EUROPEAN CENTRAL BANK

EUROSYSTEM

# Working Paper Series

Gabriel Bobeică, Silviu Oprică

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

No 3233

**Abstract**

We empirically analyse the role of judgement in assigning overall scores by the euro area supervisors as part of the yearly Supervisory Review and Evaluation Process (SREP), which evaluates banks' risks and sets supervisory actions. We also analyse its role in shaping the drivers of the Pillar 2 capital requirement (P2R) that banks must fulfil.

We find that supervisors actively adjust the weight of the components of the overall score to reflect qualitative information, thereby smoothing fluctuations in the final assessment. The analysis reveals a *common* supervisory judgement channel, which could reflect shared priorities and concerns, such as systemic vulnerabilities or macroeconomic conditions. We also show that certain risks, such as credit risk, can play a decisive role in the overall assessment of a bank's viability.

These findings underpin the critical role of judgement in adapting supervisory frameworks to evolving risks and systemic conditions, providing flexibility at both the individual and system-wide levels.

**JEL Classification**: G21, G28, C23, E58

**Keywords**: supervisory judgement, Supervisory Review and Evaluation Process (SREP), overall SREP score, Pillar 2 capital requirements (P2R), panel data

## Non-technical summary

Supervisors evaluate banks' risk profiles using composite scores, which summarize risks and guide future supervisory actions. Two prominent systems are the CAMELS rating in the US and the Overall Supervisory Review and Evaluation Process Score (OSS) in the EU. While the composite scores aim to reflect the risk situation of a given bank on the basis of objective and often quantitative information available to supervisors, qualitative elements within the supervisory process are essential for achieving a well-rounded assessment of the bank.

Supervisory judgement allows supervisors to account for qualitative aspects and forward-looking information. It also helps address emerging risks that are not fully captured by mechanical score calculations. This enables tailored evaluations, especially in complex or uncertain circumstances, by addressing limitations such as the backward-looking nature of the data or the assumption that risk dimensions are perfect substitutes. For instance, a strong performance in one risk area cannot always offset weaknesses elsewhere. Supervisors also use judgement to identify interactions between risks and adjust scores accordingly.

Despite its importance, there is limited empirical analysis of the extent to which supervisory judgement affects the overall assessment, due to data confidentiality and the relative novelty of certain supervisory frameworks. The study addresses this gap by analysing how euro area supervisors apply judgement when setting the overall SREP score. The process starts with an automatic score derived from the equal-weighted average of the four SREP element scores. Supervisors then adjust this score based on qualitative insights and specific vulnerabilities, reflecting a mix of quantitative and qualitative inputs.

The study finds that supervisory judgement significantly influences changes in the overall SREP score, contributing almost as much as the automatic score. Supervisors adjust the weights of the SREP elements to dampen fluctuations in the overall score. They directly incorporate relevant information into their overall assessment, preventing its impact from being potentially diluted during the aggregation into higher-level scores. This highlights the non-linear nature of bank viability evaluations, where certain risks, like credit risk, can play a decisive role.

Our empirical analysis fosters a better understanding of how supervisors use judgement. One part of the supervisory judgement is explained by common determinants, and a more important part remains idiosyncratic. This provides supporting evidence for a *common* supervisory judgement channel, where supervisors across the euro area apply adjustments to OSS scores in a similar manner. Thus, there is a common denominator in terms of their response and the variables they respond to. This commonality could reflect shared priorities and concerns, such as systemic vulnerabilities or macroeconomic conditions. Following crisis episodes like the Great Financial Crisis or the COVID-19 pandemic, supervisors may emphasize non-performing loans or sectoral vulnerabilities. The study also finds that common supervisory judgement can vary depending on the direction of changes of the overall SREP score (upgrades or downgrades), but it is less influenced by a bank's riskiness, size, or complexity.

Beyond overall SREP score, supervisory judgement is evident in setting Pillar 2 capital requirements (P2R), which address additional risks not covered by minimum capital requirements. Supervisory judgement in P2R is not observable, but it can be analysed by estimating what drives the P2R changes. For the time period under analysis (2016-2024), beyond the factors for which the direct link is explicitly prescribed in the applicable supervisory methodology (e.g., overall risk profile, and internal capital adequacy processes - ICAAP) the study identifies other key drivers of P2R changes, like the business model assessment. We interpret this as evidence of common supervisory judgement in setting P2R.

The findings underscore the importance of supervisory judgement in adapting evaluations to individual banks and systemic conditions. By providing flexibility, judgement helps supervisors address emerging risks without frequently revising methodologies. The study confirms the empirical impact of supervisory judgement on OSS and P2R outcomes, offering valuable insights for supervisors and stakeholders. It highlights that supervisory judgement is not only essential for tailoring assessments but also for reflecting broader system-wide characteristics. These results can help bridge information gaps among stakeholders and support the ongoing reforms of supervisory frameworks.

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

# 1 Introduction

The importance of prudential oversight to reduce information asymmetries, safeguard financial markets and reduce moral hazard has been a central topic of the academic literature over the years. The case for prudential regulation is made in seminal theoretical works such as Dewatripont and Tirole 1994, who argue regulators' main role is preventing financial intermediaries' insolvencies, minimising the risks of accounting manipulations, and thus preventing financial market failures.

Carletti et al. 2016 further show that, in a banking union, the central supervision has an incentive to enhance information collection. In practical terms, this refers to the qualitative and quantitative information supervisory assessments are based on, in order to eliminate information asymmetries that lead to increased risk-taking by supervised banks.

The empirical literature has been able to test a number of these theoretical predictions, with works such as Hirtle, Kovner, and Plosser 2020 and Hirtle and Kovner 2021 showing banks more intensely supervised exhibit less risk-taking, are less volatile, and less sensitive to downturns. For the European case, Abbassi et al. 2023 show inter alia that exercises such as the European Central Bank's Asset Quality Review (AQR), have led to a de-risking of securities and credit prior to the AQR compliance checks.

Agarwal, Luca, et al. 2014 then focus on the implementation side of banking regulation, showing that regulators may inconsistently implement identical rules due to differences in incentives and institutional design, which can adversely impact the effectiveness of the regulation.

Many academic works focus on regulatory aspects, and in particular macroprudential related topics, engaging only broadly the actual microprudential supervisory process, owing most likely to data limitations and the complexity of the microprudential review process itself. We therefore contribute to this strand of the literature by analysing how supervisors in the euro area assess the riskiness of supervised banks, and in particular zoom in on a component of these regular assessments – the supervisory judgement, which plays a key role in microprudential supervisory outcomes. We are aware of only one other work engaging this policy-relevant topic, Agarwal, Morais, et al. 2024 for the case of the US.

To better understand what supervisory judgement entails, it is useful to note that supervisors often assess the soundness of banks through scores, 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 several 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 EU supervisors to summarise the risk profile of the bank from a viability perspective (European Banking Authority 2022).

Within the supervisory process, judgemental elements are always present (Basel Committee on Banking Supervision 2019). Therefore, supervisory judgement is commonplace in the practice of supervisors scoring the banks. Compared to a pure mechanical approach, where some pre-defined rules encode hard data to risk scores, discretion allows the supervisors to avoid blind spots and to tailor their assessment to the inherent complexity and specificity of the institutions, their evolution over time, including changes in risk appetite or their niche in the market, and to also adjust scores in recognition of qualitative aspects such as risk culture and governance practices. Furthermore, most risk indicators are of backward-looking nature, as they are derived from historical data, while supervisors, through their activities, also have access to forward-looking information, as well as peer group information allowing benchmarking of risk profiles and risk management practices in the industry. In this context, discretionary elements introduced into the scores reflect supervisory knowledge and aim to better adapt to situations of heightened uncertainty and emerging new risks.

Despite its importance, there are, however, very few empirical studies analysing the judgment component of supervisory scores and quantifying its magnitude in the overall scores. This scarcity is no doubt the result of factors such as high confidentiality of data or the lack of sufficiently long time series for a robust econometric analysis, as some supervisory frameworks are still relatively recent.

The importance of supervisory judgement in banking supervision is emphasised in policy papers (see Balan et al. 2025) and external review reports (Expert Group to the Chair of the Supervisory Board of the ECB 2023). For the case of the US, Agarwal, Morais, et al. 2024 show that discretion is a sizable component of supervisory bank ratings which can deviate to a large degree from algorithmic measures.

Our concept of supervisory judgement is akin to that of supervisory discretion in Agarwal, Morais, et al. 2024. Still, there are some relevant differences. In Agarwal, Morais, et al. 2024 discretion is modelled at the examiner and examiner-time level and it reflects deviations from an optimal rule, which the authors estimate, as it is unobservable. In our case, supervisory judgement is modelled at the institution-time level and it reflects supervisors' adjustments to the starting point of a rule-based score and is, therefore, observable.

Our work aims to contribute to filling a gap in the literature, through an empirical analysis of how euro area supervisors apply supervisory judgment at the level of the OSS score. This analysis is feasible due to the fact they can choose to adjust the OSS score from the starting point of an automatic score equal to the simple average of the four encompassing element scores, to a final implementation. This final score encompasses supervisory judgment and additional insights which then forms the basis of the risk profile and, potentially, measures aimed at the bank under supervision.

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. We show that the supervisory judgement provides an important contribution to changes in OSS, close to that of the automatic score. To our knowledge, this is the first study to empirically confirm the impact supervisory judgement has on the supervisory outcomes for the euro area.

We postulate and confirm empirically the existence of a common channel in the supervisory judgement by estimating an unobserved effects model for the supervisory judgement at the bank-year level for OSS. This entails commonalities in the way supervisors adjust the overall score. Compared to the starting point of the automatic score, they change how the various aspects of the assessment are combined in the overall score. This adjustment has a stabilizing effect, causing the overall score to fluctuate less than the automatic score.[1]

In addition, as part of the common supervisory judgement, we find strong statistical evidence that supervisors perceive certain risks, such as credit risk, as playing a decisive role in the bank's overall viability. Not only is the information on these risks leveraged upon when assessing SREP elements such as capital adequacy, thereby indirectly impacting the overall score, but it is also directly taken into consideration in their adjustment of the OSS. We interpret this as evidence of non-linearity in the viability assessment of a bank. We show that this additional information is linked to qualitative aspects of the banks' activity, like internal controls. Finally, we show that 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.[2]

In the next step, 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 that common supervisory judgement can vary depending on the direction of the OSS change (for the recent SREP cycles), but not necessarily with respect to the riskiness, or the cluster of the bank, for which the evidence is much weaker and less robust.

Supervisory judgement is also a feature of other aspects of the supervisory activities, such as setting the additional capital requirements — e.g., the Pillar 2 requirement (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. For the second half (2021-2024), the approach was expanded by looking more closely at institutions' individual

[1]The interplay between the automatic score and the OSS is well exemplified by the evolution of scores in SREP 2024, in which "[t]he stability of the banks' SREP scores reflects, on the one hand, an improvement in key risk indicators and, on the other hand, the high degree of uncertainty concerning the economic outlook." (Buch 2024)

[2]Non-linearities in the functional form of the supervisory judgement may explain additional aspects of the remaining supervisory judgement; however, this will be addressed in future research.

risk drivers, while also taking into account banks' internal capital adequacy assessment process (ICAAP).[3]

Therefore, all potential factors can be considered when estimating what drove the P2R changes for the time interval in our analysis. 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[4] — 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.

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.

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 paper is structured as follows. The next section provides a simplified overview of the SREP methodology, underpinning our modelling and data set choices. The third section is dedicated to the presentation of the research methodology, discussing the results, and conducting robustness checks. The fourth section concludes.

---

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

[4]The internal capital adequacy assessment process is an important input factor in the SSM SREP, feeding into all SREP assessments and into the Pillar 2 capital determination process in accordance with the EBA Guidelines on common procedures and methodologies for the SREP (European Central Bank 2018).

## 2 Institutional background

### 2.1 Overview of SREP[5]

The ECB is one of the largest supervisors in the world. It supervises directly more than 110 banking groups.[6] It fulfils its supervisory mandate through regular assessments of banks, one of which is the Supervisory Review and Evaluation Process (SREP), and by means of qualitative and quantitative information collected from the banks under supervision.

Our investigation focuses on the overall SREP score (OSS) and the Pillar 2 Requirement (P2R), which are two of the most important outcomes of the supervisory review process. OSS reflects the risk profile of the institution from a viability perspective,[7] while the P2R covers risks that are not (sufficiently) covered by the Pillar 1 requirements. OSS is the numerical indicator of the overall SREP assessment, which forms the basis for the supervisors to decide the most appropriate set of supervisory measures that would best address the overall situation of the institution and its specific weaknesses.

The SREP's scoring system implemented by European banking supervision ranks banks from 1 (low risk) to 4 (high risk). The framework[8] includes the possibility of applying "+"/"-" qualifiers to scores 2 and 3.

Scores are mostly used as a means of summarising supervisors' views and facilitating high-level, cross-sector comparisons and communication, both within the Single Supervisory Mechanism (SSM) and with the institution itself.[9]

When setting all scores, supervisors exercise judgement based on their knowledge of the institution. More precisely, the principle of *constrained judgement*[10] applies throughout the SREP, allowing supervisors to take the specificity and complexity of an institution into account while also ensuring a consistent supervisory approach.

OSS is set starting from an automatic overall score (automatic OSS), which is computed as the simple average of the four SREP element scores: *business model assessment*, *internal*

[5]This overview follows the public available information on the SREP methodology implemented by the SSM (European Central Bank 2015, 2016, 2017, 2019, 2021, 2022, 2023a, 2023b, 2024c; Expert Group to the Chair of the Supervisory Board of the ECB 2023).

[6]At the end of 2024, there were 114 significant banking groups under European banking supervision. The banks directly supervised by the ECB hold almost 82% of banking assets in the participating countries (Montagner 2025). Banks are considered significant based on: size (total assets over €30 billion or among the three biggest in a country), importance for the economy where they are located, significance of cross-border activities and whether they have requested or received financial assistance directly from the European Stability Mechanism or the European Financial Stability Facility.

[7]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).

[8]See Expert Group to the Chair of the Supervisory Board of the ECB 2023.

[9]See European Central Bank 2024c.

[10]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).

*governance*, *capital adequacy*, and *liquidity adequacy*, each providing 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.

Supervisors[11] rely on a comprehensive set of information in their assessment. This includes a wide range of quantitative and qualitative information sources, with quantitative data being of particular importance for fostering consistency and comparability. Key sources of quantitative information include risk indicators based on financial and prudential reporting, indicators of economic and market conditions, regulatory disclosures, supervisory stress test results, market information (e.g., external ratings, investors' quantitative analyses, etc.). Among the key sources of qualitative information, we note relevant documentation reflecting the bank's governance, risk management and internal controls, supervisory findings (e.g., from inspections and meetings), reports on the operating environment of institutions.[12]

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 judgment, 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.

As an outcome of the SREP assessment, banks may be required to hold additional own funds requirements (Pillar 2 requirements – P2R). For the first half of our sample, ECB Banking

[11] We employ the general term of 'supervisor' to refer to the Joint Supervisory Team (JST) that performs the SREP. A JST is established for each significant institution. JSTs are formed of staff of the ECB and the relevant national supervisors, including the competent authorities of the countries in which credit institutions, banking subsidiaries or significant cross-border branches of a given banking group are established.

[12] See European Central Bank 2024c for a detailed presentation of the key sources of supervisory information, as well as the three phases in risk assessment for a schematic explanation (see, e.g., European Central Bank 2016) on how this translates into scores.

Supervision adopted a holistic approach to setting the P2R, 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 measures,[13] e.g., qualitative measures.

In 2021, the holistic approach used to determine the P2R was expanded by looking more closely at institutions' individual risk drivers, which should be addressed through additional capital requirements.[14] In this approach, P2R was established using a four-step procedure based primarily on its risk assessment for each bank but also taking into account banks' internal capital adequacy assessment processes.[15] In the first step, the supervisor selects the appropriate initial P2R from a bucket of possible values based on the assessment of the overall risk to the bank's capital. This assessment is performed by applying weighting factors for Pillar 2 risks to the scores of the aforementioned SREP elements, and by using expert judgement to take into account the bank's specific situation, including the reliability of the bank's internal capital adequacy assessment process. In the second step, the supervisor breaks down the initial Pillar 2 requirement into several risk-by-risk add-ons. In the next step, the supervisor challenges the initial risk-by-risk add-ons resulting from the previous step. Finally, the supervisor determines the final risk-by-risk add-ons that lead to the definitive P2R.[16] Starting with the 2026 SREP cycle the P2R is based on a revised methodology, driven more directly by relevant areas of risk, and in which higher risks will continue to result in worse SREP scores and a higher Pillar 2 Requirement.[17]

## 2.2 Data

We leverage in our analysis on a unique set of supervisory data. We focus on identifying patterns in the supervisory judgement that are valid across supervisors, as well as throughout time. The dataset is an unbalanced panel that covers a history of nine SREP cycles, from 2016 to 2024,[18] with observation unit being the bank-SREP cycle, and the cross-section consisting of all significant banks under SREP assessment each year.

[13] See European Central Bank 2019.

[14] See European Central Bank 2021.

[15] See Expert Group to the Chair of the Supervisory Board of the ECB 2023.

[16] See European Central Bank 2024c.

