Cluster Analysis of the EU Countries during the Covid-19 Period (2019–2021)

Andrii Hrabariev[1] | *Kyiv National Economic University named after Vadym Hetman, Kyiv, Ukraine*
Vasyl Derbentsev[2] | *Kyiv National Economic University named after Vadym Hetman, Kyiv, Ukraine*
Mykhailo Baraniuk[3] | *Kyiv National Economic University named after Vadym Hetman, Kyiv, Ukraine*

Received 11.8.2025 (revision received 26.11.2025), **Accepted** (reviewed) 11.12.2025, **Published** 11.9.2026

Abstract

This study investigates structural shifts in European Union countries banking system typologies during the COVID-19 period from 2019 to 2021. Using the k-means clustering method (compared with hierarchical agglomerative clustering and Gaussian mixture models) on 10 financial and macroeconomic indicators (e.g., non-performing loans, z-score, GDP growth) at the country level for 25 European Union countries, we identified four distinct clusters with stable typological definitions but evolving country compositions. Comparison of cluster profiles reveals the pandemic period as a structural catalyst, reinforcing the resilience of Nordic Resilience Systems) and Core Stability Systems clusters. Significant compositional shifts occurred: Dynamic Growth Systems absorbed advanced economies, balancing growth and stability, while High Intermediation Systems transformed into a vulnerable typology, marked by high fiscal risk and low stability buffers. These findings highlight structural shifts associated with the pandemic period and differentiated policy responses, and inform regulatory strategies for enhancing EU banking resilience.

Keywords	*DOI*	*JEL code*
EU banking systems, COVID-19, cluster analysis, financial resilience, banking and macroeconomic indicators, structural shifts, temporal analysis.	*https://doi.org/10.54694/stat.2025.40*	*G21, G28, C38*

INTRODUCTION

The stability and efficiency of banking systems are paramount to the economic health and resilience of any region, particularly within a highly integrated economic bloc like the European Union (EU).

[1] Andrii Hrabariev, PhD, Associated Professor, Department of Informatics and Systemology, Kyiv National Economic University named after Vadym Hetman, Prospekt Beresteiskyi 54/1, Kyiv, 03057, Ukraine. E-mail: andrgrab@gmail.com, phone: +380 68 158 8558. ORCID: https://orcid.org/0000-0001-6165-0996.

[2] Vasyl Derbentsev, PhD, Professor, Department of Informatics and Systemology, Kyiv National Economic University named after Vadym Hetman, Prospekt Beresteiskyi 54/1, Kyiv, 03057, Ukraine. Corresponding author: e-mail: derbv@kneu.edu.ua, phone: +380 67 407 6179. ORCID: https://orcid.org/0000-0002-8988-2526.

[3] Mykhailo Baraniuk, PhD student, Department of Informatics and Systemology, Kyiv National Economic University named after Vadym Hetman, Prospekt Beresteiskyi 54/1, Kyiv, 03057, Ukraine.

Defined by a complex interplay of macroeconomic conditions, regulatory frameworks, and inherent financial sector characteristics, the performance and robustness of national banking sectors significantly influence overall financial stability (Demirgüç-Kunt et al., 2021).

The period between 2019 and 2021 presents a unique and critical juncture for examining these dynamics. The year 2019 serves as a baseline, representing the pre-pandemic financial landscape, while 2021 captures the immediate aftermath and initial recovery phase from the unprecedented COVID-19 shock. This period was also characterized by massive monetary and fiscal policy interventions, including the European Central Bank's Pandemic Emergency Purchase Programme (PEPP) and extensive loan moratoriums, designed to buffer banking systems from the pandemic's direct impacts (ECB, 2021; EBA, 2020). Understanding the shifts and resilience within EU banking systems during this turbulent period is crucial for policymakers, regulators, and market participants alike (EBA, 2020).

Traditional assessments of financial stability often rely on aggregate indicators or country-specific analyses. However, a more granular understanding can be achieved through unsupervised machine learning techniques, specifically clustering, which group countries based on their inherent financial similarities (Mihalech & Košíková, 2021; Poznyak & Kolyada, 2023). Such methods, applied to comprehensive financial and macro-financial indicators, can reveal distinct banking system typologies and underlying structural similarities not immediately apparent through conventional analysis.

This study aims to investigate the stability and evolution of banking system typologies across EU countries (excluding Greece and Luxembourg, identified as consistent outliers) over the 2019–2021 period. Utilizing the k-means clustering method (K-Means), compared with hierarchical agglomerative clustering (HAC) and Gaussian mixture models (GMM), we identify four robust typologies based on ten key financial and macro-financial indicators. Our approach examines how these typologies, defined by characteristic profiles of asset quality, capitalization, and macroeconomic exposure, transformed in country composition and risk characteristics during the pandemic period. Country-level analysis captures systemic patterns informing macroprudential policy, though it necessarily aggregates over within-country heterogeneity.

Considering these objectives, we propose the following *working hypotheses:*

H1: Resilience of Core Clusters. Despite the exogenous shock of the COVID-19 pandemic and associated policy responses, core clusters of financially stable and mature Western European banking systems, as well as highly resilient Nordic banking systems, will demonstrate a significant stability in their composition from 2019 to 2021, exhibiting robust financial health indicators throughout the period.

H2: Emergence of Differentiated Vulnerabilities. The pandemic period will be associated with re-clustering of countries with pre-existing or newly exposed vulnerabilities, particularly among Central and Eastern European (CEE) nations and smaller, tourism-dependent economies. These newly formed or significantly altered clusters will exhibit higher levels of non-performing loans (NPLs), increased government debt, and potentially lower z-scores in 2021 compared to 2019.

H3: Dynamic Adaptation of Specialized Systems. Countries previously identified as having "specialized" or "transitional" banking systems will exhibit diverse adaptive responses to the pandemic period, leading to their re-alignment into either more stable core groups or new clusters with altered risk profiles defined by specific emerging risks or opportunities. This will manifest in notable shifts in their cluster membership and indicators mean values.

By empirically testing these hypotheses through a comparative analysis of clustering results and the dynamics of key financial indicators, this research aims to contribute significantly to the understanding of banking system resilience and the evolving financial architecture of the European Union during periods of major economic shocks combined with unprecedented policy interventions.

1 LITERATURE SURVEY

The COVID-19 pandemic significantly impacted EU banking systems, prompting extensive research on their resilience, typologies, and policy responses during 2019–2021.

Robust pre-crisis capital buffers (CET1 ~15% in 2019 vs. ~7% in 2008) mitigated losses during the pandemic (EBA, 2020). Dursun-de Neef and Schandlbauer (2021) found that worse-capitalized banks increased lending by ~10% to avoid loan loss recognition, while better-capitalized banks faced a ~15% rise in delinquent loans. Cao and Chou (2022) confirmed that banks with higher regulatory capital (CAR >12%) lent more resiliently, particularly in Nordic and Core Western European clusters, maintaining NPL ratios below 2% (Demirgüç-Kunt et al., 2021). Clichici and Zeldea (2021) highlighted the resilience of Central and Eastern European (CEE) banks, with low NPL ratios (~2%) and high capitalization.

Alessi et al. (2022) noted challenges in NPL provisioning in Southern banks due to GDPG declines (~6.1% in 2020), while Ari et al. (2021) reported a ~20% increase in delinquent loans in vulnerable systems. Kozak (2021) emphasized equity ratios' role in the CEE resilience, and Salazar et al. (2023) linked IFRS 9 adoption to improved z-score stability in Western European banks. Foglia et al. (2022) underscored high volatility connectedness in Eurozone banks.

