ISSN 2409-2665
Journal of Logistics, Informatics and Service Science
Vol. 13 (2026) No. 11, pp.262-278
DOI:10.33168/JLISS.2026.1113

Sustainability-Oriented Legal Compliance Risk Screening for Belt and Road Cross-Border E-Commerce Logistics: An Auditable Decision-Support Framework

Pengpeng Jia[1], Toktobaev Bolot Toktomyshevich[1], aidarbekova Gulnara[1], Tynaeva Nazgul Talantbekovna[1], Shaozheng Guo[1], Guiling Liao[3], Lei Mei[2,3,*]

[1] The Kyrgyz National University named after Zhusup Balasagyn, Bishkek, Kyrgyzstan.
[2] Farabi International Business School, Al-Farabi Kazakh National University, Kazakhstan.
[3] School of Business, Geely University of China, ChengDu, China.
598589134@qq.com (Corresponding author)

Abstract. Cross-border e-commerce expands market access across Belt and Road economies, but platform-mediated logistics exposes firms to heterogeneous customs, data, consumer-protection, taxation, cybersecurity, and environmental requirements. This study develops a transparent country-level screening and decision-support framework that integrates governance quality, logistics performance, economic exposure, and digital readiness. After correcting country-name matching with ISO3 identifiers, the analysis covers 66 economies using 2023 Worldwide Governance Indicators, Logistics Performance Index, and World Development Indicators data. A weighted proxy (governance 0.50, logistics 0.30, digital readiness 0.20) is evaluated with logistic regression, random forest, and gradient boosting under a stratified 49/17 holdout and random seed 42. Logistic regression and random forest each obtain 0.82 accuracy, 0.73 precision, 1.00 recall, 0.84 F1-score, and 1.00 AUC; gradient boosting obtains 0.59 accuracy, 0.70 F1-score, and 0.94 AUC. These metrics measure recovery of an internally constructed median-split proxy, not prediction of independently adjudicated violations. Regulatory quality, rule of law, government effectiveness, logistics competence, and tracking capability provide the strongest screening signals. The framework converts relative risk evidence into proportionate due-diligence, documentation, routing, monitoring, and escalation controls while remaining subordinate to transaction- and jurisdiction-specific legal review.

Keywords: Belt and Road; cross-border e-commerce; legal compliance; logistics risk; decision support; supply chain governance; machine learning.

1. Introduction

Cross-border e-commerce has altered the geography and tempo of international trade. Small consignments, frequent transactions, rapid product turnover, platform intermediation, and consumer-facing fulfillment compress decisions that conventional trade could distribute across longer cycles. Under the Belt and Road Initiative, expanded transport corridors, digital marketplaces, and overseas warehouses have improved access to markets across Asia, Europe, the Middle East, and Africa. Yet these gains also multiply the interfaces at which a shipment, data record, payment, product claim, or return can become non-compliant. Logistics is therefore not a neutral carrier of legal obligations; it is the operational system through which those obligations are performed.

Compliance exposure arises from several interacting regimes. Customs rules determine classification, valuation, origin, and release; product and consumer rules govern labeling, safety, warranties, and returns; tax and payment rules shape settlement and reporting; data and cybersecurity rules constrain the collection and transfer of customer and shipment information; and environmental rules increasingly affect packaging, transport emissions, and market entry. Differences in rule of law and administrative capacity influence how consistently these formal requirements are interpreted and enforced. Institutional uncertainty can consequently appear as clearance delay, repeated inspection, document rejection, inventory detention, inconsistent fees, or a dispute over responsibility (Anderson & Marcouiller, 2002; Kaufmann et al., 2011).

The networked structure of cross-border e-commerce amplifies these effects. A single order may involve a marketplace, seller, payment provider, freight forwarder, customs broker, bonded facility, last-mile carrier, and return processor. Information generated by one actor is reused by others, while liability may remain contractually fragmented. Kawa and Zdrenka (2016) describe the need for integrative coordination in cross-border e-commerce, and Giuffrida et al. (2020) show that logistics service quality is central to cross-border business-to-consumer performance. Operational delay and legal uncertainty are thus mutually reinforcing: weak information quality increases compliance risk, while regulatory intervention increases lead-time variability and cost.

Existing research provides strong foundations in institutional economics, trade logistics, supply chain resilience, and predictive analytics (Baryannis et al., 2019; Choi et al., 2018). Nevertheless, three gaps remain. First, legal compliance is often discussed as a list of rules rather than as a measurable operating environment. Second, country-screening studies seldom integrate institutional, logistics, and digital conditions in one auditable model. Third, predictive studies may report accuracy without translating model output into proportionate controls for market entry, route selection, or partner governance. These gaps are especially important in Belt and Road networks, where institutional and infrastructural heterogeneity is not peripheral but constitutive of the network (Herrero & Xu, 2017; Lee & Shen, 2020).

This study therefore asks three questions: (1) Which country-level institutional, logistics, economic, and digital indicators provide useful signals of relative legal compliance vulnerability? (2) How consistently do linear and nonlinear classifiers distinguish higher- and lower-risk jurisdictions? (3) How can the resulting evidence be converted into proportionate supply chain controls without overstating the model as legal advice? The objective is to develop a transparent screening framework for 66 Belt and Road economies using public, comparable indicators and ISO3-based entity matching.

The contribution is fourfold. Conceptually, the paper links legal compliance to the institutional and operational interfaces of a logistics network. Methodologically, it documents the proxy construction, standardization, model comparison, and interpretation steps so that the framework can be replicated and challenged. Empirically, it identifies governance and logistics capabilities that dominate relative-risk classification. Managerially, it connects risk tiers to due diligence, documentation, routing, monitoring, and escalation decisions. Importantly, the paper treats the model as a screening device: a country score cannot determine whether a specific product, transaction, or contractual arrangement is lawful.

2. Literature Review

2.1. Cross-Border E-Commerce and Logistics Dependence

Cross-border e-commerce combines digital market access with physical fulfillment. Its apparent dematerialization ends when a product enters a warehouse, border crossing, inspection channel, or return network. Logistics performance therefore affects delivery reliability, consumer trust, and the feasibility of a platform promise. Giuffrida et al. (2020) identify customs, international transport, warehousing, last-mile delivery, and reverse logistics as interdependent components of cross-border business-to-consumer service. Earlier work on developing economies also shows that weak infrastructure, payment frictions, and institutional barriers constrain viable e-commerce models (Kshetri, 2007).

