



European Economic Integration

Dmytro ZHERLITSYN,
Nataliia REKOVA

DIGITAL DIVIDE AND TYPOLOGY OF EUROPEAN UNION COUNTRIES ACROSS TWO MEASUREMENT REGIMES OF THE DIGITAL ECONOMY AND SOCIETY INDEX

Abstract

The article assesses the digital divide and develops a typology of EU countries by the level and dynamics of digital development across two measurement regimes: the composite Digital Economy and Society Index (DESI), 2016–2021, and the Digital Decade indicator system, 2021–2025. The discontinuation of the composite DESI after 2022 and the transition to the Digital Decade system created a structural incompatibility in the dynamics, preventing direct comparison across the reform. Indicators are placed on a common dimensionless scale, the divide is measured by the cross-country standard deviation of the domain score, and the typology is built by two clustering methods whose agreement is checked. The direction of change in the digital divide varies by domain and regime; the gap widened in 3 of 4 DESI domains, whereas in the Digital Decade regime Digital Public Services and Human Capital converge fastest and Integration diverges. A typology of 3 country groups is identified and reproduced in both regimes, all groups clearing the stability threshold under the main resampling scheme; the transition between regimes was accompanied by 8 of 27 countries changing group. The digital level is closely related to GDP per capita, whereas grouping by digital features does not coincide with the income partition. A country typology compared across the two regimes and a quantitative estimate of the structural shift between them are proposed for the first time. The results give an instrument for comparative analysis that stays valid when the measurement system changes, because it compares partitions rather than indicator values.

© Dmytro Zherlitsyn, Nataliia Reкова, 2026.

Zherlitsyn Dmytro, DSc (Economics), Professor, Researcher R4, University of National and World Economy, Sofia, Bulgaria. ORCID: 0000-0002-2331-8690 Email: d.zherlitsyn@unwe.bg
Rekova Nataliia, DSc (Economics), Professor, Researcher, Faculty of Mathematics and Informatics, Sofia University "St. Kliment Ohridski", Sofia, Bulgaria. ORCID: 0000-0003-0956-6564 Email: rekova@uni-sofia.bg

Key Words:

cluster analysis, convergence, Digital Decade, digital divide, Digital Economy and Society Index (DESI), European Union.

JEL: C38, O33, C43, L86.

4 tables, 5 figures, 9 formulas, 33 references.

Problem Statement

The digital transformation of the European Union economy remains one of the Union's policy priorities (Troitiño, 2022), and from 2014 to 2022 its quantitative measurement was provided by the composite Digital Economy and Society Index (DESI). In 2022, the European Commission discontinued the publication of DESI in its previous form and moved to the Digital Decade indicator system oriented toward the targets of the Digital Decade Policy Programme to 2030. This change creates a methodological problem, since the set of underlying indicators, their definitions, and units after 2021 differ substantially from the previous ones, and two Eurostat series underwent documented structural breaks, making a direct comparison of values before and after the reform invalid.

The relevance of the study follows from the fact that assessing the digital divide between member states and their typology requires a method that remains valid when the measurement system changes. Existing studies mostly rely on the continuous DESI series up to 2022 and do not account for the structural incompatibility caused by the transition to the Digital Decade system. It remains an open question whether the typology of countries by the level and dynamics of digital development is preserved at the transition between the two measurement regimes, and whether the speed of change of the digital divide differs across domains. Without answers to these questions, the comparative analysis of EU countries' digital development risks resting on incompatible data.

The aim of the article is to assess the digital divide and to construct a typology of the European Union countries by the level and dynamics of digital development across two measurement regimes, the regime of the composite DESI for 2016–2021 and the regime of the Digital Decade indicator system for 2021–2025, applying a method that remains valid under the structural incompatibility between these regimes.

Literature Review

The first research direction concerns DESI as a composite index, its construction methodology, and its critique. The methodological foundations of composite indices, particularly the issues of weighting, aggregation, and robustness,

are systematised by Greco et al. (2019), and a general critique of the subjectivity of indicator selection and weight choice is developed by Kuc-Czarnecka et al., 2020. With respect to DESI directly, Olczyk & Kuc-Czarnecka (2022) use sensitivity and uncertainty analysis to show that the index is excessively determined by the connectivity domain, whereas the use of internet services and digital public services add limited informational value. The structural redundancy of DESI indicators is confirmed by Bruno et al. (2023), where principal component analysis is used to build a reduced version of the index that reproduces the original rankings with fewer indicators. The statistical analysis of the DESI dimensions by Bánhidi et al. (2020) revealed a substantial correlation between domains and separately highlighted the problem of the index's frequently changing composition, which hampers comparison over time. Stavitsky et al. (2019) present an econometric analysis of the drivers of DESI, which shows the index's dependence on its own prior dynamics and on macroeconomic variables.

Despite the considerable attention to the DESI methodology, existing studies analyse the index within a single continuous series up to 2022 and do not consider the methodological consequences of its discontinuation and the transition to the Digital Decade indicator system. The compatibility of the two measurement regimes and the validity of their comparison remain unresolved, which defines the first research task of this work.

The second research direction covers the digital divide, digital convergence, and the typology of EU countries. The convergence of digital development based on DESI is examined by Kovács et al. (2022), where signs of a narrowing gap together with a Matthew effect are found, and by Andrei et al. (2023), where sigma- and beta-convergence methods confirm the narrowing of the digital divide with the leading role of economic development. In Borowiecki et al. (2021) work, cluster analysis of DESI data records overall convergence across most components of the index, along with strengthening polarisation in the Integration of Digital Technology domain. Kolupaieva & Tiesheva (2023) analyse the asymmetry and unevenness of EU countries' digital development, showing that the digital divide can persist even amid the rapid diffusion of technologies.

Several works develop a typology of countries and citizens by level of digital development. In the seminal study, Cruz-Jesus et al. (2012) use factor and cluster analysis to single out several groups of EU countries along two latent dimensions of digital development. Gomes & Dias (2025) build a typology of EU citizens based on their use of digital technologies, identifying a concentration of the digital divide in particular countries. Dobos & Bánhidi (2025) present a grouping of Central and Eastern European countries by digital readiness, and Zeinali & Mussari (2025) conduct a cluster analysis of digital development using DESI data from the Digital Decade period. Szeles & Simionescu (2020) examine regional differences in the EU digital economy and their drivers. Closest in subject is the study of de Maya Matallana et al. (2026), which builds a distance-based composite for EU countries and analyses partial convergence under persistent structural differences.

Masoura & Malefaki (2023) trace the dynamics of the DESI index itself in a socioeconomic dimension, while Zherlitsyn & Rekova (2026) examine the link between digital development and labour markets in selected European economies. Bibliometric mapping of digital-divide research in the EU and the Western Balkans over 2018-2023 is performed by Kovač et al. (2024), with digital-transformation trends across the EU are mapped bibliometrically in Zherlitsyn et al. (2025), which together

allow the state of the field to be outlined and the research gap to be justified. Szajt et al. (2024) assess the convergence of digital development on DESI data, where the catch-up effect proved weaker than expected.

A common limitation of the cited studies is that they assess typology and convergence within a single measurement regime, without testing the stability of country groupings when the indicator set changes. It remains unresolved to what extent the typology of countries by the level and dynamics of digital development is preserved in the transition from the composite DESI to the Digital Decade indicator system, and whether this typology is determined by countries' income level. Here and below, two measurement regimes of digital development are considered: the regime of the composite DESI and the regime of the Digital Decade indicators, which correspond to two different methodological configurations for monitoring the EU's digital transformation.

A synthesis of the cited studies makes it possible to define the unresolved problems and formulate the hypotheses of this work.

H1. The direction of change of the digital divide differs across the domains of digital development.

H2. EU countries form stable groups by the level and dynamics of digital development.

H3. The grouping of countries by the features of digital development does not coincide with their partition by the level of economic development.

This study tests these hypotheses in the two measurement regimes, using a method robust to their structural incompatibility. A deeper analysis of the methodological properties of the DESI index and the convergence of digital development outlines a broader set of problems that define directions for further research.

Methodology

The analysis rests on two measurement regimes, the period of the discontinued DESI composite (2016 to 2021, with the observation year equal to the edition year minus 1, because the index carries a one-year publication lag), and the current period of the Digital Decade dashboard (2021 to 2025). The reason for this separation is twofold: the publication lag and the documented structural change in two Eurostat series, digital skills under the revised Digital Competence Framework from 2021 and ICT specialists after the revision of the Labour Force Survey, which may render the values incomparable across the reform.

