

ICT ADOPTION AND THE BUSINESS ENVIRONMENT FOR ENTERPRISE PERFORMANCE: A CROSS-COUNTRY TYPOLOGY

Ben T. Ocra^{1*}, Godfred Yaw Koi-Akrofi², Henry Akwetey Martey³, Isaac Asampana⁴, Bansah Promise⁵

¹*Department of Information Technology/University of Professional Studies, Accra, Ghana*

²*Faculty of Information Technology and Communication Studies/University of Professional Studies Accra, Ghana*

^{3&4}*Department of Information Technology/University of Professional Studies, Accra, Ghana*

⁵*Kumasi Centre for Collaborative Research in Tropical Medicine/Kwame Nkrumah University of Science and Technology, Kumasi, Ghana*

* Corresponding Author Email: ben.ocra@upsamail.edu.gh

ABSTRACT

This research provides a data-driven country typology, which is based on empirically derived clusters of ICT adoption, business environment and firm performance, instead of on predefined income or regional categories. The data at country level were gathered for 178 countries, based on 15 indicators across three areas: digital infrastructure, business environment, and firm performance. Digital infrastructure and credit-access indicators were sourced from World Bank World Development Indicators. The indicators were reduced to four component scores by domain-specific principal component analysis and clustered by k-means and ward hierarchical clustering. Stability of the clusters was evaluated by bootstrap consensus clustering, and was tested with the Adjusted Rand Index, PERMANOVA, Kruskal-Wallis tests and sensitivity analyses. Three empirically derived clusters were identified. Cluster 1 showed less favorable digital-infrastructure indicators and firm-level innovation indicators, alongside higher financial-constraint and infrastructure-reliability burdens. Cluster 2 combined more favorable digital-infrastructure and finance-access indicators with comparatively low firm-level innovation indicators. Cluster 3 showed favorable values across nearly all indicators. The findings indicated that stronger digital infrastructure indicators were not always correlated with stronger firm-innovation indicators, suggesting enterprise-development strategies may benefit from taking a holistic view of technological, business-environment, and organizational conditions, with longitudinal data needed to establish developmental trajectories.

Keywords: ICT Adoption, Business Environment, Firm Performance, Country Typology, Consensus Clustering, Sub-Saharan Africa, Digital Divide

1. INTRODUCTION

Information and communication technology (ICT) has been integral component of national economic development policy strategies and governments are increasingly allocating resources to enhance the digital infrastructure to boost productivity, competitiveness of the business and to promote the growth of Small and Medium Sized Enterprises (SMEs). The overall impact of ICT on economic growth is positive, although the impact is different across countries and stages of development (Bahrini & Qaffas, 2019; Kumari & Singh, 2023; Pradhan et al., 2022). Evidence from both macroeconomic and firm-level studies suggest that the impact of ICT on economic performance is not uniform. Bahrini and Qaffas (2019) found positive average impact of ICTs on growth in developing economies and high regional and technology heterogeneity. In the same way, Kumari and Singh (2023) found that financial development and trade openness influence the degree of complementarities between ICT infrastructure and the growth process, and that the complementarities are greater in low-income than in high-income economies. A similar conclusion was made by Karim et al. (2022) in a meta-analysis of 147 empirical studies. These literatures, are

individual, business-environment and firm level research which provide support for considering the adoption of ICTs, business-environment quality, and firm performance as interdependent dimensions of national business development and should be examined jointly rather than in isolation.

However, there are comparatively few studies that have combined these dimensions within a single comparative framework. The existing cross-country comparisons are mainly based on income classifications or regional grouping, which are useful for making a broad comparison but do not reflect meaningful differences in the countries in the same income group or in the countries in the same geographic region (Fantom & Serajuddin, 2016; Khan et al., 2022). Similar limitations apply to single-domain indices, like regulatory or business-environment indices, that can be very different even between neighboring countries (Rogge & Archer, 2021). This research fills this gap by developing a data-driven typology of 178 countries, which is based on indicators of ICT adoption, quality of the business environment and firm performance, combined within a single analytical framework. The research is guided by two research questions. First, do countries have distinct, multidimensional profiles when analyzing the adoption of ICTs together with the conditions of the business environment and firm performance, and if they do, how are digital infrastructure and firm-level outcomes distributed across these profiles? Second, is there any sub-regional structure in Sub-Saharan Africa that is not reflected in classifications using the region as an analytical unit? The Research provides an analytical approach that can be used in future comparative studies and evidence for policymakers to identify countries with structurally conditions in designing ICT and SME development strategies. The remainder of this paper presents the literature review, methodology deployed, findings, conclusions and implications of the Research.

2. LITERATURE REVIEW

2.1 Theoretical Framework

The research was based on the Technology–Organization–Environment (TOE) framework (Tornatzky et al., 1991) that identifies three contexts: technological, organizational, environmental, that influence outcomes of technology adoption. TOE believes that outcomes are a function of the technology available and the organizational capacity and external context. This is of relevance here because countries are not only different in terms of access to digital infrastructure, they are also different in terms of conditions under which firms use digital infrastructure, which Myovella et al., (2021) examined in relation to internet and broadband disparities across 41 Sub-Saharan African countries, but which does not represent an empirical test of TOE specifically at the country level. The research's domains are organized by the three TOE contexts with each context having a stated limitation. Digital-infrastructure indicators are examples of the technological context. Business-environment indicators (finance access, infrastructure reliability) are indicators of the environment. The other indicators (tax e-filing, certification, product/process/R&D activity) refer to the organizational context, but are based on firm practices and not on firm performance, such as profitability; aggregated to the country level, they proxy organizational capacity, not individual-firm performance.

The two moves from the original TOE framing (as a firm-level, performance-based approach) are explicitly mentioned, not suggested. Based on the synthesis of 147 studies, Karim et al., (2022) concluded that the impact of ICT on the performance of firms is dependent on the level of ICT development in the country. Digital technology use is not the same as adoption and innovation, and adoption is not the same as productivity in a study of 711 South African manufacturing micro/small enterprises, Gaglio et al., (2022) found that digital technology use was associated with innovation, which in turn was associated with productivity. TOE's function here is not confirmatory, but organizational, it is justifying an analysis of three domains together, rather than separately, and not testing a country-level causal mechanism, which this study does not test. This is especially

important for Sub-Saharan Africa, where country-level variation can be significant and could be lost in the analysis.

2.2 Empirical Review

The digital divide has been seen as a lack of equal access to communication technologies, but recent studies have considered it a more complex construct, which includes the cost of technology, connectivity, digital literacy, usage patterns, and the economic benefits that companies and people get from digital technologies (Robinson et al., 2020). These dimensions have been consistently linked with structural factors such as income, institutional quality, electricity infrastructure and governance capacity across countries (Mardikyan et al., 2015; Myovella et al., 2021). Firm-level studies also report positive effects of digital technologies on SME productivity, innovation, and organizational performance (Chege et al., 2019; Eller et al., 2020; Gaglio et al., 2022).

This literature justifies the three domains of analysis employed in this research. Research on business-environment quality shows that technology is only one part of the equation; research on the digital divide shows that access is not enough; and firm-level research shows that organizational capability is the key link between digital investment and performance. The three dimensions; adoption of ICTs, the quality of the business environment and firm performance are conceptually interdependent, yet existing research has typically examined them separately rather than jointly. They have rarely been combined in a single typology, however: the studies that classify countries do so at one level at a time, either on digital readiness or on general development indicators; no typology, to the best of authors' knowledge, classifies countries at all three levels simultaneously, at the country level. This gap directly drives the multidimensional country typology developed in this research with the help of principal component analysis and consensus clustering, focusing on the Sub-Saharan African economies within the global pattern.

