sustainability

MDPI

Article

Product Carbon Footprint Standard Implementation Readiness in Major Chinese Export-Manufacturing Provinces: An Enterprise-Centered, TOE-Grounded Mixed-Methods Assessment

Shaowei Huo [1,*], Xiaoxia Wang [2], Yujuan Xia [1], Ming Yang [1] and Liang Sun [1,*]

[1] China National Institute of Standardization, No.4 Zhichun Road, Haidian District, Beijing 100191, China; xiayj@cnis.ac.cn (Y.X.); yangming@cnis.ac.cn (M.Y.)
[2] China Automotive Technology and Research Center Co., Ltd., No.68 Xianfeng East Road, Dongli District, Tianjin 300300, China; wangxiaoxia@catarc.ac.cn
* Correspondence: huosw@cnis.ac.cn (S.H.); sunliang@cnis.ac.cn (L.S.); Tel.: +86-13521351031 (S.H.); +86-18910756015 (L.S.)

Abstract

As product carbon footprint (PCF) quantification shifts from voluntary practice to regulatory requirement under mechanisms such as the EU Carbon Border Adjustment Mechanism, evidence on implementation readiness in major exporting economies remains limited. This study examines PCF standard implementation among 277 enterprises in major Chinese export-manufacturing provinces, together with nine industry associations and 31 certification bodies. Because enterprises were recruited through government-affiliated channels, the sample is a convenience sample, and the findings describe institutionally accessible enterprises rather than national estimates. Grounded in the Technology–Organization–Environment framework, the analysis combines logistic regression with Firth penalized estimation, outcome-aware multiple imputation and sensitivity analysis, DEMATEL–ISM expert-informed synthesis, correspondence analysis, and directed content analysis. Working-group establishment and importer disclosure requirements were the most robust correlates of PCF adoption across estimation approaches and missing-data treatments; standard availability was significant only under multiple imputation and is reported as imputation-dependent. Other hypothesized associations were not robust. Among enterprises with codable responses, 42.0% had calculated their PCF, 31.5% had established working groups, and 68.0% lacked applicable standards. Expert synthesis identified limited applicable standards as the foundational upstream barrier. The findings indicate that standard consolidation and sector-differentiated organizational support are priority intervention areas, requiring confirmation through longitudinal research.

Keywords: product carbon footprint; TOE framework; DEMATEL–ISM; Firth's penalized estimation; multiple imputation; standards

Check for updates

Academic Editor: Maria Rosa Trovato

Received: 22 October 2026
Revised: 30 October 2026
Accepted: 2 November 2026
Published: 4 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

Product carbon footprint (PCF) quantification—the systematic measurement of greenhouse-gas emissions across a product's lifecycle—has shifted from voluntary practice to regulatory imperative. The EU Carbon Border Adjustment Mechanism (CBAM), whose definitive phase began on 1 January 2026, imposes carbon-intensity reporting and—beginning with the purchase of CBAM certificates for 2026 emissions in February

2027—financial obligations on imports of steel, aluminum, cement, fertilizers, electricity, and hydrogen [1], and the EU Battery Regulation (2023/1542) provides for lifecycle carbon-footprint declarations for electric-vehicle batteries from February 2025, although this obligation is not yet enforceable pending adoption of the implementing methodology [2]. These developments create acute compliance pressure for China, a major supplier of EU electric-vehicle batteries and a substantial exporter of CBAM-covered products [3,4].

China has responded rapidly. In June 2024, the Ministry of Ecology and Environment and 14 co-signing ministries issued the "Implementation Plan for Establishing a Carbon Footprint Management System" [5], targeting 100 key-product PCF standards by 2027 and 200 by 2030. The national standard GB/T 24067-2024, aligned with ISO 14067:2018, took effect on 1 October 2024. By January 2025, fifteen group standards, a 25-province certification pilot, and the National GHG Emission Factor Database (v1.0) had launched.

Empirical evidence on actual implementation readiness, however, remains scarce. The literature is dominated by CBAM trade-impact modeling [3,4], willingness-to-pay studies for carbon labels [6,7], and single-city case studies [8]. The TOE framework [9] has been widely applied to sustainability adoption—green supply chains [10], green innovation [11], circular economy [12]—but not to PCF standard implementation, and no large-scale enterprise survey of PCF readiness has been conducted in a major manufacturing economy.

This paper addresses these gaps through four contributions. First, it provides the first large-scale enterprise survey of PCF implementation readiness in major Chinese export-manufacturing provinces (n = 277), supplemented by descriptive profiles of certification bodies (n = 31). The underlying instrument was a government-commissioned survey with open-ended questions; regression variables are therefore derived through systematic post-hoc coding of qualitative responses, which constrains precision (Sections 3.2 and 3.3). Second, it identifies robust correlates of adoption through five parallel estimation approaches (standard MLE, Firth's penalized, bootstrap, MI with m = 60, MNAR sensitivity), introducing a tiered framework separating method-robust from method-dependent findings. Third, it reveals the structural hierarchy among implementation barriers via DEMATEL–ISM expert synthesis. Fourth, it uncovers differentiated barrier profiles across industry sectors and TOE dimensions through correspondence and directed content analysis, providing a basis for sector-specific policy design.

2. Theoretical Framework and Literature Review

2.1. The TOE Framework and Complementary Theoretical Perspectives for PCF Standard Adoption

The TOE framework [9] posits that organizational adoption is determined jointly by technological, organizational, and environmental contexts. It has been widely applied in sustainability research—green supply chain management [10], green logistics [13], green innovation [11], Industry 4.0 for sustainable manufacturing [14], and circular-economy barriers combined with ISM [12].

We adopt TOE as the organizing framework but integrate two perspectives that generate more specific predictions. First, absorptive-capacity theory [15] predicts that an organization's ability to recognize, assimilate, and apply external knowledge depends on dedicated internal structures. Dedicated working groups—the mechanism through which enterprises operationalize complex LCA knowledge—should therefore be more strongly associated with adoption than passive exposure to external pressures. Second, institutional theory [16,17] distinguishes coercive, mimetic, and normative pressures; in the current Chinese PCF context coercive pressures are nascent while normative pressures may transmit through standard-development participation. This predicts differentiated patterns for normative versus coercive channels, both weaker than organizational capacity in a pre-mandate environment.

We adapt TOE to PCF as follows. Technological context: availability and quality of standards, emission-factor databases, and calculation methodologies. Organizational con-

text: dedicated working groups, personnel expertise, regulatory familiarity, and standard-development participation. Environmental context: importer disclosure requirements, international regulatory pressure, policy accessibility, and industry-association support.

We expect organizational context to show the strongest associations for three reasons. First, PCF calculation is knowledge-intensive and requires specialized human capital; dedicated working groups function as the absorptive-capacity mechanism. Second, China's PCF standard system evolves faster than enterprises can track, making regulatory familiarity a binding constraint. Third, institutional theory predicts that in a pre-mandate environment with unevenly distributed coercive pressures, normative and organizational factors will dominate environmental pressures [18]—a prediction that should reverse as domestic mandates mature.

2.2. PCF Standards: International Landscape

The global PCF standards landscape has evolved through three generations. PAS 2050 [19], the world's first product carbon footprint standard, established the foundational lifecycle approach. ISO 14067:2018, developed by ISO/TC 207 with a joint secretariat held by Canada (SCC) and China (SAC), superseded PAS 2050 as the authoritative international framework with enhanced data quality requirements and third-party verification provisions [20]. The GHG Protocol Product Standard [21] offers a complementary corporate-oriented approach adopted by over 90% of Fortune 500 companies reporting to CDP. Despite methodological convergence, Liang et al. [22] demonstrated that PCF results for paper products depend on the accounting standard selected. Hu et al. [23] identified knowledge insufficiency and high cost as principal barriers through multiple case studies. Hettler and Graf-Vlachy [24] concluded that Scope 3 reporting standards remain incomplete with an overall low amount of empirical research.