Data sources

The digital indicators are based on the European Commission Digital Decade DESI dashboard, which publishes the DESI editions and the underlying indicator series for EU member states (European Commission, 2026a). From this source, the analysis takes 10 DESI sub-dimensions for the composite period (up to 2021) and 17 Digital Decade indicators available for at least 25 of the 27 member states in each year. For the Digital Decade regime, the dashboard carries no dimension attribute, so the assignment of the indicators to 4 domains is made by the authors by the unit of observation: 3 indicators for Human Capital (ICT specialists, internet use, overall internet take-up), 6 for Connectivity, 2 for Integration of Digital Technology and 6 for Digital Public Services.

In both regimes, the domain map only orders the indicators for the gap analysis and the clustering features. In the discontinued composite, household connectivity belonged to Connectivity, so the domain scores of the two regimes are not interchangeable. The coverage requirement excludes indicators that the Commission publishes once every two years, in particular at least basic digital skills and the digital intensity of small and medium-sized enterprises, so the only qualification indicator retained in the panel is the share of ICT specialists, and the other two domain indicators are contextual rather than target indicators. The Digital Decade panel has one gap out of 2,295 country–year pairs: household internet access for Ireland in 2022, so the Human Capital domain score for this country in that year is computed from 2 of the 3 indicators. The DESI sub-dimension panel is fully populated, with no imputation. The complete list of indicators is given in Appendix A. All indicators are increasing in desirability, and the published value is not transformed before standardisation. Standardisation is a linear transformation and therefore does not change the shape of an indicator's distribution. The European Commission's methodological note defines the number of unicorn companies as the sum of companies under points 11(a) and 11(b) of Article 2 of Decision (EU) 2022/2481, attributed to a member state by the location of the headquarters at the time the data were obtained from the commercial platform Dealroom (European Commission, 2026b). This is a single cross-section of the commercial base in absolute units, whereas the other indicator of the domain is expressed as a percentage.

The macroeconomic indicators are based on Eurostat data and matched to the corresponding observation year of each regime. Four dataset series are used: GDP per capita in purchasing power standards from tec00114, real labour productivity per person from nama_10_lp_ulc under the position RLPR_PER as an index with 2015 equal to 100, the unemployment rate of the population aged 15 to 74 from une_rt_a, and gross domestic expenditure on research and development as a share of GDP from rd_e_gerdot (Eurostat, n.d). The macroeconomic indicators are not included in the clustering and are used only to describe the link between digital and economic development.

For the data panels under the European Commission's official methodology, the edition year, observation year, and reference year are distinguished. The observation year for the digital development features is taken from the Commission data, normally with a 1-year lag. The reference year for the macroeconomic indicators is the common year available for all 4 indicators across 27 countries. In the Digital Decade regime, this is 2024; in the DESI regime, 2021.

Standardisation and the features of level and dynamics

Indicators are reported in different units, shares, scores and counts, so they are placed on one dimensionless scale before any distance or correlation is computed. For indicator i and country c in observation year t with the raw value x , the standardised score pooled across years is the z-score of equation (1) (Milligan & Cooper, 1988):

$$z_{ict} = \frac{x_{ict} - \bar{x}_i}{\sigma_i} \quad (1)$$

where x_{ict} is the raw value of indicator i for country c in year t ; $\bar{x}_i$ is the mean of indicator i across all countries and years; σ_i is the population standard deviation of indicator i across all countries and years.

Pooling observations across years rather than within each year removes the unit of measurement, preserves the time signal needed for subsequent trend estimation, and provides a common scale with zero mean and unit variance.

The study is retrospective, so standardisation over the entire time window ensures inter-temporal comparability without out-of-sample processing. The choice of standardisation rather than mapping onto a fixed interval follows from the state of the data source. For the discontinued composite, the European Commission published the bounds of such a mapping, whereas for the Digital Decade indicator system they are not made public, and the methodological note explicitly states that no normalisation is applied (European Commission, 2026b). A mapping onto an interval would require assigning bounds for 6 of the 17 indicators, including the number of unicorn companies. Standardisation requires no such bounds.

The domain score of a country each year is the mean of the standardised scores of the indicators assigned to that domain, equation (2), where the set I_D collects the indicators of domain D . The overall score aggregates the 4 domains under two weighting schemes reported in parallel, the equal-weight scheme and the scheme proportional to the number of indicators per domain, equation (3), so that the result does not depend on a single convention (Greco et al., 2019):

$$Z_{Dct} = \frac{1}{|I_D|} \sum_{i \in I_D} z_{ict} \quad (2)$$

$$Z_{ct}^{\text{eq}} = \frac{1}{4} \sum_D Z_{Dct}$$

$$Z_{ct}^{\text{prop}} = \sum_D \frac{|I_D|}{N} Z_{Dct}, \quad N = \sum_D |I_D| \quad (3)$$

where z_{ict} is the standardised score of the indicator from equation (1); Z_{ct}^{eq} is the country's overall score under the equal-weight scheme; Z_{ct}^{prop} is the country's overall score under the proportional scheme; I_D is the set of indicators of domain D ; $|I_D|$ is the number of indicators in it; Z_{Dct} is the score of domain D for country c in year t ; N is the total number of indicators across all domains.

Below, the main results are presented under the equal-weight scheme.

Each country is then summarised by two scalar features. The level is the standardised score in the final year of the regime. The dynamics is the slope of that score over the whole period of the regime, estimated by ordinary least squares as the coefficient β_1 in the linear trend of equation (4), so that the whole interval affects the trend rather than only the two end points, and the estimate is robust to the noise of a single year (Zherlitsyn, 2024):

$$y_t = \beta_0 + \beta_1 t + \varepsilon \quad (4)$$

where y_t is the series to which the trend is applied; t is the observation year; β_0 is the intercept of the trend; β_1 is the slope of the trend, which serves as the dynamics feature; ε is the residual.

A positive slope means growth of the country score in units of the pooled standard deviation.

Measuring the gap

The digital divide in a domain and year is the cross-country spread of the domain score, measured by the sample standard deviation in equation (5). The change in this spread over time is interpreted as convergence or divergence in the sigma sense, that is, as a narrowing or widening of the gap between member states. Ordering the 4 domains by the slope of the spread, estimated by equation (4), answers the question of which parts of the digital economy converge and which diverge, separately in each regime, so that the effect of the reform is visible (Kovács et al., 2022):

$$SD_{Dt} = \sqrt{\frac{1}{n-1} \sum_c (Z_{Dct} - \bar{Z}_{Dt})^2} \quad (5)$$

where n is the number of countries; Z_{Dct} is the score of domain D for country c in year t ; $\bar{Z}_{Dt}$ is the cross-country mean of the score of domains D in year t ; SD_{Dt} is the standard deviation of the domain score across countries, that is, the measure of the gap.

Typology by clustering and its validation

The typology groups countries by their level and dynamics. Two clustering methods with different logics are run in parallel and compared: k-means, which assigns each country to the nearest of k centroids by minimising the within-cluster sum of squares, and Ward agglomerative hierarchical clustering, which at each step merges the pair of groups that yields the smallest increase of the within-cluster sum of squares. Both methods are applied to the standardised feature matrix (Arbelaitz et al., 2013; Murtagh & Legendre, 2014; Zherlitsyn, 2024). The agreement between these two methods is a reference characteristic, since both minimise the within-group sum of squares in Euclidean space and, on a small sample with one dominant axis, may coincide tautologically.

The number of groups is chosen by statistical criteria complemented by the economic content of the partition. Since a partition into two groups opposes the leaders to the remaining countries and carries no typological information, the primary number of groups k is chosen among the solutions with three or more groups as the value that maximises the average silhouette coefficient of the Ward partition and at which the two algorithms agree, equation (6). The silhouette of a country is computed by equation (7), and the score of the partition is the mean across all countries (Arbelaitz et al., 2013; Hennig, 2007):

$$k^* = \arg \max_{k \in K} S(k) \quad (6)$$

where k is a candidate number of groups; K is the set of candidates with three or more groups for which both algorithms give an identical partition; $S(k)$ is the average silhouette coefficient of the partition into k groups; k^* is the number of groups chosen.

$$s_i = \frac{b_i - a_i}{\max(a_i, b_i)} \quad (7)$$

where a_i is the mean distance of country i to the other countries of its cluster; b_i is the smallest mean distance of country i to the countries of another cluster; s_i is the silhouette of country i .

The choice is additionally examined through three indices. The Calinski-Harabasz index reports the ratio of between- to within-cluster dispersion, where a higher value is better. The Davies-Bouldin index averages, across clusters, the ratio of dispersion to the distance to the worst-neighbour centroid; a lower value is better. For the hierarchy, the cophenetic correlation between the dendrogram distances and the original distances is computed, which measures how accurately the tree preserves the geometry of the data (Arbelaitz et al., 2013; Saraçlı et al., 2013).

