International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(11s) 2026 pp. 968-979
https://doi.org/10.70917/ijcisim-2026-3827
Article

Digital Connectivity and GovTech Maturity Across 162 Economies:Development-Adjusted Analysis and Interpretable Machine Learning for Public Information Systems

Juan Moshe M. Duyan

Kalinga State University, Tabuk City, Kalinga, Philippines
Corresponding author: jmduyan@ksu.edu.ph

Abstract: Digital government operates within a connectivity environment that can facilitate or hinder the use of public information systems. This study examines the cross-national association between digital connectivity and GovTech maturity while explicitly accounting for income and regional structure. A cross-sectional secondary-data design integrates the International Telecommunication Union ICT Development Index (IDI) 2025, the World Bank GovTech Maturity Index (GTMI) 2025, and United Nations E-Government Survey 2024 indicators. After ISO-3 matching, 162 economies had valid GTMI observations. Pearson correlations assessed linear relationships and Spearman correlations assessed monotonic rank relationships. HC3-robust regression controlled for World Bank income group and ITU region. Ridge regression and random forest were evaluated with 20 repeated five-fold out-of-fold runs and a leave-one-ITU-region-out robustness check. IDI was associated with GTMI ($r = 0.601$, $p < 0.001$); after income and region adjustment, a 10-point higher IDI remained associated with a 0.108 higher GTMI ($p < 0.001$), with an adjusted partial correlation of $r = 0.404$. Cross-source convergence was strong for public-service delivery versus the UN Online Service Index ($r = 0.813$) and digital citizen engagement versus the UN E-Participation Index ($r = 0.819$). Random forest achieved modestly higher repeated out-of-fold R^2 than ridge (0.367 vs. 0.294; mean $\Delta R^2 = 0.073$, 95% CI 0.057–0.090), and region-grouped validation yielded $R^2 = 0.334$. Held-out permutation analysis and grouped sensitivity analysis identified affordability as the strongest model-level signal, although correlated predictors limit individual attribution. The findings show that connectivity remains meaningfully associated with GovTech maturity after broad development controls, but accounts for only part of cross-national variation. The results support affordability-aware, mobile-first, low-bandwidth, interoperable, and citizen-centered public information-system design.

Keywords: digital divide; GovTech; digital government; computer information systems; ICT connectivity; interpretable machine learning; public-service delivery; e-participation

1. Introduction

Digital government has evolved from publishing information on government websites toward integrated public information systems that support digital identity, payments, registries, case management, data exchange, participation, and end-to-end public services. This transition makes digital inclusion an information-systems problem as well as a social-policy concern. A technically mature service may still generate uneven public value when users face expensive data, weak household connectivity, unreliable networks, or service interfaces that assume continuous high-bandwidth access. The World Bank GovTech framework reflects this wider transformation through core government systems, public-service delivery, digital citizen engagement, and GovTech enablers [1].

Digital-divide research has long argued that inequality extends beyond device ownership to differences in access quality, skills, effective use, and outcomes [4]. Cross-country studies also show that national income and development conditions are strongly related to ICT disparities [6], which means that any observed connectivity–

digital-government relationship may partly reflect broader socioeconomic development. E-government research has similarly found that telecommunications and institutional conditions are associated with online services and e-participation [5]. These findings make confounding by development level a central methodological issue rather than a secondary limitation.

At the same time, contemporary public-sector digital maturity is multidimensional. Recent work on digital service transformation emphasizes not only technology and connectivity but also organizational processes, capabilities, governance, service design, and public value [7]. This broader view is consistent with the GTMI architecture, in which service delivery and citizen engagement sit alongside core systems and enabling institutions. The relevant research question is therefore not whether connectivity alone determines GovTech maturity, but whether connectivity remains meaningfully associated with GovTech outcomes after accounting for major development differences, and whether specific connectivity characteristics contain reproducible predictive information.

This study contributes in three ways. First, it integrates current ITU, World Bank, and UN datasets to examine the connectivity–GovTech relationship across 162 economies. Second, it tests robustness to income group and geographic region instead of relying only on bivariate correlation. Third, it combines repeated and region-grouped predictive validation with permutation-based interpretation, explicitly addressing predictor correlation and geographic dependence. This positions the study within computer information systems, data analytics, information governance, and public-sector digital transformation rather than treating the digital divide solely as a local access problem.

The study addresses five research questions:

RQ1. How strongly is national ICT connectivity associated with GovTech maturity and its major information-system dimensions?

RQ2. Does the IDI–GTMI association remain meaningful after accounting for income group and world region?

RQ3. Do connectivity indicators provide stable predictive information for GTMI under repeated and region-grouped validation?

RQ4. Which connectivity features or feature groups contribute most to the predictive model, and how should correlated predictors affect interpretation?

RQ5. How does the Philippines compare with the ITU Asia-Pacific region and the merged international sample?

2. Conceptual and Information-Systems Framework

2.1 Digital divide as a layered access environment

The digital divide is layered rather than binary. Coverage and device access create the technical possibility of connection, while affordability, reliability, household access, ownership, and effective use determine whether that connection becomes meaningful. The ITU IDI operationalizes universal and meaningful connectivity through comparable indicators covering internet use, household internet access, mobile broadband, network coverage, traffic intensity, affordability, and mobile ownership [2]. These conditions characterize the environment in which citizens interact with digital public systems.

