future transportation

MDPI

Article

Comparing Critical Factors in Indonesia's Sea Toll Program: An Exploratory Fuzzy DEMATEL Analysis

Belva Agustin [1], Nisa Ekşili [2,*], Artuğ Eren Coşkun [3,*] and Pınar Çelik [4]

[1] Andhika Group, Jakarta 13730, Indonesia; belva.agustin@andhika.com
[2] Department of Business Administration, Faculty of Economics and Administrative Sciences, Akdeniz University, 07058 Antalya, Türkiye
[3] Department of International Trade and Logistics, Faculty of Applied Sciences, Akdeniz University, 07058 Antalya, Türkiye
[4] Department of Tourism Management, Serik Faculty of Business Administration, Akdeniz University, 07058 Antalya, Türkiye
* Correspondence: nisaeksili@akdeniz.edu.tr (N.E.); erencoskun@akdeniz.edu.tr (A.E.C.)

Abstract

Indonesia's archipelagic geography creates persistent logistics disparities that the Sea Toll Program has sought to address since 2015. This study integrates stakeholder-derived factor identification with a comparative assessment of expert judgments. Ten factors derived from the responses of 42 stakeholders were evaluated by 10 experts. One-step summaries provide the principal quantitative comparison, with Fuzzy DEMATEL used to explore how higher-order weighting affects the same judgments. K10 (Improving Logistics Activities in Underdeveloped Regions) and K8 (Product Supply and Demand) had the largest positive balances in both summaries, while K1 (Port Infrastructure) had the most negative balance. Descriptive agreement was weak (Kendall's W = 0.069), and all ten 95% bootstrap intervals for the transformed net indices included zero. Fleet availability (K2) and human resources (K5) occupied boundary positions. Geometric aggregation preserved nine baseline signs but changed prominence rankings. Excluding K9 (Corruption and Transparency) reversed the transformed signs of the two boundary factors, K2 and K5. The study contributes a contextual factor framework and identifies where comparative model results are sensitive to expert composition and analytical choices. It specifies route-level indicators for evaluating literature-informed proposals concerning cargo, connectivity, infrastructure, and information systems, without presenting those proposals as an empirically established intervention sequence.

Keywords: Sea Toll Program; Fuzzy DEMATEL method; logistics performance; critical factors; regional inequality; maritime policy

Check for updates

Received: 28 September 2026
Revised: 10 November 2026
Accepted: 16 November 2026
Published: 19 November 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Indonesia is the world's largest archipelagic country. The Geospatial Information Agency reported 17,380 named and georeferenced islands for 2024 [1]. This geography creates persistent logistics challenges because long inter-island distances, uneven infrastructure, and weak intermodal connections increase distribution costs and contribute to large regional price differentials. Earlier official evidence reported that a bag of cement in parts of Papua could cost up to twenty times the price in Java and linked provincial price differences to remoteness and inadequate transport infrastructure [2,3]. National logistics costs have also remained high; a macro-level study estimated them at 23.62% of GDP in 2015 while emphasizing important measurement limitations [4].

The Indonesian government launched the Sea Toll Program (Tol Laut) in 2015 to strengthen scheduled maritime connectivity, reduce scarcity and price disparities, and serve disadvantaged, frontier, outermost, and border regions [5,6]. Official Ministry of Transportation figures show that the program expanded from three routes and eleven ports in 2015 to thirty-nine routes and 114 ports in 2024. The same source reports that handled cargo increased from 30 tons and 88 TEUs in 2015 to 989.75 tons and 31,878 TEUs in 2023 [7]. The ministry reports tonnage and TEUs as separate measures; the figure of 989.75 tons therefore denotes reported cargo throughput rather than nominal vessel or port capacity, and it should not be combined with the TEU figure. The route/port and cargo statistics also refer to different reporting years. Despite this expansion, infrastructure gaps, route economics, institutional coordination, and local-government participation continue to constrain implementation [8–12].

This study uses a sequential exploratory design to identify factors perceived by stakeholders to shape the Sea Toll Program and to examine their comparative positions through an exploratory Fuzzy DEMATEL transformation. Here, 'program performance' is a policy-outcome domain encompassing service regularity, cargo utilization, delivery time, regional price accessibility, port and last-mile efficiency, subsidy effectiveness, and route continuity, not a directly observed dependent variable. The analysis addresses three questions: RQ1, which factors emerge from stakeholder accounts and the relevant literature? RQ2, what patterns emerge in the one-step comparative summaries, and how do they differ when higher-order weighting is applied? RQ3, how stable are these indices under changes in expert composition, weights, TFN width, defuzzification, and aggregation? The contribution lies in integrating qualitative factor development with an uncertainty-qualified comparison of expert judgments. The indices are computational summaries of the supplied matrices; they do not establish the existence, strength, or sequence of directed relationships between the factors.

2. Literature Review

2.1. Archipelagic Logistics and Regional Inequality

The complexity of Indonesia's logistics system stems largely from the country's unique geographical structure. Rodrigue's The Geography of Transport Systems [13,14] theoretically demonstrates that transportation is not merely the movement of goods and people but also a fundamental economic lever connecting regions to one another. Applied to the Indonesian context, Rodrigue's concepts of connectivity and efficiency provide a decisive framework for explaining the structural difficulties encountered in delivering regular and cost-effective transport services to remote regions.

Myrdal's Theory of Cumulative Causation [15] explains how regional inequalities may reinforce themselves through backwash and spread mechanisms. In the Sea Toll context, thin local cargo bases, high logistics costs, and limited infrastructure can mutually reinforce peripheral disadvantage: weak production reduces return cargo, low utilization raises unit transport costs, and high costs further constrain local market development. Conversely, more reliable connectivity and local logistics capacity may support spread effects by enabling production, market access, and investment. The study therefore treats cargo generation (K8), regional logistics capacity (K10), infrastructure (K1), and intermodal integration (K7) as mutually connected conditions rather than as isolated bottlenecks [16,17].

The Bappenas Sea Toll development framework identifies port infrastructure and multimodal connectivity as core components of national logistics integration [18]. Earlier analysis of Indonesia's maritime transport system by Ridwan and Hartini [19], together with more recent evidence on the Sea Toll Program and logistics efficiency [20], underscores the difficulty of designing reliable freight networks across a geographically dispersed

archipelago. These constraints are not only physical: they interact with cargo concentration, institutional arrangements, and the organization of inter-island services.

2.2. The Sea Toll Program: Development, Objectives, and Performance

The Sea Toll concept was publicly advanced in 2014, and the scheduled public-service freight program was formally launched on 4 November 2015 [5,21]. Its initial legal basis was Presidential Regulation No. 106 of 2015 and Minister of Transportation Regulation No. PM 161 of 2015 [22,23]. These instruments established public-service obligations for regular maritime freight to disadvantaged, remote, outermost, and border regions; both were subsequently superseded by later regulations. In operational terms, the program combines scheduled trunk and feeder services, designated cargo, vessel assignments, and tariff subsidies across western and eastern Indonesia.

The pioneering assessment by Prihartono et al. [18] defines the program's primary objective as connecting large and small islands through regular voyages, reducing distribution costs, and eliminating inter-regional price disparities. A subsequent analytical review by Nur et al. [24] reports price reductions of approximately eleven to twenty percent for selected necessities in destination areas, while also emphasizing the need for continued evaluation. Fahmiasari and Parikesit [25] compare Sea Tollway and Nusantara Pendulum network configurations and show that maritime network design is central to east–west connectivity.

At the same time, studies documenting the program's structural and institutional problems occupy a prominent place in the literature. The four-year assessment by Sa'adah et al. [26] shows that, despite observable improvements in connectivity, return cargo, institutional coordination, and distribution sustainability remain unresolved. Ratnawati [27] identifies supply-chain bottlenecks and inadequate port infrastructure, while Kristini et al. [28], using AHP evidence from the Tahuna route, prioritize multimodal links, port-adjacent infrastructure, and container costs. Together, these studies indicate that physical connectivity alone does not guarantee efficient and durable service.

More recent studies deepen these structural concerns. Fauzi et al. [8] identify the continuity of business financing, the inadequacy of inland logistics infrastructure, and the shortage of qualified human resources as the primary factors threatening the program's sustainability. Rahayu et al. [10] argue that the effective participation of local governments is indispensable to the program's success and that the lack of coordination between central and local government creates administrative and technical obstacles. Apriani et al. [29], meanwhile, position the Sea Toll not merely as a logistics strategy but as a "blue economy" approach that integrally supports the fisheries, tourism, and maritime industries.

2.3. Supply Chain Management and Logistics Performance

Supply chain management (SCM) provides a process-integration lens rather than a single testable theory in this study. Mentzer et al. [30] conceptualize SCM as the systemic coordination of business functions across organizations, while Lambert et al. [31] emphasize cross-organizational process integration. Empirical supply-chain research further shows that internal, supplier, customer, information, and logistics integration operate jointly rather than as isolated capabilities [32,33]. Applied here, this perspective explains why port, fleet, cargo, information, subsidy, and hinterland conditions should be assessed as an interdependent network. The framework does not treat D-R as an objective KPI or causal coefficient; instead, it uses prominence and net comparative index to organize stakeholder perceptions about where coordination failures and enabling conditions are located.

In the Indonesian context, SCM plays a decisive role in establishing inter-regional connectivity. Humang et al. [34] show that integration among Sea Tollway, Pelni, and Pioneer services depends jointly on time, cost, freight, port and vessel characteristics, gov-

ernment policy, and operating conditions. At the port-system level, Song and Panayides [35] demonstrate that intermodal services, organizational relationships, information and communication systems, value-added services, and logistics practices jointly shape supply-chain integration and competitiveness.

Objective performance indicators remain important for subsequent validation. Relevant route-level measures include schedule reliability, vessel utilization, return-cargo share, port waiting and turnaround time, last-mile cost, subsidy per unit carried, price dispersion, and route continuity. Multi-stakeholder port-performance research similarly emphasizes that performance indicators are interdependent and should reflect different stakeholder perspectives [36], while performance-measurement guidance requires indicators to be linked to explicit decision purposes [37]. These indicators were not used as outcome data in the present model; they are specified as an external-validation agenda and as monitoring variables for the policy hypotheses developed below.

2.4. Factors Affecting the Performance of the Sea Toll Program

Port infrastructure is repeatedly treated as an enabling condition rather than a standalone outcome. The initial policy design linked government-defined routes and tariffs with public-service obligations, state compensation, and the assignment of PT PELNI as the operator [22]. At route level, Kristini et al. [28] identify multimodal access, infrastructure around the port, and container costs as priority subfactors for improving effectiveness and efficiency. Port readiness is therefore considered together with hinterland connections, cargo availability, and operating arrangements in the present factor system.

Fleet management and route utilization remain fundamental operational concerns. A route-level evaluation by Mawikere and Herijanto [38] documents uneven berth utilization, loading-unloading productivity, and ship-service performance across three eastern Indonesian ports. Through return-cargo factor analyses, Gugat et al. [39] show that thin backhaul volumes undermine vessel utilization, while Ratnawati [27] argues that stronger return cargo could alter the program's cost structure. This evidence supports treating fleet availability, cargo balance, and port performance as interdependent rather than isolated conditions.

UNCTAD's [40] comprehensive assessment emphasizes that digital transformation accompanies physical modernization in improving port and maritime-logistics performance. In the Sea Toll context, the official SITOLAUT platform provides route, schedule, tariff, cargo, and participant information and supports online access to subsidized logistics services [41]. The present study does not audit the route-level coverage or performance of this platform; rather, it treats data visibility and traceability as elements of the governance construct.

Institutional coordination is another recurring concern. Prihartono et al. [18] emphasize the division of responsibilities among national ministries, local governments, operators, and cargo actors. Rahayu et al. [10] show more broadly that Indonesian maritime governance spans port-to-port, within-port, and port-to-hinterland subsystems whose governance arrangements are not always aligned. This fragmentation helps explain why formal authority does not automatically produce integrated implementation.

Mubarak et al. [9] likewise identify overlapping authority and policy inconsistency as implementation constraints. Rahayu et al. [10] caution that coordination must account for the practical differences among maritime-transport subsystems, while Fauzi et al. [8] identify infrastructure, synchronization, and stakeholder participation as continuing challenges. These studies support modeling coordination and regulation as interdependent institutional conditions rather than as exogenous guarantees of performance.

The effectiveness of government subsidies also remains contested. For 2024, the government reported an aggregate IDR 1.87 trillion allocation for pioneer maritime transport, covering 117 pioneer routes, six livestock routes, thirty-nine Sea Toll routes, and sixteen rede

services [42]. Separately, PT PELNI reported an IDR 214 billion budget for its 2024 Sea Toll assignment [43]. Supply Chain Indonesia [44] also records the policy debate surrounding subsidy continuity and return-cargo requirements. These figures describe different budget scopes and should not be combined as if they referred to a single program allocation.

Qualified human resources are repeatedly identified as important for maritime logistics implementation [8,18,26]. Transparency and integrity are also relevant because route, cargo, price, and subsidy decisions depend on complete, verifiable information. Port-governance research links transparency to accountability and stakeholder trust [45], while Indonesian evidence describes digital systems that strengthen cargo visibility and coordination [46]. These studies motivate the analytical governance interpretation of K9. However, the expert instrument used the label 'Corruption and Transparency'; its numerical position cannot establish that every expert evaluated an identically defined, positively oriented governance construct.

The lack of integration among transport modes is persistently emphasized in the literature as one of the program's key bottlenecks. Ratnawati [27] foregrounds the critical importance of linking different transport modalities and shows that inadequate road infrastructure poses a serious obstacle to the distribution of goods. Kristini et al. [28] similarly rank multimodal transport among the priority subfactors in their AHP assessment of the Tahuna route. Sa'adah et al. [26] emphasize that the Sea Toll's success is directly tied to the integration of main and feeder ports and to the effectiveness of hinterland logistics systems.

Amin et al. [47] show that the Sea Toll Program has opened new economic potential for maritime transport, while noting that sustaining these gains depends on stronger multimodal integration infrastructure. This requirement is also consistent with Rodrigue's [14] theoretical framework on the geography of transport systems, which holds that effective logistics networks must combine seamless connectivity with high efficiency. Chronic inadequacies in inland transport connections, particularly in eastern Indonesia, continue to sustain the last-mile distribution problem as a structural obstacle, considerably limiting the program's regional impact.

2.5. Gaps in the Literature

Prior Sea Toll research has documented connectivity, price disparity, port readiness, return cargo, governance, and local economic development, but the literature uses heterogeneous data sources and analytical levels. Policy evaluations describe implementation constraints [26,27], network studies compare shipping configurations [25], port-impact research evaluates resilience and competitiveness [6], econometric work estimates price effects [48], and system-dynamics studies examine feedback mechanisms [49,50]. These approaches provide complementary evidence. The present study adds stakeholder-derived factor development and a transparent examination of comparative expert judgments, rather than an independent empirical test of the directed relationships investigated by those approaches.

The present study addresses this contextual gap by integrating a stakeholder-derived factor set, an exploratory transformation of comparative judgments, and explicit diagnostics of dependence on panel composition and modeling choices. The analysis distinguishes the original factor labels from their reporting interpretations, including unresolved directional ambiguity in K9. Robustness is evaluated as a property of the specified computational model, not as evidence that the questionnaire directly measured factor-to-factor influence.

Table 1 summarizes the analytical relation between selected prior studies and the present research. The comparison shows that the contribution lies in integrating qualitative factor development, fuzzy comparative modeling, and explicit robustness diagnostics, while recognizing that the outputs remain expert-perception evidence requiring validation against operational data.

Table 1. Research focus and unresolved issues in selected studies on the Sea Toll Program.

Study	Data/Method	Main Focus	Unresolved Issues
Sa'adah et al. [26]	Policy evaluation	Connectivity and distribution implementation	How are operational and institutional factors perceived to interact?
Fahmiasari and Parikesit [25]	Network-efficiency comparison	Container network design and east–west connectivity	How do network, cargo, and institutional conditions relate in one factor system?
Arisusanty et al. [51]	Expert AHP	Feeder-port selection for Sea Toll services	How do the broader Sea Toll factors compare in prominence and net comparative index?
Sunitiyoso et al. [49]	Systems-thinking analysis	Feedback among port-performance factors	Which Sea Toll-specific factors appear prominent, and how stable are their positions?
Sofiyandi et al. [48]	Econometric policy evaluation	Subsidy effects on food prices	Which perceived conditions may qualify subsidy effectiveness?
Amin et al. [47,50]	System dynamics and efficiency analysis	Local economy, cargo volume, and maritime costs	How are regional logistics capacity and cargo balance positioned in the comparative expert assessment?

2.6. Decision-Making Under Uncertainty and the Choice of Fuzzy DEMATEL

Fuzzy set theory permits partial membership and provides a formal representation of linguistic vagueness in expert judgments [52]. In Fuzzy DEMATEL, linguistic or comparative judgments are represented by triangular fuzzy numbers and subsequently transformed into a matrix from which prominence and net comparative indices are obtained [53]. Grey-system approaches offer an alternative for incomplete-information settings [54,55], but they are not applied in this study.

The archived expert evaluations were represented on a reciprocal-style 1/9–9 comparative scale and converted into triangular fuzzy numbers. Nine-point fuzzy comparative scales have precedent in hybrid MCDM applications involving DEMATEL [56]. However, a greater number of categories does not establish equivalence with standard DEMATEL direct-influence elicitation, which commonly assesses the two directions separately [57,58]. The original questionnaire's mirrored response layout and its subsequent comparative coding are therefore distinguished from independent bilateral effect measurement. The present model is an exploratory transformation of the archived computational matrices.

The primary empirical evidence consists of the stakeholder-derived factors and the comparative expert evaluations. The Fuzzy DEMATEL transformation provides an additional, explicitly model-dependent summary of those evaluations. A matrix entry records a coded comparative judgment, not an independently elicited effect of one factor on another. Accordingly, the resulting indices do not identify temporal causality, mediation, or policy-treatment effects.

