

## DISCUSSANT REMARKS

# **Comments on the paper by Sumit Agarwal, Bernardo Morais, Amit Seru and Kelly Shue entitled “Weighted Noise: Discretion in Regulation”**

**Contribution by Claudia Buch, Chair of the Supervisory Board of the ECB, at the tenth ECB Annual Research Conference organised jointly with Stanford University’s Hoover Institution on “The Next Financial Crisis?”**

*Frankfurt am Main, 18 September 2025*

## **Summary and findings of the paper**

The discussion on “rules versus discretion” has a long tradition in central banking.<sup>[1]</sup> Since a seminal paper by Kydland and Prescott in 1977 “Rules rather than discretion: the inconsistency of optimal plans”,<sup>[2]</sup> there has been an intensive debate over the years about which rules central banks should follow in setting their monetary policy rates, and how much discretion to allow for. Almost 50 years on from the original paper, the debate is still ongoing.<sup>[3]</sup>

Banking supervision has its own version of the debate on rules versus discretion.<sup>[4]</sup> Should supervisory assessments rely mainly on quantitative metrics or should they also incorporate qualitative judgement? How can decisions remain consistent across banks while still reflecting institution-specific circumstances? To what extent should supervisory processes be codified into rules or models, as opposed to depending on expert judgement? These debates have shaped supervisory practice over the past decades and remain highly relevant today.

The paper by Sumit Agarwal, Bernardo Morais, Amit Seru and Kelly Shue<sup>[5]</sup> makes a key contribution to this debate. It uses data for the United States to shed light on the supervisory decision-making process, analysing the determinants of CAMELS ratings.

The authors use bank-level data for the years 1998-2020 to answer the following questions:

- How does hard and soft information affect supervisory decisions? Hard information captures the observable characteristics of banks, which is used to derive an expected decision. The actual decision (an observable) differs from the expected decision, reflecting the use of soft information, supervisory discretion and statistical errors (both of which are unobservable).

- › What explains the use of judgement? Is it the result of subjective decision-making, different weights applied, or disagreement on the weights? To answer these questions, the authors focus on the fact that examiners rotate and may therefore assess the same bank differently.
- › How does judgement affect real outcomes?

The main findings of the paper by Agarwal et al. are:

- › Supervisors apply judgement. Changes in banks' ratings are often driven by changes in the supervisory assessment of banks, reflecting differences in the degree of conservatism.
- › Judgement has real implications. It affects the capitalisation and lending of banks, as banks may act in anticipation of judgement.
- › Supervisory judgement improves forecasts. It adds information to purely mechanistic, rules-based models.
- › Overall, there is a trade-off between rules and discretion. Discretion allows for the use of additional, soft information but introduces "noise" into supervisory decisions.

Overall, this is a very well-structured, thoroughly researched paper on a highly relevant topic. My comments focus on the drivers of discretion, implications for the evaluation of supervisory effectiveness, and implications for the use of supervisory judgement.

