

Quagliariello, Mario

Working Paper

Are stress tests beauty contests? (and what we can do about it)

EBA Staff Paper Series, No. 4

Provided in Cooperation with:

European Banking Authority (EBA), Paris La Défense

Suggested Citation: Quagliariello, Mario (2019) : Are stress tests beauty contests? (and what we can do about it), EBA Staff Paper Series, No. 4, ISBN 978-92-9245-641-2, European Banking Authority (EBA), Paris La Défense,
<https://doi.org/10.2853/29499>

This Version is available at:

<https://hdl.handle.net/10419/299509>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

N. 4 – JULY/2019

ARE STRESS TESTS BEAUTY CONTESTS?

(AND WHAT WE CAN DO ABOUT IT)

by Mario Quagliariello

ABSTRACT

There is increasing debate on whether official-sector stress tests are fit for purpose and deliver what they are designed for. In the EU, this discussion focuses on the costs and benefits of such a large-scale bottom-up exercise and on the possible options to improve its usefulness for banks, supervisors and the general public, but also on exploring ways to make it less burdensome.

This paper discusses whether some improvements to the design of the EU-wide stress test could better align the incentives of the different stakeholders involved. In particular, the objective is to understand whether there is room for mitigating the beauty contest problem, namely the problem that banks are more interested in showing that they are outperforming others than in identifying actual risks in an adverse scenario.

KEYWORDS

Stress test; bank; EBA; supervision; beauty contest; incentives.

... professional investment may be likened to those newspaper competitions in which the competitors have to pick out the six prettiest faces from a hundred photographs, the prize being awarded to the competitor whose choice most nearly corresponds to the average preferences of the competitors as a whole; so that each competitor has to pick, not those faces which he himself finds prettiest, but those which he thinks likeliest to catch the fancy of the other competitors, all of whom are looking at the problem from the same point of view. It is not a case of choosing those which, to the best of one's judgment, are really the prettiest, nor even those which average opinion genuinely thinks the prettiest. We have reached the third degree where we devote our intelligences to anticipating what average opinion expects the average opinion to be.

J. M. Keynes, The General Theory of Employment, Interest, and Money

Introduction

This paper deals with supervisory bottom-up stress tests — specifically the European Union (EU)-wide stress-testing exercise — and discusses whether some improvements to the design could better align the incentives of the different stakeholders involved. In particular, the objective is to understand whether there is room for mitigating the beauty contest problem, namely the problem that banks are more interested in showing that they are outperforming others than in identifying actual risks in an adverse scenario.

Stress tests have been applied in the banking industry since the early 1990s, in response to regulatory requirements. In the EU, in line with the Capital Requirements Directive (CRD) and Capital Requirements Regulation (CRR), banks use stress testing to assess their ability to absorb losses and maintain sufficient capital levels in adverse conditions.

Beyond institutions' risk management, stress tests play an important role as a component of the toolkit available to supervisors and macroprudential authorities. The International Monetary Fund (IMF) relies on stress tests as part of its Financial Sector Assessment Program; likewise, stress tests were used during the great financial crisis as a tool for quantifying the recapitalisation needs of the banking sector in both the US and the EU. In normal times, central banks and supervisory authorities make extensive use of these techniques to identify weaknesses in the financial system and possible threats to financial stability¹.

¹ Schuermann (2016).

Irrespective of the institution that actually runs the stress test, the main goals remain broadly the same, that is, to assess the resilience of the banking sector to extreme but plausible shocks and to make decisions to prevent such shocks or mitigate their impact.

In the EU, as the 2008 financial turmoil evolved into a sovereign crisis, public authorities deployed stress tests with the objective of dissipating uncertainty on the magnitude and distribution of losses across banks. Increased market scrutiny and public expectations led to a call for greater transparency as regards individual banks' exposures, to address concerns that bad news was being hidden². Transparency was considered a way to restore the confidence of market participants in the EU banking sector, following the successful example of the US.

