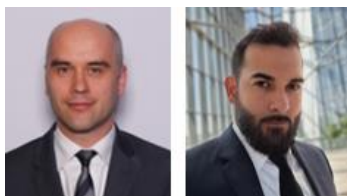


The role of judgement in supervisory scores and additional capital requirements assigned to banks



Gabriel Bobeică | European Central Bank (ECB)

Silviu Oprică | European Central Bank (ECB)

Keywords: Supervisory judgement, Supervisory Review and Evaluation Process (SREP), overall SREP score, Pillar 2 Capital Requirements (P2R), panel data

JEL codes: G21, G28, C23, E58

Abstract

Judgement is always present in the supervisory process. Yet, there are only few studies which analyse the extent to which it affects the supervisory outcomes.

Our empirical analysis shows that supervisors actively use judgement. They adjust the weights of the overall score components to reflect qualitative information, thereby smoothing fluctuations in the final assessment. This hints at a common supervisory judgement channel, which could reflect shared priorities and concerns, such as systemic vulnerabilities or macroeconomic conditions.

These findings underpin the critical role of judgement in adapting supervisory frameworks to evolving risks and systemic conditions, providing flexibility at both the bank and system-wide levels. The sizeable idiosyncratic component of supervisory judgement also confirms the importance of second line of defence supervisory functions in reducing unwarranted variability in supervisory outcomes.

Disclaimer: This policy brief is based on [ECB Working Paper Series, No 3233](#). This paper should not be reported as representing the views of the European Central Bank (ECB). The views expressed are those of the authors and do not necessarily reflect those of the ECB.

What is supervisory judgement?

Within the supervisory process, judgemental elements are always present.¹

To better understand what supervisory judgement entails, it is useful to note that supervisors often assess the soundness of banks through scores², which reflect the overall risks, and which may entail significant consequences for future bank activity through further supervisory actions. Among the more prominent examples are the CAMELS rating system, used in the US by supervisory agencies to provide a summary of bank conditions (Gaul and Jones, 2021), and the Overall Supervisory Review and Evaluation Process Score (OSS), which is used by the supervisors in the EU to summarise the risk profile of the bank from a viability perspective (European Banking Authority, 2022).

OSS is the numerical indicator of the overall Supervisory Review and Evaluation Process (SREP) assessment, which forms the basis for the supervisors to decide the most appropriate set of measures that would best address the overall situation of the institution and its specific weaknesses. OSS reflects the risk profile of the institution from a viability perspective. The viability of an institution is defined as its proximity to a point of non-viability on the basis of the adequacy of its own funds and liquidity resources, governance, controls, and/or business model or strategy to cover the risks to which it is or may be exposed (European Banking Authority 2022).

OSS is set starting from an automatic overall score (automatic OSS), which is computed as the simple average of the scores for the four elements assessed in SREP: business model, internal governance, capital adequacy, and liquidity adequacy. Each score provides an indication of the risk to the institution's viability stemming from the respective element. Capital and liquidity adequacy scores reflect considerations on the availability of capital and, respectively, liquidity and funding, including under stress conditions, and have an important forward-looking component. They also leverage on the assessment of more granular, individual risks, such as credit, market, operational, and interest rate risk in the banking book in the case of capital adequacy, and liquidity and funding risks in the case of liquidity adequacy.

Automatic OSS has its limitations and cannot always fully encapsulate a bank's viability. Firstly, it is based on a limited set of indicators. Secondly, by design, it equally weighs the four SREP elements. This approach implicitly assumes that they are perfect substitutes across all banks, and therefore that strong performance in one risk area can offset weaknesses in another. Furthermore, risks often interact and amplify one another, which is not captured by a linear aggregation mechanism. Supervisory judgement can address these shortcomings when supervisors adjust the OSS to reflect banks' specific characteristics, vulnerabilities and deficiencies more accurately.

Supervisors adjust the automatic OSS based on supervisory judgment³, incorporating their knowledge of the institution, peer comparisons, business model, macro environment, and other factors. They may also highlight weaknesses in the OSS that are considered particularly significant for the bank. While the automatic score can fluctuate from year to year depending on its components, the supervisory judgement part can offset or in some instances amplify this pattern.