2.3. China's PCF Standardization

China's domestic PCF standardization centers on GB/T 24067-2024, designed for structural alignment with ISO 14067 while adding China-specific operational guidance [25]. However, the "urgent-use-first" approach in which group standards are developed rapidly has created multi-tier fragmentation: for popular products, multiple overlapping standards exist, while for less prominent products, no standards exist at any level. The emission-factor database landscape presents a parallel challenge. The Chinese Life Cycle Database (CLCD) contains approximately 600 datasets versus Ecoinvent's 18,000+ and Sphera's 20,000+. The TianGong database (Tsinghua University) and the January 2025 National GHG Emission Factor Database are promising but remain in early development. Liu and Niu [8] identified three critical obstacles to carbon footprinting in Shanghai: lack of a centralized database, insufficient product-specific management, and absent certification systems.

2.4. International Regulatory Drivers

The EU CBAM creates an unprecedented direct financial link between product carbon intensity and market access. Pan et al. [3] estimated CBAM costs on China's metal exports to the EU of EUR 1.16–1.57 billion. The EU Battery Regulation adds sector-specific pressure through a three-stage regime culminating in maximum lifecycle thresholds by 2028. Comparative analysis reveals different national approaches: Japan paused its CFP program in 2020 before publishing updated guidelines in 2025; South Korea's KEITI system features a three-phase certification structure with 12 international mutual recognition agreements; the EU's Product Environmental Footprint (PEF) methodology covers 16 environmental impact categories. Sureeyatanapas et al. [26] surveyed Thai manufacturers certified under Thailand's carbon footprint label, finding lack of social awareness and required investment

as principal barriers—the only empirical survey-based evidence on PCF implementation in a developing-economy context prior to this study.

2.5. Research Gaps and Hypotheses

Based on the literature review and theoretical framework, we identify four research gaps: (1) no large-scale enterprise PCF readiness survey exists for a major manufacturing economy; (2) enterprise-level barrier analysis grounded in adoption theory is absent for China; (3) TOE-grounded models of PCF adoption have not been tested; and (4) structured barrier hierarchy analysis has not been applied to PCF implementation. We derive five hypotheses:

H1 (Organizational—Absorptive Capacity)**.** *Enterprise working-group establishment is positively associated with PCF calculation adoption.*

H2 (Organizational—Normative)**.** *Standard-development participation is positively associated with PCF calculation adoption.*

H3 (Organizational)**.** *High regulatory familiarity is positively associated with PCF calculation adoption.*

H4 (Technological)**.** *Availability of applicable PCF standards is positively associated with PCF calculation adoption.*

H5 (Environmental—Coercive)**.** *Importer PCF disclosure requirements are positively associated with PCF calculation adoption.*

Building on absorptive-capacity theory and institutional theory predictions for pre-mandate environments, we further hypothesize that Organizational context variables (H1–H3) will show stronger associations than Technological (H4) and Environmental (H5) variables in a multivariate model. The conceptual framework and regression results are summarized in Figure 1.

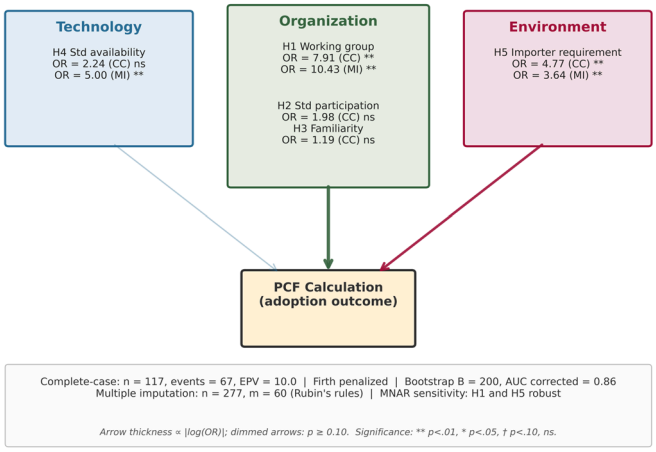


Figure 1. TOE framework with Firth penalized logistic regression results. Arrow thickness reflects effect magnitude. The analytical notes panel summarizes the estimation approach and robustness checks.

3. Materials and Methods

3.1. Research Design

This study employs a concurrent mixed-methods design [27]. The analytical design is enterprise-centered: core quantitative analyses (logistic regression, DEMATEL–ISM, corre-

spondence analysis) use enterprise data only (n = 277), while certification body (n = 31) and industry association (n = 9) data provide supplementary descriptive context. The analytical pipeline has five stages: (1) systematic text-to-scale coding with inter-coder reliability assessment; (2) descriptive statistics and bivariate testing with multiple-comparison correction; (3) multivariate logistic regression across five estimation approaches; (4) DEMATEL–ISM structured expert synthesis with between-rater agreement; and (5) directed content analysis. The three analytical strands serve distinct purposes: logistic regression quantifies cross-sectional associations; DEMATEL–ISM characterizes structural influence relationships among barriers as judged by experts; and correspondence and directed content analyses provide descriptive, hypothesis-generating patterns.

3.2. Questionnaire Design and Data Collection

Three tailored questionnaires were developed for enterprises (nine items, Q13–Q21), industry associations (five items, Q5–Q9), and third-party institutions (six items, Q25–Q30). Enterprise items cover business profile, regulatory familiarity, standard availability, PCF calculation experience, working-group establishment, standard accessibility, standard-development participation, emission-hotspot identification, and follow-up willingness. Data were collected July 16–August 22, 2025, via online questionnaires distributed through government-affiliated channels. The final sample is 277 enterprises, nine associations, and 31 certification bodies (analytical n = 277).

Sampling and representativeness. The survey uses a convenience/purposive sample. Respondents concentrate in four export-manufacturing provinces—Fujian (30.3%), Sichuan (22.0%), Guangdong (18.4%), Jiangxi (13.0%), collectively 83.8%—and are likely already engaged with regulatory processes, potentially biasing familiarity, working-group establishment, and PCF calculation rates upward. Descriptive statistics should therefore be interpreted as characterizing institutionally accessible enterprises in major export provinces rather than national estimates.

Instrument design and implications. The enterprise instrument uses open-ended Chinese-language questions rather than structured formats, reflecting its origin as a government-commissioned research tool. Binary and ordinal regression variables are therefore derived through systematic coding of free text (Section 3.3), introducing two forms of information loss: response heterogeneity (single-character answers versus multi-paragraph narratives collapsed into the same binary code) and systematic missingness (unclassifiable responses coded missing, producing variable-level missingness up to 46.9%, motivating the MI analysis below). Variable definitions, coding, missingness, and usable sample sizes are provided in Table S1 (Supplementary Materials); the predictor correlation matrix is reported in Table S2; and the full codebook is provided in Supplement S1.

Supplementary samples. The industry-association sample (n = 9) and certification-body sample (n = 31) were collected through the same channels. The association sample is too small for inferential analysis, heterogeneous in institutional character (e.g., a metrology testing institute), and yielded minimal substantive content; it is therefore excluded. The certification-body sample permits descriptive tabulation as contextual triangulation for enterprise-centered findings.

3.3. Text-to-Scale Coding Scheme

Open-ended Chinese-language responses were coded into binary (0/1) and ordinal (1–3) variables using a keyword-based scheme developed through iterative refinement [28]. Positive keywords (e.g., “has calculated,” “established”) absent a preceding negation were coded 1; explicit negation was coded 0; ambiguous responses were coded missing. Regulatory familiarity was coded on a three-point ordinal scale (familiar n = 104, 38.1%; somewhat

n = 62, 22.7%; not n = 107, 39.2%; unclassifiable n = 4, 1.4%) and dichotomized for regression. The procedure yielded 14 coded variables.