3. Materials and Methods

As illustrated in Figure 1, the study employed a sequential exploratory design [59]. Open-ended stakeholder responses were first analyzed to develop a context-specific factor set; the resulting factor labels were then transferred to the expert instrument. Integration therefore occurred at the connecting stage, where the qualitative themes became the criteria evaluated in the quantitative phase, and again at interpretation, where the matrix results

were compared with qualitative evidence and prior literature. Because the first phase used an online open-ended questionnaire rather than interviews, it offered broad stakeholder coverage but less opportunity for probing; this trade-off is acknowledged as a limitation.

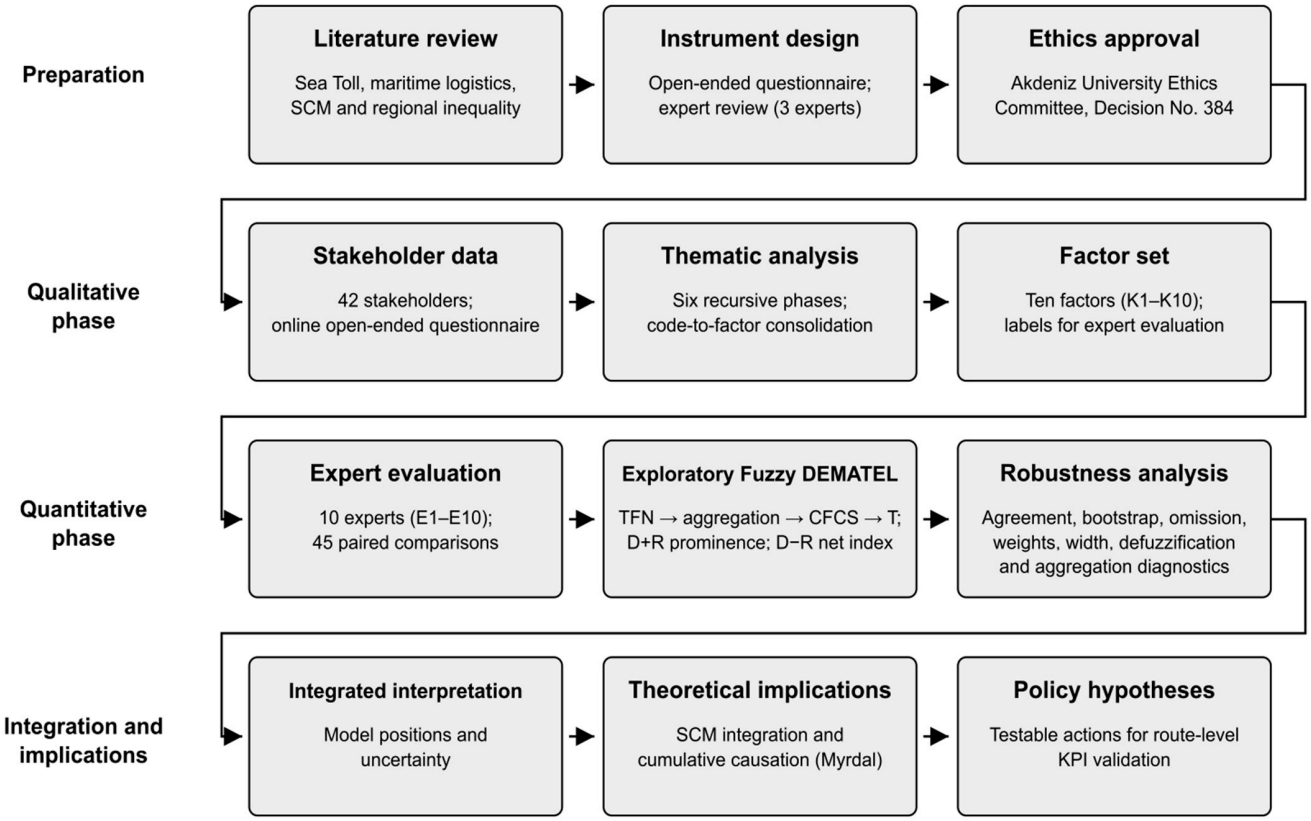


Figure 1. Sequential exploratory research design for the Sea Toll Program. Source: Agustin [60] (Sequential exploratory mixed-methods design adapted from Creswell [59]).

3.1. Sampling Procedure and Data Collection

The qualitative phase used convenience and purposive access to recruit adults who were directly involved in, professionally exposed to, or institutionally knowledgeable about maritime logistics and the Sea Toll Program. The 42 respondents comprised 17 shipping-company representatives, 9 logistics providers, 6 government officials, 4 seafaring or ship-board personnel, 2 maritime academics/instructors, and one respondent from each of the following categories: port authority, PT PELNI, cargo owner/shipper, and human-resources practice. Seventeen respondents reported more than five years of professional experience, five reported three to five years, eleven reported one to three years, and nine reported less than one year (see Table 2). The online instrument contained demographic questions and open-ended items on program objectives, benefits, infrastructure, fleet adequacy, intermodal integration, subsidies, regulation, local economic conditions, workforce, coordination, external conditions, and critical barriers; its construction followed established questionnaire-design principles concerning clarity, relevance, and respondent burden [61].

The factor-development process emphasized information adequacy and iterative review of the responses. No sequential saturation log or complete code-by-case frequency matrix was retained; formal saturation could therefore not be established [62]. Three maritime-logistics experts reviewed the questionnaire and factor structure for clarity, relevance, overlap, and coverage. This review was distinct from member checking by the original participants. Appendix G presents the retained open-ended questions and illustrative coding paths linking participant responses to the ten factors.

Table 2. Profile of the 42 qualitative-phase respondents.

Characteristic	Category	*n*	%
Stakeholder role	Shipping-company representatives	17	40.5
	Logistics providers	9	21.4
	Government officials	6	14.3
	Seafaring/shipboard personnel	4	9.5
	Maritime academics/instructors	2	4.8
	Port authority	1	2.4
	PT PELNI	1	2.4
	Cargo owner/shipper	1	2.4
	Human-resources professional	1	2.4
Professional experience	More than 5 years	17	40.5
	3–5 years	5	11.9
	1–3 years	11	26.2
	Less than 1 year	9	21.4

In the quantitative phase, ten experts were selected to provide institutional and operational breadth: two academics, one port-operations manager, six shipping/logistics specialists or analysts, and one government transport inspector (see Table 3). Eligibility required a role that provided substantive knowledge of maritime logistics, port/shipping operations, policy, or Sea Toll implementation. Experts were coded consistently as E1–E10. Six experts had between 15 and 19 years of experience, while four had 20 years or more. Data for the qualitative and expert-evaluation phases were collected between 10 and 24 June 2025, following ethical approval granted on 10 June 2025. Equal weights were used because no externally validated competence scores were available; deterministic and Dirichlet weight analyses were used to examine dependence on this choice. The small stakeholder strata (including a one-person government stratum and a two-person academic stratum) do not support statistically meaningful group-specific DEMATEL models; individual leave-one-out and weight perturbation analyses provide the more defensible heterogeneity diagnostics.

Table 3. Expert profiles, analytical identifiers, and rationale for selection.

Participant	Role	Reason for Expert Selection
E1	Logistics Faculty Member—Academic	Possesses deep theoretical knowledge of logistics and understands national program policies.
E2	Senior port operations manager	Direct experience with port operations in the remote regions served by the Sea Toll.
E3	Shipping Agency Director	Experienced in managing shipping operations and the efficient use of Sea Toll routes.
E4	Senior executive in an inter-island shipping company	CEO of a maritime transport company distributing goods between islands through the Sea Toll.
E5	General Manager, Logistics Company	Logistics expert with a managerial perspective on the flow of goods in the maritime logistics network.
E6	Logistics Faculty Member—Academic	Academic knowledge of logistics integration and the effects of government policy.
E7	Ship Captain	Experienced in maritime operations and understands the technical challenges of Sea Toll implementation.
E8	Maritime Advisor	Maritime expert with an international perspective on maritime policy and logistics.

Table 3. *Cont.*

Participant	Role	Reason for Expert Selection
E9	Senior executive in a state-owned shipping operator	Involved in strategic decision-making at the state maritime transport company.
E10	Government maritime transport inspector	Oversees and regulates the government implementation of the Sea Toll Program.

3.2. Data Analysis

The qualitative analysis followed six recursive phases: familiarization, meaning-unit identification, initial coding, candidate-theme development, review against the response set, and definition of final factors [63]. One researcher conducted the primary coding and discussed the factor structure with the research team and three maritime-logistics experts. Codes were consolidated when they concerned the same operational issue and separated when they represented distinct resources, processes, institutional conditions, or contextual capacities. Appendix G distinguishes content supported by participant statements from literature-informed elaboration.

Credibility was supported by comparing stakeholder accounts across roles, examining their relation to previous research and official evidence, and obtaining expert review of the questionnaire and factor structure. The findings were not returned to the original 42 respondents, so member checking was not undertaken. The first author collected the data in Indonesia between 10 and 24 June 2025. The procedural sequence comprised questionnaire review, qualitative data collection and coding, factor review, and expert evaluation. The records identify the overall collection window but do not separately date each activity. The quantitative phase comprised individual fuzzification, equal-weight aggregation, CFCS defuzzification, normalization, and calculation of the model indices.

3.3. Fuzzy DEMATEL Method

DEMATEL was developed to represent interdependencies in complex systems [64–66]. Standard elicitation assesses the influence of factor i on factor j separately from the reverse direction. In this study, ten experts each selected one side and intensity category per unordered factor pair for 45 pairs in a mirrored layout. Their computational records comprise ten 10 × 10 matrices, each with a zero diagonal and 90 off-diagonal comparative entries. Only one direction was elicited for each pair; the opposite matrix entry was calculated from the same mark.

Table 4 presents the original factor labels, the corresponding reporting labels, and the scope supported by the retained qualitative responses. The common information supplied to experts is stated once in the table note. For K6–K10, the reporting labels organize the relevant domains rather than establish identical operational definitions across experts. In particular, the original wording of K9 leaves its evaluative direction ambiguous, while the broader capacity interpretation of K10 extends beyond the explicitly retained statements.

The questionnaire displayed intensity codes 0–4 on both sides, including the label "No Impact" for code 0. The marked side and category were converted at the analysis stage to reciprocal comparative values using Table A11. The rule maps code 0 to 1 in both cells; this is a coding convention, not an elicited judgment of equal influence. No comparison was marked at code 0, so this convention did not affect the numerical results. Table 5 gives the TFNs assigned to the coded scores. Appendix H reproduces the recorded response layout and states what the questionnaire documentation supports.

Table 4. Critical factors, original instrument wording, and qualitative scope of interpretation.

Code	Original Instrument Wording	Reporting Label	Qualitative Scope and Interpretation
K1	Port Infrastructure	Port Infrastructure	Port facilities in remote regions; inland access is a related but distinct concern.
K2	Fleet Availability	Fleet Availability	Vessel provision; deployment arrangements extend the explicit fleet-availability wording.
K3	Coordination between Stakeholders	Stakeholder Coordination	Coordination among institutions and stakeholders.
K4	Government Policies and Regulations	Government Policies and Regulations	Legal and regulatory framework; coherence and enforceability are analytical elaborations.
K5	Human Resources	Human Resources	Workforce, knowledge, and skills; digital and administrative competencies were not separately established.
K6	Effectiveness of Government Subsidies	Subsidy Effectiveness	Financial support and subsidy administration; a common outcome-based effectiveness measure was not specified.
K7	Integration of Transport Modes	Intermodal Transport Integration	Connections between sea services and inland delivery.
K8	Product Supply and Demand	Cargo Supply–Demand Balance	Local commodities and return-cargo availability; balance and continuity summarize this domain.
K9	Corruption and Transparency	Governance Transparency and Integrity	Transparency and corruption concerns; the combined original label does not establish a uniform evaluative direction.
K10	Improving Logistics Activities in Underdeveloped Regions	Logistics Capacity Development in Lagging Regions	Local economic activity and shipping access for producers; consolidation and storage are analytical extensions.

Note: Experts received factor labels without written operational definitions. The scope descriptions summarize retained qualitative evidence rather than a standardized definition supplied to the panel. Table A9 presents the participant-linked coding examples. These descriptions do not resolve the original directional ambiguity of K9 or establish the expanded capacity dimensions of K10.

Table 5. One-to-one mapping of the reciprocal comparative scale and triangular fuzzy numbers.

Numerical Value (Saaty)	Linguistic Term	Triangular Fuzzy Number (l, m, u)	Reciprocal Value	Reciprocal TFN (l, m, u)
1	Neutral analytical ratio	(1, 1, 1)	1	(1, 1, 1)
3	Moderate dominance	(2, 3, 4)	1/3	(1/4, 1/3, 1/2)
5	Strong dominance	(4, 5, 6)	1/5	(1/6, 1/5, 1/4)
7	Very strong dominance	(6, 7, 8)	1/7	(1/8, 1/7, 1/6)
9	Extreme dominance	(9, 9, 9)	1/9	(1/9, 1/9, 1/9)

Note: This table gives analytical scale values and TFNs, not the questionnaire response wording. The value 1 is the neutral numeric ratio assigned to code 0 on either side; "No Impact" was the form label for that code. Zero is used only on the matrix diagonal. Table A11 maps each selectable option to its two entries.

For example, with Port Infrastructure on the left and Product Supply and Demand on the right, a right-hand code 1 mark yields 1/3 in the left-to-right entry and 3 in the reverse entry. Neither entry is a separately elicited measure of an effect of one factor on the other.

The 450 expert–factor-pair records were verified against the questionnaire selections using the coding rule in Table A11. Of the 210 pairs with at least one discrepancy between the questionnaire selections and the archived matrix entries, 64 also violated the reciprocity condition. The analytical matrices are based on the questionnaire selections and their corresponding reciprocal entries. For E3's K2/K9 comparison, the selection documented in the original response corresponds to a forward score of 7 and a reciprocal entry of 1/7. Original response records are kept confidential because they contain participants' personal data; the corresponding anonymized analytical entries are included in the shared research materials. Further verification details are provided in Appendix H, Table A12.

The principal quantitative summary uses the row-plus-column total and row-minus-column balance of the normalized comparison matrix X. This one-step summary retains the most direct connection to the coded selections. The transformation in Equation (9) is a secondary check: it tests whether prominence and sign depend on weighting comparisons by the rest of the comparison structure.

For example, if the normalized entries for A/B and B/C are 0.20 and 0.30, their product contributes 0.06 to the A/C entry of X^2. The product is a model weight and does not establish an empirical transmission pathway.

Agreement in direction and relative standing between X and T supports describing a pattern as insensitive to this particular representation. A material divergence instead identifies a finding that depends on higher-order weighting and should not guide substantive interpretation on its own.

$$a_{ij} = \frac{1}{n}\sum_{k=1}^{n} a_{ij}^{(k)} \tag{1}$$

In Equations (1) and (2), k indexes the experts and n denotes the total number of experts (n = 10), while i and j index the ten factors. Because the analytical sequence fuzzifies individual judgments before aggregation, Equation (1) is reported only as the descriptive equal-weight aggregation of the individual expert judgments; the operative fuzzy aggregation is given in Equation (2). Diagonal elements are set to zero and are excluded from all CFCS extrema.

Each expert judgment was converted to a triangular fuzzy number using Table 5, after which the expert-specific TFNs were aggregated by the equal-weight arithmetic mean. For a positive TFN (l, m, u), the reciprocal scale entry was formed as $(1/u, 1/m, 1/l)$, and the ordering condition l ≤ m ≤ u was checked. This reciprocal transformation defines the scale categories; it does not impose reciprocal equality on the arithmetic group matrix. The TFN scale originates in fuzzy AHP and is applied here to comparative judgments [56,67].

After individual fuzzification, the TFNs were aggregated for each linked comparison under equal expert weights ($w_k = 0.10$). The sensitivity analysis examined the stability of the results under alternative expert-weight configurations.

$$l_{ij} = \frac{1}{n}\sum_{k=1}^{n} l_{ij}^{(k)},\ m_{ij} = \frac{1}{n}\sum_{k=1}^{n} m_{ij}^{(k)},\ u_{ij} = \frac{1}{n}\sum_{k=1}^{n} u_{ij}^{(k)}\ (n = 10) \tag{2}$$

The aggregated fuzzy matrix was defuzzified using the CFCS procedure [68]. The off-diagonal minimum lower bound and maximum upper bound were used to define the global range; diagonal zeros were excluded so that self-influence did not determine the scaling interval. The CFCS steps were normalization, calculation of left and right normalized values, calculation of the total normalized crisp value, and rescaling to the original range. From Equation (3) onward, p denotes the total number of factors ($p = 10$), while i and j denote the corresponding factor indices.

$$x_{ij}^{l} = \frac{l_{ij} - \Delta_{min}}{\Delta}, x_{ij}^{m} = \frac{m_{ij} - \Delta_{min}}{\Delta}, x_{ij}^{u} = \frac{u_{ij} - \Delta_{min}}{\Delta}, \Delta = \Delta_{max} - \Delta_{min} \tag{3}$$

$$x_{ij}^{ls} = \frac{x_{ij}^{m}}{1 + x_{ij}^{m} - x_{ij}^{l}}, x_{ij}^{rs} = \frac{x_{ij}^{u}}{1 + x_{ij}^{u} - x_{ij}^{m}} \quad (4)$$

$$x_{ij}^{c} = \frac{x_{ij}^{ls}\left(1 - x_{ij}^{ls}\right) + \left(x_{ij}^{rs}\right)^{2}}{1 - x_{ij}^{ls} + x_{ij}^{rs}} \quad (5)$$

$$z_{ij} = \Delta_{min} + x_{ij}^{c}\Delta \quad (6)$$

The CFCS range is defined by the minimum lower bound and maximum upper bound across the off-diagonal cells of the aggregated fuzzy matrix. Diagonal zeros are excluded from these extrema. Rows and columns retain the orientation of the comparative entries (see Table 6); their totals are inputs to the exploratory transformation rather than separately elicited outgoing and incoming effects.