In 2011, the then newly established European Banking Authority (EBA) was tasked with coordinating the stress tests, supplying the methodology, working with the European Systemic Risk Board (the EU macroprudential authority) to provide a common scenario and publishing the results in a coordinated fashion. Crucially, the EBA did not have — and still does not have — enforcement powers to assure the quality of the results of individual banks; this is instead the responsibility of the relevant supervisors (national competent authorities and, since 2014, the European Central Bank-Banking Supervision for the euro area).

In the first EU-wide stress test, the EBA was responsible for communication on all of the steps of the exercise, from the announcement to the publication of the results. The methodology, the scenario and detailed bank-by-bank outcomes were published and widely disseminated. Since then, granular disclosure of banks' results and of underlying exposures has been a key feature of the European exercise. This has allowed market participants to make their own assessments of banks' solvency and to compare and contrast banks' results.

Stress testing has undoubtedly been a key tool in managing the crisis and it is now part of the supervisory toolkit at the global level. There is, however, increasing debate on whether or not official-sector stress tests are fit for purpose and deliver what they are designed for. In the EU, this discussion focuses on the costs and benefits of such a large-scale exercise and on the possible options to improve the usefulness of the exercise for banks, supervisors and the general public, but also on exploring ways to make it less burdensome.

Objectives of the EU-wide stress test

Over time, the EU-wide stress test has served many purposes, such as improvement of transparency, crisis management, identification of vulnerabilities and benchmarking. More importantly, it has been consistently used for determining banks' capital needs — either immediate recapitalisation or medium-term supervisory expectations — and to disseminate detailed information on banks' exposures to various sources of risk. The methodology has always aimed to strike a balance

² Haben and Quagliariello (2015).

between realism and conservatism, taking into account banks' business models, while ensuring consistent treatment across banks in the EU and comparability in the results. This is in line with other system-wide exercises, but the EU setting required additional effort to ensure a level playing field across institutions originating from different jurisdictions that were subject to accounting rules and supervisory practices that were not fully harmonised.

During the crisis, stress tests were used to identify the capital needs of weak institutions, with subsequent recapitalisation efforts to restore their resilience. These exercises were, thus, pass/fail: banks that were not able to meet a predefined capital threshold 'failed' the test and had to be recapitalised immediately. Strictly speaking, the capital thresholds identified were not a legal requirement, but banks failing to meet these thresholds were expected to agree with their supervisors the appropriate remedial measures and their implementation within a given timeline.

In 2016, the EBA moved away from the pass/fail setting. This was a time when capital ratios had improved and it was no longer deemed necessary to focus on whether or not banks needed immediate recapitalisation. The stress test was instead used by supervisors as an input to the supervisory review and evaluation process and for setting capital expectations (Pillar 2 guidance). This required a recalibration of the framework to ensure that it primarily informed supervisors about potential vulnerabilities in individual banks, challenged banks' own risk management and provided some insight into aggregate systemic risk. The objective was also to help market participants understand the sensitivity of banks to hypothetical adverse market developments and to quantify the possible depletion of capital under such scenarios.

Notwithstanding criticisms and limitations, the EU-wide stress test has facilitated the consistent assessment of banks in the Single Market, highlighting risks that may be growing in isolated pockets of vulnerability but, if crystallised, might generate spill-over effects in the wider banking system. Particularly in the light of the significant heterogeneity in the EU banking sector, the exercise — which used a single scenario, was conducted simultaneously in the entire EU and used a consistent methodology — provided a unique snapshot of the soundness of EU banks. The EU-wide scope also gave credibility to stress testing at a time when there were concerns that national supervisors might shy away from addressing important risks. The announcement of the stress tests also led to a considerable degree of pre-emptive action. Overall, the EBA exercises have contributed to the significant strengthening of the capital positions of European banks, identifying poor risk management practices and recognising non-performing loans and actions to reduce them.

While these achievements are undeniable, the question remains whether the EU-wide stress test could be made more incentive-compatible for participating banks. However, it is equally important to assess what kind of reforms could instead facilitate opportunistic behaviour.