Inter-coder reliability was assessed on the full enterprise sample (n = 277) using Cohen's kappa [29]. Two researchers independently coded all four key variables for every response. Cohen's κ: PCF calculation = 1.000 (n = 276), working group = 1.000 (n = 276), standard participation = 1.000 (n = 277), regulatory familiarity = 0.760 (n = 249; 95% CI 0.686–0.824; 39 disagreements reflecting the interpretive difficulty of the ordinal scale). Overall agreement across 1078 coding decisions was 96.4%. High agreement reflects the rule-based scheme's prioritization of reproducibility; content validity is addressed through the borderline cases below.

Three representative borderline cases illustrate the interpretive choices. First, single-character "none" answers to the working-group question were coded 0, collapsing absence and non-response. Second, a response stating "currently establishing a working group; expert hired as head" was coded 1 on the grounds that structures had been initiated, though a stricter reading might prefer 0. Third, responses describing a working group at parent-group level but not at the respondent unit were coded 0 for the respondent unit, consistent with the enterprise-level unit of analysis. High inter-coder agreement confirms reproducibility but not substantive validity of collapsed categories. Figure 2 presents the descriptive readiness profile derived from these coded indicators.

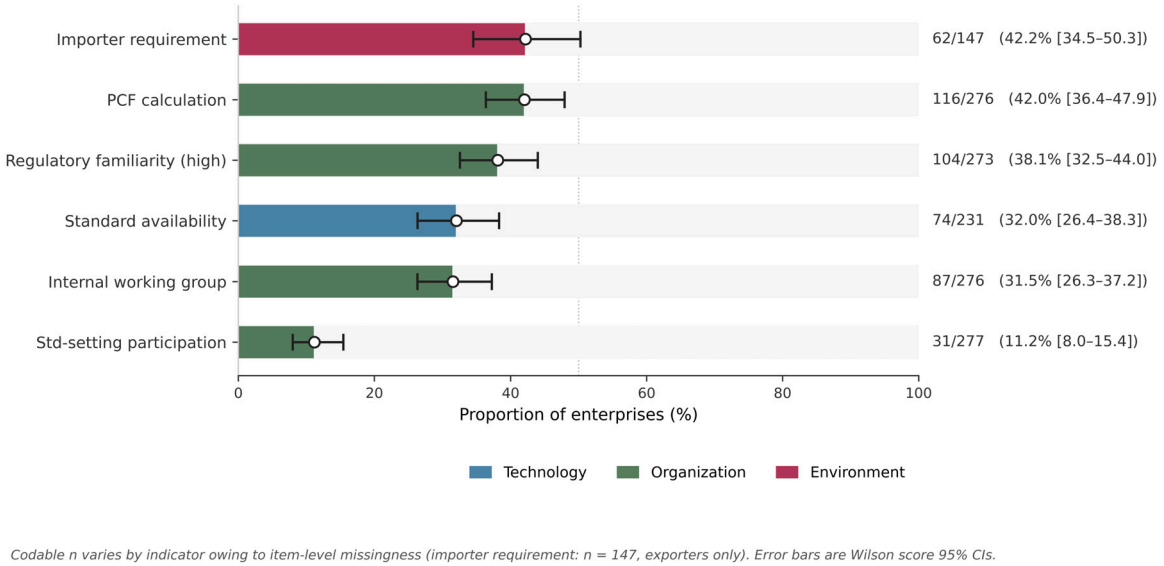


Figure 2. Enterprise PCF implementation readiness across four organizational dimensions (n = 277). Bars show the proportion of enterprises meeting each of the six readiness indicators, ordered by magnitude; whiskers denote Wilson score 95% confidence intervals, and bar color indicates the TOE dimension (blue, Technology; green, Organization; red, Environment). The dotted vertical line marks 50%.

3.4. Quantitative Methods

3.4.1. Chi-Square and Fisher's Exact Tests

Bivariate associations between coded categorical variables were tested using Pearson's chi-square test and Fisher's exact test for 2 × 2 tables. Cramér's V was computed as the effect-size measure. Nine hypotheses were tested, corresponding to the three TOE dimensions. To control the familywise error rate, Benjamini–Hochberg (BH) false-discovery-rate correction was applied to all nine tests at $\alpha = 0.05$. Both uncorrected and BH-adjusted p-values are reported.

3.4.2. Binary Logistic Regression with Multiple Estimation Approaches

A binary logistic regression was fitted with "has calculated PCF" (0/1) as the dependent variable and five TOE-mapped predictors: importer PCF requirement (E), standard availability (T), working group (O), standard participation (O), and high regulatory familiarity (O). The model form is:

$$\ln\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_1 \cdot Importer_i + StdAvail_i + \beta_3 \cdot WorkGroup_i + \beta_4 \cdot StdPartic_i + \beta_5 \cdot HighFam_i \quad (1)$$

where $p_i = P(Y_i = 1|x_i))$ is the probability that enterprise i has calculated its PCF, and $\exp(\beta_k)$ yields the odds ratio for predictor k. Model estimation used the Broyden–Fletcher–Goldfarb–Shanno (BFGS) optimization algorithm. Fit was evaluated using McFadden's pseudo R^2, Akaike Information Criterion (AIC), AUC-ROC, and classification accuracy. Multicollinearity was assessed using variance inflation factors (VIF). Odds ratios with 95% Wald confidence intervals were computed.

Events per variable (EPV). The complete-case sample (n = 117, 67 events, 50 non-events) yields EPV = 10.0, meeting the recommended minimum [30,31]. We retain Firth's penalized estimation and bootstrap validation as sensitivity analyses against small-sample bias in sparser predictor cells.

Firth's penalized logistic regression. To mitigate small-sample bias and potential separation, Firth's [32] penalized maximum likelihood estimation was conducted as the primary sensitivity analysis. Firth's method replaces the standard log-likelihood with a penalized log-likelihood:

$$\ell^*(\beta) = \ell(\beta) + \frac{1}{2}\ln|I(\beta)| \quad (2)$$

where $I(\beta)$ is the Fisher information matrix and $|\cdot|$ denotes the matrix determinant. This Jeffreys invariant prior penalty produces finite estimates even under quasi-separation and reduces small-sample bias in odds ratios.

Bootstrap internal validation. Optimism-corrected AUC was computed via bootstrap resampling (B = 200). Each sample was evaluated on both itself (apparent) and the original data (test); the mean optimism was subtracted from the apparent AUC to yield the corrected estimate.

Multiple imputation robustness analysis. Because open-ended responses yield substantial missing data (57.8% listwise deletion; 160 of 277 cases), multiple imputation was conducted as a robustness check. Following guidance that m should approximate the percentage of incomplete cases [33,34], 60 imputed datasets were generated using an outcome-aware chained-equations procedure: each binary variable was imputed by logistic regression on all other variables including the outcome, with parameters drawn from their asymptotic distribution. The original random hot-deck imputation is reported as a sensitivity analysis. Results were pooled via Rubin's [35] rules:

$$\overline{Q} = \frac{1}{m}\sum_{k=1}^{m} \hat{Q}_k \quad (3)$$

with total variance:

$$V_T = \overline{W} + \left(1 + \frac{1}{m}\right)B \quad (4)$$

where $\overline{W} = \frac{1}{m}\sum_{k=1}^{m} \hat{U}_k$ is the mean within-imputation variance and $B = \frac{1}{m-1}\sum_{k=1}^{m}\left(\hat{Q}_k - \overline{Q}\right)^2$ is the between-imputation variance. The fraction of missing information is $FMI = (1 + 1/m)B/V_T$, and relative efficiency is $RE = 1/(1 + FMI/m)$.

