask for the source documentation.

"You're having calls with them every 24 to 48 hours and sending them information, either the very next day or two days later. Those two to four weeks are very intense. We are an enormous organisation with tens of millions of customers, so the sample sizes are very big.

"As an example, in a recent exam their sample size was 770 customer accounts. So in 14 days they went from several tens of millions down to 770. And then they said you have 48 hours to find us the 770 underlying transaction data, whether it's retail or trading book or corporate banking – here's all the different data, customer documentation that we want.

"And we provide it to them. That requires enormous capability across the bank to have the people at the back end of the organisation, where they analyse data and produce analysis on risk measurement, to go to the front-line people who own know-your-customer and anti-money laundering data, account opening, transaction booking system of origin system of records."

Covid scenarios

Finding the worst worst case

As the Covid-19 pandemic trashes historical data, a *Risk.net* tie-up with Ron Dembo's new outfit, riskthinking.AI, tests the promise of polling. By Tom Osborn

Need to know

- The path of Covid-19 – and the damage it wreaks on the economy – depends on a blend of health policy, politics, human behaviour, and efforts to find a treatment or vaccine.
- It can be difficult to generate robust stress scenarios when facing this level of uncertainty.
- One solution – touted by Algorithmics founder Ron Dembo – is to build scenarios using a large-scale poll of informed respondents.
- *Risk.net* has joined forces with Dembo's new firm, riskthinking.AI, to test the idea.
- The approach is controversial, not least for the difficulty faced in gathering enough sincerely held radical views.
- But if it gains support, the range of potential applications is broad, from helping to model the impact of climate change to gauging exposure to cyber terrorism.

The still-spreading coronavirus has extracted a heavy toll from the US economy, and triggered a catastrophic collapse in demand for US assets. The dollar will have lost 20.25% against the euro and 25.5% against the yen; the S&P 500 will plunge by a third. Meanwhile, the 10-year US Treasury will be yielding 0.1%.

It's a bleak, bruising scenario – it may not seem very likely today – but that's precisely the point.

The scenario is one of 64 derived from a survey of more than 300 *Risk.net* readers, carried out in late March. Their six-month forecasts of individual risk factors were knitted into multi-factor scenarios by riskthinking.AI – the start-up identified and combined the population's more extreme views.

To put it another way, this is crowd-sourced stress – and it's an attempt to tackle one of the key challenges that arises during periods of pervasive uncertainty, when backward-looking risk models are rudderless and the traditional fallback is to ask small groups of in-house experts for their judgement on how bad things could plausibly get.

That fallback is where many portfolio and risk managers currently find themselves, and some are not enjoying it.

A senior quant involved in scenario design at one US bank describes scenario construction for the pandemic as "an excruciatingly impossible task – I mean, I really, really don't like it". His beef with the process is that the bank is attaching finger-in-the-air loss estimates to a range of scenarios, rather than discussing and analysing its vulnerability to those scenarios.

A simpler objection comes from a former regulator: "I think most people within banks haven't got a bloody clue how bad this is going to get."

At one large Asia-Pacific bank, the regional head of model validation describes having to make Covid-19-enforced adjustments to a sea of indicators – from probability of default modelling for loans, to net interest income models – because the bank expects the eventual impact of the virus to be far worse than the worst-case losses implied by any recession in its dataset.

"Right now, we're in a unique situation: a lot of the analytical tools and models that are seen as business as usual in stress-testing are either not relevant, or giving outputs that don't make a lot of sense, simply because of the unique movements we've seen," he says. "So, you're going to have a lot of uncertainty with the traditional models you've got in place – that's understood. Management decision-making and overrides, a revisiting of assumptions, is a key process that's currently underway."

In that context, he suggests, simply asking a large number of people how bad they expect things to get seems a reasonable exercise.

That's where the survey comes in. Polling outsiders is not the way stress scenarios are normally constructed, and some dismiss it as a solution. Ron Dembo – founder of riskthinking.AI and, in 1989, of modelling vendor Algorithmics – argues it's an idea whose time has come. Even before the pandemic injected uncertainty into a host of critical risk factors, market participants were struggling to work out their exposure to climate change, to cyber crime and to technological change. Those struggles will continue once the questions associated with Covid-19 have been answered.

Outlier views

The predictive capacity of the pilot survey can already be tested.

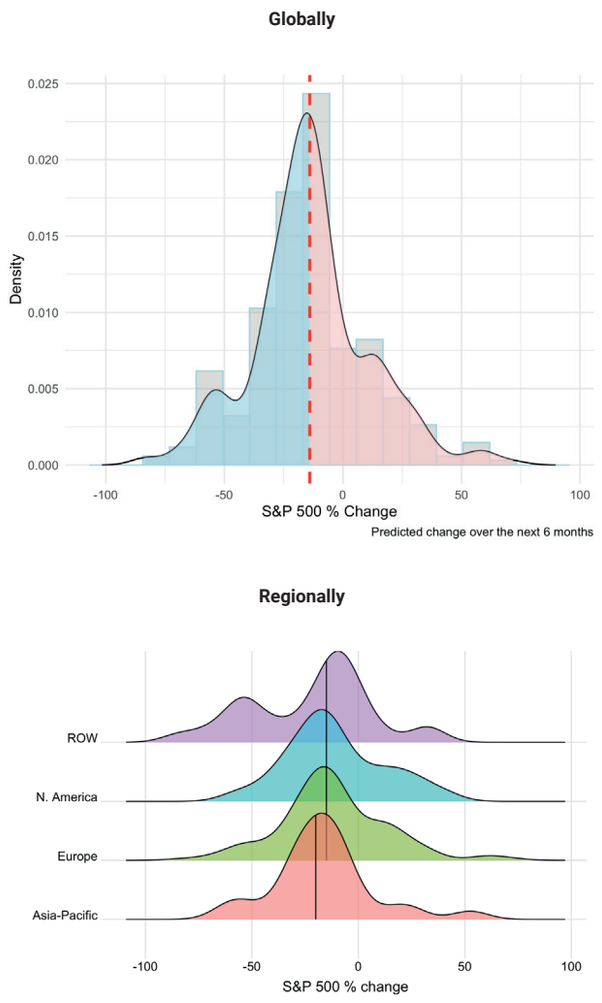
Respondents in Asia-Pacific expected, correctly, to be locked down for less time than their peers in the US and Europe. Risk managers were marginally less pessimistic about the ensuing first-quarter hit to US GDP than their peers in banking, broking and consulting, but everyone expected it to be pretty bad: a median average of all those surveyed predicted a 5% drop in GDP – significantly worse than the 3.5% being predicted by economists at the time, and almost bang on the actual 4.8% fall published by US statisticians in May. By the end of September, almost all respondents expected the S&P 500 would be some way below its March 26 closing level of 2,630, with a 15% drawdown being the average view. The index was at 3,363 on September 30.

The point of the exercise was not to find a consensus, however: the shape of the distribution of responses is the important part – and, within them, the extremes weighted against the overall distribution. These outlier views serve as the inputs for stress scenarios that are built from the survey, based on combinations of multiple factors arrived at via a decision tree analysis.

Dembo acknowledges the approach is not without controversy – particularly, given the importance placed on them, the crucial process of distinguishing between sincerely held outlier views and mischief-making. The former are plugged into the scenarios, while the latter is discarded. As Dembo puts it: "It's an art, weighing up whether someone is trolling us, or for real."

Even as a method of reliably spotting black swans, "people will find it controversial, and we accept that", Dembo adds.

1 S&P 500 change



Source: *Risk.net*/riskthinking.AI

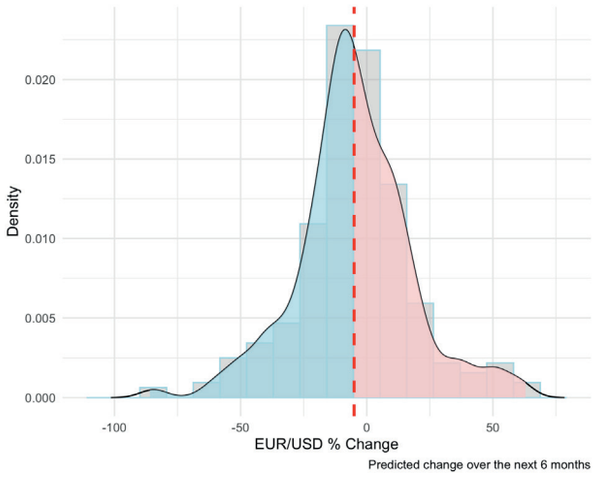
He's not wrong. Those used to being given a scenario expressed in terms of its impact on financial markers express bewilderment at being asked to provide the opposite.

"How do you 'detect' something that's used as input?" asks the chief risk officer at a large European asset manager. "A black swan is an unexpected event. If you start by defining an event, and then ask experts how it impacts markets, your model doesn't 'find' black swans. It is the input of the model, not the output."

Dembo responds: "If I want to find genuine black swans, I need an extreme range of views. We as individuals are pretty bad at generating extremes. But on single factors with a diversity of opinions, we might uncover more extreme views than the consensus. We are completely useless at dreaming up scenarios on multiple factors, however. That's why we poll a broad range of experts and seek well-justified extreme views.

"We believe people are much better at capturing that uncertainty in a single factor," he adds.

2 Euro/dollar change viewed globally



Source: *Risk.net*/riskthinking.AI

He contrasts the approach with the way scenario analysis is usually conducted in banking, wherein an end-state is pre-defined – either by dedicated teams within the organisation, or set by its supervisors – and the lender then sets about decomposing its impact on its portfolios and loan books to work out its all-in exposure, usually as a means of assessing capital adequacy to those events.

That process itself, though based on quantitative techniques, relies a lot on expert judgement and layering of assumptions to arrive at an agreed impact estimate. More importantly, Dembo contends, if firms are using it to establish a true picture of their exposure to genuinely extreme events, then they're going about it the wrong way.

There is some sympathy for that view. The Asia-Pacific bank's regional head of model validation sees polling as a way to benchmark or challenge the in-house views, rather than replacing them. "The scenarios give you a range of current market views, which – on top of what you're doing anyway – will hopefully give you added visibility, and hopefully allow better financial decisions to be made, anticipating how many reserves to set aside for the credit books, things of that variety. From that point of view, it's a good exercise to go through," he says.