Clustering methodologies reveal EU banking heterogeneity. Mihalech and Košíková (2021) applied clustering to EBA risk indicators, identifying groups (eastern, southern, northern, core). Danko and Suchý (2021) used clustering to assess financial integration, identifying leaders and lagging systems, supporting our typology.

Kaminskyi and Nehrey (2023) applied clustering to assess US banks' resilience to COVID-19 shocks, using indicators like total assets and recovery rates, reinforcing clustering applicability for identifying resilience patterns, despite focusing on a non-EU context. Our study extends these by using a comprehensive set of financial and macroeconomic indicators (see Table A1), capturing longitudinal shifts across 2019–2021.

Monetary and fiscal interventions stabilized EU banks. The European Central Bank's Pandemic Emergency Purchase Programme (PEPP, €1.85 trillion) reduced funding pressures (ECB, 2021). EBA (2020) reported that loan moratoriums increased CAR by ~1.5% and liquidity by ~10%, stabilizing NPL at ~2.5% by Q4 2020 (EBA, 2021). Feyen et al. (2021) highlighted PEPP's role in mitigating GDEBT pressures (>100% in Southern Europe), though IMF (2021) noted that €2 trillion in loan guarantees raised fiscal risks. Demir and Danisman (2021) emphasized fiscal support's role in liquidity stability. Colak and Öztekin (2021) found ECB policies boosted lending by ~15% in resilient banks.

The macroeconomic environment shaped banking responses. The World Bank (2022) reported a 6.1% GDP contraction in 2020, with unemployment rate (UNEMP) peaking at 7.8% in 2020. Chodnicka-Jaworska (2022) found that GDPG declines and UNEMP spikes delayed bank rating downgrades by 6–12 months. Plikas et al. (2024) showed that lockdown strictness increased SME NPLs by ~30% in Southern Europe.

Existing literature provides insights but has gaps. Studies like Dursun-de Neef and Schandlbauer (2021), Alessi et al. (2022), and Foglia et al. (2022) focus on bank performance but overlook structural shifts. Mihalech and Košíková (2021), Danko and Suchý (2021), and Kaminskyi and Nehrey (2023) apply clustering but use fewer indicators and different contexts.

Our study addresses these gaps by clustering the EU-25 (excluding Greece and Luxembourg) using 10 indicators from TheGlobalEconomy.com (2025) and the World Bank's Global Financial Development Database (2025). We applied K-Means compared with the results obtained by other methods to identify dynamic typologies and provide a basis for developing regulatory strategies.

2 METHODS

This study applies a multistage methodology designed to explore the structural characteristics and resilience of the banking systems of the EU member states under the influence of the COVID-19 pandemic. The methodological approach is focused on multivariate statistical analysis, including feature selection techniques and unsupervised clustering algorithms, with the goal of identifying stable groupings of countries and detecting any systemic shifts across selected time periods (Hastie et al., 2009).

2.1 Research Design and Periodization

The analytical design is structured around two reference years: 2019 (pre-COVID baseline) and 2021 (post-pandemic adaptation). This temporal segmentation allows for comparison of structural changes in national banking systems across distinct phases of external pressure. By employing a uniform set of indicators and methodological procedures across these years, we ensure internal consistency and comparability of the clustering results.

The overall process consists of the following key stages: (1) selection and grouping of relevant indicators; (2) data collection and pre-processing; (3) dimensionality reduction and indicators filtering; and (4) cluster analysis and validation.

2.2 Data Sources and Pre-processing

Data were sourced from TheGlobalEconomy.com, an aggregator providing access to harmonized socio-economic indicators from official databases, including the World Bank's Global Financial Development Database, the European Central Bank's Statistical Data Warehouse, and the IMF DataMapper.

The analysis covers 25 EU member states, excluding Greece and Luxembourg. Preliminary clustering of all EU27 countries revealed that Luxembourg consistently acted as a standalone outlier cluster across all three methods (K-Means, HAC, and GMM). Similarly, Greece was identified as an outlier by both the K-Means and HAC algorithms. These exclusions were methodologically justified, as Luxembourg's exceptionally large banking sector relative to its domestic economy (deposits at 413.21% of GDP) and Greece's unique post-sovereign debt crisis financial profile (high non-performing loans at 38.07% in 2019) would disproportionately skew cluster formation and distort the analysis of the remaining EU member states.

Missing values (e.g., bank assets for Cyprus, bank concentration for Germany) were imputed using regional medians. Outliers were identified via the interquartile range (IQR) method and winsorized to preserve variance.

All indicators x_{ij} were normalized to the range [0,1] using the min-max normalization:

$$x_{ij}^{norm} = \frac{x_{ij} - \min\left(x_j\right)}{\max\left(x_j\right) - \min\left(x_j\right)},\ i = 1,2,\ldots,n, j = 1,2,\ldots,m, \tag{1}$$

where $\max(x_j)$, $\min(x_j)$ – are the minimum and maximum values of the indicators across all countries n – number of countries, m – number of indicators.

2.3 Indicator Selection

This study initially considered 15 indicators to capture the resilience, efficiency, and macroeconomic context of EU banking systems, grouped into three analytical blocks: (1) banking system resilience and stability (e.g., z-score, non-performing loans ratio (NPL)), (2) access to and efficiency of financial intermediation (e.g., bank credit to private sector), and (3) macroeconomic and institutional environment (e.g., GDP growth, government debt). These indicators were selected based on their theoretical significance, empirical use in prior financial stability studies (Alessi et al., 2022; Chodnicka-Jaworska, 2022; Colak & Öztekin, 2021; Demirgüç-Kunt et al., 2021), and data availability across EU members.

The selection of the final set of indicators was a critical step to ensure the reliability and interpretability of the cluster analysis, as excessive correlation (multicollinearity) between indicators can distort clustering results (Hastie et al., 2009). To mitigate multicollinearity and enhance clustering performance, a hybrid feature selection approach was applied:

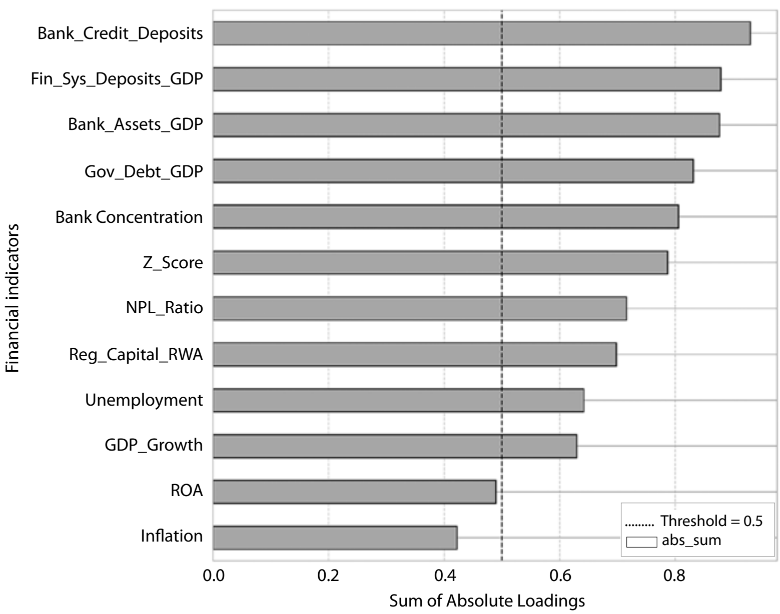
Stage 1: Correlation Analysis and Indicator Elimination. We conducted a detailed Pearson correlation analysis among all 15 indicators for both periods (2019 and 2021) to identify the redundant. Indicators with a correlation coefficient exceeding $|\rho| > 0.7$ with others were identified as candidates for exclusion. In such cases, we retained the indicator with greater theoretical relevance or a higher contribution to the variance. Through this process, three indicators were excluded:

1. Bank Credit to Private Sector (% GDP) – high correlation with Bank Assets to GDP ($\rho = 0.83$). We retained BAGDP as it more directly captures overall banking sector size relative to the economy.
2. Current Account Balance (% GDP) – high correlation with Government Debt ($\rho = -0.68$) and FDGDP ($\rho = 0.71$), reflecting fiscal-financial linkages. We retained GDEBT as it more directly captures sovereign risk affecting banking systems.
3. Foreign Direct Investment (% GDP) – moderate correlations with multiple indicators (GDPG: $\rho = 0.65$, BAGDP: $\rho = 0.58$) and less direct impact on banking system resilience compared to domestic credit dynamics.