A quantitative analysis of supervisory judgement

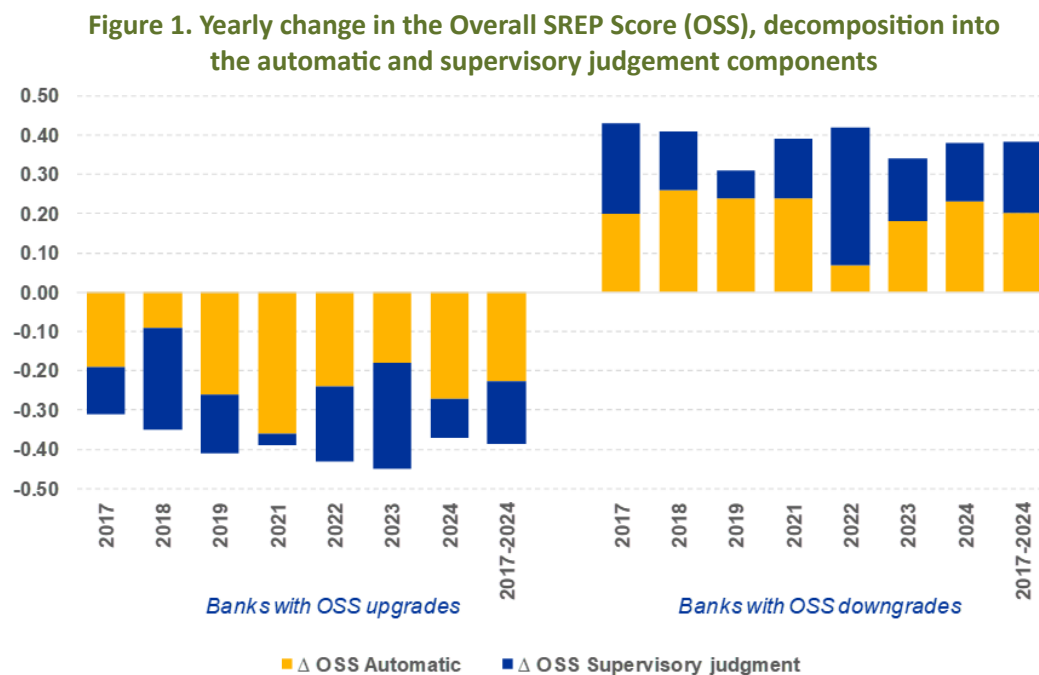
Leveraging on the supervisory dataset spanning nine evaluation cycles (SREP), from 2016 to 2024, we conduct the first thorough quantitative analysis of determinants driving changes in the overall supervisory score (OSS) and Pillar 2 Requirements (P2R) across significant banking institutions in the euro area.

¹ "The supervision of banks is not an exact science, and therefore, discretionary elements within the supervisory review process are inevitable" (Basel Committee on Banking Supervision, 2019).

² Supervisory scores are conceptually akin to ratings, in the sense they use a standardised scale to summarise the assessment, with the obvious differences in terms of their scope, function and confidentiality regime.

³ In the euro area, according to the principle of constrained judgement, in a four-grade scale, compared to the starting point the score can be improved by one notch and worsened by two notches on supervisory judgement (see European Central Bank, 2015).

We start by breaking down the changes in the OSS into the changes in the automatic score and supervisory judgement and by calculating the contribution of each component to the total change. The breakdown is illustrated in Figure 1.



Notes: The figure presents the average change in OSS vs the previous SREP cycle, broken down by bank groups with upgrades (negative numerical difference) and banks which received downgrades (positive numerical difference), respectively. SREP 2020 not represented, given the “pragmatic approach” applied in the context of the COVID19 pandemic. Analysis based on a variable sample of banks, depending on the number of significant banking institutions in each SREP cycle. The unit on the vertical axis is the score notch.