2.2 GovTech as multidimensional public information-systems maturity

GovTech maturity extends beyond the presence of websites or portals. The GTMI evaluates core government systems, public-service delivery, digital citizen engagement, and GovTech enablers [1]. Public-sector maturity models similarly conceptualize digital transformation as a combination of technology, process capability, organization, governance, and continuous service improvement [7]. From an information-systems perspective, mature GovTech therefore reflects architecture, integration, process digitization, user-facing capability, information governance, organizational readiness, and the ability to create public value.

2.3 Analytical proposition and confounding structure

The analytical proposition is deliberately associational: stronger connectivity should be positively related to GovTech maturity because connectivity expands the practical operating environment for digital services, but the relationship is expected to be incomplete because institutional and organizational factors also shape outcomes. Income and regional context are potential confounders because they are related to both connectivity and public-sector digital capacity [6]. The analysis therefore estimates the unadjusted association, then evaluates whether the relationship persists after broad controls for income group and ITU region. Predictive models are treated as complementary tools for pattern detection rather than causal estimators.

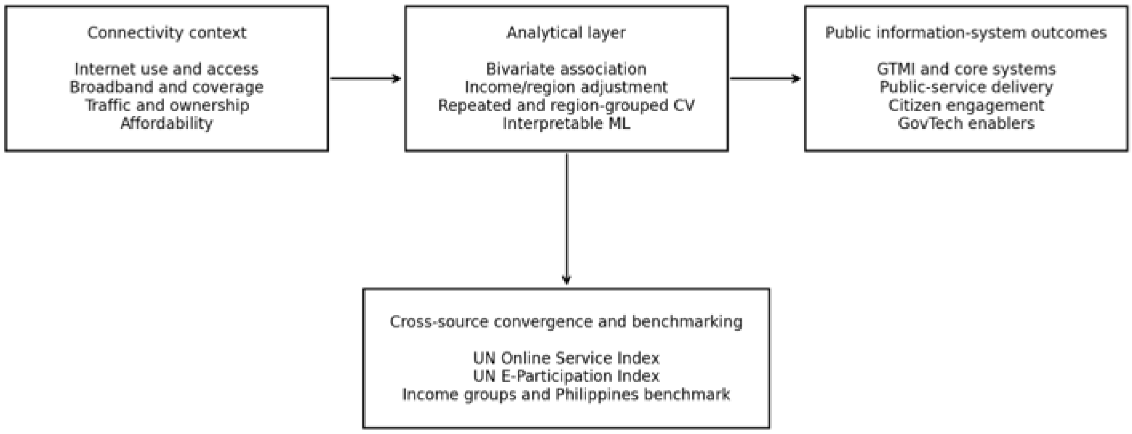

Figure 1. Revised analytical framework linking connectivity conditions, development-adjusted and predictive analysis, GovTech outcomes, and cross-source benchmarking.

3. Materials and Methods

3.1 Research design and data sources

The study uses a cross-sectional, country-level secondary-data design. Official datasets were downloaded and integrated in July 2026. The World Bank GovTech Dataset (December 2025) provides the 2025 GTMI and four focus-area indices. The ITU ICT Development Index 2025 provides comparable connectivity indicators for 164 economies; most underlying IDI indicators refer to 2023 conditions. UN E-Government Survey 2024 measures were used for cross-source convergence and benchmarking rather than being described as model validation [1]–[3], [10].

Source	Reference period	Measures used	Analytical role
ITU ICT Development Index 2025	Mostly 2023 indicators; 2025 release	IDI and nine normalized connectivity components	Connectivity exposure and ML predictors
World Bank GTMI 2025	2025	GTMI, CGSI, PSDI, DCEI, GTEI	Primary public information-system outcomes
UN E-Government Survey	2024	EGDI, Online Service Index, E-Participation Index	Cross-source convergence and benchmarking

Table 1. Secondary data sources and analytical roles.

3.2 Sample construction, missingness, and temporal alignment

Datasets were merged using ISO alpha-3 economy codes. The ITU and World Bank files yielded 163 overlapping economies. Nicaragua lacked a numeric 2025 GTMI value and was excluded from primary GTMI analyses, leaving N = 162. UN OSI and EPI values were unavailable for Hong Kong SAR, China and Macao SAR, China, producing N = 161 for IDI–UN correlations and N = 160 for cross-source pairs involving GTMI dimensions because Nicaragua lacked GTMI. Among the 162 GTMI observations, missing values across the nine ML predictors were: internet use 4, household internet 8, mobile broadband 0, network coverage 0, mobile traffic 15, fixed traffic 40, mobile affordability 1, fixed affordability 4, and mobile ownership 6. Missing predictor values were imputed only within model training folds using the training-set median.

The sources are not perfectly synchronized in time: the IDI 2025 release relies mainly on 2023 connectivity indicators, the UN survey measures 2024 e-government conditions, and GTMI reflects 2025. The design is therefore interpreted as a cross-source structural comparison rather than a same-year causal model. No individual-level or personally identifiable data were used.

3.3 Variables

The primary outcome was GTMI (0–1). Secondary World Bank outcomes were the Core Government Systems Index (CGSI), Public Service Delivery Index (PSDI), Digital Citizen Engagement Index (DCEI), and GovTech Enablers Index (GTEI). The primary explanatory construct was the ITU ICT Development Index (IDI; 0–100). The nine machine-learning features were normalized ITU scores for internet use, household internet access, mobile-broadband subscriptions, 3G/4G coverage, mobile internet traffic, fixed internet traffic, mobile-service affordability, fixed-broadband affordability, and mobile-phone ownership. World Bank income-group classification embedded in the ITU dataset and the six ITU regions were used as broad development and geographic controls.