[17] While information from banks' internal capital adequacy assessment process outcomes will no longer directly affect the Pillar 2 Requirement, the supervisory assessment of the quality of banks' internal capital adequacy process will continue to feed into the SREP assessments of business models, internal governance and overall risk management (Buch 2025b).

[18] SREP was first conducted on a common approach by the SSM in 2015. However, we choose not to include the respective cycle in our sample, as the gains from having one additional time data point would be outweighed by the costs attached to ad-hoc assumptions needed to deal with very significant changes in the approach and methodology that were introduced in 2016 such as the introduction of the Pillar 2 requirement and Pillar 2 guidance, and the introduction of score qualifiers for the OSS (European Central Bank 2015, 2016).

We restrict our dataset to a much smaller subset of the supervisory information, for reasons of feasibility, while also aiming to have a relatively good time coverage of the bank cross-section. We include in our dataset three types of data: macroeconomic indicators, key risk indicators calculated based on supervisory data (e.g., financial reporting - FINREP and common reporting - COREP), and risk scores that result in the SREP assessment.

Macroeconomic factors are selected to reflect the real sector, general price developments, as well as the financial sector developments. We opted for real GDP growth and the inflation rate (based on HICP index) in the home country of each bank, the national stock index (in log), and VSTOXX.

We use key risk indicators, with a double role. First, they can pinpoint deficiencies or vulnerabilities that the supervisors consider directly relevant to the overall score. Second, they account for bank characteristics, such as size and risk profile, that can also play a role in the supervisory judgment.

Key risk indicators capture general bank characteristics, such as size and composition of the balance sheet, as well as information on general risk profile, such as the risk weights density, and on individual risks, such as the non-performing loans (NPL) ratio. The full list of such key risk indicators includes: *total assets* (log transformation), loan book calculated as the *share of loans in total assets*, *government loans ratio* calculated as the share in total loans, *NPL ratio*, *coverage ratio*, *risk weighted assets (RWA) density* calculated as the ratio between the total risk exposure amount and total assets, *credit risk*, *market risk* and *operational risk shares of RWA*, *net stable funding ratio* (NSFR), *liquidity coverage ratio* (LCR), *deposits ratio*, *loans to deposits*, *cost-to-income ratio*, *return on equity*, *return on assets*, *net interest income as percentage of total assets* and the *change in the economic value of equity as a result of a 200 basis points (bps) shock*.

In the list of our explanatory variables, we use SREP element and combined scores to capture the importance supervisors would give to the respective risks and corresponding weaknesses, in line with what the SSM's supervisory methodology prescribes. In this context, element scores refer to the four elements assessed in the SREP, and combined scores refer to individual risks to capital. Element scores include *business model and profitability assessment* (BMA), *governance and internal management* (GOV), *capital adequacy* (CAP) and *liquidity adequacy* (LIQ) scores,[19] and combined scores denote *credit risk* (CRED), *market risk* (MK), *operational risk* (OP) and *interest rate in the banking book* (IRRBB) scores.[20] Additionally, the P2R estimations incorporate a constructed overall score variable and the ICAAP score. The overall score variable used in the P2R regression is meant to proxy the overall risk profile of the institution, as reflected and considered in the SREP methodology. The series combines

[19] Also referred to in the paper as *viability scores*.
[20] Also referred to in the paper as *risk scores*.

the values for the overall SREP score before 2021, and those of the overall risk score afterwards, to account for the change in P2R methodology that took place in 2021. The ICAAP score is incorporated into the P2R model to reflect its significance in determining the P2R.

Dummy variables are added at the country, peer group, and cluster level[21] to account for bank-specific factors and capture qualitative information.

The dependent variables in our analysis are the supervisory judgement component of the OSS and the P2R. To obtain the supervisory judgement component of the OSS, we subtract the automatic OSS score from the final OSS before estimation. P2R values used in the estimation are calculated by excluding the component related to the non-performing exposures (NPE) shortfall add-on, which has a dedicated calculation methodology. They also refer to total capital requirements.

Score series are expressed in integers from 1 to 4. If individual scores included qualifiers, they were replaced with the corresponding scores without qualifiers. All other series, except for the log of total assets, are expressed as percentages.

The variables used in the econometric analysis and their summary statistics are presented in Table A.1.

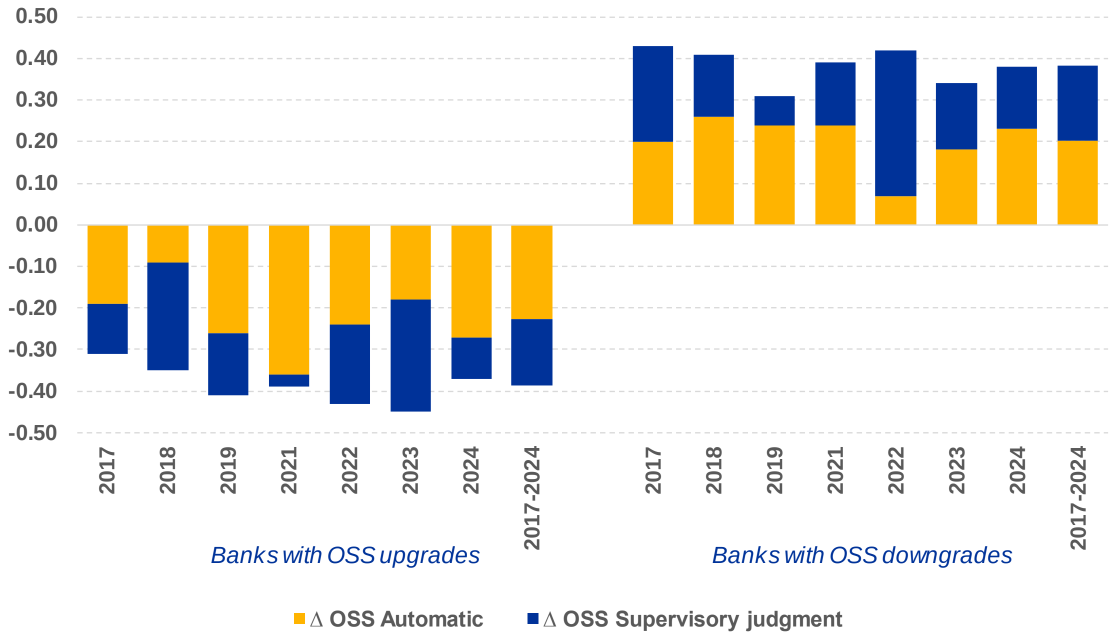
# 3 Estimation methodology and results

## 3.1 Estimation methodology

We base our approach on the SREP methodology implemented within the SSM. Through this lens, the OSS reflects the risk of the bank in the eyes of its supervisor, while various types of factors are at play and can be observed through key risk indicators, risk controls and macroeconomic variables. These are channelled into the overall assessment through the "automatic score", which is the simple average of the four SREP element scores (business model, governance, capital adequacy and liquidity adequacy). Therefore, based on a mechanistic rule that applies in the same way to all the banks, these elements are factored into a preliminary "automatic score," which is then catered to the specificities of each institution through supervisory judgement to better reflect its risk profile. This includes adjusting the weights of the SREP elements to address weaknesses identified as particularly significant for the institution under review. The supervisory adjustment additionally considers bank-specific characteristics, such as business model, as well as other variables that describe the environment in which the banks operate. Our analysis aims to assess which of these aspects prevail in the supervisors' decision to worsen or improve their view on a bank's risks.

[21] Peer groups reflect the business model specificities, while the clusters reflect the size and complexity, e.g., "Cluster 1" covers the largest and most complex banks (European Central Bank 2024b).

Figure 1: Yearly changes in the Overall SREP Score (OSS), decomposition into the automatic and supervisory judgement components



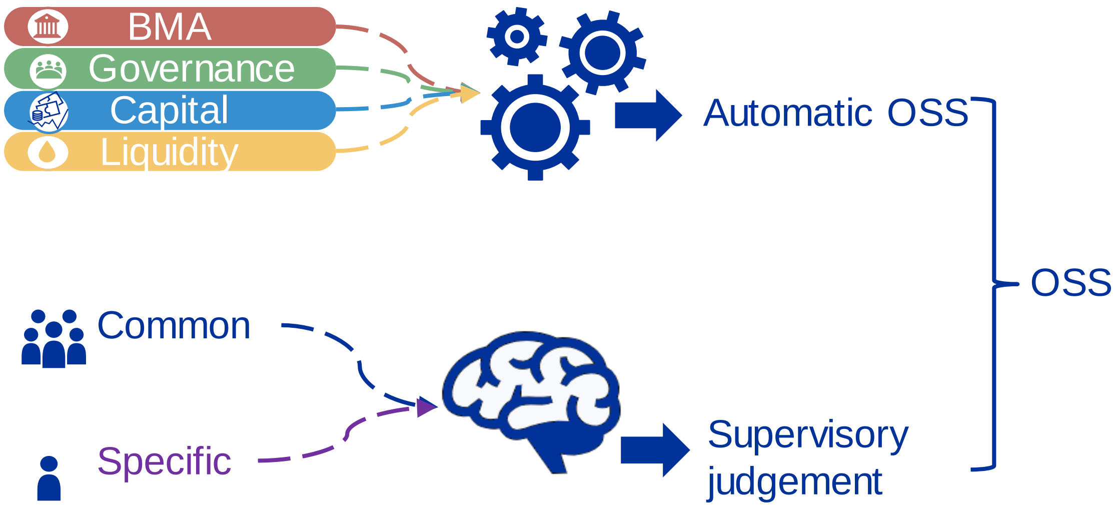

Notes: *The figure presents the average change in OSS vs the previous SREP cycle, broken down by banks with upgrades (negative numerical difference) and banks which received downgrades (positive numerical difference), respectively. SREP 2020 not represented, given the "pragmatic approach" applied in the context of the COVID-19 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.*

We focus our analysis on the supervisory judgement component, which reflects a significant part of the overall score beyond the mechanical starting point of the automatic score. Figure 1 illustrates the breakdown of the yearly changes in OSS due to the automatic score and supervisory judgement. OSS upgrades are 59% due to changes in the automatic score and 41% due to changes in supervisory judgement.[22] Conversely, OSS downgrades are 53% due to changes in automatic score and 47% due to changes in 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. Supervisory judgement has in fact a higher contribution, as the methodology provides for its application also at the level of the SREP element scores (which give the automatic OSS), as well as at the combined and risk level scores (which inform some of the element scores).

We test the existence of a common channel of the supervisory judgement, i.e., we examine the hypothesis that there are common factors driving the supervisors' judgement. Our

[22] Average, across all the SREP cycles from 2017 to 2024; SREP 2020 not included, due to the pragmatic approach applied in the context of the COVID19 pandemic.

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 judgement, 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. The framework we use to confirm the existence of the common supervisory judgement component is presented in equation (2), corresponding to the part of the supervisory judgement that is explained by scores, key risk indicators and macro controls.*

approach is illustrated in Figure 2. As argued in the Introduction, the remainder also includes functional forms of supervisory judgement that account for non-linear aspects, such as overweighing of some risks and interactions between risks.

We adopt for both the supervisory judgement component of the OSS and the P2R the following unobserved effects model:

$$y_{it} = \boldsymbol{x}_{it} \cdot \boldsymbol{\beta} + c_i + u_{it}; t = 1, 2, \cdots, T. \tag{1}$$

Where $y_{it}$ is supervisory judgement in the OSS, respectively in the P2R; $\boldsymbol{x}_{it}$ contains observable variables that vary across institution $i$ and over time period $t$, such as the SREP element scores, individual risk scores, key risk indicators (KRIs), macro variables (common for banks in the same country), variables that change with $t$, such as the SREP cycle dummy and macro variables that are common across banks, and variables that change with each bank institution $i$, such as country, peer group, and cluster dummy variables; $c_i$ denote unobservable effects (individual heterogeneity), treated as random, and $u_{it}$ denote the idiosyncratic errors.

The left-hand side (LHS) variable is exogenous by construction, as it is a variable constructed by supervisors on the basis of historical data available to them at the time of the evaluation. By the nature of the SREP process, scores are set in a sequence from more granular, individual scores, towards higher level, more aggregate scores. Therefore, no reverse-causality or simultaneity should exist, enabling the use of standard panel data estimators. Due to the fact that the LHS variable as well as financial and accounting time series data used as explanatory variables generally tend to show persistency, we will use a first difference (FD) estimator to account for any potential serial correlation present in the idiosyncratic error term. Other estimators yield results similar to those in the main tables presented in the forthcoming sections.

To retrieve the supervisory judgement component of the OSS, we subtract the automatic OSS score from the final OSS. To operationalise the model in equation (1), we then use this remainder as left-hand-side (LHS) variable in a regression specified in first difference:

$$
\begin{aligned}
\Delta Supervisory\_judgement\_OSS_{it} = {} & \boldsymbol{\beta} \cdot \Delta SREP\_Element\_Scores_{it} \\
& + \boldsymbol{\gamma} \cdot \Delta Combined\_Scores_{it} \\
& + \boldsymbol{\delta} \cdot \Delta Key\_Risk\_Indicators_{it} \\
& + \boldsymbol{\theta} \cdot \Delta Macro\_controls_{k(i)t} + \varepsilon_{it}.
\end{aligned}
\tag{2}
$$

where the observation unit is at the bank $i$ and SREP cycle $t$ level, $k$ is the home country of bank $i$, and $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$, $\boldsymbol{\delta}$, $\boldsymbol{\theta}$ are vectors of coefficients.

A similar estimation is employed for the first difference in P2R:

$$
\begin{aligned}
\Delta P2R_{it} = {} & \alpha \cdot \Delta Overall\_Score_{it} \\
& + \boldsymbol{\beta} \cdot \Delta SREP\_Element\_Scores_{it} \\
& + \boldsymbol{\gamma} \cdot \Delta Combined\_Scores_{it} \\
& + \omega \cdot \Delta ICAAP\_Score_{it} \\
& + \boldsymbol{\delta} \cdot \Delta Key\_Risk\_Indicators_{it} \\
& + \boldsymbol{\theta} \cdot \Delta Macro\_controls_{k(i)t} + \varepsilon_{it}.
\end{aligned}
\tag{3}
$$

Our modelling choices are underpinned by the SREP methodology applicable during the time interval under analysis, as summarised in section 2. We use the SREP element scores, which are part of the automatic score, also to explain the supervisory judgement, and to reflect the possibility of supervisors to adjust their weights in the OSS as part of supervisory judgement.

The estimations are conducted for the full sample, 2016-2024, as well as for two subsamples: 2016-2020 and 2021-2024. The decision to split the sample at 2021 was based on the respective SREP cycle being the first when major changes in the SREP methodology were implemented, e.g., regarding the determination of P2R, or the extension of qualifiers that started in the same year with the internal governance (GOV) score and extended in the following years to other element and individual risk scores.

### 3.2 Empirical results

This sub-section presents the main findings. The key results were tested across a wide range of robustness checks, which are discussed at the end.

We interpret and discuss robust and unambiguous results, meaning that the coefficients are significant also in the alternative estimations performed in the robustness check part and the estimates have the expected sign, e.g., a worsening of an individual score is used by the supervisors to adjust toward a similarly worse overall score.

#### Are there commonalities in supervisory judgement?

Checking the hypothesis regarding commonalities in supervisory judgement is equivalent to testing whether coefficients estimated in equation (2) are significantly different from zero. Components of the vector $\boldsymbol{\beta}$ statistically different from zero provide empirical evidence that as part of supervisory judgement, supervisors adjust the weights given a priori in the automatic score to the corresponding SREP element (BMA, GOV, CAP, LIQ), with the sign indicating whether the fluctuations in the automatic score are amplified (in case of "+") or reduced (in case of "-") by this adjustment. Statistical significance for the components of the other coefficient vectors, namely individual risk aspects ($\boldsymbol{\gamma}$), bank specific characteristics ($\boldsymbol{\delta}$) or macro factors ($\boldsymbol{\theta}$), hints at aspects that are directly relevant for the viability of banks. These factors reflect the supervisors' views after (or conditional on) the indirect influence which is channelled through the element scores that is accounted for. This construction implicitly provides us with a more formal definition of common supervisory judgement. Having the estimated coefficient not significant for a certain factor (e.g. SREP element, combined risk score, KRI, macro) does not imply that the respective factor is not considered by a supervisor (or a sub-set) or in a SREP cycle (or a sub-sample), but rather that it is not relevant as an adjustment factor from the perspective of the whole sample of banks and the entirety of the period analysed.