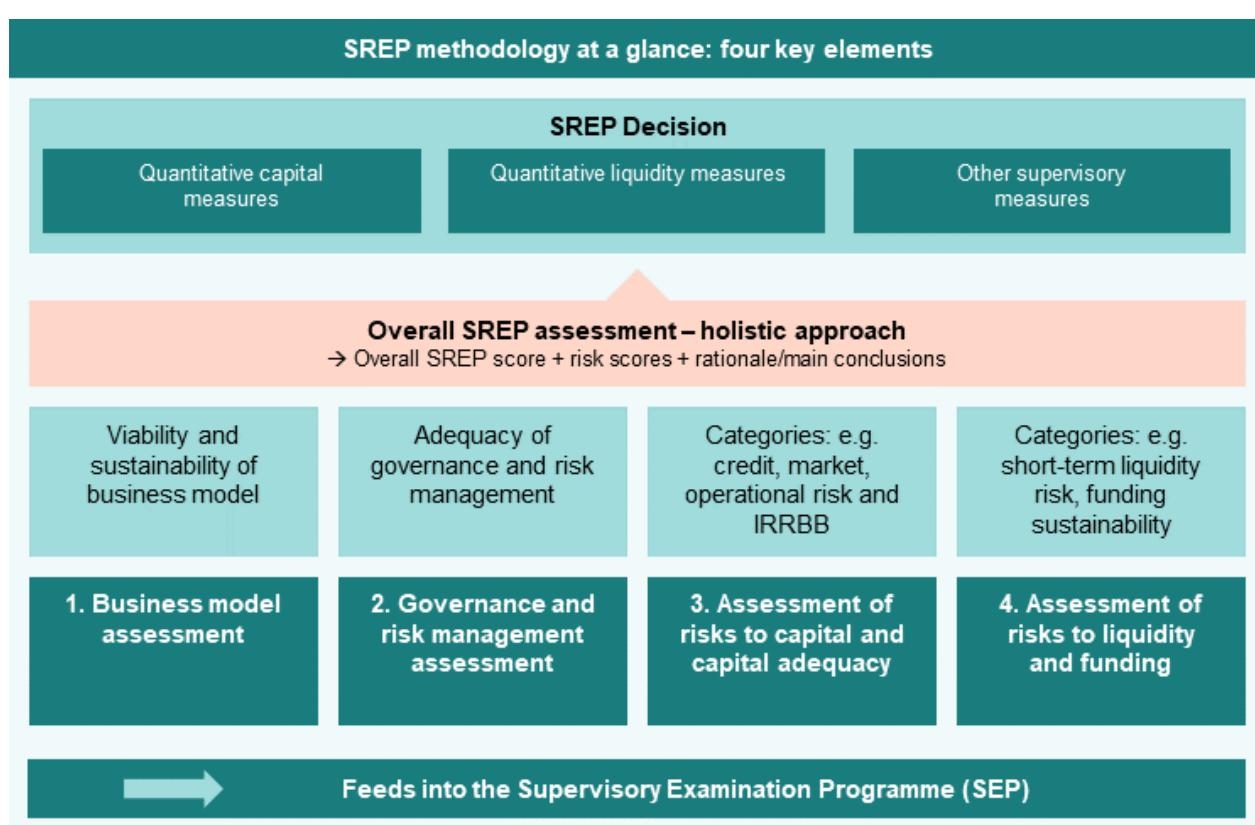
## Use of supervisory ratings in Europe

The discussion on rules versus discretion in supervisory decision-making is very relevant for European banking supervision.

In the United States, CAMELS ratings provide summary statistics on banks' health based on information on **Capital Adequacy, Asset Quality, Management, Earnings, Liquidity, and Sensitivity to Market Risk**.<sup>[6]</sup> Each component is assigned a score on a five-point scale, and supervisors assign a composite rating that reflects the bank's overall condition. CAMELS ratings inform policy decisions such as deposit insurance premia or access to the Federal Reserve's lender of last resort function.

In the European Union, supervisory processes are governed by a harmonised legal and methodological framework.<sup>[7]</sup> ECB Banking Supervision uses scores that are defined in the Supervisory Review and Evaluation Process (SREP scores). These scores capture the bank's business model, governance, risks to capital, and liquidity. The elements of the SREP are assessed on a four-point scale, with +/- qualifiers. These scores are then combined into an overall SREP score, which reflects the supervisor's view of the institution's viability.

Figure 1: SREP methodology



SREP scores include quantitative and qualitative elements. Judging the soundness of banks through capital and liquidity indicators alone would not do justice to the fact that weaknesses in governance and business models are often the root cause of bank distress.

In the SREP, supervisors employ a principle of constrained judgement by going through three phases (Figure 2). Phase 1 is about gathering information from bank reports. In Phase 2, an automated score is generated for the risk level. In Phase 3, supervisors adjust these anchored scores within a defined range.

Figure 2: The three phases of the SREP assessment

Phase 1 Data gathering	Phase 2 Automated anchoring score	Phase 3 Supervisory judgement
<p><b>Main sources:</b></p> <ul style="list-style-type: none"> <li>• EBA regulatory reporting</li> <li>• ECB supervisory reporting</li> <li>• Materiality</li> </ul>	Scoring risk level	Adjustments based on additional factors and considering banks' specificities and complexity

Notes: EBA regulatory reporting refers to the implementing technical standards on supervisory reporting, commonly known as COREP and FINREP<sup>[8]</sup> ECB supervisory reporting refers to the [ECB short-term exercise](#) to collect additional information in the context of the SREP.

The SREP scores form the basis for Pillar 2 capital requirements and they have binding prudential consequences. The SREP integrates forward-looking analysis, including stress tests, capital and liquidity planning, and the sustainability of business models over the medium term.