3.4 Statistical association and development-adjusted robustness

Pearson correlation quantified linear association, while Spearman correlation quantified monotonic rank association. Cross-source convergent evidence was assessed by correlating conceptually corresponding measures from different sources, particularly PSDI with the UN Online Service Index and DCEI with the UN E-Participation Index. To address development confounding, ordinary least-squares models regressed GTMI on IDI expressed in 10-point units. Model 1 was unadjusted; Model 2 added income-group indicators; Model 3 added both income-group and ITU-region indicators. HC3 heteroskedasticity-robust standard errors were used. An income-and-region-adjusted partial correlation was also obtained by correlating residuals from separate IDI and GTMI regressions on the same categorical controls. These analyses estimate conditional association, not causal effect.

3.5 Predictive modeling, validation, and interpretability

Ridge regression served as a regularized linear baseline and random-forest regression as a flexible nonlinear comparator. Ridge used $\alpha = 1.0$ with standardized features after median imputation. Random forest used 500 trees, minimum samples per leaf = 3, maximum depth = unrestricted, and max_features = 1.0 after median imputation. Hyperparameters were fixed a priori; no hyperparameter search was performed, reducing the risk of tuning to a small cross-national sample.

Predictive stability was evaluated with 20 repetitions of shuffled five-fold cross-validation. Within each repetition, every economy was held out exactly once; R^2, MAE, and RMSE were calculated from one pooled set of out-of-fold predictions across all 162 economies, and the reported values are the mean and standard deviation across the 20 repetitions. The paired difference $\Delta R^2 = R^2RF - R^2Ridge$ was summarized across repetitions. To assess sensitivity to geographic dependence, a six-fold GroupKFold analysis held out one entire ITU region per fold and calculated pooled out-of-region predictions. This responds to evidence that ordinary random CV can be optimistic when observations share geographic structure [9].

Random-forest interpretation used held-out permutation importance in a separate five-fold run. All nine predictors are displayed; error bars represent the standard deviation of fold-level mean importance across the five held-out folds. Because correlated predictors can redistribute or suppress individual permutation importance [8], the study also reports a predictor correlation matrix and grouped permutation sensitivity in which related variables were jointly permuted as three conceptual sets: access/ownership, network/traffic, and affordability. A one-dimensional partial-dependence profile for mobile affordability was generated from the full fitted forest only as a descriptive visualization of model behavior; it is not a causal dose-response estimate.

3.6 Income and Philippines benchmarking

Income-group means were calculated for low-, lower-middle-, upper-middle-, and high-income economies. “Asia-Pacific” follows the ITU regional classification. The primary Asia-Pacific GTMI sample contains 29 economies; UN-based regional medians use 27 because Hong Kong SAR, China and Macao SAR, China lack OSI/EPI values in the merged file. Philippines percentile positions use the empirical cumulative distribution function within the nonmissing merged sample for each metric, with ties included at or below the Philippines score. They are therefore labeled merged-sample percentiles rather than world rankings.

4. Results

4.1 Descriptive pattern

Across the 163-economy merged file, mean IDI was 77.6 (median 84.7). Among the 162 valid GTMI cases, mean GTMI was 0.646 (median 0.715). Mean PSDI (0.720) exceeded mean DCEI (0.531), indicating that citizen-facing service delivery was, on average, more mature than digital participation mechanisms in this sample.