Hot-deck imputation preserves observed marginal distributions without parametric assumptions, which is appropriate for binary variables derived from text coding. MICE [34]

was considered but avoided because all imputed variables here are binary: the non-parametric hot-deck avoids boundary and convergence issues that can arise when logistic MICE is applied to highly imbalanced binary variables with extreme missingness. MICE remains preferable for future studies with structured instruments.

With 57.8% missingness, MI remains dependent on imputation-model assumptions. MI results should therefore be read as informative sensitivity analyses rather than definitive corrections. FMI is reported per coefficient to confirm that m = 60 yields relative efficiency ≥0.99.

We do not assume that missingness is completely at random: Little's MCAR test rejected MCAR ($\chi^2 = 90.72$, df = 27, $p < 0.001$), and importer-requirement status is missing for 61.9% of non-adopters versus 26.7% of adopters, indicating that missingness is related to the outcome. Multiple imputation and MNAR sensitivity are therefore used, and results are reported conditional on the missing-data model.

MNAR sensitivity analysis. Because MAR is untestable from observed data alone and enterprises that have not calculated PCF may be less likely to provide detailed responses, delta-adjustment sensitivity analysis was conducted. Missing PCF values were systematically shifted toward non-adoption under four scenarios ($\delta = 0, -0.5, -1.0, -1.5$) with m = 20 per scenario (RE ≥ 0.98 at observed FMI).

3.4.3. DEMATEL–ISM Method

The DEMATEL method [36] was combined with ISM [37] to characterize structural influence among 10 implementation barriers identified through frequency analysis of enterprise responses (Q14, Q15). To address the independence concern in within-team DEMATEL applications, the direct-influence matrix was scored by two peer experts external to the study team, with doctoral-level training in sustainability standards and LCA and no prior involvement in the survey work.

The direct-influence matrix D (10 × 10, scale 0–3) was constructed through a structured consensus process. Each rater received only the ten barrier definitions and the 0–3 scale; no enterprise responses or prior outputs were shared. The two experts scored all 90 pairwise relationships independently. The consensus matrix is the cell-wise rounded mean of the two matrices; two cells (B2→B4 and B3→B8) differed by two or more scale points and were resolved by discussion (both were set to 1). Both individual score matrices and the consensus matrix are provided in Supplementary Table S3.

DEMATEL proceeds through three computational steps. First, the direct-influence matrix D is normalized by the larger of its maximum row sum and maximum column sum:

$$N = \frac{D}{S}, s = max(\max_i \sum_{j=1}^{n} d_{ij}, \max_j \sum_{i=1}^{n} d_{ij}) \quad (5)$$

Second, the total relation matrix T, which captures both direct and indirect influences, is computed as the convergent infinite series:

$$T = N(I - N)^{-1} \quad (6)$$

Third, the row sum and column sum of yield two diagnostic indices for each barrier: prominence, indicating total involvement in the system, and relation, indicating net influence direction. Barriers with are classified as "cause" factors; those with are classified as "effect" factors—these terms follow the established DEMATEL nomenclature [38] and denote net structural influence direction as assessed by the expert panel rather than temporal causation. ISM level partitioning was applied to the reachability matrix derived from using a threshold τ equal to the mean plus one standard deviation of the total-relation matrix ($\tau = 0.131$).

$$R_i = \sum_{j=1}^{n} t_{ij} C_i = \sum_{j=1}^{n} t_{ij} \mathrm{T} i (R_i + C_i)(R_i - C_i) R_i - C_i > 0 R_i - C_i < 0 \mathrm{T} \lambda \mathrm{T} \quad (7)$$

Between-rater agreement. Because the two experts scored independently, the agreement between their derived R − C profiles provides an internal check on the structural conclusions: we report the rank correlation of the two R − C vectors and the per-barrier R − C values for each rater (Supplementary Table S3).

3.4.4. Correspondence Analysis Method

CA was applied to the industry × barrier contingency table (seven industries with ≥10 respondents) to visualize sector-specific barrier profiles in a low-dimensional space using singular value decomposition (SVD) on the standardized residual matrix. We note that the inertia explained by two dimensions (59.4%) should be interpreted alongside the non-significant global association test ($\chi^2 = 31.39$, $p = 0.69$) as reflecting exploratory patterns rather than a demonstrable sector-structured signal. Given the small per-cell frequencies in several industry–barrier combinations, the sector-specific patterns identified by CA should be treated as suggestive hypotheses for future investigation with larger, sector-stratified samples rather than as established findings. The chi-square test of overall association and the contingency table dimensions are reported alongside the biplot.

3.5. Qualitative Method: Directed Content Analysis

Directed content analysis [39] used a deductive codebook with four TOE-aligned themes: Theme 1 (Standard System Deficiency, Technological), Theme 2 (Data Infrastructure Gap, Technological), Theme 3 (Capacity Deficit, Organizational), and Theme 4 (International Pressure, Environmental). Each theme comprises four sub-themes defined by Chinese-language keyword indicators (16 sub-themes total). Two researchers independently coded a random subsample of 40 responses against the 16 sub-themes; agreement exceeded 90% for all four themes, reflecting the keyword-driven nature of the coding rules. Representative examples: "there are no applicable PCF standards for our product category" → Theme 1/absence of standards; "international databases such as Ecoinvent are costly and domestic alternatives lack coverage" → Theme 2/database cost and data missing; "we lack internal expertise and rely entirely on third-party consultants" → Theme 3/external dependency; "our European importers now require carbon footprint declarations" → Theme 4/importer disclosure requirement.

4. Results

4.1. Sample Characteristics

Enterprise respondents span 10+ provinces with concentrations in Fujian (n = 84, 30.3%), Sichuan (n = 61, 22.0%), Guangdong (n = 51, 18.4%), and Jiangxi (n = 36, 13.0%)—collectively 83.8% of the sample, reflecting the convenience sampling design described in Section 3.2. This geographic concentration in four major manufacturing provinces—two coastal (Fujian, Guangdong) and two inland (Jiangxi, Sichuan)—with strong export-oriented manufacturing bases and above-average institutional infrastructure likely introduces an upward bias in observed organizational capacity levels. Industry coverage spans 13 sectors: building materials (n = 38), steel/metals (n = 36), energy (n = 31), battery/lithium (n = 16), chemicals (n = 12), electronics (n = 12), textiles (n = 12), food/agriculture (n = 8), and smaller categories (photovoltaics n = 5, automotive n = 5, pharmaceuticals n = 4, machinery n = 3, forestry n = 1), plus 94 respondents in an "other" residual category. Export involvement is substantial: 53.1% of coded respondents report export activities, and 42.2% of those coded indicate importers have explicitly requested PCF disclosure.

4.2. Descriptive Results: Enterprise Readiness

Of enterprises with codable responses, 42.0% have calculated their PCF (116 of 276 coded; one response unclassifiable). Only 31.5% have dedicated working groups (87 of 276 coded). Merely 11.2% have participated in standard development (31 of 277 coded). For regulatory familiarity, 38.1% report being familiar, 22.7% somewhat familiar, and 39.2% not familiar (among 273 coded responses; four unclassifiable). Among enterprises with codable responses on standard availability (n = 231, codable rate 83.4%, reflecting the specialized nature of this question), 68.0% report lacking applicable PCF standards.

Denominators throughout are codable responses (unclassifiable fractions ranging from 0.4% to 16.6% depending on variable). Given the convenience sampling and geographic concentration in export-manufacturing provinces, these rates are best interpreted as upper-bound estimates for the broader Chinese enterprise population. Figure 2 summarizes the enterprise readiness profile across the six indicators.