The distributions

A look at the distributions shows some respondents foresee further heavy losses across equity, bond and currencies portfolios.

Figure 1 shows respondents' predictions for percentage gains or losses for the S&P 500, relative to its closing level on March 26 (2,630). A mean average of respondents predicted a 14% decline for the blue-chip index – but a small group at the lower bound predicted a drop of more than 80%. A few optimistic souls, currently closer to being in the money, predicted a 72% gain.

A wide spread of responses might appear surprising, but Evan Sekeris, who previously oversaw operational risk at the US Federal Reserve, sees this as a key advantage of the crowd-sourcing approach: there is information in the shape of the distribution, not just in the extremes that are used to generate the scenario.

A relatively narrow distribution of responses – say, between -10% and -20% – would suggest less uncertainty on the topic being polled, he argues. "But if you have a distribution like you have here, where some people are saying it could go to -80%, while others are telling you it could double, or go up 50% or 100%, then that's a completely different picture, right? Which is: nobody really knows. The consensus seems to be around this number, but there are small sub-buckets – one in the positive and one in the negative – that strongly think otherwise. That information alone is very important."

The picture is similar for expectations of the euro's value against the dollar. An average of 303 respondents predict a 5% decline for the single currency against the greenback over the next six months – but those at the lower bound predicted an 86% decline, while those at the other end of the distribution expected a 63% gain.

Interestingly, while regional expectations were broadly in line with this picture, the distribution of responses from North America, consisting of 60 votes, was tighter: a mean average expected a 4.4% decline for the euro, while the most extreme predictions were of a 57% decline – and, at the other end of the scale, a 61% decline for the dollar. In Europe, the more extreme predictions of a 161 sample count were more severe: a mean decline of 3.1% was anticipated, with a median of -5%; but some predicted an 83% decline for the euro, while others saw a 58% gain versus the dollar.

This is surprising, says one leading academic on the construction of stress scenarios; ordinarily, with a larger sample size, a distribution pattern would tend to be tighter. In fact, with respect to the shape of the distribution's tails, the opposite appears to be the case – again, lending support to Dembo's premise that if one garners enough views, a statistically significant number of respondents in a sample will predict an extreme outcome.

"Usually by n=300, you should have a pretty normal distribution – and you didn't have that, across a lot of things. You had a pretty decent sample size, and yet you had some pretty funky distributions. I was surprised you didn't have a greater convergence around 'normality'," he says.

The shape of those distributions suggests "there's a lot there to tease out about what is at the edge of people's beliefs," he adds. "Even if you just say 'there's always a few outliers', that doesn't explain why you still have, at one or two standard deviations, quite a number of professional investors and the like who believe something that seems to be quite outlandish, relative to current asset prices. That's a train of thought that I think would be really useful for investors, for your readers, for policy-makers."

The survey also highlights the starkly different realities respondents in different parts of the globe expect to be living with while the pandemic lasts. Asked how many weeks it would be until the majority of schools, businesses and other places of work are "functioning normally" in their own region, expectations were again wildly different. While the mean expectation was of an 18-week lockdown, a significant chunk of respondents said they expected it to last for 52 weeks – although here, much might have turned on respondents' interpretation of "functioning normally".

Viewed regionally, a clear split on lockdown expectations emerges between respondents in the US and Europe versus those in Asia-Pacific and the rest of the world. All four regional distributions are similarly shaped, with the mean expectation falling in a tight range between 17 and 20 weeks, but the positioning of the lower bound for each varies hugely, on a proportional basis: in Europe and the US, the shortest lockdown anyone anticipated in late March was four weeks; in Asia-Pacific, where South Korea avoided a complete shutdown of the economy altogether, the lower bound fell at one week.

The scenarios

In Dembo's conceit, it is the positioning of the lower and upper bounds that is crucial as a means of challenging consensus expectations; the stress scenarios designed to capture black swans – unlikely but plausible events combining in unexpected ways with hard-to-predict consequences – are themselves based on combinations of these extreme views, constructed via a decision tree analysis.

Sekeris agrees with that approach. He too argues banks should start by determining the path that leads to a loss or an event by decomposing it – so, by finding the different variables in the decision tree that lead to that final outcome. The tree itself then becomes hundreds of final outcomes that are the various possible combinations of all the different nodes, and the different paths events might take.

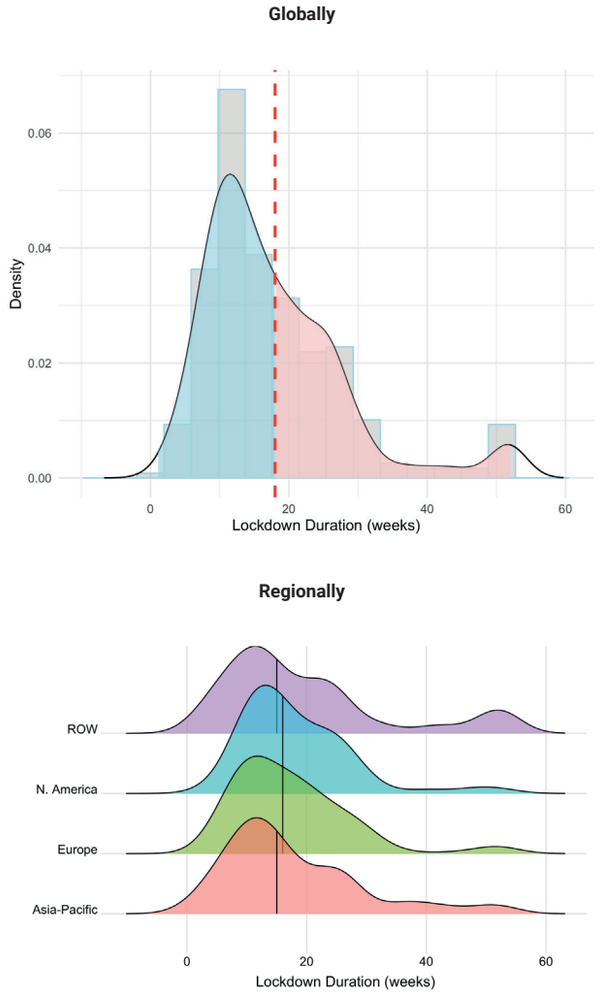
"But the idea is that then when you go to the experts, you don't ask them to tell you what the final outcome is. Because the point is, no matter how much of an expert you are, when you're asking just the final outcome, you're basically asking them to mentally do all the calculations of that tree – to go through all these three steps and say, 'Okay, if the S&P goes to here, the next step is, this happens, then this happens – and here's my outcome.' And what happens then is, you lose that all-important range of numbers at the end, because they'll say, 'our worst-case expectation is a loss of X'. Well, plus or minus how much? It's never going to be a loss of exactly X," he says.

Worse, he adds, the way the expert arrived at that loss figure is highly likely to be prone to error and bias – a product of the computational shortcuts that someone is taking in trying to come up with that final number, without focusing on the elements before.

"What you're doing here with your structure is forcing people to ignore that final number and to focus on the elements of the tree, the nodes. And that's very important, because those nodes are things that they understand better, and they can provide you information much more clearly about," he says.

Of the scenarios contained in the decision tree, 64 are then combined into a table. Each is assigned a likelihood score, ranging from 2.99% to 0.73%.

3 Lockdown view



Source: *Risk.net*/riskthinking.AI

The scenario with the highest likelihood – number 47 on the table – posits the following: a 12-week lockdown; a 7.5% gain for the dollar versus the euro through September 26; a 2.5% gain versus the yen; a 25% gain for the S&P 500; a yield of 0.9% on US 10-year bonds – and a -21% first-quarter hit to US growth.

The least likely scenario – summarised at the start of this article – encompasses: a 30-week lockdown; a 20.25% decline for the dollar against the euro; 25.5% decline against the yen; a 33% decline for the S&P 500; a yield of 0.1% for US Treasuries; and a 3% first-quarter decline in GDP.

At press time, scenario 47 certainly appears more true to life: most Group of 20 economies remain in some form of lockdown, though roughly 12 weeks on, many of the conditions attached are steadily being relaxed; the euro is currently up 1.3% on the dollar, while the greenback is up 21.3% on the yen. The S&P is currently up 15.7%, while the yield on US Treasuries stands at 0.65%. US GDP declined by 4.8% during the first quarter; current estimates put the hit to Q2 GDP anywhere between -12% and -40%. In other words, as Sekeris says, "nobody really knows". ■

Previously published on Risk.net

A. Covid-19 stress scenarios

For the four financial indicators shown, the value given is a six-month estimate of the % change versus the market's closing price on March 26