Results for the supervisory judgement component of the OSS, presented in Table 1, suggest

Table 1. OSS Supervisory judgement. First-differencing (FD) estimation

| Dependent variable: | Sample | | |
|---|---|---|---|
| OSS Supervisory judgement | 2016-2024 | 2016-2020 | 2021-2024 |
| BMA Score | -0.1509*** | -0.1718* | -0.1122*** |
| | (-3.93) | (-1.99) | (-2.63) |
| GOV Score | -0.0836** | -0.1543** | -0.0742* |
| | (-2.04) | (-2.18) | (-1.90) |
| CAP Score | -0.0906** | -0.1186 | -0.0833* |
| | (-2.05) | (-1.51) | (-1.96) |
| LIQ Score | -0.1765*** | -0.1645* | -0.1833** |
| | (-2.69) | (-1.78) | (-2.58) |
| CRED Score | 0.1127*** | 0.1225** | 0.0890** |
| | (3.26) | (2.13) | (2.22) |
| MK Score | 0.0240 | 0.1036 | -0.0039 |
| | (0.77) | (1.40) | (-0.11) |
| OP Score | 0.0678* | 0.1420** | 0.0042 |
| | (1.86) | (2.22) | (0.10) |
| IRRBB Score | 0.0252 | 0.0543 | 0.0264 |
| | (0.94) | (0.87) | (0.91) |
| Macro controls | Yes | Yes | Yes |
| Key Risk Indicators | Yes | Yes | Yes |
| SREP cycle dummy variables | Yes | Yes | Yes |
| Observations | 474 | 162 | 312 |
| Banks | 99 | 63 | 98 |
| Adj. R-squared | 0.21 | 0.28 | 0.24 |

The table presents the main results of the estimation based on equation (2), in which the dependent variable is the 'yearly change of the supervisory judgement in the OSS,' which is measured as the difference between OSS set by supervisors and the automatic OSS, the latter being calculated as the simple average of the four SREP viability scores. Individual scores are employed in their numerical form, with 1 representing the full difference between two main score categories (e.g., '2' and '3'). The estimated coefficients for the four SREP viability scores (BMA, GOV, CAP and LIQ) should be interpreted jointly with the value of 0.25 assumed a priori in the automatic score, which equals the simple average of the four viability scores. Explanatory variables (except for SREP cycle dummy) are in first difference. Robust t statistics, calculated based on standard errors clustered by banks, in parentheses. * $p<0.10$, ** $p< 0.05$, *** $p<0.01$. A more detailed version of the results is available in the Annex (Table A.3).

that, compared to the automatic score, supervisors adjust downwards the weights of the SREP elements, and that additional information, such as that linked to credit risk, is considered directly relevant for the OSS. In other words, the estimated coefficients for the SREP element scores are all significant and negative. This holds true for the estimation conducted using the full sample and the two sub-periods, except for the coefficient of capital adequacy (CAP) in the 2016-2020 sample estimates. Given the smaller sample size, these results should be interpreted with caution. The credit risk score coefficient estimate is positive and significant across all three samples. On average, over the 2016-2024 period, while credit does not feature in the composition of the automatic score directly, it has been given a 0.11 relative coefficient in the final score, at the expense of the other elements which have negative coefficients. For example, liquidity has experienced one of the largest relative reductions in importance (of approximately 0.18) among the 4 elements in the automatic score. This result is in line with expectations given the importance of credit risk for banking institutions in general, as well as the fact ample liquidity has been available in the system over the 2016-2024 period following multiple liquidity support measures, alleviating concerns over bank liquidity issues in the studied period. The effect is sizable in magnitude. It implies a greater impact than both liquidity and business model assessment (BMA), though slightly less than the impact of the capital score component on the final score.

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.

Common behaviour may be conditional on a sample and/or time horizon. A comparison of estimation results across samples suggests that the factors considered by supervisors in their adjustments vary over time, as is the case with the internal governance (GOV) score, with an increase in the negative weight in the adjustment from -0.15 in the early SREP cycles to -0.07 in recent years, suggesting governance is playing an increasingly important role in the final SREP score assigned to banks, closer to the weight implied by the automatic indicators. Since 2021, when the process was revised with the aim of concentrating supervisory work on fewer priorities,[23] governance featured in every year among the European banking supervision supervisory priorities.

Other coefficients that are significant in the estimation, such as the operational risk (OP) score, are less robust when changing the estimation method. Finally, we note the common component explains roughly 20%-30% of the changes in the OSS supervisory judgement, suggesting that a significant proportion of judgement is bank-specific or idiosyncratic in nature.

[23] See Buch 2025a.

**Are 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 the first differencing estimation for the supervisory judgement of the OSS from equation (2) with interaction terms aimed to disentangle coefficient estimates for the element and combined scores in cases of: (i) upgrades, downgrades and no changes in the OSS; (ii) low and medium-low risk (OSS 1, 2+, 2 and 2-), medium-high and high risk (OSS 3+, 3, 3- and 4), and (iii) large and complex (cluster 1, 2 and 3) and small and less complex (cluster 4 and 5) banks.

For each of the three augmented estimation models we test the equality of the estimated coefficients for the interaction terms involving the element and combined risk scores. If the null hypothesis is rejected, it is interpreted as evidence supporting the corresponding sub-pattern in the supervisory adjustment of the OSS. We perform this statistical testing using both the Likelihood Ratio (LR) and Wald tests, and we conclude only when the two cross-validate their results. Estimation results are presented in Tables 2, 3 and 4.

Wald and LR tests in Table 2 indicate that the way business model assessment (BMA) and credit risk (CRED) scores are used as part of supervisory judgement are significantly different for the period 2021-2024 between cases when OSS is upgraded and cases when OSS is downgraded, with the strongest effect for credit risk. The BMA score is taken into account more in the cases of OSS upgrades, while CRED is focused on in cases of downgrades in the 2021-2024 sub-sample.

The result for CRED is particularly interesting, as lending is the main economic objective of banks. Therefore, considering the aforementioned limitations of the linear aggregation functional forms in composite indicators, such as the automatic SREP score starting point, the results in Table 2 suggest that supervisors have paid close attention to downside credit risk in the post-pandemic period, when deciding the score of a supervised bank. The reason could be that improvements in other risk areas could conceal deteriorations in credit risk when relying solely on the automatic SREP score and its equal weighting scheme, which would otherwise warrant a more prudent stance. This behaviour is in line with the main findings of the paper, that supervisory judgement is employed to incorporate additional information into SREP assessments.

Table 3 presents the selected coefficients obtained in the estimation of equation (2) which was augmented by adding variables interacting the change in the element (BMA, GOV, CAP and LIQ) and combined (CRED, MK, OP and IRRBB) scores with dummy variables differentiating between low and medium-low risk, respectively medium-high and high risk OSS in OSS. The null hypothesis that low/medium-low and high/medium-high risk banks are not treated

Table 2. OSS supervisory judgement. First-differencing (FD) estimation with interaction on the OSS upgrade/downgrade

| | Sample | | |
|---|---|---|---|
| | 2016-2024 | 2016-2020 | 2021-2024 |
| BMA Score | | | |
| OSS no change | -0.2172 | -0.2261 | -0.1963 |
| OSS upgrades | -0.2012 | -0.2348 | -0.0390 |
| OSS downgrades | -0.2720 | -0.3466 | -0.2453 |
| [prob. H0: OSS upgrades = OSS downgrades] | | | |
| Wald | [0.1742] | [0.1029] | [0.0197] |
| LR | [0.1434] | [0.0820] | [0.0075] |
| GOV Score | | | |
| OSS no change | -0.1438 | -0.1653 | -0.1467 |
| OSS upgrades | -0.0150 | -0.1478 | 0.0188 |
| OSS downgrades | -0.0874 | -0.1607 | -0.0643 |
| [prob. H0: OSS upgrades = OSS downgrades] | | | |
| Wald | [0.2075] | [0.9355] | [0.3518] |
| LR | [0.1862] | [0.9032] | [0.3206] |
| CAP Score | | | |
| OSS no change | -0.1368 | -0.1584 | -0.0915 |
| OSS upgrades | -0.1258 | -0.1929 | -0.1099 |
| OSS downgrades | -0.0429 | -0.0575 | -0.0209 |
| [prob. H0: OSS upgrades = OSS downgrades] | | | |
| Wald | [0.1007] | [0.2230] | [0.1955] |
| LR | [0.0630] | [0.0430] | [0.1600] |
| LIQ Score | | | |
| OSS no change | -0.1837 | -0.1829 | -0.2126 |
| OSS upgrades | -0.2949 | -0.5506 | -0.1054 |
| OSS downgrades | -0.2922 | -0.4709 | -0.2930 |
| [prob. H0: OSS upgrades = OSS downgrades] | | | |
| Wald | [0.9795] | [0.7042] | [0.1349] |
| LR | [0.9737] | [0.5143] | [0.1073] |
| CRED Score | | | |
| OSS no change | 0.1383 | 0.1261 | 0.1271 |
| OSS upgrades | 0.0515 | 0.1685 | -0.0558 |
| OSS downgrades | 0.1107 | 0.1172 | 0.1182 |
| [prob. H0: OSS upgrades = OSS downgrades] | | | |
| Wald | [0.3003] | [0.5530] | [0.0061] |
| LR | [0.2086] | [0.5533] | [0.0097] |

differently by supervisors when assigning the OSS cannot be rejected.

Table 4 presents the selected coefficient estimates obtained in the estimation of equation (2) which was augmented by adding variables interacting the change in the element (BMA, GOV, CAP and LIQ) and combined (CRED, MK, OP and IRRBB) with dummy variables

Table 2 (cont'd) OSS supervisory judgement. First-differencing (FD) estimation with interaction on the OSS upgrade/downgrade

| | Sample | | |
|---|---|---|---|
| | 2016-2024 | 2016-2020 | 2021-2024 |
| MK Score | | | |
| OSS no change | -0.0031 | 0.0056 | -0.0344 |
| OSS upgrades | -0.0477 | -0.0192 | -0.0550 |
| OSS downgrades | -0.0184 | 0.0424 | -0.0543 |
| [prob. H0: OSS upgrades = OSS downgrades] | | | |
| Wald | [0.5382] | [0.5187] | [0.9912] |
| LR | [0.4778] | [0.3645] | [0.9917] |
| OP Score | | | |
| OSS no change | 0.0256 | 0.0781 | -0.0154 |
| OSS upgrades | -0.0546 | 0.2451 | -0.2677 |
| OSS downgrades | 0.0409 | 0.2189 | -0.0087 |
| [prob. H0: OSS upgrades = OSS downgrades] | | | |
| Wald | [0.2636] | [0.8671] | [0.0124] |
| LR | [0.1369] | [0.8178] | [0.0038] |
| IRRBB Score | | | |
| OSS no change | -0.0020 | 0.0034 | 0.0108 |
| OSS upgrades | 0.0762 | 0.0513 | 0.0516 |
| OSS downgrades | 0.0755 | 0.1257 | 0.0525 |
| [prob. H0: OSS upgrades = OSS downgrades] | | | |
| Wald | [0.9896] | [0.4075] | [0.9905] |
| LR | [0.9885] | [0.2334] | [0.9909] |
| Macro controls | Yes | Yes | Yes |
| Key risk indicators | Yes | Yes | Yes |
| SREP cycle dummy variables | Yes | Yes | Yes |
| Observations | 452 | 160 | 292 |
| Banks | 94 | 62 | 91 |
| Adj. R-squared | 0.41 | 0.55 | 0.43 |

The table presents the selected coefficient estimates obtained in the estimation of equation (2) which was augmented by adding variables interacting the change in the element (BMA, GOV, CAP and LIQ) and risk (CRED, MK, OP and IRRBB) scores with dummy variables differentiating between upgrade ($\Delta OSS < 0$) and downgrade ($\Delta OSS > 0$) in OSS. The dependent variable is the supervisory judgement in OSS. In the brackets we report the probabilities associated with the null of equality between the coefficients of the interaction variables for the Wald and Likelihood Ratio (LR) tests.

differentiating larger and more complex institutions (Cluster 1-3), respectively smaller and less complex (Cluster 4-5). Looking at the size of the supervised institutions, we find no systematic evidence of statistical difference between the judgement adjustments done to small versus large

Table 3. OSS supervisory judgement. First-differencing (FD) estimation with interaction on the low and medium-low risk / medium-high and high risk OSS

| | Sample | | |
|---|---|---|---|
| | 2016-2024 | 2016-2020 | 2021-2024 |
| BMA Score | | | |
| Low and medium-low risk OSS | -0.1490 | -0.1001 | -0.1298 |
| High and medium-high risk OSS | -0.1988 | -0.2038 | -0.1777 |
| [prob. of H0: Low and medium-low risk OSS= High and medium-high risk OSS] | | | |
| Wald | [0.3602] | [0.4368] | [0.4505] |
| LR | [0.3043] | [0.2014] | [0.4348] |
| GOV Score | | | |
| Low and medium-low risk OSS | -0.0535 | -0.0771 | -0.0754 |
| High and medium-high risk OSS | -0.1328 | -0.3120 | -0.0897 |
| [prob. of H0: Low and medium-low risk OSS= High and medium-high risk OSS] | | | |
| Wald | [0.1633] | [0.1428] | [0.7837] |
| LR | [0.1095] | [0.0345] | [0.7985] |
| CAP Score | | | |
| Low and medium-low risk OSS | -0.1392 | -0.2298 | -0.0953 |
| High and medium-high risk OSS | -0.1111 | -0.1074 | -0.1179 |
| [prob. of H0: Low and medium-low risk OSS= High and medium-high risk OSS] | | | |
| Wald | [0.6948] | [0.3234] | [0.7999] |
| LR | [0.6190] | [0.1948] | [0.7335] |
| LIQ Score | | | |
| Low and medium-low risk OSS | -0.2123 | -0.2845 | -0.0882 |
| High and medium-high risk OSS | -0.1757 | -0.1325 | -0.2365 |
| [prob. of H0: Low and medium-low risk OSS= High and medium-high risk OSS] | | | |
| Wald | [0.6829] | [0.2954] | [0.1465] |
| LR | [0.5600] | [0.0628] | [0.0996] |
| CRED Score | | | |
| Low and medium-low risk OSS | 0.0735 | 0.1208 | 0.0499 |
| High and medium-high risk OSS | 0.1368 | 0.0232 | 0.1478 |
| [prob. of H0: Low and medium-low risk OSS= High and medium-high risk OSS] | | | |
| Wald | [0.3230] | [0.4568] | [0.1625] |
| LR | [0.1333] | [0.2111] | [0.0493] |

banking institutions.

**What type of information features prominently in the common supervisory judgement?**

According to the SREP methodology, combined risk scores (CRED, MK, OP and IRRBB) are based on quantitative information (e.g., KRIs that enter the calculation of the corresponding

Table 3 (cont'd) OSS supervisory judgement. First-differencing (FD) estimation with interaction on the low and medium-low risk / medium-high and high risk OSS

| | Sample | | |
|---|---|---|---|
| | 2016-2024 | 2016-2020 | 2021-2024 |
| MK Score | | | |
| Low and medium-low risk OSS | 0.0644 | 0.0303 | 0.0313 |
| High and medium-high risk OSS | 0.0163 | 0.2053 | -0.0198 |
| [prob. of H0: Low and medium-low risk OSS= High and medium-high risk OSS] | | | |
| Wald | [0.1638] | [0.2135] | [0.2389] |
| LR | [0.2020] | [0.0731] | [0.2088] |
| OP Score | | | |
| Low and medium-low risk OSS | -0.0130 | 0.1550 | -0.1108 |
| High and medium-high risk OSS | 0.1201 | 0.1267 | 0.0772 |
| [prob. of H0: Low and medium-low risk OSS= High and medium-high risk OSS] | | | |
| Wald | [0.0101] | [0.7963] | [0.0026] |
| LR | [0.0078] | [0.7301] | [0.0017] |
| IRRBB Score | | | |
| Low and medium-low risk OSS | 0.0296 | -0.0044 | 0.0065 |
| High and medium-high risk OSS | 0.0198 | 0.1471 | 0.0380 |
| [prob. of H0: Low and medium-low risk OSS= High and medium-high risk OSS] | | | |
| Wald | [0.7970] | [0.1505] | [0.4883] |
| LR | [0.7975] | [0.0523] | [0.4970] |
| Macro controls | Yes | Yes | Yes |
| Key risk indicators | Yes | Yes | Yes |
| SREP cycle dummy variables | Yes | Yes | Yes |
| Observations | 474 | 162 | 312 |
| Banks | 99 | 63 | 98 |
| Adj. R-squared | 0.25 | 0.33 | 0.31 |

The table presents the coefficients estimates from the regression in equation (2). To assess the impact of risk brackets on supervisory judgement, interaction terms are added between the change in the viability score (BMA, GOV, CAP and LIQ), as well as risk scores (CRED, MK, OP and IRRBB) and dummy variables differentiating between low and medium-low risk OSS ('1' and '2' including qualifiers) and medium-high and high risk OSS ('3' including qualifiers and '4'). The dependent variable is the supervisory judgement in OSS. In the brackets we report the probabilities associated with the null of equality between the coefficients of the interaction variables for the Wald and Likelihood Ratio (LR) tests.

automatic scores), as well as on information of a qualitative nature (e.g., corresponding risk control scores). We can test which type of information features directly in the common component of the adjustment of OSS by supervisors by replacing in the OSS supervisory judgement estimation (equation 2) the individual risk scores with the corresponding risk control scores.