The literature increasingly emphasizes coordination rather than isolated transport efficiency. Omni-channel and e-commerce systems require inventory visibility, responsive fulfillment, coherent returns, and cross-firm integration (Hübner et al., 2016; Huo et al., 2016). Danh et al. (2025) illustrate how waiting time, actor linkages, and quality management jointly shape supply-chain performance, while Liu and Walsh (2019) show that logistics organization and standardization are central to fresh-product e-commerce development. These studies support an interface view: the unit of vulnerability is not only the firm or country, but the transfer of goods, documents, data, and accountability between actors.

2.2. Institutional Environment and Compliance Exposure

Institutional theory explains why identical commercial practices can generate different outcomes across jurisdictions. Formal rules define permissible action, but governance quality influences consistency, transparency, and enforceability. Rule of law affects contract expectations and dispute resolution; regulatory quality affects the credibility and design of policy; government effectiveness shapes administrative execution; and control of corruption affects whether official procedures operate predictably (North, 1990; Kaufmann et al., 2011). For cross-border operators, these dimensions influence the variance around a formal compliance process as much as its nominal requirement.

Environmental requirements illustrate how compliance pressure can also create strategic capability. Syafrianita et al. (2025) show in Indonesia's courier sector that green technology and green transportation strengthen green supply chain performance, with agility translating green capabilities into operational outcomes. For cross-border e-commerce, packaging, transport mode, emissions reporting, and overseas warehousing are increasingly subject to environmental rules (Mangiaracina et al., 2015). Compliance screening should therefore distinguish the stringency of a rule from the predictability of its administration: a demanding but transparent jurisdiction may be easier to operate in than a nominally permissive but inconsistently administered one.

2.3. Supply Chain Risk, Resilience, and Network Vulnerability

Supply chain risk research distinguishes the probability of disruption from the network's capacity to anticipate, absorb, and recover. Jüttner et al. (2003) and Tang (2006) frame risk as exposure created by supply, demand, process, control, and environmental sources. Christopher and Peck (2004), Sheffi (2005), and Sheffi and Rice (2005) emphasize visibility, collaboration, agility, and enterprise resilience, while Ivanov and Dolgui (2020) extend resilience toward viability in intertwined networks. These perspectives imply that compliance capability is not merely preventive. Firms also need detection, containment, alternative routing, evidence preservation, and recovery mechanisms when regulatory events occur.

Recent work places prediction and adaptation closer together. Abdulwili et al. (2025) propose an anticipatory and ethically grounded architecture for resilient logistics in crisis-affected regions. Asim et al. (2024) find that resilience depends on capabilities that reduce vulnerability in large-scale manufacturing supply chains. Although these contexts differ from e-commerce, both reinforce the need to combine an early-warning signal with decentralized response capacity. A country risk score has managerial value only when it triggers a defined action by a responsible actor.

2.4. Digital Readiness, Traceability, and Cyber Risk

Digital readiness reduces some compliance frictions by improving traceability, document availability, communication speed, and transaction visibility. The UNCTAD B2C E-Commerce Index treats Internet use and secure-server density as components of e-commerce readiness (UNCTAD, 2018). Secure servers, interoperable records, and reliable connectivity can support electronic customs filing, shipment tracking, audit trails, and identity verification. At the same time, digitalization creates new dependencies: compromised customer data, payment information, trade documents, or platform services can interrupt both compliance and fulfillment. The NIST Cybersecurity Framework 2.0 accordingly treats cyber-risk governance and supply-chain risk management as enterprise responsibilities (NIST, 2024).

The implication is a dual role for digital maturity. It is a protective capability when it improves visibility and control, but it can enlarge the attack surface when governance and security lag behind adoption. Izunobi and Burinskiene (2025) similarly find that digital maturity and change-management capability must develop together in global supply chains. Accordingly, the present framework uses Internet access and secure-server density as readiness signals, not as proof that digital activity is compliant or secure.

2.5. Belt and Road Context and Research Gap

The Belt and Road Initiative creates a setting in which connectivity gains coexist with institutional distance. Transport investment may reduce physical travel time while legal and administrative differences continue to affect border processing, data exchange, product entry, and dispute handling. Lee and Shen (2020) argue that Belt and Road logistics innovation requires coordinated network design, and Butt and Ali (2020) show that sustainability concerns are embedded in supply-chain development. A useful screening model must therefore integrate opportunity-related capacity with exposure-related institutional conditions.

The literature review supports a four-layer mechanism. Institutional quality affects the predictability of rules and enforcement. Logistics capability affects whether requirements can be executed reliably. Digital readiness affects visibility, record integrity, and coordination. Economic exposure affects transaction volume and incentives but does not, by itself, guarantee compliance quality. The gap is an auditable country-level framework that integrates these layers and maps relative risk to operational controls. Fig. 1 summarizes that logic.

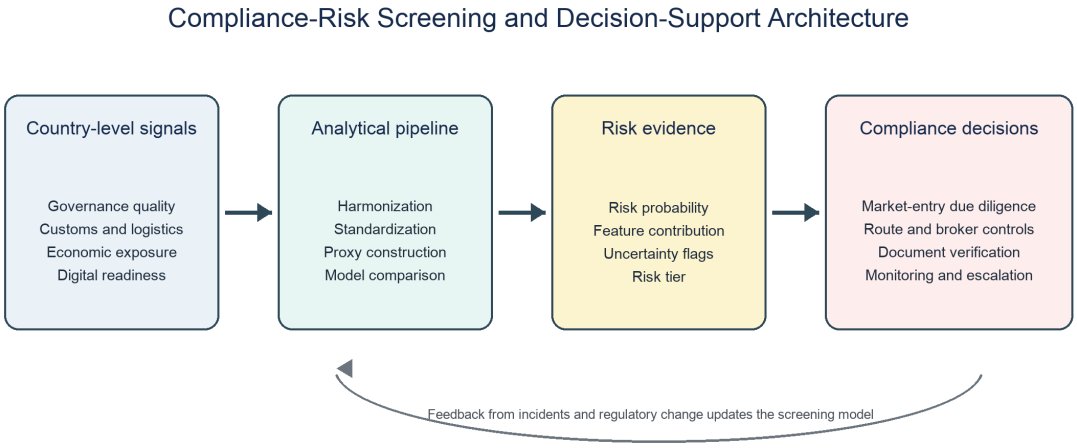


Fig. 1: Compliance-Risk Screening and Decision-Support Architecture

The architecture separates evidence generation from the ultimate legal decision. Country indicators pass through harmonization, standardization, proxy construction, and model comparison. The outputs are a relative score, a probability, feature contributions, and uncertainty flags. Those outputs support controls, while feedback from incidents and regulatory change updates the model. This separation is essential because legal compliance remains transaction-, product-, and jurisdiction-specific.