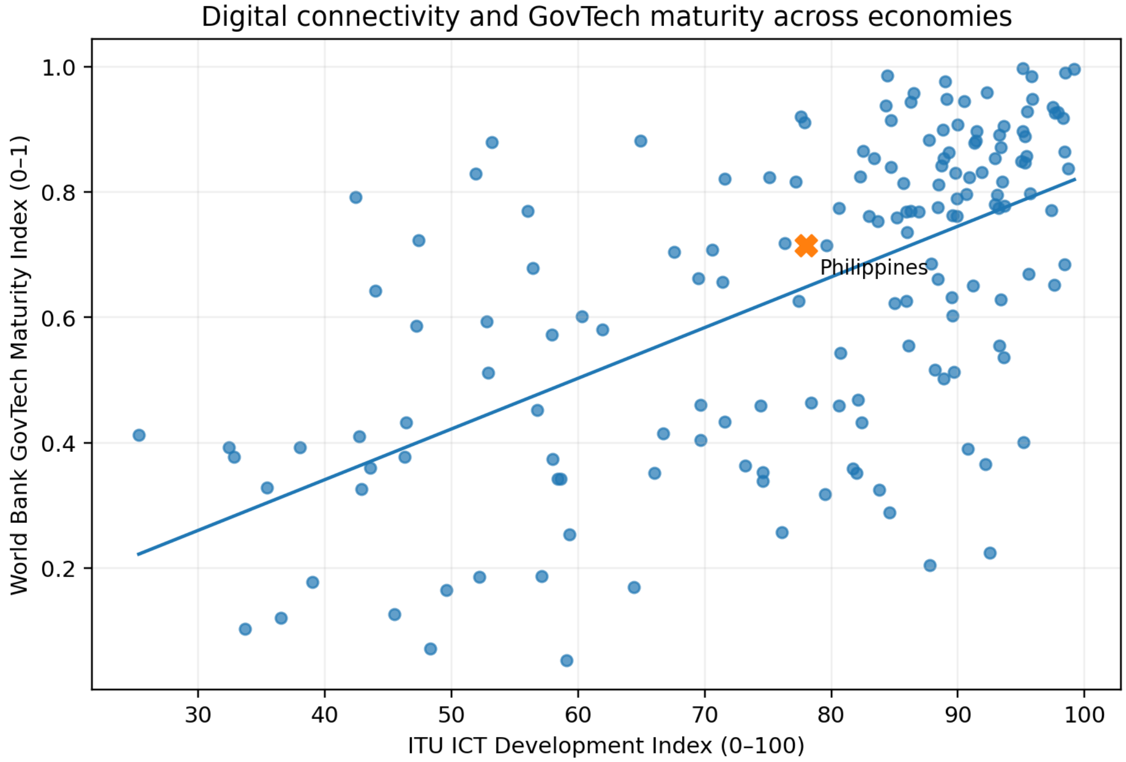

Figure 2. IDI and GTMI across 162 economies. The line is an ordinary least-squares trend for visualization only; the Philippines is marked separately.

4.2 Bivariate relationships and cross-source convergent evidence

Relationship	N	Pearson r	p	Spearman ρ	p
IDI–GTMI	162	0.601	<0.001	0.601	<0.001
IDI–CGSI	162	0.525	<0.001	0.502	<0.001
IDI–PSDI	162	0.593	<0.001	0.570	<0.001
IDI–DCEI	162	0.579	<0.001	0.595	<0.001
IDI–GTEI	162	0.574	<0.001	0.594	<0.001
IDI–UN OSI	161	0.702	<0.001	0.739	<0.001
IDI–UN EPI	161	0.641	<0.001	0.690	<0.001
PSDI–UN OSI	160	0.813	<0.001	0.807	<0.001
DCEI–UN EPI	160	0.819	<0.001	0.816	<0.001
GTMI–UN EGDI	160	0.791	<0.001	0.777	<0.001

Table 2. Connectivity associations and cross-source convergent evidence.

IDI was positively associated with overall GTMI ($r = 0.601$), public-service delivery ($r = 0.593$), digital citizen engagement ($r = 0.579$), and GovTech enablers ($r = 0.574$), with all $p < 0.001$. The relationship with the UN Online Service Index was $r = 0.702$, and the relationship with UN E-Participation was $r = 0.641$. These are associations and do not establish that connectivity is a necessary or sufficient condition for GovTech maturity.

The separate cross-source tests showed strong convergence between conceptually corresponding digital-government measures: PSDI correlated with the UN Online Service Index at $r = 0.813$, DCEI correlated with the UN E-Participation Index at $r = 0.819$, and GTMI correlated with the UN E-Government Development Index at $r = 0.791$ (all $p < 0.001$). These results support cross-source comparability of the outcome constructs more directly than correlating IDI with multiple outcomes.

4.3 Development-adjusted association

Model	IDI coefficient per 10 points	HC3 SE	p	R^2	Adjusted R^2
Unadjusted	0.081	0.008	<0.001	0.361	0.357
Income-adjusted	0.095	0.018	<0.001	0.368	0.348
Income + region adjusted	0.108	0.018	<0.001	0.406	0.366

Table 3. Robustness of the IDI–GTMI association to income-group and regional adjustment.

The unadjusted model estimated a 0.081 increase in GTMI for each 10-point higher IDI. Rather than weakening to zero after adjustment, the coefficient was 0.095 with income-group controls and 0.108 after adding ITU-region controls (HC3 $p < 0.001$ in both adjusted models). The residual-based partial correlation controlling for both income group and region was $r = 0.404$ ($p < 0.001$). Thus, the original $r = 0.601$ relationship is partly embedded in broad development structure, but a moderate conditional association remains. The result should still not be interpreted causally because income group and region do not exhaust institutional, historical, fiscal, or governance confounding.

4.4 Predictive performance and geographic robustness

Validation scheme	Model	R^2	MAE	RMSE
Repeated 5-fold (20 repeats)	Ridge regression	0.294 ± 0.040	0.164 ± 0.004	0.207 ± 0.006
Repeated 5-fold (20 repeats)	Random forest	0.367 ± 0.028	0.153 ± 0.004	0.196 ± 0.004
Leave-one-ITU-region-out	Ridge regression	0.216	0.176	0.218
Leave-one-ITU-region-out	Random forest	0.334	0.154	0.201

Table 4. Predictive performance. Repeated-CV values are mean ± SD of pooled out-of-fold metrics across 20 repetitions; region-grouped values are pooled across six held-out ITU regions.

Across 20 repeated five-fold runs, random forest achieved mean pooled out-of-fold $R^2 = 0.367$ compared with 0.294 for ridge regression. The paired mean ΔR^2 was 0.073 (95% CI 0.057–0.090), and random forest exceeded ridge in 20 of 20 repetitions. This supports a modest and stable predictive advantage, but it does not by itself prove that the underlying relationship is nonlinear. Under leave-one-ITU-region-out validation, the random forest retained pooled $R^2 = 0.334$ (MAE = 0.154; RMSE = 0.201), close to its repeated-random-CV performance, whereas ridge yielded $R^2 = 0.216$. Performance varied across individual held-out regions, so the regional result is best treated as a robustness check rather than evidence of universal transferability.