Additional estimates are also reported for k^*+1 and k^*-1 groups, since this may correspond to an additional ranking level among the differentiated indicators of EU countries; this partition is economically motivated and not presented as a statistical optimum. In each partition, the groups are relabelled in descending order of mean digital level, so group 1 is the leading one regardless of the arbitrary internal numbering.

Sensitivity checks close the typology. The agreement between two partitions of the same countries is measured by the adjusted Rand index, the share of country pairs that the two partitions classify identically, corrected for chance by the marginal pair counts, equation (8) (Warrens & van der Hoef, 2022):

$$ARI = \frac{\sum_{ij} \binom{n_{ij}}{2} - \frac{\sum_i \binom{a_i}{2} \sum_j \binom{b_j}{2}}{\binom{n}{2}}}{\frac{1}{2} \left[\sum_i \binom{a_i}{2} + \sum_j \binom{b_j}{2} \right] - \frac{\sum_i \binom{a_i}{2} \sum_j \binom{b_j}{2}}{\binom{n}{2}}} \quad (8)$$

where n_{ij} is the number of countries shared by group i of the first partition and group j of the second; a_i is the number of countries in group i of the first partition; b_j is the number of countries in group j of the second partition; n is the total number of countries; $\binom{n}{2}$ is the binomial coefficient, that is, the number of pairs.

The index equals one for identical partitions and approaches zero for chance agreement. Within each regime, it assesses the sensitivity to the weighting convention by comparing the groupings under the two aggregation schemes. The sensitivity of the Integration domain, the narrowest in composition, is examined separately. The slope of the gap is computed on the published values, after the logarithm, after winsorising at the 90th percentile and with the number of unicorn companies excluded, and for the extended composition of the domain with 6 dashboard indicators available on the biennial grid. Across the regimes, the same index computed between the typologies of the two regimes on the shared countries measures the extent to which the change of the indicator set rebuilt the grouping.

Stability is measured at three levels. The stability of the partition as a whole is measured by the average adjusted Rand index between the reference partition and the partitions obtained on resampled sets of countries. The stability of an individual group is measured by the Jaccard index, and a group is considered stable if its average index is no lower than 0.75 (Hennig, 2007). The stability of an individual observation is measured by the share of replications in which a country retains the same group as in the baseline partition.

Because standardisation is performed over the entire set of observations, a resample inherits the mean and spread computed with the excluded countries

included, which inflates the stability estimates. The estimate is therefore duplicated under a strict scheme, in which each replication is recomputed from scratch, from the published values through the level and the slope, relying only on the retained countries. The partition is additionally perturbed with random noise at the level of 10% of the standard deviation of the features, and each configuration is checked on 3 independent sampling sequences.

Link to the economy and reproducibility

The link between digital development and the economy is described by the Spearman rank correlation between the digital level and each macroeconomic indicator:

$$\rho = 1 - \frac{6 \sum_i d_i^2}{n(n^2 - 1)} \quad (9)$$

where d_i is the rank difference of country i between the two indicators compared; n is the number of countries; ρ is the Spearman rank correlation coefficient.

The typology is compared with the partition of countries into GDP per capita terciles using the adjusted Rand index in equation (8) (Kovács et al., 2022). Spearman's rank correlation measures a monotonic association, whereas the adjusted Rand index measures agreement between two partitions by country pairs; neither measure depends on the magnitude of the known upper outliers in GDP per capita for Luxembourg and Ireland.

The significance level is set at 5 %. Since 8 correlations with the digital level of the corresponding regime are tested simultaneously across the two regimes, p-values corrected by the Holm step-down procedure are reported alongside the uncorrected ones (Holm, 1979).

The analysis was carried out with the analytical tools and libraries of the Python 3.12 programming language (libraries NumPy 1.26, pandas 2.3, scikit-learn 1.7, SciPy 1.16 and Matplotlib 3.10).

Research Results

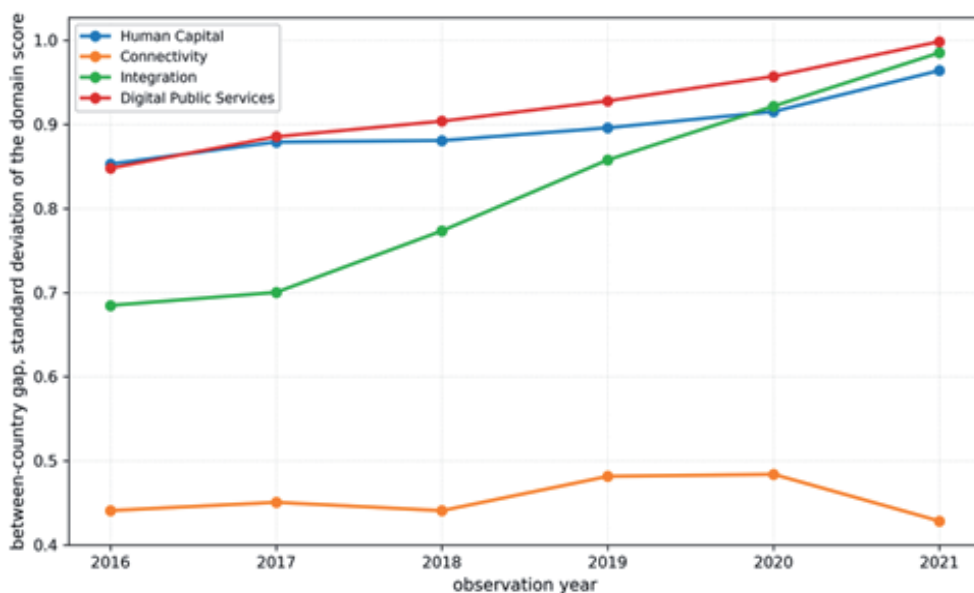
Change of the gap across the four domains

The first direction of the analysis concerns whether the digital divide changes at the same speed across the domains, and whether this dynamic changed after the European Commission revised its measurement approach. Figure 1 shows the annual dynamics of the between-country gap in each domain, and Table 1 reports the trend slope of that change.

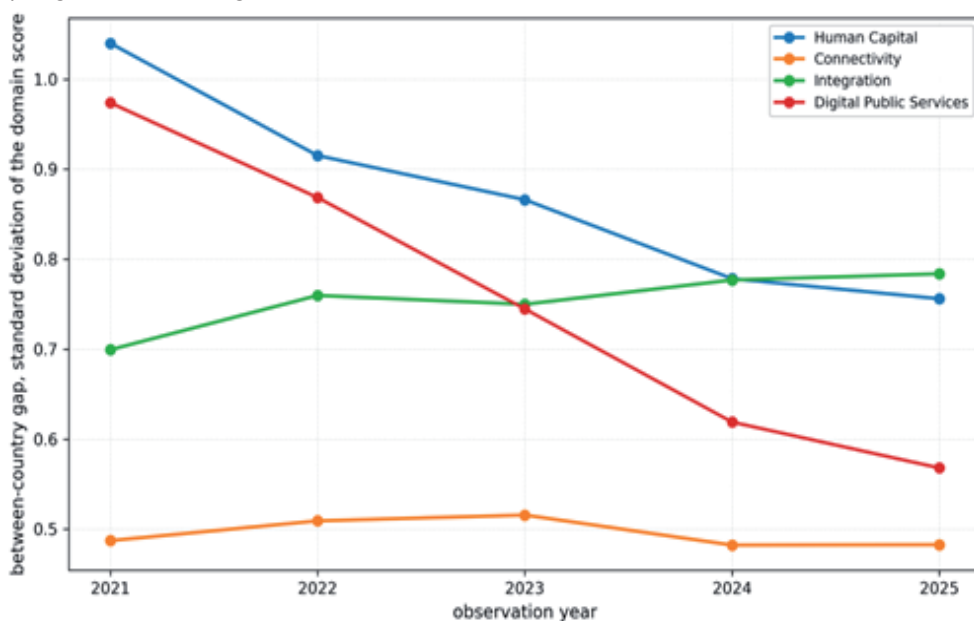
Figure 1

The between-country gap across the four domains and observation years

a) DESI regime



b) Digital Decade regime



Source: computed by the authors on DESI data (European Commission, 2026a).