OSS upgrades are 59% supported by changes in the automatic score and 41% by changes in supervisory judgement. Conversely, OSS downgrades are 53% supported by changes in the automatic score and 47% by changes in the supervisory judgement. Therefore, the contribution of supervisory judgement to OSS changes, be it upgrades or downgrades, is sizable, being close to that of the automatic score.

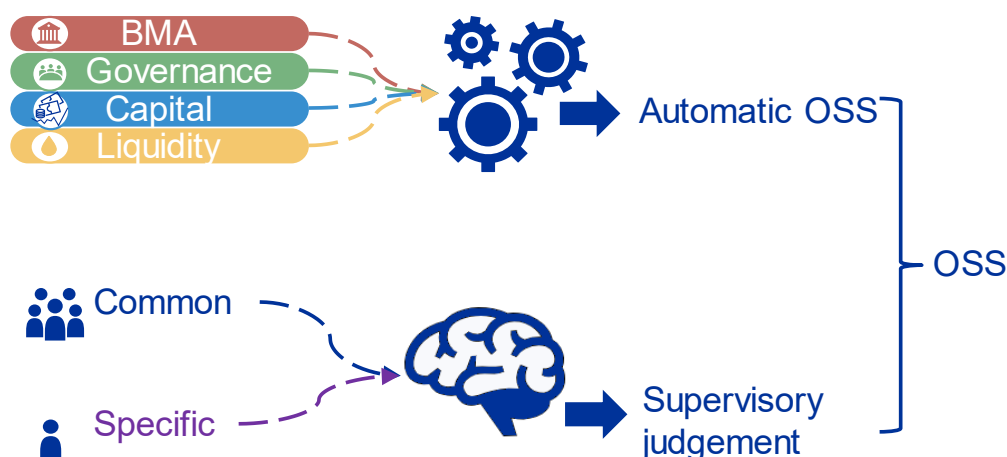
Are there commonalities in supervisory judgement?

Next, we focus on identifying patterns in the supervisory judgement that are valid across supervisors, as well as throughout time.

We postulate the existence of a common channel in the supervisory judgement, which entails commonalities in the way supervisors adjust the overall score. To test the hypothesis of common factors driving the supervisors’ judgment, we estimate an unobserved effects model for the supervisory judgement at the bank-year level for OSS.

Our approach is illustrated in Figure 2. The common component corresponds to the part of the supervisory judgment that is explained by scores, key risk indicators and macro controls. The remainder reflects the bank and supervisor specific component, but it also includes functional forms that could account for non-linear aspects, such as overweighing of some risks and interactions between risks.

Figure 2. Stylized representation of OSS driver categories



Notes: The figure provides a simplified representation of the framework that underpins our empirical estimation approach, inspired by the SREP methodology. We break down the OSS into the automatic component (simple average of the four SREP viability scores: BMA, governance, capital and liquidity) and supervisory judgment, which we further disentangle in a component that is common across a set of variables, time and supervisors, and one that is bank-specific for at least one of these three dimensions.

Panel data estimates explaining changes in the supervisory judgement component of the OSS show significant negative coefficients for the four SREP element scores. This suggests that, compared to the starting point of the automatic score, the supervisors change how the various aspects of the assessment are combined in the overall score and that the adjustment has a stabilizing effect, causing the overall score to fluctuate less than the automatic score.

In addition, as part of the common supervisory judgement, we find strong statistical evidence that supervisors perceive certain risks, such as credit risk (risk control aspects), as playing a decisive role in the bank's overall viability. We interpret this as evidence of non-linearity in the viability assessment of a bank. Testing what type of information features prominently in the common supervisory judgement, we find that it is linked to qualitative aspects of the banks' activity, like internal controls.

From a high-level policy perspective, credit risk has also been a consistent focus in the priorities set by European banking supervision from 2016 to 2024, further consolidating the observed dynamics. Initially, the emphasis was on addressing high levels of non-performing loans (NPLs) and enhancing banks' credit risk management practices. Over time, the focus has expanded to include emerging risks such as deteriorating credit quality due to economic uncertainties, climate-related credit risks, and vulnerabilities arising from geopolitical tensions and market disruptions.