Incentives in stress tests

In any supervisory stress test, banks have an incentive to minimise the impact of the assumed shock, as it can affect — more or less directly — the scale and intensity of follow-up actions. The way in which the stress test framework is designed may strengthen such incentives.

An important starting point here is to specify the role that banks are expected to play in the exercise. In a top-down stress test run directly by the supervisors, there is little room for banks to influence the outcome. In a bottom-up exercise, banks use their own internal models and, thus, may deliver results that are either not conservative enough or too lenient³. If the outcome of the stress test directly affects the level of capital, the incentive to underestimate the effect of the adverse scenario could be more pronounced. Finally, if the results are publicly disclosed — triggering additional market pressure on weaker banks — there is a further incentive for banks to provide the most reassuring estimates of the impact.

This is the beauty contest problem: banks want to look ‘the prettiest’, by trying to anticipate what the prettiest is in the opinion of supervisors and market participants.

The EU-wide stress test is a bottom-up exercise; it affects supervisory decisions on banks’ capital and its outcome is communicated very transparently to the public. Without safeguards, the incentive structure is such that the objective for banks could shift from identifying actual and potential risks to simply looking good to their supervisors and stakeholders.

The ambition of looking good is not bad *per se* and can have positive consequences: it may give banks a genuine reason to perform well in a stress test and therefore prepare accordingly. *Ex-ante* preparation for a stress test and adjustments in balance sheets are not necessarily an attempt to game the rules. Pre-emptive actions, such as capital raised by banks ahead of the 2011 EU-wide stress test, are a prudent and safe way to face the exercise with a stronger starting point⁴. Disposal of lower quality or non-core assets as part of longer term business planning are also actions that show that stress tests can play a role in banks’ risk management.

On the other hand, some actions taken ahead of the stress test could purposefully attempt to misrepresent risks or exposures, particularly around the reference date. Some research argues, for instance, that detailed *ex-post* disclosure might adversely affect the *ex-ante* incentives of bank managers and lead them to take myopic actions to pass the test. Because of these incentives, banks may pick sub-optimal portfolios that increase their probability of passing the stress tests but that reduce the fundamental value of the bank. If the supervisory stress test model or methodology is known and stable over time, this behaviour could become more pernicious, as banks could find it even easier to pass the exercise⁵.

³ In a somewhat negative assessment of central bank stress testing, Dowd (2015) claims that ‘the reason is that banks want bad models because they understate their risks, and the regulatory system endorses bad models because it is captured by the banks. Most risk modelling is then just a game: banks pretend to model risks, but they are really gaming the risk numbers’.

⁴ EBA (2011).

⁵ Goldstein and Sapra (2014).

The potential incentives for window dressing and circumventing the purpose of the stress test have been part of the consideration in the design of the EU-wide exercise. Therefore, since the first stress test in 2011, the effort to mitigate beauty contest incentives has been a top priority and contributed to several methodological elements that are still used today. In particular, the methodology developed by the EBA aimed to constrain what banks can and cannot do using internal models. Backstops in the form of caps to profits and floors to losses in the adverse scenario have become a distinctive element of the exercises. The decision to rely on such a ‘constrained bottom-up’ approach tried to strike a balance between a willingness to encourage risk management by banks and a determination to ensure sufficient conservatism of the results and a level playing field across banks.

Methodological constraints are the first line of defence in a bottom-up exercise. They have been gradually upgraded to ensure that overly optimistic internal estimates are prevented. Not surprisingly, these constraints have also been criticised for being too strict, conservative and not particularly realistic. Clearly, banks would be better off if the methodology was unconstrained, since the constraints are designed to reduce the degrees of freedom and to impose some conservatism on their assumptions.

However, the question remains whether or not such methodological safeguards have been effective in preventing beauty contest behaviour, and whether or not they have generated some unintended consequences. The point is not that the constraints may lead to excessive conservatism in the estimation of impacts, but rather that — being calibrated on historical aggregate data — they may not reflect the true risk sensitivity of a bank, resulting in the outcome of the exercise being of little use for internal risk management and planning.