3. Methodology

3.1. Research Design and Unit of Analysis

The study adopts a cross-sectional, country-level comparative design. The unit of analysis is a Belt and Road economy for which the selected governance, logistics, economic, and digital indicators can be aligned. The author-defined inclusion list contains 66 economies. ISO3 identifiers are used for the analytical merge, preventing country-name variants—notably Turkey/Turkiye—from silently removing an intended observation. The framework is intended for ex ante screening: it ranks relative operating vulnerability before product- and transaction-specific legal review. It does not estimate the causal effect of a single law or predict adjudicated violations.

The workflow has six stages: define the compliance mechanism; collect comparable public indicators; harmonize country identifiers and observation windows; transform skewed variables and standardize scales; construct a transparent proxy and binary screening label; and compare classification models using a holdout sample. The final stage translates model evidence into proportionate controls. Every transformation is designed to be auditable so that a compliance team can understand why a jurisdiction is flagged.

3.2. Data Sources and Temporal Alignment

Governance indicators are the 2023 estimates from the Worldwide Governance Indicators (WGI); logistics measures are from the 2023 International Logistics Performance Index (LPI); and economic and digital indicators are 2023 observations from the World Development Indicators (WDI). The governance block comprises regulatory quality, rule of law, control of corruption, government effectiveness, and political stability. The logistics block comprises customs, infrastructure, logistics competence, tracking and tracing, and timeliness. Trade openness and gross domestic product per capita represent economic context, while Internet use and secure-server density represent digital readiness (Arvis et al., 2023; UNCTAD, 2018; World Bank, 2023a, 2023b).

All requested indicators are aligned to 2023. Country entities are merged by ISO3 code, duplicate records are removed, and secure-server density is reconstructed as the number of secure Internet servers divided by population and multiplied by one million. The source year, indicator code, raw value, missingness flag, and official source URL are retained for every country-variable observation. The supplementary materials provide the 66-economy list and a 924-row country-variable source-year matrix alongside the raw and processed analytical files.

3.3. Variable Construction

The dependent construct, legal compliance risk proxy, captures relative vulnerability rather than observed legal guilt. Each continuous input is standardized so that variables measured in governance units, index scores, percentages, and monetary values can be combined. For variable x in country i, the z-score is:

$$z_{ij} = \frac{x_{ij} - \bar{x}_j}{s_j} \quad (1)$$

where x_{ij} is the observed value of indicator j for country i; $\bar{x}_j$ is the sample mean of indicator j; and s_j is its sample standard deviation.

Secure-server density is transformed with log(1 + x) before it is standardized within the digital-readiness proxy. Gross domestic product per capita remains in current US dollars in the recovered model and is standardized by the logistic-regression pipeline. Higher governance, logistics, and digital-readiness scores imply lower vulnerability, so the composite proxy reverses their direction. The implemented weights are 0.50 for governance, 0.30 for logistics, and 0.20 for digital readiness; these weights are transparent modeling choices rather than empirically identified causal effects.

$$LCR_i = -(w^G G_i + w^L L_i + w^D D_i) \quad (2)$$

where G, L, and D are standardized block scores for governance, logistics, and digital readiness. The continuous score supports ranking. For model comparison, economies above the sample median

are coded as higher relative risk (1) and those at or below the median as lower relative risk (0). Median splitting is a screening convention rather than a natural legal threshold, and the same conceptual indicator system contributes to both the label and predictors.

Table 1. Constructs, Operational Indicators, and Expected Risk Direction

Construct	Operational measure	Role	Expected association with risk
Legal compliance risk proxy	Reverse composite of institutional, logistics, and digital block scores	Outcome	Higher = greater vulnerability
Regulatory quality	Worldwide Governance Indicators estimate	Governance	Negative
Rule of law	Worldwide Governance Indicators estimate	Governance	Negative
Control of corruption	Worldwide Governance Indicators estimate	Governance	Negative
Government effectiveness	Worldwide Governance Indicators estimate	Governance	Negative
Political stability	Worldwide Governance Indicators estimate	Governance	Negative
Customs performance	Logistics Performance Index score	Logistics	Negative
Infrastructure quality	Logistics Performance Index score	Logistics	Negative
Logistics competence	Logistics Performance Index score	Logistics	Negative
Tracking and tracing	Logistics Performance Index score	Logistics	Negative
Timeliness	Logistics Performance Index score	Logistics	Negative
Trade openness	Trade as a percentage of gross domestic product	Economic control	Ambiguous
Income level	Gross domestic product per capita (current US dollars; untransformed in the recovered model)	Economic control	Model covariate; no imposed sign
Internet use	Individuals using the Internet (% of population)	Digital	Negative
Secure-server density	Secure Internet servers per million people; log(1 + x) in the digital proxy	Digital	Negative
E-commerce readiness	Composite of Internet use and secure-server density	Digital	Negative

3.4. Preprocessing and Missing-Data Controls

Numeric fields are coerced consistently and processed in the alphabetical API order recovered from the original workflow. To preserve the intended 66-economy coverage, missing numeric inputs are replaced with their full-sample arithmetic mean before index construction. This affects the five LPI dimensions for 15 economies, four trade-openness observations, and three Internet-use observations; the five WGI indicators are complete. A variable-level and country-level imputation audit is supplied with the data. Because this procedure uses full-sample information and attenuates dispersion where an entire block is absent, the resulting holdout metrics are treated only as internal structural diagnostics.