Table 4. OSS supervisory judgement. First-differencing (FD) estimation with interaction on the large/small banks

| | Sample | | |
|---|---|---|---|
| | 2016-2024 | 2016-2020 | 2021-2024 |
| BMA Score | | | |
| Cluster 1-3 | -0.1459 | -0.0077 | -0.1535 |
| Cluster 4-5 | -0.1741 | -0.2326 | -0.0733 |
| [prob. of H0: Cluster 1-3=Cluster 4-5] | | | |
| Wald | [0.7427] | [0.1032] | [0.3907] |
| LR | [0.6272] | [0.0717] | [0.2497] |
| GOV Score | | | |
| Cluster 1-3 | -0.1245 | -0.1340 | -0.1162 |
| Cluster 4-5 | -0.0388 | -0.1730 | -0.0298 |
| [prob. of H0: Cluster 1-3=Cluster 4-5] | | | |
| Wald | [0.2938] | [0.7443] | [0.2859] |
| LR | [0.1510] | [0.7250] | [0.2313] |
| CAP Score | | | |
| Cluster 1-3 | -0.0261 | -0.0754 | -0.0068 |
| Cluster 4-5 | -0.1620 | -0.1341 | -0.1449 |
| [prob. of H0: Cluster 1-3=Cluster 4-5] | | | |
| Wald | [0.1116] | [0.6942] | [0.0833] |
| LR | [0.0246] | [0.5940] | [0.0589] |
| LIQ Score | | | |
| Cluster 1-3 | -0.1752 | -0.1434 | -0.1935 |
| Cluster 4-5 | -0.1811 | -0.1717 | -0.1864 |
| [prob. of H0: Cluster 1-3=Cluster 4-5] | | | |
| Wald | [0.9643] | [0.8743] | [0.9504] |
| LR | [0.9335] | [0.7699] | [0.9447] |
| CRED Score | | | |
| Cluster 1-3 | 0.1072 | 0.0812 | 0.0961 |
| Cluster 4-5 | 0.1068 | 0.1714 | 0.0724 |
| [prob. of H0: Cluster 1-3=Cluster 4-5] | | | |
| Wald | [0.9940] | [0.5031] | [0.7448] |
| LR | [0.9923] | [0.3181] | [0.6871] |

The quantitative information is already included as KRIs, bank controls and macro variables. Therefore, significant coefficients for the risk control scores in the resulting estimation would indicate that qualitative information is the one that features prominently in common supervisory judgement.

The results are presented in Table 5. Credit and operational risk control scores have significant positive coefficients, indicating that additional information considered in the supervisory

Table 4 (cont'd) OSS supervisory judgement. First-differencing (FD) estimation with interaction on the large/small banks

| | Sample | | |
|---|---|---|---|
| | 2016-2024 | 2016-2020 | 2021-2024 |
| MK Score | | | |
| Cluster 1-3 | 0.0216 | 0.1524 | 0.0036 |
| Cluster 4-5 | 0.0251 | 0.1162 | -0.0011 |
| [prob. of H0: Cluster 1-3=Cluster 4-5] | | | |
| Wald | [0.9631] | [0.7939] | [0.9535] |
| LR | [0.9568] | [0.7617] | [0.9493] |
| OP Score | | | |
| Cluster 1-3 | 0.0620 | 0.0523 | 0.0490 |
| Cluster 4-5 | 0.0684 | 0.2106 | -0.0197 |
| [prob. of H0: Cluster 1-3=Cluster 4-5] | | | |
| Wald | [0.9184] | [0.1863] | [0.3423] |
| LR | [0.9285] | [0.1459] | [0.4428] |
| IRRBB Score | | | |
| Cluster 1-3 | 0.0387 | 0.0451 | 0.0387 |
| Cluster 4-5 | 0.0161 | 0.0191 | 0.0147 |
| [prob. of H0: Cluster 1-3=Cluster 4-5] | | | |
| Wald | [0.6585] | [0.8107] | [0.6597] |
| LR | [0.6247] | [0.7977] | [0.6557] |
| Macro controls | Yes | Yes | Yes |
| Key risk indicators | Yes | Yes | Yes |
| SREP cycle dummy variables | Yes | Yes | Yes |
| Observations | 474 | 162 | 312 |
| Banks | 99 | 63 | 98 |
| Adj. R-squared | 0.20 | 0.26 | 0.23 |

The table presents the coefficients estimates from the regression in equation (2). To assess the impact of size and complexity of banking groups on supervisory judgement, interaction terms are added between the change in the viability score (BMA, GOV, CAP and LIQ), as well as risk scores (CRED, MK, OP and IRRBB) and dummy variables differentiating larger and more complex institutions, respectively smaller and less complex ones. The dependent variable is the supervisory judgement in OSS. In the brackets we report the probabilities associated with the null of equality between the coefficients of the interaction variables for the Wald and Likelihood Ration (LR) tests.

judgement is linked to qualitative (risk control) aspects of the two risks.

Table 5. OSS supervisory judgement. Alternative specification of the first-differencing (FD) estimation

| | Sample | | |
|---|---|---|---|
| | 2016-2024 | 2016-2020 | 2021-2024 |
| BMA Score | -0.1766*** | -0.1923** | -0.1382*** |
| | (-4.37) | (-2.49) | (-3.10) |
| GOV Score | -0.1000** | -0.1158 | -0.1005** |
| | (-2.28) | (-1.59) | (-2.42) |
| CAP Score | -0.0807* | -0.0928 | -0.0688 |
| | (-1.77) | (-1.40) | (-1.33) |
| LIQ Score | -0.1612** | -0.1752* | -0.1257* |
| | (-2.51) | (-1.93) | (-1.74) |
| CRED Risk Control Score | 0.0684** | 0.1307 | 0.0494 |
| | (2.14) | (1.46) | (1.39) |
| MK Risk Control Score | 0.0117 | 0.0173 | -0.0320 |
| | (0.32) | (0.24) | (-0.91) |
| OP Risk Control Score | 0.0896*** | 0.2084*** | 0.0340 |
| | (2.97) | (3.37) | (1.14) |
| IRRBB Risk Control Score | 0.0299 | 0.0030 | 0.0477 |
| | (1.03) | (0.05) | (1.58) |
| Macro controls | Yes | Yes | Yes |
| Key Risk Indicators | Yes | Yes | Yes |
| SREP cycle dummy variables | Yes | Yes | Yes |
| Observations | 455 | 161 | 294 |
| Banks | 99 | 63 | 98 |
| Adj. R-squared | 0.20 | 0.32 | 0.24 |

The table presents the main results of the estimation based on equation (2), in which combined risk scores (CRED, MK, OP and IRRBB) are replaced in the list of explanatory variables with the corresponding risk control scores. The dependent variable is the supervisory judgement in OSS. Robust t statistics, calculated based on standard errors clustered by banks, in parentheses. * p<0.10, ** p<0.05, *** p<0.01.

### What was driving P2R?[24]

Results of the estimates for the P2R model in difference (equation 3) are presented in Table 6. The overall score coefficient is significant for the full sample estimation and for the sub-period starting with 2021, but not before, which could suggest that the link between the changes in the risk profile and those in P2R became stronger starting with 2021, when the methodology for setting P2R changed. SREP scores also seem to play a direct role: business model assessment (BMA) for the full sample, internal governance (GOV) in the recent years, as well as internal

[24] Our analysis covers the time interval 2016-2024. A new P2R methodology is applied starting with SREP 2026 cycle (Buch 2025b).

capital adequacy (ICAAP) score for the entire sample and for the sub-sample starting in 2021. This is consistent with the prominent role played by the ICAAP score in the P2R methodology change implemented in 2021. The additional coefficients which are statistically significant have relatively small values, suggesting a lower contribution to the supervisory judgment in P2R setting. The common component explains roughly 30% of the changes in the P2R. Similar to supervisory judgment for the OSS, the majority of this component remains bank-specific.

Table 6. P2R. First-differencing (FD) estimation

| Dependent variable: | Sample | | |
|---|---|---|---|
| P2R | 2016-2024 | 2016-2020 | 2021-2024 |
| Overall Score | 0.2713*** | 0.0202 | 0.3097*** |
| | (3.49) | (0.10) | (3.50) |
| BMA Score | 0.0870** | 0.2696 | 0.0556* |
| | (2.19) | (1.49) | (1.95) |
| GOV Score | 0.0276 | -0.0106 | 0.0473** |
| | (0.68) | (-0.07) | (2.04) |
| CAP Score | 0.0347 | 0.0419 | 0.0517 |
| | (1.26) | (0.50) | (1.64) |
| LIQ Score | 0.1003 | 0.1646* | 0.0273 |
| | (1.65) | (1.85) | (0.44) |
| CRED Score | 0.0140 | -0.0038 | 0.0198 |
| | (0.61) | (-0.08) | (0.79) |
| MK Score | 0.0156 | 0.1015 | -0.0153 |
| | (0.80) | (1.63) | (-0.79) |
| OP Score | 0.0321 | -0.0361 | 0.0627* |
| | (0.76) | (-0.43) | (1.89) |
| IRRBB Score | -0.0130 | 0.0425 | -0.0153 |
| | (-0.79) | (0.72) | (-0.79) |
| ICAAP Score | 0.0701*** | 0.0773 | 0.0834*** |
| | (3.45) | (1.53) | (2.78) |
| Macro controls | Yes | Yes | Yes |
| Key Risk Indicators | Yes | Yes | Yes |
| SREP cycle dummy variables | Yes | Yes | Yes |
| Observations | 465 | 159 | 306 |
| Banks | 99 | 62 | 98 |
| Adj. R-squared | 0.30 | 0.31 | 0.39 |

The table presents the main results of the estimation based on equation (3), in which the dependent variable is the yearly change in the P2R expressed in percentage. Scores are employed in their numerical form, with 1 representing the full difference between two main score categories (e.g., '2' and '3'). The difference accounting for score categories with qualifiers amounts to roughly one third. Robust t statistics, calculated based on standard errors clustered by banks, in parentheses. * p<0.10, ** p<0.05, *** p<0.01. A more detailed version of the results is available in the Annex (Table Table A.4).

**Robustness checks**

We conduct robustness checks along several dimensions. First, we analyse the structure of the correlation between supervisory scores. This is shown in Table A.2. Total correlation coefficients range from approximately 0.3 to 0.6 for the viability scores (BMA, GOV, CAP and LIQ) and from 0.0 to 0.1 for the intrinsic risk scores (CRED, MK, OP and IRRBB), driven by the 'between' correlation, which is much higher than the 'within' correlation. This is consistent with the existence of bank-specific unobservable factors that are correlated with the supervisory scores, especially the viability scores. The values of the correlation coefficients, statistically significant in almost all cases, indicate that each supervisory score we employ as explanatory variable has additional informational content in the regression, especially the intrinsic risk scores. Together with the previously mentioned serial correlation in the variables, this argues in favour of the preferred estimation technique employed for our study, which is first difference estimation.

To ensure the robustness of the findings, we test alternative estimation methods. Tables A.5 and A.6 present the results of the estimation using pooled OLS (POLS), random effects (RE), fixed effects (FE), correlated random effects (CRE) and fixed effects assuming AR (1) serial correlation in the idiosyncratic errors $\{u_{it}\}$ (FEAR).

The results are qualitatively similar to those based on first-differencing, with the difference in the size of the estimated coefficients likely explained by the underlying assumptions of the POLS and RE estimators. More detailed results are available in the appendices.

Second, in order to account for the fact scores are categorical variables and may have an ordinal interpretation, we estimate panel ordered logit and probit having supervisory judgement as the left-hand side (LHS) variable. The panel ordered logit and probit estimations also confirm the key findings of the results of the model for the OSS supervisory judgement. Estimation results are presented in Tables A.7 and A.8. The statistically significant estimates for the changes in viability scores (BMA, GOV, CAP and LIQ) suggest that if these scores increase (worsen from a risk perspective), there is a higher likelihood of a smaller (possibly negative) OSS expert judgement adjustment, which, in fact, compensates the increase in OSS as a result of the automatic score, which would otherwise have been higher. In a similar manner, when the CRED score worsens, there is higher likelihood to have a higher (possibly positive) OSS expert judgement adjustment, in the direction of a worse overall score.

Third, a score's level can be significantly influenced by its past realization, implying an autoregressive formulation for the variable of interest. To capture this persistence, we specify for both OSS and P2R a dynamic model, for which we employ an Arellano-Bond estimator. This comes with the caveat that it may produce less precise estimates due to limited temporal data, which constrains the lag order of the model and affects the estimated variance when considering the large number of regressors that need to be controlled for.

A dynamic model estimated for the OSS confirms the key findings of the supervisory judgment model for the OSS. Estimation results are presented in Table A.9. As expected, the coefficient of the OSS lag is significant. However, the lag coefficient, estimated at 0.3–0.35, does not indicate a high degree of persistence. Estimating the same model without risk scores yields a lag coefficient of around 0.6-0.8, which aligns with the annual variability observed in the OSS.[25] Taken together, the two results suggest that approximately half of the persistence in the OSS is attributable to the stability of the risk scores. All four element scores have estimated coefficients significantly different from 0, but also significantly lower than the 0.25 implied by their contribution to the automatic score. The CRED score coefficient is also significant, as in the supervisory judgement model. Additionally, the NPL ratio and the size of the balance sheet (with a negative sign, indicating that, ceteris paribus, higher total assets imply better OSS) are significant.

The P2R dynamic model confirms the role played by the overall risk profile and the ICAAP score in the determination of P2R. Estimation results are presented in Table A.10. The P2R own lag is significant, with the estimated coefficient ranging between 0.3-0.6 depending on the sample, which indicates a greater persistence in the P2R compared to the OSS. Overall score coefficient is significant for all three estimated samples and, as was the case in the first difference model, the ICAAP score is significant for 2021-2024 and for the entire sample. Additionally, the size of the balance sheet coefficient is significant for the entire sample, with negative sign. One caveat of this analysis is that qualitative measures (e.g., number and severity) are not considered. They are used by the supervisors to complement or to add to the quantitative measures, represented by the P2R.

Finally, we explore how sensitive our results are to the set of explanatory variables. Tables A.11 and A.12 present the results of the first-differencing estimation for the OSS supervisory judgement and the P2R using the following explanatory variables: (i) only the supervisory scores; (ii) supervisory scores and macro variables; (iii) supervisory scores and bank indicators. The results indicate that, compared to a more parsimonious option that would include only the supervisory scores, employing a wider range of explanatory variables does indeed help improve the accuracy in the estimation of some coefficients (e.g., such as business model assessment – BMA, in the OSS supervisory judgment estimation, and capital adequacy – CAP, in the P2R estimation).

[25]For example, in SREP 2024 74% of OSS remained unchanged (European Central Bank 2024a).

# 4 Concluding remarks

Leveraging on a novel supervisory dataset spanning nine evaluation cycles (SREP), we conduct the first thorough quantitative analysis of determinants driving changes in the overall supervisory score (OSS) and P2R across significant banking institutions in the euro area. We focus our analysis on supervisory judgement, as stylised facts show that it plays an important role in the OSS, and we identify commonalities in how joint supervisory teams across the SSM universe assess risks throughout multiple SREP cycles.

Panel data estimates of changes in the supervisory judgement component of the OSS show significant negative coefficients for the four SREP element scores, alongside positive coefficients for additional risk indicators. This suggests supervisors dampen the relative contribution of the standard SREP elements to the overall bank scores, in favour of incorporating additional relevant information into the overall assessment of bank riskiness.

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.

Specifically, we find a significant coefficient for credit risk, indicating that information linked to this risk, which we further analyze to encompass qualitative risk control aspects, is deemed directly relevant by supervisors for the assessment of the overall viability of the supervised institution.

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 sub-sample). 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, 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.

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.

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.

To our knowledge, this is the first study to empirically confirm the impact of supervisory judgement on supervisory outcomes within the euro area. We provide evidence that European supervisors actively use judgement when setting the OSS and the P2R capital requirements. Our results can help bridge gaps in understanding caused by asymmetric information among various stakeholders regarding the process.

We confirm the existence of a *common* supervisory judgement channel that has implications for how supervisory judgement should be analysed empirically. It highlights the role of supervisory judgement in providing the flexibility needed not only at the level of individual banks, but also at the level of the entire system, to reflect, for example, supervisory priorities.

In addition, 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.

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.

## References

Abbassi, Puriya et al. (2023). *Stressed Banks? Evidence from the Largest-Ever Supervisory Review*. Finance and Economics Discussion Series 2023-021. Board of Governors of the Federal Reserve System (U.S.)

Agarwal, S., D. Luca, et al. (2014). "Inconsistent Regulators: Evidence from Banking". In: *The Quarterly Journal of Economics* 129 (2), pp. 889–938.

Agarwal, S., B. C. Morais, et al. (2024). *Noisy Experts? Discretion in Regulation*. Working Paper 32344. NBER. URL: https://www.nber.org/papers/w32344.

Balan, M. et al. (2025). *Act early or pay later: the role of qualitative measures in effective supervisory frameworks*. Insights on policy implementation 66. FSI. URL: https://www.bis.org/fsi/publ/insights66.pdf.

Basel Committee on Banking Supervision (2019). *Supervisory review process*. URL: https://www.bis.org/basel_framework/standard/SRP.htm?tldate=20250422.

Buch, C. (2024). *"Introductory statement" at the press conference on the 2024 SREP results and the supervisory priorities for 2025-27*. URL: https://www.bankingsupervision.europa.eu/press/speeches/date/2024/html/ssm.sp241217~d7af2d9e29.en.html.

— (2025a). *Comments on the paper by Sumit Agarwal, Bernardo Morais, Amit Seru and Kelly Shue entitled "Weighted Noise: Discretion in Regulation" contribution at the tenth ECB Annual Research Conference organised jointly with Stanford University's Hoover Institution on "The Next Financial Crisis?* URL: https://www.bankingsupervision.europa.eu/press/speeches/date/2025/html/ssm.sp250918~5e9f04035e.en.html.