The asymmetric pass-through of the impact of the stress on the net interest income — a constraint designed to prevent banks from fully passing to borrowers the increase in interest rates, while at the same time fully internalising the loss from the increase in the interest paid — is a good example of what banks tend to consider too conservative and as not correctly representing the evolution of the cost of funding under stress. Of course, there are strong arguments in favour of this assumption, not least the fact that banks would claim that — in times of crisis — they are all able to access funding and to charge higher rates to their customers. Nevertheless, this does not imply that there is no room for thinking of better solutions.

The static balance sheet assumption — a constraint that implies the freezing of banks’ balance sheets at the reference date so that no managerial actions are allowed — is the application in stress testing of the *ceteris paribus* assumption. As the IMF pointed out as part of its 2013 EU Financial Sector Assessment Program⁶, the use of a static balance sheet over a 3-year period may be justified mainly on the grounds of tractability and the desire to facilitate comparability across banks. However, the IMF also notes that, because it facilitates comparability, the assumption may also lead investors to interpret the results as a beauty contest. Indeed, even after the EBA decided to remove any references to the pass/fail capital threshold, several observers continued to focus on banks’ final capital positions rather than on the overall process.

⁶ IMF (2013).

The quality assurance process is an additional and important tool for limiting beauty contest incentives and for ensuring credible outcomes. In this process, competent authorities assess the credibility of banks' assumptions, models and results and try to identify strategic behaviours and attempts to circumvent the methodology. A bottom-up stress test could not work without an intense review of banks' submissions, and the outcome of this process must be measurable. In 2018, for instance, according to the results submitted by banks in the first round of the exercise, the impact of the adverse scenario was, on average, 100 basis points lower than in the results following the quality assurance.

However, the authorities responsible for the stress test are often operating within a short time frame when carrying out their assessment. At the same time, the models used by banks are not validated by supervisors. This means that the quality assurance process needs some standardisation and automatic tools to prioritise work and focus on the more material aspects. This stage is followed by tailor-made discussions with the banks, but at some point supervisors may have to resort to shortcuts and to apply conservative assumptions or benchmarks if they are not convinced that what banks have provided is reliable and if they remain uncomfortable with banks' explanations. While this process is an ongoing dialogue, the bias of supervisors towards conservatism is understandable and may come at the cost of reduced risk sensitivity. In that respect, the obsession of the media and the analyst community with the headline figures in terms of capital ratios and capital depletion is not helpful, because it may reinforce the perception that some specific numbers matter more than the entire process of risk identification.

From the banks' point of view, since quality assurance tends to focus on outliers, it may be optimal to deliver results that would not attract too much attention, that is, not too pessimistic, as investors can become concerned and penalise underperformers, but also not so optimistic as to trigger closer supervisory scrutiny and difficult questions. The incentive here is not minimisation of the impact. However, showing that own figures sit neatly around the median of the distribution — or perhaps slightly above — is a comfortable position to be in in this setting.

For this reason, the EBA has been always very cautious in disclosing information — even at the aggregate level — on preliminary stress test outcomes during the quality assurance process. Descriptive statistics have been distributed to the competent authorities to support their own assessments, but have never been disseminated beyond the supervisory community. Lately, however, private-sector initiatives have allowed banks to pool their anonymised results ahead of submission to the authorities and receive, in return, information on how they perform compared with the other participating institutions. These tools could have beneficial effects in terms of the quality of banks' submissions, thus making the entire quality assurance smoother. However, they could also be used for wrong reasons, thus providing a further boost to flawed incentives.

Finally, full transparency is the last line of defence against the circumvention of the methodology. A detailed disclosure of risk exposures, the composition of capital and profitability is a way to increase market discipline and allows investors and analysts to understand the risk drivers for each institution and to undertake their own assessment of the scenario, assumptions and outcomes⁷. Disclosure of the

⁷ Ong and Pazarbasioglu (2013).

stress test results improves the overall trust in the banking sector by alleviating concerns about regulatory forbearance in concealing bad news from the market⁸.