2.3 Research Propositions

Existing empirical evidence provides support for this multidimensional perspective. Studies generally report positive relationships between ICT development and economic or firm-level performance, but the magnitude and direction of these relationships vary considerably across countries and firms. Bahrini and Qaffas (2019), Kumari and Singh (2023), and Pradhan et al. (2022), for example, demonstrate that the economic benefits associated with ICT infrastructure differ across development contexts. The literature therefore indicates that similar levels of technological development need not produce similar enterprise outcomes. This argument is consistent with Karim et al. (2022), whose meta-analysis shows that the relationship between ICT use and firm performance varies according to countries' ICT development status.

Consequently:

Proposition 1 (P1): Countries will exhibit distinct multidimensional profiles based on the joint configuration of digital infrastructure, business-environment conditions, and firm-level innovation and organizational practices.

The literature further suggests that digital infrastructure alone is insufficient to explain differences in enterprise outcomes.

Proposition 2 (P2): Higher levels of digital infrastructure will not necessarily coincide with equally favorable business-environment conditions or stronger firm-level innovation and organizational practices across countries.

The business environment provides another important source of variation. Véganzonès-Varoudakis and Nguyễn (2018) show that regulatory quality, infrastructure, governance and finance affect firms' capacity to compete and innovate. The literature reviewed similarly indicates that, institutional and infrastructural conditions influence whether digital investment translates into productive outcomes.

Proposition 3 (P3): Countries with comparable levels of digital infrastructure will exhibit different firm-level innovation and organizational profiles where their business-environment conditions differ.

The organizational dimension further strengthens this argument. Research increasingly demonstrates that possession or adoption of ICT does not automatically produce superior organizational outcomes. Rather, complementary capabilities; including innovation capacity, managerial practices, human capital and the capacity to redesign business processes.

3. METHODS

3.1 Study Design and Analytical Strategy

This research used an unsupervised, cross-national design to identify country typologies based on digital infrastructure, business environment, and firm performance letting groupings emerge from the data rather than relying on predefined income or regional categories. The workflow had eight steps: (1) compilation of indicators from publicly available data, which were merged by country code; (2) cleaning of data and screening for missing values and outliers; (3) z-scoring of indicators and conducting PCA for each domain to reduce dimensionality; (4) testing cluster-number solutions and cluster using k-means and Ward hierarchical methods; (5) stability testing via bootstrap consensus clustering; (6) internal tests and sensitivity analysis of the solution; (7) profiling, labelling and mapping of clusters; and (8) conducting a sub-study of indicators for the sub-region of Sub-Saharan Africa to capture within-region structure that was not captured globally. All method decisions (diagnostic, algorithm, validation steps) were made prior to the finalization of clusters; labels were assigned after (Sub-section 3.5). The sample includes 178 countries (47 in Sub-Saharan Africa). The typology was evaluated for four parameters: statistical performance, stability of the resampling, consistency of the algorithms, and interpretive usefulness. Cluster analysis approach is particularly appropriate where multiple interrelated indicators are used to identify relatively homogeneous groups within heterogeneous observations (Everitt et al., 2011; Kaufman & Rousseeuw, 1990). This approach therefore allows countries to be classified according to their observed combinations of digital infrastructure, business-environment conditions, and firm-level innovation/organizational practices rather than imposing conventional income or geographic categories.

3.2 Data Sources and Country Matching

Country-level indicators came from three sources. Digital infrastructure and credit-access indicators (mobile subscriptions, internet use, broadband, private-sector credit) came from World Bank WDI (2026), via Our World in Data, for reference year 2021. Business-environment and firm-level indicators came from World Bank Enterprise Surveys and the Business Ready programme (2026), using each country's most recent survey round. These two sources aren't perfectly time-aligned: Enterprise Surveys run on a rolling, country-specific cycle (2009–2026 in this sample, with 64% surveyed 2023–2025), while WDI indicators are fixed at 2021 for all countries. Aligning both to 2021 was not feasible it would have left only one usable country, so the mismatch is treated as a limitation (see Limitations). Risk is highest for 14 countries with pre-2020 surveys (e.g., Belarus, Russia, Ukraine, Venezuela, Yemen); since two of these (Sudan, Timor-Leste) fall within the Sub-Saharan Africa sub-study, this is flagged again in Sub-section 3.7. Countries were matched via ISO3 codes. Of 198 countries with partial data, 20 were dropped for missing over 50% of the 15 indicators (e.g., Cuba, Iran, North Korea, UAE, several microstates). Most exclusions fall into two groups: small states/territories with no Enterprise Survey (11 countries) and states with limited international statistical cooperation, often due to sanctions or conflict (6 countries); the rest (Algeria, Haiti, UAE) are one-off gaps. Exclusions were spread across regions concentrated somewhat in Europe & Central Asia, MENA, and East Asia & Pacific/North America with no evidence of bias by income or performance level. The final sample: 178 countries, including 47 in Sub-Saharan Africa, used for the regional sub-study in Sub-section 3.7.

3.3 Indicator Selection

Fifteen indicators represent three domains digital infrastructure, business environment, and firm-level innovation/organizational practice chosen for relevance to the TOE framework (Sub-section 2.1) and data availability. Full definitions, sources, units, coding, and missingness are presented in Supplementary Table S1. The third domain (tax e-filing, quality certification, product/process/R&D activity) is called "firm-level innovation and organizational practice" rather than "firm performance," since it captures firm behaviours and capabilities not direct outcomes like profitability or revenue growth. Per the TOE framework (Sub-section 2.3), it serves as a proxy for firms' aggregate capacity to adopt digital tools and innovate, a plausible but imperfect correlate of performance. The underlying WBES/B-READY sample isn't SME-only. Enterprise Surveys cover formal, registered, non-agricultural firms with 5+ employees, stratified by size (small, medium, large, with an extra stratum for largest firms in some economies). Firms under 5 employees and informal-sector businesses are excluded. So while this data is often used in SME research, the indicators here reflect the full formal-sector size distribution, not an SME subsample a scope consideration given that firm dynamics can vary by size.

3.4 Missing data, transformation, and outlier detection

Missing data and imputation. Missingness ranged from 0–12.9% per indicator, mostly in Enterprise Survey items not covered by every country's latest round. Countries missing over 50% of indicators were already excluded (Sub-section 3.2). Remaining gaps were filled using each variable's sample median (computed across all 178 countries, not by region/income) chosen over the mean because several indicators were skewed even before transformation, and medians resist distortion by extreme values. Eleven variables received imputed values (e.g., domestic credit, internet-related and firm-innovation indicators); four (finance constraint, electrical outages, mobile subscriptions, quality certification) had no missing data. A sensitivity check re-ran the full pipeline (domain PCA + k-means, $k=3$) using mean instead of median imputation. The two solutions agreed closely (Adjusted Rand Index = 0.934), showing cluster assignments were robust to this choice. Transformation. Indicators with $|\text{skewness}| > 1.0$ were transformed: electrical outages, internet-connection days, and domestic credit (all right-skewed) via $\log(x+1)$; e-filing rate (left-skewed) via reflection then $\log(x+1)$. $\log(x+1)$ was used throughout to handle zero values without excluding them. No other indicators needed transformation. Outlier detection. Multivariate outliers were checked via Mahalanobis distance across all 15 standardized indicators, tested against a chi-square distribution ($df=15$, $\alpha=0.001$, critical value=37.70). No countries exceeded this threshold, so none were flagged, removed, or down-weighted.

3.5 Domain-Specific Principal Component Analysis

Because the 15 indicators span three conceptually distinct domains measured on different units and scales, PCA was applied separately within each domain rather than jointly across all indicators. Within each domain, all variables were standardized to z-scores (mean = 0, standard deviation = 1) prior to extraction, so that PCA was performed on the correlation matrix rather than the covariance matrix; this was necessary given the substantially different numeric ranges across indicators within each domain (for example, within the business-environment domain, finance constraint ranges 0–100 while days internet connection ranges 1–30 days).