Table 6. Defuzzified (CFCS) comparison matrix (Z).

Performance Criteria	K1	K2	K3	K4	K5	K6	K7	K8	K9	K10	Row Total
K1	0.000	4.638	1.724	2.201	3.965	1.518	2.595	2.395	2.399	2.196	23.630
K2	3.680	0.000	3.081	5.521	3.962	4.168	2.595	3.760	5.320	4.644	36.731
K3	6.005	4.849	0.000	2.689	3.106	3.379	2.405	3.294	3.766	2.399	31.892
K4	5.527	2.592	4.858	0.000	3.085	3.278	3.288	1.722	3.764	2.892	31.006
K5	3.965	4.165	4.251	4.647	0.000	4.051	4.055	3.091	2.208	3.569	34.002
K6	5.999	3.764	3.779	4.846	3.293	0.000	3.282	2.198	3.097	3.078	33.335
K7	5.317	5.317	5.123	4.446	3.098	4.644	0.000	3.078	3.281	3.959	38.261
K8	5.524	4.371	4.048	5.593	4.449	5.536	5.051	0.000	3.283	3.075	40.931
K9	5.320	2.399	3.971	4.168	4.926	4.251	4.849	4.649	0.000	2.198	36.730
K10	5.530	3.282	5.320	4.455	3.974	5.051	4.365	5.253	5.733	0.000	42.963
Column Total	46.867	35.375	36.155	38.565	33.858	35.877	32.484	29.441	32.850	28.008	

The defuzzified comparison matrix Z was normalized in Table 7. The maximum row total (42.963) and maximum column total (46.867) were calculated, and the reciprocal of the larger value was used as the scaling coefficient. Diagonal values remained zero. The resulting matrix satisfied the convergence condition $\rho(X) < 1$.

$$x_{ij} = \frac{z_{ij}}{\max\left\{\max\limits_{1\le i\le p}\sum_{j=1}^{p} z_{ij}, \max\limits_{1\le j\le p}\sum_{i=1}^{p} z_{ij}\right\}} \quad (7)$$

Table 7. Normalized comparison matrix (X).

Performance Criteria	K1	K2	K3	K4	K5	K6	K7	K8	K9	K10
K1	0.000	0.099	0.037	0.047	0.085	0.032	0.055	0.051	0.051	0.047
K2	0.079	0.000	0.066	0.118	0.085	0.089	0.055	0.080	0.114	0.099
K3	0.128	0.103	0.000	0.057	0.066	0.072	0.051	0.070	0.080	0.051
K4	0.118	0.055	0.104	0.000	0.066	0.070	0.070	0.037	0.080	0.062
K5	0.085	0.089	0.091	0.099	0.000	0.086	0.087	0.066	0.047	0.076
K6	0.128	0.080	0.081	0.103	0.070	0.000	0.070	0.047	0.066	0.066
K7	0.113	0.113	0.109	0.095	0.066	0.099	0.000	0.066	0.070	0.084
K8	0.118	0.093	0.086	0.119	0.095	0.118	0.108	0.000	0.070	0.066
K9	0.114	0.051	0.085	0.089	0.105	0.091	0.103	0.099	0.000	0.047
K10	0.118	0.070	0.114	0.095	0.085	0.108	0.093	0.112	0.122	0.000

A 10 × 10 identity matrix I was used to form I-X. The resulting matrix $(I - X)^{-1}$ is presented in Table 8; it existed for the baseline and every sensitivity scenario.

$$(I - X)_{ij} = I_{ij} - x_{ij}, i, j = 1, \ldots, p \quad (8)$$

Table 8. Inverse of the matrix $I - X$.

Performance Criteria	K1	K2	K3	K4	K5	K6	K7	K8	K9	K10
K1	1.193	0.243	0.187	0.208	0.224	0.182	0.190	0.175	0.189	0.168
K2	0.367	1.226	0.291	0.349	0.297	0.307	0.261	0.263	0.314	0.272
K3	0.366	0.291	1.192	0.262	0.251	0.258	0.226	0.227	0.254	0.205
K4	0.351	0.244	0.282	1.198	0.243	0.249	0.236	0.193	0.248	0.208
K5	0.350	0.293	0.294	0.314	1.200	0.286	0.269	0.234	0.240	0.239
K6	0.376	0.278	0.276	0.308	0.259	1.197	0.247	0.211	0.248	0.223
K7	0.404	0.338	0.331	0.334	0.285	0.319	1.210	0.254	0.282	0.265
K8	0.426	0.335	0.327	0.371	0.323	0.350	0.322	1.202	0.293	0.260
K9	0.392	0.276	0.302	0.320	0.310	0.303	0.298	0.274	1.203	0.224
K10	0.445	0.328	0.364	0.364	0.329	0.355	0.323	0.318	0.350	1.207

The transformed comparison matrix T was calculated as $X(I - X)^{-1}$. When the spectral radius of X is below one, this equals $X + X^2 + X^3 + \ldots$. The higher-order terms are algebraic weights under the chosen representation, not observed indirect effects.

$$T = X(I - X)^{-1} \tag{9}$$

For factor i, D is the row sum of T and R the corresponding column sum. $D + R$ is termed prominence and $D - R$ the net comparative index. The indices summarize the matrix and do not measure causal effects or establish an intervention order.

$$D_i = \sum_{j=1}^{p} T_{ij}, i = 1, \ldots, p; \quad R_j = \sum_{i=1}^{p} T_{ij}, j = 1, \ldots, p \tag{10}$$

$$Prominence_i = D_i + R_i, Net\ comparative\ index_i = D_i - R_i \tag{11}$$

Positive and negative $D - R$ values indicate which modeled marginal total is greater. A near-zero value indicates similar row and column totals in T.

3.4. Sensitivity and Robustness Analyses

The sensitivity analyses used the same ten expert matrices. Each run recomputed TFN conversion, aggregation, defuzzification, normalization, and the D, R, $D + R$, and $D - R$ indices. Agreement was assessed using one consistently oriented entry per unordered factor pair: $i < j$, with the lower-numbered factor defining the forward orientation. This gives 45 entries per expert without counting the reverse orientation as an additional observation. Because pairs share factors and the panel was purposively selected, the agreement coefficients are descriptive; no inferential tests or ICC confidence intervals are reported. Calculations used double precision without intermediate reporting rounding.

Tie-corrected Kendall's W [69], absolute-agreement ICC point estimates [70–72], and pairwise Spearman and Kendall tau-b coefficients summarize the oriented comparisons. Natural logarithms were used for ICC and log-distance diagnostics, and negative ICC estimates were retained. Appendix I verifies that reconstructing the forward-oriented comparisons from the reciprocal entries reproduces the agreement estimates. Expert-level leave-one-out and weight analyses examine panel heterogeneity, as the small stakeholder strata do not support group-specific inference.

Ten leave-one-expert-out scenarios and 10,000 non-parametric expert-bootstrap resamples were calculated [73]. Each bootstrap draw selects ten complete expert matrices with replacement, preserving all entries within a selected expert's record. Repeated selections give that expert proportionally greater representation. Percentile intervals and sign frequencies describe variation under resampling of this observed panel, not population probabilities or uncertainty about objective causal effects. They are reported separately from parameter-perturbation scenarios.

Equal expert weights of 0.10 form the baseline. Deterministic scenarios set one expert's weight to 0.05, 0.15, or 0.20 and rescale the other nine proportionally. Separate Dirichlet simulations use concentration parameters of 100, 20, and 5, with 5000 draws per level; these are design choices, not estimated priors. TFN width multipliers of 0.50, 0.75, 1.00, 1.25, and 1.50 expand or contract positive-side triangles before reciprocal scale entries are regenerated. CFCS is compared with COG and GMIR. Appendix J gives the exact transformations and scenario counts. Appendix I compares arithmetic and component-wise geometric aggregation [74]; the latter preserves reciprocity only when the input comparisons themselves satisfy it.

The combined collection contains 6165 scenarios: 15 equal-weight combinations, 150 leave-one-out combinations, and 6000 sampled Dirichlet stress combinations. The last category constitutes 97.32% of the collection; pooled retention percentages therefore depend strongly on that chosen representation. They are descriptive scenario frequencies, not general robustness probabilities. The computations used Python 3.12.14 and NumPy 2.3.5. Bootstrap, Dirichlet, and combined simulations used seeds 20260904, 20260906, and 20260907, respectively. Access arrangements for the computational materials are specified in the Data Availability Statement.

4. Results and Discussion

4.1. Critical Factors

The qualitative analysis identified ten factors covering ports, vessels, coordination, policy, personnel, subsidies, intermodal connections, cargo, governance, and regional logistics development. Table 4 relates the original instrument labels to their analytical descriptions and supporting evidence. Appendix G presents the retained questions and contextualized coding examples. The factor domains are supported by these accounts, while the interpretation of K9 and the expanded capacity dimensions of K10 remain qualified by the wording presented to experts.

4.2. Comparative Results and Fuzzy DEMATEL Diagnostic

The largest row and column totals of the coded-comparison matrix Z are 42.963 for K10 and 46.867 for K1, respectively. Table 9 reports the transformed matrix T, and Table 10, Panel B, reports its row and column indices. In the principal one-step comparison (Table 10, Panel A), K2 and K10 have the largest row-plus-column totals (1.539 and 1.514). K10 and K8 have the largest positive balances (+0.319 and +0.245), and K1 has the most negative balance (−0.496). The transformed summary shows the same pattern.

Table 9. Transformed comparison matrix (T).

Performance Criteria	K1	K2	K3	K4	K5	K6	K7	K8	K9	K10
K1	0.193	0.243	0.187	0.208	0.224	0.182	0.190	0.175	0.189	0.168
K2	0.367	0.226	0.291	0.349	0.297	0.307	0.261	0.263	0.314	0.272
K3	0.366	0.291	0.192	0.262	0.251	0.258	0.226	0.227	0.254	0.205
K4	0.351	0.244	0.282	0.198	0.243	0.249	0.236	0.193	0.248	0.208
K5	0.350	0.293	0.294	0.314	0.200	0.286	0.269	0.234	0.240	0.239
K6	0.376	0.278	0.276	0.308	0.259	0.197	0.247	0.211	0.248	0.223
K7	0.404	0.338	0.331	0.334	0.285	0.319	0.210	0.254	0.282	0.265
K8	0.426	0.335	0.327	0.371	0.323	0.350	0.322	0.202	0.293	0.260
K9	0.392	0.276	0.302	0.320	0.310	0.303	0.298	0.274	0.203	0.224
K10	0.445	0.328	0.364	0.364	0.329	0.355	0.323	0.318	0.350	0.207

The largest off-diagonal entries of T were K10/K1 (0.445), K8/K1 (0.426), K7/K1 (0.404), and K9/K1 (0.392). Their common column identifies a feature of the transformed comparison profile. Qualitative statements about inland connections and cargo availability are considered separately in Section 4.3. Tables 10 and A16 report both summaries; the comparison does not assume that T is the more accurate measure of program performance.

Table 10. Principal one-step comparative summaries (Panel A) and secondary transformed indices (Panel B).

Panel A. One-step totals, balances, and ranks of X.			
Factor	**One-Step Total (X)**	**One-Step Balance (X)**	**One-Step Rank**
K1	1.50419	−0.49581	4
K2	1.53852	+0.02894	1
K3	1.45190	−0.09096	9
K4	1.48442	−0.16130	7
K5	1.44792	+0.00309	10
K6	1.47677	−0.05424	8
K7	1.50948	+0.12325	3
K8	1.50151	+0.24515	5
K9	1.48461	+0.08279	6
K10	1.51430	+0.31909	2

Panel B. Baseline transformed indices of T, using the original factor labels.					
Performance Criteria	***Di***	***Ri***	***Di + Ri***	***Di − Ri***	**Index Position**
Port Infrastructure (K1)	1.958	3.671	5.628	−1.7132	Negative
Fleet Availability (K2)	2.946	2.852	5.798	0.0933	Boundary
Coordination between Stakeholders (K3)	2.533	2.846	5.379	−0.3129	Negative
Government Policies and Regulations (K4)	2.452	3.029	5.481	−0.5767	Negative
Human Resources (K5)	2.720	2.721	5.441	−0.0010	Boundary
Effectiveness of Government Subsidies (K6)	2.624	2.806	5.430	−0.1829	Negative
Integration of Transport Modes (K7)	3.022	2.581	5.603	0.4408	Positive
Product Supply and Demand (K8)	3.209	2.351	5.561	0.8583	Positive
Corruption and Transparency (K9)	2.903	2.621	5.524	0.2816	Positive
Improving Logistics Activities in Underdeveloped Regions (K10)	3.384	2.271	5.656	1.1128	Positive

Note: Panel A ranks are descending row-plus-column totals of X; balances are row-minus-column totals. In Panel B, "Boundary" identifies K2 and K5 because their baseline net comparative indices are near zero and their signs vary in the sensitivity analyses. Positive and negative labels, including the markers in Figure 2, describe the baseline arithmetic computation only, not stable factor classes. All ten 95% expert-bootstrap intervals for $D - R$ include zero.

The transformed D − R index is positive for K10 (1.11279), K8 (0.85825), K7 (0.44078), K9 (0.28155), and K2 (0.09334), and negative for K1 (−1.71324), K4 (−0.57666), K3 (−0.31290), K6 (−0.18292), and K5 (−0.00099). K2 and K5 are near the sign boundary; their baseline signs do not establish a stable directional classification.

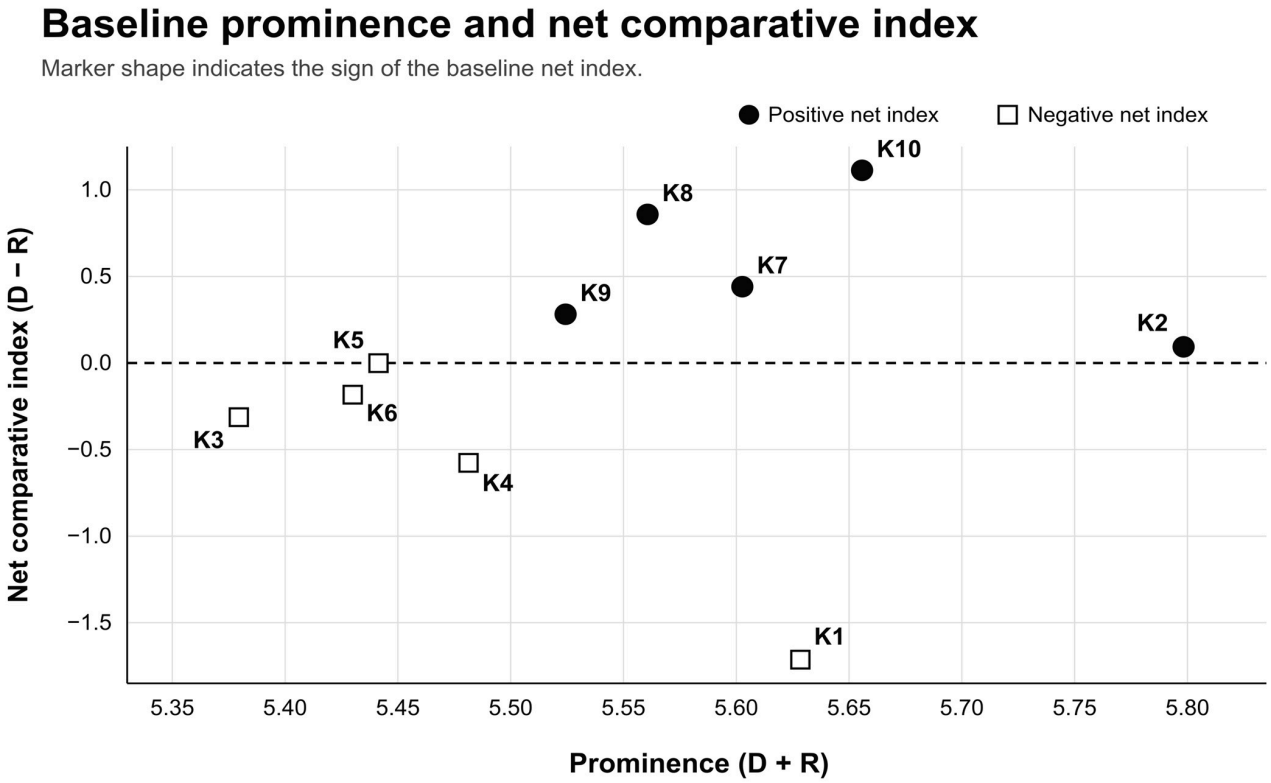


Figure 2. Baseline prominence and net comparative index map. Filled circles indicate positive $D - R$; open squares indicate negative $D - R$. K2 and K5 lie close to the zero boundary. Marker signs refer to the baseline computation; all ten 95% expert-bootstrap intervals include zero.

In the transformed arithmetic model, K2 (5.798), K10 (5.656), K1 (5.628), and K7 (5.603) have the four highest $D + R$ values. Scores span 5.379–5.798, and K5 and K6 differ by only 0.011; adjacent ranks should not be read as substantively distinct priorities. K10 has the largest D (3.384), and K1 has the largest R (3.671).

The one-step and transformed rankings are closely aligned (Spearman = 0.952; Table A16), and nine of ten signs agree. K2 remains first, while K1 moves from fourth to third and K5 from tenth to eighth. K5 changes from +0.00309 in X to −0.00099 in T, making its sign sensitive to representation. The transformed results are therefore useful as a check on the one-step findings, not as a basis for giving greater weight to small rank or sign differences.

Sensitivity Analysis Findings

Numerical checks confirmed that the normalized matrices satisfied $\rho(X) < 1$ and that $I - X$ was invertible throughout the baseline and sensitivity analyses. D and R were verified against the row and column totals of T, respectively; numerical comparisons used unrounded values.