Notwithstanding these safeguards, the EU-wide stress test is still perceived by many banks as a beauty contest. According to a recent survey⁹, 4 out of 10 senior bank managers consider the stress test relevant, either as a means of fostering competition or as an industry-wide transparency exercise. At the same time, 6 out of 10 respondents feel that the published results do not fully reflect their bank's resilience.

In a nutshell, incentives are such that, when results are published, banks want to look good and public authorities want to look tough. The identification of risks and gaining an understanding of banks' sensitivity to shocks — the main goals of the exercise — may thus become a side story.

Some considerations

The EU-wide stress tests have evolved over time. The current framework is the result of progressive adjustments to the methodology, improvements in the process and refinements in the interaction between banks and supervisors. This progress has taken place under the limits imposed by the current regulation and the checks and balances embedded in the EBA's governance structure. Still, the overall framework is not yet completely satisfactory. In particular, the question remains whether it is possible to better align the incentives to avoid beauty contest behaviour, as this feature continues to dominate the public perception of the exercise. Providing the right incentives for banks would also allow some of the constraints in the methodology to be relaxed, making the exercise more useful and possibly less operationally burdensome.

In this section, I do not aspire to provide a final answer; instead, I list a few aspects that may deserve some further consideration. The different options presented below are not neutral in terms of costs, either for banks or for supervisors, and the final choice would depend on the preferred burden-sharing scenario if the changes are deemed to be desirable.

As discussed in the previous section, while the EU-wide exercise serves many purposes, it can essentially be considered a supervisory stress test for the purpose of setting capital targets. It also serves as a disclosure and benchmarking exercise. These features reflect the nature and objectives of the stress test, but they also affect the way that the different players choose to behave.

The first aspect to consider is the link between the stress test and the quantification of the capital needs. If the main incentive for banks is to minimise the capital impact, decoupling the stress test and Pillar 2 guidance could be an option. Admittedly, there is no need for the EU-wide stress test to act as a supervisory exercise, but unbundling the two concepts would not necessarily address the issue. Any

⁸ Goldstein and Sapra (2014).

⁹ Oliver Wyman (2018).

supervisory stress test would have to deal with similar problems, as the very purpose of the stress test is to facilitate evidence-based supervisory decisions, including on banks' capital¹⁰. In addition, the EU framework — although it introduces a mechanistic tie between capital depletion in the stress test and Pillar 2 guidance — remains flexible, and judgement plays a significant role in supervisory decisions (in terms of the assessment of credible mitigating actions beyond capital, the possible offsetting effect of the macroprudential buffers, etc.).

If, instead, it is considered that public disclosure deters banks from providing an accurate picture of their risk exposures, one possibility would be to reduce the level of transparency of individual results and, in the extreme, to discontinue the publication of bank-by-bank results. The value added of the granular disclosure and its contribution to market discipline are important in the EU-wide stress test. However, some are of the view that, while transparency is critical in periods of turmoil, it may create trade-offs in normal times. Schuermann (2016), for instance, argues that supervisors ought to be very transparent and have a rich disclosure in a crisis. Nonetheless, in normal times, a more limited disclosure may be desirable, provided that banks report comprehensively to their own stakeholders and, in this way, promote monitoring and market discipline.

In accordance with the EBA's founding regulation, transparency of individual results is a legal requirement for the EU-wide stress test and it is part of the EBA's established practice. Indeed, transparency is among the few elements of the EU stress-testing framework that are generally praised. Analysts, market participants and banks themselves do value bank-by-bank disclosures and benefit from the large-scale benchmarking exercise. While this principle has never been questioned, transparency can be delivered in different ways.

One option is to reduce the precision of bank-by-bank results. For instance, instead of publishing point estimates of the capital position in the adverse scenario, banks could be grouped in buckets depending on the capital depletion or the post-stress capital ratios¹¹. Alternatively, transparency could be limited to the starting points (i.e. with no disclosure of stress test results), but with very generous disclosure of supervisory decisions and follow-up actions, including Pillar 2 requirements and guidance.