3.5 Cluster Identification and Validation

Candidate cluster numbers ($k = 2\text{--}10$) were evaluated on the four PCA component scores using three diagnostics: within-cluster sum of squares (elbow), average silhouette width (cluster: silhouette), and the gap statistic (cluster: clusGap, 500 bootstrap replicates, max $k = 10$, first-SE rule). K-means used the Hartigan–Wong algorithm (stats: kmeans, 100 random starts, max 300 iterations; lowest-WCSS solution retained). Ward hierarchical clustering used Ward.D2 (stats: hclust, method = "ward. D2") on Euclidean distances from the same four components. Agreement

between the two methods was assessed via Adjusted Rand Index (`mclust::adjustedRandIndex`). Bootstrap consensus clustering (custom R procedure) tested stability across $k = 2-6$: 500 resamples, each 80% of the 178 countries (sampled with replacement), clustered via both k -means (`nstart = 50`) and Ward.D2. A consensus matrix recorded co-clustering proportions per country pair; PAC was calculated as the share of off-diagonal values strictly between 0.10 and 0.90, with lower PAC indicating a more stable solution. "Contestability measures" (undefined in the original draft) has been replaced with this specified consensus-matrix/PAC procedure.

3.6 Geographic Visualizations

Country cluster assignments were matched to spatial boundaries using ISO3 codes and visualized as thematic world maps in R version 4.6.1. Country boundary polygons (medium-resolution scale) were obtained from the `rnaturalearth` package; spatial data were processed using `sf` and plotted using `ggplot2`. Countries in the analytic sample without a corresponding boundary match in the `rnaturalearth` dataset specifically Kosovo (OWID_KOS), which lacked a standard ISO3-matched boundary in this dataset were excluded from the map visualization only, not from the statistical analysis; all other countries in the 178-country analytic sample matched successfully.

3.7 Sub-study: Within-Sub-Saharan Africa Typology

A regional sub-analysis was run because the global three-cluster solution placed most SSA countries into a single cluster (Sub-section 5.7), raising the question of whether finer sub-groups exist within the region. The 47 SSA countries were re-clustered via k -means on their four global PCA component scores (`digital_pc1`, `business_pc1`, `business_pc2`, `firm_pc1`).

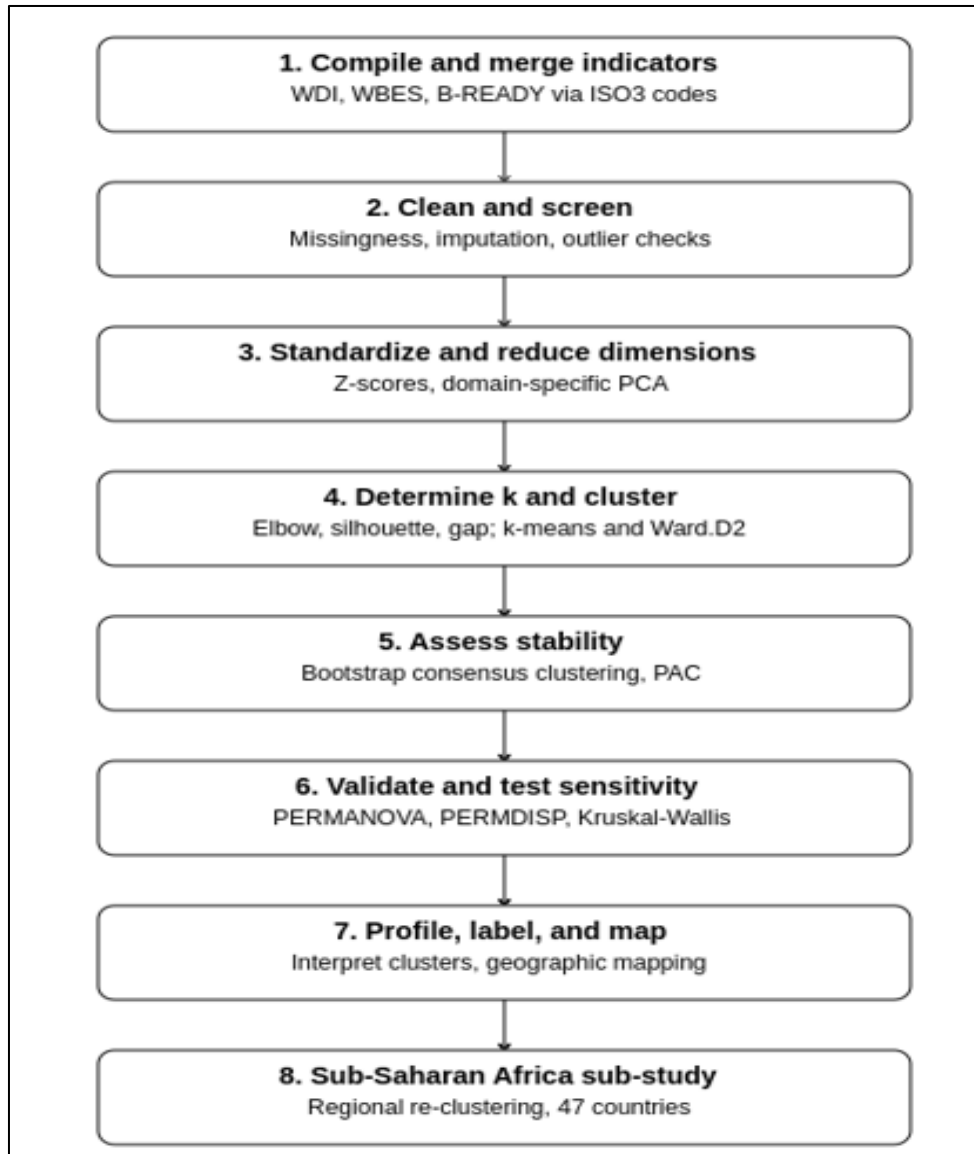
Why global (not SSA-specific) components were used: First, re-running PCA on just 47 countries would yield roughly 3 observations per component below accepted minimums for stable PCA. Second, using the same components as the global analysis keeps SSA sub-clusters comparable to global clusters (needed for the comparison in Sub-section 5.7). The trade-off, stated explicitly: global components reflect the full 178-country variance structure, not SSA-specific patterns, so they may not capture the dimensions along which SSA countries actually differ as efficiently as a region-specific PCA would. This is a stated limitation, not resolved here.

Cluster-number diagnostics: For $k = 2-7$ on the SSA subset, silhouette width peaked at $k = 2$ (0.320) and declined thereafter ($k=3$: 0.225; $k=4$: 0.235; $k=5$: 0.234) mirroring the global pattern where silhouette favors $k=2$ despite $k=3$ being retained on other grounds. WCSS declined steadily with no clear elbow at $k=3$.

Why $k = 3$ was retained: Not because regional diagnostics favoured it, but for direct comparability with the global three-cluster solution, enabling cross-tabulation of SSA countries' global vs. regional assignments (Sub-section 5.7). This is explicitly a design choice for comparability, not an empirically optimal SSA solution.

Stability: With only 47 countries, stability was tested via 200 bootstrap resamples (47 countries, with replacement), fitting k -means ($k=3$) to each and comparing to the full-sample assignment via ARI. Mean ARI = 0.484 (SD = 0.219) moderate stability, notably lower than the global solution's (Sub-section 5.5), consistent with reduced power in a small regional sample. This is reported as a limitation on the precise sub-cluster boundaries; however, the broader qualitative finding that SSA is not homogeneous is considered more robust than the exact three-way split, and is supported by independent evidence (Saba & David, 2020).