— (2025b). *Reviewing the Pillar 2 requirement methodology*. URL: https://www.bankingsupervision.europa.eu/press/blog/2025/html/ssm.blog20250310~0a07f5d777.en.html.

Carletti, E. et al. (2016). *Supervisory Incentives in a Banking Union*. Working Paper 16/186. IMF.

Dewatripont, M. and J. Tirole (1994). *The Prudential Regulation of Banks*. The MIT Press.

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/supervisory-review-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.

— (2016). *SSM SREP Methodology Booklet – 2016 edition*. URL: https://www.bankingsupervision.europa.eu/ecb/pub/pdf/srep_methodology_booklet_2016.en.pdf.

— (2017). *SSM SREP Methodology Booklet – 2017 edition*. URL: https://www.bankingsupervision.europa.eu/ecb/pub/pdf/ssm.srep_methodology_booklet_2017.en.pdf.

European Central Bank (2018). *ECB Guide to the internal capital adequacy assessment process (ICAAP)*. URL: https://www.bankingsupervision.europa.eu/ecb/pub/pdf/ssm.icaap_guide_201811.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.

— (2021). *Supervisory methodology 2020*. URL: https://www.bankingsupervision.europa.eu/activities/srep/2021/html/ssm.srep202101_supervisorymethodology2021.en.html.

— (2022). *Supervisory methodology 2021*. URL: https://www.bankingsupervision.europa.eu/activities/srep/2022/html/ssm.srep202202_supervisorymethodology2022.en.html.

— (2023a). *Supervisory methodology 2022*. URL: https://www.bankingsupervision.europa.eu/activities/srep/2023/html/ssm.srep202302_supervisorymethodology2023.en.html.

— (2023b). *Supervisory methodology 2023*. URL: https://www.bankingsupervision.europa.eu/activities/srep/2023/html/ssm.srep202312_supervisorymethodology2023.en.html.

— (2024a). *Aggregated results of SREP 2024*. URL: https://www.bankingsupervision.europa.eu/activities/srep/2024/html/ssm.srep202412_aggregatedresults2024.en.html.

— (2024b). *Supervisory Manual*. URL: https://www.bankingsupervision.europa.eu/ecb/pub/pdf/ssm.supervisory_guides202401_manual.en.pdf.

— (2024c). *Supervisory methodology 2024*. URL: https://www.bankingsupervision.europa.eu/activities/srep/2024/html/ssm.srep202412_supervisorymethodology2024.en.html.

Expert Group to the Chair of the Supervisory Board of the ECB (2023). *Assessment of the European Central Bank's Supervisory Review and Evaluation Process*. URL: https://www.bankingsupervision.europa.eu/ecb/pub/pdf/annex/ssm.pr230417_annex.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-perf-reg/pub-econ-working-paper-camels-ratings.pdf.

Hirtle, B. and A. Kovner (2021). *Bank Supervision*. Staff Reports 952. Federal Reserve Bank of New York.

Hirtle, B., A. Kovner, and M. Plosser (2020). "The Impact of Supervision on Bank Performance". In: *The Journal of Finance* 75 (5), pp. 2765–2808.

Montagner, P. (2025). *SREP reform: towards simpler and more effective supervision*. URL: https://www.bankingsupervision.europa.eu/press/interviews/date/2025/html/ssm.in250703~0f642ce705.en.html.

Wooldridge, J. M. (2002). *Econometric Analysis of Cross Section and Panel Data*. The MIT Press.

# Appendices

Table A.1 Descriptive statistics

| | N | Mean | Std. dev. | Min. | p25 | Median | p75 | Max. |
|---|---|---|---|---|---|---|---|---|
| *SREP scores* | | | | | | | | |
| Overall SREP score | 981 | 2.64 | 0.60 | 1.00 | 2.30 | 2.60 | 3.00 | 4.00 |
| BMA score | 981 | 2.58 | 0.67 | 1.00 | 2.00 | 3.00 | 3.00 | 4.00 |
| GOV score | 981 | 2.86 | 0.52 | 1.00 | 3.00 | 3.00 | 3.00 | 4.00 |
| CAP score | 981 | 2.33 | 0.76 | 1.00 | 2.00 | 2.00 | 3.00 | 4.00 |
| LIQ score | 981 | 2.14 | 0.58 | 1.00 | 2.00 | 2.00 | 2.00 | 4.00 |
| CRED score | 981 | 2.65 | 0.74 | 1.00 | 2.00 | 3.00 | 3.00 | 4.00 |
| MK score | 981 | 2.09 | 0.63 | 0.00 | 2.00 | 2.00 | 2.00 | 4.00 |
| OP score | 981 | 2.94 | 0.53 | 1.00 | 3.00 | 3.00 | 3.00 | 4.00 |
| IRRBB score | 981 | 2.34 | 0.55 | 1.00 | 2.00 | 2.00 | 3.00 | 4.00 |
| ICAAP score | 975 | 2.56 | 0.54 | 1.00 | 2.00 | 3.00 | 3.00 | 4.00 |
| *Bank Indicators* | | | | | | | | |
| Ln(TA) | 978 | 4.27 | 1.47 | 0.85 | 3.51 | 4.17 | 5.11 | 7.79 |
| Loan book | 978 | 65.43 | 17.78 | 0.00 | 58.35 | 67.71 | 76.29 | 100 |
| Government loans ratio | 978 | 6.58 | 14.03 | 0.00 | 0.57 | 2.14 | 5.43 | 94.80 |
| NPL ratio | 978 | 5.44 | 8.78 | 0.00 | 1.26 | 2.67 | 4.97 | 59.05 |
| Coverage ratio | 887 | 39.03 | 17.88 | 0.00 | 29.12 | 40.71 | 49.88 | 100.00 |
| RWA density | 980 | 40.04 | 17.40 | 2.19 | 29.40 | 37.87 | 49.31 | 150.06 |
| Credit risk share of RWA | 980 | 84.51 | 9.25 | 34.36 | 82.19 | 86.93 | 90.19 | 99.92 |
| Market risk share of RWA | 980 | 2.75 | 4.42 | 0.00 | 0.17 | 1.33 | 3.63 | 38.11 |
| Operational risk share of RWA | 980 | 10.11 | 5.43 | 0.08 | 6.98 | 9.02 | 11.50 | 43.16 |
| NSFR | 849 | 129.66 | 33.48 | 70.71 | 111.97 | 124.21 | 138.04 | 460.55 |
| LCR | 838 | 214.24 | 130.56 | 37.94 | 141.48 | 172.45 | 228.59 | 1,000.00 |
| Deposits ratio | 980 | 75.63 | 23.03 | 0.00 | 66.79 | 83.63 | 92.39 | 99.59 |
| L2D | 968 | 122.97 | 152.88 | 0.00 | 77.38 | 93.87 | 110.25 | 1,658.35 |

Table A.1 (cont'd) Descriptive statistics

| | N | Mean | Std. dev. | Min. | p25 | Median | p75 | Max. |
|---|---|---|---|---|---|---|---|---|
| *Bank Indicators* | | | | | | | | |
| Cost-to-Income | 761 | 62.01 | 18.17 | 9.05 | 51.09 | 61.22 | 71.68 | 140.00 |
| RoE | 887 | 5.31 | 9.25 | -68.10 | 3.01 | 5.89 | 9.09 | 48.43 |
| RoA | 887 | 0.43 | 0.82 | -5.33 | 0.18 | 0.39 | 0.71 | 8.91 |
| NII in Total assets | 789 | 0.42 | 0.69 | -0.29 | 0.00 | 0.09 | 0.65 | 5.16 |
| Change in EVE 200 bps | 709 | -47.08 | 26.01 | -100.00 | -67.80 | -46.26 | -22.74 | -10.00 |
| *Macro variables* | | | | | | | | |
| GDP growth | 981 | 1.87 | 3.62 | -10.36 | 0.51 | 1.63 | 3.30 | 20.55 |
| Inflation | 981 | 2.58 | 3.74 | -2.17 | 0.50 | 1.37 | 3.16 | 34.15 |
| Ln (stock index) | 977 | 8.38 | 1.30 | 3.84 | 7.41 | 8.61 | 9.34 | 10.28 |
| VSTOXX | 981 | 18.62 | 3.80 | 12.73 | 15.03 | 19.45 | 21.75 | 23.15 |

This table presents the descriptive statistics for the data used in the estimations. SREP scores are employed in their numerical form, with 1 representing the full difference between two main score categories (e.g., '2' and '3'). The difference accounting for score categories with qualifiers amounts to roughly one third. Except for variables in log, all others are expressed in percentages. Source of data is supervisory data set for scores, including bank indicators that are calculated based on standardised financial and regulatory reporting (FINREP/COREP) and SDW for macro variables. Differences in the total number of observations are the result of some banks not being included in the SREP sample in every year (for example as in the respective years they were not classified as "significant institutions") or of some data series having no observations in certain years for some banks.

Table A.2 Pairwise correlation between supervisory scores

| | BMA | GOV | CAP | LIQ | CRED | MK | OP | IRRBB |
|---|---|---|---|---|---|---|---|---|
| BMA | | | | | | | | |
| Total | | 0.3624*** | 0.5831*** | 0.4223*** | 0.5011*** | 0.0639** | 0.3547*** | 0.0661** |
| Within | 1.0000 | 0.1413*** | 0.2872*** | 0.2138*** | 0.2434*** | -0.0030 | 0.1289*** | -0.1215*** |
| Between | | 0.4825*** | 0.6639*** | 0.4968*** | 0.5889*** | 0.0917*** | 0.4561*** | 0.1669*** |
| GOV | | | | | | | | |
| Total | | | 0.3727*** | 0.3134*** | 0.4029*** | 0.1272*** | 0.3639*** | 0.1719*** |
| Within | | 1.0000 | 0.1822*** | 0.1842*** | 0.1565*** | 0.0827*** | 0.2429*** | 0.0072 |
| Between | | | 0.4754*** | 0.3925*** | 0.5465*** | 0.1571*** | 0.4476*** | 0.3105*** |
| CAP | | | | | | | | |
| Total | | | | 0.4864*** | 0.4217*** | 0.1816*** | 0.3826*** | 0.0581* |
| Within | | | 1.0000 | 0.1574*** | 0.2447*** | 0.0400 | 0.1699*** | -0.0254 |
| Between | | | | 0.5975*** | 0.7343*** | 0.1099*** | 0.4747*** | 0.1017*** |
| LIQ | | | | | | | | |
| Total | | | | | 0.4217*** | 0.1816*** | 0.3200*** | 0.1033*** |
| Within | | | | 1.0000 | 0.1193*** | 0.0027 | 0.1475*** | -0.0379 |
| Between | | | | | 0.5420*** | 0.2673*** | 0.4073*** | 0.1927*** |
| CRED | | | | | | | | |
| Total | | | | | | 0.0265 | 0.3404 | 0.0548* |
| Within | | | | | 1.0000 | 0.0810** | 0.1715*** | -0.0272 |
| Between | | | | | | 0.0017 | 0.4237*** | 0.1050 |
| MK | | | | | | | | |
| Total | | | | | | | 0.1170*** | 0.0282 |
| Within | | | | | | 1.0000 | 0.0016 | 0.0329 |
| Between | | | | | | | 0.1829*** | 0.0253 |
| OP | | | | | | | | |
| Total | | | | | | | | 0.1419** |
| Within | | | | | | | 1.0000 | 0.0845*** |
| Between | | | | | | | | 0.1854*** |
| IRRBB | | | | | | | | |
| | | | | | | | | 1.0000 |

This table presents the correlation structure in the supervisory scores used as right-hand-side variables. First line represents the total correlation (i.e., $corr[x_{it}, y_{it}]$ for each pair of variables $\{x_{it}, y_{it}\}$), second line the 'within' correlation (i.e., $corr[x_{it} - \overline{x}_i, y_{it} - \overline{y}_i]$, where $\overline{x}_i = \sum_t x_{it}$) and the third line the 'between' correlation (i.e., $corr[\overline{x}_i, \overline{y}_i]$). *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively. The null hypothesis of no serial correlation in the first-order serial correlation test proposed by Wooldridge 2002 is rejected with a 1% level of significance.

Table A.3 OSS supervisory judgement. First-differencing (FD) estimation

| Dependent variable: OSS Supervisory judgement | Sample | | |
|---|---|---|---|
| | 2016-2024 | 2016-2020 | 2021-2024 |
| BMA Score | -0.1509*** | -0.1718* | -0.1122*** |
| | (-3.93) | (-1.99) | (-2.63) |
| GOV Score | -0.0836** | -0.1543** | -0.0742* |
| | (-2.04) | (-2.18) | (-1.90) |
| CAP Score | -0.0906** | -0.1186 | -0.0833* |
| | (-2.05) | (-1.51) | (-1.96) |
| LIQ Score | -0.1765*** | -0.1645* | -0.1833** |
| | (-2.69) | (-1.78) | (-2.58) |
| CRED Score | 0.1127*** | 0.1225** | 0.0890** |
| | (3.26) | (2.13) | (2.22) |
| MK Score | 0.0240 | 0.1036 | -0.0039 |
| | (0.77) | (1.40) | (-0.11) |
| OP Score | 0.0678* | 0.1420** | 0.0042 |
| | (1.86) | (2.22) | (0.10) |
| IRRBB Score | 0.0252 | 0.0543 | 0.0264 |
| | (0.94) | (0.87) | (0.91) |
| GDP growth | -0.0014 | 0.0094 | -0.0006 |
| | (-0.30) | (0.83) | (-0.12) |
| Inflation | -0.0048 | 0.0317 | -0.0067 |
| | (-1.06) | (1.11) | (-1.31) |
| Ln (stock index) | 0.1191 | -0.0152 | 0.1346 |
| | (0.78) | (-0.07) | (0.71) |
| VSTOXX | 0.0024 | 0.0029 | 0.0064 |
| | (0.55) | (0.31) | (1.22) |
| Ln (Total assets) | 0.0576 | -0.0035 | 0.0987 |
| | (0.86) | (-0.02) | (0.89) |
| Loan book | 0.0047** | 0.0013 | 0.0091*** |
| | (2.17) | (0.33) | (2.75) |
| Government loans ratio | -0.0094 | -0.0062 | -0.0021 |
| | (-1.59) | (-0.77) | (-0.22) |
| NPL ratio | 0.0075 | 0.0130** | 0.0036 |
| | (1.19) | (2.15) | (0.44) |
| Coverage ratio | 0.0016 | 0.0037*** | -0.0002 |
| | (1.56) | (3.50) | (-0.16) |
| RWA density | -0.0065 | -0.0055 | -0.0042 |
| | (-1.50) | (-1.33) | (-0.74) |
| Credit risk share of RWA | 0.0013 | -0.0052 | 0.0024 |
| | (0.48) | (-0.77) | (0.71) |
| Market risk share of RWA | -0.0043 | 0.0062 | -0.0089 |
| | (-0.64) | (0.49) | (-1.10) |
| Operational risk share of RWA | 0.0057 | -0.0050 | 0.0015 |
| | (0.80) | (-0.42) | (0.18) |

Table A.3 (cont'd) OSS supervisory judgment. First-differencing (FD) estimation

| Dependent variable: | Sample | | |
|---|---|---|---|
| OSS Supervisory judgement | 2016-2024 | 2016-2020 | 2021-2024 |
| NSFR | 0.0004 | -0.0002 | 0.0010 |
| | (0.62) | (-0.34) | (1.38) |
| LCR | -0.0003** | 0.0001 | -0.0005*** |
| | (-2.44) | (0.41) | (-3.01) |
| Deposits ratio | -0.0011 | -0.0002 | -0.0028 |
| | (-0.34) | (-0.04) | (-0.70) |
| L2D | -0.0004*** | -0.0003* | -0.0006*** |
| | (-2.57) | (-1.71) | (-2.86) |
| Cost to Income | -0.0020** | -0.0032*** | -0.0007 |
| | (-2.09) | (-3.74) | (-0.48) |
| RoE | 0.0012 | 0.0035 | -0.0098* |
| | (0.27) | (1.05) | (-1.69) |
| RoA | -0.0835 | -0.0351 | 0.0150 |
| | (-1.36) | (-0.56) | (0.19) |
| NII in Total assets | -0.0113 | 0.0048 | -0.0270 |
| | (-0.77) | (0.25) | (-1.40) |
| Change in EVE 200 bps | 0.0004 | 0.0002 | 0.0008* |
| | (1.00) | (0.45) | (1.86) |
| SREP cycle dummy variables | Yes | Yes | Yes |
| Observations | 474 | 162 | 312 |
| Banks | 99 | 63 | 98 |
| Adj. R-squared | 0.21 | 0.28 | 0.24 |