Scenario	Lockdown duration (weeks)	EUR/USD	USD/JPY	S&P 500	US Treasury 10-year note	US GDP (%)	Likelihood (%)
Scenario 47	12	-7.5	2.5	25	0.9	-21	2.99%
Scenario 58	9.33	-15.33	-11	3.33	1.2	-11.67	2.73%
Scenario 62	7.67	-12.33	3	-21	1.27	-21.67	2.69%
Scenario 40	11.6	-9.4	9	-9.2	0.22	-18	2.62%
Scenario 59	11	-10	-11.5	-20	1	-20	2.46%
Scenario 46	11.2	-14.2	-14.6	-3.6	0.16	-14.8	2.40%
Scenario 38	12.83	-22.67	23.17	-24.83	0.2	-22	2.37%
Scenario 64	4	17	26	36	3	-25	2.36%
Scenario 43	10.75	-35.75	-41.88	-39.63	0.29	-25.25	2.16%
Scenario 45	9.75	7.5	-12.25	-4.75	1.55	-17.75	2.16%
Scenario 35	14.4	10.8	13.6	-32.4	1.14	-15.2	2.13%
Scenario 54	12.5	7.5	3.5	15	0.55	-10	2.08%
Scenario 26	23	-9.5	7	15	1.65	-15.5	2.02%
Scenario 39	12	14.5	-10	-18	1.05	-14.5	1.95%
Scenario 56	10	10	-17	-1.33	0.4	-18	1.90%
Scenario 51	8.79	-10.93	10.79	10.93	1.27	-2.43	1.89%
Scenario 49	12.6	16.2	18.2	-40.8	0.41	-27.4	1.87%
Scenario 30	20	-15	-17.5	10	1.15	-22.5	1.84%
Scenario 2	29.8	-25.8	23.8	-43.4	2.06	-19.4	1.82%
Scenario 4	39	-25.5	4.5	34	0.3	-10	1.77%
Scenario 55	9.86	-23.14	-17	9.86	0.99	-0.86	1.72%
Scenario 61	8	19.33	-30.33	-41	0.2	-22.67	1.71%
Scenario 50	10.75	-20.63	15.13	-26.63	1.07	-1.5	1.70%
Scenario 21	24.75	-33.5	-33.5	-30.75	1.05	-17.75	1.66%
Scenario 42	12.44	-16.89	15.67	6.78	0.26	-1.78	1.66%
Scenario 18	28	-25	-30	-5	0	-15	1.62%
Scenario 12	29.4	-22.4	23	-25.2	0.04	-16.4	1.60%
Scenario 29	22	31.5	28	1	0.85	-15.5	1.60%
Scenario 44	13	-24.6	-18.4	-22.8	1.4	-2.4	1.55%
Scenario 57	10.5	-14.25	-9.25	-3	0.28	-1.5	1.51%
Scenario 48	12.83	-19	8.33	-25.17	0.27	-3.5	1.49%
Scenario 36	10.08	8.38	9.08	3.85	1.16	-0.69	1.49%
Scenario 7	26.91	-30.27	-21.91	-33.82	0.24	-18.45	1.46%
Scenario 11	30.67	25.83	-24	12.67	1.73	-16.17	1.46%
Scenario 19	26	27.67	22.67	-37.33	1.2	-19	1.44%
Scenario 10	52	40	8	41	0.5	-50	1.40%
Scenario 37	13.29	-22.93	-19.93	-25.71	0.26	-0.5	1.37%
Scenario 41	11.9	7.1	-10.8	5.1	1.27	-2.2	1.36%
Scenario 53	11.25	15.5	11.25	-30.75	1.2	1	1.35%
Scenario 22	23.5	35.5	-27	-30	1.5	-12.5	1.32%
Scenario 63	8	1	-3	-14	0.1	2	1.31%
Scenario 25	23	10.5	-15.5	20.5	0.25	-17.5	1.28%
Scenario 8	30.67	-21.33	12.67	7.33	1.27	-3.33	1.27%
Scenario 13	27.75	16	1.5	-36.5	0.23	-26	1.27%
Scenario 60	10.33	10.33	-13.33	-25.67	0.87	2.67	1.23%
Scenario 34	14.2	12.2	-13.8	-3.6	0.26	-1.6	1.20%
Scenario 33	12.5	8	7.5	-29.75	0.23	-2.75	1.18%
Scenario 24	23.33	-12	-15	9.33	1.6	-1	1.16%
Scenario 23	25.33	44.33	-44.67	-45.67	0.07	-28.33	1.16%
Scenario 32	19	-12.5	15	-22.5	1.5	-4.5	1.15%
Scenario 9	31.33	-15.17	14.33	-2.33	0.27	-1.33	1.12%
Scenario 52	11.25	11.42	-21.83	-28.25	0.22	-2	1.08%
Scenario 16	30.8	-11.4	-14	-25	1.26	-4.2	1.05%
Scenario 5	34	-16.83	-17.17	-10.17	0.18	-1.17	1.02%
Scenario 6	22.67	-15.33	4.33	-22.67	0.33	-3.67	1.01%
Scenario 14	26	15.83	7.83	13.33	1.42	5.33	1.01%
Scenario 15	25.71	-27.71	-20.43	-38.43	0.09	0.57	0.92%
Scenario 28	22.83	18.33	-20.67	6.33	1.07	-2.67	0.92%
Scenario 20	23.5	1.5	0.5	-27.5	1.4	-0.5	0.91%
Scenario 27	23.33	0.33	1	19.67	0.17	-1	0.89%
Scenario 31	20	10	-10	-25	2	-1	0.83%
Scenario 17	28.67	6	-16.67	-0.67	0.17	-0.33	0.81%
Scenario 3	28	9.5	14.33	-33.5	0.32	-3.17	0.80%
Scenario 1	30	20.25	-25.5	-33	0.1	-3	0.73%

Risk.net | Follow Topics

Build your own email alerts

Follow our coverage on individual organisations like the ECB or the Federal Reserve. Or pick specific topics such as CCPs, FRTB or Mifid II. Plus many more. Select your topics and you'll receive everything you are interested in, all in one email.

Visit risk.net/follow

Apocalypse how?

Risk.net and riskthinking.AI's second crowd-sourced scenario exercise reveals polarised views in equities and foreign exchange. By Tom Osborn and James Ryder

Financial markets have been living with Covid-19 for more than eight months, giving participants some time to get to grips with the impact and implications of the disease – but, over that period, their predictions have become more polarised, rather than less.

That is the key finding of *Risk.net*'s second set of crowd-sourced stress scenarios, constructed by polling readers – and run once again in conjunction with riskthinking.AI.

Some poll respondents expect a vaccine to be available within a year, while others expect it to take five or more. Similar extremes can be seen in forecasts of financial benchmark levels – a 50% rise for the S&P 500, for example, or a 68% collapse; a 100% widening in credit spreads, or a 61% decline. That array of paths is making life difficult for portfolio and risk managers.

"Our attitude is: we need to have different scenarios for totally different potential future states of the world, because we don't know what the world will look like in the middle of 2021," says the chief risk officer (CRO) of one large European bank.

This is where the crowd-sourcing exercise comes in – it's an attempt to generate a set of genuinely taxing scenarios, which are rooted in the views of market participants without being beholden to any particular perspective. As with March's pilot survey, *Risk.net*'s audience of finance professionals were asked to give a six-month forecast on the path of a series of key financial indicators, relative to their values in late June. These views were used to construct the multi-factor stress scenarios found towards the end of this article.

Results

Markets are certainly calmer than they were in March, and individual respondents may feel more certain about their direction, but a lumpy distribution on most of the polled factors shows there is little consensus – with the tails providing a far wider set of extreme views than last time around.

Evan Sekeris, who previously oversaw operational risk at the US Federal Reserve Board, is not surprised.

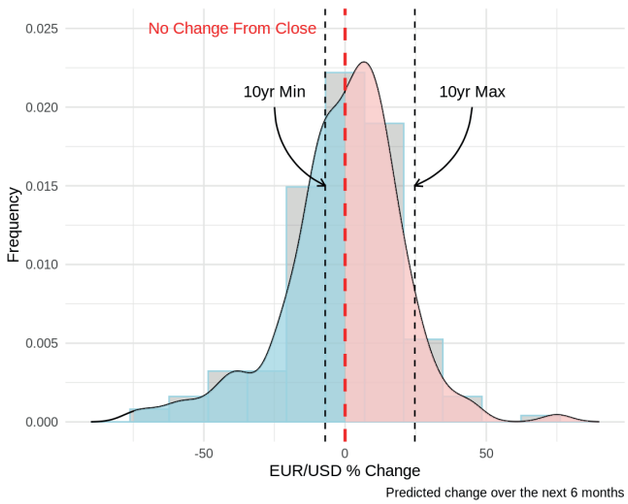
"We face not just an uncertain future, but one in which we could be looking at states of the world, in the next few months, that are unimaginable. I could point to the more extreme scenarios, which have low probabilities, and say, looking at certain elements of the data, that they actually have a much higher probability," he adds, in reference to the scenarios' likelihood scores, which reflect the compounded joint probability distribution of each risk factor.

"The fact that it has never been seen does not make it next-to-impossible – it's just that we don't understand how to measure it," he adds.

Take, for instance, respondents' views on the dollar's value against the euro. Relative to the June 22 price of $1.12, a mean average of respondents foresee little change, predicting a decline of less than 1% for the dollar against the single currency from that level. The median average of views was broadly similar, pointing to a 2% gain.

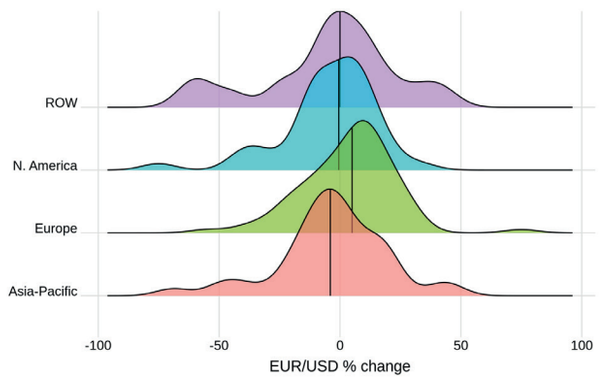
The extremes of the distributions, however, lie a long way from these averages. The only consensus between dollar bears and dollar bulls is their belief in just how far the greenback could rise or fall: some see an increase of 75% relative to that $1.12 baseline price, while others predict the rate falling by the same percentage. Both extremes would take the single currency beyond the bounds of history: its record high of $1.60 came in July 2008, while its all-time low of $0.82 was in October 2000, soon after launch.

1 EUR/USD change – globally



Source: *Risk.net*/riskthinking.AI

2 EUR/USD change – regionally



Source: *Risk.net*/riskthinking.AI

Similarly wide extremes are seen in forecasts for the S&P 500. Relative to its closing value on June 22, when it stood at 3,117, the most pessimistic respondents predicted the index collapsing by two-thirds over the next six months to just under 1,000 – levels not seen since the aftermath of the global financial crisis that began in 2007–08. The most optimistic believe it could continue its post-Covid-19 rally, and extend its gains above 4,500 – a rise of more than 50% from its March lows, and comfortably into record territory.

Backtest

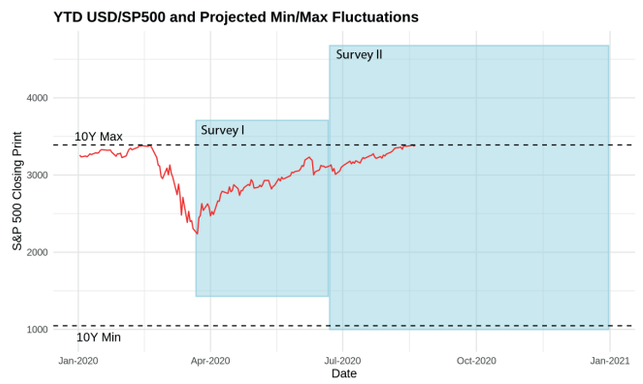
The backtest on the S&P 500 between this survey and the last underlines how views have become more extreme. When *Risk.net* first asked participants to predict the impact of Covid-19 in late March, the S&P was nearing a multi-year low of 2,191.