Following correlation-based elimination, 12 indicators remained for further analysis.

Stage 2: Principal Component Analysis (PCA) Validation. Second, PCA assessed the variance explained by the indicators (Figure 1).

Figure 1 Sum of Absolute PCA Loadings for 12 Indicators



Source: Own construction

The effectiveness of PCA for dimensionality reduction in economic data has been empirically confirmed by recent comparative studies. Poznyak and Kolyada (2023) evaluated 11 dimensionality reduction methods and found PCA to demonstrate superior performance in preserving data structure while reducing complexity, particularly for macroeconomic and financial indicators. This supports our choice of PCA for feature selection validation in the present study.

PCA was applied to the remaining 12 indicators to assess the variance explained (Figure 1). The first four PCA components accounted for approximately 80% of the total variance, with key contributions from CAR, financial system deposits to GDP ratio (FDGDP), bank concentration (BCONC), and confirming the relevance of the selected indicators.

Two indicators with the lowest cumulative loadings – Return on Assets (ROA) and Inflation Rate (CPI) – were excluded. ROA's exclusion is justified as it is already incorporated as a component of the z-score, while CPI showed less direct relevance to banking system stability during the COVID-19 period characterized by supply shocks. This yielded a final set of 10 indicators (see Table A1 in Appendix).

This two-stage approach ensures that each of the final 10 indicators contributes to unique information while maintaining low multicollinearity (all pairwise $|\rho| < 0.7$), providing accurate and interpretable cluster demarcation.

2.4 Cluster Analysis and Validation of Results

The core stage of the analysis involved the application of unsupervised clustering algorithms to the normalized data (Jain, 2010). To ensure robustness of the identified country groupings, we conducted experimental validation using three clustering methods: K-Means, Hierarchical agglomerative clustering (HAC) with Ward's linkage, and Gaussian Mixture Models (GMM).

Each method embodies different assumptions, serving as validation where no ground truth exists (Hastie et al., 2009):

- *K-Means*, a centroid-based method, minimizes the sum of squared distances between data points and cluster centroids. The algorithm used the Euclidean distance and random centroid initialization.
- *HAC* constructs a dendrogram with Ward's linkage and squared Euclidean distance, minimizing within-cluster variance (Rousseeuw & Kaufman, 1990). The dendrogram was cut at $k = 4$.
- *GMM* – modelled data as a mixture of multivariate normal distributions, capturing potential elliptical clusters (McLachlan & Peel, 2000). A full covariance matrix was used, with $k = 4$ selected based on the Bayesian Information Criterion (BIC).

  Each method embodies different assumptions: the k-means method assumes spherical clusters, hierarchical agglomerative clustering assumes hierarchical structure, and Gaussian mixture models assume elliptical, probabilistic clusters.

Each country is represented as a data point – a vector of normalized indicator values. Cluster centroids represent mean indicator values for countries within each cluster, defining typical banking system profiles. Preliminary experiments demonstrated that all three methods produced largely consistent country groupings, with only minor variations at cluster boundaries. When $k = 4$ clusters were specified for both 2019 and 2021 periods, the methods achieved over 85% agreement in country assignments. This cross--method convergence confirms strong evidence that the identified clusters represent genuine structural patterns in EU banking systems rather than methodological artifacts.

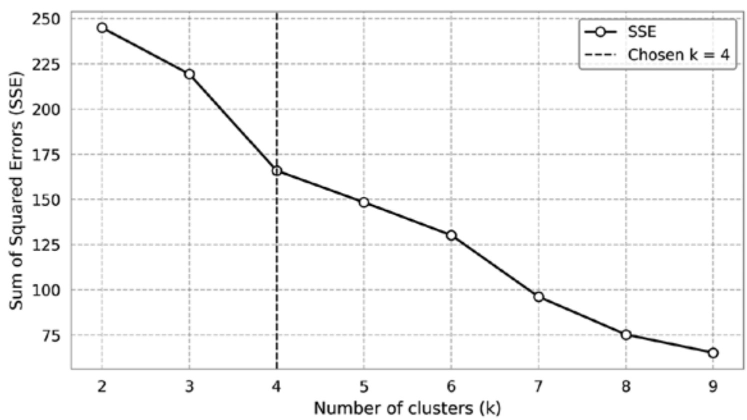
K-Means was selected as the principal method for: (1) Clear and interpretable partition through distance-based logic and defined centroids representing "typical" banking system profiles (Jain, 2010); (2) Computational efficiency and reproducibility, allowing iterative comparison across periods (Hastie et al., 2009); (3) Superior empirical performance: (Poznyak & Kolyada, 2023; Osadcha et al., 2025) found K-Means the most effective, with agglomerative clustering ranking second; (4) Euclidean distance and normalized indicators satisfied K-Means assumptions; (5) Nearly identical results to GMM and HAC, confirming robustness; (6) Established precedent in banking sector research (Mihalech & Košíková, 2021; Kaminskyi & Nehrey, 2023).; (7) The final typological classification is based on K-Means, which combines interpretability, computational simplicity, and empirical robustness.

We chose $k = 4$ as the optimal number of clusters despite the silhouette score suggesting higher values ($k = 8$) because the elbow method indicated a significant reduction in the sum of squared errors (SSE) at this point (Figure 2), marking a clear "elbow" in the curve (Rousseeuw & Kaufman, 1990).

Additionally, considering the specific context of clustering EU countries, a smaller number of clusters offers a more interpretable and actionable segmentation, balancing model simplicity with meaningful differentiation among groups. This decision aligns with the practical goal of deriving clear insights for policy or economic analysis, where overly granular clusters might obscure broader patterns.

Clustering results were validated by: (1) Cluster profiles: average indicator values characterized financial profiles (Table 1, Table 2); (2) Temporal mapping: country transitions from 2019 to 2021 assessed stability and reconfiguration during the pandemic period (Figure 5, detailed country migrations in Table A4); (3) method consistency: high agreement (>85%) among K-Means, HAC, and GMM were confirmed robust typologies (Jain, 2010).

Figure 2 Results of the elbow method for determining the optimal number of clusters based on the k-means method



Source: Own construction

3 EMPIRICAL RESULTS

This section presents the clustering analysis results for EU banking systems in 2019 (pre-pandemic baseline) and 2021 (post-pandemic), followed by a comprehensive temporal analysis of structural transformations induced by the COVID-19 crisis.

Data processing and analysis were performed in Google Colab using Python 3.10 with libraries: *pandas* and *numpy* for data manipulation; *scikit-learn* for normalization, PCA, K-Means, HAC, GMM, and validation metrics; *matplotlib* and *seaborn* for visualization; *statsmodels* and *scipy* for correlation and statistical validation.

3.1 Baseline Clusters Analysis (2019)

The application of K-Means to the 2019 data consistently reveals four distinct banking system typologies within the EU. The results of K-Means are shown in Figure 3 (projection on the two principal components), and Table 1 presents the cluster composition and key characteristic profiles for each identified banking system typology.