The table presents selected results from the estimation based on equation (2), in which the dependent variable is the yearly change in the supervisory judgement in the OSS, which is measurable as the difference between OSS set by supervisors and the automatic OSS, calculated as simple average of the four SREP viability scores. Scores are employed in their numerical form, with 1 representing the full difference between two main score categories (e.g., '2' and '3'). The estimated coefficients for the four SREP viability scores (BMA, GOV, CAP and LIQ) should be interpreted in conjunction with their value of 0.25 assumed a priori in the automatic score. Robust t statistics, calculated based on standard errors clustered by banks, in parentheses. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Table A.4 P2R. First-differencing (FD) estimation

| Dependent variable: | Sample | | |
|---|---|---|---|
| P2R | 2016-2024 | 2016-2020 | 2021-2024 |
| Overall Score | 0.2713*** | 0.0202 | 0.3097*** |
| | (3.49) | (0.10) | (3.50) |
| BMA Score | 0.0870** | 0.2696 | 0.0556* |
| | (2.19) | (1.49) | (1.95) |
| GOV Score | 0.0276 | -0.0106 | 0.0473** |
| | (0.68) | (-0.07) | (2.04) |
| CAP Score | 0.0347 | 0.0419 | 0.0517 |
| | (1.26) | (0.50) | (1.64) |
| LIQ Score | 0.1003 | 0.1646* | 0.0273 |
| | (1.65) | (1.85) | (0.44) |
| CRED Score | 0.0140 | -0.0038 | 0.0198 |
| | (0.61) | (-0.08) | (0.79) |
| MK Score | 0.0156 | 0.1015 | -0.0153 |
| | (0.80) | (1.63) | (-0.79) |
| OP Score | 0.0321 | -0.0361 | 0.0627* |
| | (0.76) | (-0.43) | (1.89) |
| IRRBB Score | -0.0130 | 0.0425 | -0.0153 |
| | (-0.79) | (0.72) | (-0.79) |
| ICAAP Score | 0.0701*** | 0.0773 | 0.0834*** |
| | (3.45) | (1.53) | (2.78) |
| GDP growth | -0.0003 | -0.0315* | 0.0018 |
| | (-0.12) | (-1.71) | (0.80) |
| Inflation | 0.0004 | -0.0291 | 0.0003 |
| | (0.13) | (-0.88) | (0.07) |
| Ln (stock index) | -0.0260 | 0.3039 | -0.0465 |
| | (-0.29) | (0.88) | (-0.51) |
| VSTOXX | 0.0037 | -0.0195* | 0.0028 |
| | (1.15) | (-1.79) | (0.89) |
| Ln (Total assets) | 0.0466 | 0.3177 | -0.0460 |
| | (0.62) | (1.61) | (-0.73) |
| Loan book | 0.0013 | 0.0014 | -0.0017 |
| | (0.61) | (0.54) | (-0.78) |
| Government loans ratio | -0.0076 | -0.0150 | -0.0063 |
| | (-0.98) | (-1.56) | (-0.71) |
| NPL ratio | 0.0097* | 0.0249** | 0.0050 |
| | (1.91) | (2.57) | (0.59) |
| Coverage ratio | -0.0010* | 0.0005 | -0.0006 |
| | (-1.91) | (0.24) | (-0.95) |

Table A.4 (cont'd) P2R. First-differencing (FD) estimation

| Dependent variable: P2R | Sample 2016-2024 | 2016-2020 | 2021-2024 |
|---|---|---|---|
| RWA density | -0.0042 | -0.0155* | -0.0017 |
| | (-1.62) | (-1.95) | (-0.51) |
| Credit risk share of RWA | -0.0021 | 0.0009 | -0.0026 |
| | (-1.39) | (0.13) | (-1.30) |
| Market risk share of RWA | -0.0119** | -0.0204 | -0.0084* |
| | (-1.99) | (-1.10) | (-1.78) |
| Operational risk share of RWA | -0.0040 | 0.0092 | -0.0057 |
| | (-0.77) | (0.60) | (-1.29) |
| NSFR | 0.0005 | -0.0007 | 0.0004 |
| | (1.41) | (-0.62) | (0.91) |
| LCR | 0.0001 | -0.0001 | -0.0000 |
| | (0.67) | (-0.44) | (-0.39) |
| Deposits ratio | 0.0004 | -0.0020 | 0.0012 |
| | (0.21) | (-0.62) | (0.50) |
| L2D | 0.0003* | 0.0002 | 0.0004* |
| | (1.95) | (1.20) | (1.87) |
| Cost to Income | -0.0001 | -0.0017 | 0.0004 |
| | (-0.26) | (-1.17) | (0.44) |
| RoE | -0.0015 | -0.0014 | 0.0043 |
| | (-0.81) | (-0.58) | (1.26) |
| RoA | 0.0184 | 0.0371 | -0.0524 |
| | (0.60) | (0.92) | (-1.17) |
| NII in Total assets | -0.0023 | 0.0050 | -0.0128 |
| | (-0.29) | (0.25) | (-1.15) |
| Change in EVE 200 bps | 0.0001 | -0.0005 | 0.0004 |
| | (0.52) | (-0.90) | (1.32) |
| SREP Cycle dummy variables | Yes | Yes | Yes |
| Observations | 465 | 159 | 306 |
| Banks | 99 | 62 | 98 |
| Adj. R-squared | 0.30 | 0.31 | 0.39 |

The table presents selected results of the estimation based on equation (3), in which the dependent variable is the yearly change in the P2R expressed in percentage. Scores are employed in their numerical form, with 1 representing the full difference between two main score categories (e.g., '2' and '3'). The difference accounting for score categories with qualifiers amounts to roughly one third. Robust t statistics, calculated based on standard errors clustered by banks, in parentheses. * p<0.10, ** p<0.05, *** p<0.01.

Table A.5 OSS Supervisory judgement. Alternative methods of estimation

| Dependent variable: OSS supervisory judgement Sample: 2016-2024 | Estimation method | | | | | |
|---|---|---|---|---|---|---|
| | Pooled OLS (POLS) | Random effects (RE) | Fixed effects (FE) | Correlated random effects (CRE) | FE assuming AR(1) serial correlation | First-differencing (FD) |
| BMA Score | -0.0543* | -0.0908*** | -0.1261*** | -0.1242*** | -0.1391*** | -0.1509*** |
| | (-1.89) | (-3.15) | (-3.86) | (-3.71) | (-4.23) | (-3.93) |
| GOV Score | -0.0764*** | -0.0738*** | -0.0741*** | -0.0773*** | -0.0509 | -0.0836** |
| | (-3.07) | (-2.99) | (-2.75) | (-2.59) | (-1.65) | (-2.04) |
| CAP Score | -0.0741** | -0.0532 | -0.0571 | -0.0488 | -0.0612* | -0.0906** |
| | (-2.05) | (-1.57) | (-1.55) | (-1.30) | (-1.90) | (-2.05) |
| LIQ Score | -0.1400*** | -0.1351*** | -0.1344*** | -0.1546*** | -0.1921*** | -0.1765*** |
| | (-3.19) | (-3.04) | (-2.77) | (-3.16) | (-4.33) | (-2.69) |
| CRED Score | 0.1537*** | 0.1339*** | 0.1241*** | 0.1283*** | 0.1021*** | 0.1127*** |
| | (5.37) | (4.99) | (4.39) | (4.39) | (3.64) | (3.26) |
| MK Score | 0.0125 | 0.0083 | 0.0138 | 0.0132 | -0.0169 | 0.0240 |
| | (0.56) | (0.37) | (0.51) | (0.46) | (-0.62) | (0.77) |
| OP Score | 0.0449 | 0.0467 | 0.0530 | 0.0433 | 0.0485 | 0.0678* |
| | (1.23) | (1.37) | (1.49) | (1.12) | (1.33) | (1.86) |
| IRRBB Score | 0.0503** | 0.0547*** | 0.0459** | 0.0576*** | 0.0181 | 0.0252 |
| | (1.99) | (2.61) | (2.20) | (2.63) | (0.70) | (0.94) |
| GDP growth | -0.0017 | -0.0015 | -0.0007 | -0.0022 | 0.0012 | -0.0014 |
| | (-0.37) | (-0.32) | (-0.14) | (-0.46) | (0.32) | (-0.30) |
| Inflation | -0.0114** | -0.0115** | -0.0121** | -0.0117** | -0.0073 | -0.0048 |
| | (-2.09) | (-2.24) | (-2.44) | (-2.22) | (-1.40) | (-1.06) |
| Lnstockidx | 0.0772 | 0.0625 | 0.0483 | 0.0470 | 0.1027 | 0.1191 |
| | (0.64) | (0.58) | (0.46) | (0.42) | (1.02) | (0.78) |
| VSTOXX | -0.0013 | 0.0128* | 0.0119* | 0.0127* | -0.0145 | 0.0024 |
| | (-0.26) | (1.89) | (1.81) | (1.86) | (-0.15) | (0.55) |
| Ln (Total assets) | -0.0964*** | -0.0744** | -0.0104 | -0.0217 | 0.1137 | 0.0576 |
| | (-3.25) | (-2.24) | (-0.22) | (-0.43) | (1.41) | (0.86) |
| Loan book | 0.0002 | 0.0013 | 0.0030 | 0.0029 | 0.0035 | 0.0047** |
| | (0.11) | (0.95) | (1.58) | (1.44) | (1.41) | (2.17) |
| Government loans ratio | 0.0021* | 0.0013 | -0.0010 | -0.0036 | -0.0022 | -0.0094 |
| | (1.76) | (1.07) | (-0.21) | (-0.70) | (-0.30) | (-1.59) |
| NPL ratio | 0.0120** | 0.0119** | 0.0112* | 0.0125** | 0.0324*** | 0.0075 |
| | (2.22) | (2.16) | (1.95) | (2.19) | (4.06) | (1.19) |
| Coverage ratio | -0.0001 | 0.0004 | 0.0011 | 0.0011 | 0.0010 | 0.0016 |
| | (-0.15) | (0.45) | (1.16) | (1.08) | (0.95) | (1.56) |
| RWA density | -0.0039** | -0.0043** | -0.0047* | -0.0061** | -0.0069* | -0.0065 |
| | (-2.51) | (-2.45) | (-1.70) | (-2.06) | (-1.89) | (-1.50) |

Table A.5 (cont'd) OSS Supervisory judgement. Alternative methods of estimation

| Dependent variable: OSS supervisory judgement Sample: 2016-2024 | Estimation method | | | | | |
|---|---|---|---|---|---|---|
| | Pooled OLS (POLS) | Random effects (RE) | Fixed effects (FE) | Correlated random effects (CRE) | FE assuming AR(1) serial correlation | First-differencing (FD) |
| Credit risk share of RWA | 0.0020 | 0.0002 | -0.0011 | -0.0010 | 0.0000 | 0.0013 |
| | (0.61) | (0.08) | (-0.39) | (-0.35) | (0.01) | (0.48) |
| Market risk share of RWA | 0.0038 | 0.0019 | -0.0018 | -0.0014 | -0.0021 | -0.0043 |
| | (0.63) | (0.34) | (-0.33) | (-0.23) | (-0.28) | (-0.64) |
| Operational risk share of RWA | 0.0057 | 0.0019 | -0.0030 | -0.0015 | -0.0137* | 0.0057 |
| | (1.00) | (0.41) | (-0.59) | (-0.30) | (-1.75) | (0.80) |
| NSFR | 0.0004 | -0.0002 | -0.0006 | -0.0005 | 0.0001 | 0.0004 |
| | (0.70) | (-0.37) | (-1.18) | (-0.94) | (0.18) | (0.62) |
| LCR | -0.0000 | -0.0000 | -0.0000 | -0.0001 | -0.0002 | -0.0003** |
| | (-0.18) | (-0.23) | (-0.22) | (-0.53) | (-1.40) | (-2.44) |
| Deposits ratio | -0.0002 | 0.0005 | 0.0019 | 0.0020 | 0.0006 | -0.0011 |
| | (-0.17) | (0.33) | (0.82) | (0.85) | (0.21) | (-0.34) |
| L2D | -0.0002* | -0.0003** | -0.0003** | -0.0003** | -0.0004** | -0.0004*** |
| | (-1.82) | (-2.46) | (-2.09) | (-2.16) | (-2.07) | (-2.57) |
| Cost to Income | -0.0004 | -0.0006 | -0.0008 | -0.0006 | -0.0017 | -0.0020** |
| | (-0.35) | (-0.58) | (-0.82) | (-0.52) | (-1.59) | (-2.09) |
| RoE | -0.0032 | 0.0001 | 0.0014 | 0.0014 | -0.0061 | 0.0012 |
| | (-1.01) | (0.03) | (0.43) | (0.45) | (-1.59) | (0.27) |
| RoA | 0.0070 | -0.0457 | -0.0669 | -0.0623 | 0.0146 | -0.0835 |
| | (0.18) | (-1.09) | (-1.45) | (-1.33) | (0.27) | (-1.36) |
| NII in Total assets | -0.0116 | -0.0084 | -0.0060 | -0.0060 | -0.0164 | -0.0113 |
| | (-0.81) | (-0.65) | (-0.45) | (-0.44) | (-1.07) | (-0.77) |
| Change in EVE 200bps | 0.0006 | 0.0006* | 0.0006* | 0.0006* | 0.0006* | 0.0004 |
| | (1.71) | (1.80) | (1.93) | (1.85) | (1.90) | (1.00) |

Table A.5 (cont'd) OSS Supervisory judgement. Alternative methods of estimation

| Dependent variable: | Estimation method | | | | | |
|---|---|---|---|---|---|---|
| OSS supervisory judgement<br>Sample: 2016-2024 | Pooled OLS (POLS) | Random effects (RE) | Fixed effects (FE) | Correlated random effects (CRE) | FE assuming AR(1) serial correlation | First-differencing (FD) |
| SREP cycle dummy variables | Yes | Yes | Yes | Yes | Yes | Yes |
| Country dummy variables | Yes | Yes | | Yes | | |
| Cluster dummy variables | Yes | Yes | | Yes | | |
| Peer group dummy variables | Yes | Yes | | Yes | | |
| Time average variables | | | | Yes | | |
| Constant | -0.138<br>(-0.11) | 0.6631<br>(0.77) | 0.0732<br>(0.07) | 16.3809**<br>(1.99) | -0.2599*<br>(-1.84) | |
| Observations | 584 | 584 | 584 | 584 | 488 | 474 |
| Banks | 96 | 96 | 96 | 96 | | |
| R-squared | 0.49 | | 0.31 | | | 0.21 |

This table presents selected results from the alternative methods of estimating equation (1), in which the dependent variable is supervisory judgement in OSS, measured as difference vs the automatic OSS calculated as simple average of the four SREP viability scores, estimated with various alternative techniques. Scores are employed in their numerical form, with 1 representing the full difference between two main score categories (e.g., '2' and '3'). The estimated coefficients for the four SREP viability scores (BMA, GOV, CAP and LIQ) should be interpreted in conjunction with their value of 0.25 assumed a priori in the automatic score. Other explanatory variables, except for those in logs, are expressed in percentages. In parentheses we report t and robust t statistics, based on the standard errors clustered by banks. * p<0.10, ** p<0.05, *** p<0.01.

Table A.6 P2R. Alternative methods of estimation

| Dependent variable: P2R<br>Sample: 2016-2024 | Estimation method: Pooled OLS (POLS) | Random effects (RE) | Fixed effects (FE) | Correlated random effects (CRE) | FE assuming AR(1) serial correlation | First-differencing (FD) |
|---|---|---|---|---|---|---|
| Overall Score | 0.6646*** | 0.5467*** | 0.4760*** | 0.5203*** | 0.2848*** | 0.2713*** |
| | (8.87) | (7.91) | (6.16) | (6.66) | (5.54) | (3.49) |
| BMA Score | 0.1105*** | 0.0980*** | 0.0786* | 0.0555 | 0.1044*** | 0.0870** |
| | (3.07) | (2.62) | (1.96) | (1.28) | (3.57) | (2.19) |
| GOV Score | 0.0741* | 0.0892** | 0.0937** | 0.0951** | 0.0178 | 0.0276 |
| | (1.67) | (2.25) | (2.45) | (2.36) | (0.67) | (0.68) |
| CAP Score | 0.0716** | 0.0728** | 0.0702* | 0.0728* | 0.0423 | 0.0347 |
| | (2.02) | (2.03) | (1.85) | (1.84) | (1.53) | (1.26) |
| LIQ Score | -0.0282 | 0.0110 | 0.0296 | 0.0323 | 0.0754** | 0.1003 |
| | (-0.55) | (0.23) | (0.59) | (0.61) | (2.09) | (1.65) |
| CRED Score | 0.0180 | 0.0049 | 0.0019 | -0.0067 | 0.0135 | 0.0140 |
| | (0.54) | (0.17) | (0.07) | (-0.23) | (0.57) | (0.61) |
| MK Score | -0.0304 | -0.0229 | -0.0213 | -0.0220 | -0.0106 | 0.0156 |
| | (-0.90) | (-0.88) | (-0.77) | (-0.74) | (-0.46) | (0.80) |
| OP Score | -0.0017 | 0.0093 | 0.0111 | 0.0095 | -0.0044 | 0.0321 |
| | (-0.04) | (0.20) | (0.24) | (0.20) | (-0.14) | (0.76) |
| IRRBB Score | -0.0144 | -0.0349 | -0.0429 | -0.0352 | -0.0392* | -0.0130 |
| | (-0.40) | (-1.15) | (-1.55) | (-1.12) | (-1.80) | (-0.79) |
| ICAAP Score | 0.0974*** | 0.0680*** | 0.0720*** | 0.0562** | 0.0549** | 0.0701*** |
| | (3.14) | (2.74) | (2.83) | (2.05) | (2.53) | (3.45) |
| GDP growth | -0.0010 | -0.0034 | -0.0037 | -0.0024 | 0.0016 | -0.0003 |
| | (-0.21) | (-0.80) | (-0.90) | (-0.58) | (0.52) | (-0.12) |
| Inflation | 0.0090 | 0.0051 | 0.0035 | 0.0007 | -0.0009 | 0.0004 |
| | (1.63) | (1.22) | (0.93) | (0.18) | (-0.22) | (0.13) |
| Ln (stock index) | 0.2284 | 0.1817 | 0.1785 | 0.1781 | -0.0602 | -0.0260 |
| | (1.31) | (1.14) | (1.18) | (1.11) | (-0.68) | (-0.29) |
| VSTOXX | -0.0336 | -0.0009 | -0.0004 | 0.0047 | 0.1368 | 0.0037 |
| | (-0.52) | (-0.15) | (-0.06) | (0.73) | (0.93) | (1.15) |
| Ln (Total assets) | -0.0758 | 0.0012 | 0.0857 | 0.1094 | 0.1875** | 0.0466 |
| | (-1.52) | (0.02) | (1.32) | (1.63) | (2.55) | (0.62) |
| Loan book | 0.0005 | 0.0004 | 0.0020 | 0.0023 | 0.0011 | 0.0013 |
| | (0.26) | (0.27) | (0.92) | (1.00) | (0.54) | (0.61) |
| Government loans ratio | -0.0033 | -0.0060* | 0.0020 | -0.0018 | -0.0153** | -0.0076 |
| | (-1.10) | (-1.75) | (0.29) | (-0.25) | (-2.39) | (-0.98) |
| NPL ratio | 0.0048 | 0.0023 | 0.0019 | 0.0003 | -0.0033 | 0.0097* |
| | (1.09) | (0.66) | (0.49) | (0.09) | (-0.40) | (1.91) |
| Coverage ratio | 0.0006 | -0.0009 | -0.0009 | -0.0010 | -0.0009 | -0.0010* |
| | (0.48) | (-0.99) | (-1.02) | (-1.07) | (-1.03) | (-1.91) |
| RWA density | -0.0056** | -0.0056** | -0.0056 | -0.0043 | -0.0028 | -0.0042 |
| | (-2.52) | (-2.37) | (-1.63) | (-1.25) | (-0.83) | (-1.62) |