CAMELS ratings and SREP scores are therefore different institutional responses to the same underlying challenge, i.e. how to condense a wide body of quantitative and qualitative information into a holistic supervisory assessment, while preserving space for supervisory judgement.

## Comment 1 — What determines the choice between rules and discretion?

Any scoring system needs to strike a balance between consistency and comparability across banks and the need to take bank-specific information into account. How much to rely on mechanical rules and how much judgement to exercise is a difficult choice that also depends on the environment in which banks operate.

In a highly uncertain environment, purely discretionary decision-making may add to uncertainty and reduce the predictability of supervisory decisions. Overly constrained, rules-based decision-making may reduce the flexibility available to adjust decisions to take bank-specific circumstances into account.

Moreover, any rules-based decision-making process must assume that underlying fundamentals and economic relationships do not change much over time. However, banks are operating in a period of transformation, characterised by the digitalisation of financial services and structural change in the real economy. Applying the same rules over time may therefore lead to different and potentially unintended outcomes. For example, during the Covid pandemic and the energy crisis, fiscal policy helped to buffer the impact of these shocks on the real economy. When assessing the future resilience of the financial sector, it is therefore important to take such factors into account.<sup>[9]</sup>

Particularly in this environment, supervisory judgement is essential to provide a forward-looking risk assessment. Financial reporting and risk indicators are by their nature backward-looking, capturing what

has already happened. If supervision is to be genuinely risk-based and forward-looking – identifying vulnerabilities *before* they crystallise into problems for society – it must draw on qualitative insights and soft information that go beyond what is visible in the numbers.

The empirical setting in this paper relies on a long panel of bank-level data ranging from 1998 to 2020. This period covers several very different macro-financial environments, including the dot-com recession, the global financial crisis and the COVID-19 shock. It also spans a period of significant regulatory reform, including the introduction of Basel II or the post-crisis reform agenda. Hence, the estimates capture supervisory behaviour under both calm and stressed conditions. However, the paper does not explicitly model uncertainty or structural breaks by, for example, using sub-samples or regime-switching models that would make it possible to distinguish behaviour in periods of high and low uncertainty.

As a result, the findings should be interpreted as average effects across regimes. It would be interesting to explore whether the effects of supervisory judgement are more or less pronounced in periods of heightened uncertainty when forward-looking judgement could be of most value.

Moreover, it would be interesting if the authors could discuss the implications for supervisory transparency. Is more transparency on supervisory methodology good or bad, and what elements of the methodology should be published? Supervised institutions and the general public have a legitimate interest in understanding the framework supervisors use. Using judgement may introduce some element of “noise” into the supervisory process – different decision-makers assessing the same case may come to different conclusions whereas, in a purely rules-based system, automatic scores would be generated. European banking supervision responds to this need to be transparent and accountable by publishing its methodologies and guidance.

## **Comment 2 — How can the effects and effectiveness of supervision be evaluated?**

Currently, there is a very active policy debate on the benefits of post-crisis financial sector reforms. This debate relates not only to regulation but also to supervision.

To assess the effects and potential unintended side effects of the post-crisis policy agenda, evaluation frameworks are needed. As regards supervision, this means assessing whether supervisory judgement improves the trade-off between rules and discretion. Judgement may introduce bias into supervisory decisions (which would be an unintended outcome) but it also addresses the shortcomings of overly simplistic or overly complex rules that do not reflect the underlying risks (which would be an intended outcome).

The paper by Agarwal, Morais, Seru and Shue provides an excellent framework for such evaluations, which need to be supported by a sound infrastructure for evidence-based policymaking.<sup>[10]</sup>

Within European banking supervision, we are assessing the impact of the use of judgement on SREP outcomes. A forthcoming paper by Bobeică and Oprică shows that:<sup>[11]</sup>

- › First, supervisors do not simply take the automatic overall score – the equal-weighted average of the SREP elements – at face value. Instead, they frequently adjust it. These adjustments tend to smooth out fluctuations in the automatic score and ensure that the final assessment is more stable. Supervisors can thus bring in qualitative insights and emerging risks in a forward-looking way that may not be captured by the rules-based component;
- › Second, the adjustments are not random. Supervisors tend to respond in similar ways to common signals, such as shifts in the macroeconomic environment or sector-wide vulnerabilities. For example, supervisors respond to rising non-performing loans or risks related to commercial real estate to make sure that scores better reflect the implications for risk. This points to a “common supervisory judgement” channel: discretion is not idiosyncratic but reflects shared priorities. European banking supervision has indeed been setting supervisory priorities since its creation. In 2021, the process was revised with the aim of concentrating supervisory work on fewer priorities;
- › Third, certain risks are given particular weight in the use of supervisory judgement. The automatic score assumes that strong performance in one risk area can offset weaknesses in another. Yet, supervisors may exercise their judgement to adjust the score. They may, for example, consider credit risk to be decisive for a bank’s viability: if it increases, supervisors can adjust the overall score, even if other risk indicators look sound. Supervisory discretion can therefore capture non-linearities and interactions between risks.