4. ANALYSIS



Source: Eight-Step Analytical framework developed and applied by the authors

5. FINDINGS/RESULTS

5.1 Descriptive Statistics

Descriptive statistics for the 15 indicators used in the clustering analysis are given in Table 1, calculated on observed values prior to imputation (Sub-section 3.4); the missingness percentages reported in Table 1 therefore reflect data availability in the final 178-country analytic sample before the median-imputation procedure was applied, not a residual level of missingness remaining after data preparation. Overall coverage was strong for the digital-infrastructure indicators (0.0%–1.1% missing). Indicators drawn from the World Bank Enterprise Surveys showed higher missingness, ranging from 0.0% to 12.9%, reflecting differences in survey timing and question coverage across countries (Sub-section 3.2). This missingness is reported here as an observed data characteristic rather than as an evaluation of whether the level is acceptable in the abstract; whether it is adequately addressed depends on the imputation procedure applied, not on the raw percentage alone. As detailed in Sub-section 3.4, values were imputed using the sample median for each affected variable, and the sensitivity of the clustering solution to this choice was tested directly by re-running the full analytic pipeline using mean imputation in place of median imputation; the resulting three-cluster solutions showed high agreement (Adjusted Rand Index = 0.934, Sub-section 3.4), indicating that cluster assignments were not substantially sensitive to the specific imputation method used for the missingness levels observed in this sample. This comparison, rather than a

general threshold for "acceptable" missingness, is offered as the basis for treating the imputation procedure as adequate for the clustering analysis reported here.

Table 1. Descriptive statistics and missingness for clustering indicators, final analytical sample (n = 178), prior to imputation.

Indicator	Mean	SD	Min	Max	% Missing
Mobile subscriptions per 100 people	107.33	34.32	19.71	199.24	0.0
Internet users (%)	65.13	27.25	6.67	100.00	1.1
Fixed broadband per 100 people	15.59	14.88	0.00	47.94	1.1
Domestic credit to private sector (% GDP)	53.88	42.69	2.75	219.13	12.9
Finance constraint index	34.10	15.82	6.40	73.00	0.0
Electrical outages (monthly)	3.30	6.30	0.00	41.90	0.0
Competition change index	75.87	8.75	47.60	94.90	9.0
Internet-switching difficulty	71.99	12.16	20.90	96.00	9.0
Days to internet connection	5.89	6.05	1.00	30.00	11.2
Firms reporting internet disruption (%)	34.97	19.63	0.70	82.40	9.0
Firms filing taxes electronically (%)	76.03	30.20	0.00	100.00	9.0
Firms with quality certification (%)	13.03	10.44	0.00	45.80	0.0
Firms introducing new product (%)	34.15	21.66	0.10	87.10	0.6
Firms introducing process innovation (%)	25.58	19.69	0.00	86.00	0.6
Firms reporting R&D spend (%)	19.16	13.01	0.00	59.50	0.6

Note: Statistics reflect the final 178-country analytical sample described in Sub-section 3.2, calculated on observed values prior to median imputation of remaining missing values (Sub-section 3.4). SD=Standard Deviation, Min=minimum, Max= maximum

5.2 Correlation analysis

To assess the appropriateness of dimension reduction, within-domain Pearson product-moment correlation matrices were computed for the 15 clustering indicators prior to PCA (Figure 1). Pearson's r was used because PCA itself operates on linear correlation structure (the correlation matrix, given the z-score standardization applied in Sub-section 3.5); a linear correlation measure was therefore the appropriate diagnostic for assessing whether PCA's linearity assumption was reasonably met for each domain, rather than a rank-based measure such as Spearman's ρ , which would not directly correspond to the relationships PCA decomposes. All three digital-infrastructure indicators were positively correlated ($r = .45$ to $.73$), consistent with a single dominant dimension and supporting the retention of one component in this domain (Sub-section 5.3). In the business-environment domain, correlations were more heterogeneous: domestic credit was negatively correlated with finance constraint ($r = -.48$) and electrical outages ($r = -.51$), while internet-switching difficulty and connection days correlated weakly with the remaining business-environment indicators ($r = .16$ to $.31$). This heterogeneity rather than a single dominant correlation pattern foreshadowed the two-component solution retained for this domain, since indicators that do not correlate strongly with the rest of their domain are unlikely to load together onto a single component. The three innovation indicators (new product introduction, process innovation, R&D spending) correlated with one another ($r = .61$ to $.63$), while quality certification and tax e-filing showed weaker or negative correlations with this group, prefiguring the loading pattern in which tax e-filing separated from the innovation indicators on the domain's second component (Sub-section 5.3).

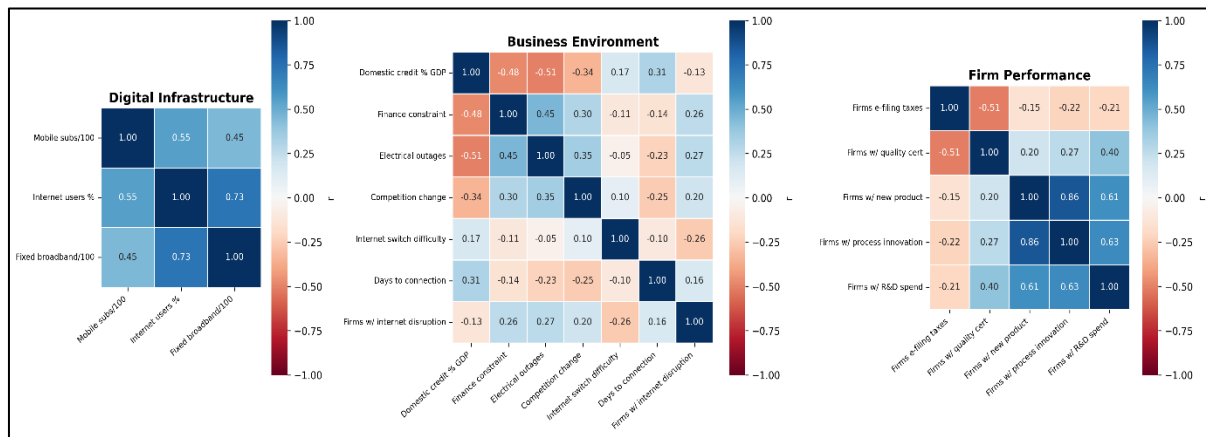


Figure 1. Correlation matrices for clustering indicators.

5.3 Domain principal component analysis

PCA was run separately for each domain after z-score standardization, using Kaiser's eigenvalue >1 rule for retention. The digital-infrastructure domain yielded one component (eigenvalue 2.16, 72.1% variance). The business-environment domain yielded two (2.49, 35.6%; 1.38, 19.8%): the first driven by finance constraint, outages, competition, and credit; the second by internet-switching difficulty, connection days, and disruption. The firm-level domain yielded one retained component (2.70, 54.1%), though a second component (eigenvalue 1.25, 24.9%) also exceeded the cutoff but was excluded on interpretive grounds rather than the stated rule an inconsistency we flag rather than hide. Re-running clustering with this second component included showed moderate agreement with the main result ($ARI = 0.87$; cluster sizes shifted from 66/62/50 to 62/54/62), indicating the choice has some but not drastic effect. All reported results use the original four-component solution (digital PC1, business PC1–PC2, firm PC1), consistent with the three-domain design.

5.4 Cluster-number selection

Elbow, silhouette, and gap statistic were computed for $k = 2$ –10. WSS declined from 924.6 ($k=2$) to 659.4 ($k=3$); gap statistic peaked at $k=3$ (0.4375 vs. 0.4050 at $k=2$); silhouette favoured $k=2$ (0.343 vs. 0.332). Because these three diagnostics disagreed, the pre-specified tie-breaker was bootstrap consensus stability (PAC, Sub-section 3.6), computed for all candidate k values as part of the same protocol: PAC was lower indicating more stable, less ambiguous clustering at $k=3$ (0.362) than at $k=2$ (0.406), independently supporting $k=3$ over $k=2$ on a criterion set prior to inspecting cluster content. To assess what is lost at $k=2$ substantively, we computed the $k=2$ solution directly and cross-tabulated it against the retained $k=3$ solution. The $k=2$ solution ($n=72, 106$) collapses almost entirely into: Nascent (62 of 62 countries) plus 10 Transitional countries in one group, and Transitional (56 of 66) merged with Advanced (all 50) in the other. In other words, $k=2$ erases the distinction between Transitional and Advanced Digital Economies precisely the contrast underlying this research's central finding that strong digital infrastructure does not consistently correspond to strong firm-level innovation (Sub-section 5.6). This confirms, with cluster-level evidence rather than assertion, that $k=2$ obscures a substantively important distinction the data otherwise support; $k=3$ is retained on this basis, together with the gap statistic and PAC evidence above.