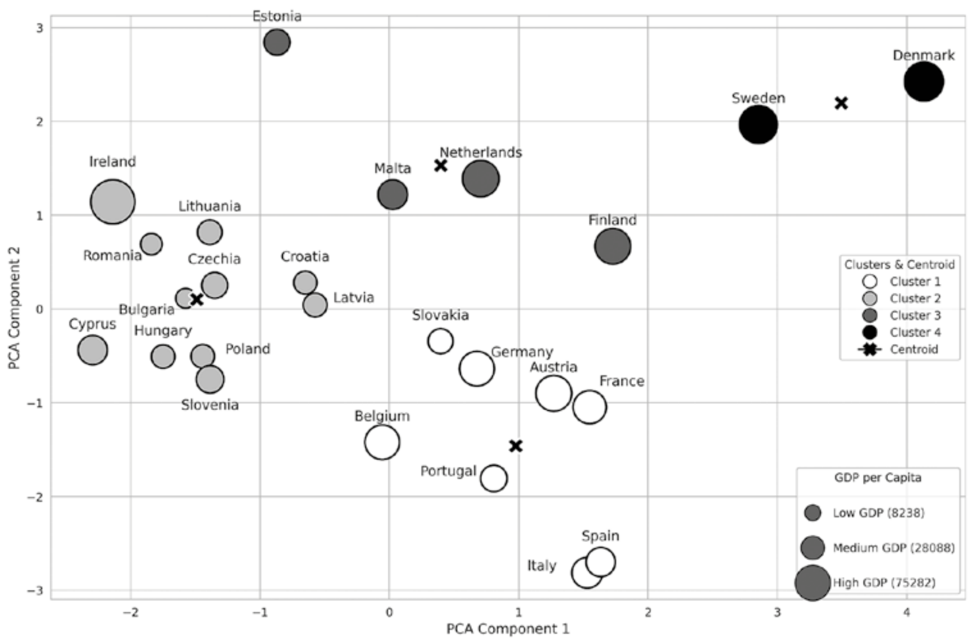
The Core Stability System (CSS) cluster (Cluster 1), comprising major Eurozone economies (Germany, France, Italy), represents the stable backbone of the EU financial system. This cluster demonstrates exceptional macroeconomic stability with the lowest unemployment (4.28%) and moderate government debt (50.44%). Their banking systems exhibit healthy asset quality (NPL: 2.96%), robust capital adequacy (CAR: 20.38%), and solid stability buffers (z-score: 14.44), establishing them as the benchmark for mature, well-regulated financial intermediation.

Table 1 EU Banking System Cluster Composition and Profiles (2019)

Name/Abbreviation		Countries	Profile
Cluster 1	Core Stability Systems (CSS)	Austria, Belgium, Germany, Spain, France, Italy, Portugal, Slovakia	Mature, stable: NPL 2.96%, CAR 20.38%, z-score 14.44, GDEBT 50.44%
Cluster 2	Dynamic Growth Systems (DGS)	Bulgaria, Cyprus, Czechia, Croatia, Hungary, Ireland, Lithuania, Latvia, Poland, Romania, Slovenia	Growth-oriented: GDPG 3.53%, NPL 5.44%, CAR 19.47%, z-score 13.56
Cluster 3	High Intermediation Systems (HIS)	Estonia, Finland, Malta, Netherlands	Deep markets: BAGDP 100.45%, LADR 114.00%, NPL 2.42%, z-score 14.62
Cluster 4	Resilience Systems (RS)	Denmark, Sweden	Exceptional: z-score 31.54, NPL 0.98%, GDEBT 34.65%

Source: Own construction. Note: All indicators represent cluster means. For detailed values of all 10 indicators, see Table A2 in ANNEX,

Figure 3 K-Means Cluster Visualization on PCA Components (Excluding Luxembourg and Greece as outliers, Point Size by GDP per Capita, 2019)



Source: Own construction

The Dynamic Growth System (DGS, Cluster 2) reflects economies balancing growth dynamism with elevated credit risk (Cluster 2). This group, including most Central and Eastern European countries plus

Ireland, demonstrates the highest GDP growth (3.53%) but also the highest NPL ratio (5.44%), slightly exceeding optimal ranges. Despite strong capital adequacy (CAR: 19.47%), their lower z-score (13.56) indicates reduced stability buffers. Ireland's consistent inclusion highlights its unique financial system characteristics that align with developmental banking patterns despite high GDP per capita.

The High Intermediation Systems (HIS, Cluster 3) features highly sophisticated financial sectors with deep market penetration. Comprising Estonia, Finland, Malta, and Netherlands, this cluster is characterized by significantly elevated banking sector depth relative to GDP. These economies show significantly higher Bank Assets to GDP (100.45%) and elevated loan-to-deposit ratios (114.00%), indicating extensive financial intermediation, sophisticated non-deposit funding sources, and significant cross-border banking operations. Despite sector complexity, they maintain excellent asset quality (NPL: 2.42%) and strong capital positions (CAR: 20.40%), resulting in healthy z-scores (14.62). The combination of deep financial markets, low credit risk, and robust capitalization distinguishes this cluster from other banking system typologies.

The Resilience Systems (NRS, Cluster 4), consisting solely of Denmark and Sweden, represent the pinnacle of financial resilience. This archetype exhibits exceptional characteristics: the lowest government debt (34.65%), minimal NPLs (0.98%), and an outstanding z-score (31.54) indicating virtually no insolvency risk. Their massive banking sectors (BAGDP: 154.93%, LADR: 227.56%) reflect sophisticated international operations managed under superior prudential frameworks, establishing them as the gold standard for banking system stability.

3.2 Post-Pandemic Cluster Evolution (2021)

To ensure robustness and comparability, the clustering analysis for 2021 was conducted using the same set of indicators and methodology (K-Means) as the 2019 baseline. Furthermore, Greece and Luxembourg were again excluded from the primary analysis due to their persistent status as outliers, often forming singleton clusters or significantly skewing the results.

It is important to emphasize that cluster numbering and naming remain consistent across both periods (2019 and 2021), as presented in Tables 1 and 2. This approach allows us to trace structural transformations in the EU banking systems by observing changes in country composition within fixed cluster typologies. The movement of countries between clusters reveals the pandemic's role as a structural catalyst, reshaping banking system alignments rather than merely causing temporary disruptions.

By comparing the clustering solutions for 2019 and 2021, we can assess the stability of the identified clusters and pinpoint which country groupings remained resilient and which were reshuffled due to the crisis. This analysis provides key insights into the differential effects of the pandemic on various banking systems.

The results of K-Means for 2021 are presented in Figure 4, and the full results of the cluster assignments by method are presented in ANNEX (Table A3). For the main analysis, however, we focus on the consolidated profiles and their temporal evolution, with detailed cluster composition and profiles in Table 2.

The 2021 analysis, employing identical methodology, reveals significant structural reconfiguration following the COVID-19 shock. While maintaining four clusters, country compositions and characteristic profiles transformed substantially, reflecting differentiated pandemic impacts and recovery trajectories.

The Core Stability Systems (CSS, Cluster 1) expanded to include, demonstrating the core's attractive stability during crisis recovery. This group achieved a strong economic rebound (GDPG: 5.96%) while maintaining sound banking fundamentals. Despite increased fiscal burden (GDEBT rose from 50.44% to 66.61% due to pandemic support measures), they improved asset quality (NPL decreased to 2.14%) and strengthened capital positions (CAR: 21.94%). The modest z-score decline (from 14.44 to 13.68) reflects controlled absorption of pandemic shock.