Table 11 reports descriptive agreement across the 45 consistently oriented stored comparisons. Kendall's W summarizes ordering similarity, while ICC point estimates summarize absolute agreement on log-transformed values. No inferential tests or confidence intervals are attached to these coefficients. The panel is treated as the observed purposive group, and

the comparisons are not assumed to be statistically independent simply because duplicated reverse entries are excluded.

Table 11. Descriptive agreement across 45 consistently oriented stored comparisons.

Descriptive Metric	Estimate	Stored Entries	Orientation
Kendall W	0.0685	45 per expert	$i < j$
ICC (2,1), ln	−0.0272	45 per expert	$i < j$
ICC (2,K), ln	−0.3595	45 per expert	$i < j$

For the $i < j$ orientation, Kendall's W was 0.0685, ICC(2,1) was −0.0272, and ICC(2,K) was −0.3595. These estimates indicate weak common ordering and poor absolute agreement within the observed panel, not a negative amount of reliability to be interpreted literally. The reconstructed orientation gives identical estimates, as expected from the reciprocal coding (Appendix I). Accordingly, the arithmetic aggregate is interpreted as an average of heterogeneous judgments rather than a consensus matrix.

Leave-one-out results show marked expert sensitivity. K2 turned negative when E5, E8, or E9 was omitted; K5 turned positive when E2, E5, E7, E8, or E9 was omitted. K6 and K9 changed sign in three and four cases, respectively. K1, K3, K4, K7, K8, and K10 retained their baseline signs in all ten leave-one-out scenarios. Bootstrap resampling provides the broader expert-composition check shown in Figure 3.

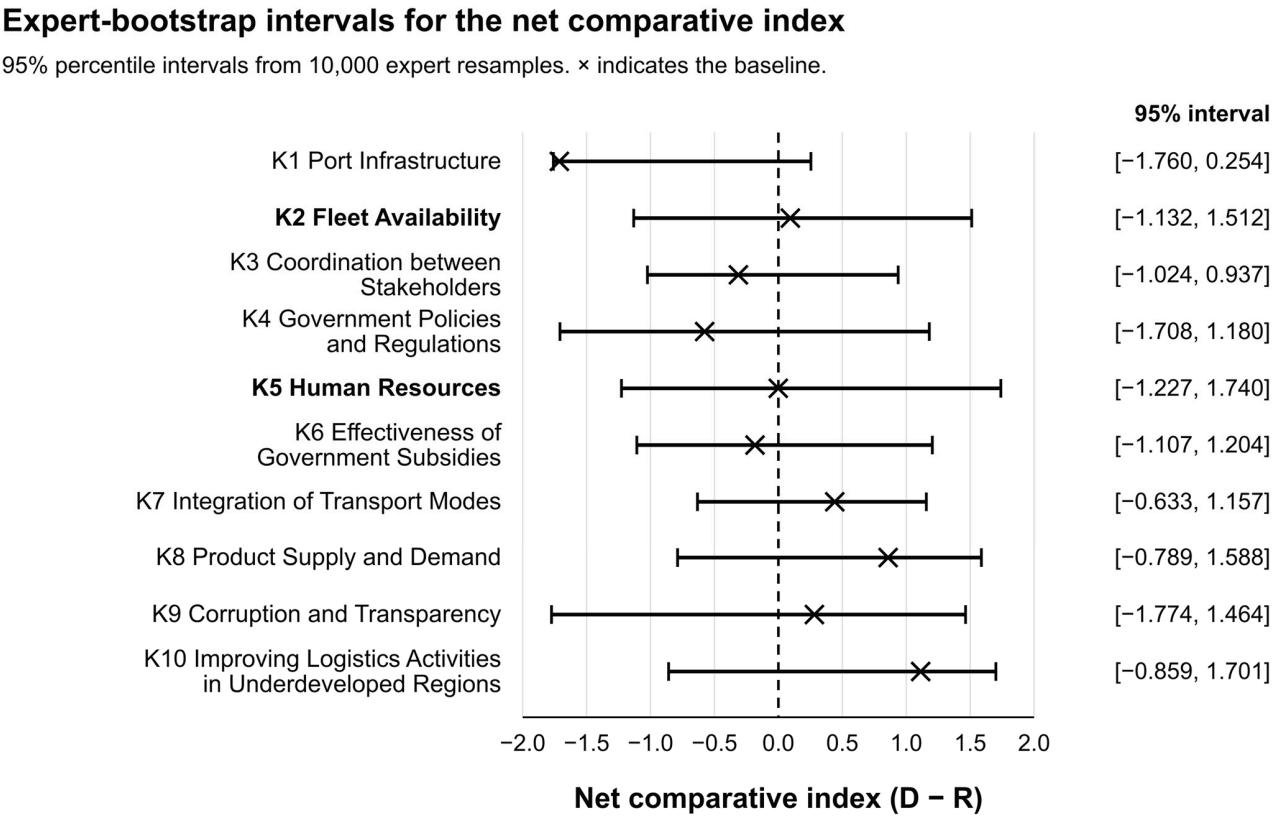


Figure 3. Expert-bootstrap 95% percentile intervals for the net comparative index $(D - R)$, based on 10,000 resamples.

All ten expert-bootstrap 95% percentile intervals include zero. K1 has the most frequently negative index (95.66%; interval [−1.760, 0.254]). Positive indices occur in 53.31% of resamples for K2 and 50.05% for K5, consistent with their boundary positions. Positive

frequencies are 79.57% for K7, 87.06% for K8, 60.22% for K9, and 86.75% for K10. These frequencies describe resampling of the observed expert panel.

TFN-width and defuzzification scenarios preserved all baseline signs and $D + R$ ranks within their specified ranges. In the deterministic expert-weight scenarios, K2, K5, and K9 changed signs in 11, 15, and four of 30 cases, respectively; K6 changed in two cases, and K3 and K4 each changed once. Geometric aggregation adds a distinct sensitivity result: K5 changes sign and the $D + R$ rank correlation with the arithmetic model is 0.685 (Appendix I). Thus, stability to TFN width and defuzzification should not be generalized to all modeling choices. Expert composition, relative weighting, and aggregation must be considered separately.

Table 12 separates the three combined scenario families in Appendix J. Each retention percentage is the share of cases in that family with the baseline $D - R$ sign; each rank range is the minimum to maximum $D + R$ rank within that family. Bootstrap resampling and geometric aggregation are reported separately in Appendices C and I, respectively.

Table 12. Baseline sign retention and prominence-rank ranges by scenario family.

Factor	Baseline $D - R$	Equal Weight Sign Retained	Equal Weight $D + R$ Ranks	Omission Sign Retained	Omission $D + R$ Ranks	Dirichlet Stress Sign Retained	Dirichlet Stress $D + R$ Ranks
K1	−1.71324	100.00%	3–3	100.00%	2–4	99.98%	1–6
K2	0.09334	100.00%	1–1	70.00%	1–1	58.75%	1–4
K3	−0.31290	100.00%	10–10	100.00%	8–10	89.03%	5–10
K4	−0.57666	100.00%	7–7	100.00%	6–9	93.27%	3–10
K5	−0.00099	100.00%	8–8	50.00%	7–10	50.02%	3–10
K6	−0.18292	100.00%	9–9	70.00%	7–10	75.65%	5–10
K7	0.44078	100.00%	4–4	100.00%	2–5	97.62%	1–10
K8	0.85825	100.00%	5–5	100.00%	3–7	98.98%	1–10
K9	0.28155	100.00%	6–6	60.00%	4–9	72.07%	1–10
K10	1.11279	100.00%	2–2	100.00%	2–5	99.42%	1–7

All signs and ranks are retained in the 15 equal-weight parameter cases. Sign retention differs under expert omission and Dirichlet stress. In the stress cases, K1 and K10 retain their signs in 99.98% and 99.42% of cases, respectively, while K2 and K5 retain theirs in 58.75% and 50.02%. K7 and K8 retain their signs in more than 97% of stress cases but span ranks 1–10. K9 also spans ranks 1–10 and retains its sign in 72.07% of these cases. Sign retention therefore does not establish rank stability.

4.3. Discussion

The stakeholder accounts and prior literature support considering Sea Toll performance across regional, operational, and institutional domains. The discussion distinguishes comparative model results, relationships reported by stakeholders, and mechanisms proposed in prior research.

K10 has the largest positive baseline $D - R$ and is positive in 86.75% of expert-bootstrap resamples, although its two-sided interval crosses zero. Its original wording refers broadly to improving logistics activities in underdeveloped regions. Amin et al. [50] model feedback between local economic activity, cargo, and maritime costs, while Amin et al. [47] connect maritime logistics and island development. These studies provide a literature-based explanation to investigate, not a mechanism demonstrated by the present comparisons. Storage, consolidation, and distribution capacity remain possible elaborations of K10 rather than dimensions shown to have been uniformly assessed by all experts.

K8 has a positive baseline $D - R$ and is positive in 87.06% of bootstrap resamples, although its interval includes zero. The qualitative archive directly links a lack of return cargo to a lack of local commodities (A1K3; Appendix G). This is consistent with the return-cargo constraints discussed by Sa'adah et al. [26] and Gugat et al. [39]. Sofiyandi et al. [48] additionally show that subsidy outcomes vary with complementary conditions. Together, the qualitative and published evidence motivates testing whether production and cargo consolidation improve route economics; the DEMATEL transformation does not estimate the magnitude or timing of that relationship.

K7's positive baseline index and K1's negative index can be considered alongside port-regionalization and hinterland-integration research [75,76]. A qualitative respondent also explicitly linked port improvements to better inland transport (A1K1; Appendix G). These are independent grounds for examining port and hinterland conditions jointly. K1's negative $D - R$ alone does not demonstrate dependence on feeder services, cargo, or governance. K1 remains predominantly negative in the resamples (95.66%), although its 95% interval includes zero.

K2 has the highest prominence under arithmetic aggregation but a near-zero $D - R$ and an approximately even bootstrap sign split. Its prominence rank is 3 under geometric aggregation. Thus, its ranking is sensitive to how heterogeneous judgments are combined, while neither sign provides a stable basis for a fixed directional role. K5 similarly lies near zero and changes sign under the geometric diagnostic. These factors warrant assessment against route-specific vessel and workforce indicators rather than assignment to an established cause or outcome category.

All expert-bootstrap intervals include zero, so baseline signs should not be read as uniform expert support. K9's positive index concerns the original label 'Corruption and Transparency'. Research on port governance [45] and Indonesian digital logistics [46] supports investigating transparency and information visibility; it does not resolve how every expert interpreted that combined label. Excluding K9 changes the signs of K2 and K5, while the other seven retained factors keep their signs (Table A17). This indicates that the ambiguous item affects the two boundary classifications without resolving the meaning of K9.

4.4. Theoretical and Analytical Contributions

The study contributes an empirical factor framework and a qualified comparative assessment to cumulative-causation and supply-chain-integration discussions. Myrdal's framework suggests that weak production, thin cargo, and high logistics costs can reinforce regional disadvantage [15]. The prominence of regional logistics development and cargo in the archived assessment helps identify relevant topics for empirical testing, without confirming that theory. Longitudinal regional and route-level studies should examine whether stronger local logistics activity is associated with improved cargo balance, whether cargo balance precedes higher route utilization and lower transport costs, and whether those changes improve accessibility.

From an SCM perspective, the qualitative factor set places ports, vessels, inland connections, cargo, information, institutions, and subsidies within a common program-performance discussion [30,31]. This contextual contribution does not depend on treating $D - R$ as an observed causal mechanism. The aggregation-sensitive prominence of K2 and the negative model position of K1 caution against deriving intervention priorities from a single matrix ranking. Whether physical resources function as conversion capacities whose outcomes depend on cargo and coordination remains a literature-informed proposition requiring objective performance data.

Analytically, the study illustrates why elicitation design, agreement, and aggregation diagnostics must accompany an expert-based transformation. TFN-width and defuzzification stability coexist with weak descriptive agreement, expert-resampling uncertainty, and changes under geometric aggregation. Reporting those differences preserves the distinction between a reproducible computation and a validated measurement of directed effects. The archived comparative instrument is not proposed as a substitute for standard independent direct-influence DEMATEL elicitation.

4.5. Practical Contributions

The practical contribution is a set of literature-informed, testable policy proposals organized around the stakeholder-derived factors. Matrix positions identify topics requiring scrutiny rather than an empirically established sequence of interventions. The comparative summaries highlight K1, K10, and K8, but their expert-bootstrap intervals include zero. These properties do not demonstrate which investment should occur first. Table 13 links the proposals to relevant factors and measurable indicators, with their evidential status stated explicitly.

Table 13. Staged policy hypotheses, related factors, and monitoring indicators.

Horizon	Proposed Action	Related Factors	Monitoring Indicators	Evidential Status
Short term	Common route-level indicator set; data-sharing standards; subsidy trigger conditions	K9, K6	Load factor, return-cargo ratio, schedule reliability, subsidy per unit carried, data completeness	Enabling measures; no causal assumption required
Medium term	Cargo consolidation, storage, and feeder-connectivity pilots in selected lagging regions	K10, K8, K7	Return-cargo share, delivery time, price dispersion, pilot vs. comparison routes	Literature-informed pilots; comparative signs do not establish sequencing
Long term	Port investment appraised jointly with cargo and hinterland conditions; shared performance dashboard; adaptive subsidy rules	K1, K4, K3	Port turnaround, route continuity, cost per unit, regional accessibility	Joint appraisal supported by qualitative and published evidence; no dependency inferred from $D - R$
Continuous	Threshold-based monitoring of fleet and workforce	K2, K5	Vessel availability vs. demand, cancellations, documented skill gaps	Boundary factors; direction unresolved

Short term: measurement and information baseline. The most defensible immediate actions are those that require limited capital and generate the evidence needed to test the expert pattern. First, a common route-level indicator set should be defined and reported for every Sea Toll route. It should include outbound and return load factors, the return-cargo ratio, schedule reliability, port waiting and turnaround time, subsidy per tonne or TEU carried, and destination price dispersion for designated goods. Second, data-sharing standards should specify which cargo, schedule, tariff, and subsidy records operators, ports, and local governments submit, and at what frequency. The existing SITOLAUT platform [41] provides a natural basis for this, and it would operationalize the information visibility discussed under the analytical interpretation of K9 [45,46]. Third, explicit subsidy trigger conditions should be specified, for example minimum return-cargo thresholds or service-reliability criteria that prompt route review. These conditions would make subsidy

effectiveness (K6) verifiable rather than assumed [48]. Indicators should be linked to explicit decision purposes and should reflect the perspectives of different stakeholders [36,37].

Medium term: route-level pilots. The qualitative accounts and published regional-logistics evidence motivate pilots that strengthen cargo consolidation, storage, and local production linkages in selected lagging regions. Pilots should be combined with feeder and inland connectivity improvements (K7), because return-cargo constraints and weak hinterland links have been documented on several routes [26,27,39]. Pilot routes should be compared with similar non-pilot routes, and their indicators should be assessed before and after implementation. The hypothesis being tested is whether improvements in regional capacity and cargo balance precede gains in load factor, return-cargo share, delivery time, and price dispersion. Because the expert-bootstrap intervals for K10 and K8 include zero, pilot outcomes, rather than $D - R$ values, should determine whether the approach is scaled up.

Long term: structural and institutional alignment. Joint appraisal of port investment, cargo potential, and inland access is supported by the qualitative account linking port facilities to land transport and by route-level evidence on multimodal and port-adjacent constraints [28]. It is not inferred from K1's negative $D - R$. A shared monitoring dashboard involving the Ministry of Transportation, local governments, ports, carriers, and cargo owners could build on the short-term indicators and support evaluation of coordination and regulatory arrangements [8,10]. These proposals require institutional feasibility assessment and route-level validation.

Adaptive monitoring of boundary factors. Fleet availability (K2) and human resources (K5) should not be assigned fixed priority. Their near-zero $D - R$ values and even bootstrap splits indicate that their role may differ across routes and stakeholder perspectives. They are better addressed through monitoring thresholds, such as vessel availability relative to scheduled demand, voyage cancellations, and documented skill gaps in port, shipboard, and digital-logistics functions. Investment would be triggered where these thresholds are breached.

Overall, the practical contribution is a monitoring-and-learning framework: qualitative findings and published evidence motivate candidate mechanisms, measurement makes them testable, pilots evaluate them, and subsequent priorities depend on operational evidence. The exploratory indices provide contextual comparison, not verification of the mechanisms or their implementation order.

5. Conclusions

This study identifies stakeholder-derived factors relevant to Indonesia's Sea Toll Program and summarizes comparative expert judgments using one-step indices and a secondary Fuzzy DEMATEL analysis. K10 and K8 have the largest positive balances in both the one-step and transformed summaries, and K1 has the most negative balance. K1 is negative in 95.66% of expert-bootstrap samples and K10 is positive in 86.75%; all ten 95% intervals include zero. These conditional model results do not establish directed substantive dependencies. Geometric aggregation preserves nine baseline signs but changes prominence ranks and reverses K5's sign. K2 and K5 are close to the sign boundary in the full model, and both signs reverse when K9 is excluded.

The contribution lies in organizing regional, cargo, connectivity, resource, governance, and institutional concerns within a traceable comparative assessment. The findings support specifying policy hypotheses and monitoring indicators, not an optimal investment order. The proposed short-, medium-, and long-term actions draw on qualitative accounts and prior literature and require validation using route-level operational, cost, price, and regional-development evidence.

Several limitations should be considered when interpreting the findings. Convenience-based online recruitment limited follow-up probing, and no sequential saturation log or complete code-by-case frequency matrix was retained. The expanded analytical descriptions do not establish that all experts evaluated identically defined constructs, particularly for K9 and K10. The small, heterogeneous panel showed weak descriptive agreement. The instrument elicited one marked side and intensity per pair, not two independent directional assessments. The reciprocal coding was applied during analysis, and the questionnaire wording does not establish that a code 0 selection expressed equality. Arithmetic aggregation can also amplify strongly opposing judgments. The indices therefore summarize comparative expert assessments rather than temporal causal effects or policy-treatment impacts. The focus on one national program provides contextual depth but may limit transferability; the proposed relationships require validation against objective route-level performance indicators.