Table A.6 (cont'd) P2R. Alternative methods of estimation

| Dependent variable: P2R<br>Sample: 2016-2024 | Estimation method | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
| | Pooled OLS (POLS) | Random effects (RE) | Fixed effects (FE) | Correlated random effects (CRE) | FE assuming AR(1) serial correlation | First-differencing (FD) |
| Credit risk share of RWA | -0.0045 | -0.0013 | 0.0014 | 0.0054 | -0.0015 | -0.0021 |
| | (-1.04) | (-0.48) | (0.45) | (1.41) | (-0.46) | (-1.39) |
| Market risk share of RWA | -0.0048 | -0.0124* | -0.0180** | -0.0166* | -0.0065 | -0.0119** |
| | (-0.64) | (-1.79) | (-2.30) | (-1.90) | (-1.05) | (-1.99) |
| Operational risk share of RWA | 0.0066 | 0.0074 | 0.0043 | 0.0116 | -0.0065 | -0.0040 |
| | (1.02) | (1.10) | (0.51) | (1.26) | (-1.00) | (-0.77) |
| NSFR | 0.0014 | -0.0006 | -0.0011 | -0.0013 | 0.0003 | 0.0005 |
| | (1.15) | (-0.55) | (-1.02) | (-1.23) | (0.55) | (1.41) |
| LCR | -0.0003** | -0.0000 | 0.0002 | 0.0002 | 0.0000 | 0.0001 |
| | (-2.05) | (-0.03) | (1.25) | (1.26) | (0.28) | (0.67) |
| Deposits ratio | -0.0012 | -0.0007 | -0.0006 | -0.0025 | -0.0006 | 0.0004 |
| | (-0.46) | (-0.27) | (-0.22) | (-0.83) | (-0.23) | (0.21) |
| L2D | -0.0002 | -0.0000 | 0.0002 | 0.0001 | 0.0001 | 0.0003* |
| | (-0.62) | (-0.22) | (1.03) | (0.33) | (0.69) | (1.95) |
| Cost to Income | -0.0024* | -0.0008 | -0.0004 | 0.0002 | -0.0009 | -0.0001 |
| | (-1.84) | (-0.86) | (-0.42) | (0.21) | (-1.11) | (-0.26) |
| RoE | -0.0034 | 0.0033 | 0.0033 | 0.0030 | -0.0028 | -0.0015 |
| | (-0.75) | (1.06) | (1.21) | (1.07) | (-0.95) | (-0.81) |
| RoA | 0.0839 | -0.0197 | -0.0216 | -0.0107 | 0.0283 | 0.0184 |
| | (1.25) | (-0.39) | (-0.48) | (-0.23) | (0.71) | (0.60) |
| NII in Total assets | -0.0046 | -0.0038 | -0.0008 | 0.0024 | -0.0019 | -0.0023 |
| | (-0.27) | (-0.28) | (-0.06) | (0.17) | (-0.16) | (-0.29) |
| Change in EVE 200 bps | 0.0003 | -0.0000 | -0.0001 | 0.0001 | 0.0001 | 0.0001 |
| | (0.80) | (-0.13) | (-0.21) | (0.31) | (0.39) | (0.52) |

Table A.6 (cont'd) P2R. Alternative methods of estimation

| Dependent variable: P2R<br><br>Sample: 2016-2024 | Estimation method<br>Pooled OLS (POLS) | Random effects (RE) | Fixed effects (FE) | Correlated random effects (CRE) | FE assuming AR(1) serial correlation | First-differencing (FD) |
|---|---|---|---|---|---|---|
| SREP Cycle dummy variables | Yes | Yes | Yes | Yes | Yes | Yes |
| Country dummy variables | Yes | Yes | | Yes | | |
| Cluster dummy variables | Yes | Yes | | Yes | | |
| Peer group dummy variables | Yes | Yes | | Yes | | |
| Time average variables | | | | Yes | | |
| Constant | | -0.3255<br>(-0.26) | -1.7049<br>(-1.09) | 10.2328<br>(1.16) | -0.0122<br>(-0.15) | |
| Observations | 577 | 577 | 577 | 577 | 481 | 465 |
| R-squared | 0.99 | | 0.57 | | | 0.30 |

The table presents selected results of the estimation based on equation (1), in which the dependent variable is the P2R expressed in percentage. Scores are employed in their numerical form, with 1 representing the full difference between two main score categories (e.g., '2' and '3'). The difference accounting for score categories with qualifiers amounts to roughly one third. Other explanatory variables, except for those in logs, are expressed in percentages. In parentheses we report t and robust t statistics, based on standard errors clustered by banks. * p<0.10, ** p<0.05, *** p<0.01.

Table A.7 OSS Supervisory judgement. Panel ordered logit estimation

| Dependent variable Δ OSS Supervisory judgement | Sample 2016-2024 | 2016-2020 | 2021-2024 |
|---|---|---|---|
| Δ BMA Score | -1.4309*** | -2.2995*** | -0.9708** |
| | (-4.33) | (-3.04) | (-2.42) |
| Δ GOV Score | -0.8894** | -2.4602** | -0.6826* |
| | (-2.20) | (-2.43) | (-1.93) |
| Δ CAP Score | -1.0164*** | -2.6092*** | -0.7006** |
| | (-2.56) | (-2.59) | (-1.90) |
| Δ LIQ Score | -1.6996*** | -2.5101*** | -1.6588*** |
| | (-3.05) | (-2.64) | (-2.74) |
| Δ CRED Score | 1.0018*** | 1.7920** | 0.7005* |
| | (2.92) | (2.42) | (1.79) |
| Δ MK Score | 0.2502 | 1.8718 | -0.0034 |
| | (0.73) | (1.57) | (-0.01) |
| Δ OP Score | 0.5161 | 1.1609 | 0.0772 |
| | (1.49) | (1.51) | (0.18) |
| Δ IRRBB Score | 0.2034 | 0.5943 | 0.1717 |
| | (0.74) | (0.69) | (0.57) |
| Δ GDP growth | 0.0173 | -0.0209 | 0.0149 |
| | (0.44) | (-0.13) | (0.36) |
| Δ Inflation | -0.0260 | 0.6719 | -0.0318 |
| | (1.2860) | (1.57) | (-1.01) |
| Δ Ln (stock index) | 1.2860 | 1.1307 | 1.7258 |
| | (0.98) | (0.29) | (1.14) |
| Δ VSTOXX | 0.0212 | 0.0260 | 0.0255 |
| | (0.66) | (0.24) | (0.63) |
| Δ Ln (TA) | 0.2761 | -0.8738 | 0.7512 |
| | (0.40) | (-0.49) | (0.65) |
| Δ Loan book | 0.0427* | 0.0046 | 0.0884** |
| | (1.91) | (0.06) | (2.67) |
| Δ Government loans ratio | -0.0573 | -0.0010 | -0.0149 |
| | (-0.96) | (-0.01) | (-0.16) |
| Δ NPL ratio | 0.0447 | 0.2163** | 0.0225 |
| | (0.76) | (2.33) | (0.28) |
| Δ Coverage ratio | 0.0205** | 0.0580*** | -0.0001 |
| | (2.20) | (3.91) | (-0.01) |
| Δ RWA density | -0.0664 | -0.0961 | -0.0625 |
| | (-1.51) | (-1.15) | (-1.19) |

Table A.7 (cont'd) OSS Supervisory judgement. Panel ordered logit estimation

| Dependent variable | Sample | | |
|---|---|---|---|
| Δ OSS Supervisory judgement | 2016-2024 | 2016-2020 | 2021-2024 |
| Δ Credit risk share of RWA | 0.0234 | -0.0780 | 0.0255 |
| | (0.76) | (-0.96) | (0.77) |
| Δ Market risk share of RWA | -0.0307 | -0.0315 | -0.0687 |
| | (-0.41) | (-0.16) | (-0.87) |
| Δ Operational risk share of RWA | 0.0155 | -0.1304 | -0.0253 |
| | (0.20) | (-0.71) | (-0.30) |
| Δ NSFR | 0.0081 | -0.0013 | 0.0117** |
| | (1.09) | (-0.14) | (2.13) |
| Δ LCR | -0.0031*** | 0.0017 | -0.0041*** |
| | (-2.63) | (0.46) | (-2.93) |
| Δ Deposits ratio | -0.0120 | -0.0487 | -0.0104 |
| | (-0.34) | (-0.88) | (-0.26) |
| Δ L2D | -0.0031** | -0.0042 | -0.0046** |
| | (-2.34) | (-1.36) | (-2.34) |
| Δ Cost to Income | -0.0158* | -0.0475*** | -0.0062 |
| | (-1.83) | (-4.67) | (-0.48) |
| Δ RoE | 0.0002 | 0.0630 | -0.0941 |
| | (0.01) | (1.31) | (-1.56) |
| Δ RoA | -0.6097 | -0.9322 | 0.2355 |
| | (-0.98) | (-1.05) | (0.28) |
| Δ NII in Total assets | -0.0870 | 0.0242 | -0.2237 |
| | (-0.58) | (0.10) | (-1.19) |
| Δ Change in EVE 200 bps | 0.0035 | 0.0055 | 0.0063 |
| | (1.02) | (0.97) | (1.61) |
| DummySREP2020 | -0.3063 | 0.4860 | - |
| | (-1.31) | (0.49) | |
| Observations | 474 | 162 | 312 |

The table presents the results of the estimation of a panel ordered logit model, in which dependent variable is the change in the OSS supervisory judgement treated as a categorical variable (e.g., in number of score notches). t statistics in parentheses. * p<0.10, ** p<0.05, *** p<0.01.

Table A.8 OSS Supervisory judgement. Panel ordered probit estimation

| Dependent variable Δ OSS Supervisory judgement | Sample 2016-2024 | 2016-2020 | 2021-2024 |
|---|---|---|---|
| Δ BMA Score | -0.7297*** | -1.2130*** | -0.5388*** |
| | (-4.43) | (-3.43) | (-2.81) |
| Δ GOV Score | -0.4514** | -1.1051** | -0.3771** |
| | (-2.26) | (-2.57) | (-2.06) |
| Δ CAP Score | -0.5096** | -1.1791** | -0.4192** |
| | (-2.49) | (-2.55) | (-2.21) |
| Δ LIQ Score | -0.8593*** | -1.0861** | -0.9237*** |
| | (-2.88) | (-2.49) | (-3.06) |
| Δ CRED Score | 0.5387*** | 0.9789*** | 0.4138** |
| | (3.32) | (2.94) | (2.25) |
| Δ MK Score | 0.1235 | 0.6632 | -0.0339 |
| | (0.77) | (1.27) | (-0.21) |
| Δ OP Score | 0.3056* | 0.7054* | 0.0497 |
| | (1.86) | (1.91) | (0.24) |
| Δ IRRBB Score | 0.1374 | 0.4483 | 0.1381 |
| | (0.94) | (1.13) | (0.95) |
| Δ GDP growth | 0.0133 | 0.0070 | 0.0145 |
| | (0.73) | (0.10) | (0.73) |
| Δ Inflation | -0.0131 | 0.3268* | -0.0178 |
| | (-0.91) | (1.71) | (-1.15) |
| Δ Ln (stock index) | 0.7335 | 0.8831 | 0.9260 |
| | (1.20) | (0.54) | (1.26) |
| Δ VSTOXX | 0.0141 | 0.0236 | 0.0189 |
| | (0.87) | (0.50) | (0.87) |
| Δ Ln (TA) | 0.2381 | -0.2748 | 0.5142 |
| | (0.68) | (-0.28) | (0.92) |
| Δ Loan book | 0.0194* | 0.0016 | 0.0424*** |
| | (1.77) | (0.07) | (2.82) |
| Δ Government loans ratio | -0.0351 | -0.0231 | -0.0108 |
| | (-1.20) | (-0.42) | (-0.23) |
| Δ NPL ratio | 0.0410 | 0.0991*** | 0.0209 |
| | (1.50) | (2.59) | (0.56) |
| Δ Coverage ratio | 0.0093* | 0.0287*** | -0.0001 |
| | (1.94) | (3.97) | (-0.01) |
| Δ RWA density | -0.0295 | -0.0340 | -0.0242 |
| | (-1.39) | (-1.15) | (-0.94) |

Table A.8 (cont'd) OSS Supervisory judgement. Panel ordered probit estimation

| Dependent variable<br>Δ OSS Supervisory judgement | Sample<br>2016-2024 | <br>2016-2020 | <br>2021-2024 |
|---|---|---|---|
| Δ Credit risk share of RWA | 0.0074 | -0.0447 | 0.0095 |
| | (0.46) | (-1.08) | (0.52) |
| Δ Market risk share of RWA | -0.0234 | 0.0319 | -0.0448 |
| | (-0.63) | (0.36) | (-1.06) |
| Δ Operational risk share of RWA | 0.0136 | -0.0581 | -0.0088 |
| | (0.35) | (-0.65) | (-0.20) |
| Δ NSFR | 0.0031 | -0.0031 | 0.0059** |
| | (1.01) | (-0.66) | (2.08) |
| Δ LCR | -0.0016** | 0.0017 | -0.0023*** |
| | (-2.45) | (0.87) | (-3.20) |
| Δ Deposits ratio | -0.0030 | -0.0276 | -0.0038 |
| | (-0.18) | (-1.00) | (-0.19) |
| Δ L2D | -0.0016** | -0.0025* | -0.0023** |
| | (-2.41) | (-1.73) | (-2.46) |
| Δ Cost to Income | -0.0092** | -0.0223*** | -0.0037 |
| | (-2.39) | (-4.69) | (-0.58) |
| Δ RoE | 0.0030 | 0.0334 | -0.0568* |
| | (0.14) | (1.35) | (-1.96) |
| Δ RoA | -0.3411 | -0.4372 | 0.2031 |
| | (-1.16) | (-0.96) | (0.52) |
| Δ NII in Total assets | -0.0475 | 0.0436 | -0.1373 |
| | (-0.61) | (0.35) | (-1.45) |
| Δ Change in EVE 200 bps | 0.0022 | 0.0025 | 0.0040* |
| | (1.18) | (0.85) | (1.89) |
| DummySREP2020 | -0.2051 | 0.2899 | |
| | (-1.41) | (0.65) | |
| Observations | 474 | 162 | 312 |

The table presents the results of the estimation of a panel ordered probit model, in which dependent variable is the change in the OSS supervisory judgement treated as a categorical variable (e.g., in number of score notches). t statistics in parentheses. * p<0.10, ** p<0.05, *** p<0.01.