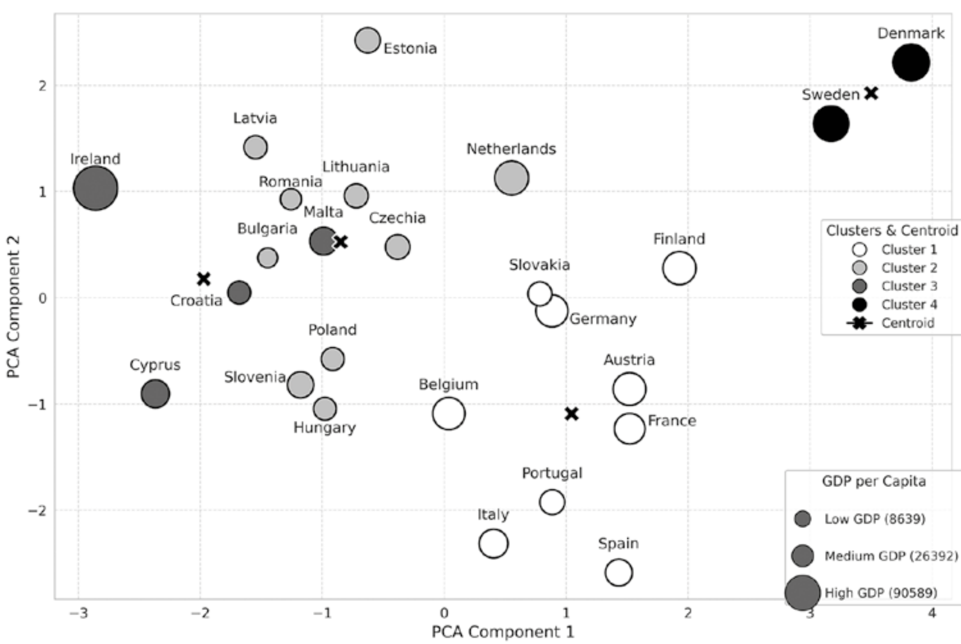
Table 2 EU Banking System Cluster Composition and Profiles (2021)

Name/Abbreviation		Countries	Profile
Cluster 1	Core Stability Systems (CSS)	Austria, Belgium, Germany, Spain, Finland, France, Italy, Portugal, Slovakia	Recovery: GDPG 5.96%, NPL 2.14%, CAR 21.94%, z-score 13.68, GDEBT 66.61%
Cluster 2	Dynamic Growth Systems (DGS)	Bulgaria, Czechia, Estonia, Hungary, Lithuania, Latvia, Netherlands, Poland, Romania, Slovenia	High growth: GDPG 8.27%, NPL 3.71%, CAR 20.68%, z-score 12.46
Cluster 3	High Intermediation Systems (HIS)	Cyprus, Croatia, Ireland, Malta	Vulnerable: GDPG 8.76%, GDEBT 88.25%, z-score 10.17, NPL 3.59%
Cluster 4	Nordic Resilience Systems (NRS)	Denmark, Sweden	Strengthened: z-score 32.89, NPL 0.83%, GDEBT 36.35%

Source: Own construction. Note: Italic indicates new cluster members in 2021 compared to 2019. All indicators represent cluster means. For detailed values of all 10 indicators, see Table A3 in ANNEX.

A critical reconfiguration occurred within the Dynamic Growth Systems (DGS, Cluster 2) – it began to contain countries from the former 2019 DGS cluster alongside Estonia and Netherlands, which migrated from the 2019 HIS cluster. This convergence was driven by shared rapid recovery dynamics (GDPG: 8.27%), and represents a fundamental typological shift where advanced economies aligned with high-growth markets. However, this dynamism carries costs: elevated NPLs (3.71%) and the lowest z-score among stable clusters (12.46%) indicate reduced stability buffers despite strong growth momentum.

Figure 4 K-Means Cluster Visualization on PCA Components (Excluding Luxembourg and Greece as outliers, Point Size by GDP per Capita, 2021)



Source: Own construction

The High Intermediation Systems (HIS, Cluster 3) experienced the most substantial reconfiguration. In 2019, it consisted of highly integrated and low-risk banking systems (Estonia, Finland, Malta, Netherlands). By 2021, its composition shifted almost entirely toward peripheral and island economies (Cyprus, Croatia, Ireland, Malta), with only Malta remaining from the original group. The 2021 HIS cluster exhibits high government debt levels (88.25%) and markedly weaker stability buffers (z-score: 10.17). This shift reflects not a change in the cluster's structural definition – high financial intermediation relative to economic size – but rather uneven pandemic-era impact on countries with greater exposure to financial-sector fragilities.

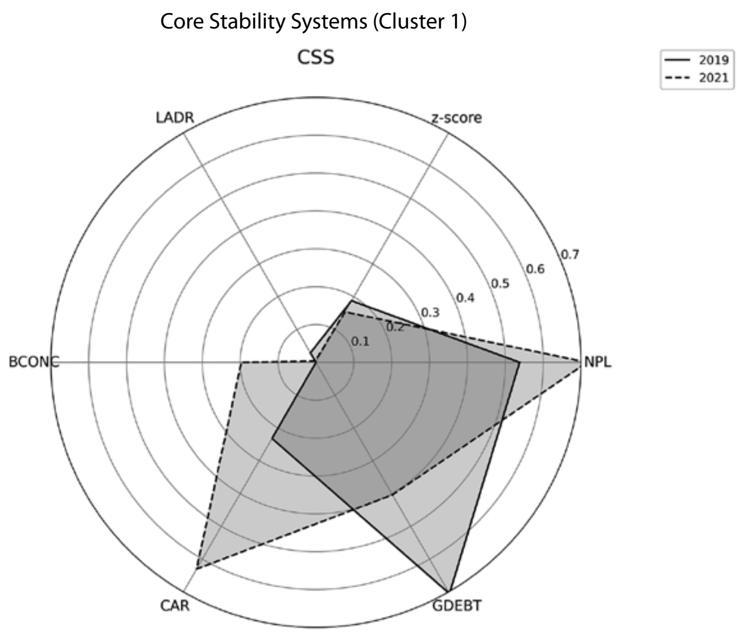
The Nordic Resilience Systems (NRS, Cluster 4) remained unchanged and strengthened its exceptional position. Post-pandemic metrics show further improvement: z-score increased to 32.89, NPLs decreased to 0.83%, while maintaining minimal fiscal burden (GDEBT: 36.35%). This paradoxical strengthening during global crisis confirms their unique resilience and superior risk management capabilities.

3.3 Temporal Analysis of Cluster Profile Changes (2019 vs. 2021)

The comparative analysis of 2019–2021 cluster profiles reveals the COVID-19 pandemic's role as a structural catalyst, creating both reinforcement of existing strengths and emergence of new vulnerabilities. Figure 5 visualizes these transformations across six key indicators, demonstrating how different banking system archetypes absorbed and adapted to the exogenous shock.

As established in Section 3.2, cluster numbering and naming remain fixed across both periods, enabling direct comparison of how identical cluster typologies evolved in their characteristic profiles and country composition.