5.5 Cluster Solution

K-means clustering (100 random starts, four domain component scores) produced three clusters: 62 countries (34.8%), 66 (37.1%), 50 (28.1%). Agreement with an independently derived Ward hierarchical solution was moderate ($ARI = 0.524$); two of three clusters aligned closely across algorithms, while the third split differently between methods. Bootstrap consensus clustering (500 resamples, both algorithms) produced a non-monotonic PAC pattern across k : 0.406 ($k=2$), 0.362 ($k=3$), 0.385 ($k=4$), 0.338 ($k=5$), 0.314 ($k=6$) PAC is in fact lower, not higher, at $k=5$ and $k=6$ than at $k=3$. PAC alone therefore does not support $k=3$ as optimal; taken in isolation it would favor $k=6$.

The $k=3$ solution was instead retained on the combined weight of the gap statistic, cluster-level substantive evidence, and pre-specified stability criteria detailed in Sub-section 5.4, alongside the observation that the $k=6$ solution showed markedly uneven bootstrap stability across its sub-clusters (mean within-cluster consensus ranging 0.446–0.947) (Figure 2) a pattern consistent with over-partitioning rather than genuine additional structure. This solution is presented as the best-supported partition given the available diagnostics, not as evidence of a single "true" underlying k , consistent with the exploratory nature of this clustering analysis.

5.6 Cluster Profiles

Table 2 reports mean z-scores for each of the 15 original indicators by cluster, computed on untransformed values (no indicators were reverse-coded prior to this table; raw indicator direction is preserved throughout). Because indicator direction differs across the set higher values indicate more favorable conditions for nine indicators (mobile subscriptions, internet users, fixed broadband, domestic credit, competition change, tax e-filing, quality certification, new product, process innovation, R&D spend) and less favorable conditions for the remaining six (finance constraint, electrical outages, internet-switching difficulty, days to internet connection, firms reporting internet disruption) cluster comparisons below are described according to each indicator's actual direction rather than by z-score sign alone, and "favorable" is used in place of "higher/stronger" where an indicator is adverse-when-high. Clusters were labelled Nascent (Cluster 1, $n=62$), Transitional (Cluster 2, $n=66$), and Advanced (Cluster 3, $n=50$) Digital Economies based on these profiles. Nascent Digital Economies showed the least favorable values on all three digital-infrastructure indicators ($z = -0.81$ to -1.14), the least favorable finance-constraint score ($z = 0.67$, indicating the highest constraint) and electrical-outages score ($z = 0.71$, indicating the most frequent outages), and the least favorable values on most firm-level innovation/practice indicators.

Transitional Digital Economies showed favorable digital-infrastructure scores ($z = 0.30$ – 0.56) and more favorable finance-constraint and outages scores than Cluster 1 ($z = -0.36$ and -0.38 respectively, indicating lower constraint and fewer outages), but the least favorable values of the three clusters on new product introduction ($z = -0.64$) and process innovation ($z = -0.61$). Advanced Digital Economies showed favorable values on most indicators, including the most favorable firm-level innovation/practice scores (new product $z=1.15$; process innovation $z=1.12$; R&D spend $z=1.07$) and favorable digital-infrastructure and credit-access scores. This cluster did not show the most favorable value on every indicator: its finance-constraint ($z=-0.36$) and electrical-outages ($z=-0.38$) scores were statistically indistinguishable from Cluster 2 at two decimal places, and its competition-change score ($z=-0.37$) was less favorable, in direction, than Cluster 1's $z=0.57$ would suggest under a naive higher-is-better reading correctly interpreted, `competition_change_index` is coded such that higher values reflect more reported change in competition, not unambiguously "better" conditions, and is treated in this study as a directionally ambiguous business-environment indicator rather than one with a clear favorable pole.

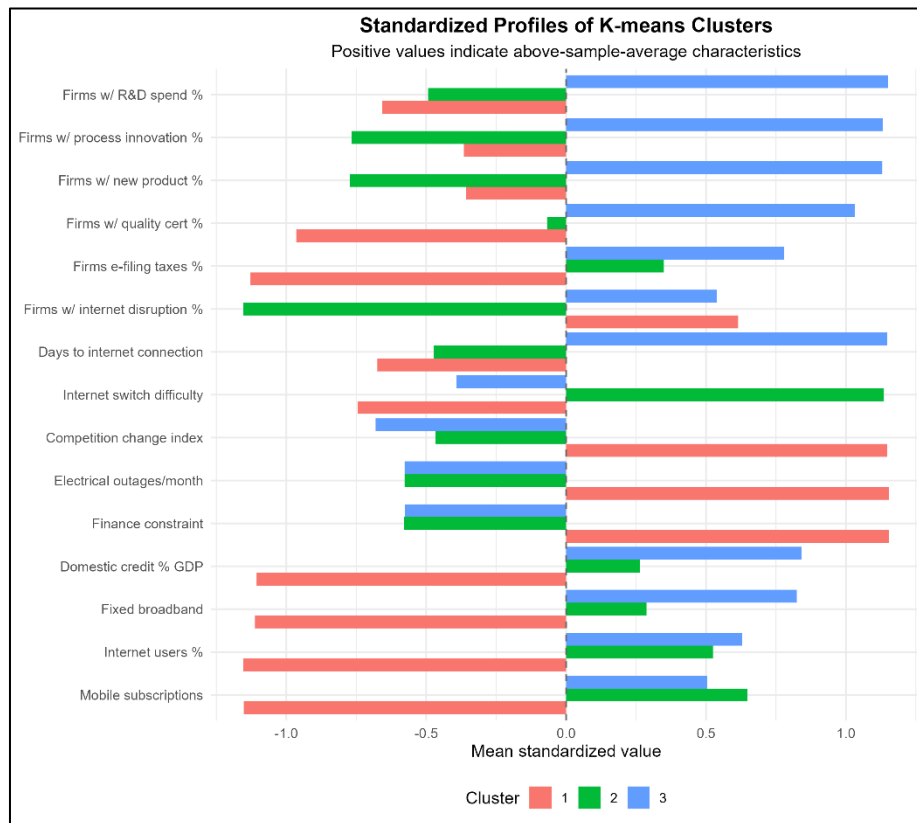


Figure 2. Within-domain correlation matrices for the fifteen clustering indicators.

Table 2. Mean standardized indicator profiles by final cluster.

Indicator	Nascent Digital Economies	Transitional Digital Economies	Advanced Digital Economies
Mobile subscriptions per 100 people	-0.81	0.48	0.37
Internet users (%)	-1.14	0.56	0.67
Fixed broadband per 100 people	-0.94	0.30	0.77
Domestic credit to private sector (% GDP)	-0.68	0.20	0.58
Finance constraint index	0.67	-0.36	-0.36
Electrical outages (monthly)	0.71	-0.38	-0.38
Competition change index	0.57	-0.26	-0.37
Internet-switching difficulty	-0.26	0.36	-0.15
Days to internet connection	-0.33	-0.22	0.69
Firms reporting internet disruption (%)	0.33	-0.54	0.29
Firms filing taxes electronically (%)	-0.69	0.25	0.52
Firms with quality certification (%)	-0.58	0.00	0.72
Firms introducing new product (%)	-0.25	-0.64	1.15
Firms introducing process innovation (%)	-0.25	-0.61	1.12
Firms reporting R&D spend (%)	-0.49	-0.35	1.07

Note: Values are z-scores calculated across the full 178-country sample. n = 62 for Nascent Digital Economies, n = 66 for Transitional Digital Economies, n = 50 for Advanced Digital Economies.