Multicollinearity is expected because governance and logistics indicators measure related institutional capabilities. Logistic coefficients are therefore interpreted cautiously, and the nonlinear models are used only to examine whether interactions and thresholds alter internal classification. Correlation and feature contribution are descriptive, not causal. The label is constructed from related public indicators; high performance therefore reflects construct coherence and cannot be interpreted as validation against an independent database of violations.

3.5. Model Specification and Internal Validation

Logistic regression provides an interpretable benchmark for binary outcomes and model diagnostics (Hosmer et al., 2013). For country i with predictor vector X_i, the probability of higher relative risk is estimated as:

$$P(Y_i = 1 \mid X_i) = \frac{1}{1+exp[-(\beta_0+\beta^\top X_i)]} \quad (3)$$

where $P(Y_i = 1 / X_i)$ is the estimated probability that country i belongs to the higher-risk group; X_i is the predictor vector; β_0 is the intercept; and β is the vector of estimated coefficients.

Random forest aggregates decorrelated decision trees and can represent nonlinear relations and interactions without imposing a single functional form (Breiman, 2001). Gradient boosting builds trees sequentially to correct prior residual error (Friedman, 2001). The recovered specifications are logistic regression with a StandardScaler and 2,000 maximum iterations, random forest with 300 trees and maximum depth 6, and gradient boosting with 200 trees and learning rate 0.05. Each model uses random seed 42.

$$\hat{y}_i^{RF} = \frac{1}{T}\sum_{t=1}^{T} h_t(X_i) \quad (4)$$

where T is the number of trees and $h_t(X_i)$ is the class-probability prediction produced by tree t for country i. Averaging decorrelated trees reduces variance and captures nonlinear interactions among the predictors.

$$F_m(X_i) = F_{m-1}(X_i) + \eta f_m(X_i) \quad (5)$$

where $F_m(X_i)$ is the updated prediction at boosting iteration m; $f_m(X_i)$ is the new weak learner fitted to the preceding residual error; and η is the learning rate controlling the contribution of each new tree.

A stratified 25% holdout preserves class balance and produces 49 training and 17 test economies. Models are compared using accuracy, precision, recall, F1-score, and area under the receiver operating characteristic curve. Accuracy alone can conceal asymmetric error: a false negative may leave a higher-vulnerability environment under-reviewed, whereas a false positive may impose unnecessary control cost. This is internal validation against a label derived from the same conceptual system; it does not test prediction of independent legal events. With only 17 test observations, one case changes accuracy by 5.88 percentage points.

3.6. Decision-Support Logic and Ethical Safeguards

The model output is converted into risk tiers based on relative rank and uncertainty. A high score triggers enhanced due diligence, not automatic market exclusion. Human review checks product rules, sanctions, customs classification, tax registration, data transfer, consumer remedies, and contractual allocation of responsibility. Evidence should be documented so that a decision can be challenged and updated. Automated scores must not be used to discriminate against a country, supplier, or customer without a transaction-relevant rationale.

Governance safeguards include version control, source-year disclosure, model monitoring, override logging, and periodic back-testing against actual clearance events, disputes, penalties, and document rejections. Data-minimization and cybersecurity controls are necessary where firm-level operational data are added. These safeguards align predictive analytics with accountable decision-making and enterprise cyber-risk governance (Waller & Fawcett, 2013; NIST, 2024).

4. Results and Analysis

4.1. Descriptive Statistics

The corrected 66-economy sample shows substantial heterogeneity. Governance estimates cover broad negative-to-positive ranges, and logistics indicators range from weak to advanced performance. Trade openness ranges from 20.60% to 323.70% of gross domestic product. Internet use ranges from 9.29% to full population coverage, while secure-server density ranges from 17.47 to 228,648.44 per million people. This variation supports comparative screening but cautions against a one-size-fits-all compliance policy.

Table 2. Descriptive Statistics

Variable	Mean	Standard deviation	Minimum	Maximum
Legal compliance risk proxy	0.00	0.80	−2.53	1.58
Regulatory quality	−0.04	0.71	−1.47	2.13
Rule of law	−0.25	0.66	−1.72	1.42
Control of corruption	−0.23	0.70	−1.42	1.92
Government effectiveness	0.06	0.71	−1.58	2.26
Political stability	−0.23	0.78	−2.22	1.33
Customs performance	2.83	0.40	2.10	4.20
Infrastructure quality	2.99	0.53	2.10	4.60
Logistics competence	3.06	0.49	2.10	4.40
Tracking and tracing	3.11	0.50	2.00	4.40
Timeliness	3.32	0.40	2.40	4.30
Trade openness	98.57	56.13	20.60	323.70
Gross domestic product per capita	14,258.43	17,365.16	1,002.31	86,382.59
Internet use	78.53	20.08	9.29	100.00
Secure-server density	12,649.67	31,759.47	17.47	228,648.44
E-commerce readiness	0.00	0.88	−2.51	1.49

Note: N = 66. Statistics are calculated after the documented full-sample mean imputation; raw missingness and imputed values are reported in the supplementary audit.

The dispersion is substantively important. Two jurisdictions with similar trade openness may differ sharply in customs reliability or rule of law; two digitally connected markets may differ in administrative predictability. The framework therefore avoids using market size or connectivity as a shortcut for compliance quality. Risk emerges from the configuration of capabilities across blocks.