Yet the range of views then – when markets were still in freefall, and central banks and governments had yet to firm up pledges of trillions of dollars in ultra-cheap credit and other support measures – was far narrower than it is now.

"Our attitude is: we need to have different scenarios for totally different potential future states of the world, because we don't know what the world will look like in the middle of 2021"

Chief risk officer at a large European bank

3 S&P 500 – survey backtest



Source: *Risk.net*/RiskThinking.AI

"I'm not surprised," says a stress-testing expert at one financial think-tank. "A lot of the time when we're thinking about the transaction costs of crises, we think about economic dislocation – which is horrendous, and I don't mean to belittle it. But the raw uncertainty in financial markets drives risk premia a lot. When share prices fell in February, was that a function of assumptions around future cashflows? Or was it higher discounting of future cashflows because of the uncertainty?"

Scenarios

The first step to take when dealing with uncertainty, is to acknowledge it. The European bank CRO says his institution's approach is to compile scenarios that envisage very different paths for the major drivers of a portfolio. For instance, the bank might construct a base scenario where US assets retain their haven allure in a crisis, he says, and it might then construct a rival scenario in which that allure is lost. This alternative future might involve the Swiss franc rallying to levels not seen since the aftermath of the financial crisis, or the euro strengthening to pre-crisis levels, as it assumes the role of a genuine global reserve currency.

Crowd-sourcing scenarios have some appeal, the CRO says; the informed audience of respondents will have access to the same public information on the same portfolio drivers his team is considering – but will also be factoring non-public data into their own opinions, based on the composition of their own portfolios, for instance.

"I think [this approach] absolutely has merit. You have 179 responses, [and] in these responses you will probably have the most diversified underlying sources of wisdom… And I think the broad range of answers – on currencies, for instance – shows the full spectrum is incorporated. So I think your approach is probably superior to many standalone studies where just one – or maybe two, or three, or five – experts look at the problem and create common sense narratives about everything they know and have read and experienced about the current environment," he says.

On the face of it, common sense is missing from some of the latest batch of scenarios. In scenario 42, for instance, while US GDP posts a huge upside surprise – no decline in output at all – the US dollar loses 5% of its value against the euro, the S&P 500 rises by the same amount, credit spreads tighten by 10% and the 10-year US Treasury yield more than doubles to 1.5%. That is an odd combination of events.

Still, looking back to the previous survey, conducted in the dark days of early lockdown in the US and Europe when markets were in freefall, very few posited scenarios in which restrictions on movement lasted more than 30 weeks: this regime saw a significant decline in the dollar against the single currency, a 7.3% gain for the S&P 500, and the US 10-year hitting 1.3%.

Come June, at least some of those estimates were looking good: the Fed's spectacular monetary easing programme had driven the dollar down against the euro, and driven a stonking recovery in US equities. The early estimate of the first-quarter hit to GDP was almost bang on, at -3.3%.

At the time, this particular scenario – number 8 – was given just a 1.9% chance of occurring.

This is the default way the scenarios below are grouped, too. Likelihood is estimated as a function of the joint probability distribution of each risk factor relative to its modal consensus value. These are then multiplied by one another to give a likelihood score.

Constructing the scenarios in this manner also means scenarios that no single respondent posited – equities falling in value in lockstep with Treasuries, for instance as briefly occurred in March – can be generated.

The score does not reflect incidence of past occurrence, nor does it take into account historical cross-asset correlations, both of which would ordinarily be cornerstones of scenario design for risk managers. As riskthinking.AI's founder Ron Dembo sees it, though, that's the point: to move away from traditional approaches to stress-testing, or at least subject them to challenge via the wisdom of the crowd.

Readers are free to evaluate the usefulness of the likelihood function for themselves, or to group the scenarios by any other theme. One way of doing so is by forecasts for second-quarter US GDP; for instance, the most pessimistic scenario, number 35, posits a -37% fall in GDP – the current official estimate published subsequently is a 32.9% contraction – and imagines a catastrophic 68% collapse in equity prices by year-end, and a two-thirds widening in credit spreads.[1]

Conversely, such a doomsday outcome appears to have led to a rally for dollar assets, with the greenback gaining by almost half versus the single currency. Such a scenario is ascribed a likelihood score of 0.37%.

Those who saw an upside surprise to GDP – a contraction of just 20%, as in Scenario 14, say – foresaw a market rally: a continued rise in the S&P 500 of 16% from its June 22 level, and a dramatic rush for US Treasuries sparking a collapse in yields to -1.3%. This scenario is ascribed a likelihood score of 2.04%. ■

Previously published on Risk.net

[1] *Bureau of Economic Analysis, US Department of Commerce (July 2020),* GDP, Q2 2020 (advance estimate) and annual update, *https://bit.ly/2F5H5m5*

A. Covid-19 stress scenarios – Three months on

For the four financial indicators shown, the value given is a six-month estimate of the % change versus the market's closing price on June 22

Scenario	Years until a Covid-19 vaccine	EUR/USD	CDX IG index	S&P 500	US Treasury 10-year note	US GDP (%)	Likelihood
Scenario 17	1	33	50	-34	-0.5	-30.5	5.83%
Scenario 13	1	41	30	-30	-2.6	0	5.76%
Scenario 12	1	-29	50	-45	-1.7	-28	5.39%
Scenario 2	2	-25	30	-35	-0.2	0	5.33%
Scenario 18	1	20	-33	-31	-1.2	-25	4.60%
Scenario 25	1	28	-45	-36	-0.5	2	4.55%
Scenario 6	1	-75	-61	-68	-2.4	-33.5	4.26%
Scenario 8	1	-58	-49	-64	-0.5	0	4.21%
Scenario 9	4	27	40	-45	-1	-25	2.94%
Scenario 15	5	43	100	-49	-0.5	0	2.90%
Scenario 3	2	17	20	15	-0.3	-15	2.79%
Scenario 21	1	17	59	41	-0.9	2.5	2.76%
Scenario 24	5	-21	43	-36	-1.2	-23.5	2.72%
Scenario 37	5	-41	20	-25	0.5	-5	2.69%
Scenario 26	2	-40	30	20	-2	-20	2.58%
Scenario 19	1	-15	79	15	-1.3	2	2.55%
Scenario 16	5	15	-20	-20	0	-35	2.32%
Scenario 5	5	14	-9	-14	-0.2	0	2.29%
Scenario 10	1	75	-20	20	-0.4	-30	2.21%
Scenario 30	1	33	-37	50	-1	2	2.18%
Scenario 28	5	-40	-45	-38	-0.7	-23	2.15%
Scenario 20	4	-15	-15	-15	0	0	2.12%
Scenario 14	2	-16	-10	16	-1.3	-20	2.04%
Scenario 11	1	-10	-10	15	0.3	0	2.02%
Scenario 33	3	25	35	20	-0.6	-19	1.41%
Scenario 53	5	-75	-9	-3	1	1	1.39%
Scenario 22	4	-25	30	50	-0.5	-20	1.30%
Scenario 41	4	-14	22	8	0	0	1.29%
Scenario 4	3	15	-10	5	-0.7	-14	1.11%
Scenario 45	5	10	-5	5	0	0	1.10%
Scenario 27	5	-20	-10	10	-1	-16	1.03%
Scenario 57	2	-67	-45	-29	-1	1	1.02%
Scenario 32	1	10	15	-20	1.2	-15	0.73%
Scenario 31	2	2	8	-7	1	0	0.72%
Scenario 34	2	-15	40	-25	1.2	-20	0.68%
Scenario 61	5	41	-51	3	1	2	0.67%
Scenario 56	5	74	10	-18	0	1	0.58%
Scenario 38	2	16	-23	-16	1	-11	0.57%
Scenario 54	4	-1	3	48	1	2	0.54%
Scenario 63	2	69	-33	49	2	2	0.53%
Scenario 23	3	25	30	-25	0.9	-4.5	0.37%
Scenario 35	3	46	63	-68	1.4	-37	0.37%
Scenario 36	1	12	19	27	1	2	0.35%
Scenario 40	2	24	3	41	1.1	-22.5	0.35%
Scenario 50	3	5	95	44	1	2	0.34%
Scenario 52	2	-9	96	-19	1	1	0.34%
Scenario 29	2	-5	16	35	2	-4.5	0.32%
Scenario 60	1	-28	-14	-53	1	2	0.32%
Scenario 47	4	6	-27	-36	1	1	0.29%
Scenario 55	5	7	-59	-6	-2	2	0.29%
Scenario 1	2	5	-10	10	1	-14.5	0.28%
Scenario 39	3	-15	-20	-10	0.9	-5	0.27%
Scenario 43	2	24	-10	7	1.3	-1	0.27%
Scenario 58	3	22	-41	10	0	1	0.27%
Scenario 51	2	-67	85	-61	0	1	0.26%
Scenario 42	1	-5	-10	5	1.5	0	0.25%
Scenario 46	5	5	5	15	0.9	-12	0.18%
Scenario 59	2	-45	-52	50	2	2	0.18%
Scenario 48	3	75	19	10	0	2	0.16%
Scenario 49	1	-5	67	43	1	2	0.16%
Scenario 7	3	5	0	11	0.9	-10.5	0.14%
Scenario 44	3	3	-10	5	0.9	-14	0.14%
Scenario 62	2	5	51	-3	-2	2	0.13%
Scenario 64	1	-60	-21	44	0	1	0.13%

Ron Dembo on crowd-spotting black swans

Veteran quant and Yale professor turned serial risk software entrepreneur Ron Dembo argues that large groups are better at gauging extreme uncertainty than small teams of experts. By Tom Osborn

The future isn't what it used to be – at least, not if judged by the financial industry's efforts at modelling it.

Time and again, the standard method of estimating losses by looking at what has happened before – imagining the future by selectively replaying the past – has been found wanting.

Covid-19 is just the latest example. For every day that passes with the world's largest economies under lockdown, the outlook darkens – tens of billions of dollars in loan losses, double-digit hits to growth, and runaway unemployment.