The two regimes give different pictures of the change in the gap. In the DESI regime, the gap between countries widened in 3 of 4 domains, from Human Capital (+0.0194 per year) to Integration, which diverged fastest (+0.0643 per year). The Connectivity domain remained practically unchanged, with a trend slope of +0.0022 per year, and the gap in the last year is lower than in the first. The signs differ because the series is non-monotone, namely the gap reaches its maximum of 0.484 in 2020 and falls to 0.428 in 2021. In the Digital Decade regime, the sign of the slope reverses in Digital Public Services and Human Capital, Connectivity stays close to zero in both regimes, and Integration keeps a positive sign; the domains also differ in composition between the two regimes. The gap narrows fastest in the Digital Public Services domain, -0.1061 per year, followed by Human Capital, -0.0705 , Connectivity remained almost unchanged, -0.0036 , and only Integration continued to diverge, $+0.0186$. The Integration domain in the Digital Decade regime rests on only 2 indicators, and one of them, the number of unicorn companies, is an absolute count concentrated in a few large economies and dependent on the size of the economy rather than on the intensity of the use of digital technologies by enterprises. The skewness of this indicator in the computation window equals 2.148, and its maximum reaches 4.07 standard deviations; in a domain of 2 indicators, it carries half of the weight, so the properties of its distribution pass into the domain score. The slope of the gap changes sign between treatments, from $+0.0186$ on the published values to -0.0142 with the indicator excluded, -0.0027 under the logarithm, and -0.0001 under winsorising at the 90th percentile; under such a composition, the direction is not stable. Under the extended composition with 6 dashboard indicators on the grid of 2021, 2023 and 2025, determined by the biennial periodicity of the enterprise surveys, the divergence is preserved under any treatment, from $+0.0173$ to $+0.0220$, so the composition of the domain determines whether the direction withstands the alternative treatments. The incompatibility of the units is acknowledged by the source itself (European Commission, 2026b) and is systematic, since 3 domains of the Digital Decade are composed of scales bounded from above and it is precisely these that converge, whereas the only domain that diverges contains an unbounded count. The sign reversal in the Human Capital domain follows from its composition. In the Digital Decade regime, 2 of 3 indicators measure internet access, and the gap narrows by 0.0973 and 0.0827 per year at mean levels of 93.0% and 94.6% and maxima of 99.7% and 99.3%, that is, at the limit of the scale. By contrast, the share of ICT specialists widens the gap by 0.0135 per year, and the indicator of at least basic digital skills, which does not pass the coverage requirement of the annual grid, by 0.0150. Skills diverge in both regimes, and in the DESI regime the domain rests on them alone, whereas in the Digital Decade regime the sign of the domain is set by the access indicators. The direction of change of the gap differs across domains, so H1 is confirmed. Since the direction of the conclusion depends on the indicator system, the two regimes cannot be treated as one continuous series. This raises the task of forming a comparable list of core indicators for the whole period, which lies beyond the scope of this study.

Table 1

Estimated annual trend slope of the between-country gap by domain and regime, with the gap in the first and last years

Regime	Domain*	Estimated trend slope per year	Gap, start	Gap, end**
DESI	Connectivity	+0.0022	0.441	0.428
DESI	Human Capital	+0.0194	0.853	0.964
DESI	Digital Public Services	+0.0283	0.848	0.998
DESI	Integration	+0.0643	0.685	0.985
Digital Decade	Digital Public Services	-0.1061	0.974	0.568
Digital Decade	Human Capital	-0.0705	1.040	0.756
Digital Decade	Connectivity	-0.0036	0.487	0.483
Digital Decade	Integration	+0.0186	0.699	0.784

Note: * The domains differ in composition between the two regimes, so they are not directly comparable. ** For a non-monotone series, the trend slope and the difference between the last and the first year may have different signs, as in the Connectivity domain of the DESI regime, so both quantities are reported.

Source: computed by the authors on DESI data (European Commission, 2026a).

Choice of the number of groups and partition quality

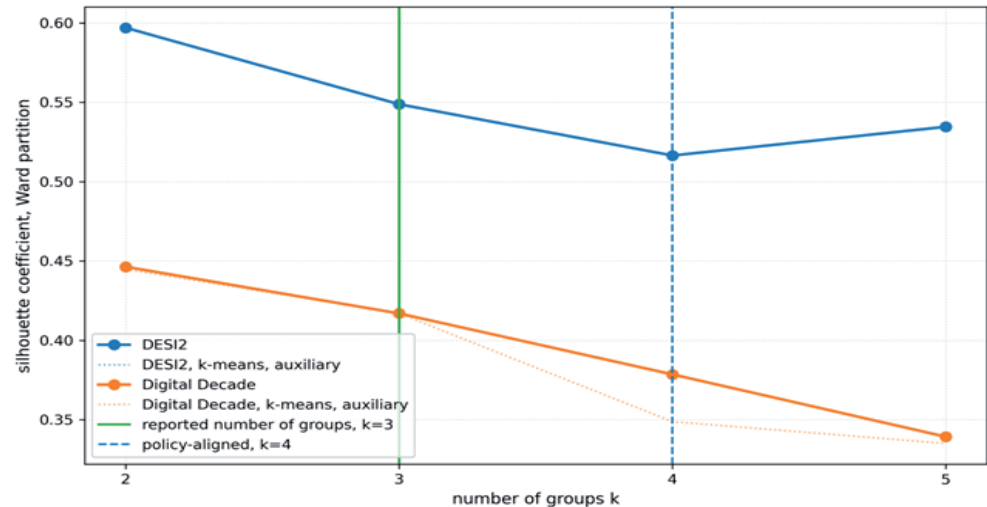
The next direction of the analysis concerns partitioning EU countries into stable groups by the level and dynamics of digital development in each measurement regime. The number of groups is assessed against three criteria: the quality of the partition, the agreement between the two algorithms, and the stability of each group.

Figure 2 shows the silhouette of the Ward partition as a solid line and the auxiliary silhouette of k-means as a dashed line, both under the equal-weight aggregation scheme.

In the DESI regime, the two algorithms produce identical partitions for any number of groups in the range examined, from 2 to 5, whereas in the Digital Decade regime they coincide only at 3 groups. The silhouette at 3 groups equals 0.417 in the Digital Decade regime, corresponding to moderate separation, and 0.549 in the DESI regime; the two methods agree completely, and the adjusted Rand index equals 1.000 in both regimes. Table 2 reports the partition quality indices.

Figure 2

Partition quality as a function of the number of groups, both regimes



Source: developed by the authors on DESI data (European Commission, 2026a).

The data in Table 2 show that the rank correlation of the level between the equal-weight and the proportional schemes equals 0.976 in the DESI regime and 0.949 in the Digital Decade regime, that is, the ordering of the countries by level practically does not depend on the weighting scheme. The composition of the groups, by contrast, does depend on it, since the adjusted Rand index between the same partitions equals 0.625 and 0.367, respectively, and 4 and 7 countries change membership. The proportional scheme gives each indicator the same weight and is therefore identical to aggregation without the domain layer, so the reported quantities measure the extent to which the typology depends on the partition of indicators into domains.

Table 2

Partition quality into 3 groups by regime and aggregation scheme

Regime	Silhouette (Ward)	Method agreement (ARI)	Calinski-Harabasz	Davies-Bouldin	Cophenetic
DESI	0.549	1.000	60.87	0.640	0.777
Digital Decade	0.417	1.000	37.27	0.741	0.642
DESI (proportional weighting)	0.537	0.702	48.47	0.764	0.791
Digital Decade (proportional weighting)	0.467	0.883	40.41	0.737	0.737

Source: developed by the authors on DESI data (European Commission, 2026a).

Stability of the partition under resampling

The next direction of the analysis concerns whether the three-group partition is preserved when the composition of the country sample changes – see Table 3. As Table 3 shows, under the strict scheme, the average agreement of the partition in the Digital Decade regime equals 0.697, and the Jaccard indices of the 3 groups equal 0.809, 0.774 and 0.867, so all three clear the threshold of 0.75. Feature resampling gives more optimistic values of 0.761 and 0.816, and the difference between the two schemes averages 0.034. In the DESI regime, the partition is more stable; the agreement equals 0.974 at a minimum Jaccard index of 0.958. The stability of the partition into 4 groups depends on the regime. In the DESI regime, the minimum Jaccard index equals 0.896, and the threshold is cleared; in the Digital Decade regime, it equals 0.592, and the threshold is not cleared. The partition into 5 groups clears the threshold in neither regime. A separate resampling run, in which the features are additionally perturbed by random noise at the level of 10 % of their standard deviation, gives an agreement of 0.918 in the DESI regime, where all 3 groups clear the threshold, and 0.626 in the Digital Decade regime, where 2 groups out of 3 clear it. Hence, among the solutions with three or more groups, only the three-group solution withstands the check in both regimes under the main scheme, whereas under the perturbation of the features its stability in the Digital Decade regime declines. The conclusion does not change when the sampling sequence is changed.