4.2. Correlation Structure

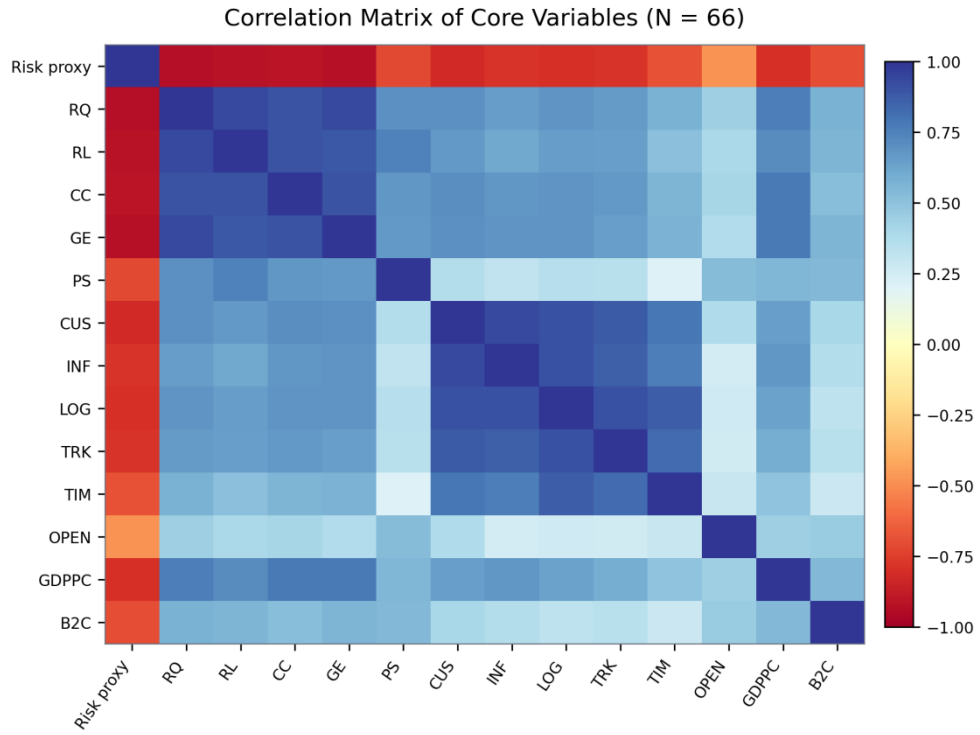


Fig. 2: Correlation Matrix of Core Variables in the Corrected 66-Economy Sample

The correlation matrix in Fig. 2 shows a strong negative association between the proxy and governance quality. Rule of law, regulatory quality, control of corruption, and government effectiveness move together, indicating a common institutional-capacity dimension. Logistics measures also cluster strongly, especially customs, infrastructure, logistics competence, tracking, and timeliness. This pattern is consistent with the idea that border performance depends on an ecosystem rather than a single agency.

The correlations also reveal an interpretation constraint. Because the proxy is assembled from related institutional, logistics, and digital information, the matrix confirms internal coherence but cannot establish that improving one indicator will causally reduce legal incidents. The practical reading is diagnostic: weak scores across multiple blocks indicate a need for stronger controls and more transaction-specific evidence.

4.3. Internal Classification Performance

Against the internally constructed label, logistic regression and random forest each classify 82.35% of the 17 corrected-sample test cases correctly, with precision of 0.73, recall of 1.00, F1-score of 0.84, and AUC of 1.00. Gradient boosting records 58.82% accuracy, 0.53 precision, 1.00 recall, 0.70 F1-score, and 0.94 AUC. These values primarily measure recoverability of the proxy boundary; they do not demonstrate equivalent performance for observed legal violations.

Table 3. Internal Holdout Classification Performance

Model	Accuracy	Precision	Recall	F1-score	AUC
Logistic regression	0.82	0.73	1.00	0.84	1.00
Random forest	0.82	0.73	1.00	0.84	1.00
Gradient boosting	0.59	0.53	1.00	0.70	0.94

Note: Train n = 49; test n = 17; random seed = 42. Labels are an internal median split, not external legal events.

The results favor parsimony. Logistic regression matches random forest on all reported corrected-sample metrics and outperforms gradient boosting on accuracy and F1-score. For organizational

adoption, an interpretable linear benchmark is easier to audit, explain, and challenge. Random forest remains useful as a nonlinear comparison and source of feature contribution, while the poorer gradient-boosting result illustrates the instability of a more complex learner in a small sample.

4.4. Feature Contribution and Discrimination

As Fig. 3 shows, full-sample random-forest feature contribution ranks regulatory quality and rule of law first, followed by government effectiveness, logistics competence, tracking and tracing, control of corruption, customs, and infrastructure. The ordering supports the theoretical mechanism: predictable governance defines the institutional envelope, while logistics capability determines whether formal requirements can be executed and evidenced in routine operations. Trade openness contributes least, indicating that transaction intensity is not equivalent to compliance reliability.

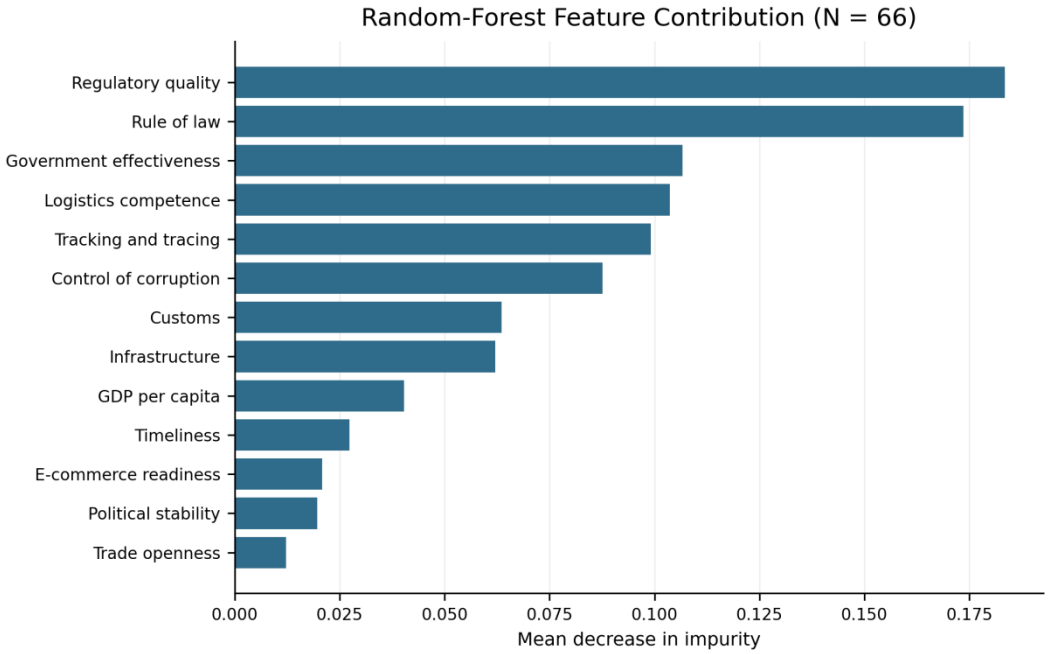


Fig. 3: Random-Forest Feature Contribution in the Corrected 66-Economy Sample

Feature contribution is model-specific and does not measure causal effect. Correlated variables can divide importance, and rankings may change under alternative samples or proxy weights. The ROC curve in Fig. 4 shows perfect separation for the 17-case internal random-forest holdout, but the result is conditional on the constructed label, selected split, imputation procedure, and random seed.