Ultimately, the question of whether judgement “distorts” supervisory decisions requires a framework for assessing the welfare implications of supervisory outcomes. If judgement brings a decision more closely in line with societal preferences, then it is not distortionary. Theoretical research suggests that some “noise” in supervisory decisions may indeed lead to more careful lending behaviour and less risk taking.<sup>[12]</sup>

Agarwal et al. show that, possibly as a result of supervisory discretion, banks may behave in a more conservative way than they would in a system where ratings are set in a purely mechanistic way. From a welfare perspective, this might be a desirable outcome. Geopolitical risks, digitalisation and other structural trends change the risk environment in which banks operate in an unknown way. Basing supervisory decisions on observables and historic correlations alone would not take due account of these changes and may, in the end, lead to weaker resilience and higher risk taking.

### **Comment 3 — How should supervisory judgement and discretion be used?**

The results of this paper have concrete implications for how best to exercise supervisory judgement. Generally, scoring systems need to include rules-based components while allowing for the use of supervisory judgement. But there is a trade-off: relying on quantitative models does not allow judgement and expertise to be used. Excessive reliance on judgement may not sufficiently address cognitive biases.

The results of the paper by Agarwal et al. suggest several ways to improve the use of judgement in supervisory discretion. First, standardised guidelines for how examiners should weigh the components of

the rating systems can reduce variability and enhance consistency. Second, rather than eliminating discretion, the paper advocates modest constraints that preserve the ability to process soft information while mitigating the costs of excessive variability. Third, algorithms and machine learning tools can complement human judgement. These tools can provide initial assessments or identify patterns in the data, helping to reduce bias and improve consistency while retaining scope for human discretion.

These proposals resonate with the constrained judgement approach used in the SREP. In Europe, supervisors are equipped to incorporate soft information and forward-looking insights, but their decisions are supported by a unified methodology, benchmarking and second-line oversight. Supervisory outcomes proposed by Joint Supervisory Teams for each bank are benchmarked against those of peers, both within and across countries, to identify outliers and align assessments with a common methodology. This benchmarking is conducted by a second line of defence located in a directorate which is independent from day-to-day supervision. In this way, supervisory judgement remains anchored by shared standards, reducing unwarranted variability – a concern well recognised in the wider literature on judgement and decision-making<sup>[13]</sup> – while preserving flexibility to reflect bank-specific circumstances.

Since 2023 the SREP has been reformed with the introduction of a multi-year approach, allowing supervisors to conduct more focused risk assessments and drawing on synergies across different supervisory activities. This year, we have tested a pilot for a revised and simplified Pillar 2 requirement (P2R) methodology, which will be applied from 2026 onwards.<sup>[14]</sup> The new methodology builds on the existing framework by streamlining procedures and further clarifying how risks identified in the SREP feed into capital requirements.

Supervisory judgement is central to the new approach. It is applied in two ways: in scoring risks, setting capital requirements for specific risk areas; and in assessing overall risk profiles, particularly where risks interact or exceed the sum of their parts.

Figure 3: Current and revised methodologies for setting Pillar 2 requirements



Empowering supervisors to use judgement and apply the risk tolerance framework is an integral part of the new SREP. The reform has therefore been complemented by a dedicated strategy to change supervisory culture and train supervisors to use judgement, including to assess emerging risks. The use of judgement is not about relying less on analytics or methodological rigour. Instead, it is about drawing on all relevant information, in an era of uncertainty and disruption, to scan the horizon and capture developments that are insufficiently addressed by common rules and methodologies. This is particularly important as supervisory work increasingly relies on artificial intelligence and advanced analytical tools: human judgement remains essential to interpret outputs, assess context and ensure that supervision adapts to risks that cannot be reduced to data alone. To provide space for these analyses of emerging risks, we are also streamlining supervisory procedures.<sup>[15]</sup>

Overall, our objective is to make European supervision more efficient, effective, and risk based to maintain strong supervisory standards. In an environment characterised by a high degree of geopolitical uncertainty and structural change, we remain clearly focused on keeping the banking sector resilient, both operationally and financially. This requires the continued use of constrained judgement and a clear agenda for assessing the effectiveness of our work, to which we remain committed. The empirical methodologies developed in this paper and work of the Basel Committee on Banking Supervision on supervisory effectiveness<sup>[16]</sup> will therefore inform and inspire our work.