5.7 Geographic Distribution and Sub-Saharan Africa sub-study

The geographic distribution of the three global clusters is presented in Figure 3. Nascent Digital Economies were concentrated in Sub-Saharan Africa, South Asia, and parts of the Middle East. Advanced Digital Economies were concentrated in North America, Western Europe, East Asia, and Oceania, together with a smaller number of upper-middle-income economies. Transitional Digital Economies spanned Eastern Europe, Latin America, Southeast Asia, and parts of North and Southern Africa. Of the 47 Sub-Saharan African countries in the analytic sample, the majority were

classified as Nascent Digital Economies, a smaller number as Transitional Digital Economies, and none as Advanced Digital Economies. This concentration of SSA countries within a single global cluster motivated a supplementary regional analysis, re-clustering the 47 SSA countries using their global PCA component scores (Sub-section 3.7). Regional cluster-number diagnostics for the SSA subset, considered independently of the global analysis, favoured a two-cluster solution (silhouette width: $k=2$, 0.320; $k=3$, 0.225); $k=3$ was nonetheless retained for this sub-study specifically to allow direct comparison between each country's global cluster assignment and its SSA-specific sub-cluster assignment, rather than because regional diagnostics independently identified three as the best-supported number of SSA sub-groups. Bootstrap resampling (200 replicates) indicated moderate stability for the retained three-sub-group solution (mean Adjusted Rand Index = 0.484, SD = 0.219), lower than the stability observed for the global three-cluster solution and consistent with the reduced statistical power of a 47-country regional subsample. Given this, the specific boundaries of the three SSA sub-clusters reported below should be interpreted with caution; the broader, more robust finding is that SSA does not form a single homogeneous group under this typology, independent of the exact number of sub-groups selected.

With this caveat, the $k=3$ procedure identified three SSA sub-groups (Figure 4). Nascent SSA Economies ($n=20$) showed the lowest digital-infrastructure and firm-level innovation/practice scores within the region. Transitional SSA Economies ($n=12$), including South Africa, Namibia, and Botswana, showed comparatively higher firm-level innovation/practice scores despite digital-infrastructure scores close to the regional average. Adaptive SSA Economies ($n=15$), including Nigeria, Kenya, and Zambia, showed digital-infrastructure and business-environment scores comparable to the Nascent SSA group, but higher firm-level innovation/practice scores than either of the other two SSA sub-groups. The majority of these countries were pooled within a single cluster in the global three-cluster solution, and this regional sub-structure was not distinguished at the global scale, though the precise three-way partition reported here reflects a design choice made for comparability rather than an independently optimal regional solution (Sub-section 3.7).

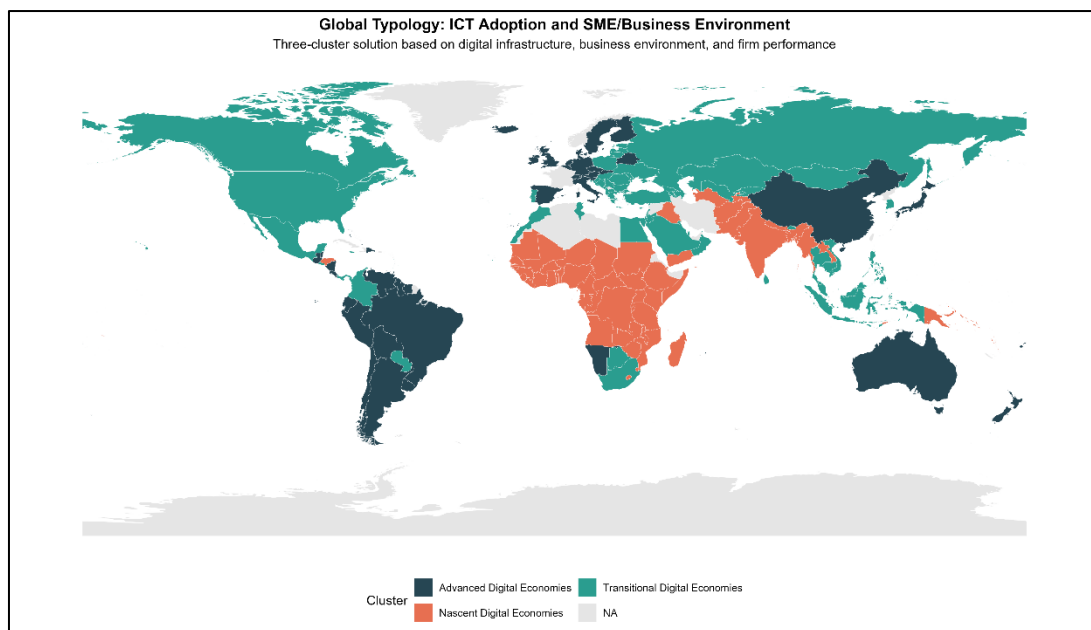


Figure 3. Geographic distribution of the global three-cluster typology

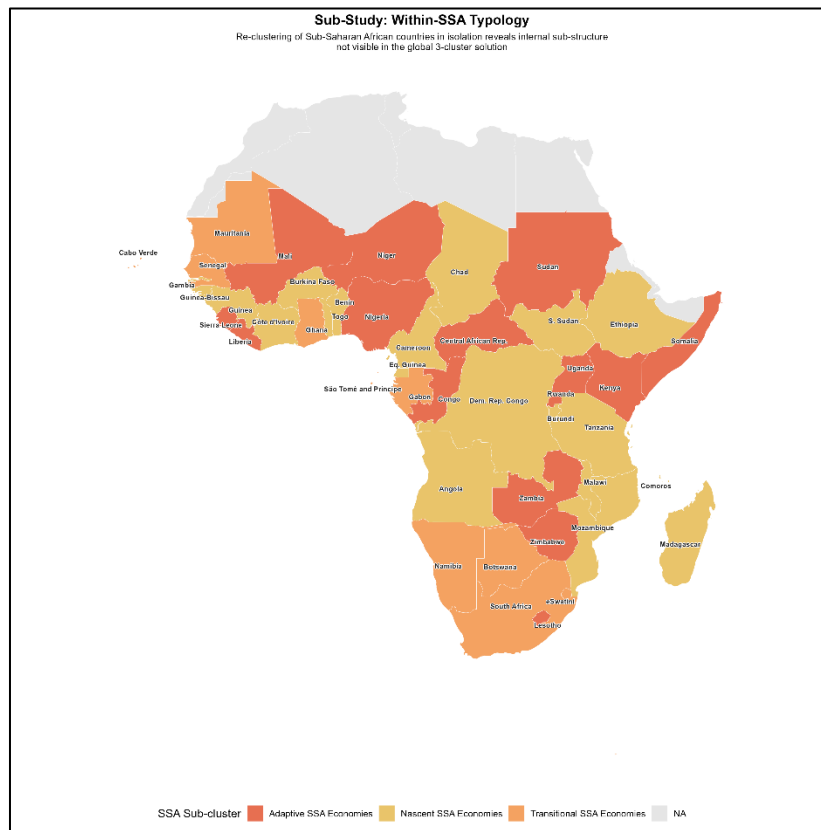


Figure 4. Geographic distribution of the within-Sub-Saharan Africa sub-study typology.