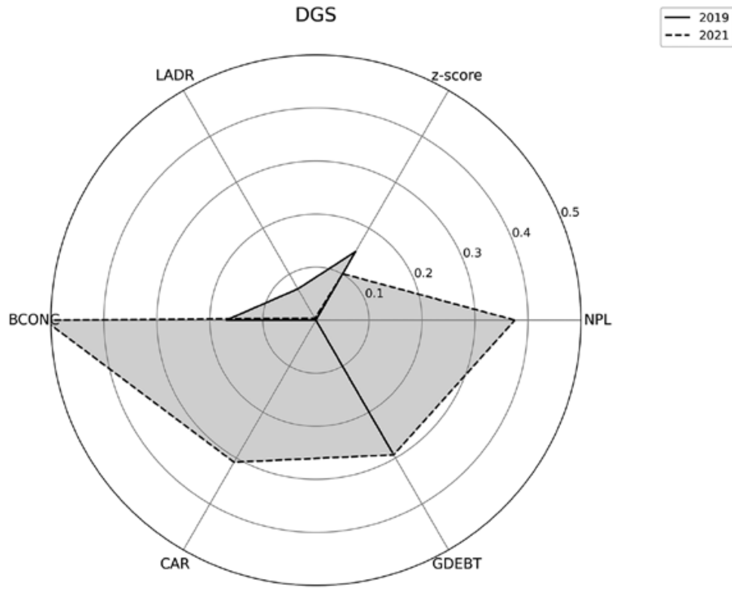
Figure 5 Temporal profile changes of EU Banking Clusters (2019 vs. 2021, Normalized values with inverted NPL and GDEBT ratios)



Source: Own construction

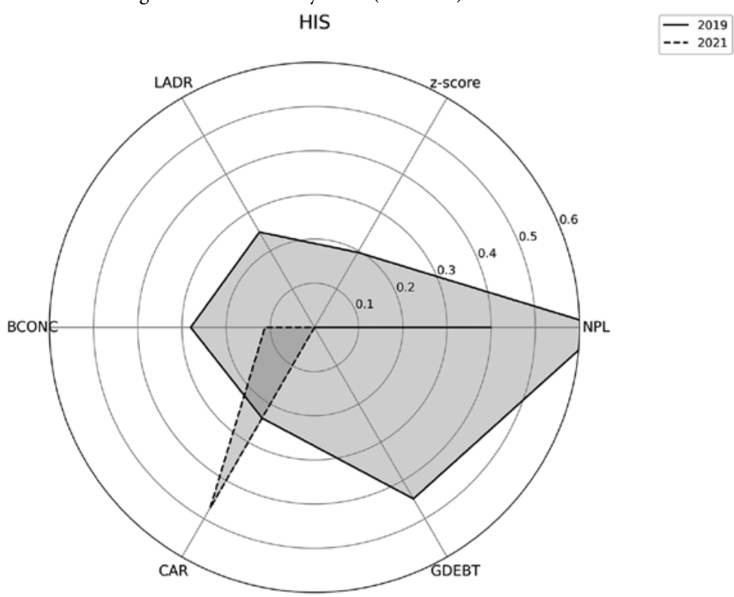
Figure 5 (continuation)

Dynamic Growth Systems (Cluster 2)



Source: Own construction

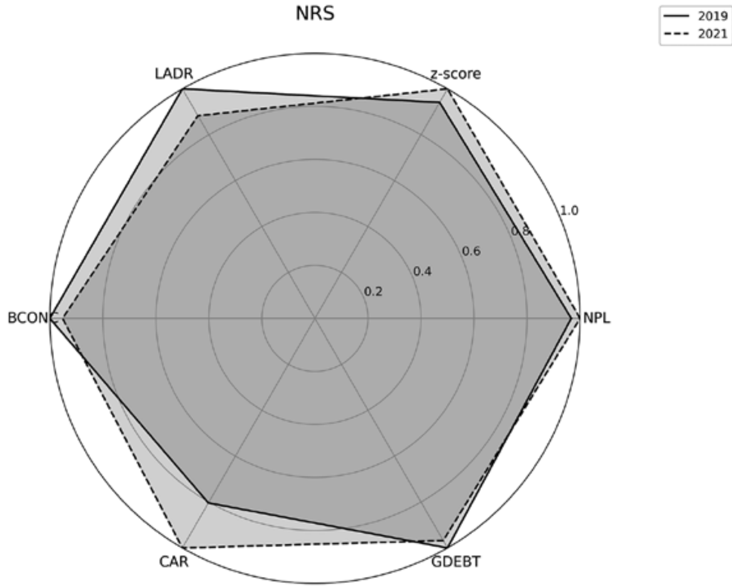
High Intermediation Systems (Cluster 3)



Source: Own construction

Figure 5 (continuation)

Nordic Resilience Systems (Cluster 4)



Source: Own construction

The radar chart presents normalized profiles for each of the four clusters, allowing direct visual comparison of structural changes. The most striking transformations occurred in Clusters 3 (HIS) and 2 (DGS), while Clusters 1 (CSS) and 4 (NRS) demonstrated remarkable stability.

The Core Stability Systems evolution (Cluster 1, top-left chart) demonstrates controlled shock absorption by mature economies. While z-score declined moderately by 5.3% (14.44→13.68), reflecting reduced stability buffers, these systems successfully managed credit risk with NPL improvements of 27.7% (2.96%→2.14%). The substantial GDEBT increase of 32.1% (50.44%→66.61%) reflects massive fiscal intervention, highlighting the trade-off between immediate economic support and long-term fiscal health. Decreased concentration and deleveraging (LADR decline) suggest strategic market adjustments.

The Dynamic Growth Systems (Cluster 2) retained its core CEE composition while absorbing Estonia and the Netherlands from the former HIS group. This configuration shows the strongest growth (GDPG: 8.27%) but with reduced stability buffers (z-score: 13.56→12.46), indicating that rapid recovery was accompanied by increased systemic sensitivity. Asset quality improved (NPLs: 5.44%→3.71%), suggesting genuine economic rebound rather than excessive credit expansion. The convergence of advanced financial systems (Estonia, Netherlands) with CEE economies represents a structural realignment driven by post-crisis growth dynamics.

The High Intermediation Systems (Cluster 3) experienced the most substantial transformation. All original 2019 members except Malta migrated to other clusters, while Cyprus, Croatia, and Ireland joined. In 2021, government debt rose to 88.25%, stability buffers weakened sharply (z-score: 10.17), and NPLs increased (2.42%→3.59%). The reduction in loan-to-deposit ratios (114.00%→76.44%) signals forced deleveraging and diminished intermediation capacity. The cluster shifted from a low-risk profile to a high-risk, fiscally constrained configuration requiring closer supervisory attention, while maintaining high financial intermediation (BAGDP: 88.88%) (see Table A4).

The Nordic Resilience Systems (Cluster 4, top-left chart) defies conventional crisis expectations. Denmark and Sweden not only maintained their exceptional position but paradoxically strengthened it: z-score increased by 4.2% (31.54→32.89), NPLs decreased by 15.3% (0.98%→0.83%), while fiscal position remained virtually unchanged (+4.9% GDEBT increase). This counter-cyclical performance establishes Nordic systems as uniquely resilient, capable of strengthening during the global financial stress through superior institutional frameworks and risk management.

These temporal dynamics confirm all three research hypotheses: Nordic and Core Stability clusters demonstrated resilience (H1), new vulnerabilities emerged in reconfigured groupings (H2), and specialized systems underwent dynamic adaptation through cluster realignment (H3). The pandemic acted not merely as a temporary shock but as a structural transformation catalyst, creating new banking system typologies that reflect post-crisis realities of differentiated resilience, fiscal burden, and growth-stability trade-offs across the European Union.

4 DISCUSSION

This study identifies clear structural shifts in the EU banking system typologies between 2019 and 2021. These changes must be interpreted considering the unprecedented fiscal and monetary interventions implemented during the COVID-19 period, including the ECB's PEPP programme, loan moratoria, and extensive guarantee schemes. As these measures mitigated direct balance-sheet deterioration, the transformations captured in our analysis represent the *combined* effect of pandemic pressures and policy support rather than the pandemic in isolation.

Country-level aggregation allows us to detect systemic patterns and cross-national typologies, though it necessarily masks the within-country heterogeneity. For our purpose, identifying structural banking system configurations, this macro-level approach is appropriate, but bank-level analyses would complement our findings.