Future research should independently elicit $i- > j$ and $j- > i$ influence on a standard direct-influence Fuzzy DEMATEL scale, recruit larger balanced panels that permit stakeholder and regional subgroup analysis, and pre-register consensus procedures. The proposed relationships should be validated against vessel utilization, return cargo, schedule reliability, port turnaround, subsidy efficiency, price dispersion, and regional production data. Longitudinal system-dynamics, structural-equation, Bayesian, and alternative MCDM models could then test whether the perceived links persist over time and whether proposed policy sequences produce measurable outcomes.

Author Contributions: Conceptualization, N.E. and B.A.; methodology, N.E. and A.E.C.; software, B.A.; validation, N.E. and B.A.; formal analysis, A.E.C. and P.Ç.; investigation, N.E. and B.A.; resources, B.A.; data curation, B.A. and A.E.C.; writing—original draft preparation, N.E., B.A. and P.Ç.; writing—review and editing, N.E., B.A., A.E.C. and P.Ç.; visualization, A.E.C. and P.Ç.; supervision, N.E.; project administration, N.E. and A.E.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki and approved by the Social and Human Sciences Ethics Committee of Akdeniz University (decision no. 384).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The de-identified response-to-matrix mapping, calculation workbooks, and executable reproduction workflow are available for confidential peer review and from the corresponding author upon reasonable request. These materials are not publicly released because, in a small purposive expert panel, individual judgment records read together with participant profiles could allow participants to be identified. Participant profile information beyond that reported in Table 3 and the original open-ended responses beyond the excerpts in Appendix G are restricted on the same grounds, under the conditions of the ethics approval (Akdeniz University Social and Human Sciences Ethics Committee, decision no. 384).

Acknowledgments: This article is derived from the master's thesis of Belva Agustin called "Identification of critical factors affecting the performance of the Sea Toll Programme: An analysis with the DEMATEL method" completed at Akdeniz University. Generative AI tools (Claude Opus 5.5, Grammarly v.1.184.0.0. and GPT-6-1 Sol) were used exclusively for language-related support, including grammar correction and improving clarity and readability in selected sections of the manuscript. No AI tools were used for data analysis, interpretation, or reference generation. All content, conceptual development, and final revisions were performed by the authors, who take full responsibility for the manuscript.

Conflicts of Interest: Author Belva Agustin is employed by the company Andhika Group. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Appendix A. Inter-Expert Agreement and Distributional Findings

This appendix summarizes expert comparison patterns using one consistently oriented entry per unordered factor pair ($i < j$). Table A1 reports the least similar expert pairs, while Table A2 identifies the unique factor pairs with the greatest dispersion on the logarithmic scale. Reverse directions are not counted as additional observations.

Table A1. Expert pairs with the lowest pairwise agreement.

Expert Pair	Spearman	Kendall Tau-b	Mean Absolute ln Difference
E2–E10	−0.484	−0.434	3.329
E4–E9	−0.420	−0.378	2.937
E3–E10	−0.305	−0.293	2.966
E6–E9	−0.282	−0.246	2.403
E1–E5	−0.267	−0.240	1.929
E3–E5	−0.247	−0.223	2.062
E6–E8	−0.235	−0.208	2.278
E4–E5	−0.181	−0.163	2.376
E7–E10	−0.172	−0.167	3.298
E1–E8	−0.170	−0.151	2.504

Note: Expert pairs are ordered by Spearman coefficient across the 45 oriented comparisons. Negative coefficients indicate opposing ordering patterns in the stored data. No significance tests are attached to these descriptive values.

The negative pairwise coefficients in Table A1 show that the weak shared pattern observed in the overall agreement results becomes more pronounced for certain expert pairs. Pairwise correlations were not, however, used on their own as a criterion for excluding an expert; expert influence was instead tested through the recomputation scenarios in Appendix B.

Table A2. Unique oriented factor pairs with the largest expert dispersion.

Pair	Mean	Median	SD(ln)	>1	=1	<1
K7/K10	3.959	3.571	2.213	5	0	5
K2/K5	3.962	3.571	2.187	5	0	5
K1/K2	4.644	7.000	2.185	6	0	4
K2/K8	3.759	2.571	2.182	5	0	5
K1/K5	3.965	3.571	2.161	5	0	5

Note: Each unordered factor pair is included once, oriented from the lower-numbered factor to the higher-numbered factor. Mean and median refer to stored comparative values; SD(ln) is their natural-log standard deviation. The counts >1, =1, and <1 describe the two coded orientations and the neutral numeric category.

Figure A1 presents the expert Spearman matrix for these same 45 oriented stored comparisons. The orientation is fixed by factor codes rather than selected from the results. Shared factors across comparisons and purposive panel selection preclude treating these coefficients as population-level inference.

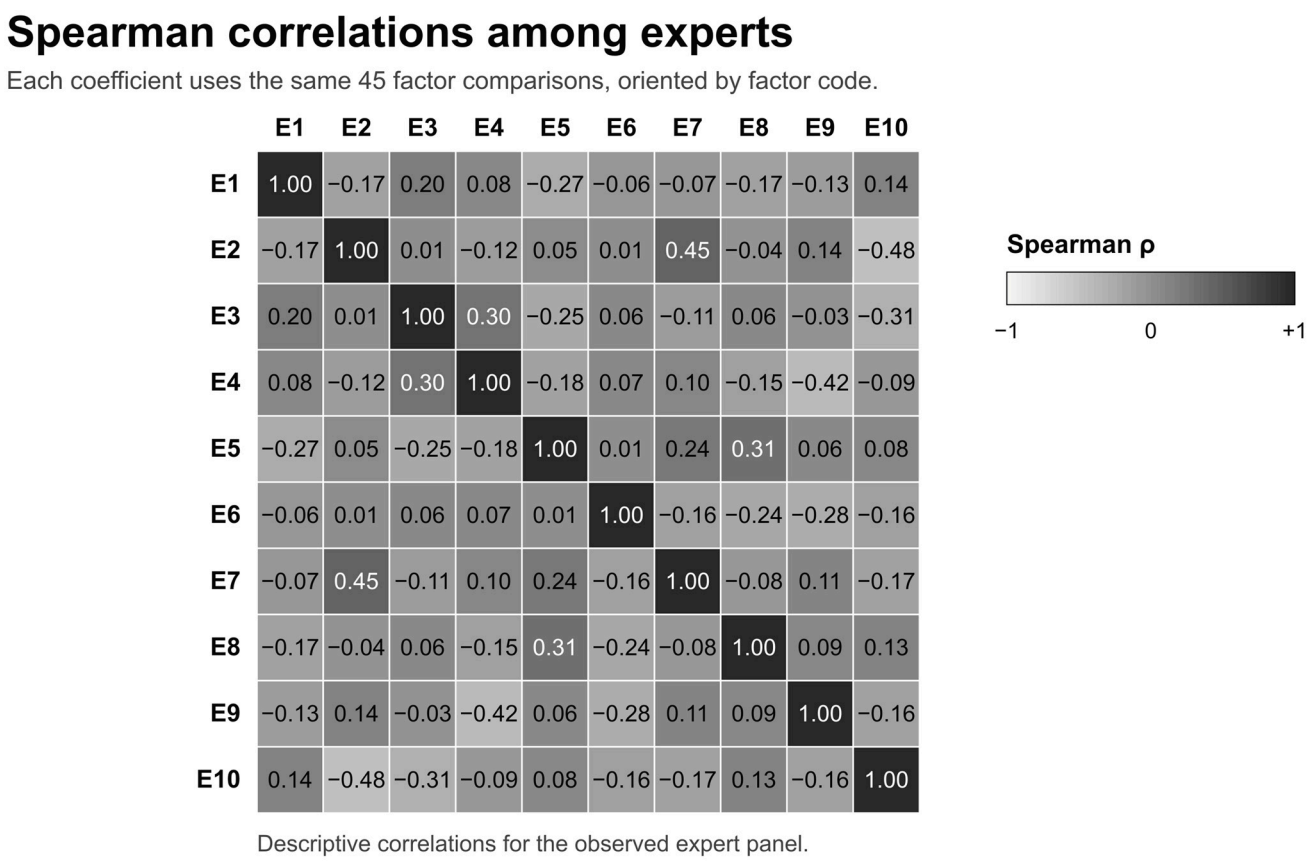


Figure A1. Descriptive Spearman correlations across 45 oriented comparisons.

Appendix B. Leave-One-Expert-Out Sensitivity Analysis

In each row, one expert was removed from the data set and the Fuzzy DEMATEL processing chain was recalculated. In Table A3, sign changes are assessed relative to the sign of the baseline net comparative index.

Table A3. Factor $D - R$ values across all leave-one-expert-out scenarios.

Excluded	K1	K2	K3	K4	K5	K6	K7	K8	K9	K10
E1	−1.667	0.734	−0.042	−0.598	−0.102	−0.311	0.546	0.831	−0.229	0.839
E2	−1.702	0.046	−0.348	−0.184	0.055	0.097	0.430	0.675	0.197	0.733
E3	−1.706	0.326	−0.076	−0.784	−0.451	0.089	0.523	0.773	−0.163	1.469
E4	−1.708	0.442	−0.440	−0.498	−0.430	−0.272	0.573	1.176	−0.085	1.243
E5	−1.692	−0.196	−0.382	−0.545	0.208	−0.374	0.091	0.881	0.716	1.294
E6	−1.730	0.114	−0.276	−0.248	−0.181	−0.141	0.378	0.750	0.333	1.000
E7	−1.704	0.376	−0.161	−0.505	0.250	−0.151	0.495	0.857	−0.156	0.698
E8	−1.591	−0.174	−0.553	−0.834	0.195	−0.140	0.179	0.830	0.465	1.624
E9	−1.723	−0.448	−0.118	−0.482	0.465	0.007	0.517	0.503	0.509	0.770
E10	−1.656	0.021	−0.599	−1.092	−0.073	−0.599	0.601	1.235	0.813	1.350

Note: Because the baseline $D - R$ values of K2 and K5 are close to zero, group classification changes in some leave-one-expert-out scenarios.

Figure A2 reports the range of $D - R$ values across the ten omission scenarios and marks the baseline. Ranges that cross the zero line indicate that the classification is sensitive to expert composition.

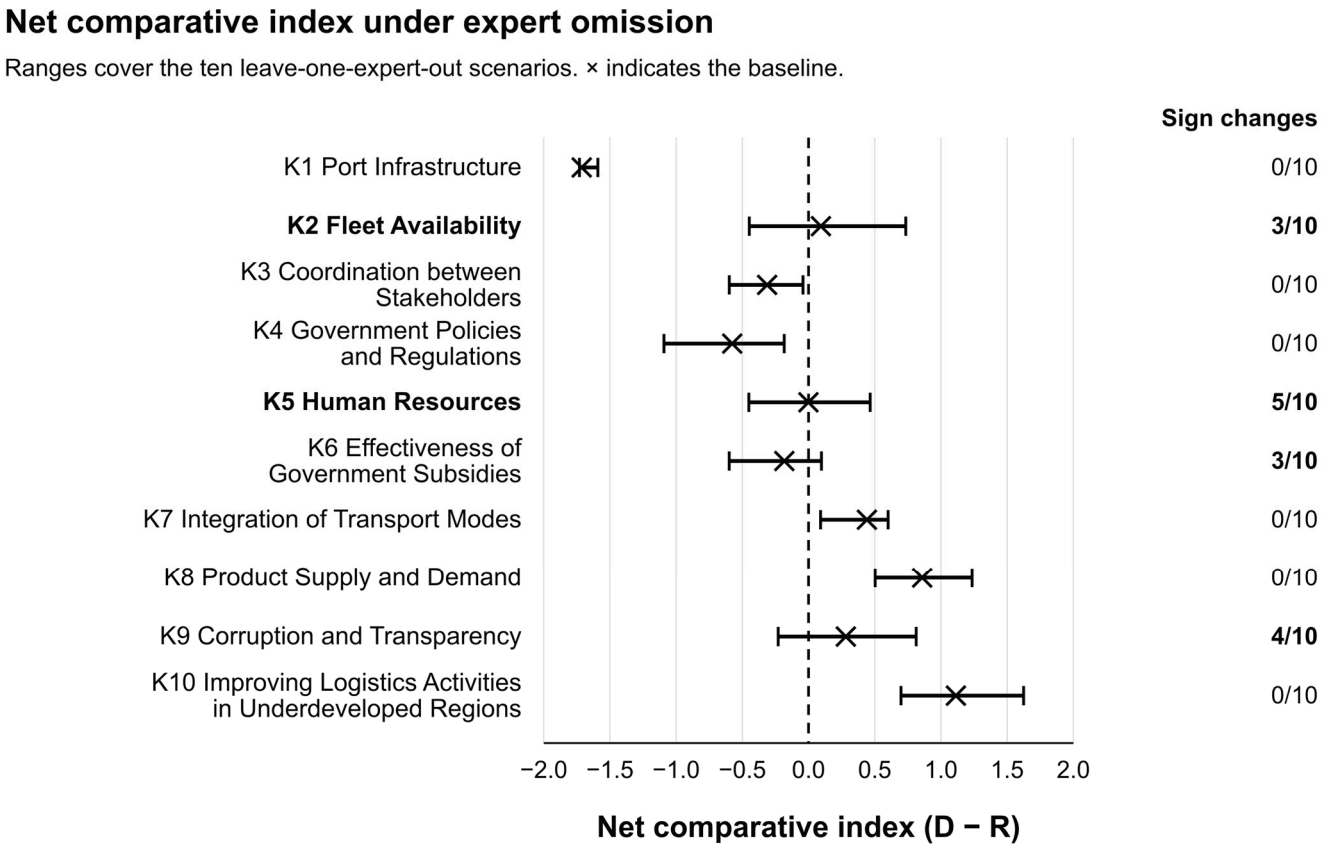


Figure A2. Change in $D - R$ across leave-one-expert-out scenarios.

Appendix C. Bootstrap Robustness Analysis

Table A4 reports summaries of 10,000 resamples drawn with replacement from the observed expert panel. Whole matrices are resampled, retaining all linked entries within each selected expert. Percentile intervals and sign frequencies describe this conditional resampling distribution, not population probabilities of substantive or causal factor classes.

Table A4. Expert-bootstrap percentile intervals, conditional sign frequencies, and $D + R$ rank summaries.

Factor	Baseline $D - R$	95% CI	Positive Resamples (%)	Negative Resamples (%)	$D + R$ Rank 1 (%)	$D + R$ Top 3 (%)	Mean $D + R$ Rank (SD)	Assessment
K1	−1.7132	[−1.760; 0.254]	4.34%	95.66%	8.21%	58.29%	3.43 (1.57)	Predominantly negative; interval includes zero
K2	0.0933	[−1.132; 1.512]	53.31%	46.69%	67.98%	91.97%	1.64 (1.26)	Boundary factor
K3	−0.3129	[−1.024; 0.937]	29.35%	70.65%	0.16%	1.24%	8.52 (1.68)	Interval includes zero
K4	−0.5767	[−1.708; 1.180]	23.72%	76.28%	0.46%	7.56%	6.69 (1.99)	Interval includes zero
K5	−0.0010	[−1.227; 1.740]	50.05%	49.95%	0.02%	2.01%	7.66 (1.76)	Boundary factor
K6	−0.1829	[−1.107; 1.204]	37.22%	62.78%	0.05%	0.67%	8.24 (1.52)	Interval includes zero
K7	0.4408	[−0.633; 1.157]	79.57%	20.43%	5.82%	33.11%	4.65 (2.22)	Frequently positive; interval includes zero

Table A4. *Cont.*

Factor	Baseline $D - R$	95% CI	Positive Resamples (%)	Negative Resamples (%)	$D + R$ Rank 1 (%)	$D + R$ Top 3 (%)	Mean $D + R$ Rank (SD)	Assessment
K8	0.8583	[−0.789; 1.588]	87.06%	12.94%	1.72%	21.92%	5.27 (2.05)	Frequently positive; interval includes zero
K9	0.2816	[−1.774; 1.464]	60.22%	39.78%	5.15%	27.33%	5.48 (2.65)	Interval includes zero
K10	1.1128	[−0.859; 1.701]	86.75%	13.25%	10.43%	55.90%	3.43 (1.78)	Frequently positive; interval includes zero

Note: Percentile intervals are the 2.5th and 97.5th percentiles of the expert-resampling distribution. All ranking statistics use $D + R$; $D - R$ is used for sign summaries only. Positive and negative percentages are frequencies conditional on resampling the observed ten-expert panel, not probabilities of a true causal class.

Figure A3 shows the percentage of resamples with positive $D - R$. Values near 50% indicate sign sensitivity to panel resampling, while values near the extremes indicate greater conditional sign retention. These percentages are not generalized beyond the observed panel and specified model.