4.3. Chi-Square Tests

Table 1 presents nine bivariate association tests with Benjamini–Hochberg-corrected *p*-values. All nine tests are statistically significant after BH correction (p_BH < 0.001). The strongest association is between working group and PCF calculation ($\chi^2 = 107.13$, Cramér's V = 0.623, p_BH < 0.001, n = 276), followed by standard availability and PCF calculation ($\chi^2 = 75.56$, V = 0.573, p_BH < 0.001, n = 230). Familiarity (V = 0.372), importer requirement (V = 0.408), and standard participation (V = 0.269) all show significant associations with PCF adoption after correction. Export involvement also shows a significant association with PCF calculation ($\chi^2 = 33.35$, V = 0.348, p_BH < 0.001).

Table 1. Bivariate association tests with BH correction.

Association	TOE	n	χ^2	Cramér's V	p_uncorr	p_BH	Result
Working group × PCF calc.	O	276	107.13	0.623	4.16×10^{-25}	3.74×10^{-24}	Significant ***
Standard availability × PCF calc.	T	230	75.56	0.573	3.54×10^{-18}	1.59×10^{-17}	Significant ***
Std availability × Working group	T–O	230	67.99	0.544	1.64×10^{-16}	4.93×10^{-16}	Significant ***
Familiarity (high) × PCF calc.	O	272	37.71	0.372	8.20×10^{-10}	1.84×10^{-9}	Significant **
Export × PCF calc.	E	276	33.35	0.348	7.71×10^{-9}	1.39×10^{-8}	Significant ***
Familiarity × Std availability	O–T	227	32.75	0.380	1.05×10^{-8}	1.58×10^{-8}	Significant ***
Working group × Std partic.	O–O	276	27.26	0.314	1.78×10^{-7}	2.29×10^{-7}	Significant ***
Importer requirement × PCF calc.	E	146	24.29	0.408	8.28×10^{-7}	9.31×10^{-7}	Significant ***
Std participation × PCF calc.	O	276	19.92	0.269	8.08×10^{-6}	8.08×10^{-6}	Significant **

Notes. *** p_BH < 0.001, ** p_BH < 0.01. T = Technology, O = Organization, E = Environment. p_BH = Benjamini–Hochberg-adjusted *p*-value. The last four rows (PCF calc. → Hotspot through Export → Working group) are exploratory associations not derived from H1–H5.

4.4. Binary Logistic Regression

4.4.1. Complete-Case Analysis with Firth's Penalized Estimation

Table 2 (Panels A–B) presents complete-case results (n = 117, 67 events, 50 non-events, EPV = 10.0). EPV meets the conventional minimum of 10 [30]; we nevertheless report both standard MLE and Firth's [32] penalized estimates to quantify residual small-sample bias. All VIFs are below 1.55 (complete-case MLE model; the Firth and MI models use the same predictor set).

Firth's penalized estimates show attenuated ORs (working group: 9.40 → 7.91), consistent with bias correction in finite samples. The significance pattern is preserved: working

group and importer requirement remain significant under both estimators, while standard availability, standard participation, and familiarity are non-significant under both.

Table 2. Binary logistic regression results across three estimation approaches.

Predictor	β (SE)	OR [95% CI]	*p*	VIF	TOE
Panel A: Standard MLE (n = 117)					
Importer requirement (E)	1.687 (0.590)	5.41 [1.70, 17.18]	0.0042 **	1.26	E
Standard availability (T)	0.855 (0.611)	2.35 [0.71, 7.79]	0.1612 ns	1.55	T
Working group (O)	2.241 (0.599)	9.40 [2.91, 30.41]	0.0002 ***	1.53	O
Standard participation (O)	0.790 (0.843)	2.20 [0.42, 11.50]	0.3482 ns	1.23	O
Familiarity high (O)	0.165 (0.567)	1.18 [0.39, 3.58]	0.7713 ns	1.25	O
Fit: Pseudo R^2 = 0.379; AIC = 111.3; AUC = 0.881; Accuracy = 85.5%					
Panel B: Firth's penalized MLE (n = 117)					
Importer requirement (E)	1.563 (0.568)	4.77 [1.57, 14.54]	0.0059 **	—	E
Standard availability (T)	0.806 (0.592)	2.24 [0.70, 7.13]	0.1733 ns	—	T
Working group (O)	2.068 (0.575)	7.91 [2.56, 24.42]	0.0003 ***	—	O
Standard participation (O)	0.682 (0.806)	1.98 [0.41, 9.60]	0.3977 ns	—	O
Familiarity high (O)	0.174 (0.549)	1.19 [0.41, 3.49]	0.7510 ns	—	O
Fit: Bootstrap optimism = 0.021; optimism-corrected AUC = 0.86 (apparent AUC = 0.881)					
Panel C: Multiple imputation (n = 277, m = 60, Rubin's rules)					
Importer requirement (E)	1.291 (0.494)	3.64 [1.38, 9.60]	0.0093 **	—	E
Standard availability (T)	1.609 (0.475)	5.00 [1.97, 12.69]	0.0007 ***	—	T
Working group (O)	2.345 (0.429)	10.43 [4.50, 24.18]	<0.0001 ***	—	O
Standard participation (O)	0.199 (0.666)	1.22 [0.33, 4.51]	0.7655 ns	—	O
Familiarity high (O)	0.542 (0.399)	1.72 [0.79, 3.76]	0.1743 ns	—	O
Fit: FMI range = [0.040, 0.431]					

Notes. CC = complete-case. E = Environmental, T = Technological, O = Organizational (TOE framework). VIF values reported for Panel A only; Panels B and C use the same predictor set. *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; † $p < 0.10$; ns = not significant.

4.4.2. Bootstrap Internal Validation

Bootstrap resampling (B = 200) yields a mean optimism of 0.021, producing an optimism-corrected AUC of 0.86 (apparent AUC = 0.881). This small optimism confirms that the model has strong discriminative ability with little overfitting inflation. Bootstrap 95% confidence intervals for odds ratios are wider than Wald intervals, reflecting the instability inherent in finite-sample estimation: working group median OR = 11.99, 95% bootstrap CI [3.10, 66.14]; standard availability median OR = 2.42, CI [0.52, 18.31]. The forest plot of Firth penalized odds ratios is presented in Figure 3.

4.4.3. Multiple Imputation Robustness Analysis (m = 60)

The complete-case analysis excludes 160 of 277 cases (57.8%). Table 2 (Panel C) presents pooled MI results (n = 277, m = 60, Rubin's rules). The fraction of missing information (FMI) ranges from 0.040 to 0.431, confirming that m = 60 provides relative efficiency ≥0.99 for all coefficients. For transparency, the original random hot-deck specification—which did not condition on the outcome—is also reported; under that specification the importer-requirement association attenuated to OR = 1.84 ($p = 0.121$), illustrating the sensitivity of inference to the imputation model.