4.5 Feature dependence and interpretable machine-learning results

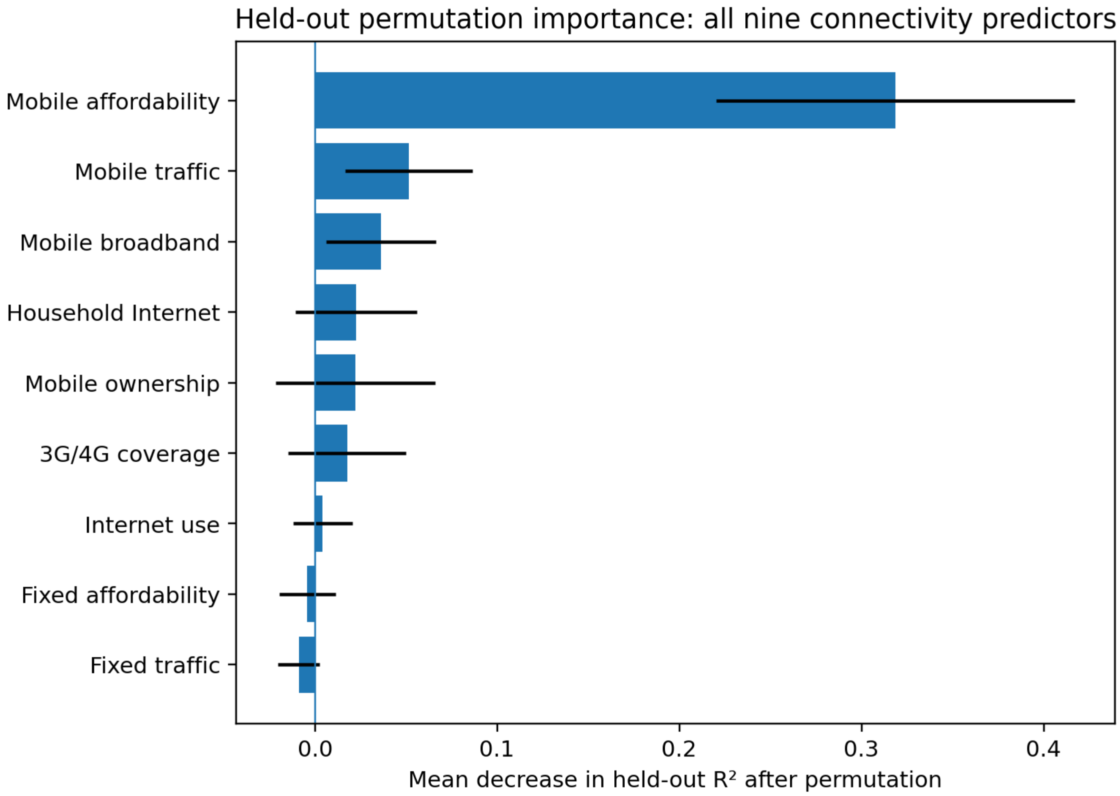

Figure 3. Held-out permutation importance for all nine random-forest predictors. Error bars are the standard deviation of fold-level mean importance across five held-out folds; negative values indicate that permutation did not reduce held-out R^2 on average.

Mobile-service affordability had the largest individual held-out permutation importance (mean $\Delta R^2 = 0.319$, SD = 0.098). Mobile traffic and mobile broadband were distant second and third signals. Fixed affordability and fixed traffic had slightly negative mean importance, indicating no stable incremental contribution under individual permutation in this correlated feature set. The result therefore differs from a causal ranking of policy levers.