5.8 Post-Clustering Diagnostics and Sensitivity Analysis

PERMANOVA (Euclidean distance, 999 permutations) indicated significant multivariate separation among clusters (pseudo- $F = 118.9$, $R^2 = .576$, $p = .001$), with cluster membership accounting for 57.6% of variance in the four domain component scores. Because PERMANOVA can be sensitive to unequal within-group dispersion rather than centroid separation alone, a dispersion-homogeneity check (Levene's test on each country's distance to its cluster centroid, analogous to PERMDISP) indicated significantly unequal dispersion across clusters (Levene's statistic = 3.98, $p = .020$; mean distance-to-centroid: Cluster 1 = 1.87, Cluster 2 = 1.64, Cluster 3 = 2.00). Part of the PERMANOVA effect size may therefore reflect differences in within-cluster spread rather than centroid location alone, and the $R^2 = .576$ result should be read as evidence of overall multivariate difference among clusters rather than a dispersion-independent test of group centroids. Kruskal-Wallis tests indicated significant differences across clusters for all 15 indicators; because 15 tests were conducted on clustering indicators, raw p -values were corrected using Benjamini-Hochberg FDR ($\alpha = .05$), and all 15 remained significant after correction. Effect sizes (epsilon-squared) ranged from small-to-moderate for several business-environment indicators (e.g., internet-switching difficulty, $\epsilon^2 = .073$) too large for most digital-infrastructure and firm-level innovation indicators (e.g., fixed broadband, $\epsilon^2 = .612$), indicating that cluster membership is a stronger discriminator for some indicators than others. Dunn's post-hoc tests with Benjamini-Hochberg correction were used to identify which specific cluster pairs differed for each indicator (Supplementary Table A1).

Two sensitivity analyses were conducted. First, the 15 countries with the lowest bootstrap own-cluster consensus were excluded and cluster-mean profiles recalculated; means shifted by an average of 0.033 SD across all 15 indicators and three clusters (maximum single-indicator shift = 0.089 SD), indicating the typology is not substantially dependent on the least stable cases. Second, the clustering procedure was independently re-applied to the 47 Sub-Saharan African countries (Sub-section 5.7), identifying three regional sub-groups not distinguished by the global solution; this result indicates that the global three-cluster typology does not capture region-specific structure

within Sub-Saharan Africa, and is reported as a substantive finding of the regional sub-study rather than as evidence of an absence of structure in the region the opposite reading appears in the previous draft and is corrected here, since three coherent sub-groups were in fact identified. Table 3 summarises this evidence.

Table 3. Validation and Sensitivity Summary

Procedures	Result
Cluster-number diagnostics	Elbow and gap statistic supported $k = 3$; silhouette width was highest at $k = 2$
K-means vs. hierarchical clustering	ARI = 0.524
Consensus clustering (PAC), $k = 2-6$	0.406 / 0.362 / 0.385 / 0.338 / 0.314
PERMANOVA	Pseudo-F = 118.9; $R^2 = 0.576$; $p = 0.001$
Kruskal-Wallis, 15 indicators	$p < 0.05$ for all indicators
Sensitivity: exclusion of least stable countries	Mean shift = 0.033 SD; maximum shift = 0.089 SD
Sensitivity: within-SSA re-clustering	Three sub-groups identified within the 47-country SSA sample

6. DISCUSSION OF RESULTS

Based on the indicators of digital infrastructure, business environment and firm-level innovation/practice, this study established three clusters: Nascent Digital Economy, Transitional Digital Economy, and Advanced Digital Economy, in this sample of 178 countries. These are patterns that have been observed in this sample and not a typology in general. The Transitional cluster combined favorable digital-infrastructure and credit indicators with the least favorable firm-level innovation indicators of the three groups an association specific to the indicators used, not a general claim about "finance" or "innovation.". The SSA sub-study identified three sub-groups within a region which was almost entirely in one global cluster, one of which was a mix of weak digital/business indicators, and comparatively strong firm-level indicators. The Transitional pattern is consistent with prior evidence that ICT's association with firm outcomes varies: Karim et al. (2022) found this association differs by a country's ICT development level; Kumari & Singh, (2023) found it varies by income group; Gaglio et al. (2022) found digital technologies were linked to productivity mainly through innovation, not adoption directly. The results support an interpretation of the infrastructure and credit indicators as preceding the capability of firms in this cluster, but it is not possible to determine from cross-sectional data whether this is a sequence of development.

The SSA result is an extension of the Saba and David (2020) study which revealed that SSA is not a single region, but rather several ICT-convergence clubs. The analysis in the previous study focused on ICT convergence only, while this research also reveals the pattern of the combination of ICT, business and firm-level indicators for each sub-group, which was not discussed in the previous study. The Adaptive SSA sub-group (e.g., Nigeria, Kenya, Zambia) is in line with Kuteyi & Winkler, (2022) who describe firms filling infrastructure gaps with mobile alternatives. This is an interpretation and not a tested finding: The clustering is an identification of the pattern, not a test of the adaptation mechanism. The TOE framework (Sub-section 2) used in this study is typically a firm-level framework, and the country-level clusters created from aggregated survey indicators are subject to a real ecological-inference risk, as is not a firm-level TOE mechanism. The contribution is a structural analogy, where the three domains represent the three contexts of TOE: technological, environmental and organizational, and the pattern of the Transitional cluster, favorable technology/environment, weaker organizational indicators, is in line with TOE's premise that these contexts do not necessarily move together. This does not show that TOE's organizational level mechanisms are at play at the country level; it motivates the design of the study.

These patterns are policy considerations that are suggested by the profiles, but have not been tested. In line with Tödtling & Trippel, (2005) who argue that innovation policy should be based on regional capacities, not uniform approaches: Nascent economies: favorable infrastructure/finance indicators, suggesting this as relevant; Transitional economies: favorable infrastructure indicators, but weak firm-capability indicators, suggesting skills and innovation-support as relevant; Advanced economies: favorable across domains, suggesting continued upgrading as relevant. These are associations, not evidence that a specific intervention would be effective in improving outcomes causal evaluation would be required to confirm that. As with SSA-specific implications, sub-regional differentiation is a sensible design consideration, rather than a proven solution. The Transitional cluster relates to SDG 9 (infrastructure, innovation) and SDG 8 (productive employment): The pattern suggests that investment in infrastructure could be linked to less innovation gains when the organizational capability is low.

7. CONCLUSION

The principal component analysis and consensus clustering were used in this research to create a data-driven typology for 178 countries using digital infrastructure, business-environment quality and firm-level innovation/practice indicators. Three clusters were identified (Nascent, Transitional, Advanced Digital Economies) and validated internally and through sensitivity analysis. In addition to the differences in digital access, the typology reveals differences in the combination of digital infrastructure, business conditions and indicators of the firms. A supplemental Sub-Saharan Africa analysis revealed three additional sub-groups, which were not distinguished in the global classification, suggesting significant within-region variation in the region that is frequently classified as a single region. There are two trends. First, there was no consistent relationship between the indicators of digital infrastructure and finance access and the indicators of innovation at the firm level: the Transitional cluster includes the strong indicators of digital infrastructure and finance access, but comparatively weak indicators of innovation at the firm level, suggesting that digital infrastructure and finance access alone do not explain the capacity of firms in this sample to innovate. Second, there is significant variation within the broad regional categories: the Sub-Saharan Africa sub-study reveals that there is considerable variation between countries that are grouped in the same global-cluster, highlighting the importance of country-level analysis, not regional analysis. These are cross sectional correlations, not developmental sequence or causal effects. In addition to empirical results, this research presents a reproducible analytical framework to build multidimensional country typologies based on development indicators.

8. IMPLICATIONS

The results provide a specific TOE insight: technological and organizational conditions can be empirically separated, not only interdependently, but also in a positive way (the Transitional cluster has good infrastructure and credit conditions and poor firm-level innovation conditions). This illustrates that digital development is not only reflected in income, geography, and readiness indicators, but in the combination of these, not in any single indicator. The three global clusters and three sub-clusters in SSA are indicative of different policy considerations, not of estimated effects. The infrastructure and finance investment is weak in nascent economies, in line with the weak indicators in all domains. Transitional economies have strong indicators for infrastructure, but weak indicators for innovation, which is consistent with the skills and innovation-support as a relevant complement. Advanced economies are good in all areas, in line with further upgrading. There are sub-groups that differ in the conditions that are constrained, indicating sub-regional targeting instead of region-wide targeting. Digital-policy monitoring should also monitor the growth of innovation, as well as that of infrastructure, because here, connectivity did not match with good

innovation. The implications (SDG 8 and 9) are associations that are suggested by cluster profiles, not causal estimates.