Table 3

Stability of the partition under country resampling, by regime and number of groups*

Regime	Groups	Strict scheme, mean agreement (ARI)	Strict scheme, lowest Jaccard index	Groups at or above 0.75, strict scheme	Feature resampling, ARI	Feature resampling, Jaccard
DESI	2	0.997	0.998	2 of 2	0.997	0.998
DESI	3	0.974	0.958	3 of 3	0.975	0.958
DESI	4	0.959	0.896	4 of 4	0.957	0.900
DESI	5	0.894	0.721	4 of 5	0.883	0.732
Digital Decade	2	0.663	0.807	2 of 2	0.729	0.846
Digital Decade	3	0.697	0.774	3 of 3	0.761	0.816
Digital Decade	4	0.639	0.592	2 of 4	0.704	0.641
Digital Decade	5	0.613	0.504	2 of 5	0.684	0.559

Note: * Resampling without replacement at the level of 90 % of the countries, 400 replications. A group is considered stable if it has an average Jaccard index of at least 0.75 (Hennig, 2007).

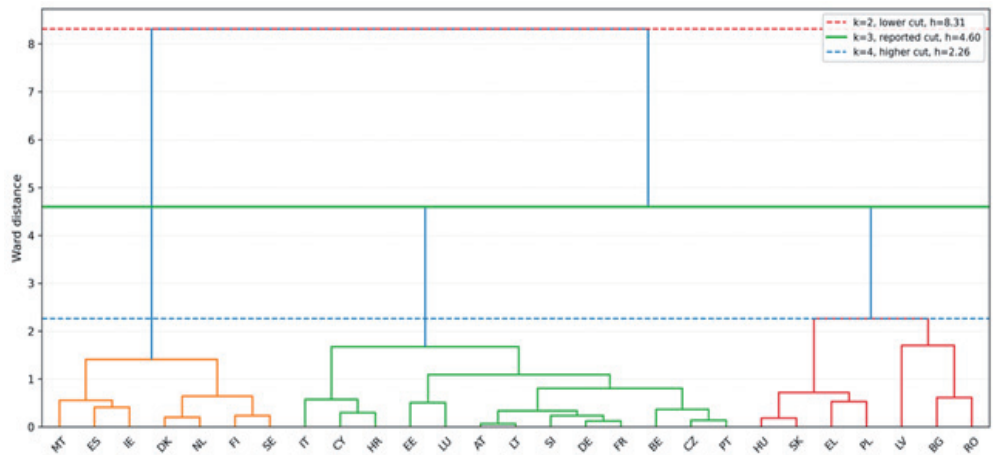
Source: computed by the authors.

The clustering trees in Figure 3 show the partition structure and support the identification of 3 groups. At the same time, the branch composition and merge heights of the two trees differ, reflecting the structural shift analysed in detail below.

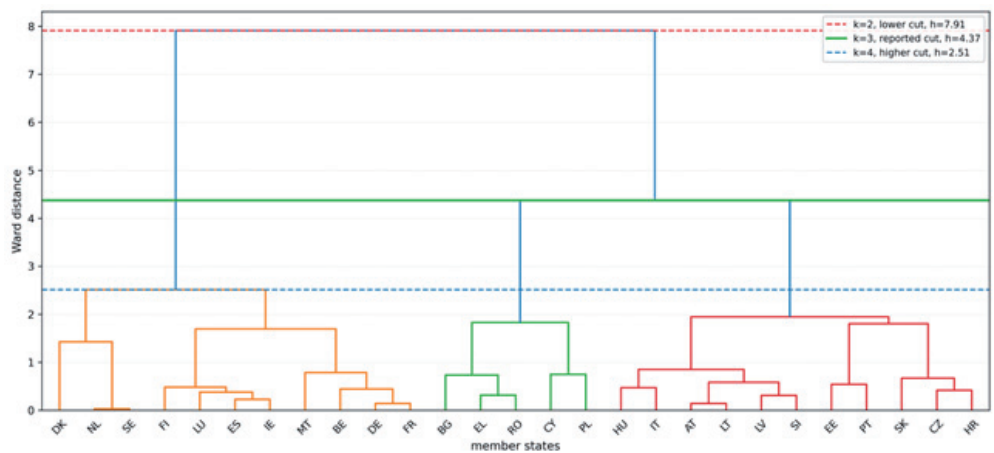
Figure 4 additionally illustrates the 3 clusters of the Digital Decade regime on a plane where the level attained is placed on the Ox axis and the dynamics on the Oy axis. The groups are renumbered so that the first is always the leading one. The marker size corresponds to the share of resamples in which a country keeps the membership shown, so a small marker indicates a country on the boundary between groups.

Figure 3
Clustering trees of the member states by Ward's method

a) DESI regime



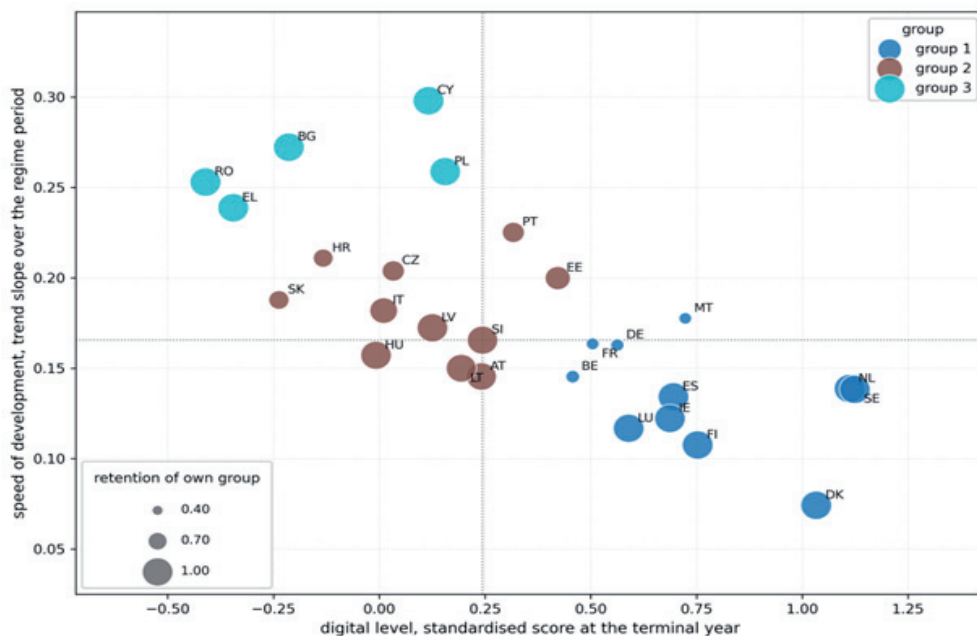
b) Digital Decade regime



Source: developed by the authors on DESI data (European Commission, 2026a).

Figure 4

Three groups of countries in the plane of level and dynamics, Digital Decade regime, groups ordered by descending mean level



Source: developed by the authors on DESI data (European Commission, 2026a).

The lowest retention share is held by Malta at 0.475, France at 0.482, Germany at 0.503 and Belgium at 0.518, whereas 8 countries, namely Bulgaria, Cyprus, Denmark, Greece, Finland, the Netherlands, Romania and Sweden, hold the highest share, equal to one. The blurred boundary runs within the upper tier, not between the leading and trailing groups. The leading group comprises Denmark, Sweden, the Netherlands, Finland, Ireland, Malta, Luxembourg, Germany and Belgium, as well as Spain and France, whose GDP per capita in PPS is below the EU average. The middle group comprises Estonia, Portugal, Slovenia, Austria, Lithuania, Latvia, Czechia, Italy, Hungary, Croatia, and Slovakia, with Austria's GDP per capita exceeding the EU average. Bulgaria, Romania, Greece, Cyprus and Poland close the ranking. The groups are separated primarily by dynamics, since the slope ranges of the middle and trailing groups do not overlap, whereas by level they partly overlap, and Cyprus and Poland stand above several countries in the middle group. The leading group has the highest level and the lowest dynamics, the trailing group the opposite, so by the overall score the gap between them narrows over time.

A partition into 4 groups is reported separately. This reading is not a statistical optimum, since at 4 groups the partition quality is lower than at 3 in both regimes. Resampling refines this conclusion: at 4 groups, the stability threshold is cleared in the DESI regime but not in the Digital Decade regime. The fourth partition splits one of the existing groups in two, and in the two regimes these are different groups.

In the Digital Decade regime, it separates Denmark, the Netherlands, and Sweden from the leading group, whereas in the DESI regime it splits the trailing group into Greece, Hungary, Poland, and Slovakia, and Bulgaria, Latvia, and Romania. The middle group remains undivided in both regimes, so the additional group does not correspond to a separate level of development.