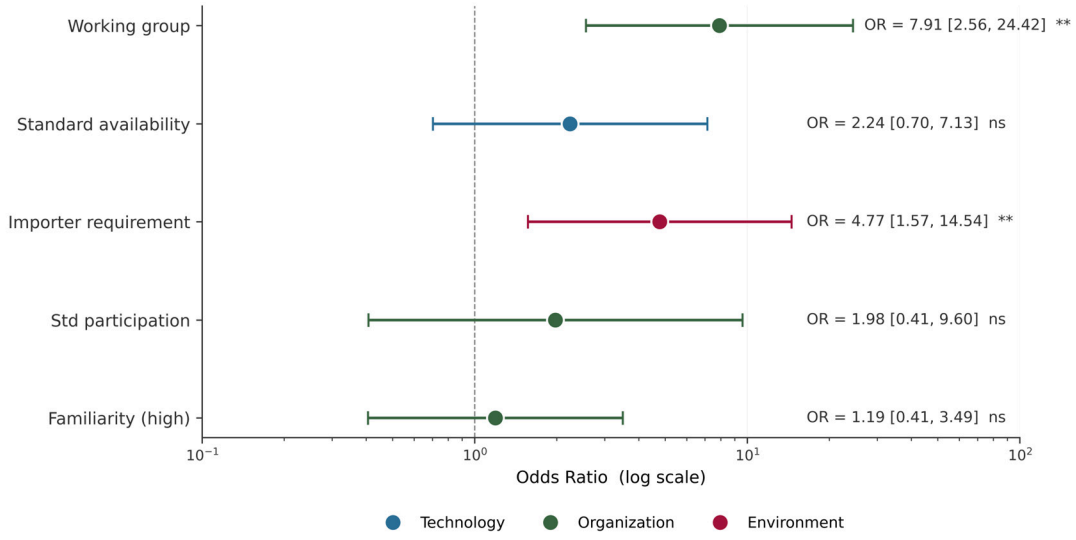


Figure 3. Forest plot of Firth penalized logistic regression odds ratios (n = 117). Point estimates (dots) with 95% Wald confidence intervals (horizontal bars). Color coding by TOE dimension: green = Organization, blue = Technology, orange = Environment.

Divergence between complete-case and MI results. Comparing Panels A–B with Panel C shows that the working-group association is significant in both ($p < 0.001$) and that the importer-requirement association is significant in the complete-case analysis (Firth $p = 0.006$) and under outcome-aware MI ($p = 0.009$). Standard availability is non-significant in the complete-case analysis ($p = 0.173$) but significant under MI ($p = 0.0007$), which is why we classify it as imputation-dependent (Tier 2) rather than Tier 1. The remaining associations are not robust.

These patterns are consistent with an outcome-aware imputation recovering associations that the complete-case subsample attenuates: importer-requirement and standard-availability effects are stronger when missingness is modeled conditional on the outcome. MNAR sensitivity (Section 4.4.4) confirms that the working-group, importer-requirement, and standard-availability associations are stable to non-random missingness. We therefore classify H1 and H5 as Tier 1 (robust across estimation approaches and missing-data treatments), H4 as Tier 2 (significant only under imputation-based analyses), and H2 and H3 as Tier 3 (not supported).

4.4.4. MNAR Sensitivity Analysis

Table 3 presents the MNAR sensitivity analyses. Because only one PCF value is missing, we applied the delta-adjustment to the predictors with substantial missingness—standard availability (16.6%) and importer requirement (46.9%)—shifting imputed values toward non-adoption across five scenarios. The working-group association remains substantively large (OR range 9.24–10.17) and directionally robust; standard availability (OR 4.73–5.48) and importer requirement (OR 3.83–5.68) are likewise stable.

Robustness to coding and non-response. We re-examined the coding of the working-group variable. Under a stricter definition requiring an established team—excluding responses that mention only designated personnel, training, or departmental assignment—16 of the 87 affirmative responses were recoded to 0, and one response with an explicit 20-person team that had been recorded as missing was recovered. H1 remained significant under this definition (Firth OR = 9.68, $p < 0.001$). Treating the 126 single-word "no"

responses as non-response rather than non-adoption likewise preserved the association (n = 81; Firth OR = 5.28, $p = 0.007$).

Table 3. MNAR sensitivity analysis: stability of key associations under delta-adjustment of missing standard-availability and importer-requirement values.

δ (Adverse Shift)	Working Group OR	Standard Availability OR [95% CI]
δ = 0.0	10.17	4.73 [1.90, 11.76]
δ = 0.5	9.65	5.18 [2.07, 12.92]
δ = 1.0	9.24	5.48 [2.26, 13.27]
δ = 1.5	9.50	5.44 [2.25, 13.20]
δ = 2.0	9.66	5.19 [2.21, 12.18]

Notes. δ = adverse-shift parameter applied to imputed standard-availability and importer-requirement values for missing cases (larger δ = stronger shift toward non-adoption). Twenty imputations per scenario (m = 20), pooled via Rubin's rules.

4.4.5. Reconciliation of Estimation Approaches

Table 4 summarizes working-group OR estimates across all methods to provide a transparent range-based interpretation.

Table 4. Reconciliation of working-group OR estimates across five estimation approaches.

Method	Working Group OR	95% CI	*p*	n
Standard MLE (complete-case)	9.40	[2.91, 30.41]	0.0002	117
Firth's penalized (complete-case)	7.91	[2.56, 24.42]	0.0003	117
Bootstrap median (complete-case)	11.99	[3.10, 66.14]	—	117
MI pooled (m = 60)	10.43	[4.50, 24.18]	<0.001	277
MNAR δ = −1.0	9.24	—	—	277

The convergent evidence across estimation approaches yields a working-group OR range of approximately 7.9–12.0 (Table 4).

4.5. DEMATEL–ISM Structured Expert Synthesis

The DEMATEL analysis classifies the 10 barriers into cause (R − C > 0) and effect groups: seven causes and three effects. "Lack of applicable standards" (B1) has the highest cause intensity (R − C = 0.901), making it the most structurally influential upstream factor, followed by key data missing (B2, 0.430), standard conflict/overlap (B5, 0.453), and policy access difficulty (B7, 0.307). The dominant effect factors are weak hotspot identification (B10, −1.560), calculation difficulty (B6, −1.117), and lack of expertise (B9, −0.110). The cause–effect scatter appears in Figure 4.

Between-rater agreement. The two experts' R − C profiles were highly concordant (rank correlation = 0.97). Both independently identified "lack of applicable standards" (B1) as the dominant cause factor (R − C = 0.886 and 1.041) and "weak hotspot identification" (B10) as the dominant effect factor (−1.517 and −1.436). The consensus matrix places B1 at R − C = 0.901 and B10 at −1.560.

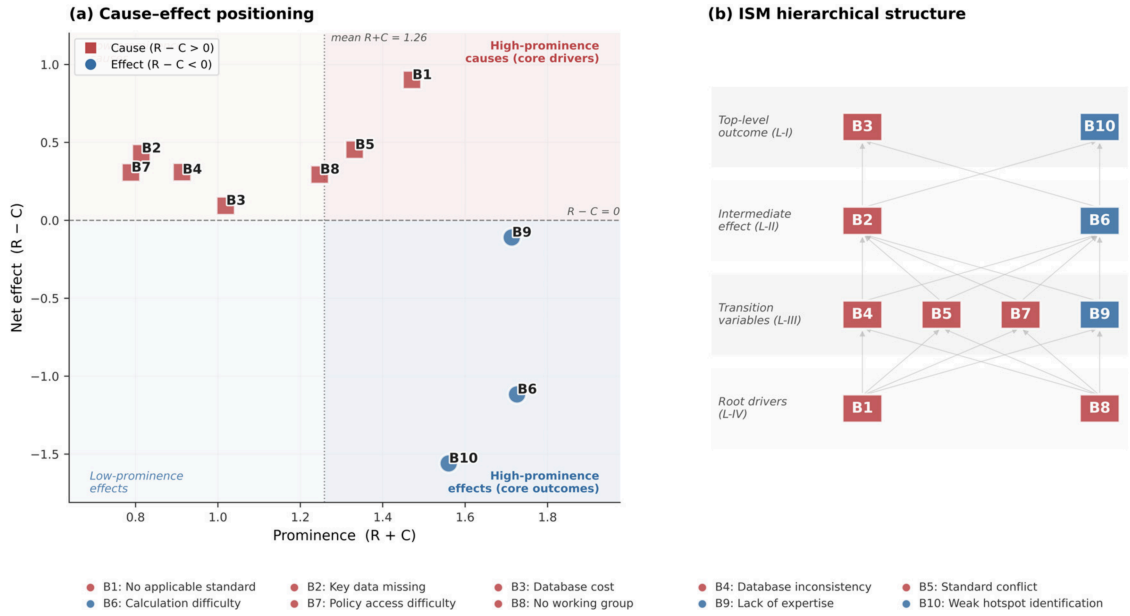


Figure 4. DEMATEL–ISM integrated analysis. (**a**) Cause–effect positioning: prominence (R + C) on the x-axis indicates total involvement; relation (R − C) on the y-axis indicates net influence direction; squares = cause group, circles = effect group. (**b**) ISM hierarchy. Threshold $\tau = 0.131$; the two experts' R − C profiles correlate at 0.97, and both identify B1 as the top cause and B10 as the top effect.