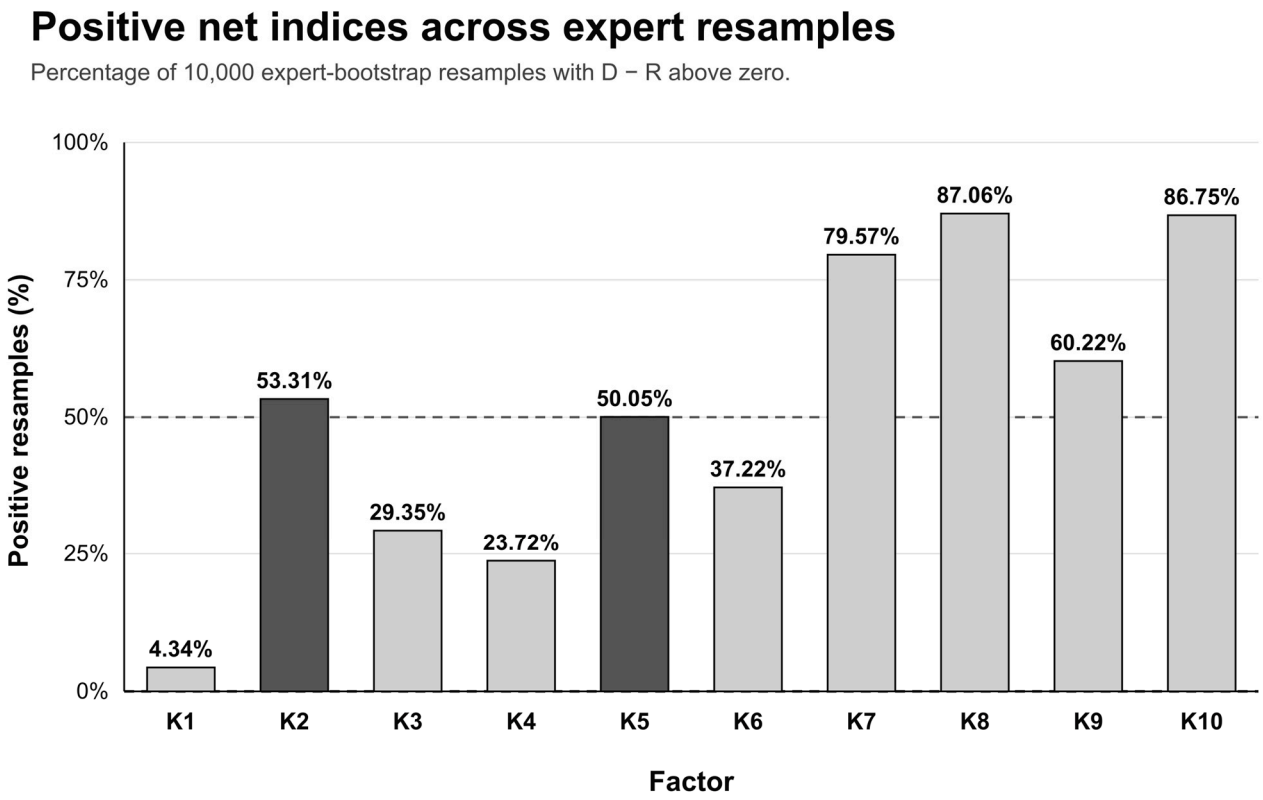


Figure A3. Percentage of observed-panel bootstrap resamples with positive $D - R$.

The $D - R$ distributions for K2 and K5 are presented in detail in Figure A4. Positive $D - R$ occurs in 53.31% of the K2 resamples and 50.05% of the K5 resamples. Both distributions span zero, consistent with their boundary positions.

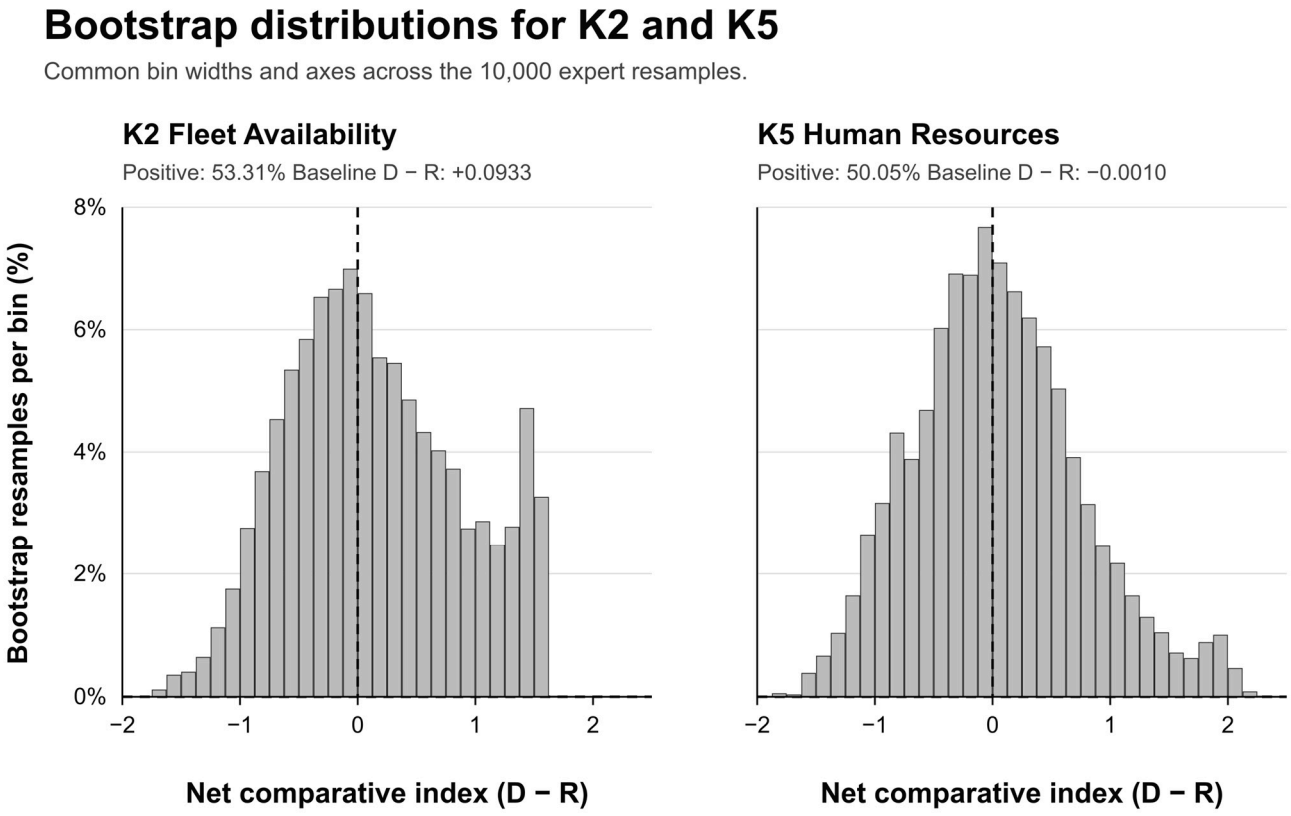


Figure A4. Bootstrap $D - R$ distributions for K2 and K5.

Appendix D. Expert-Weight Sensitivity Analysis

Table A5 summarizes 30 deterministic scenarios: each expert in turn receives a weight of 0.05, 0.15, or 0.20. The remaining nine weights are $(1 - w)/9$, namely 0.105556, 0.094444, or 0.088889. Thus the total remains one. Sign changes are counted relative to the arithmetic baseline; the scenarios are not sampled probabilities.

Table A5. Summary of deterministic expert-weight scenarios.

Factor	Baseline $D - R$	Min	Max	Sign Changes	Positive	Negative
K1	−1.7132	−1.732	−1.660	0	0.0%	100.0%
K2	0.0933	−0.321	0.925	11	63.3%	36.7%
K3	−0.3129	−0.606	0.103	1	3.3%	96.7%
K4	−0.5767	−1.317	0.176	1	3.3%	96.7%
K5	−0.0010	−0.716	0.408	15	50.0%	50.0%
K6	−0.1829	−0.477	0.416	2	6.7%	93.3%
K7	0.4408	0.199	0.721	0	100.0%	0.0%
K8	0.8583	0.293	1.383	0	100.0%	0.0%
K9	0.2816	−0.486	0.619	4	86.7%	13.3%
K10	1.1128	0.581	1.617	0	100.0%	0.0%

Note: The number of sign changes shows how many of the 30 deterministic weight scenarios produced a result different from the baseline classification. K6 also changed from a negative to a positive net comparative index in two of the 30 deterministic expert-weight scenarios; nevertheless, it retained its baseline negative classification in 93.3% of the scenarios.

In addition to the deterministic variations, expert weights were also generated randomly at three different Dirichlet concentration levels. In Table A6, a smaller α parameter indicates more variable weights and, correspondingly, a greater relative influence of expert composition.

Table A6. Factor-level results of the Dirichlet expert-weight simulations.

α	Factor	$D - R$ 95% Range	Mean $D - R$	Positive Draws (%)	Negative Draws (%)	$D + R$ Rank Range	Mean $D + R$ Rank
100	K1	[−1.724; −1.699]	−1.712	0.0%	100.0%	2–4	2.96
100	K2	[−0.079; 0.287]	0.096	85.3%	14.7%	1–1	1.00
100	K3	[−0.416; −0.203]	−0.313	0.0%	100.0%	10–10	10.00
100	K4	[−0.753; −0.410]	−0.578	0.0%	100.0%	6–8	6.99
100	K5	[−0.163; 0.167]	0.001	50.2%	49.8%	7–9	8.13
100	K6	[−0.308; −0.059]	−0.184	0.3%	99.7%	8–9	8.87
100	K7	[0.355; 0.528]	0.441	100.0%	0.0%	2–5	4.00
100	K8	[0.728; 0.991]	0.859	100.0%	0.0%	4–6	5.02
100	K9	[0.067; 0.470]	0.278	99.4%	0.6%	5–7	5.98
100	K10	[0.927; 1.299]	1.111	100.0%	0.0%	2–4	2.04
20	K1	[−1.736; −1.674]	−1.708	0.0%	100.0%	2–5	2.92
20	K2	[−0.275; 0.564]	0.101	67.1%	32.9%	1–1	1.00
20	K3	[−0.550; −0.055]	−0.310	0.8%	99.2%	8–10	9.98
20	K4	[−0.977; −0.198]	−0.579	0.2%	99.8%	5–9	6.95
20	K5	[−0.365; 0.363]	0.001	50.4%	49.6%	6–10	8.19
20	K6	[−0.454; 0.118]	−0.180	10.7%	89.3%	6–10	8.69
20	K7	[0.247; 0.633]	0.438	100.0%	0.0%	2–7	3.93
20	K8	[0.555; 1.162]	0.857	100.0%	0.0%	3–7	5.08
20	K9	[−0.233; 0.668]	0.267	87.9%	12.1%	3–9	5.94
20	K10	[0.706; 1.535]	1.114	100.0%	0.0%	2–5	2.32
5	K1	[−1.749; −1.180]	−1.656	0.0%	100.0%	1–7	3.03
5	K2	[−0.566; 1.103]	0.128	58.6%	41.4%	1–4	1.02
5	K3	[−0.743; 0.253]	−0.294	12.0%	88.0%	5–10	9.67
5	K4	[−1.374; 0.243]	−0.578	7.1%	92.9%	3–10	6.93
5	K5	[−0.749; 0.792]	0.008	50.6%	49.4%	4–10	8.11
5	K6	[−0.709; 0.509]	−0.168	25.5%	74.5%	5–10	8.61
5	K7	[−0.004; 0.824]	0.416	97.5%	2.5%	1–10	4.05
5	K8	[0.193; 1.402]	0.835	99.2%	0.8%	1–10	5.02
5	K9	[−1.025; 0.973]	0.214	72.4%	27.6%	1–10	5.86
5	K10	[0.286; 1.656]	1.095	99.6%	0.4%	1–7	2.71

Note: 5000 simulations were performed at each α level; $\alpha = 100$ represents low, $\alpha = 20$ moderate, and $\alpha = 5$ high weight variability.

Appendix E. TFN-Width Sensitivity Analysis

TFN widths were varied at $\lambda = 0.50$, 0.75, 1.00, 1.25, and 1.50 using the positive-side transformation in Appendix J. Reciprocal scale TFNs were generated after each width transformation, and the complete calculation chain was repeated. Table A7 presents the resulting $D - R$ values and $D + R$ rank ranges.

Table A7. $D - R$ values and rank stability across TFN-width scenarios.

Factor	$\lambda = 0.50$	$\lambda = 0.75$	$\lambda = 1.00$	$\lambda = 1.25$	$\lambda = 1.50$	$D + R$ Rank Range	Sign Changes
K1	−1.7102	−1.7116	−1.7132	−1.7150	−1.7168	3–3	0
K2	0.0933	0.0933	0.0933	0.0934	0.0934	1–1	0
K3	−0.3134	−0.3132	−0.3129	−0.3127	−0.3124	10–10	0
K4	−0.5752	−0.5759	−0.5767	−0.5776	−0.5786	7–7	0
K5	−0.0014	−0.0012	−0.0010	−0.0007	−0.0005	8–8	0
K6	−0.1815	−0.1822	−0.1829	−0.1837	−0.1846	9–9	0
K7	0.4410	0.4409	0.4408	0.4407	0.4408	4–4	0
K8	0.8567	0.8574	0.8583	0.8592	0.8602	5–5	0
K9	0.2809	0.2812	0.2816	0.2820	0.2824	6–6	0
K10	1.1099	1.1112	1.1128	1.1144	1.1160	2–2	0

Note: $\lambda = 1.00$ represents the baseline TFN conversion. No factor's group classification changed over the range examined.

Although Table A7 shows limited variation in the magnitude of $D - R$, the preservation of signs and of the $D + R$ ranking shows that the baseline findings are stable against the range of TFN widths considered.

Appendix F. Defuzzification-Method Sensitivity Analysis

The baseline CFCS results were compared against the center-of-gravity (COG) and graded mean integration representation (GMIR) methods. Table A8 jointly presents the $D - R$ values and $D + R$ ranks obtained under each method.

Table A8. $D - R$ values and $D + R$ ranks under alternative defuzzification methods.

Factor	CFCS $D - R$	COG $D - R$	GMIR $D - R$	CFCS $D + R$ Rank	COG $D + R$ Rank	GMIR $D + R$ Rank	Sign Changes
K1	−1.7132	−1.7088	−1.7088	3	3	3	0
K2	0.0933	0.0932	0.0932	1	1	1	0
K3	−0.3129	−0.3137	−0.3137	10	10	10	0
K4	−0.5767	−0.5744	−0.5745	7	7	7	0
K5	−0.0010	−0.0019	−0.0018	8	8	8	0
K6	−0.1829	−0.1808	−0.1808	9	9	9	0
K7	0.4408	0.4411	0.4411	4	4	4	0
K8	0.8583	0.8560	0.8560	5	5	5	0
K9	0.2816	0.2806	0.2806	6	6	6	0
K10	1.1128	1.1087	1.1085	2	2	2	0

Note: Ranks are determined by $D + R$ values. The CFCS classification and ranking were preserved under both alternative methods.

Figure A5 presents $D + R$ rank agreement across the three defuzzification methods. Their rankings coincide within this limited comparison. This is not an intervention-priority order and should not be generalized to aggregation changes, for which Appendix I shows a different ranking pattern.

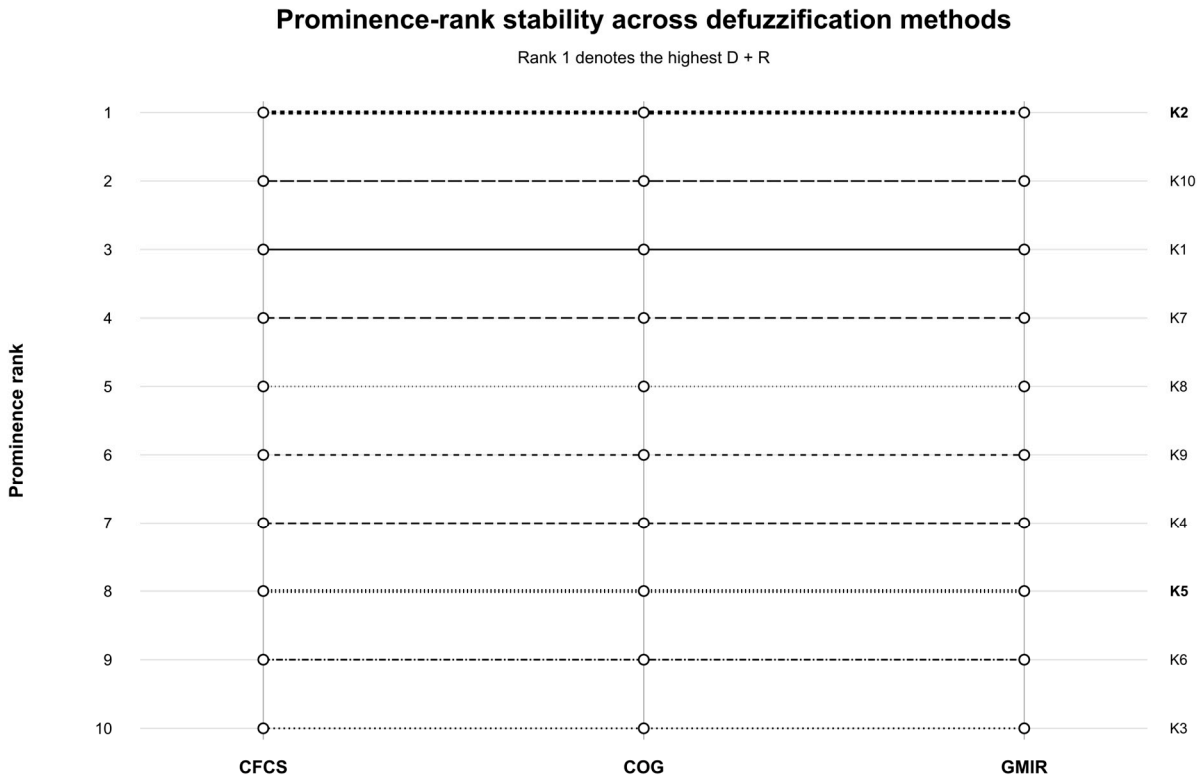


Figure A5. $D + R$ rank stability across defuzzification methods.

Appendix G. Qualitative Instrument and Coding Audit Trail

The archived thesis [60] retains three open-ended questions, presented here in English translation: 'What do you consider to be the main purpose of the Sea Toll Program?'; 'What suggestions do you have for improving the future performance of the Sea Toll Program?'; and 'What do you consider to be the most critical factors affecting the performance of the Sea Toll Program?' Table A9 provides illustrative coding paths from the retained responses to initial codes, candidate themes, and final factors, distinguishing respondent-supported content from broader literature-informed elaboration. Participant codes A1K1–A1K42 identify the 42 qualitative-phase respondents and are distinct from the quantitative expert codes E1–E10.