Table A.9 OSS: dynamic panel estimation

| Dependent variable<br>Overall SREP Score | Sample<br>2016-2024 | <br>2016-2020 | <br>2021-2024 |
|---|---|---|---|
| Lag (Overall SREP Score) | 0.2862*** | 0.3304*** | 0.2802*** |
| | (6.53) | (4.49) | (4.71) |
| BMA Score | 0.1517*** | 0.1377*** | 0.1507*** |
| | (4.42) | (2.71) | (3.14) |
| GOV Score | 0.1317*** | 0.1271*** | 0.1050*** |
| | (4.62) | (3.12) | (3.06) |
| CAP Score | 0.1440*** | 0.1889*** | 0.1459*** |
| | (3.48) | (3.50) | (3.26) |
| LIQ Score | 0.1220*** | 0.1322*** | 0.1289* |
| | (2.62) | (2.59) | (1.82) |
| CRED Score | 0.1250*** | 0.0990** | 0.1104*** |
| | (3.68) | (1.99) | (2.59) |
| MK Score | -0.0449 | 0.0432 | -0.0372 |
| | (-1.29) | (0.92) | (-0.85) |
| OP Score | -0.0212 | 0.0206 | -0.0661 |
| | (-0.56) | (0.41) | (-1.52) |
| IRRBB Score | 0.0409 | -0.0415 | 0.0720** |
| | (1.59) | (-0.78) | (2.38) |
| Ln (TA) | -0.0581** | -0.0876*** | -0.0350 |
| | (-2.25) | (-3.36) | (-1.17) |
| Loan book | -0.0023 | -0.0010 | 0.0002 |
| | (-1.25) | (-0.55) | (0.08) |
| Government loans ratio | -0.0011 | -0.0012 | -0.0063 |
| | (-0.45) | (-0.50) | (-1.43) |
| NPL ratio | 0.0154*** | 0.0129** | 0.0094 |
| | (3.05) | (2.12) | (1.35) |
| Coverage ratio | -0.0004 | 0.0005 | -0.0012 |
| | (-0.35) | (0.39) | (-0.89) |
| RWA density | -0.0014 | -0.0069** | 0.0005 |
| | (-0.56) | (-2.35) | (0.21) |
| Credit risk share of RWA | 0.0011 | 0.0025 | -0.0005 |
| | (0.44) | (0.58) | (-0.14) |
| Market risk share of RWA | 0.0080 | 0.0074 | 0.0092 |
| | (1.11) | (1.12) | (1.02) |
| Operational risk share of RWA | -0.0066 | 0.0034 | -0.0027 |
| | (-1.19) | (0.44) | (-0.41) |
| NSFR | 0.0001 | -0.0010 | 0.0006 |
| | (0.29) | (-1.30) | (0.82) |
| LCR | 0.0002 | -0.0003 | 0.0001 |
| | (0.80) | (-1.33) | (0.61) |
| Deposits ratio | -0.0012 | -0.0005 | -0.0013 |
| | (-0.73) | (-0.27) | (-0.55) |

Table A.9 (cont'd) OSS: dynamic panel estimation

| Dependent variable<br>Overall SREP Score | 2016-2024 | Sample<br>2016-2020 | 2021-2024 |
|---|---|---|---|
| L2D | 0.0000 | 0.0001 | 0.0001 |
| | (0.42) | (0.42) | (0.56) |
| Cost to Income | -0.0008 | -0.0011 | 0.0004 |
| | (-0.80) | (-1.04) | (0.25) |
| RoE | 0.0046 | 0.0072*** | -0.0065 |
| | (1.30) | (3.10) | (-1.32) |
| RoA | -0.0681 | -0.0691* | 0.0118 |
| | (-1.59) | (-1.77) | (0.19) |
| NII in Total assets | -0.0064 | 0.0062 | -0.0120 |
| | (-0.47) | (0.37) | (-0.77) |
| Change in EVE 200bps | 0.0003 | -0.0001 | 0.0008 |
| | (1.01) | (-0.35) | (1.62) |
| GDP growth | -0.0010 | 0.0185* | -0.0013 |
| | (-0.25) | (1.94) | (-0.30) |
| Inflation | -0.0073 | -0.0047 | -0.0076 |
| | (-1.52) | (-0.23) | (-1.51) |
| Ln (stock index) | -0.0170 | -0.0302 | -0.0545* |
| | (-0.75) | (-1.46) | (-1.81) |
| Constant | 0.9467** | 0.8703* | 1.2050** |
| | (2.09) | (1.93) | (1.99) |
| Observations | 452 | 160 | 292 |

The table presents the results of a dynamic model for OSS, estimated using the Arellano-Bond technique. t statistics in parentheses. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Table A.10 P2R: dynamic panel estimation

| Dependent variable | Sample | | |
|---|---|---|---|
| P2R | 2016-2024 | 2016-2020 | 2021-2024 |
| Lag (P2R) | 0.4039*** | 0.2883** | 0.5736*** |
| | (4.91) | (2.46) | (10.76) |
| Overall Score | 0.4296*** | 0.4045*** | 0.3979*** |
| | (5.25) | (2.80) | (4.61) |
| BMA Score | 0.0343 | 0.1337 | 0.0112 |
| | (0.76) | (0.99) | (0.34) |
| GOV Score | 0.0364 | 0.0108 | 0.0247 |
| | (0.84) | (0.12) | (0.73) |
| CAP Score | -0.0334 | -0.0529 | 0.0242 |
| | (-0.80) | (-0.68) | (0.68) |
| LIQS core | 0.0617 | 0.0781 | 0.0201 |
| | (1.02) | (1.06) | (0.30) |
| CRED Score | 0.0109 | -0.0553 | -0.0185 |
| | (0.29) | (-0.98) | (-0.64) |
| MKS core | 0.0128 | 0.1169* | -0.0359 |
| | (0.40) | (1.91) | (-1.39) |
| OP Score | -0.0244 | 0.0465 | 0.0354 |
| | (-0.44) | (0.63) | (0.84) |
| IRRBB Score | -0.0257 | -0.1651** | -0.0006 |
| | (-1.01) | (-2.23) | (-0.02) |
| ICAAP Score | 0.0707*** | 0.0689 | 0.0939*** |
| | (2.63) | (1.23) | (2.78) |
| Ln (TA) | -0.0664*** | -0.0508 | -0.0225 |
| | (-2.69) | (-1.24) | (-1.28) |
| Loan book | -0.0006 | -0.0022 | -0.0021 |
| | (-0.30) | (-0.60) | (-1.19) |
| Government loans ratio | -0.0051 | -0.0057 | 0.0012 |
| | (-1.48) | (-1.23) | (0.65) |
| NPL ratio | -0.0011 | 0.0058 | -0.0098 |
| | (-0.19) | (0.59) | (-1.17) |
| Coverage ratio | -0.0011 | -0.0031* | -0.0005 |
| | (-1.35) | (-1.65) | (-0.79) |
| RWA density | -0.0039* | 0.0006 | 0.0010 |
| | (-1.66) | (0.17) | (0.64) |
| Credit risk share of RWA | 0.0011 | -0.0097 | -0.0007 |
| | (0.32) | (-1.03) | (-0.25) |
| Market risk share of RWA | 0.0004 | -0.0275** | -0.0008 |
| | (0.07) | (-2.29) | (-0.14) |
| Operational risk share of RWA | -0.0016 | -0.0023 | -0.0032 |
| | (-0.27) | (-0.21) | (-0.69) |
| NSFR | 0.0003 | 0.0024 | 0.0008* |
| | (0.50) | (1.09) | (1.67) |
| LCR | -0.0001 | -0.0009** | -0.0002 |
| | (-0.75) | (-2.28) | (-1.11) |

Table A.10 (cont'd) P2R: dynamic panel estimation

| Dependent variable | Sample | | |
|---|---|---|---|
| P2R | 2016-2024 | 2016-2020 | 2021-2024 |
| Deposits ratio | -0.0013 | 0.0009 | -0.0022 |
| | (-0.58) | (0.26) | (-1.27) |
| L2D | 0.0001 | 0.0004 | 0.0002 |
| | (0.48) | (1.37) | (1.13) |
| Cost to Income | 0.0006 | -0.0014 | 0.0009 |
| | (0.69) | (-0.80) | (1.20) |
| RoE | -0.0040 | -0.0046 | -0.0001 |
| | (-1.50) | (-0.78) | (-0.03) |
| RoA | 0.0712 | 0.1189 | 0.0289 |
| | (1.53) | (0.94) | (0.55) |
| NII in Total assets | -0.0005 | 0.0392* | -0.0097 |
| | (-0.04) | (1.88) | (-0.78) |
| Change in EVE 200 bps | 0.0001 | -0.0005 | 0.0002 |
| | (0.24) | (-0.74) | (0.59) |
| GDP growth | 0.0010 | -0.0063 | 0.0025 |
| | (0.34) | (-0.29) | (1.08) |
| Inflation | 0.0019 | 0.0381 | 0.0030 |
| | (0.43) | (1.27) | (0.71) |
| Ln (stock index) | -0.0007 | -0.0465 | -0.0335 |
| | (-0.03) | (-1.34) | (-1.61) |
| Constant | 0.3413 | 1.5101 | 0.0560 |
| | (0.59) | (1.39) | (0.12) |
| Observations | 436 | 154 | 282 |

The table presents the results of a dynamic model for P2R, estimated using the Arellano-Bond technique. t statistics in parentheses. * p<0.10, ** p<0.05, *** p<0.01.

Table A.11 OSS Supervisory judgement. First-differencing (FD) estimation, using different sets of explanatory variables

| Dependent variable: OSS Supervisory judgement | Explanatory variables | | |
|---|---|---|---|
| | Supervisory scores | Supervisory scores and macro variables | Supervisory scores and bank indicators |
| BMA Score | -0.1152*** | -0.1108*** | -0.1540*** |
| | (-4.10) | (-3.81) | (-4.12) |
| GOV Score | -0.1075*** | -0.1061*** | -0.0860** |
| | (-3.71) | (-3.68) | (-2.10) |
| CAP Score | -0.1043*** | -0.1057*** | -0.0894** |
| | (-3.21) | (-3.26) | (-2.02) |
| LIQ Score | -0.2002*** | -0.1992*** | -0.1768*** |
| | (-5.56) | (-5.51) | (-2.69) |
| CRED Score | 0.1110*** | 0.1139*** | 0.1079*** |
| | (4.78) | (5.00) | (3.06) |
| MK Score | 0.0066 | 0.0071 | 0.0254 |
| | (0.31) | (0.34) | (0.79) |
| OP Score | 0.0730*** | 0.0754*** | 0.0643* |
| | (3.08) | (3.18) | (1.73) |
| IRRBB Score | 0.0290 | 0.0271 | 0.0280 |
| | (1.19) | (1.11) | (1.04) |
| GDP growth | | -0.0036 | |
| | | (-0.96) | |
| Inflation | | -0.0048 | |
| | | (-1.30) | |
| Ln (stock index) | | 0.0543 | |
| | | (0.62) | |
| VSTOXX | | -0.0021 | |
| | | (-0.80) | |
| Ln (Total assets) | | | 0.0611 |
| | | | (0.90) |
| Loan book | | | 0.0047** |
| | | | (2.16) |
| Government loans ratio | | | -0.0092 |
| | | | (-1.53) |
| NPL ratio | | | 0.0076 |
| | | | (1.18) |
| Coverage ratio | | | 0.0016 |
| | | | (1.63) |
| RWA density | | | -0.0066 |
| | | | (-1.51) |
| Credit risk share of RWA | | | 0.0010 |
| | | | (0.34) |
| Market risk share of RWA | | | -0.0049 |
| | | | (-0.73) |
| Operational risk share of RWA | | | 0.0055 |
| | | | (0.76) |

Table A.11 (cont'd) OSS Supervisory judgement. First-differencing (FD) estimation, using different sets of explanatory variables

| Dependent variable: OSS Supervisory judgement | Explanatory variables: Supervisory scores | Supervisory scores and macro variables | Supervisory scores and bank indicators |
|---|---|---|---|
| NSFR | | | 0.0004 |
| | | | (0.66) |
| LCR | | | -0.0003** |
| | | | (-2.39) |
| Deposits ratio | | | -0.0012 |
| | | | (-0.36) |
| L2D | | | -0.0004** |
| | | | (-2.55) |
| Cost to Income | | | -0.0021** |
| | | | (-2.30) |
| RoE | | | 0.0010 |
| | | | (0.23) |
| RoA | | | -0.0786 |
| | | | (-1.32) |
| NII in Total assets | | | -0.0107 |
| | | | (-0.77) |
| Change in EVE 200 bps | | | 0.0004 |
| | | | (1.09) |
| SREP cycle dummy variables | Yes | Yes | Yes |
| Observations | 821 | 818 | 474 |
| Banks | 139 | 138 | 99 |
| Adj. R-squared | 0.22 | 0.22 | 0.21 |

The table presents selected results of the first differencing estimation for supervisory judgement in OSS using various sets of explanatory variables. t statistics, calculated based on standard errors clustered by bank, in parentheses. * p<0.10, ** p<0.05, *** p<0.01.

Table A.12 P2R. First-differencing (FD) estimation, using different sets of explanatory variables

| Dependent variable: P2R | Explanatory variables | | |
|---|---|---|---|
| | Supervisory scores | Supervisory scores and macro variables | Supervisory scores and bank indicators |
| Overall Score | 0.3663*** | 0.3664*** | 0.2693*** |
| | (4.03) | (3.95) | (3.50) |
| BMA Score | 0.0691** | 0.0705* | 0.0876** |
| | (2.00) | (1.97) | (2.25) |
| GOV Score | 0.0416 | 0.0418 | 0.0280 |
| | (1.42) | (1.42) | (0.69) |
| CAP Score | 0.0645** | 0.0638** | 0.0350 |
| | (2.40) | (2.37) | (1.26) |
| LIQS Score | 0.0500* | 0.0503* | 0.1006 |
| | (1.66) | (1.66) | (1.66) |
| CRED Score | 0.0204 | 0.0215 | 0.0151 |
| | (0.92) | (0.95) | (0.68) |
| MK Score | 0.0215 | 0.0211 | 0.0153 |
| | (1.43) | (1.40) | (0.79) |
| OP Score | 0.0422 | 0.0433 | 0.0330 |
| | (1.24) | (1.27) | (0.78) |
| IRRBB Score | 0.0305 | 0.0307 | -0.0134 |
| | (0.96) | (0.95) | (-0.83) |
| ICAAP Score | 0.0605*** | 0.0601*** | 0.0699*** |
| | (3.22) | (3.14) | (3.51) |
| GDP growth | | -0.0012 | |
| | | (-0.68) | |
| Inflation | | -0.0002 | |
| | | (-0.09) | |
| Ln (stock index) | | 0.0353 | |
| | | (0.43) | |
| VSTOXX | | -0.0131*** | |
| | | (-3.04) | |
| Ln (Total assets) | | | 0.0471 |
| | | | (0.62) |
| Loan book | | | 0.0013 |
| | | | (0.62) |
| Government loans ratio | | | -0.0077 |
| | | | (-0.99) |
| NPL ratio | | | 0.0096* |
| | | | (1.89) |
| Coverage ratio | | | -0.0010* |
| | | | (-1.90) |

Table A.12 (cont'd) P2R. First-differencing (FD) estimation, using different sets of explanatory variables

| Dependent variable: P2R | Explanatory variables: Supervisory scores | Supervisory scores and macro variables | Supervisory scores and bank indicators |
|---|---|---|---|
| RWA density | | | -0.0042 |
| | | | (-1.63) |
| Credit risk share of RWA | | | -0.0021 |
| | | | (-1.47) |
| Market risk share of RWA | | | -0.0119* |
| | | | (-1.98) |
| Operational risk share of RWA | | | -0.0040 |
| | | | (-0.77) |
| NSFR | | | 0.0005 |
| | | | (1.43) |
| LCR | | | 0.0001 |
| | | | (0.67) |
| Deposits ratio | | | 0.0005 |
| | | | (0.22) |
| L2D | | | 0.0003* |
| | | | (1.97) |
| Cost to Income | | | -0.0001 |
| | | | (-0.25) |
| RoE | | | -0.0015 |
| | | | (-0.81) |
| RoA | | | 0.0184 |
| | | | (0.59) |
| NII in Total assets | | | -0.0024 |
| | | | (-0.31) |
| Change in EVE 200 bps | | | 0.0001 |
| | | | (0.52) |
| SREP Cycle dummy variables | Yes | Yes | Yes |
| Observations | 797 | 794 | 465 |
| Banks | 137 | 136 | 99 |
| Adj. R-squared | 0.27 | 0.27 | 0.30 |

The table presents selected results of the first differencing estimation for P2R using various sets of explanatory variables. t statistics, calculated based on standard errors clustered by bank, in parentheses. * p<0.10, ** p<0.05, *** p<0.01.

**Acknowledgements**

This paper has benefited from feedback and comments from Olga Arratibel, Thomas Prehofer and Mario Quagliariello, as well as from the participants in the Banking Supervision Internal Research Seminar, SSR and SMD network members, and from bilateral exchanges with colleagues from ACPR, Deutsche Bundesbank, Banco de España, and Oesterreichische Nationalbank.
The views expressed are those of the authors and do not necessarily reflect those of the ECB.

**Gabriel Bobeică (corresponding author)**
European Central Bank, Frankfurt am Main, Germany; email: gabriel.bobeica@ecb.europa.eu

**Silviu Oprică**
European Central Bank, Frankfurt am Main, Germany; email: silviu.oprica@ecb.europa.eu



Postal address 60640 Frankfurt am Main, Germany
Telephone +49 69 1344 0
Website www.ecb.europa.eu



PDF ISBN 978-92-899-7845-3 ISSN 1725-2806 doi:10.2866/6754206 QB-01-26-116-EN-N