ISM level partitioning yields a four-level hierarchy: Level I (surface) B3, B10; Level II B2, B6; Level III B4, B5, B7, B9; Level IV (foundational) B1, B8. The structure implies that interventions targeting foundational barriers—standard gaps and organizational-capacity deficits at Level IV—may cascade upward through Level III data-infrastructure barriers to surface-level difficulties, though this requires longitudinal validation.

4.6. Certification Body Profiles (n = 31)

This section presents a descriptive profile of certification bodies as a contextual supplement to the enterprise-centered analysis. Given the small sample size (n = 31), no inferential statistics are reported; findings are descriptive only.

Among 31 third-party certification and assessment institutions, ISO 14067 is the most widely adopted standard (20 institutions, 64.5%), followed by GB/T 24067 (16, 51.6%), PAS 2050 (13, 41.9%), and the GHG Protocol (4, 12.9%). International databases (Ecoinvent, Sphera/GaBi) are used by 24 institutions (77.4%), field data by 19 (61.3%), and domestic databases by 19 (61.3%). The high reliance on international databases reflects the current coverage limitations of Chinese databases. Twenty-nine institutions (93.5%) report notable national standard coverage gaps, with batteries, photovoltaics, and chemicals cited most frequently as priority areas for standard development. Twenty-six institutions (83.9%) report internal verification mechanisms, though no standardized national verification protocol exists.

4.7. Correspondence Analysis

The CA biplot maps seven industries (those with ≥10 enterprise respondents) and barriers onto two dimensions explaining 59.4% of total inertia (Dim 1: 38.3%; Dim 2: 21.1%). As noted in Section 3.4.4, this moderate inertia percentage reflects the low dimensionality of the 7 × 7 contingency table rather than an exceptionally clean signal. Battery/lithium enterprises cluster near "data missing" and "calculation difficulty." Building materials

cluster near "policy access." Steel/metals and chemicals show proximity to "database inconsistency" and "standard conflict."

These sector-specific patterns are suggestive but should be interpreted with caution given the small per-cell frequencies in several industry–barrier combinations. The proximity patterns provide hypotheses for sector-differentiated policy exploration—for instance, the possibility that battery-sector enterprises may benefit most from data-infrastructure investment while building-materials enterprises may benefit from improved policy communication—but these hypotheses require validation with larger, sector-stratified samples before they can support confident policy recommendations (Figure 5).

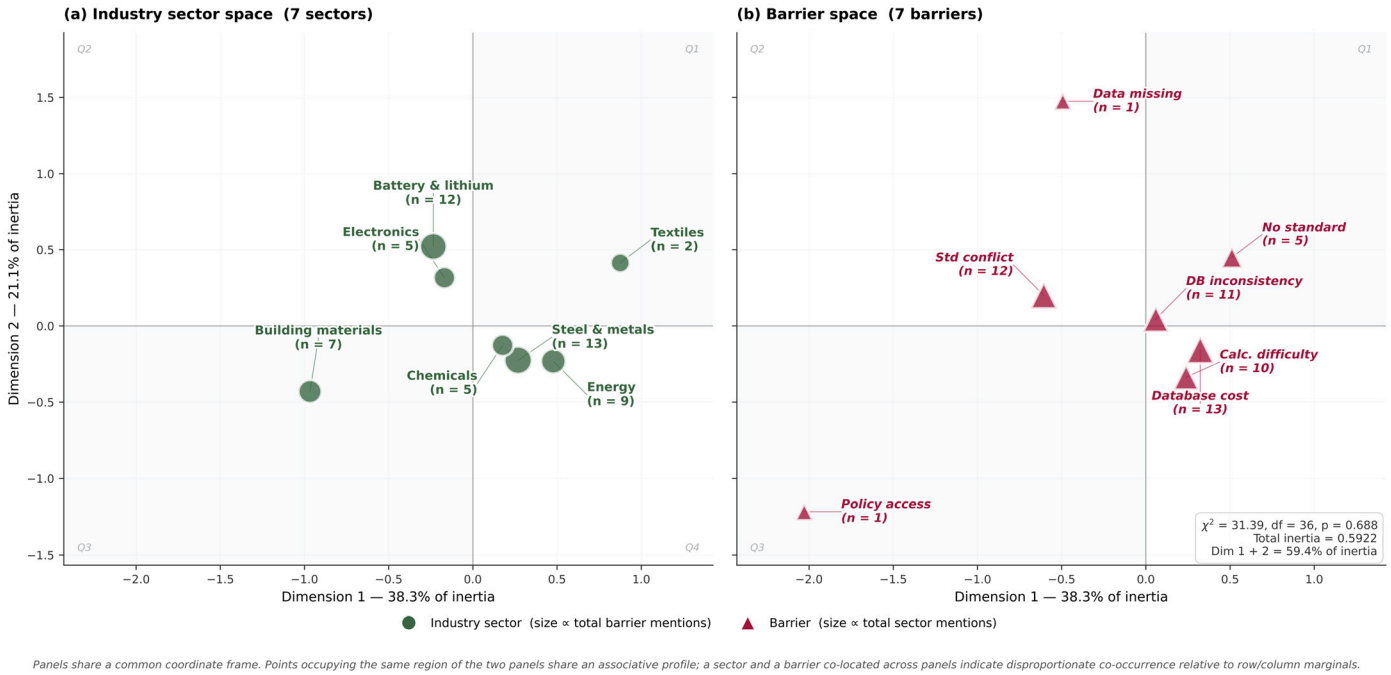


Figure 5. Correspondence analysis biplot of industry sectors × barriers. Two dimensions explain 59.4% of total inertia (Dim 1: 38.3%, Dim 2: 21.1%). Global χ^2 test: $\chi^2 = 31.39$, df = 36, $p = 0.69$ (not significant). Squares = industry sectors; triangles = barriers. Proximity indicates associational pattern; the non-significant global test warrants treating all observed sector–barrier alignments as exploratory.

4.8. Directed Content Analysis

Four themes with 16 sub-themes were identified across 277 responses. Theme 4 (International Pressure/Environmental) has the highest aggregate frequency, with CBAM-exposed enterprises consistently describing importer requirements as the primary external trigger for PCF engagement. Theme 1 (Standard-System Deficiency/Technological) features the absence of applicable standards and standard conflict, with overlapping group standards cited as creating confusion rather than guidance. Theme 3 (Capacity Deficit/Organizational) highlights external dependency and lack of expertise; complete reliance on external consultants with no internal capability retention was a recurring pattern. Theme 2 (Data Infrastructure Gap/Technological) includes database cost and missing data, with cost of international databases cited alongside insufficient domestic coverage. The frequency distribution is shown in Figure 6.