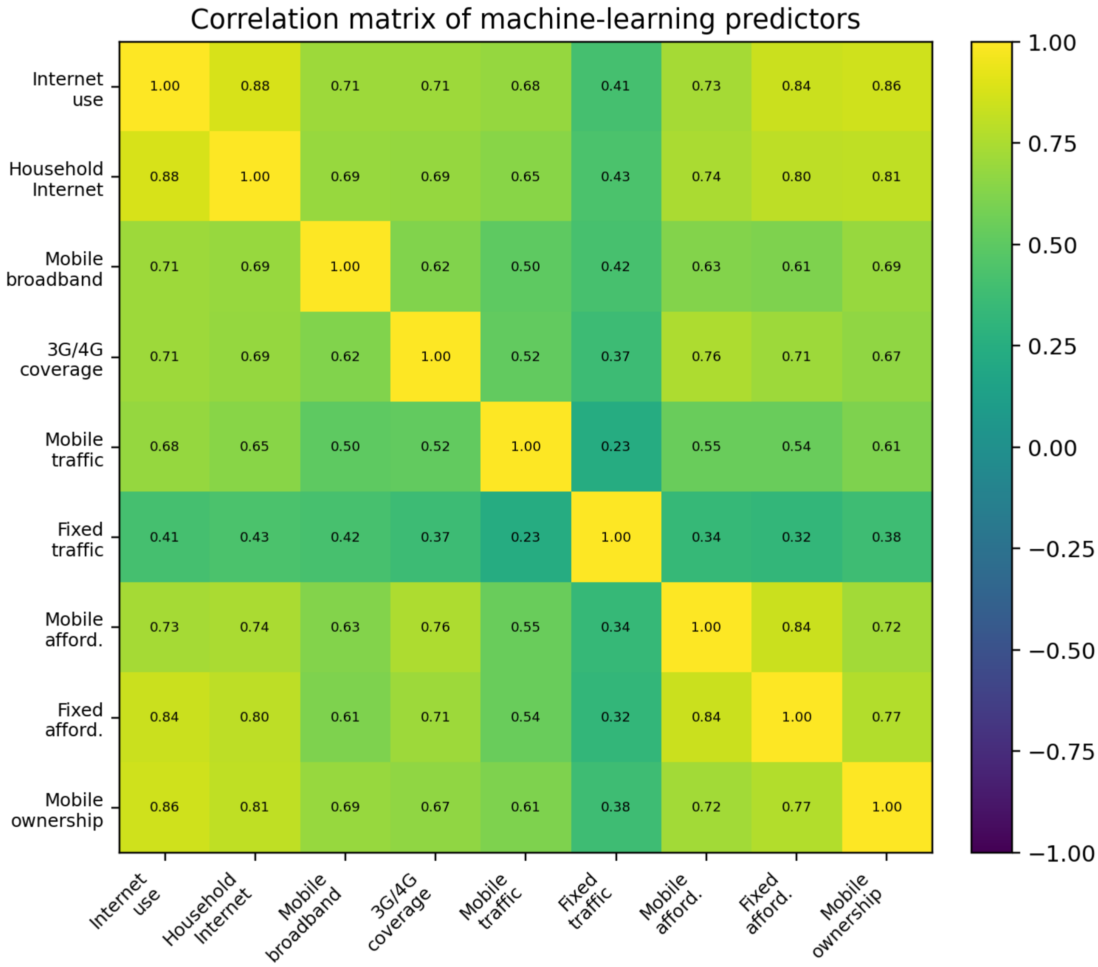

Figure 4. Pearson correlation matrix of the nine machine-learning predictors. Several access and affordability indicators are strongly correlated.

Strong predictor dependence was evident. Pairwise correlations at or above |0.80| were: Internet use–Household Internet: r = 0.881; Internet use–Mobile ownership: r = 0.859; Mobile affordability–Fixed affordability: r = 0.842; Internet use–Fixed affordability: r = 0.838; Household Internet–Mobile ownership: r = 0.812. This confirms the reviewer-relevant concern that individual permutation importance can be redistributed among correlated variables [8].

Jointly permuted feature group	**Mean held-out ΔR^2**	**SD across folds**
Affordability	0.370	0.095
Network / traffic	0.126	0.084
Access / ownership	0.043	0.051

Table 5. Grouped permutation sensitivity analysis. Related predictors were permuted jointly within each held-out fold.

Grouped permutation produced the same broad substantive pattern without requiring individual attribution: jointly permuting the two affordability indicators reduced held-out R^2 by 0.370 on average, compared with 0.126 for network/traffic variables and 0.043 for access/ownership variables. This supports interpreting affordability as a robust model-level signal while avoiding the claim that one affordability indicator independently causes GovTech maturity.

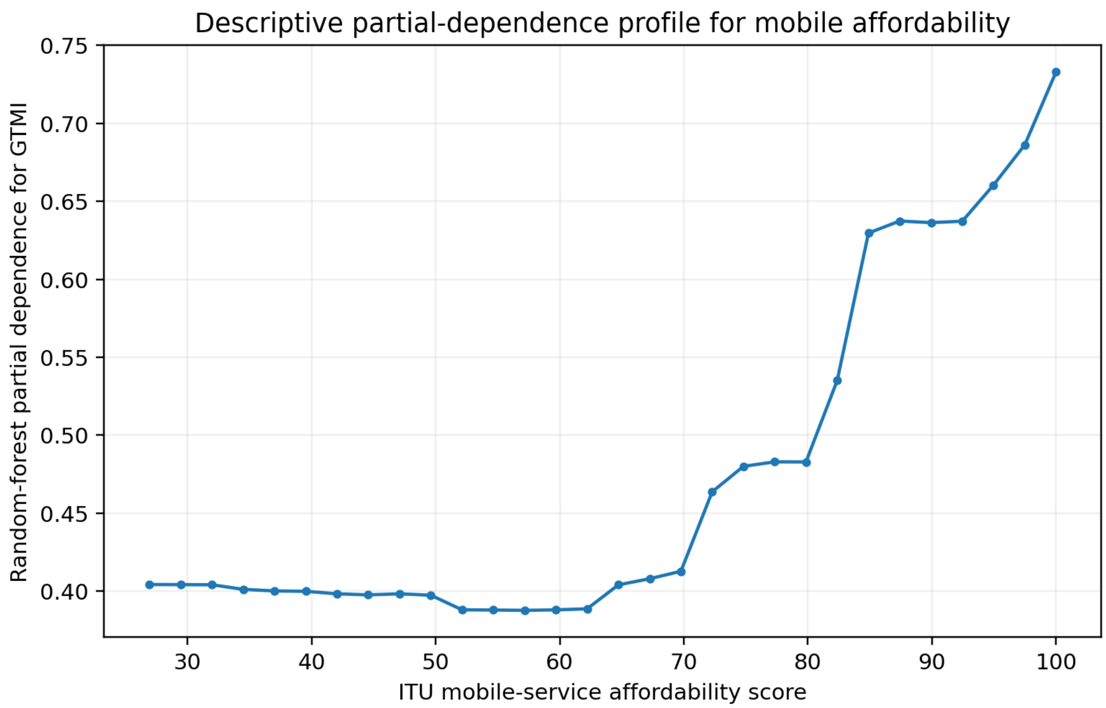

Figure 5. Descriptive partial-dependence profile for mobile-service affordability in the full random forest. The curve describes fitted model behavior under marginalization and should not be read as a causal effect.

The partial-dependence profile was relatively flat across lower and middle affordability scores and rose more sharply at higher scores. Because correlated predictors can make marginal partial-dependence profiles sensitive to combinations that are uncommon in the observed data, this pattern is interpreted only as evidence of a possible threshold-like model relationship requiring confirmation in longitudinal or quasi-experimental work.