Could banks have seen any of this coming? Ron Dembo thinks so.

The veteran quant – a former Yale professor turned serial risk software entrepreneur – has a new venture, riskthinking.AI, which is trying to bring about a mindset shift in the field of scenario generation.

The basic aim of scenario generation is to build a forward-looking gauge of risks for which there are few or no precedents to rely on. Various quantitative approaches are used to estimate losses by working out what could go wrong, and how bad the consequences could be. The outputs of the analysis are then used to put a dollar value on a firm's risk exposures.

But for events whose impact carries a high degree of uncertainty, such as climate change, cyber risk or indeed pandemics, Dembo argues that the classical approach – using a small cadre of in-house experts to work out what could go wrong, and then working backwards to see how the firm could be affected – is misguided at best, and dangerously misleading at worst.

His instinct is instead to trust in the wisdom of crowds: polling a broad audience of risk and finance professionals – *Risk.net* readers, in the case of a recent survey on Covid-19 – to generate estimates of an event's impact on a range of key financial indicators. These are then layered together in different combinations to form scenarios. The basic premise being, if one garners enough views, a statistically significant number of respondents in a sample will predict an extreme outcome.

Dembo accepts many in finance, particularly quants, will find the approach hard to stomach. But he insists it is the best hope firms have of getting a handle on risks that classical statistical techniques probably can't handle.

"The point we're making is, by seeking enough expectations, you can derive extreme events. When you have extreme uncertainty, you want to capture that – that's the ultimate driver of our scenarios. We believe people are much better at capturing that uncertainty in a single factor," he says. "This is not the way people normally construct scenarios. People will find it controversial, and we accept that."

As evidence, he points to how few firms or governments anticipated the impact of Covid-19. Yet pandemics have happened before, many of them more deadly, he points out. If a firm had cast the net wide enough in January, it could have captured the outlier views of those who believed the pandemic could spread rapidly, that countries would enter lengthy lockdowns, and that extreme economic damage would be the result.

"If you'd just polled epidemiologists, they'd have said: 'Here's one possible scenario: this thing goes nuts, and flies all over the world.' You can extrapolate from that: 'What happens if we have to shut down entire economies?' But the question is: would you have extrapolated from that starting point and given enough thought to the potential consequences without polling a broader number of financial professionals?"

A tricky part of Dembo's approach is weighting the extreme possibilities suggested within a broad sample, giving an indication of how likely respondents believe they are to occur. This is achieved by weighting outlier views against the overall distribution of the results.

Before that, though, the firm's quants must apply judgement to distinguish between genuine outlier views, and people not taking the exercise seriously – an approach they acknowledge is open to accusations of bias.

"It's an art, weighing up whether someone is trolling us, or for real. Sometimes, when you look at the responses, you see 100% changes for a given factor. Sometimes, we take those into account – because they're not crazy. But when you look at distributions and see a 50% change in a currency – when did we last see that in the euro in a few months? It couldn't happen. But it is an indication of the wild extremes people expect. You can then weight the responses any way you like, but the reality is there's a real person who said this – and in the case of the poll we conducted with *Risk.net* readers, they probably do work for a bank."

Why is polling large groups of professionals capturing better response to extreme uncertainty than asking a small group of in-house experts?

Ron Dembo: Starting with first principles: why do banks and financial firms generate scenarios? They do it to somehow get their arms around future uncertainty. No-one believes most of the scenarios they build will come to pass – they are there to guide future strategy in the face of uncertainty.

But there's a fundamental misunderstanding of scenario generation within banks and public bodies that it's just another form of forecasting, one that can be done by a small cadre of dedicated experts. It's not. In the biggest of banks, in the highest of positions, scenario thinking is not well understood. To quote one such discussion I had, the person in charge said: "We generate many scenarios, and then select the best one."

Why is asking more people necessarily a better guide? Don't you just arrive at a broader market consensus?

Ron Dembo: If I want to find genuine black swans, I need an extreme range of views. We as individuals are pretty bad at generating extremes. But on single factors with a diversity of opinions, we might uncover more extreme views than the consensus. We are completely useless at dreaming up scenarios on multiple factors, however. That's why we poll a broad range of experts and seek well-justified extreme views. And that's why we have developed an algorithm for combining individual factor uncertainty into scenarios on multiple factors.

For example, the recent bushfires in Australia were caused by a combination of unusual heat, high winds and drought. That's a particular combination – if you'd just looked at the likelihood of each of those in isolation, you wouldn't have uncovered useful information. Estimation of multi-factor black swans is the challenge.

But the black swans you're looking for are based on a combination of the single-factor inputs offered by survey respondents. How does your approach detect a black swan from single factors?

Ron Dembo: The inputs are the uncertainty reflected in those single factors – scenarios constructed on combinations of multiple factors are the outputs.

Put another way, when is a black swan a black swan? What if I'd told you, in December, that some observers had detected an outbreak of some sort of strange flu in China? As it ripples out across the rest of the world, there's a small chance that's the start of a pandemic. It's not like you'd never have thought a pandemic could spread in such a way. But if you'd asked anyone then to predict what we're experiencing now, they'd have said: "Oh, that's a black swan."

Yet if, in January, you knew there was a virus in Wuhan, and you'd then looked at data from the 1918 Spanish flu, you'd have been able to construct a scenario that looked a lot like what we're facing today. If you'd just polled epidemiologists, they'd have said: "Here's one possible scenario: this thing goes nuts, and flies all over the world." You can extrapolate from that: "What happens if we have to shut down entire economies?" But the question is, would you have extrapolated from that starting point and given enough thought to the potential consequences without polling a broader number of financial professionals?

The point we're making is, by seeking enough expectations, you can derive extreme events. When you have extreme uncertainty, you want to capture that – that's the ultimate driver of our scenarios. We believe people are much better at capturing that uncertainty in a single factor. This is not the way people normally construct scenarios. People will find it controversial, and we accept that.

Even if you had polled a broader number of financial professionals and concluded this was one possible outcome, what probability would you have assigned to it?

Ron Dembo: It depends how the views that informed that scenario compared to the overall shape of responses – and how you then weight the distribution of views above and below the consensus value. That's what's reflected in the 'likelihood' column on the scenario table (we prefer to use the term 'likelihood' because, in radically uncertain situations, people debate the use of probability).

However, the important thing is to ask: "Do I want to ignore this scenario or take it into account, independent of its likelihood?" If I ignore it, then I am taking an out-and-out bet that something similar will not happen. If I decide to take it into account, then I will need to ask the question: "What is the hedge?"

How are the upper and lower bounds set?

Ron Dembo: We're looking for the worst-case scenario, so we take the extreme values in that range as the upper and lower bounds. Zero is the consensus value – either side of that, you find the range of impacts. When we look at both tail distributions – and if we're looking for the worst- and best-case outcomes – my contention is we should choose the extremes: we choose the lower bound from the lower part of the range, and vice versa for the upper part. However, these extremes must be feasible at the very least.

We've now got a distribution that captures future uncertainty. All that interests us at this point is an upward or downward move – Brexit happening, or not, for instance. We're only interested in a binary outcome. Now, what value do we give to the upward move at a node in the binomial tree? I need four numbers at each node: I need the value up or down, and I need the probability of each.

The upper and lower bounds are ultimately found by making sure we do not include any values that are not possible.

Ron Dembo, riskthinking.AI

Usually, when constructing stress scenarios, a lot of emphasis would be placed on the correlations between factors. Why does your approach not consider correlations in this way?

Ron Dembo: When people stress-test, they usually make certain correlation assumptions. In the real world, when stresses occur, correlations go out the window. Unless you've got a way of stressing covariance matrices, which I'd like to see, I don't know how you can use correlations in a stress test. Correlations are also inherently backward-looking. What we're trying to build is a forward-looking measure – going back and working out correlation is pretty much the antithesis of what we're trying to do here.

What the approach offers instead is implied correlations. When you look at the scenario tree, it implies a correlation matrix. We might want to assume that, for instance, if the US 10-year rate moves, it will affect the S&P 500. So, you can use conditional probabilities on the branches – in other words, what's the conditional probability of one usually moving down when the other moves up?

What number of respondents is considered statistically significant in order to get the desired distributions?

Ron Dembo: Much work has been done on this in the field of expert elicitation. In many cases, when constructing scenarios now, people use samples as low as 40 respondents. The challenge is to get good estimates of the extremes.

Doesn't a human input constitute a human bias by definition, contradicting the intent of minimising such biases?

Ron Dembo: Human input is definitely human bias – but that bias is minimised by looking at a very broad range of opinions on single factors. We can't eliminate human bias, but we can minimise it by combining factors in a given scenario using our algorithm.

How do you differentiate between the kind of extreme views you need to capture in order for this to work and someone trying to game the outcome?

Ron Dembo: It's an art, weighing up whether someone is trolling us, or for real. Sometimes, when you look at the responses, you see 100% changes for a given factor. Sometimes, we take those into account – because they're not crazy. But when you look at distributions and see a 50% change in a currency – when did we last see that in the euro in a few months? It couldn't happen. But it is an indication of the wild extremes people expect. You can then weight the responses any way you like, but the reality is there's a real person who said this – and in the case of the poll we conducted with *Risk.net* readers, they probably do work for a bank.

We use expert judgement to remove genuine outliers. As an example of this, some respondents predicted a value of zero for the S&P 500. Clearly, this is nonsensical, so we eliminated it. Things are not always so black and white, but a close look at the extremes can eliminate impossible outcomes, or outcomes that lack any theoretical basis. Clearly, this can introduce some bias if it is not done carefully. But, certainly, elimination of events that have never happened before in recorded history will not necessarily be a good criterion for elimination. ■

Previously published on Risk.net

Funds turn to stress-testing in fast-forward and reverse

The Covid-19 pandemic is changing the way investors think about stress tests, according to a *Risk.net* survey. By Robert Mackenzie Smith

In February and March, a public health crisis that began in the Chinese city of Wuhan spilled over into the global financial markets, wiping out more than three years of stock market gains in a matter of weeks. As investors dashed for cash, cracks began to appear in the typically safe and liquid market for US Treasuries. Banks with fortress balance sheets saw their commercial paper trade down 10–15 points. It was an extreme event, by any standards. But Morgan Stanley Investment Management (MSIM) had a stress test for that.