The debate is open. In the EU, where bank data is scarce compared, for instance, with the US, the response to the negative externalities could involve more, rather than less, disclosure. Providing even more granular information on underlying exposures, as suggested by Goldstein and Sapra (2014), could help market participants to interpret the results and limit the risk that banks can game the system.

Improved transparency could also take the form of changes to the methodological framework. A simple improvement — but one not without cost — would be to run the stress test on multiple scenarios. This would address the criticism that reliance on a single scenario makes the methodology easier to circumvent and does not necessarily allow risks to be identified across banks with different business models¹². Supervisors could therefore design alternative scenarios or ask banks to run their own scenarios, ensuring that at least one of them is able to catch the bank's specific vulnerabilities. It would also be possible to require banks to run — as already envisaged in EU regulation — and publish the outcomes of reverse stress testing, in which banks identify possible shocks that could endanger

¹⁰ Basel Committee for Banking Supervision (2018).

¹¹ de Guindos (2018).

¹² Dowd (2015); Schuermann (2016).

their survival. This could be more challenging to communicate and explain to the public, but it would increase transparency as regards what events might be particularly detrimental for each bank and what the banks could do about it.

Transparency — or lack thereof — can also be considered from a different and more provocative angle. If, as suggested by recent research¹³, a stable and transparent methodology allows banks to adjust their models over time to minimise the impact of the stress, a simple option would be to make the methodology unstable, unpredictable and possibly opaque. This is, to some extent, the rationale for not disclosing the details of supervisory top-down models in the US, although this approach has received criticism and is currently under review¹⁴. It is also important to consider this possibility in the light of the accountability faced by supervisory authorities, particularly when they need to exercise judgement¹⁵. Still, it can be argued that at least parts of the methodology should be made less predictable or that key details — such as the scenarios — should be disclosed only at short notice to reduce the opportunities for window dressing¹⁶.

Turning to the methodological aspects, the question is whether the static balance sheet assumption could be relaxed. The rationale of this assumption is to enhance the comparability of the results and to make the entire exercise manageable from a quality assurance perspective. However, it has been argued that, even if excluding managerial mitigating actions is important in a crisis to facilitate comparisons, some dynamic reaction could be allowed in stress tests in normal times, provided it remains limited in scope, clearly framed and explained to the external stakeholders¹⁷.

Moving to a stress test in which behavioural responses are permitted would make the exercise more realistic. It would also force banks to think through their internal crisis response strategies. Even though the incentive to look good would probably remain, banks would be more inclined to provide an accurate identification and quantification of the risks, to the extent that they demonstrate that they have a credible plan in place in case the risks materialise. On the other hand, a dynamic balance sheet approach should be cautiously designed and guidance should be provided to ensure, using the IMF's words, that banks avoid 'strategies that rely on *deus ex machina* (such as the sale of an unprofitable business at a handsome price)'¹⁸.

The inclusion of banks' reaction to an adverse shock would be a significant departure from the approach followed by the EBA to date, but it is not the only possible — or the most important — way to review the methodology. A more fundamental change would be to progress towards a pure bottom-up approach in which banks can use internal models without any (or with very limited) *ex-ante* constraints, while being subject to more intense *ex-post* scrutiny. Banks could, for instance, be required to submit the outcomes of the internal models, with the constraints kicking in only at a later stage as part of the supervisory quality assurance process. Depending on the confidence one has on

¹³ Niepmann and Stebunovs (2018).

¹⁴ Quarles (2018).

¹⁵ Enria (2019).

¹⁶ Bolton et al (2019).

¹⁷ Ong and Pazarbasioglu (2013).

¹⁸ IMF (2013).

the effectiveness of market discipline, both unconstrained and constrained results could be published, leaving the onus on banks to explain any differences.

Such a change could be implemented with additional precautions. As not all banks are actually able to model the impact of the adverse scenario on all risks and asset classes, internal models could be allowed only for the more sophisticated banks. This would also require the bar of eligible stress test models to be raised. A more constrained, or even standardised, approach would be applied for the rest of the banks, which would conversely benefit from a streamlined quality assurance process and basic data requirements. It is also possible to combine top-down and bottom-up approaches by using supervisory models for relatively standardised asset classes, where sensitivities are not expected to change significantly across institutions, and letting banks project the impacts for specific business lines where they can really provide added value.