Macroeconomic factors and bank control variables have lower direct contributions to the common supervisory adjustment of the OSS, while the largest part of supervisory judgement remains specific.

Are there systematic differences in the way supervisors adjust the overall score, depending on the bank categories?

We explore further the existence of sub-patterns within the common supervisory judgement. For this, we augment our model with interaction terms suitably chosen to explore the existence of sub-patterns within the common supervisory judgement, e.g., depending on the direction of the OSS change, riskiness, or the bank's size and complexity.

We find evidence that common aspects of the judgement can differ, depending on the direction of change in OSS (upgrade vs downgrade). Regarding the upgrade vs downgrade difference, and considering that the main role banks

play in the economy is lending, results show that supervisors have paid close attention to downside credit risk (CRED) during the post-pandemic period (2021-2024). This attention has been more pronounced when deciding whether to downgrade the overall score of a supervised bank, compared to upgrade decisions. This is in line with the main findings: supervisory judgement is used to recognize interaction between risks and prevent being blindsided by improvements in other risk areas that could conceal deteriorations in, e.g., credit risk (risk control aspects), due to limitations in the linearity of the automatic composite indicator starting point. The cautious behaviour seems warranted when placing the results in historical context. The period 2021-2024 marks a time when banks have emerged from the pandemic and care has been given by supervisors to ensure balance sheets are not accumulating non-performing loans.

What was driving P2R?⁴

As an outcome of the SREP assessment, banks may be required to hold additional own funds requirements (Pillar 2 requirements – P2R). Supervisory judgement is also a feature of setting the additional capital requirements - e.g., the P2R.

Unlike the OSS, there is no automatic P2R. For the first half of the period analysed in this paper (2016-2020) the ECB established P2R in a holistic approach, on the basis of the OSS, but without a mechanical link. Although the two were strongly correlated, the correlation cannot reach 100% due to the fact that risks can also be addressed by other measures⁵, such as qualitative ones. For the second half (2021-2024), the approach was expanded by looking more closely at institutions' individual risk drivers, while also taking into account banks' internal capital adequacy assessment process (ICAAP).

The overall risk profile is the most important common driver for the changes in P2R for the analysed period from 2016 to 2024. The relation between overall risk and P2R strengthened after 2021, with the first introduction of a new P2R methodology.

Drivers that are identified beyond those for which the supervision methodology explicitly described a direct link — for example, the changes in the overall risk profile and the internal capital adequacy score — can be regarded as the manifestation of the common supervisory judgement at the level of P2R. Among these elements, we identify a direct impact from the business model assessment (in addition to the effect captured already through the overall score) as an important driver. Nevertheless, the larger part in the change of P2R remains bank-specific.

Conclusions

Our findings hold valuable insights for supervisors and market participants alike and show that European supervisors actively use judgement when setting the OSS and the P2R capital requirements.

We identify significant similarities in how European supervisors apply their judgement to determine the OSS, lending strength to a theorized common supervisory judgement channel hypothesis.

We consider this identified effect to be an important channel through which supervisors remain on top of latest developments, evolving business models, and changing financial environments. At the same time, it addresses limitations such as the implicit assumptions of risk offsetting or substitutability, which are hard coded in all linear functional forms of composite indicators. This includes the automatic scores forming the start point for SREP assessments. As a result, supervisors appear to use judgement to account for interactions between risks, recognizing that strong performance in one risk area does not always compensate for weaknesses in others, and adjust scores accordingly.

⁴ Our analysis covers a period that ends in 2024. The ECB applies a new P2R methodology as of the 2026 SREP cycle.

⁵ See European Central Bank (2019).

The implication of the common supervisory judgement channel is that it provides the flexibility needed not only at the level of individual banks, but also at the level of the entire system. This flexibility allows supervisors to reflect, for example, supervisory priorities, without too frequent changes of the methodological priors on the scores. Finally, common supervisory judgement can be seen as having informational content on the overall status of the system. Specifically, changing the weights of the elements initially considered in the automatic score and directly including additional information in the assessment of overall risk can reveal shared concerns among supervisors, such as high levels of non-performing loans as observed at the end of the Great Financial Crisis and during the Sovereign Debt Crisis, macroeconomic conditions woes seen after the COVID-19 pandemic, or vulnerabilities in certain sectors, like commercial real estate more recently.