4.6 Income-group pattern

Income group	N	IDI	GTMI	PSDI	DCEI
Low-income	16	40.8	0.373	0.439	0.220
Lower-middle-income	39	66.4	0.564	0.637	0.402
Upper-middle-income	45	81.1	0.661	0.738	0.558
High-income	61	92.5	0.764	0.836	0.683

Table 6. Mean connectivity and GovTech indicators by income group.

The descriptive income gradient remained substantial: low-income economies averaged IDI = 40.8 and GTMI = 0.373, compared with IDI = 92.5 and GTMI = 0.764 among high-income economies. This pattern reinforces why adjustment for general development was necessary. Importantly, the adjusted analysis in Section 4.3 shows that the IDI–GTMI relationship is not reducible to these broad income categories alone.

4.7 Philippines benchmark

Metric	Philippines	ITU Asia-Pacific median	Merged-sample median	Merged-sample percentile
IDI	78.0	85.7	84.7	38th
GTMI	0.715	0.679	0.715	50th

CGSI	0.749	0.749	0.723	55th
PSDI	0.801	0.784	0.788	53rd
DCEI	0.522	0.510	0.564	48th
GTEI	0.789	0.703	0.746	56th
OSI	0.805	0.704	0.655	71st
EPI	0.726	0.493	0.534	70th

Table 7. Philippines benchmark. Asia-Pacific follows ITU classification; percentile denominators vary by metric because of missing values.

The Philippines had IDI = 78.0, below the ITU Asia-Pacific median of 85.7 and the merged-sample median of 84.7. Its GTMI of 0.715 was approximately the merged-sample median and above the regional median. The country occupied a relatively stronger position on the UN Online Service Index (0.805; 71st merged-sample percentile) and E-Participation Index (0.726; 70th percentile) than on IDI (38th percentile). The defensible interpretation is therefore positional rather than mechanistic: within this merged sample, the Philippines ranks relatively better on online-service and e-participation indicators than on underlying connectivity.

5. Discussion

5.1 Connectivity remains associated after broad development adjustment

The principal result is narrower, but stronger, than the original manuscript's "necessary but incomplete" claim. Connectivity is moderately associated with GovTech maturity at the bivariate level and remains meaningfully associated after income-group and regional adjustment. The fully adjusted coefficient of 0.108 GTMI points per 10 IDI points and partial r of 0.404 indicate that broad development differences do not fully explain the relationship. At the same time, these estimates do not establish necessity, sufficiency, or causal direction. They are consistent with connectivity functioning as an important enabling condition within a wider institutional and organizational system.

This framing aligns with cross-country digital-divide evidence showing that ICT inequalities are intertwined with national development [6] and with public-sector maturity research that treats digital transformation as multidimensional [7]. The contribution of the present study is therefore not to isolate infrastructure as the determinant of GovTech, but to quantify how much connectivity covaries with contemporary public information-system maturity under explicit development controls and to test whether that relationship carries predictive information across validation schemes.

5.2 Predictive structure and affordability

Random forest showed a modest but stable predictive advantage over ridge in repeated cross-validation, and the advantage persisted under region-grouped validation. The correct interpretation is that nonlinear or interaction structure may be present; the comparison does not prove nonlinearity because model classes differ in more than functional form. The repeated out-of-fold analysis substantially reduces dependence on a single arbitrary split and resolves the earlier ambiguity between mean fold R^2 and pooled out-of-fold R^2.

Affordability remained the most important model-level signal under both individual and grouped permutation. This matters for information-systems management because effective access depends not only on network availability but on whether citizens can repeatedly afford data-intensive transactions. Nevertheless, the high correlation among access and affordability predictors means that individual feature rankings are not unique decompositions of explanatory power [8]. The grouped sensitivity result is therefore more defensible than claiming that mobile affordability alone is the dominant causal driver.

5.3 Information-systems management implications

The practical implication is to design digital public services for meaningful rather than nominal connectivity. Service-level requirements should include mobile responsiveness, low-bandwidth performance, recovery from interrupted sessions, lightweight authentication steps, compressed documents, accessibility, and assisted channels.

Shared digital infrastructure and interoperability can reduce repeated registration, document upload, and cross-agency navigation, thereby lowering the connectivity burden placed on users.

The strong cross-source convergence between PSDI and the UN Online Service Index, and between DCEI and UN E-Participation, also suggests that service delivery and citizen engagement should be managed as distinct but related information-system capabilities. Citizen feedback, consultation, grievance, and participatory systems should not be treated as optional website features. They require governance, process ownership, usability engineering, privacy and security controls, and organizational capacity alongside connectivity [1], [7], [11].

5.4 Implications for the Philippines

The Philippines benchmark should be read as a relative profile rather than evidence of “conversion efficiency.” The country’s online-service and e-participation positions are stronger than its connectivity position in the merged sample. This creates a practical inclusion question: nationally mature platforms may still impose disproportionate transaction costs on users facing weak or expensive connectivity. Philippine agencies can therefore complement system-availability metrics with completion success under constrained bandwidth, mobile data consumption per transaction, device compatibility, language accessibility, abandonment rates, and availability of assisted or offline-compatible pathways.

5.5 Limitations

Several limitations remain. First, the design is cross-sectional and cannot establish causal direction, necessity, or sufficiency. Second, the income-group and region controls are broad; unmeasured fiscal capacity, institutional quality, education, regulatory effectiveness, urbanization, and political factors may confound the relationship. Third, the source years are asynchronous: IDI 2025 largely reflects 2023 connectivity, UN outcomes reflect 2024, and GTMI reflects 2025. This can introduce temporal mismatch even if infrastructure changes slowly relative to annual service indicators. Fourth, composite indices reflect methodological and reporting choices, and some GTMI observations include government-reported inputs.