The results provide strong support for *Hypothesis 1:* the Nordic Resilience Systems (NRS) and Core Stability Systems (CSS) clusters maintained robust stability buffers, with improved asset quality and limited deterioration in intermediation metrics. These systems demonstrate the capacity of mature institutional environments to absorb large shocks when combined with substantial policy support.

Hypothesis 2 is confirmed by the transformation of the High Intermediation Systems (HIS). In 2021, this cluster (Cyprus, Croatia, Ireland, Malta) exhibited the highest sovereign debt levels and weakest stability indicators. Countries that migrated into HIS were characterized by a strong sovereign–bank nexus and elevated exposure to pandemic-related fiscal expansion, making the cluster the most risk-sensitive typology.

Hypothesis 3 is supported by the adaptive evolution of the Dynamic Growth Systems (DGS), which retained its CEE core while integrating Estonia and the Netherlands. This convergence produced a fast-growing but moderately fragile typology, combining high post-pandemic expansion with slightly reduced stability buffers.

Limitations. Several limitations must be acknowledged: (1) Aggregate country data obscure within-country banking heterogeneity; (2) a two-period observational design prevents causal attribution, as structural changes may also reflect pre-existing trends or heterogeneous policy capacities; (3) Greece and Luxembourg were excluded due to outliers' extreme characteristics, limiting generalizability; (4) the period 2019–2021 captures only immediate adjustments, with longer-term effects uncertain; and (5) qualitative factors (supervisory intensity, regulatory quality) are not incorporated, though they likely influence resilience.

CONCLUSION

This study provides an empirical assessment of structural shifts in the EU banking system typologies during 2019–2021. Using K-Means clustering, we identify four stable archetypes whose characteristics persist, though country composition changed during the pandemic period.

The findings confirm that external shocks, can act as structural catalysts reshaping banking system configurations. The pandemic reinforced resilient systems, exposed latent vulnerabilities in high-intermediation structures, and facilitated convergence patterns transcending traditional geographic boundaries.

These results have implications for macroprudential policy. They support a cluster-based regulatory approach tailoring supervisory intensity to risk characteristics of each typology rather than applying uniform requirements across the EU. Specifically, the vulnerable HIS cluster requires stricter capital buffers and enhanced sovereign-bank risk monitoring, while the dynamic DGS cluster warrants close credit expansion oversight to prevent asset quality deterioration.

ACKNOWLEDGMENT

We would like to thank TheGlobalEconomy.com, and, in particular, the team leader Neven Valev, for providing easy access to harmonized socioeconomic data, which facilitated the collection and analysis of indicators for this study. We also acknowledge compliance with the publication ethics rules of the journal Statistika: Statistics and Economy Journal (available at: https://www.csu.gov.cz/statistika_journal).

References

ALESSI, L., BRUNO, B., CARLETTI, E., NEUGEBAUER, K., & WOLFSKEIL, I. (2021). Cover your assets: Non-performing loans and coverage ratios in Europe. *Economic Policy*, 36(108): 685–733. <https://doi.org/10.1093/epolic/eiab013>.

ARI, A., CHEN, S., & RATNOVSKI, L. (2021). The dynamics of non-performing loans during banking crises: A new database with post-COVID-19 implications. *Journal of Banking and Finance*, 133: 106–140. <https://doi.org/10.1016/j.jbankfin.2021.106140>.

CAO, YIFEI & CHOU, JEN-YU. (2022). Bank resilience over the COVID-19 crisis: The role of regulatory capital. *Finance Research Letters*, 48(C). <https://doi.org/1 10.1016/j.frl.2022.102891>.

CHODNICKA-JAWORSKA, P. (2022). Impact of COVID-19 on European banks' credit ratings. *Economic Research-Ekonomska Istraživanja*, 36(3): 1–17. <https://doi.org/10.1080/1331677X.2022.2153717>.

CLICHICI, D., & ZELDEA, C.-G. (2021). The resilience of central and eastern european banking systems during the covid-19 crisis. *Economy and Sociology*, 2: 20–29. <https://doi.org/10.36004/nier.es.2021.2–02>.

COLAK, G., & ÖZTEKIN, O. (2021). The impact of COVID-19 pandemic on bank lending around the world. *Journal of Banking and Finance*, 133: 106207. <https://doi.org/10.1016/j.jbankfin.2021.106207>.

DANKO, J., & SUCHÝ, E. (2021). The Financial Integration in the European Capital Market Using a Clustering Approach on Financial Data. *Economies*, *9*(2), 89. <https://doi.org/10.3390/economies9020089>.

DEMIR, E., & DANISMAN, G. O. (2021). Banking sector reactions to COVID-19: The role of bank-specific factors and government policy responses. *Research in International Business and Finance*, 58: 101508. <https://doi.org/10.1016/j.ribaf.2021.101508>.

DEMIRGÜÇ-KUNT, A., PEDRAZA, A., & RUIZ-ORTEGA, C. (2021). Banking sector performance during the COVID-19 crisis. *Journal of Banking and Finance*, 133: 106305. <https://doi.org/10.1016/j.jbankfin.2021.106305>.

DURSUN-DE NEEF, H. Ö., & SCHANDLBAUER, A. (2021). COVID-19 and lending responses of European banks. *Journal of Banking and Finance*, 133: 106236. <https://doi.org/10.1016/j.jbankfin.2021.106236>.

EUROPEAN BANKING AUTHORITY (EBA). (2020). *The EU banking sector: First insights into the COVID-19 impacts* (EBA/REP/2020/17) [online]. <https://service.betterregulation.com/document/442177>.

EUROPEAN BANKING AUTHORITY (EBA). (2021). *Risk dashboard: Q4 2020* [online]. <https://www.eba.europa.eu/risk-analysis-and-data/risk-dashboard>.

EUROPEAN CENTRAL BANK (ECB). (2021). *Financial stability review: May 2021* [online]. <https://www.ecb.europa.eu/pub/financial-stability/fsr/html/ecb.fsr202105~c9f6a47f42.en.html>.