The structural shift between the two regimes

A separate subject of the analysis is the extent to which the change of the indicator set rebuilt the distribution of countries across groups. The comparison is valid because the same method, the same number of groups and the same 27 countries are used in both regimes, and only the indicator set differs, that is, precisely the factor whose effect is to be measured. The level of each country is determined by the final year of the corresponding regime, so the two maps are anchored on different reference years, 2021 for DESI and 2025 for the Digital Decade, together with the corresponding windows for the slope estimation, so the agreement obtained reflects the joint effect of the change of the indicator set and of the four-year interval between the reference years, and not only the replacement of the indicators.

The agreement between the three-group map of DESI and the three-group map of the Digital Decade in the shared countries equals 0.291, and 8 of 27 countries changed groups. The sizes of the 3 groups changed from 7, 13 and 7 in the DESI regime to 11, 11 and 5 in the Digital Decade regime. The leading group expanded because Belgium, Germany, France and Luxembourg moved up into it, the middle group was joined by Hungary, Latvia and Slovakia, which moved up from the trailing group, and Cyprus moved down to the trailing group.

To separate the contribution of the change in the indicator system from that of the temporal shift, two benchmarks are computed, both based on the level, since for 2021 the Digital Decade regime has no retrospective series sufficient to estimate the slope. The temporal benchmark holds the indicator set unchanged and changes only the observation year; it reassigns 5 countries at an agreement of 0.608. The indicator benchmark holds the common year 2021 and changes only the indicator set; it reassigns 10 countries at an agreement of 0.220. For comparability, the joint change is computed in the same construction; it reassigns 5 countries at an agreement of 0.463. The joint quantity is not the sum of the two separate ones, since the adjusted Rand index is not additive and the two factors are not orthogonal, so the decomposition serves as a diagnostic rather than a causal separation. The indicator benchmark yields more reassignments and a lower agreement than the temporal one, indicating a larger structural shift associated with the change of the indicator system. A separate attribution of each country's transition is not possible.

Among the 8 reassigned countries, the average share retaining membership in their own group is 0.778, compared with 0.922 for the rest; however, the difference does not reach the conventional significance level, as the Mann-Whitney test yields a p-value of 0.103. The picture is heterogeneous within the group of transitions itself: France, Germany and Belgium are among the 4 least stable countries of the regime, whereas Luxembourg, Latvia and Cyprus keep their own group practically unchanged. Because the test has low power at these group sizes, it does not prove the absence of a link, but the available evidence does not show that boundary instability explains the cross-regime shift. The two regimes can be compared only at the level of method and of partition agreement, not at the level of the underlying indicator values.

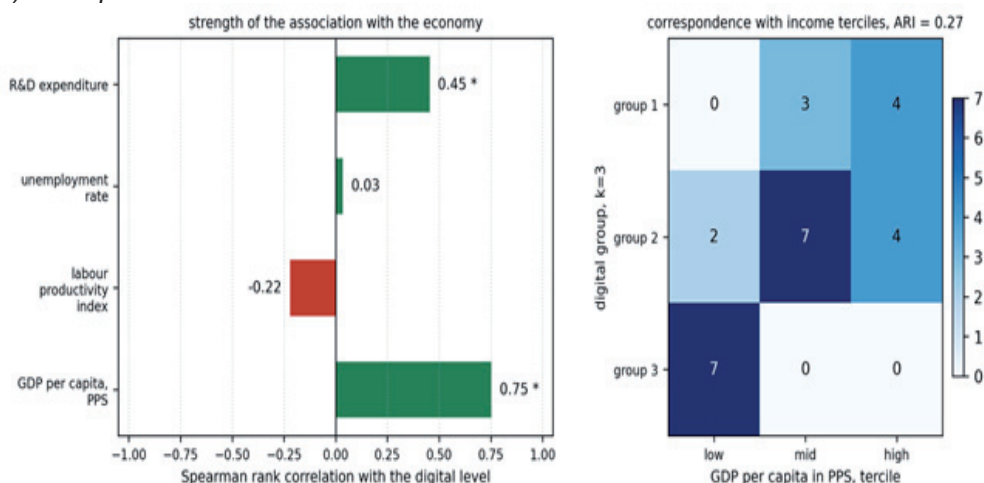
The link between the digital typology and economic development

H3 states that grouping countries by features of digital development does not coincide with their partition by the level of economic development. Figure 5 presents two pieces of evidence for each regime: on the left, the strength of the link between the digital level and 4 macroeconomic indicators; on the right, the agreement of the three-group partition with the partition of countries into GDP per capita terciles.

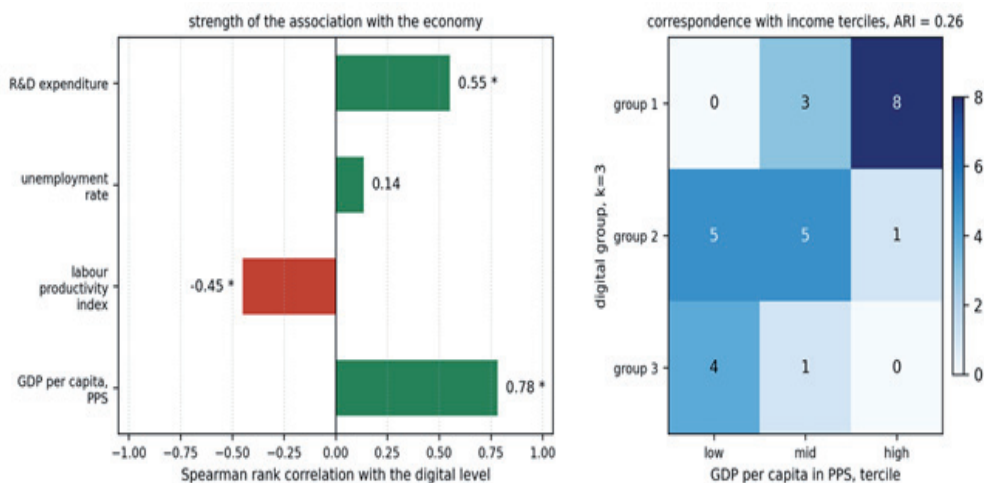
Figure 5

Link of the digital level with macroeconomic indicators and the agreement of the typology with the income terciles

a) DESI panel



b) Digital Decade panel



Source: developed by the authors on Eurostat (n.d) and DESI data (European Commission, 2026a).

As Figure 5 shows, the link with prosperity is indeed strong, the rank correlation of the digital level with GDP per capita in purchasing power standards equals 0.782 in the Digital Decade regime and 0.752 in the DESI regime, both significant also after the correction for multiplicity, and the level is also positively correlated with expenditure on research and development, 0.552 and 0.454, a link that withstands the correction in the Digital Decade regime but not in the DESI regime. At the same time, the digital partition does not duplicate the income ranking, since the agreement of the digital groups with the partition by income is weak, equal to 0.260 in the Digital Decade regime and 0.268 in the DESI regime; that is, the typology is not fully reproduced by the partition of the countries by income. For the four-group reading, this agreement is even lower, 0.172 and 0.166, so the conclusion does not depend on the number of groups chosen.

The corresponding correlation values are summarised in Table 4.

Table 4

Strength of the rank link of the digital level with macroeconomic indicators by regime, with uncorrected and Holm-corrected p-values, 27 countries

Macroeconomic indicator	ρ (DESI)	p-value (DESI)	Holm p-value (DESI)	ρ (Digital Decade)	p-value (Digital Decade)	Holm p-value (Digital Decade)
GDP per capita, PPS	0.752	< 0.001	< 0.001	0.782	< 0.001	< 0.001
Labour productivity, index	-0.221	0.268	0.804	-0.449	0.019	0.087
Unemployment rate	0.034	0.868	1.000	0.135	0.503	1.000
R&D intensity	0.454	0.018	0.087	0.552	0.003	0.017

Source: developed by the authors on Eurostat (2026) and DESI data (European Commission, 2026a).

As Table 4 shows, the labour productivity index is linked to the digital level by an inverse relation in the Digital Decade regime, significant under the uncorrected p-value and not significant after the correction, whereas the unemployment rate shows no significant link in either regime. The inverse link with productivity follows from the base of comparison, since the series is indexed to 2015 and catching-up economies that start from a lower base show a higher growth rate of the index than the leaders with a higher level of productivity. The correct statement is therefore narrow: the digital level moves inversely to the index of the cumulative change in productivity, not to its level. A strong rank correlation, together with weak agreement between the groups, confirms H3.

Discussion

The results obtained refine three directions of previous research. Regarding the convergence of digital development, our result does not contradict Andrei et al. (2023) and Kovács et al. (2022), yet it shows that the direction of the conclusion is not homogeneous across domains or measurement regimes. Within the discontinued composite, the gap widened in 3 of 4 domains, whereas in the Digital Decade indicator system, Digital Public Services and Human Capital converged fastest. Hence, a general statement about convergence is valid only with an indication of the indicator system on which it was obtained.