Considering institutions' risk management further, banks could run the stress test as a part of — and with the very same toolbox as — the Internal Capital Adequacy Assessment Process (ICAAP), simulating the impact of their own scenarios as well as of the common supervisory scenario. Banks would be expected to use the outcome of these tests for internal purposes, proving that the figures they provide to supervisors are so reliable to pass a “use test”. The cost in terms of lower comparability would be offset by reduced incentives to bend the rules, since this would not only affect the supervisory process but also the internal stress-testing process.

Where to set the bar for allowing the use of internal models would be left to the supervisors. A set of large and complex banks able to carry out the bottom-up stress test could be identified, while the remaining banks would be subject to a simplified top-down exercise. Provided ICAAP is of good quality and reliable, the benefit of this approach is that banks would run a useful, realistic and tailor-made stress test.

Top-down models, benchmarking with peers and simple sensitivity tests would still be used by supervisors as tools for challenging banks' own estimates. Their outcomes could even be published to allow investors and analysts to make their own judgement. Disagreement between a bank and its supervisor on a specific outcome would not necessarily imply that the bottom-up results are overridden. However, this could trigger follow-up supervisory actions in the form, for instance, of additional off-site assessments or on-site visits. This would not be equivalent to *ex-ante* validation, but would represent a possible way to review, *ex post*, stress test models if the supervisor is not satisfied with the bottom-up results but the time available for the quality assurance process is not sufficient to make a fully informed decision¹⁹.

The advantage would be that the stress tests are better integrated with the regular supervisory cycles and, in a repeated game, this would create a clear incentive for banks to improve models, be conservative and submit credible results. However, the actual feasibility of this approach depends very much on the resources available to supervisors for running the additional post-publication review.

¹⁹ This could also take the form of the Bank of England's (2015) qualitative review of stress-testing practices. Another option, although possibly the topic for a future paper, is to operationalise the pre-commitment approach suggested more than 20 years ago by the Federal Reserve for setting capital requirements for market risk. Under this approach, banks could quantify their own requirement with the understanding that they will face penalties should losses exceed the pre-committed capital (Kupiec and O'Brien, 1997).

Conclusions

There is a growing debate on the future of stress testing, both in the US and in the EU. Attention has been paid to fundamental governance aspects, but also to relatively minor methodological details. What is missing is a comprehensive and candid discussion on the incentive compatibility of the different frameworks.

Top-down stress tests ensure that banks do not game the methodology, but are criticised for being opaque and for discouraging the development of risk management at financial institutions. Bottom-up stress tests are prone to beauty contest behaviour, particularly if their outcomes are directly linked to supervisory actions and publicly disclosed.

In the EU-wide stress test, the constrained bottom-up approach, severe quality assurance processes and transparency of bank exposures have been instrumental in mitigating such flawed incentives. The constraints have helped in delivering comparable and conservative results, but there are concerns that this may have come with the cost of reduced risk sensitivity.

A key issue is that the governance of the exercise and its methodological constraints have been designed to limit the impact of the existing problems of incentive incompatibility, rather than to construct an incentive-compatible exercise. Solving this dilemma is not easy and any options come with various trade-offs.

It is quite possible that a fully incentive-compatible framework cannot be achieved in the context of bottom-up stress testing and that a shift towards a top-down approach is the best available option. An alternative way forward is to move cautiously towards a wider acceptance of banks' internal practices and, at the same time, to integrate more of the follow-up actions in the supervisory processes.

Still, even assuming that a consensus could be reached on the direction of travel, it would be naive to expect any sudden shift in practices. Instead, it might be more realistic to expect a gradual evolution towards the preferred framework.