FEYEN, E., ALONSO GISPERT, T., KLIATSKOVA, T., & MARE, D. S. (2021). Financial sector policy response to COVID-19 in emerging markets and developing economies. *Journal of Banking and Finance*, 133: 106184. https://doi.org/10.1016/j.jbankfin.2021.106184.
FOGLIA, M., ADDI, A., & ANGELINI, E. (2022). The Eurozone banking sector in the time of COVID-19: Measuring volatility connectedness. *Global Finance Journal*, 51(C). <https://doi.org/ 10.1016/j.gfj.2021.100677>.
HASTIE, T., TIBSHIRANI, R., & FRIEDMAN, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. 2nd ed. Springer. <https://doi.org/10.1007/978-0-387-84858-7>.
INTERNATIONAL MONETARY FUND (IMF). (2021). Departmental Papers 2021, 008: COVID-19: How will European banks fare? [online]. <https://doi.org/10.5089/9781513572772.087>.
JAIN, A. K. (2010). Data clustering: 50 years beyond K-means. *Pattern Recognition Letters*, 31(8): 651–666. <https://doi.org/10.1016/j.patrec.2009.09.011>.
KAMINSKYI, A. & NEHREY, M. (2023). Cluster Method Applying to Covid-19 Event Study for the Largest USA Banks. CEUR, 3641. <https://ceur-ws.org/Vol-3641/paper4.pdf>
KOZAK, S. (2021). The impact of COVID-19 on bank equity and performance: The case of Central Eastern South European countries. *Sustainability*, 13(19): 11036. <https://doi.org/10.3390/su131911036>.
MCLACHLAN, G. J., & PEEL, D. (2000). *Finite Mixture Models*. Wiley. <https://doi.org/10.1002/0471721182>.
MIHALECH, P., & KOŠÍKOVÁ, M. (2021). Cluster analysis of the EU banking sector based on EBA risk indicators. In: *EDAMBA 2021: 24th International Scientific Conference for Doctoral Students and Young Researchers*. pp. 306–316. <https://doi.org/10.53465/EDAMBA.2021.9788022549301.306–316>.
OSADCHA, N., ZHYTKEVYCH, O., & MATVIYCHUK, A. (2025). Comprehensive comparative analysis of clustering methods in the context of modeling nations' decarbonization potentials. SSRN. http://dx.doi.org/10.2139/ssrn.5598455.
PLIKAS, J. H., KENOURGIOS, D., & SAVVAKIS, G. A. (2024). COVID-19 and non-performing loans in Europe. *Journal of Risk and Financial Management*, 17(7): 271. <https://doi.org/10.3390/jrfm17070271>.
POZNYAK, S., & KOLYADA, Y. (2023). Comparative analysis of the effectiveness of dimensionality reduction algorithms and clustering methods on the problem of modelling economic growth. Neuro-Fuzzy Modelling Techniques in Economics, 12, 67–110. http://doi.org/10.33111/nfmte.2023.067
ROUSSEEUW, P. J., & KAUFMAN, L. (1990). *Finding Groups in Data: An Introduction to Cluster Analysis*. Wiley. <https://doi.org/10.1002/9780470316801>.
SALAZAR, Y., MERELLO, P., & ZORIO-GRIMA, A. (2023). IFRS 9, banking risk and COVID-19: Evidence from Europe. *Finance Research Letters*, 56: 104130. <https://doi.org/10.1016/j.frl.2023.104130>.
THEGLOBALECONOMY.COM. (2025). *Economic and financial indicators database* [online]. https://www.theglobaleconomy.com.
WORLD BANK. (2022). *World Development Report 2022: Finance for an equitable recovery* [online]. <https://doi.org/10.1596/978-1-4648-1730-4>.
WORLD BANK'S GLOBAL FINANCIAL DEVELOPMENT DATABASE. (2025). [online]. <https://www.worldbank.org/en/publication/gfdr/data/global-financial-development-database>

ANNEX/ APPENDIX

Table A1 The complete list financial and macro-financial indicators used in the study

	Abbr.	Full Indicator Name	Brief Description	Motivation for Inclusion	Preferred Range / Interpretation
1	GDPG	GDP Growth Rate	Annual percentage change of Gross Domestic Product	It reflects overall economic health and potential for loan growth/repayment	Higher positive values (e.g., >2–3%) indicate robust economic activity
2	UNEMP	Unemployment Rate	Percentage of the total labor force that is unemployed	It directly affects household income and ability to service debt, impacting NPLs	Lower values (<6–7%) indicate a healthy labor market
3	GDEBT	Government Debt to GDP	General government gross debt as a percentage of GDP	It indicates sovereign risk and potential fiscal constraints affecting banking sector support	Lower to moderate (<60–90%) suggests fiscal prudence and capacity

Table A1 (continuation)

	Abbr.	Full Indicator Name	Brief Description	Motivation for Inclusion	Preferred Range / Interpretation
4	BAGDP	Bank Assets to GDP	Total banking sector assets as a percentage of GDP	It reflects the size of the banking sector relative to the economy, indicating systemic importance.	Moderate to high (e.g., 70–200%). Extremely high values may signal systemic risk.
5	FDGDP	Financial System Deposits to GDP	Total deposits in the financial system as a percentage of GDP	It indicates the stability and size of the funding base for the financial sector	Moderate to high (e.g., 60–100%) suggests a stable funding source
6	BCONC	Bank Concentration	Assets of the three largest banks as a share of total banking sector assets.	It measures market power and potential "too-big-to-fail" implications; can influence competition	Moderate (e.g., 50–80%). Very high values (>90%) may signal systemic risk or lack of competition
7	NPL	Non-Performing Loans Ratio	Non-performing loans to total gross loans	It is a key indicator of asset quality and credit risk for banks	Lower is better (<3–5%), reflecting healthy loan portfolios
8	LADR	Bank Credit to Deposits Ratio	Total bank loans to total bank deposits	It measures liquidity risk; indicates reliance on non-deposit funding	Around 70–100% indicates efficient deposit utilization; >100% signals higher reliance on other funding
9	CAR	Regulatory Capital to RWA	Regulatory Tier 1 Capital to Risk-Weighted Assets.	It is primary measure of a bank's capital adequacy and ability to absorb losses	Higher is better (>15–20%) for strong capital buffers (Basel III minimum is 8%)
10	z-score	Banking System Z-score	Composite index: (ROA + capital/asset ratio) / std. dev of ROA. Indicates banking system stability	It indicates distance to insolvency (number of standard deviations); higher values imply greater stability and lower insolvency risk	Higher is better (>10–15) for lower insolvency risk

Source: Own construction. All data sourced from TheGlobalEconomy.com (2025), aggregating harmonized indicators from Eurostat, World Bank, World Bank Global Financial Development Database (GFDD), European Central Bank (ECB), and International Monetary Fund (IMF)
Note: Preferred ranges for indicators in Table A1 are defined based on regulatory standards (e.g., Basel III, Maastricht criteria) and empirical studies [Alessi et al., 2022; Cao & Chou, 2022; Chodnicka-Jaworska, 2022; Colak & Öztekin, 2021; Demirgüç-Kunt et al., 2021; Foglia et al., 2022; IMF, 2021; Salazar et al., 2023; World Bank, 2022].

Table A2 Consolidated Average Factor Values for EU Banking System Clusters (2019 Baseline)

Cluster Name (ID)	GDPG	UNEMP	GDEBT	BAGDP	FDGDP	BCONC	NPL	LADR	CAR	z-score
CSS (Cluster 1)	2.57	4.28	50.44	64.51	69.17	67.52	2.96	80.81	20.38	14.44
DGS (Cluster 2)	3.53	6.59	72.82	70.43	69.25	70.75	5.44	86.76	19.47	13.56
HIS (Cluster 3)	2.56	5.44	64.25	100.45	76.35	72.86	2.42	114.00	20.40	14.62
NRS (Cluster 4)	2.13	5.92	34.65	154.93	65.64	86.62	0.98	227.56	22.62	31.54
Desired/Optimal Range	>2–3%	<6–7%	<60–90%	70–200%	60–100%	50–80%	<3–5%	70–100%	>15–20%	>10–15

Source: Own construction.
Note: Cluster abbreviations: Core Stability Systems (CSS), Dynamic Growth Systems (DGS), High Intermediation Systems (HIS), Nordic Resilience Systems NRS.

Table A3 Consolidated Average Factor Values for EU Banking System Clusters (2021 Post-Pandemic)

Cluster Name (ID)	GDPG	UNEMP	GDEBT	BAGDP	FDGDP	BCONC	NPL	LADR	CAR	z-score
CSS (Cluster 1)	5.96	5.67	66.61	74.90	81.22	71.29	2.14	77.00	21.94	13.68
DGS (Cluster 2)	8.27	6.88	72.52	73.15	86.52	77.29	3.71	77.04	20.68	12.46
HIS (Cluster 3)	8.76	6.34	88.25	88.88	85.35	69.66	3.59	76.44	21.32	10.17
NRS (Cluster 4)	6.66	6.88	36.35	157.86	73.72	85.70	0.83	209.86	23.39	32.89
Desired/Optimal Range	>2–3%	<6–7%	<60–90%	70–200%	60–100%	50–80%	<3–5%	70–100%	>15–20%	>10–15

Source: Own construction