The presence of a sizeable idiosyncratic component in supervisory judgement confirms the importance of second line of defence supervisory functions in reducing unwarranted variability, while preserving flexibility to reflect bank-specific circumstances.

Our results can help bridge gaps in understanding caused by asymmetric information among various stakeholders regarding the process. We hope this work paves the way for more research into supervisory judgement for the euro area, for example to understand better the idiosyncratic component and supervisory activities, and fosters effective communication with market participants and the general public.

These data-driven insights are highly relevant in the current context of the SREP reform and show that a multitude of factors are taken into consideration to explain the observed patterns in scores and P2R.

References

- Basel Committee on Banking Supervision (2019). Supervisory review process. url: https://www.bis.org/basel_framework/standard/SRP.htm?tldate=20250422.
- European Banking Authority (2022). Final report on amending Guidelines on common procedures and methodologies for SREP and supervisory stress testing. url: <https://www.eba.europa.eu/activities/single-rulebook/regulatory-activities/supervisoryreview-and-evaluation-process-srep-4>.
- European Central Bank (2015). SSM SREP Methodology Booklet. url: https://www.bankingsupervision.europa.eu/ecb/pub/pdf/ssm_srep_methodology_booklet.en.pdf.
- (2019). SSM SREP Methodology Booklet - 2018 edition - to be applied in 2019. url: https://www.bankingsupervision.europa.eu/ecb/pub/pdf/ssm.srep_methodology_booklet_2018~b0e30ced94.en.pdf.
- Gaul, L. and J. Jones (2021). CAMELS Ratings and Their Information Content. Working Paper 1. Office of the Comptroller of the Currency. url: <https://www OCC.gov/publications-and-resources/publications/economics/working-papers-banking-perfreg/pub-econ-working-paper-camels-ratings.pdf>.

About the author(s)

Gabriel Bobeică is a Principal Supervisor in the Supervisory Risk-Financial Risk Division of the European Central Bank. He holds a PhD in Finance from Bucharest University of Economic Studies, and has previously worked in the horizontal banking supervision at the ECB, as well as in academics. He has a vast experience in developing supervisory methodologies and benchmarking and analysis of supervisory outcomes. His research is currently focusing on drivers and impact of judgement in banking supervision.

Silviu Oprică is a Senior Supervisor in the Supervisory Risk-Financial Risk Division of the European Central Bank. He holds a PhD in Economics from Goethe University Frankfurt, and has previously held positions at the European Central Bank's *Research, Macroeprudential and Financial Stability*, and *Supervisory Strategy and Risk* directorates. His research interests include banking supervision, macroprudential policy and financial stability, shadow banking, panel data econometrics, and machine learning applications.

SUERF Policy Notes and Briefs disseminate SUERF Members' economic research, policy-oriented analyses, and views. They analyze relevant developments, address challenges and propose solutions to current monetary, financial and macroeconomic themes. The style is analytical yet non-technical, facilitating interaction and the exchange of ideas between researchers, policy makers and financial practitioners.

SUERF Policy Notes and Briefs are accessible to the public free of charge at <https://www.suerf.org/publications/suerf-policy-notes-and-briefs/>

The views expressed are those of the authors and not necessarily those of the institutions the authors are affiliated with.

© SUERF – The European Money and Finance Forum. Reproduction or translation for educational and non-commercial purposes is permitted provided that the source is acknowledged.

Editorial Board: Ernest Gnan, David T. Llewellyn, Donato Masciandaro

Designed by the Information Management and Services Division of the Oesterreichische Nationalbank (OeNB)

SUERF Secretariat

c/o OeNB, Otto-Wagner-Platz 3A-1090 Vienna, Austria

Phone: +43 1 40 420 7206

E-Mail: suerf@oenb.at

Website: <https://www.suerf.org/>