References

- Bank of England (2015), *"Bank of England's approach to stress testing the UK banking system"*.
- Basel Committee for Banking Supervision (2018), *"Stress testing principles"*.
- Bolton Patrick, Stephen Cecchetti, Jean-Pierre Danthine and Xavier Vives (2019), *"Sound at Last? Assessing a Decade of Financial Regulation"*, CEPR.
- de Guindos Luis (2018), *"The ECB and Financial Stability"*, speech at the joint National Bank of Belgium/Toulouse School of Economics/Solvay Brussels School of Economics and Management/European Central Bank colloquium, Brussels, November.
- Dowd Kevin (2015), *"Central Bank Stress Tests: Mad, Bad, and Dangerous"*, Cato Journal, Volume 35, No. 3.
- Enria Andrea (2018), *"What we have learnt from EU-wide Stress Tests"*, speech delivered at the National Bank of Romania, November.
- Enria Andrea (2019), *"Supervising banks – Principles and priorities"*, speech at the SSM & EBF Boardroom dialogue, Frankfurt am Main, March.
- European Banking Authority (2011), *"2011 EU-wide stress test: Aggregate Report"*, July.
- Goldstein Itay and Haresh Sapra, (2014), *"Should Banks' Stress Test Results be Disclosed? An Analysis of the Costs and Benefits"*, Foundations and Trends(R) in Finance, 8, issue 1.
- Haben Piers and Mario Quagliariello (2015), *"EU-wide Stress Tests: Rationale and Basics"*, in M. Quagliariello (ed.), *Europe's New Supervisory Toolkit: Data, Benchmarking and Stress Testing for Banks and their Regulators*, RiskBooks.
- IMF (2013), *"European Union: Publication of Financial Assessment Programme Documentation"*, Technical Note on Stress Testing of Banks.
- Kupiec Paul H. and James M. O'Brien (1997), *"The pre-commitment approach: using incentives to set market risk capital requirements"*, Finance and Economics Discussion Series, 14, Board of Governors of the Federal Reserve System.
- Niepmann Friederike and Viktors Stebunovs (2018), *"Modeling Your Stress Away"*, FRB International Finance Discussion Paper, No. 1232.
- Oliver Wyman (2018), *"Supervisory stress testing: evolving for higher value"*, presentation at EIFR, Paris, December.
- Ong Li Lian and Ceyla Pazarbasioglu (2013), *"Credibility and Crisis Stress Testing"*, IMF Working Paper, WP/13/178.
- Quarles Randal K. (2018), *"A New Chapter in Stress Testing"*, speech at the Brookings Institution, Washington DC, 9 November.
- Schuermann, Til (2016), *"Stress Testing in Wartime and in Peacetime"*, in Ronald W. Anderson (ed.), *Stress Testing and Macroprudential Regulation: A Trans-Atlantic Assessment*, CEPR eBook.

ACKNOWLEDGEMENTS

The author is affiliated with the European Banking Authority (EBA). The opinions expressed in this paper are those of the author and do not reflect those of the EBA or its Board of Supervisors.

This paper has benefited from discussions with several colleagues and different stress-test stakeholders. In particular, I want to thank the following people for their useful inputs and comments: Simone Casellina, Gaetano Chionsini, Andrea Enria, Angel Monzon, Carmelo Salleo, Til Schuermann, Patrick Van Roy and two anonymous referees. Any remaining errors are my own.

Mario Quagliariello

Director of Economic Analysis and Statistics Department, EBA

EUROPEAN BANKING AUTHORITY

Floor 27 - 20 Avenue André Prothin,
92927 Paris La Défense

Tel. +33 1 86 52 70 00
E-mail: info@eba.europa.eu

<http://www.eba.europa.eu>

ISBN 978-92-9245-641-2
ISSN 2599-7831

doi: 10.2853/29499
DZ-AH-19-002-EN-N

© European Banking Authority, 2019.

Any reproduction, publication and reprint in the form of a different publication, whether printed or produced electronically, in whole or in part, is permitted, provided the source is acknowledged. Where copyright vests in a third party, permission for reproduction must be sought directly from the copyright holder.

This paper exists in English only and it can be downloaded without charge from www.eba.europa.eu/about-us/staff-papers, where information on the EBA Staff Paper Series can also be found.