Regarding the typology of countries, our partition into 3 groups reproduces, in the current indicator system, the structure identified by Cruz-Jesus et al. (2012) and adds a quantitative estimate of stability that is absent from that work. The estimate reveals a difference between the regimes that had not been recorded before. In the later regime, the partition is less stable than in the earlier one, and the least stable countries are located within the leading group rather than at the boundary with the trailing group. The problem of frequent changes in the index composition, pointed out by Bánhidi et al. (2020), becomes measurable here, since the change in the indicator set is expressed as a shift in the grouping.

As regards the link between digital development and the economy, de Maya Matallana et al. (2026) construct levels of digital development for EU countries and compares them with economic characteristics. Our results refine this picture: a high rank correlation between the level and income, combined with low agreement across groupings, means that the typology carries information that the income-based partition does not reproduce. Separately, our finding on the Integration domain aligns with Borowiecki et al. (2021), who recorded an increase in polarisation in precisely this domain on DESI data. We add a quantitative estimate of the fragility of that finding under a composition of 2 indicators; the direction does not withstand alternative treatments, whereas under the extended composition, the divergence is confirmed.

This work differs from the studies reviewed chiefly in how the comparison is built. They assess convergence and typology within a single measurement regime, as do Zeinali & Mussari (2025) for data from the Digital Decade period, whereas here the results are obtained separately in each of the two regimes and compared only at the level of method and partition agreement. This separates a change in the phenomenon itself from a change in the instrument that measures it.

Conclusions

The article assessed the digital divide and constructed a typology of the European Union countries by the level and dynamics of digital development across two measurement regimes, applying a method robust to the structural incompatibility between the discontinued DESI composite and the Digital Decade indicator system. We tested the three hypotheses separately in each regime.

Under H1, we found that the direction of change in the digital divide differs across domains. In the DESI regime, the gap between countries widened in 3 of 4 domains, fastest in the Integration domain, whereas Connectivity remained practically

unchanged. In the Digital Decade regime, Digital Public Services and Human Capital converged fastest, whereas Integration continued to diverge. This conclusion holds under the alternative treatments only with the extended domain composition, since with 2 indicators the direction is set by the treatment of the number of unicorn companies. Convergence in the Human Capital domain is driven by the internet access indicators, whose values in the leading countries approach the scale limit, whereas the skills indicators widen the gap in both regimes.

Under H2, it was established that the partition into 3 groups is reproduced in both regimes. Among the solutions with three or more groups, the highest partition quality is achieved at 3 groups, and only the three-group solution withstands the stability check in both regimes, since the four-group solution clears the threshold only in the DESI regime. The average agreement of the partition under country resampling is 0.974 in the DESI regime and 0.697 in the Digital Decade regime, and under the main scheme the Jaccard indices of all groups exceed the 0.75 threshold in both regimes, whereas under the perturbation of the features 2 groups out of 3 clear it in the Digital Decade regime.

Under H3, the digital level is closely related to income, whereas grouping countries by digital features is not reproduced by the income partition. The rank correlation of the digital level with GDP per capita in purchasing power standards is high in both regimes, while the agreement of the digital groups with the income terciles is weak. The results support all three hypotheses within the formulated limits.

A separate practical result is the quantitative estimate of the structural shift between the two regimes. Eight of 27 countries are classified differently in the two regime analyses, with a partition agreement of 0.291. A separate level-based construction shows that the temporal benchmark with an unchanged indicator set yields 5 reassignments, whereas the indicator benchmark with an unchanged year 2021 yields 10. Since the adjusted Rand index is not additive and the two factors are not orthogonal, this construction serves as a diagnostic rather than a causal separation, and it is not possible to attribute the transition of each individual country separately. A comparison of the two regimes is valid at the level of the method and of the agreement of the partitions. The work's scientific contribution lies in a comparison design for the digital development of EU countries that remains valid under changes in the measurement system, because it compares country partitions rather than indicator values. Empirical evidence shows that the three-group structure is reproduced in both regimes, whereas the composition of the groups changes during the transition.

Further research is advisable in three directions that deepen the individual results of this work. The first direction concerns the methodological properties of the composite DESI itself, in particular verifying the internal consistency of its domain structure, which will make it possible to assess how far the conclusions depend on the index's construction. The second direction concerns a deeper analysis of the digital convergence of EU countries over time, with the estimation of beta convergence on initial levels and a test of the club convergence hypothesis, which will make it possible to establish whether the countries' convergence is preserved over a longer period. The third direction concerns extending the analysis to the regional level and including macroeconomic factors in an explanatory model of digital development.

Acknowledgments

The authors gratefully acknowledge the support provided by the UNITE project BG16RFPR002-1.014-0004, funded by PRIDST, in terms of research infrastructure, collaboration, and thematic discussions; this project did not provide direct financial support for the empirical analysis or the preparation of this manuscript.

The generative assistant Claude (Anthropic, Opus 5.0) and Grammarly Generative AI (Grammarly Inc.) were used to verify and correct the analytical code and to edit the language and style of the authors' text; they did not generate research results, and the authors take full responsibility for the content, in line with the journal policy on artificial intelligence.

References

- Andrei, J. V., Chivu, L., Sima, V., Gheorghe, I. G., Nancu, D., & Duică, M. (2023). Investigating the digital convergence in European Union: An econometric analysis of pitfalls and pivots of digital economic transformation. *Economic Research-Ekonomska Istraživanja*, 36(2), Article 2142814. <https://doi.org/10.1080/1331677X.2022.2142814>
- Arbelaitz, O., Gurrutxaga, I., Muguerza, J., Pérez, J. M., & Perona, I. (2013). An extensive comparative study of cluster validity indices. *Pattern Recognition*, 46(1), 243–256. <https://doi.org/10.1016/j.patcog.2012.07.021>
- Bánhid, Z., Dobos, I., & Nemeslaki, A. (2020). What the overall Digital Economy and Society Index reveals: A statistical analysis of the DESI EU28 dimensions. *Regional Statistics*, 10(2), 42–62. <https://doi.org/10.15196/RS100209>
- Borowiecki, R., Siuta-Tokarska, B., Maroń, J., Suder, M., Thier, A., & Žmija, K. (2021). Developing digital economy and society in the light of the issue of digital convergence of the markets in the European Union countries. *Energies*, 14(9), Article 2717. <https://doi.org/10.3390/en14092717>
- Bruno, G., Diglio, A., Piccolo, C., & Pipicelli, E. (2023). A reduced Composite Indicator for Digital Divide measurement at the regional level: An application to the Digital Economy and Society Index (DESI). *Technological Forecasting and Social Change*, 190, Article 122461. <https://doi.org/10.1016/j.techfore.2023.122461>
- Cruz-Jesus, F., Oliveira, T., & Bacao, F. (2012). Digital divide across the European Union. *Information & Management*, 49(6), 278–291. <https://doi.org/10.1016/j.im.2012.09.003>
- de Maya Matallana, M., García-Luque, O., López-Martínez, M., & Rodríguez-Pasquín, M. (2026). Digital development levels in the European Union: Measurement and analysis. *Economies*, 14(2), Article 58. <https://doi.org/10.3390/economies14020058>
- Dobos, I., & Bánhid, Z. (2025). Where Central and Eastern European countries stand in terms of digital readiness. *Society and Economy*, 47(1), 66–84. <https://doi.org/10.1556/204.2024.00012>