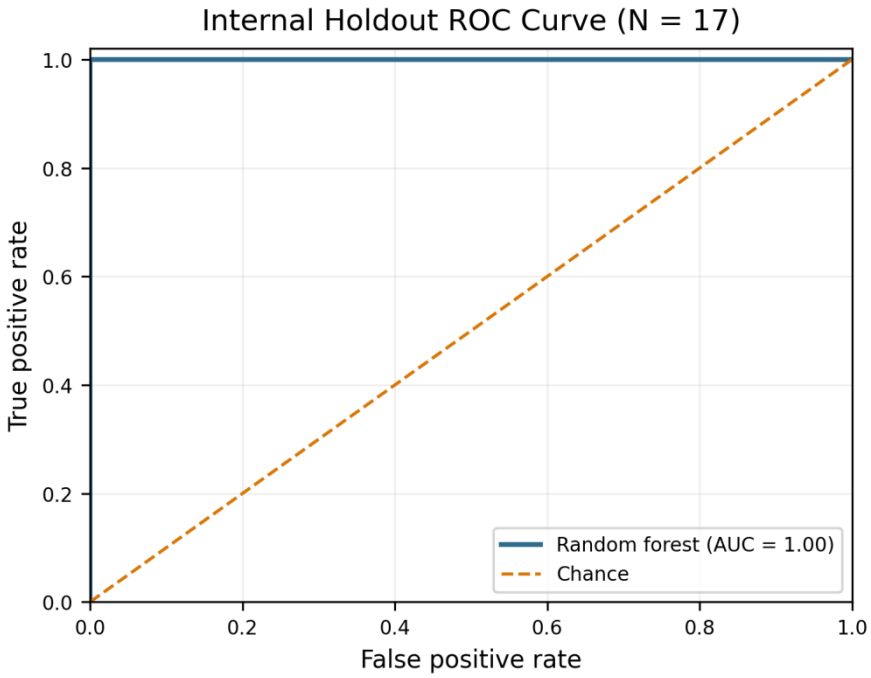


Fig. 4: Receiver Operating Characteristic Curve for the Corrected-Sample Random-Forest Model

4.5. Reproducibility Sensitivity and External-Validation Scope

The primary reproducibility sensitivity compares the recovered 65-economy name-matched sample with the corrected 66-economy ISO3-matched sample under the same random seed. Logistic-regression accuracy declines from 0.88 to 0.82; random-forest accuracy remains 0.82; and gradient-boosting accuracy declines from 0.76 to 0.59. The corrected-sample AUC values are 1.00, 1.00, and 0.94, respectively. These shifts show that headline classification metrics are sensitive to entity resolution and should not be treated as fixed operational guarantees.

Further robustness requires alternative proxy weights and thresholds, repeated resampling, and—most importantly—external validation against customs holds, document rejections, disputes, penalties, product recalls, or data incidents. Until such outcomes are available, country rankings should be reported with uncertainty flags and reviewed qualitatively rather than converted into automatic legal decisions.

5. Discussion

5.1. Theoretical Implications

The findings support an interface theory of compliance vulnerability. Formal legal requirements become operational risks at the boundaries between institutions and logistics processes: declaration and inspection, platform record and customs document, warehouse release and carrier handoff, payment settlement and tax record, or consumer return and cross-border reverse logistics. Strong rule of law and regulatory quality reduce ambiguity at these interfaces; competent customs and logistics systems reduce execution variance; and digital readiness improves visibility and evidence.

The results also extend supply chain resilience research. Resilience is often framed around physical disruption, but compliance events can propagate through the same network mechanisms. A detained shipment can create inventory shortage, missed platform commitments, chargebacks, and partner disputes. A data breach can interrupt payment and fulfillment while triggering notification duties. The anticipatory logic proposed by Abdulwili et al. (2025) and the vulnerability perspective of Asim et al. (2024) therefore apply to regulatory as well as physical shocks.

A further implication is that digitalization is contingent. Secure and mature digital systems improve documentation and coordination, yet they also require cybersecurity, data governance, and

organizational change. This interpretation is consistent with the cyber-risk governance principles in NIST (2024) and the empirical relationship between digital maturity and change-management effectiveness reported by Izunobi and Burinskiene (2025). Digital readiness should therefore be interpreted as an enabling condition, not an automatic compliance benefit.

5.2. Practical Implications for Platforms, Logistics Providers, and Exporters

For platforms, the framework can inform differentiated onboarding and monitoring. Sellers entering higher-risk jurisdictions may require verified product documentation, origin evidence, tax registration checks, enhanced screening of restricted goods, and more frequent review of customer complaints. Platform rules should identify who owns each compliance record and how quickly it must be produced. Automated flags should route cases to trained reviewers rather than silently blocking legitimate trade.

For logistics providers and customs brokers, the strongest signals point to controls at the document-execution interface. High-risk routes justify pre-clearance document review, dual verification of classification and valuation, approved-broker lists, escalation protocols for inconsistent official requests, and retention of evidence showing who changed a declaration. Tracking data should be reconciled with customs status so that an operational delay is not misclassified as a legal breach—or vice versa.

For exporters and overseas-warehouse operators, country screening can sequence market entry. A firm may pilot low-volume products, limit inventory exposure, use delivered-duty arrangements cautiously, and require local-counsel review before scaling. Environmental requirements should be included in the decision, particularly for packaging, batteries, waste, and carbon-intensive transport. Syafrianita et al. (2025) show that green technology and green transportation improve supply-chain performance through agility in the courier sector, suggesting that compliance capability can become a strategic asset rather than only a cost.

5.3. Risk-Tier Decision Matrix

Table 4 translates relative scores into proportionate actions. The tiers are illustrative and should be calibrated to the firm's risk appetite, product category, and actual incident history. The central design principle is proportionality: stronger signals lead to deeper evidence and monitoring, not an unsupported presumption that a jurisdiction or partner is unlawful.