Fifth, country observations may share regional dependence. The leave-one-region-out analysis reduces but does not eliminate this concern because six broad ITU regions contain heterogeneous countries and the CIS group is small. Sixth, missingness is concentrated in fixed-traffic data (40 of 162 observations), making median imputation consequential for that feature. Seventh, correlated predictors complicate individual feature attribution even with grouped sensitivity analysis; permutation importance and partial dependence remain model-interpretation tools rather than causal explanations [8]. Finally, national averages conceal within-country divides by geography, income, gender, disability, language, and rurality. Future research should use multi-year panels, continuous GDP/GNI and institutional controls, subnational connectivity data, and transaction-level service outcomes where available.

6. Conclusions

Across 162 economies, ICT connectivity is moderately associated with GovTech maturity and related digital-government outcomes. The association remains statistically and substantively meaningful after broad controls for income group and ITU region, but it accounts for only part of the observed cross-national variation. Strong cross-source convergence between World Bank service/engagement measures and corresponding UN indicators supports the comparability of the digital-government outcome constructs.

Repeated and region-grouped validation further shows that connectivity indicators contain reproducible predictive information for GTMI. Random forest performs modestly better than ridge, suggesting that nonlinear relationships may be present without establishing them conclusively. Affordability is the strongest model-level interpretive signal under both individual and grouped permutation, but correlated predictors prevent a causal ranking of specific connectivity variables. For public information-systems management, the evidence supports affordability-aware, mobile-first, low-bandwidth, interoperable, and citizen-centered architectures alongside institutional, governance, security, and organizational investments. For the Philippines, the main finding is positional: online-service and e-participation indicators are relatively stronger than connectivity indicators within the merged sample, highlighting the importance of ensuring that mature digital services remain usable under uneven access conditions.

Data Availability Statement

The study uses publicly available aggregate secondary data. World Bank GovTech Dataset (December 2025): https://datacatalog.worldbank.org/search/dataset/0037889/govtech-dataset. ITU ICT Development Index 2025: https://www.itu.int/itu-d/reports/statistics/idi2025/. UN E-Government Survey 2024:

https://desapublications.un.org/publications/un-e-government-survey-2024. The merged analytical dataset and reproducible analysis code should be deposited in an institutional or open repository at submission.

Ethics Statement

This secondary analysis uses country-level public datasets and contains no identifiable individual-level data.

Use of Artificial Intelligence Statement

Generative artificial intelligence (OpenAI ChatGPT) was used to assist with manuscript restructuring, language refinement, drafting of selected passages, and the development and review of analytical code. The author independently reviewed the source data, verified the statistical procedures and outputs, checked the references, and critically evaluated all interpretations and conclusions. The AI tool did not make autonomous research decisions and is not listed as an author. The author takes full responsibility for the accuracy, integrity, and final content of the manuscript.

References

1. J. R. Vasconcelos, C. Dener, H. Nii-Aponsah, C. V. Blanco, A. J. Immonen, P. Nordlund, F. Siregar, and Y. Song, “GovTech Maturity Index 2025: Tracking Public Sector Digital Transformation Worldwide,” World Bank, Washington, DC, 2025.
2. International Telecommunication Union, Measuring Digital Development: ICT Development Index 2025. Geneva: ITU, 2025.
3. United Nations Department of Economic and Social Affairs, United Nations E-Government Survey 2024: Accelerating Digital Transformation for Sustainable Development, with the Addendum on Artificial Intelligence. New York: United Nations, 2024.
4. J. van Dijk, The Network Society: Social Aspects of New Media, 2nd ed. London: Sage, 2006.
5. C. B. Williams, G. J. Gulati, and D. J. Yates, “Predictors of on-line services and e-participation: A cross-national comparison,” Government Information Quarterly, vol. 31, no. 4, pp. 526–533, 2014.
6. S. Mardikyan, E. Ayçiçek Yıldız, M. D. Ordu, and B. Şimşek, “Examining the Global Digital Divide: A Cross-Country Analysis,” Communications of the IBIMA, vol. 2015, Article ID 592253, 2015, doi: 10.5171/2015.592253.
7. N. F. Zakiuddin, S. M. Anggara, and Suhardi, “Developing Digital Service Transformation Maturity Model in Public Sector,” IEEE Access, vol. 12, pp. 174491–174506, 2024, doi: 10.1109/ACCESS.2024.3468341.
8. H. Kaneko, “Cross-validated permutation feature importance considering correlation between features,” Analytical Science Advances, vol. 3, nos. 9–10, pp. 278–287, 2022, doi: 10.1002/ansa.202200018.
9. J. Pohjankukka, T. Pahikkala, P. Nevalainen, and J. Heikkonen, “Estimating the prediction performance of spatial models via spatial k-fold cross validation,” International Journal of Geographical Information Science, vol. 31, no. 10, pp. 2001–2019, 2017, doi: 10.1080/13658816.2017.1346255.
10. World Bank, “GovTech Dataset (December 2025),” World Bank Data Catalog, 2025. Available: https://datacatalog.worldbank.org/search/dataset/0037889/govtech-dataset (accessed July 28, 2026).
11. World Bank, World Development Report 2021: Data for Better Lives. Washington, DC: World Bank, 2021.