Table A9. Contextualized archived responses, illustrative coding paths, and limits of interpretation.

Factor	Contextualized Retained Response	Illustrative Initial Code	Candidate Theme	Interpretation Boundary
K1	A1K1: "Improve port infrastructure in remote regions and connect it to better inland transport."	Port facilities; inland access	Physical logistics conditions	Port infrastructure is directly supported. The quote also describes a land connection; it does not establish a quantified K7-to-K1 effect.
K2	A1K11: "Increase the transport fleet (ships) and competent human resources."	Vessel provision; personnel	Transport resources	The fleet component supports K2. The same response may support a workforce code, without frequency weighting.
K3	A1K13: "Lack of inter-institutional coordination."	Coordination gap	Institutional organization	Supports a coordination factor; a detailed mechanism remains interpretive.

Table A9. *Cont.*

Factor	Contextualized Retained Response	Illustrative Initial Code	Candidate Theme	Interpretation Boundary
K4	A1K30: "A logistics law is needed as a legal umbrella."	Legal framework	Policy conditions	Supports a policy/regulatory factor; does not establish its causal ordering.
K5	A1K8: "The most important factor affecting this performance is the workforce itself." A1K10 recommends developing port human resources to improve skills and knowledge.	Workforce; knowledge and skills	Human capability	The two excerpts support workforce and capability concerns, not a separately measured digital or administrative competence scale.
K6	A1K42: "The subsidy process improves."	Subsidy process	Public support	Supports financial-support relevance; the expanded outcome-based definition is not independently demonstrated by this short statement.
K7	A1K11: "After the goods on board are unloaded, connections on land are still needed so that goods reach consumers/end users at an efficient price."	Sea-to-land connection	Intermodal distribution	Supports the inland-connection interpretation of K7.
K8	A1K3: "There is no return cargo because the region has no commodities."	Return-cargo shortage; commodities	Cargo availability	The respondent explicitly relates cargo to local production. Network-wide effect size and direction were not measured.
K9	A1K34: "… transparency without corruption." A1K5 recommends tariff regulation and transparency so users can see actual costs.	Transparency; corruption concerns	Governance conditions	Supports a governance theme but cannot establish a uniform expert interpretation of the original combined K9 label.
K10	A1K42: "Insufficient economic activity." A1K36 recommends helping local businesses and farmers ship products, including on return voyages.	Local activity; market access; return cargo	Regional logistics development	Supports a broad regional-development factor. Consolidation and storage remain literature-informed proposals, not dimensions proven to have been assessed by every expert.

Appendix H. Expert Comparison Instrument and Data Structure

Expert evaluations were collected between 10 and 24 June 2025. The instrument presented 45 unordered factor pairs with mirrored 0–4 intensity headings. The questionnaire was administered in English to the experts in Indonesia. The retained bilingual copy includes a Turkish translation prepared after administration and documents the factor labels, their left and right placement, and the mirrored headings. No separate contemporaneous written instruction describing relative dominance was available for verification. The layout shows that a mark records a side and an intensity, but it does not establish that participants understood the categories as ratios or that "No Impact" meant equality. Experts received the factor labels without written operational definitions; Table 4 distinguishes those labels

from the analytical descriptions and their qualitative basis. Table A10 reproduces the English-language response layout using an illustrative marked item. The mark does not represent an individual response.

Table A10. English-language response layout for a complete example item.

A	B	C	D	E	F	G	H	I	J	K	L	M
Left Factor	**No Impact 0**	**Very Less Impact 1**	**Low Impact 2**	**High Impact 3**	**Very High Impact 4**		**Very High Impact 4**	**High Impact 3**	**Low Impact 2**	**Very Less Impact 1**	**No Impact 0**	**Right Factor**
Port Infrastructure				✓								Fleet Availability

Note: Columns B–F and H–L are separated by column G. The headings reproduce the English wording used with participants. The mark in column E is illustrative.

To reproduce the analysis, locate the marked column for each factor pair and apply Table A11. If the selected column is on the left, assign its mapped magnitude to the left-to-right entry; if it is on the right, assign the reciprocal. Calculate the reverse entry as the reciprocal of the first. A code 0 mark on either side yields $(1, 1)$ under this analysis-stage convention. The ten diagonal entries of each expert matrix are zero. This procedure produces two linked matrix entries from one elicited response.

Table A11. Mapping from each selectable response to the two matrix entries.

Side	Form Code	Left-to-Right Entry	Right-to-Left Entry
Left	0	1	1
Left	1	3	1/3
Left	2	5	1/5
Left	3	7	1/7
Left	4	9	1/9
Right	4	1/9	9
Right	3	1/7	7
Right	2	1/5	5
Right	1	1/3	3
Right	0	1	1

For the illustrative example, the left-to-right entry is 7 and the reverse is 1/7. The 450 recorded marks comprise 2 selections of code 1, 30 of code 2, 241 of code 3, and 177 of code 4. Because neither code 0 option was selected, its mapping does not contribute to any computed result. The reciprocal interpretation is an analytical representation of the layout, not a separately documented instruction provided to experts.

For each archived expert-level pair, reciprocity was checked using $|a_{ij}a_{ji} - 1| > 10^{-10}$. The audit identifies 64 nonreciprocal pairs out of 450: 34 fail the expected reversal of direction, and 30 have opposite directions but unequal reciprocal magnitudes. Products among these 64 pairs range from 1/81 to 81. These are discrepancies between discrete scale entries, not rounding differences. The records do not identify the exact transcription or processing step responsible for each discrepancy.

The counts are not additive: the 64 nonreciprocal archived pairs are included in the 210 pairs with at least one questionnaire-to-matrix difference.

Table A12. Reconciliation of questionnaire selections and archived comparison records.

Audit Measure	Count or Value
Expert–factor-pair combinations	450
Nonreciprocal archived pairs	64 (14.22%)
Direction inconsistent with reciprocity	34
Opposite directions with nonreciprocal magnitudes	30
Product range among nonreciprocal pairs	1/81 to 81
Pairs with at least one score different from the analytical archive	210 (357 directional cells)
Nonreciprocal pairs in the analysis inputs	0
Maximum reciprocal residual in the analysis inputs	0 at stored precision

The analytical entries are traceable by anonymized expert code, factor pair, source row, and marked column in the confidential mapping. Entries were derived from the recorded selections rather than by averaging or forcing the two archived values to agree. For E3's K2/K9 pair, the selection in the original response corresponds to 7 and 1/7 under the coding rule in Table A11. Original response records are kept confidential because they contain participants' personal data; the corresponding anonymized analytical entries are included in the shared research materials. The archived and reconciled analyses give the same ten prominence ranks and nine of ten net-index signs. This comparison documents numerical sensitivity to the recorded inputs; it does not establish the semantic validity of the scale.

Appendix I. Aggregation and Orientation Diagnostics

The following diagnostics examine how aggregation and comparison orientation affect the model indices. Arithmetic and geometric aggregation are compared using the same expert matrices, and descriptive agreement is evaluated under two fixed orientations of the 45 factor pairs (see Table A13).

Table A13. Orientation sensitivity of descriptive agreement.

Stored Orientation	Kendall W	ICC(2,1), ln	ICC(2,K), ln
Upper triangle, $i < j$	0.0685	−0.0272	−0.3595
Reciprocal of lower triangle	0.0685	−0.0272	−0.3595

Each orientation contains 45 entries per expert. Upper-triangle entries are used in the principal descriptive analysis; reciprocals of lower-triangle entries provide the orientation check. The estimates coincide because the two representations encode the same selections.

For the aggregation diagnostic, the weighted mean of the logarithms of each positive TFN component was exponentiated separately for its lower, modal, and upper bounds [74]. Diagonal entries remained zero. CFCS, normalization, T, and its row/column indices were then recomputed (see Table A14). For reciprocal expert inputs, componentwise geometric aggregation preserves the reciprocal TFN relationship at the fuzzy stage; CFCS need not preserve it in the defuzzified scores.

For two exactly opposed reciprocal judgments, 9 and 1/9, arithmetic aggregation produces approximately 4.556 in both directions, whereas geometric aggregation produces 1 in both directions. This illustrative example explains why aggregation choice can affect prominence when expert assessments differ strongly.

Table A14. Arithmetic and geometric aggregation on the same stored entries.

Factor	AM $D - R$	GM $D - R$	AM $D + R$ Rank	GM $D + R$ Rank	Sign Change
K1	−1.71324	−1.43644	3	1	No
K2	0.09334	0.09206	1	3	No
K3	−0.31290	−0.22620	10	7	No
K4	−0.57666	−0.47933	7	6	No
K5	−0.00099	0.02786	8	10	Yes
K6	−0.18292	−0.12437	9	9	No
K7	0.44078	0.40834	4	8	No
K8	0.85825	0.70541	5	2	No
K9	0.28155	0.23190	6	5	No
K10	1.11279	0.80076	2	4	No

Nine of ten baseline signs are retained; K5 changes sign. Prominence rankings are more aggregation-sensitive (Spearman = 0.685; Kendall tau-b = 0.511). K2 moves from rank 1 to rank 3. Thus, unchanged signs in several factors do not imply stable prominence or a consensus interpretation of the arithmetic model.

Table A15. Representative stored comparison distributions.

Pair	Values in Expert Order E1 to E10	>1	=1	<1	AM Forward/ Reverse
K1/K2	0.111, 7.000, 7.000, 9.000, 7.000, 7.000, 9.000, 0.111, 0.111, 0.111	6	0	4	4.644/3.679
K2/K5	0.111, 7.000, 0.143, 0.111, 7.000, 0.143, 9.000, 7.000, 9.000, 0.111	5	0	5	3.962/4.165
K7/K10	0.111, 0.111, 0.143, 7.000, 9.000, 7.000, 0.111, 7.000, 0.111, 9.000	5	0	5	3.959/4.365
K1/K6	0.111, 7.000, 0.143, 0.143, 0.143, 7.000, 0.111, 0.143, 0.200, 0.111	2	0	8	1.510/6.029

The examples show opposing coded selections in comparisons involving K2 and provide contrasting distributions across four factor pairs (see Table A15). All 45 pair-level opposite arithmetic means exceed one. Log dispersion and the product of these means are strongly associated (descriptive Spearman = 0.957). Factor-level mean incident dispersion and arithmetic prominence are also associated (Spearman = 0.527 across ten factors). These associations support treating disagreement amplification as a contributor to prominence, not proving that it explains every ranking.

The primary comparison uses one-step summaries of X; the secondary comparison uses the corresponding summaries of T. For X, the one-step total is the sum of a factor's row and column totals, and the one-step balance is their difference. For T, these quantities are $D + R$ and $D - R$. Because normalization divides all Z entries by the same scalar, the one-step ordering and signs also represent the simple marginal summaries of Z. All calculations use the unrounded baseline inputs; no expert entry is changed.

In Equation (9), the higher-order contribution is $T - X$, comprising X^2 and all subsequent powers. These terms weight comparisons by the broader configuration in which the compared factors occur. Table A16 therefore tests the effect of that representation rather than an additional substantive relationship. A product such as $0.20 \times 0.30 = 0.06$ is a modeled comparison weight, not a measured transmission coefficient.

Table A16. One-step and transformed summaries of the baseline comparison matrix.

Factor	One-Step Total (X)	One-Step Rank	$D + R$ Rank (T)	One-Step Balance (X)	$D - R$ (T)
K1	1.50419	4	3	−0.49581	−1.71324
K2	1.53852	1	1	+0.02894	+0.09334
K3	1.45190	9	10	−0.09096	−0.31290
K4	1.48442	7	7	−0.16130	−0.57666
K5	1.44792	10	8	+0.00309	−0.00099
K6	1.47677	8	9	−0.05424	−0.18292
K7	1.50948	3	4	+0.12325	+0.44078
K8	1.50151	5	5	+0.24515	+0.85825
K9	1.48461	6	6	+0.08279	+0.28155
K10	1.51430	2	2	+0.31909	+1.11279

Note: All ranks are descending row-plus-column totals. One-step balance and $D - R$ are row-minus-column totals of X and T, respectively. The matrices use the same normalized comparative entries; T also includes higher-order terms. Values are rounded only for display. The principal one-step totals, balances, and ranks are also presented in Table 10, Panel A.

The rank correlations between X and T are 0.952 (Spearman) and 0.867 (Kendall tau-b). Five ranks and nine signs agree. K2, K10, and K8 retain ranks 1, 2, and 5, respectively. K5 changes from +0.00309 in X to −0.00099 in T. Agreement identifies findings that are not sensitive to this particular representation; the near-zero K5 reversal is a reason to qualify, not strengthen, its directional interpretation.

To assess whether the wording of K9 affects the wider comparison, its row and column were excluded from each expert matrix before aggregation. The full calculation was then repeated for the nine remaining factors. Table A17 compares ranks within the same set of nine factors and reports the change in net index.

Table A17. Comparison of the retained factors with and without K9.

Factor	Full $D - R$	$D - R$ Without K9	Change	Full Rank Among Nine	Rank Without K9	Sign Change
K1	−1.7132	−1.6774	+0.0359	3	3	No
K2	0.0933	−0.1357	−0.2291	1	1	Yes
K3	−0.3129	−0.3355	−0.0226	9	9	No
K4	−0.5767	−0.6163	−0.0397	6	7	No
K5	−0.0010	0.2235	+0.2245	7	6	Yes
K6	−0.1829	−0.1100	+0.0729	8	8	No
K7	0.4408	0.6214	+0.1806	4	4	No
K8	0.8583	1.0745	+0.2162	5	5	No
K10	1.1128	0.9555	−0.1573	2	2	No

Removing K9 reverses the signs of K2 (+0.0933 to −0.1357) and K5 (−0.0010 to +0.2235). The other seven retained factors keep their signs. K2 and K10 remain first and second in prominence, while K5 moves from seventh to sixth among the nine retained factors. This diagnostic shows how the model depends on including K9; it does not validate the remaining constructs.

Appendix J. Scenario Generation and TFN Width Specification

For a positive-side baseline TFN (l, m, u), widths are transformed around the fixed modal value before the reciprocal scale entry is generated. The same λ is applied to all

positive-side scale categories within a scenario. The neutral analytical TFN (1,1,1) and the point-valued extreme (9,9,9) remain unchanged.

$$l' = m - \lambda(m - l) \quad \text{(A1)}$$

$$u' = m + \lambda(u - m) \quad \text{(A2)}$$

The reciprocal scale TFN is formed as $(1/u', 1/m, 1/l')$, rather than by independently expanding its endpoints. For $\lambda = 0.50$, 0.75, 1.00, 1.25, and 1.50, all positive-side lower bounds remain strictly positive and $l' \leq m \leq u'$ holds. These conditions concern the TFN scale and are distinct from reciprocal equality of the expert or aggregated matrices.

Table A18. Construction of the combined scenario collection.

Category	Generation Rule	Count	Collection Share
Equal weight	5 widths × 3 defuzzification methods; ten weights = 0.10	15	0.24%
Leave one out	5 widths × 3 methods × 10 omitted experts; retained weights = 1/9	150	2.43%
Dirichlet stress	Widths 0.50 and 1.50 × 3 methods × 1000 draws from Dirichlet (5,...,5)	6000	97.32%
Total	Enumerated cases plus sampled stress cases	6165	100%

CFCS, COG and GMIR are evaluated in that order within each width. The stress draws are generated separately for each extreme-width/method combination using NumPy's default random generator and seed 20260907. The draws lie on the probability simplex: all weights are positive and sum to one, without additional narrow hard bounds. The allocation of 1000 draws to each selected stress combination controls computational coverage; it is not an estimate of scenario prevalence. The resulting pooled percentage gives each evaluated case one count and therefore chiefly reflects the stress category.

The 10,000 expert-bootstrap resamples and the separate 5000 Dirichlet draws at each of $\alpha = 100$, 20, and 5 are analyzed separately from the 6165 combined scenarios (see Table A18). All generated combined cases met the numerical validity checks and were included. Each case contributes one observation to the reported retention percentages. Table 12 reports sign frequencies and rank ranges for each family without assigning a single stability class.