The suite of worst-case scenarios the firm routinely runs across its funds includes a sped-up version of the financial crisis that began in 2007–08.

"Those who ran the 2008 crisis scenario as an instantaneous or at least a compressed shock would have had a forecast that came reasonably close to what actually occurred," says Tatiana Segal, chief risk officer (CRO) at MSIM.

"So, we did have a gauge of the potential impact of our portfolios, and I expect that is true for a lot of people."

The magnitude of the market moves in March "was largely in the same ballpark" as the financial crisis in 2008, she says. What surprised investors was the speed at which it unfolded. "The velocity of this particular crisis was very different from what we observed in 2008 – both the speed of the selloff and the speed of the recovery," says Segal.

For many investors, this sort of creative stress-testing was critical to surviving the Covid-19 crisis. In a recent *Risk.net* survey of buy-side firms, 87% of respondents said stress tests using hypothetical or custom scenarios were very or somewhat useful in helping them navigate the selloff in March. Historical stress tests, which replicate the effect of past stress events on current positions, were also deemed useful by 79% of survey respondents. By contrast, only 67% said the same about value-at-risk, the most commonly used statistical risk measure at buy-side firms.

Sudi Mariappa, global head of portfolio risk management at Pimco, suggests why stress-testing is important: "In most scenarios, such as unstressed ones, volatility measures can be the most useful. But when looking at fat-tail events, the only way you can get an idea of the magnitude of that fatter tail is really by doing stress tests."

Still, designing an effective stress test is no easy task. On this, risk managers are drawing some lessons from the Covid-19 crisis.

Andrew Chin, CRO and head of quantitative research at AllianceBernstein, says risk managers need to be more creative when constructing doomsday scenarios: "I think unfortunately we all lack imagination."

According to the *Risk.net* survey, nearly 40% of buy-side firms rely primarily on the expert judgement of in-house risk managers to generate scenarios for stress tests, while a similar number lean on a mix of risk and portfolio managers. Fewer than 15% use third-party scenario generation tools.

The problem with this is that insiders can be inward-looking and overly reliant on past experience when assessing what could go wrong. "We have to be more thoughtful about what a black swan event looks like. And that's very hard because we're coloured by our experiences," says Chin.

Look past the past

Dan Bradley, CRO at hedge fund Mariner Investing Group, agrees. He urges risk managers to look beyond past events and traditional financial indicators when constructing future scenarios.

While there were some similarities between the previous financial crisis and this one, there were also plenty of differences – starting with the cause of the turmoil. The 2008 crash was the result of failures and excesses within the financial system. Covid-19 was a truly exogenous shock.

"This was different in that it really came out of the blue," says Bradley. "It shocked the system, but it had nothing to do with the system. It was a health crisis. It wasn't part of the financial system, but it totally disrupted the system. That's one of the things that will change the way people think going forward."

One consequence is that other non-financial sources of risk – such as climate change, geopolitics and technology – may now feature more prominently in the hypothetical scenarios used in stress-testing.

Another lesson from the Covid-19 crisis is to not be overly specific when constructing scenarios. Accurately predicting the precise cause and arc of a crisis is nearly impossible.

"The Covid-19 crisis is a case study of an event no-one would have been able to model as a scenario prior to the events unfolding," says MSIM's Segal. "Thinking back to swine flu scenarios, for example, they seem just hopelessly mild in comparison."

She suggests focusing on the effects, not the cause, of a crisis when constructing scenarios: "Don't necessarily start with thinking about the cause of the downturn – think about a maximum magnitude of an adverse shock and the resulting impact on your portfolios."

MSIM runs two types of hypothetical stress tests. There are tests using conventional scenarios that start with a plausible trigger. Then, there are reverse stress tests, which consider the worst-case adverse moves that could cause its funds to breach risk limits or run against liquidity thresholds. These help the firm "identify potential vulnerabilities without necessarily rooting ourselves in causality", says Segal: "This approach helps to hedge against the failure of imagination and construct a worst-case scenario that will meet the goal of a reasonably comprehensive downside capture."

Mariappa has a similar philosophy. Pimco runs a variety of stress tests using historical data and custom scenarios sketched out by its risk and portfolio managers. But he sometimes finds the most useful exercise is simply to "shock" a portfolio to see what would happen if a particular sector fell by a certain amount.

"There's a level of precision that I fear when we try to over-specify custom scenarios," says Mariappa. "I like broader shocks to get a sense of what is 'cuspy' [and] has fatter tails."

This sort of exercise can help risk managers identify potential correlation breaks and get a sense of how assets might behave in relation to each other during periods of stress, he adds.

The third and perhaps most critical lesson from the Covid-19 pandemic and ensuing market turmoil is that market conditions – and especially liquidity – can deteriorate much faster than many investors previously assumed.

"This all happened in a pretty compressed period of time. And I think that was surprising for a lot of people," says Chris Edge, head of investment risk at T. Rowe Price.

Future stress tests will need to account for the velocity of modern markets. And they may also have to be run more frequently to be useful. According to the *Risk.net* survey, only around 20% of firms currently run stress tests on a daily basis. More than 50% of respondents said their firms ran hypothetical stress tests monthly or quarterly, while around 18% run them annually.

The results surprised some risk managers. "I was kind of shocked that the frequency was anything other than daily," says Mariner's Bradley. "Portfolios are dynamic, markets are always moving, and inputs are always changing. I would expect people to be running these daily to see how they're changing and evolving." The input to stress scenarios should be reviewed on at least a monthly basis, he adds.

T. Rowe's Edge says running hypothetical stress tests annually makes little sense. "If you're only doing it annually, I'm not really sure what the benefit of that is, as those results might not be all that meaningful any more." After all, the Covid-19 disease did not even exist in March 2019. ■

Previously published on Risk.net

Banks aim to close op risk stress test capital gap

Standardising stress drivers could help smooth the differences between bank loss estimates. By Steve Marlin

A group of banks say a project that could help reduce the gap between different lenders' estimates of operational risk losses under stress is starting to bear fruit. The estimates play a major role in determining how much capital banks must set aside against their exposures under regulatory stress-testing programmes.

The group is working on a project led by the American Bankers Association to standardise the way in which risk drivers are set for a variety of op risks, including the impact of a major cyber attack or losses stemming from rogue trading incidents.

The aim is to create a set of baseline structured scenarios that can then be applied by each bank to their own particular business mix – something that those involved hope will aid comparability between loss estimates under the US Federal Reserve's annual Comprehensive Capital Analysis and Review (CCAR) programme.

"This type of modelling moves us closer to the objective of CCAR, which is to compare how different banks perform under macroeconomic stress," says an op risk executive at one of the participating banks. "If we do it the right way, this will aid comparability."

To date, the consortium has created two dozen scenarios covering a range of threats such as mis-selling, customer data compromise, critical service disruption and supplier failure. The next step is to make its scenario analysis tool available to individual participants, which can then customise the drivers and scenario structures for their individual banks.

An online portal will house the downloadable scenarios and benchmarking data. Those involved say the data can be used by participants to better understand their risk profiles compared with their peers, and to challenge existing models, or input from subject matter experts.

One of the problems with using stress-testing to gauge op risk exposures is that losses stem from both macroeconomic variables and idiosyncratic factors specific to a bank, such as weak cyber defences or trading errors. Under CCAR, the Fed defines macroeconomic indicators such as GDP and unemployment, but it's left to each bank to define its own idiosyncratic drivers of risk.

"The focus of the Fed over the last two or three years has been [on] forcing the banks to understand their risks on a forward-looking basis, and having [a] good risk identification process. They harp on everyone to do that on op risk. Now, we have good inventories of the most material op risks," says the op risk executive.

Banks have long been frustrated by the Fed's approach to modelling op risk under CCAR, which they claim has resulted in the Fed consistently overshooting the banks' own loss forecasts. This is due to the unique way the Fed models op risk, as it relies on observations from the data it collects from all banks that are subject to CCAR, giving it a panoramic view that individual banks don't have.

Because op risk losses tend to be idiosyncratic, losses are not necessarily correlated to the macroeconomic environment, as tends to be the case with market and credit risk. Determining causality between the frequency and severity of events and actual losses is a common problem facing banks – something the banks involved hope the ABA project could help address.

"In CCAR, there's a desire to understand the truth. Any technology that can help us understand risk exposure we will explore," says an op risk executive at a second participating bank.

Under CCAR, banks are required to produce loss estimates for four sets of scenarios: the Fed-defined baseline and severely adverse scenarios; and bank holding company (BHC)-defined baseline and severely adverse scenarios.

The BHC scenarios are expected to reflect macroeconomic environments that are particularly stressful to the institution based on its geographical footprint, business mix and unique vulnerabilities. The idea is that the Fed generates a set of scenarios that reflect its broad systemic concerns – but these scenarios might not stress certain variables that a specific bank would have a particular exposure to, because of its unique characteristics.

These idiosyncratic scenarios represent a large proportion of the total op risk forecast – and some CCAR banks are known to feel aggrieved that the conservatism (or otherwise) of a lender, when putting a dollar value on the same set of exposures, has a direct impact on the capital they are expected to hold against them.

Banks fear the problem could become even more vexing with the advent of the stress capital buffer, which means a bank could see its required capital increase even if post-stress results are higher than the regulatory minimum.

The Fed produces loss estimates only on its own scenarios – it does not run its models on the BHC scenarios. The Fed-produced number is the average between the equivalent of a 1-in-60-year downturn of an empirical bootstrap of an individual bank's data, and a regression model that uses data from all CCAR banks. The watchdog does not include BHC scenarios because those would be too firm-specific, going against the Fed's desire to have an objective and consistent approach across all banks.

"The Fed estimates are not purely macro-driven, nor purely random tail-driven – they are somewhere in the middle – but with some significant tail events coming through the empirical bootstrap," says Evan Sekeris, who previously oversaw op risk at the Fed.

It is not the Fed's goal to have banks' own estimates match its own, however. Because the BHC scenarios are tailored to each bank's profile, they result in a more stressful outcome than the one observed under the equivalent regulatory scenarios, which are the ones the Fed runs its models against to produce the final CCAR results.

When it comes to op risk, the differences tend to be very significant across banks, because lenders are expected to include idiosyncratic scenarios in their BHC submissions, but not in their regulatory submissions.