- European Commission. (2026a). *Digital economy and Society indicators* [Data set]. Directorate-General for Communications Networks, Content and Technology. <https://digital-decade-desi.digital-strategy.ec.europa.eu/datasets/desi/charts>
- European Commission. (2026b). *Commission Staff Working Document: DESI 2026 methodological note* (SWD(2026) 154 final). Directorate-General for Communications Networks, Content and Technology. <https://digital-strategy.ec.europa.eu/en/news-redirect/938549>
- Eurostat. (n.d.). *Eurostat database* [Data set]. European Commission. <https://ec.europa.eu/eurostat/web/main/data/database>
- Gomes, A., & Dias, J. G. (2025). Digital divide in the European Union: A typology of EU citizens. *Social Indicators Research*, 176(1), 149–172. <https://doi.org/10.1007/s11205-024-03452-2>
- Greco, S., Ishizaka, A., Tasiou, M., & Torrìsi, G. (2019). On the methodological framework of composite indices: A review of the issues of weighting, aggregation, and robustness. *Social Indicators Research*, 141(1), 61–94. <https://doi.org/10.1007/s11205-017-1832-9>
- Hennig, C. (2007). Cluster-wise assessment of cluster stability. *Computational Statistics and Data Analysis*, 52(1), 258–271. <https://doi.org/10.1016/j.csda.2006.11.025>
- Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6(2), 65–70. <https://www.jstor.org/stable/4615733>
- Kolupaieva, I., & Tiesheva, L. (2023). Asymmetry and convergence in the development of digital technologies in the EU countries. *Equilibrium. Quarterly Journal of Economics and Economic Policy*, 18(3), 687–716. <https://economic-policy.pl/index.php/eq/article/view/2699>
- Kovač, N., Žmija, K., Roy, J. K., Kusa, R., & Duda, J. (2024). Digital divide and digitalization in Europe: A bibliometric analysis. *Equilibrium. Quarterly Journal of Economics and Economic Policy*, 19(2), 463–520. <https://economic-policy.pl/index.php/eq/article/view/2899>
- Kovács, T. Z., Bittner, B., Huzsvai, L., & Nábrádi, A. (2022). Convergence and the Matthew effect in the European Union based on the DESI index. *Mathematics*, 10(4), Article 613. <https://doi.org/10.3390/math10040613>
- Kuc-Czarnecka, M., Lo Piano, S., & Saltelli, A. (2020). Quantitative storytelling in the making of a composite indicator. *Social Indicators Research*, 149(3), 775–802. <https://doi.org/10.1007/s11205-020-02276-0>
- Masoura, M., & Malefaki, S. (2023). Evolution of the Digital Economy and Society Index in the European Union: A socioeconomic perspective. *TalTech Journal of European Studies*, 13(2), 177–203. <https://doi.org/10.2478/bjes-2023-0020>
- Milligan, G. W., & Cooper, M. C. (1988). A study of standardization of variables in cluster analysis. *Journal of Classification*, 5(2), 181–204. <https://doi.org/10.1007/BF01897163>
- Murtagh, F., & Legendre, P. (2014). Ward's hierarchical agglomerative clustering method: Which algorithms implement Ward's criterion? *Journal of Classifica-*

-
- tion, 31(3), 274–295. <https://doi.org/10.1007/s00357-014-9161-z>
- Olczyk, M., & Kuc-Czarnecka, M. (2022). Digital transformation and economic growth – DESI improvement and implementation. *Technological and Economic Development of Economy*, 28(3), 775–803. <https://doi.org/10.3846/tede.2022.16766>
- Saraçlı, S., Doğan, N., & Doğan, İ. (2013). Comparison of hierarchical cluster analysis methods by cophenetic correlation. *Journal of Inequalities and Applications*, 2013, Article 203. <https://doi.org/10.1186/1029-242X-2013-203>
- Stavytskyy, A., Kharlamova, G., & Stoica, E. A. (2019). The analysis of the Digital Economy and Society Index in the EU. *TalTech Journal of European Studies*, 9(3), 245–261. <https://doi.org/10.1515/bjes-2019-0032>
- Szajt, M., Smolag, K., & Chomiak-Orsa, I. (2024). Digitization in European Union countries. Convergence analysis based on DESI. *Procedia Computer Science*, 246, 5448–5457. <https://doi.org/10.1016/j.procs.2024.09.688>
- Szeles, M. R., & Simionescu, M. (2020). Regional patterns and drivers of the EU digital economy. *Social Indicators Research*, 150(1), 95–119. <https://doi.org/10.1007/s11205-020-02287-x>
- Troitiño, D. R. (2022). The European Union facing the 21st century: The digital revolution. *TalTech Journal of European Studies*, 12(1), 60–78. <https://doi.org/10.2478/bjes-2022-0003>
- Warrens, M. J., & van der Hoef, H. (2022). Understanding the adjusted Rand index and other partition comparison indices based on counting object pairs. *Journal of Classification*, 39(3), 487–509. <https://doi.org/10.1007/s00357-022-09413-z>
- Zeinali, Z., & Mussari, R. (2025). Expansion of the digital realm: Analyzing DESI and digital transformation across EU member states, with a focus on Spain and Italy: A cluster analysis. *Public Organization Review*, 25, 1233–1273. <https://doi.org/10.1007/s11115-025-00836-x>
- Zherlitsyn, D. (2024). *Financial data analysis using Python*. Mercury Learning and Information. <https://doi.org/10.1515/9781501521843>
- Zherlitsyn, D., Kolarov, K., & Rekova, N. (2025). Digital transformation in the EU: Bibliometric analysis and digital economy trends highlights. *Digital*, 5(1), Article 1. <https://doi.org/10.3390/digital5010001>
- Zherlitsyn, D., & Rekova, N. (2026). Comparative analysis of labor markets in Bulgaria, Italy, and the UK: Wage dynamics, labor costs, and digital development. *Economies*, 14(1), Article 13. <https://doi.org/10.3390/economies14010013>

Appendix A

Table A1

List of the indicators used, with units, coverage and missing-value counts

Regime	Domain	Indicator	Break-down	Unit	Years	Observations	Missing
Digital Decade	Human Capital	ICT specialists (desi_ict_spec)	total	pc_ind_emp	2021–2025	135	0
Digital Decade	Human Capital	Internet use (desi_iuse)	ind_total	pc_ind	2021–2025	135	0
Digital Decade	Human Capital	Overall internet take-up (desi_ci_in_h)	hh_total	pc_hh	2021–2025	134	1
Digital Decade	Connectivity	Fibre to the premises coverage (desi_ftp)	hh_deg3	pc_hh	2021–2025	135	0
Digital Decade	Connectivity	Very high capacity network coverage (desi_vhcn)	hh_deg3	pc_hh	2021–2025	135	0
Digital Decade	Connectivity	Overall 5G coverage (desi_5gcov)	hh_deg3	pc_hh	2021–2025	135	0
Digital Decade	Connectivity	5G spectrum (desi_5g_spectrum)	5g_pb	pc_spectrum_assigned	2021–2025	135	0
Digital Decade	Connectivity	Share of fixed broadband subscriptions of at least 100 Mbps (desi_bbspeed_100)	total	pc_lines	2021–2025	135	0
Digital Decade	Connectivity	Share of fixed broadband subscriptions of at least 1 Gbps (desi_bbspeed_1000)	total	pc_lines	2021–2025	135	0
Digital Decade	Integration	Number of uni-coms (desi_unicom)	total	nb_unicom	2021–2025	135	0
Digital Decade	Integration	SMEs selling on-line (desi_sme_esell)	ent_sm_xfin	pc_ent	2021–2025	135	0
Digital Decade	Digital Public Services	Digital public services for citizens (desi_dps_cit)	national	egov_score	2021–2025	135	0
Digital Decade	Digital Public Services	Digital public services for businesses (desi_dps_biz)	national	egov_score	2021–2025	135	0
Digital Decade	Digital Public Services	User support (desi_us)	national	egov_score	2021–2025	135	0
Digital Decade	Digital Public Services	Mobile friendliness (desi_mf)	egov_le_all	egov_score	2021–2025	135	0

Regime	Domain	Indicator	Break-down	Unit	Years	Observations	Missing
Digital Decade	Digital Public Services	Transparency of service delivery, design and personal data (desi_tdpd)	egov_le_all	egov_score	2021–2025	135	0
Digital Decade	Digital Public Services	Pre-filled forms (desi_pff)	egov_le_all	egov_score	2021–2025	135	0
DESI	Connectivity	Broadband price index (desi_c_bbpi)	sub-dimension score	DESI points	2016–2021	162	0
DESI	Connectivity	Fixed broad-band coverage (desi_c_fbcc)	sub-dimension score	DESI points	2016–2021	162	0
DESI	Connectivity	Fixed broad-band take-up (desi_c_fbbtc)	sub-dimension score	DESI points	2016–2021	162	0
DESI	Connectivity	Mobile broad-band (desi_c_mbb)	sub-dimension score	DESI points	2016–2021	162	0
DESI	Digital Public Services	eGovernment (desi_dps_egov)	sub-dimension score	DESI points	2016–2021	162	0
DESI	Human Capital	Advanced skills and development (desi_hc_asd)	sub-dimension score	DESI points	2016–2021	162	0
DESI	Human Capital	Internet user skills (desi_hc_bsu)	sub-dimension score	DESI points	2016–2021	162	0
DESI	Integration	Digital intensity (desi_idt_di)	sub-dimension score	DESI points	2016–2021	162	0
DESI	Integration	Digital technologies for business (desi_idt_dtb)	sub-dimension score	DESI points	2016–2021	162	0
DESI	Integration	E-commerce (desi_idt_ecomm)	sub-dimension score	DESI points	2016–2021	162	0

Note: the indicator names follow the European Commission's methodological note (European Commission, 2026b) and Eurostat (n.d).

Source: compiled by the authors.

Received: June 28, 2026.
Reviewed: September 10, 2026.
Accepted: September 18, 2026.