References

1. Geospatial Information Agency of Indonesia (BIG). Indonesia's Island Data Update: 17,380 Islands in 2024. 2024. Available online: https://sipulau.big.go.id/news/11 (accessed on 16 October 2026).
2. Herliana, L.; Parsons, S. Logistics in Indonesia. In *The Impacts and Benefits of Structural Reforms in Transport, Energy and Telecommunications Sectors*; APEC: Singapore, 2011; pp. 445–486. Available online: https://www.apec.org/docs/default-source/Publications/2011/1/The-Impacts-and-Benefits-of-Structural-Reforms-in-Transport-Energy-and-Telecommunications-Sectors/TOC/Logistics-in-Indonesia.pdf (accessed on 16 October 2026).
3. Reza, M. The Relationship between Logistics and Economic Development in Indonesia: Analysis of Time Series Data. *J. Tek. Ind.* **2013**, *15*, 119–124. [CrossRef]
4. Santoso, S.; Nurhidayat, R.; Mahmud, G.; Arijuddin, A.M. Measuring the Total Logistics Costs at the Macro Level: A Study of Indonesia. *Logistics* **2021**, *5*, 68. [CrossRef]
5. Ministry of Transportation of the Republic of Indonesia. Program Tol Laut Resmi Diluncurkan, 4 November 2015. Available online: https://dephub.go.id/post/read/program-tol-laut-resmi-diluncurkan (accessed on 16 October 2026). (In Indonesian)
6. Kusuma, L.T.W.N.; Tseng, F.-S. Analysis of the Impact of the 'Sea Toll' Program for Seaports: Resilience and Competitiveness. *Appl. Sci.* **2019**, *9*, 3407. [CrossRef]
7. Ministry of Transportation of the Republic of Indonesia. 10 Tahun Layanan Tol Laut, Trayek, Muatan, dan Armada Kapal Untuk Wilayah 3TP Naik Signifikan, 2024. Available online: https://www.kemenhub.go.id/index.php/post/read/10-tahun-layanan-tol-laut%2C-trayek%2C-muatan%2C-dan-armada-kapal-untuk-wilayah-3tp-naik-signifikan (accessed on 16 October 2026). (In Indonesian)

8. Fauzi, I.; Airawati, M.N.; Cholishoh, E.; Murtiaji, C. The Development and Challenges of Sea Tolls in Supporting Indonesia's Vision 2045. *AIP Conf. Proc.* **2024**, *3145*, 020040. [CrossRef]
9. Mubarak, M.; Nuralamsyah, N.; Hanif, M.R. Analysis of Sea Toll Policy and Its Effect on Sea Toll Effectiveness. *Sensistek* **2019**, *2*, 124–128. [CrossRef]
10. Rahayu, L.; Busscher, T.; Tillema, T.; Woltjer, J. Maritime Transport Governance Challenges in the Global South. *Mar. Policy* **2024**, *163*, 106147. [CrossRef]
11. Bahagia, S.N.; Sandee, H.; Meeuws, R. *State of Logistics Indonesia 2013*; Report No. 80871; World Bank: Washington, DC, USA, 2013. Available online: https://documents.worldbank.org/en/publication/documents-reports/documentdetail/993771468285047652 (accessed on 16 October 2026).
12. Wicaksana, I.G.W. Indonesia's Maritime Connectivity Development: Domestic and International Challenges. *Asian J. Polit. Sci.* **2017**, *25*, 212–233. [CrossRef]
13. Rodrigue, J.-P.; Comtois, C.; Slack, B. *The Geography of Transport Systems*; Routledge: London, UK, 2006; ISBN 978-0-415-35441-7.
14. Rodrigue, J.-P. *The Geography of Transport Systems*, 6th ed.; Routledge: London, UK, 2024; ISBN 978-1-032-38040-7. [CrossRef]
15. Myrdal, G. *Economic Theory and Under-Developed Regions*; Gerald Duckworth: London, UK, 1957; ISBN 978-0-416-68160-4.
16. Floerkemeier, H.; Spatafora, N.; Venables, A. *Regional Disparities, Growth, and Inclusiveness*; IMF Working Paper 2021/038; International Monetary Fund: Washington, DC, USA, 2021; ISBN 978-1-5135-6950-5. [CrossRef]
17. Arsani, A.M.; Kurniawan, M.A. Infrastructure Sharing for Acceleration of Economic Equality in Indonesia Islands Regions. *J. Perenc. Pembang. Indones. J. Dev. Plan.* **2018**, *2*, 291–308. [CrossRef]
18. Prihartono, B.; Irawan, C.; Bastian; Setyanto, W.D.W. *Konsep Tol Laut dan Implementasi 2015–2019*; Ministry of National Development Planning/Bappenas and Ministry of Transportation: Jakarta, Indonesia, 2015. Available online: https://perpustakaan.bappenas.go.id/e-library/file_upload/koleksi/migrasi-data-publikasi/file/Unit_Kerja/Deputi_Bidang_Saran_dan_Prasarana/Direktorat-Transportasi/Pengembangan%20Tol%20Laut%20Dalam%20Rpjmn%202015-2019%20Dan%20Implementasi%202015.pdf (accessed on 16 October 2026). (In Indonesian)
19. Ridwan, A.S.; Hartini, S. Analisis Transportasi Laut Indonesia. *J. Reg. City Plan.* **1997**, *8*, 49–59. Available online: https://journal-itb.org/jrcp/volumes/1997/8/issues/3/analisis-transportasi-laut-indonesia.html (accessed on 9 October 2026).
20. Febriansyah, A.; Sahara, S. Analysis of the effect of the sea toll program on logistics efficiency in Indonesia. *EKONOMIKA45* **2023**, *10*, 515–522. [CrossRef]
21. Wijaya, E. President's "Sea Toll" Program Has Been Implemented, Price Disparity Eliminated. 2015. Available online: https://setkab.go.id/program-tol-laut-presiden-telah-dijalankan-disparitas-harga-dihilangkan/ (accessed on 15 August 2026).
22. Government of Indonesia. Presidential Regulation No. 106 of 2015 on Public Service Obligations for Maritime Freight Transport. Available online: https://peraturan.bpk.go.id/Details/41855/perpres-no-106-tahun-2015 (accessed on 16 October 2026).
23. Ministry of Transportation of the Republic of Indonesia. Regulation No. PM 161 of 2015 on Public Service Obligations for Maritime Freight Transport. Available online: https://jdih.dephub.go.id/peraturan/detail?data=GZKQbJfRcdJ1JxbbuArXSN4jpaMsh9eZX8bPGhNNq6dg8cNdJb8O7Wf49Z4CRt5c2r8cPAnoKZYFJ4Pb0uqnD1nj8n3DXEPmdyo49Y1w7qPv7rDVS8ReY8UbH8zXWVTm7rQ0L5DNV0x2DfjCUnhEku6eWi (accessed on 16 October 2026). (In Indonesian)
24. Nur, H.I.; Achmadi, T.; Verdifauzi, A. Optimalisasi Program Tol Laut terhadap Penurunan Disparitas Harga: Suatu Tinjauan Analisis. *J. Penelit. Transp. Laut.* **2020**, *22*, 1–12. (In Indonesian) [CrossRef]
25. Fahmiasari, H.; Parikesit, D. Container Shipping Network Efficiency Comparison in Indonesia: Nusantara Pendulum and Sea Tollway. *Asian J. Shipp. Logist.* **2017**, *33*, 79–84. [CrossRef]
26. Sa'adah, K.; Yakti, D.; Susanto, S.R. Four-Year Evaluation of Jokowi's Sea Toll Policy: The Indonesia-Centric Connectivity Concept for Economic Equality. *Masy. Kebud. Polit.* **2019**, *32*, 290–302. [CrossRef]
27. Ratnawati, E. Sea Toll as a Means to Increase the Effectiveness of Goods and Services to the Eastern of Indonesia. *Awang Long. Law Rev.* **2019**, *1*, 120–130. [CrossRef]
28. Kristini, F.; Bambang, A.N.; Handoko, W.; Priadi, A.A. Efektivitas dan Efisiensi Program Tol Laut Berbasis AHP (Studi Kasus: Pelabuhan Tahuna). *War. Penelit. Perhub.* **2019**, *31*, 75–82. (In Indonesian) [CrossRef]
29. Apriani, E.; Arief, S.; Supriyadi, A.A. Sea Tolls as Indonesia's Blue Economy Strategy. *IOP Conf. Ser. Earth Environ. Sci.* **2025**, *1464*, 012001. [CrossRef]
30. Mentzer, J.T.; DeWitt, W.; Keebler, J.S.; Min, S.; Nix, N.W.; Smith, C.D.; Zacharia, Z.G. Defining Supply Chain Management. *J. Bus. Logist.* **2001**, *22*, 1–25. [CrossRef]
31. Lambert, D.M.; Cooper, M.C.; Pagh, J.D. Supply Chain Management: Implementation Issues and Research Opportunities. *Int. J. Logist. Manag.* **1998**, *9*, 1–19. [CrossRef]
32. Flynn, B.B.; Huo, B.; Zhao, X. The Impact of Supply Chain Integration on Performance: A Contingency and Configuration Approach. *J. Oper. Manag.* **2010**, *28*, 58–71. [CrossRef]
33. Prajogo, D.; Olhager, J. Supply Chain Integration and Performance: The Effects of Long-Term Relationships, Information Technology and Sharing, and Logistics Integration. *Int. J. Prod. Econ.* **2012**, *135*, 514–522. [CrossRef]

34. Humang, W.P.; Hadiwardoyo, S.P.; Nahry. Factors Influencing the Integration of Freight Distribution Networks in the Indonesian Archipelago: A Structural Equation Modeling Approach. *Adv. Sci. Technol. Eng. Syst. J.* **2019**, *4*, 278–286. [CrossRef]
35. Song, D.-W.; Panayides, P.M. Global Supply Chain and Port/Terminal: Integration and Competitiveness. *Marit. Policy Manag.* **2008**, *35*, 73–87. [CrossRef]
36. Ha, M.-H.; Yang, Z.; Notteboom, T.; Ng, A.K.Y.; Heo, M.-W. Revisiting Port Performance Measurement: A Hybrid Multi-Stakeholder Framework for the Modelling of Port Performance Indicators. *Transp. Res. Part E Logist. Transp. Rev.* **2017**, *103*, 1–16. [CrossRef]
37. Parmenter, D. *Key Performance Indicators: Developing, Implementing, and Using Winning KPIs*, 2nd ed.; John Wiley & Sons: Hoboken, NJ, USA, 2010; ISBN 978-0-470-54515-7.
38. Mawikere, W.A.I.; Herijanto, W. Analysis of the Performance Evaluation of the Logistics Express Liner Port on the T-14 Tanjung Perak-Larantuka-Lewoleba-Kalabahi Route in Reducing Price Disparity. *J. Civ. Eng.* **2023**, *38*, 86. [CrossRef]
39. Gugat, R.M.D.; Raga, P.; Marpaung, E.; Adety, C.S.S.Z.P. Load Factor Improvement Analysis of Return Cargo on the Tol Laut Ship Route. *Adv. Transp. Logist. Res.* **2022**, *5*, 49–69. Available online: https://proceedings.itltrisakti.ac.id/index.php/ATLR/article/download/476/515 (accessed on 16 October 2026).
40. United Nations Conference on Trade and Development. *Review of Maritime Transport 2020*; United Nations: Geneva, Switzerland, 2020; ISBN 978-92-1-112993-9. Available online: https://unctad.org/publication/review-maritime-transport-2020 (accessed on 16 October 2026).
41. Directorate General of Sea Transportation, Ministry of Transportation of the Republic of Indonesia. Portal Tol Laut Indonesia: SITOLAUT. Available online: https://sitolaut.dephub.go.id/ (accessed on 16 October 2026).
42. Junida, A.I.; Wijaya, A. Pioneer Transport Subsidy Budget Allocation for 2024 Amounts to IDR 4.1 Trillion. Antara News. 6 June 2023. Available online: https://www.antaranews.com/berita/3574422/alokasi-anggaran-subsidi-angkutan-perintis-2024-sebesar-rp41-triliun (accessed on 16 October 2026).
43. Adventy, A. Pelni Realizes IDR 214 Billion Sea Toll Subsidy Budget throughout 2024. Bisnis Indonesia. 6 January 2025. Available online: https://ekonomi.bisnis.com/read/20250106/98/1829295 (accessed on 16 October 2026).
44. Setijadi. *Analysis and Recommendations of Supply Chain Indonesia on the Sea Toll Program*; Supply Chain Indonesia: Bandung, Indonesia, 2017.
45. Brooks, M.R.; Knatz, G.; Pallis, A.A.; Wilmsmeier, G. Transparency in Port Governance: Setting a Research Agenda. *J. Shipp. Trade* **2022**, *7*, 1. [CrossRef]
46. Iman, N.; Amanda, M.T.; Angela, J. Digital Transformation for Maritime Logistics Capabilities Improvement: Cases in Indonesia. *Mar. Econ. Manag.* **2022**, *5*, 188–212. [CrossRef]
47. Amin, C.; Mulyati, H.; Anggraini, E.; Kusumastanto, T. Impact of maritime logistics on archipelagic economic development in eastern Indonesia. *Asian J. Shipp. Logist.* **2021**, *37*, 157–164. [CrossRef]
48. Sofiyandi, Y.; Kurniawan, Y.R.; Yudhistira, M.H. The Impact of Maritime Logistics Subsidy on Food Prices: Evidence from Indonesia. *Econ. Anal. Policy* **2023**, *79*, 1026–1045. [CrossRef]
49. Sunitiyoso, Y.; Nuraeni, S.; Pambudi, N.F.; Inayati, T.; Nurdayat, I.F.; Hadiansyah, F.; Tiara, A.R. Port Performance Factors and Their Interactions: A Systems Thinking Approach. *Asian J. Shipp. Logist.* **2022**, *38*, 107–123. [CrossRef]
50. Amin, C.; Hasyim, A.W.; Sun'an, M.; Hilman, R.M.; Fahmiasari, H. Impact of Increasing Local Economic Capacity on Reducing Maritime Logistics Costs in Island Provinces of Eastern Indonesia: A Dynamic System Approach. *Transp. Res. Interdiscip. Perspect.* **2024**, *27*, 101195. [CrossRef]
51. Arisusanty, D.J.; Arkeman, Y.; Rahardjo, S.; Soeboer, D.A. Analisa Menentukan Kriteria Pemilihan Pelabuhan Pengumpan Tol Laut Menggunakan Metode AHP. *ALBACORE J. Penelit. Perikan. Laut.* **2018**, *2*, 57–67. (In Indonesian) [CrossRef]
52. Zadeh, L.A. Fuzzy Sets. *Inf. Control* **1965**, *8*, 338–353. [CrossRef]
53. Kahraman, C. (Ed.) *Fuzzy Multi-Criteria Decision Making: Theory and Applications with Recent Developments*; Springer: New York, NY, USA, 2008; ISBN 978-0-387-76812-0. [CrossRef]
54. Deng, J.L. Control Problems of Grey Systems. *Syst. Control Lett.* **1982**, *1*, 288–294. [CrossRef]
55. Liu, S.; Lin, Y. *Grey Systems: Theory and Applications*; Springer: Berlin/Heidelberg, Germany, 2011; ISBN 978-3-642-16157-5. [CrossRef]
56. Sharifian, R.; Ghasemi, F.; Kharazmi, E.; Farhadi, P.; Monem, H.; Shokrpour, N. An Evaluation of the Risk Factors Associated with Implementing Projects of Health Information Technology by Fuzzy Combined ANP-DEMATEL. *PLoS ONE* **2023**, *18*, e0279819. [CrossRef] [PubMed]
57. Shieh, J.-I.; Wu, H.-H.; Huang, K.-K. A DEMATEL Method in Identifying Key Success Factors of Hospital Service Quality. *Knowl.-Based Syst.* **2010**, *23*, 277–282. [CrossRef]
58. Si, S.L.; You, X.Y.; Liu, H.C.; Zhang, P. DEMATEL technique: A systematic review of the state-of-the-art literature on methodologies and applications. *Math. Probl. Eng.* **2018**, *2018*, 3696457. [CrossRef]

59. Creswell, J.W. *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*, 4th ed.; SAGE Publications: Thousand Oaks, CA, USA, 2014; ISBN 978-1-4522-2610-1.
60. Agustin, B. Identifying Critical Factors Affecting the Performance of the Sea Toll Program: An Analysis Using the DEMATEL Method. Master's Thesis, Akdeniz University, Antalya, Türkiye, 2025.
61. Rattray, J.; Jones, M.C. Essential elements of questionnaire design and development. *J. Clin. Nurs.* **2007**, *16*, 234–243. [CrossRef] [PubMed]
62. Guest, G.; Bunce, A.; Johnson, L. How Many Interviews Are Enough? An Experiment with Data Saturation and Variability. *Field Methods* **2006**, *18*, 59–82. [CrossRef]
63. Braun, V.; Clarke, V. Using thematic analysis in psychology. *Qual. Res. Psychol.* **2006**, *3*, 77–101. [CrossRef]
64. Gabus, A.; Fontela, E. *World Problems, an Invitation to Further Thought Within the Framework of DEMATEL*; Battelle Geneva Research Centre: Geneva, Switzerland, 1972.
65. Fontela, E.; Gabus, A. *The DEMATEL Observer*; Battelle Geneva Research Centre: Geneva, Switzerland, 1976.
66. Tzeng, G.-H.; Chiang, C.-H.; Li, C.-W. Evaluating Intertwined Effects in E-Learning Programs: A Novel Hybrid MCDM Model Based on Factor Analysis and DEMATEL. *Expert Syst. Appl.* **2007**, *32*, 1028–1044. [CrossRef]
67. Chang, D.-Y. Applications of the Extent Analysis Method on Fuzzy AHP. *Eur. J. Oper. Res.* **1996**, *95*, 649–655. [CrossRef]
68. Opricovic, S.; Tzeng, G.H. Defuzzification within a multicriteria decision model. *Int. J. Uncertain. Fuzziness Knowl.-Based Syst.* **2003**, *11*, 635–652. [CrossRef]
69. Kendall, M.G.; Babington Smith, B. The problem of m rankings. *Ann. Math. Stat.* **1939**, *10*, 275–287. [CrossRef]
70. Shrout, P.E.; Fleiss, J.L. Intraclass correlations: Uses in assessing rater reliability. *Psychol. Bull.* **1979**, *86*, 420–428. [CrossRef] [PubMed]
71. McGraw, K.O.; Wong, S.P. Forming inferences about some intraclass correlation coefficients. *Psychol. Methods* **1996**, *1*, 30–46. Correction in *Psychol. Methods* **1996**, *1*, 390. https://doi.org/10.1037/1082-989X.1.4.390. [CrossRef]
72. Koo, T.K.; Li, M.Y. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J. Chiropr. Med.* **2016**, *15*, 155–163. Erratum in *J. Chiropr. Med.* **2016**, *16*, 346. https://doi.org/10.1016/j.jcm.2017.10.001. [CrossRef] [PubMed]
73. Efron, B. Bootstrap methods: Another look at the jackknife. *Ann. Stat.* **1979**, *7*, 1–26. [CrossRef]
74. Aczél, J.; Saaty, T.L. Procedures for Synthesizing Ratio Judgements. *J. Math. Psychol.* **1983**, *27*, 93–102. [CrossRef]
75. Notteboom, T.E.; Rodrigue, J.-P. Port Regionalization: Towards a New Phase in Port Development. *Marit. Policy Manag.* **2005**, *32*, 297–313. [CrossRef]
76. Franc, P.; Van der Horst, M. Understanding Hinterland Service Integration by Shipping Lines and Terminal Operators: A Theoretical and Empirical Analysis. *J. Transp. Geogr.* **2010**, *18*, 557–566. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.