Unlike the BHC submissions, the regulatory submissions do not permit the use of idiosyncratic scenarios, but instead require banks to estimate losses based solely on the Fed's baseline and severely adverse scenarios.

"Given how the numbers are calculated by the Fed on the one hand and the banks on the other, it is difficult to directly compare them. The bank estimates for the regulatory submission are missing a big item: the idiosyncratic scenario. It is fairer to compare the Fed-produced numbers to the BHC projections produced by the banks," says Sekeris. ■

Previously published on Risk.net

Dutch banks seek quantum edge for stress tests

A trio of Netherlands-headquartered banks – ABN, ING and Rabobank – are working together to explore quantum technology for stress-testing, while US quantum developer Zapata Computing is seeking a patent to apply its tech to the Comprehensive Capital Analysis and Review. By Luke Clancy

Researchers on both sides of the Atlantic – including a trio of Dutch banks – believe quantum computing can help banks produce quicker, more accurate results in regulatory stress tests.

In the Netherlands, ABN Amro, ING and Rabobank are jointly exploring quantum technology, with stress-testing one of the possible use cases. US-based quantum developer Zapata Computing, meanwhile, is seeking a patent that applies the technology to the Federal Reserve's twin-track domestic testing regime: the Comprehensive Capital Analysis and Review (CCAR) and the Dodd-Frank Act Stress Tests (DFAST).

"If you talk to chief information officers of banks in the US, one of the major demands on high-performance computing – and on the computing budget – comes from CCAR and DFAST," says Christophe Savoie, chief executive at Zapata Computing.

He claims quantum machines will be able to deliver test results "much more quickly and much more accurately".

The Dutch banks have the same conviction.

"Currently, it takes a long time to run the required number of Monte Carlo simulations. The more we can do, the better our predictions will be," says Dimitri van Esch, quantum lead at ABN Amro.

Other banks are exploring the use of quantum technology for a range of tasks, all requiring a vast amount of conventional computing power – from derivatives valuation and options pricing to settlement problems and fraud detection. A technical paper published by *Risk.net* last year showed portfolio optimisation could be carried out up to 1,000 times faster when using quantum machines.

Bit of a problem

The key to the technology's leap in performance is the quantum bit, or qubit. While a bit within a classical computer can be in one of two states – storing either a one or a zero – a qubit can exist in both states simultaneously, a phenomenon known as superposition. This means quantum computers have the capacity to run more calculations at once, which could be a big help in stress-testing.

Testing regimes vary between nations, but typically require banks to take a prescribed set of shifts in macroeconomic and market factors, and then estimate the impact on their existing portfolios. Estimations are often generated by running large numbers of time-consuming Monte Carlo simulations.

As an illustration of the computational lift, the European Union-wide stress-testing exercise carried out in 2018 required banks to complete 38 different spreadsheets. One of these spreadsheets – covering forecast credit risk losses – had more than 9,000 rows and 78 columns.

It will be some years before this kind of work can be handed over to quantum technology in its entirety, experts believe. Quantum computers cannot be bought today. The only machines that can be rented for research – from manufacturers such as IBM – are an embryonic form of the technology, known as noisy intermediate-scale quantum, or Nisq. Each contains just a few dozen qubits – as more are added, the machines become less stable and their error rate increases.

Despite these problems, researchers believe stress-testing is one of the applications that is attainable in the relatively near term.

"The speed-up depends on the number of qubits, but at least you can run it with a limited number of qubits," says ABN Amro's van Esch.

Savoie claims Zapata has found a way to get help from even today's immature quantum machines.

"We figured out how to do [CCAR] on a Nisq computer," says Savoie. "And that's significant because if you can do that, even a constant factor speed-up will significantly reduce the time that's spent doing those simulations and the money that's spent on compute."

The key is to hand over some of the underlying components of the testing to Nisq machines, he adds: "Some of the subroutines can have a quantum equivalent that can do these highly complex multivariate analyses quicker and with greater accuracy."

"Four or five years"

Banks may have to confront more than technological obstacles if they want to use quantum computers to perform stress-test calculations. The results of the exercises are used to set capital requirements and – in the US – to sign off on dividend payouts and share buybacks. The models used to generate the test results are also heavily scrutinised, and regulators may not be comfortable if the modelling is carried out on brand-new, possibly wobbly technology.

The Dutch banks have already briefed domestic regulators on their plans, in an attempt to address possible objections before the project reaches fruition.

ABN Amro's van Esch says: "We believed it would be helpful for them to be connected to this project and know what we're going to do, so they can prepare for any kind of law change or regulatory change that might be needed. Otherwise, it might take us four or five years to build an algorithm, and then we'll have a delay for another four or five years because a law needs to be changed."

At Zapata, Savoie says regulators ought to take a close look if a new form of stress simulation had been designed specifically for a quantum computer: "If there were a new quantum algorithm or technique, we would want for that to be subjected to some regulatory testing."

However, if the same calculation was just speeded up by handing off the subroutines, then it might not require the same type of scrutiny, he argues: "You're doing the same math to get to your answer." ■

Previously published on Risk.net

JP Morgan shook up market risk stress tests in Q2

JP Morgan switched stressed value at risk historical periods 60 times. By Louie Woodall

Top US dealers have to use a 12-month period of significant financial stress to calculate part of their market risk capital requirements. Over the second quarter, JP Morgan changed the starting date of its chosen stress period 60 times, having not moved it once over any of the preceding 21 quarters, regulatory filings show.

It is understood this reflects how the bank used the same lookback period for stressed value-at-risk (SVAR) as it did for the VAR-based measure over the quarter.

The SVAR capital charge, one of the six that make up the market risk requirement for advanced approaches banks, is calculated by running a firm's trading portfolio through a VAR model populated with market data from an appropriate stress period. Each bank uses a different period that fits the unique makeup of their books.

As of the end of June, JP Morgan's stress period had the starting date April 5, 2019, meaning it encompassed the Covid-19 market shock of February and March this year. Three months prior it was set to June 10, 2008. Banks are permitted to dynamically switch their internal stress periods as their portfolios change, but none of the eight global systemically important banks (G-Sibs) had changed theirs more than 11 times in one quarter before. JP Morgan did not respond to a request for comment by press time.

In the second quarter of this year, Bank of America changed the starting date of its stress period 21 times, Goldman Sachs eight times, and Citi and BNY Mellon one time each. State Street, Wells Fargo and Morgan Stanley did not alter their stress periods. Each of these G-Sibs ended the quarter using a stress period straddling 2007–09.

Aggregate SVAR charges across the eight G-Sibs increased 11% over the three months to the end of June. Citi's increased the most percentage-wise, by 37% to $1.1 billion. JP Morgan's increased less than 3% to $2.1 billion.

What is it?

Banks subject to US federal agencies' market risk rule must disclose their capital requirements in quarterly Federal Financial Institutions Examination Council 102 reports. Each bank's market risk capital requirement includes both a VAR and an SVAR component: the former based on a VAR model calibrated to a 99% confidence level over a 10-day holding period and a 12-month observation period; and the latter using the same model fed with historical data from a 12-month period of significant financial stress. The stressed VAR-based measure must be calculated at least weekly.

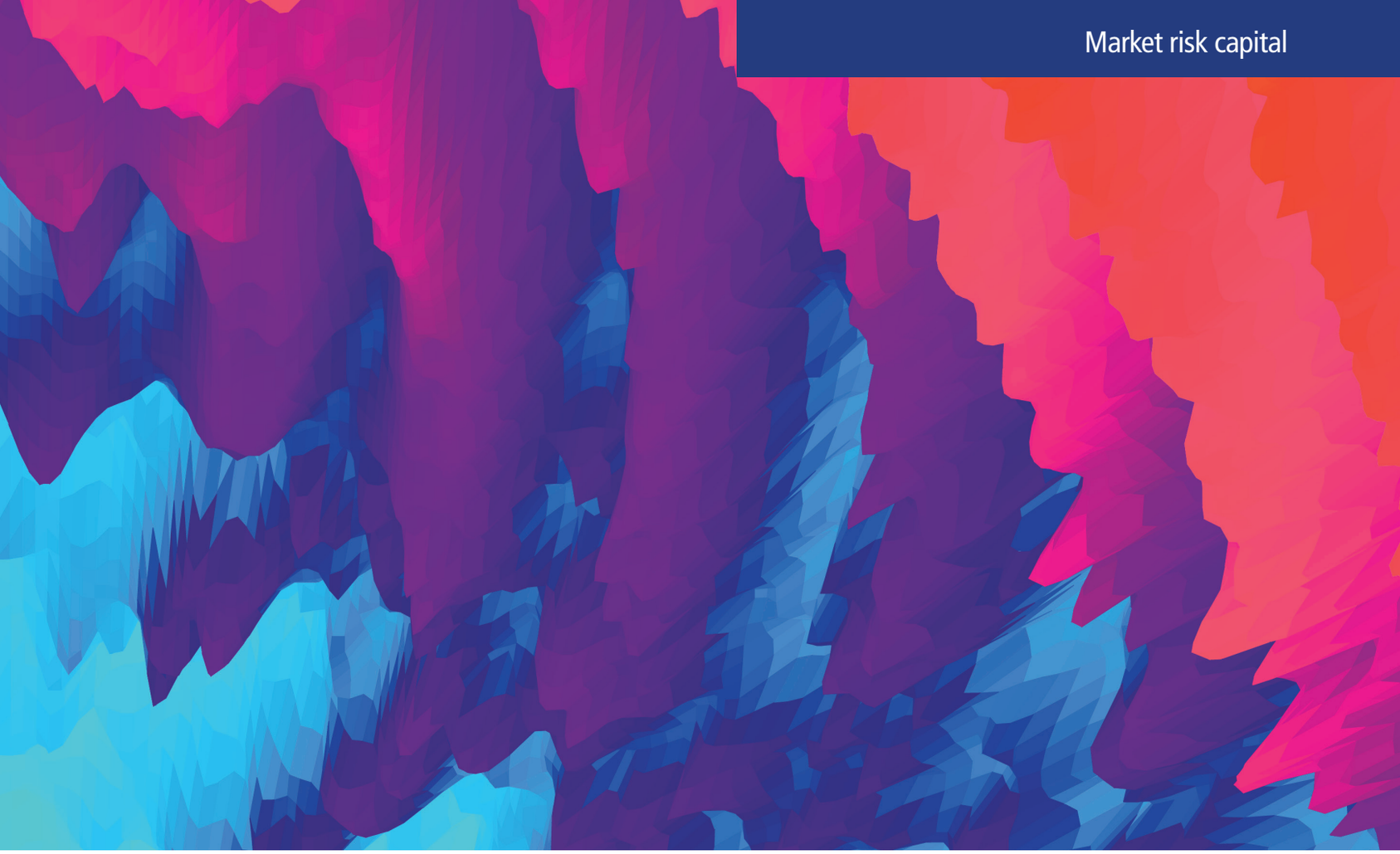
The rules also state that a bank must have policies and procedures that explain why a chosen stress period is appropriate and describe the process for "selecting, reviewing and updating the period of significant financial stress".