---

1.

I would like to thank Gabriel Bobeică, Emanuela Branca, Korbinian Ibel, Florian Narring, John Roche, Mario Quagliariello, Sebastian Scheffler, Dominik Vu and Veerle De Vuyst for their very helpful input and comments on an earlier version. All remaining errors and inaccuracies are my own.

2.

Kydland, F.E. and Prescott, E.C. (1977), "Rules Rather than Discretion: The Inconsistency of Optimal Plans", Vol. 85, No 3, *Journal of Political Economy*, pp. 473-492.

3.

Kocherlakota, N. (2016), "[Rules versus Discretion: A Reconsideration](#)", *Brookings Papers on Economic Activity*; Dellas, H. and Tavlas, G.S. (2021), "[On the Evolution of the Rules versus Discretion Debate](#)", *Economic Working Papers*, Hoover Institution, March.

4.

Maddaloni, A. and Scopelliti, A. (2019), "[Rules and discretion\(s\) in prudential regulation and supervision: evidence from EU banks in the run-up to the crisis](#)", *Working Paper Series*, No 2284, ECB, Frankfurt am Main, May.

5.

Discussion of paper "[Weighted Noise: Discretion in Regulation](#)" by Sumit Agarwal, Bernardo Morais, Amit Seru and Kelly Shue at the [Tenth ECB Annual Research Conference 2025](#).

6.

See "[Uniform Financial Institutions Rating System](#)" in the Federal Register, No 3, Vol. 6262, pp. 752, January 6, 1997, effective January 1, 1997.

7.

The main legal foundations are laid down in the [Capital Requirements Directive](#), the [Capital Requirements Regulation](#), and the [EBA Guidelines on common procedures and methodologies for the supervisory review and evaluation process \(SREP\) and supervisory stress testing \(EBA/GL/2022/03\)](#).

8.

Commission Implementing Regulation (EU) 2021/451 of 17 December 2020 laying down implementing technical standards for the application of Regulation (EU) No 575/2013 of the European Parliament and of the Council with regard to supervisory reporting of institutions and repealing Implementing Regulation (EU) No 680/2014 (OJ L 97, 19.3.2021, p. 1).

9.

Buch, C. (2025) "[Ten years of the banking union: laying the groundwork for the next decade](#)", Speech, "Finanzplatztag 2025" event organised by Börsen-Zeitung, 3 March.

10.

For a discussion of the components of such an infrastructure, see Buch, C. (2025), "[Financial stability, supervision and regulation: building a 21st century infrastructure for better, evidence-based policymaking](#)", speech at the BIS Innovation Summit, 10 September.

11.

Bobeică, G. and Oprică, S. (manuscript), "The role of judgement in supervisory scores and additional capital requirements assigned to banks", ECB.

12.

Repullo, R. (2024) "[Regulation, Supervision, and Bank Risk-Taking](#)", working paper, October

13.

Kahneman, D., Sibony, O. and Sunstein, C.R. (2021), *Noise: A Flaw in Human Judgment*.

14.

Buch, C. (2024), "[Reforming the SREP: an important milestone towards more efficient and effective supervision in a new risk environment](#)", *The Supervision Blog*, ECB, 28 May; and Buch, C. (2025),

["Reviewing the Pillar 2 requirement methodology"](#), *The Supervision Blog*, ECB, 11 March.

15.

Buch, C. (2025), "[Simplification without deregulation: European supervision, regulation and reporting in a changing environment](#)", speech at the Goldman Sachs European Financials Conference 2025, June;

Donnery, S. (2025), "[As simple as possible, but not simpler](#)", *The Supervision Blog*, ECB, 8 September.

16.

Badev, A. et al. (2025), "[Lessons on supervisory effectiveness – a literature review](#)", *Basel Committee on Banking Supervision Working Papers*, No 45, July.