Table 4. Illustrative Compliance Control Matrix

Screening outcome	Due-diligence level	Illustrative controls	Monitoring cadence
Lower relative risk	Standard country review	Normal document checklist; approved broker; periodic regulatory scan	Annual or material-change review
Moderate relative risk	Enhanced route and partner review	Pre-clearance validation; exception logging; ownership of records; sample audit	Quarterly review and incident trend
Higher relative risk	Enhanced legal and operational due diligence	Local legal review; dual approval; restricted-product check; contingency route; evidence pack	Continuous alerts and monthly governance review
High uncertainty	Human review before tier assignment	Verify missing/outdated indicators; compare alternative weights; document override rationale	Reassess when new evidence arrives

5.4. Policy Implications

For policymakers, the results suggest that transport infrastructure alone is insufficient. Predictable customs procedures, transparent guidance, interoperable digital records, secure data exchange, and accessible appeal mechanisms can reduce compliance variance. Publishing machine-readable tariff, product, and procedural information would improve both trader performance and the quality of risk models. Capacity-building should connect customs modernization with anti-corruption safeguards, cybersecurity, and professional competence.

Regional cooperation can also reduce unnecessary heterogeneity. Mutual recognition, interoperable single windows, consistent data definitions, and trusted-trader programs can lower the number of manual handoffs where errors occur. Such initiatives should preserve legitimate regulatory objectives while making evidence requirements clearer. The goal is not deregulation, but more predictable and verifiable compliance.

5.5. Limitations and Future Research

Several limitations are material. First, the dependent construct is a proxy assembled from country indicators rather than a direct count of legal violations. The same conceptual indicator system contributes to the label and predictor set, so the metrics test internal recoverability and are vulnerable to circularity. The AUC values must not be interpreted as evidence that actual enforcement events can be forecast with the same accuracy.

Second, 15 economies lack the entire 2023 LPI block and receive full-sample mean values for its five dimensions; four trade-openness and three Internet-use observations are also imputed. This preserves coverage but can compress real differences and transfers information across the holdout boundary. Future work should compare complete-case, multiple-imputation, and model-based approaches in a prospectively defined pipeline.

Third, the cross-sectional 2023 design cannot establish causal direction or capture rapid legal change. Country averages also conceal differences by border post, region, product, carrier, broker, and firm capability. A panel design with time-stamped regulatory events and hierarchical route- or transaction-level features would support more credible forecasting and explanation.

Finally, the median threshold and 0.50/0.30/0.20 weights are operational choices rather than legal standards. Alternative weighting and threshold assumptions should be reported, and model explanations should be tested with compliance officers, brokers, platform managers, and regulators before production use.

5.6. Implementation and Audit Protocol

A practical implementation should begin with a controlled pilot rather than enterprise-wide automation. The organization should appoint a model owner, a legal owner, and an operational owner; define the products, routes, and decisions within scope; and establish a baseline from recent customs holds, document corrections, complaints, payment disputes, and regulatory inquiries. Historical cases can then be scored retrospectively to determine whether the country-level signal adds information beyond existing checklists. The pilot should specify measurable outcomes, including review time, false-alert rate, avoided document errors, escalation quality, and user understanding of model explanations.

Production use requires a documented decision record. Each assessment should retain the data vintage, transformation version, model version, risk tier, principal contributing factors, reviewer, override rationale, and final control decision. Overrides are not model failures by definition; they are a source of learning when transaction-specific evidence legitimately outweighs a country-level signal. Repeated overrides in the same direction may reveal a stale indicator, an omitted product rule, or a control threshold that is poorly calibrated. Audit logs should therefore be reviewed jointly by compliance, logistics, data governance, and internal audit.

Monitoring should distinguish data drift, model drift, and regulatory drift. Data drift occurs when the distribution of country indicators changes; model drift occurs when classification performance deteriorates against observed outcomes; and regulatory drift occurs when new laws or enforcement practices change the meaning of risk faster than public indicators are updated. A material event—such as sanctions, a customs-system change, a data-transfer restriction, or a major cyber incident—should trigger an interim review rather than waiting for the annual cycle. This governance protocol keeps the framework contestable, current, and subordinate to accountable human judgment.

6. Conclusions

This study develops a transparent legal compliance risk screening and decision-support framework for Belt and Road cross-border e-commerce logistics and supply chain networks. By integrating governance, logistics, economic, and digital indicators for 66 economies and correcting country matching with ISO3 identifiers, it organizes relative vulnerability into an auditable country-level decision aid. Logistic regression and random forest provide the strongest internal classification results, while regulatory quality, rule of law, government effectiveness, logistics competence, and tracking capability provide the leading screening signals.

The principal contribution is not a claim that an algorithm can determine legality. It is a structured connection between country-level evidence and proportionate operational controls. Platforms can differentiate onboarding and monitoring; logistics providers can strengthen document and broker governance; exporters can stage market entry and overseas warehousing; and policymakers can reduce uncertainty through transparent, interoperable, and secure trade administration.

The framework should be used with explicit limits. It ranks relative vulnerability, depends on proxy data and material imputation, and requires human review of the governing law, product, transaction, and contract. Its next research stage is external validation against observed compliance events and dynamic regulatory change. With those safeguards, predictive analytics can support—rather than displace—accountable legal and supply chain judgment.

References

Abdulwili, M. S., Kamal Aldeen, D. J. & Dalati, S. (2025), Reimagining supply chain resilience: Developing RENIAS and EPI for humanitarian logistics in crisis-affected regions—Conceptual framework with an illustrative case in Aleppo, Journal of Service, Innovation and Sustainable Development, Vol. 6, No. 1, 137-160. https://doi.org/10.33168/SISD.2025.0109

Anderson, J. E. & Marcouiller, D. (2002), Insecurity and the pattern of trade: An empirical investigation, Review of Economics and Statistics, Vol. 84, No. 2, 342-352.

Arvis, J.-F., Ojala, L., Shepherd, B., Ulybina, D. & Wiederer, C. (2023), Connecting to Compete 2023: Trade Logistics in an Uncertain Global Economy, World Bank, Washington, DC.