Why it matters

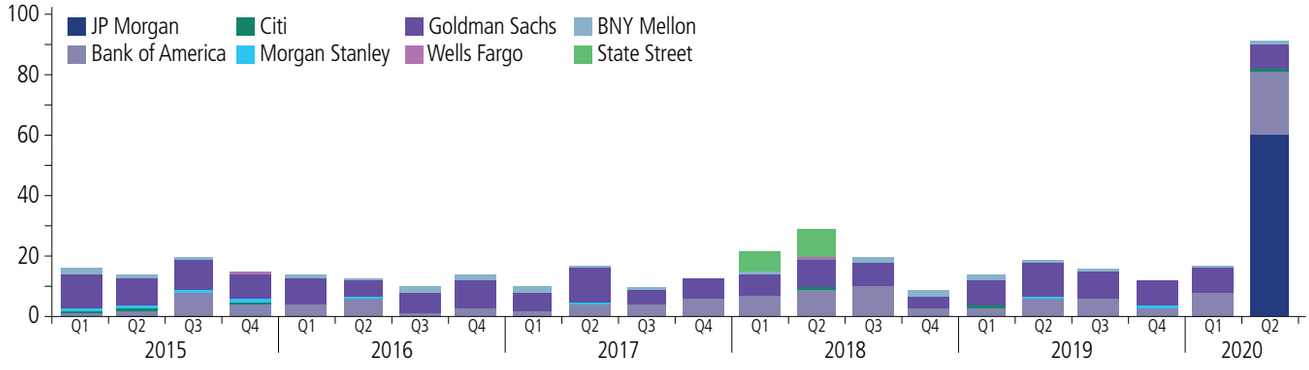
The Covid-19-induced market panic led to unprecedented market moves over Q2, and huge turnover in top banks' trading portfolios, as dealers rushed to fulfil client orders. Both factors could cause a bank to change the stress period used to set its SVAR. Banks don't disclose the starting date of each stress period used over a three-month period, only the one they end the quarter with, but it's plausible that other banks beside JP Morgan used periods encompassing the Covid shock at some point over Q2.

It is still noteworthy that only JP Morgan ended the period using an observation window referencing 2020, implying the end-June composition of the other banks' portfolios would face a tougher time under a redo of the global financial crisis than another bout of Covid-19-inspired volatility. ■

Previously published on Risk.net

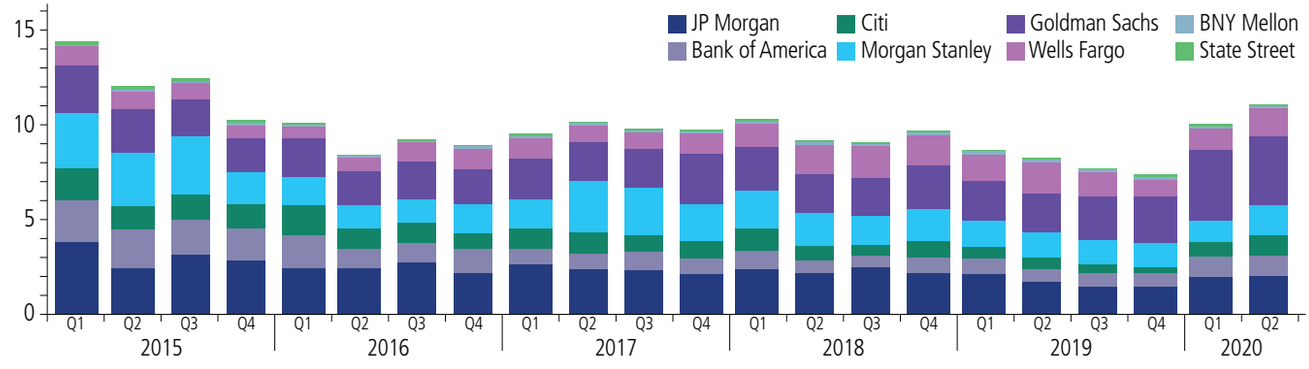


1 Number of changes to stress period starting date for the preceding 12 weeks



Source: Federal Financial Institutions Examination Council 102 filings

2 Stressed VAR charges of US G-Sibs ($ billions)



Source: Federal Financial Institutions Examination Council 102 filings

Bank investors still don't think bail-in will happen

Experts warn questions surrounding bailing in bank bondholders means the problem of being 'too big to fail' persists. By Philip Alexander

Investors in bank debt still assume governments would bail out a large bank rather than bail in bondholders, a global watchdog has been warned by analysts and former regulators.

"The market thinks you will still blink when it comes to it," said Paul Tucker, chair of the Systemic Risk Council, which focuses on the US and Europe. He was addressing a virtual workshop on September 4 organised by the Financial Stability Board (FSB).

Tucker, a former deputy governor for financial stability at the Bank of England, pointed to an FSB report in June, saying it had found a widening in spreads on the debt issued by global systemically important banks (G-Sibs).[1] But he added: "My instinct is that the magnitude of those spreads is still small."

The FSB report notes that, in the wake of too-big-to-fail reforms, bail-inable debt now yields more than "otherwise similar debt instruments to which it is subordinated". "This suggests that investors are at least partially pricing in the risk of G-Sib failure and a potential bail-in," the document reads.

Tucker warned that the "scale and rapidity" of market interventions by central banks in March this year might strengthen the conviction among investors that officials would ultimately baulk at allowing a G-Sib to fail.

Speaking at the same event, co-founder and head of banks strategy at Autonomous Research, Stuart Graham, said he was struck by end-investors' relatively limited understanding of total loss-absorbing capacity (TLAC), the new class of senior debt designed by the FSB to be easily bailed in during a bank resolution.

"I would estimate – gut feeling – probably only 25% of them understand TLAC and how it works. Most of them have a very hazy knowledge of it, and very few of them think regulators would actually allow a major financial institution to fail," Graham said.

That belief had been strengthened by the "contortions" of Italian or German authorities in recent years to save banks that are not even anywhere near the G-Sib category, but are deemed significant to specific regions or sectors, he added. Italy provided aid for the takeover of two banks in the Venetian region in 2017, on the basis that they were regionally systemic, even after the European Union's Single Resolution Board had ruled that they were not systemic. Germany intervened in 2019 to rescue NordLB – an important player in ship finance but not a G-Sib.

Tucker said his impression was that analysts and investors were not looking into the many nuances of resolution strategies, suggesting they did not think bail-in was a real possibility.

One versus many

Alberto Gallo, head of macro strategies at Algebris Investments, which manages around €12 billion ($14 billion) in assets, mostly bank capital instruments, said he thought individual banks were now easier to resolve. But he warned that "too-big-to-fail has not gone away", noting that banking systems across Europe were several times larger than national GDP.

"Banks have more capital, so it looks like you have reduced the risk from a single-institution point of view," Gallo said.

"But overall, at a top-down level, when looking at the banking system size as a whole, the problem is still there – it is very hard to believe that, in a systemic crisis, you would be able to bail in many banks in the same country," he continued, implying that imposing losses on a large number of bondholders at the same time would be too damaging to the economy.

Even at the level of individual banks, there are still "very serious problems" with bail-in, said Martin Hellwig, director of the Max Planck Institute for Research on Collective Goods and a former chair of the advisory scientific committee at the European Systemic Risk Board. Specifically, he said, it was still unclear where extra liquidity would come from for a bank in resolution, to keep it operating, once an existing EU emergency pot ran dry – a question raised publicly by the Single Resolution Board in 2018.

What's more, single-institution and systemic crises are not two distinct problems but "circles of hell", said Tucker of the Systemic Risk Council. If Bear Stearns had been bailed in early in 2008, he said, "the incentives on Lehman to take outside capital during the summer and for some others…to delever their repo and derivatives books would have been greater", potentially defusing the wider systemic crisis that unfolded in September that year.

Politics versus rules

Hellwig also suggested that the resolution of a G-Sib would encounter political difficulties: "You have political resistance which comes in anytime there are vested interests inside the political system trying to make sure that the bank remains there. What I do not see is any big change in stance on the side of the political authorities – resolution is not just a technical problem, but a political one."

This view was echoed by Thierry Philipponnat, head of research at public interest lobby group Finance Watch, who warned: "If you want someone to take an unpopular decision – and bail-in is likely to be an unpopular decision – do not ask politicians to do it."

In Italy in particular, retail investors' holdings of bank debt have made bail-in a highly controversial topic.

There is a way for regulators to take the politics out of decisions on bank recapitalisations, including through bail-in: by making greater use of stress-testing, argued Stephen Cecchetti, a finance professor at Brandeis International Business School who previously worked at the Bank for International Settlements. Stress tests could be used during a crisis to either reassure markets or show that a bank needs extra equity, he said.

"How can we use private funds to recapitalise an institution in a manner that ensures its ability to serve as a source of credit to healthy non-financial firms? I believe stress tests are an important part of the answer," he said.

The US Federal Reserve faced criticism earlier this year when its stress test add-on based on coronavirus scenarios was not used to determine whether banks should raise extra capital.

"My hope is that one morning, I am going to open my email alert from the FSB and there's going to be an announcement of the publication of key attributes of effective stress tests and automatic recapitalisation regimes – then we will know that too-big-to-fail is further in the past," Cecchetti said. ■

Previously published on Risk.net

[1] *FSB (June 2020)*, Evaluation of the effects of too-big-to-fail reforms – Consultation report, *https://bit.ly/3nLlnFs*