9. LIMITATIONS AND FUTURE RESEARCH

The study identified two limitations. The cross-sectional design of the analysis does not allow for causal inferences, as the typologies found are empirical associations, and not evidence of causal relationships between digital infrastructure, business environments, and firm performance. Longitudinal analyses would offer more information on the transition between typologies over time in countries. Second, the business environment and firm performance indicators were obtained from the World Bank Enterprise Surveys, which are largely based on non-agricultural, formal enterprises. This means that informal-sector activities and those that could be linked to resilience and innovation in many Sub-Saharan African economies are not fully captured. Future research might include longitudinal data, wider indicators of digital capability, and indicators of informal and platform-based economic activity to gain a deeper understanding of the mechanisms through which digital transformation relates to the performance of firms in various development contexts.

10. REFERENCES

- Bahrini, R., & Qaffas, A. (2019). Impact of Information and Communication Technology on Economic Growth: Evidence from Developing Countries. *Economies*. <https://doi.org/10.3390/economies7010021>
- Basel, S., Gopakumar, K., & Rao, R. (2021). Classification of countries based on development indices by using K-means and grey relational analysis. *GeoJournal*, 87, 3915–3933. <https://doi.org/10.1007/s10708-021-10479-2>
- Chege, S., Wang, D., & Suntu, S. (2019). Impact of information technology innovation on firm performance in Kenya. *Information Technology for Development*, 26, 316–345. <https://doi.org/10.1080/02681102.2019.1573717>
- Díaz-Arancibia, J., Hochstetter-Diez, J., Bustamante-Mora, A., Sepúlveda-Cuevas, S., Albayay, I., & Arango-López, J. (2024). Navigating Digital Transformation and Technology Adoption: A Literature Review from Small and Medium-Sized Enterprises in Developing Countries. *Sustainability*. <https://doi.org/10.3390/su16145946>
- Doshi, R., Kelley, J., & Simmons, B. (2019). The Power of Ranking: The Ease of Doing Business Indicator and Global Regulatory Behavior. *International Organization*, 73, 611–643. <https://doi.org/10.1017/s0020818319000158>
- Ehimuan, B., Ehimuan, B., Anyanwu, A., Olorunsogo, T., Akindote, O., Abrahams, T., & Reis, O. (2024). Digital inclusion initiatives: Bridging the connectivity gap in Africa and the USA – A review. *International Journal of Science and Research Archive*. <https://doi.org/10.30574/ijrsra.2024.11.1.0061>
- Eller, R., Alford, P., Kallmuenzer, A., & Peters, M. (2020). Antecedents, consequences, and challenges of small and medium-sized enterprise digitalization. *Journal of Business Research*, 112, 119–127. <https://doi.org/10.1016/j.jbusres.2020.03.004>
- Fantom, N., & Serajuddin, U. (2016). *The World Bank's Classification of Countries by Income*. 1–52. <https://doi.org/10.1596/1813-9450-7528>
- Gaglio, C., Kraemer-Mbula, E., & Lorenz, E. (2022). The effects of digital transformation on innovation and productivity: Firm-level evidence of South African manufacturing micro

and small enterprises. *Technological Forecasting and Social Change*.
<https://doi.org/10.1016/j.techfore.2022.121785>

Karim, M., Nahar, S., & Demirbağ, M. (2022). Resource-Based Perspective on ICT Use and Firm Performance: A Meta-analysis Investigating the Moderating Role of Cross-Country ICT Development Status. *Technological Forecasting and Social Change*.
<https://doi.org/10.1016/j.techfore.2022.121626>

Khan, T., Abimbola, S., Kyobutungi, C., & Pai, M. (2022). How we classify countries and people—And why it matters. *BMJ Global Health*, 7. <https://doi.org/10.1136/bmjgh-2022-009704>

Kumari, R., & Singh, S. (2023). Impact of ICT Infrastructure, Financial Development, and Trade Openness on Economic Growth: New Evidence from Low- and High-Income Countries. *Journal of the Knowledge Economy*, 1–30. <https://doi.org/10.1007/s13132-023-01332-7>

Kuteyi, D., & Winkler, H. (2022). Logistics Challenges in Sub-Saharan Africa and Opportunities for Digitalization. *Sustainability*. <https://doi.org/10.3390/su14042399>

Mardikyan, S., Yıldız, E., Ordu, M., & Şimşek, B. (2015). Examining the Global Digital Divide: A Cross- Country Analysis. *Communications of The IbIMA*, 2015, 1–10.
<https://doi.org/10.5171/2015.592253>

Miškuřová, M., Kořiková, M., Vařaničová, P., & Kiseřáková, D. (2025). Digitalization and Artificial Intelligence: A Comparative Study of Indices on Digital Competitiveness. *Inf.*, 16, 286. <https://doi.org/10.3390/info16040286>

Myovella, G., Karaçuka, M., & Haucap, J. (2021). Determinants of digitalization and digital divide in Sub-Saharan African economies: A spatial Durbin analysis. *Telecommunications Policy*. <https://doi.org/10.1016/j.telpol.2021.102224>

Ochieng, H. K., & Park, B. (2021). Institutional Quality, Regulatory Environment and Microeconomic Performance: Evidence from Transition and Non-transition Developing Countries. *East Asian Economic Review*. <https://doi.org/10.11644/kiep.eaer.2021.25.3.398>

Pradhan, R., Arvin, M., Nair, M., Hall, J., & Bennett, S. (2022). Institutional development in an information-driven economy: Can ICTs enhance economic growth for low- and lower middle-income countries? *Information Technology for Development*, 28, 468–487.
<https://doi.org/10.1080/02681102.2022.2051417>

Richmond, K., & Triplett, R. (2018). ICT and income inequality: A cross-national perspective. *International Review of Applied Economics*, 32, 195–214.
<https://doi.org/10.1080/02692171.2017.1338677>

Robinson, L., Schulz, J., Blank, G., Ragnedda, M., Ono, H., Hogan, B., Mesch, G., Cotten, S., Kretchmer, S., Hale, T., Yan, P., Wellman, B., Harper, M.-G., Quan-Haase, A., Dunn, H., Casilli, A., Tubaro, P., Carveth, R., Chen, W., Drabowicz, T. (2020). Digital inequalities 2.0: Legacy inequalities in the information age. *First Monday*, 25.
<https://doi.org/10.5210/fm.v25i7.10842>

- Rogge, N., & Archer, G. (2021). Measuring and analyzing country change in establishing ease of doing business using a revised version of World Bank's ease of doing business index. *Eur. J. Oper. Res.*, 290, 373–385. <https://doi.org/10.1016/j.ejor.2020.07.065>
- Saba, C., & David, O. (2020). Convergence patterns in global ICT: Fresh insights from a club clustering algorithm. *Telecommunications Policy*, 44, 102010. <https://doi.org/10.1016/j.telpol.2020.102010>
- Tödtling, F., & Trippl, M. (2005). One size fits all?: Towards a differentiated regional innovation policy approach. *Research Policy*, 34, 1203–1219. <https://doi.org/10.1016/j.respol.2005.01.018>
- Tornatzky, L., Fleischer, M., & Chakrabarti, A. (1991). *Processes of technological innovation*. <https://doi.org/10.1007/bf02371446>
- Ullah, B. (2020). Financial constraints, corruption, and SME growth in transition economies. *The Quarterly Review of Economics and Finance*, 75, 120–132. <https://doi.org/10.1016/j.qref.2019.05.009>