Asim, Z., Sorooshian, S., Vasudevan, A., Yosof, Y. & Huang, T. (2024), Navigating supply chain vulnerability: A case study of resilience in Pakistan's large-scale manufacturing sector, Journal of Infrastructure Policy and Development, Vol. 8, No. 10, Article 5539.

Baryannis, G., Validi, S., Dani, S. & Antoniou, G. (2019), Supply chain risk management and artificial intelligence: State of the art and future research directions, International Journal of Production Research, Vol. 57, No. 7, 2179-2202.

Breiman, L. (2001), Random forests, Machine Learning, Vol. 45, No. 1, 5-32.

Butt, A. S. & Ali, I. (2020), Understanding the implications of Belt and Road Initiative for sustainable supply chains: An environmental perspective, Benchmarking: An International Journal, Vol. 27, No. 9, 2631-2648.

Choi, T.-M., Wallace, S. W. & Wang, Y. (2018), Big data analytics in operations management, Production and Operations Management, Vol. 27, No. 10, 1868-1883.

Christopher, M. & Peck, H. (2004), Building the resilient supply chain, The International Journal of Logistics Management, Vol. 15, No. 2, 1-14.

Danh, L. N., Truc, N. T. T. & Nhan, D. H. (2025), Quantitative analysis and upgrade solutions for the mud crab supply chain in Mekong Delta Vietnam, Journal of System and Management Sciences, Vol. 15, No. 1, 307-325. https://doi.org/10.33168/JSMS.2025.0119

Friedman, J. H. (2001), Greedy function approximation: A gradient boosting machine, Annals of Statistics, Vol. 29, No. 5, 1189-1232.

Giuffrida, M., Mangiaracina, R., Perego, A. & Tumino, A. (2020), Cross-border B2C e-commerce to Greater China and the role of logistics: A literature review, International Journal of Physical Distribution & Logistics Management, Vol. 50, No. 9/10, 895-931.

Herrero, A. G. & Xu, J. (2017), China's Belt and Road Initiative: Can Europe expect trade gains?, China & World Economy, Vol. 25, No. 6, 84-99.

Hosmer, D. W., Lemeshow, S. & Sturdivant, R. X. (2013), Applied Logistic Regression, 3rd ed., Wiley, New York.

Hübner, A., Holzapfel, A. & Kuhn, H. (2016), Distribution systems in omni-channel retailing, Business Research, Vol. 9, No. 2, 255-296.

Huo, B., Ye, Y., Zhao, X. & Shou, Y. (2016), The impact of human capital on supply chain integration and competitive performance, International Journal of Production Economics, Vol. 178, 132-143.

Ivanov, D. & Dolgui, A. (2020), Viability of intertwined supply networks: Extending the supply chain resilience angles towards survivability, International Journal of Production Research, Vol. 58, No. 10, 2904-2915.

Izunobi, D.-A. & Burinskiene, A. (2025), Digital maturity and change management effectiveness in global supply chain operations: An empirical investigation of international businesses, Journal of Management Changes in the Digital Era, Vol. 2, 249-271. https://doi.org/10.33168/JMCDE.2025.0119

Jüttner, U., Peck, H. & Christopher, M. (2003), Supply chain risk management: Outlining an agenda for future research, International Journal of Logistics: Research and Applications, Vol. 6, No. 4, 197-210.

Kaufmann, D., Kraay, A. & Mastruzzi, M. (2011), The worldwide governance indicators: Methodology and analytical issues, Hague Journal on the Rule of Law, Vol. 3, No. 2, 220-246.

Kawa, A. & Zdrenka, W. (2016), Conception of integrator in cross-border e-commerce, LogForum, Vol. 12, No. 1, 63-73.

Kshetri, N. (2007), Barriers to e-commerce and competitive business models in developing countries: A case study, Electronic Commerce Research and Applications, Vol. 6, No. 4, 443-452.

Lee, H. L. & Shen, Z.-J. M. (2020), Supply chain and logistics innovations with the Belt and Road Initiative, Journal of Management Science and Engineering, Vol. 5, No. 2, 77-86.

Liu, X. & Walsh, J. (2019), Study on development strategies of fresh agricultural products e-commerce in China, International Business Research, Vol. 12, No. 8, 61-70.

Mangiaracina, R., Marchet, G., Perotti, S. & Tumino, A. (2015), A review of the environmental implications of B2C e-commerce: A logistics perspective, International Journal of Physical Distribution & Logistics Management, Vol. 45, No. 6, 565-591.

Marti, L., Puertas, R. & García, L. (2014), The importance of the Logistics Performance Index in international trade, Applied Economics, Vol. 46, No. 24, 2982-2992.

NIST. (2024), The NIST Cybersecurity Framework (CSF) 2.0, NIST CSWP 29, U.S. Department of Commerce, Gaithersburg, MD.

North, D. C. (1990), Institutions, Institutional Change and Economic Performance, Cambridge University Press, Cambridge.

Sheffi, Y. (2005), The Resilient Enterprise: Overcoming Vulnerability for Competitive Advantage, MIT Press, Cambridge, MA.

Sheffi, Y. & Rice, J. B. Jr. (2005), A supply chain view of the resilient enterprise, MIT Sloan Management Review, Vol. 47, No. 1, 41-48.

Syafrianita, Purnomo, A., Amran, K. M., Chaidar, D. F. R. & Pakpahan, H. M. (2025), Green technology and transportation in supply chain management: The mediating role of agility on performance in Indonesia's courier industry, Journal of Logistics, Informatics and Service Science, Vol. 12, No. 4, 338-359. https://doi.org/10.33168/JLISS.2025.0418

Tang, C. S. (2006), Perspectives in supply chain risk management, International Journal of Production Economics, Vol. 103, No. 2, 451-488.

UNCTAD. (2018), UNCTAD B2C E-Commerce Index 2018: Focus on Africa, Technical Notes on ICT for Development No. 12, United Nations Conference on Trade and Development, Geneva.

Waller, M. A. & Fawcett, S. E. (2013), Data science, predictive analytics, and big data: A revolution that will transform supply chain design and management, Journal of Business Logistics, Vol. 34, No. 2, 77-84.

World Bank. (2023a), Worldwide Governance Indicators: 2023 Estimates, World Bank, Washington, DC.

World Bank. (2023b), World Development Indicators: 2023 Data, World Bank, Washington, DC.