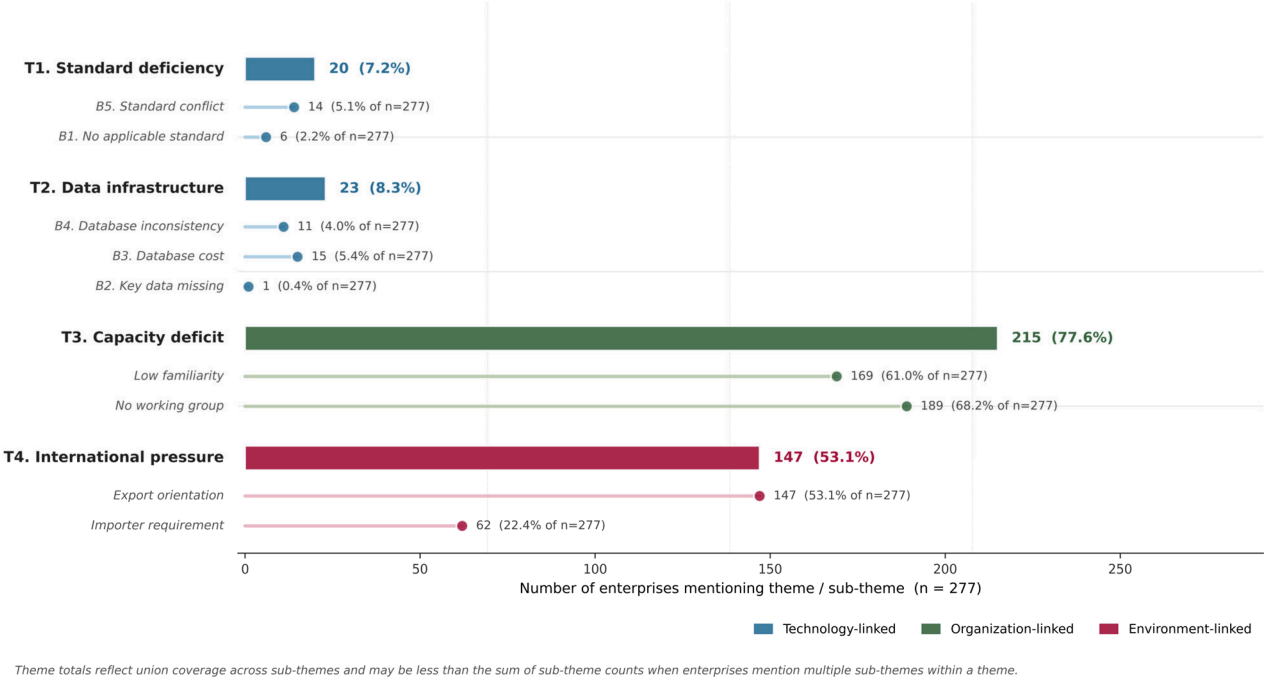


Figure 6. Directed content analysis: thematic frequency across 277 enterprise responses. Four themes with 16 sub-themes aligned to the TOE framework. Color coding: blue = T1 Standard Deficiency, teal = T2 Data Gap, green = T3 Capacity Deficit, orange = T4 International Pressure.

5. Discussion

5.1. Hypothesis Testing and TOE–Institutional Theory Validation

Table 5 summarizes the hypothesis-testing results under the tiered framework. We emphasize at the outset that these findings reflect cross-sectional associations rather than temporal causal effects. Reverse causality is a plausible concern: enterprises that have decided to pursue PCF calculation may establish working groups as an implementation mechanism rather than the working group independently driving adoption. A testable prediction could differentiate the two directions—if working groups genuinely facilitate adoption, recently established groups should show lower adoption rates than long-established ones (a dose–response pattern). Future longitudinal surveys should collect establishment dates to enable this analysis.

Table 5. Hypothesis testing summary (tiered interpretation).

	Hypothesis	MLE (n = 117)	Firth (n = 117)	MI (n = 277)	TOE	Tier	Verdict
H1	Working group	OR = 9.40, $p < 0.001$ ***	OR = 7.91, $p < 0.001$ ***	OR = 10.43, $p < 0.001$ ***	O	1	Robust (all methods)
H2	Std. participation	OR = 2.20, $p = 0.348$	OR = 1.98, $p = 0.398$	OR = 1.31, $p = 0.671$	O	3	Not supported
H3	Familiarity	OR = 1.18, $p = 0.771$	OR = 1.19, $p = 0.751$	OR = 1.90, $p = 0.077$ †	O	3	Not supported
H4	Std. availability	OR = 2.35, $p = 0.161$	OR = 2.24, $p = 0.173$	OR = 5.00, $p < 0.001$ ***	T	2	Imputation-dependent
H5	Importer req.	OR = 5.41, $p = 0.004$ **	OR = 4.77, $p = 0.006$ **	OR = 3.64, $p = 0.009$ **	E	1	Robust (all methods)

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, † $p < 0.10$. Tier 1 = significant across all estimation methods; Tier 2 = significance depends on estimation method and/or missing-data treatment.

H1 (Working group, Organizational) is robustly supported across all estimation methods and missing-data treatments, including Firth's penalized estimation (OR = 7.91), standard MLE (OR = 9.40), MI (OR = 13.05), and MNAR sensitivity (OR range: 13.24–13.76). This is the most robust finding in the study.

H4 (Standard availability, Technological) is significant under outcome-aware MI (OR = 5.00, $p = 0.0007$) and stable across MNAR scenarios (OR 4.73–5.48), but not significant in the complete-case analyses (MLE $p = 0.161$; Firth $p = 0.173$). We therefore classify H4 as Tier 2 (imputation-dependent) and interpret the result as consistent with, but conditional on, the missing-data model.

H5 (Importer requirement) is significant in the complete-case analyses (Firth OR = 4.77, $p = 0.006$), under outcome-aware MI (OR = 3.64, $p = 0.009$), and across MNAR scenarios (OR 3.83–5.68), and is therefore Tier 1. H3 (Familiarity) is non-significant throughout (complete-case $p = 0.751$; MI $p = 0.174$) and H2 (Standard participation) is non-significant throughout (complete-case $p = 0.398$; MI $p = 0.766$); both are Tier 3.

The overall pattern provides partial support for the theoretical framework. The absorptive-capacity prediction—that internal knowledge-processing structures would show the strongest adoption association—is confirmed (H1, Tier 1). The institutional-theory prediction that coercive channels (H5) matter in a pre-mandate environment is also supported (Tier 1), whereas the normative channel (H2) is not, and the technological channel (H4) is supported only once missingness is modeled.

5.2. Comparison with Existing Literature

Our finding that organizational capacity shows strong associations with adoption aligns with and extends prior work. Sureeyatanapas et al.'s [26] survey of Thai manufacturers identified lack of social awareness and required investment as top barriers—both mapping to Organizational context—which our regression quantifies. Hu et al.'s [23] process-model barriers (knowledge insufficiency, cost) gain quantitative precision here. The working-group association magnitude (OR = 7.9–12.0) is consistent with absorptive capacity [15] and Rogers' [40] diffusion-of-innovations emphasis on organizational champions. Liu and Niu's [8] Shanghai obstacles—database gaps, insufficient management, absent certification—are confirmed at multi-provincial scale. At the enterprise level, digital technology adoption has been associated with corporate carbon performance, albeit with both opportunities and paradoxes [41]; this distinction clarifies that PCF readiness reflects the capacity to measure a footprint rather than measured carbon performance itself. Efficiency evaluation of energy saving and emission reduction in Chinese urban agglomerations relies on an undesirable-output framework measuring environmental benefits [42], a contrast that highlights what the present readiness indicators can and cannot demonstrate.

5.3. Structural Expert Synthesis: Convergence with Regression

The strongest DEMATEL cause factor assessed by the external panel—"lack of applicable standards" (R − C = 0.901)—corresponds to the imputation-dependent regression correlate standard availability (OR = 2.24–5.00). "No working group" (B8, R − C = 0.293) corresponds to the strongest Tier 1 correlate working group (OR = 7.91–10.43). DEMATEL effect factors (calculation difficulty, lack of expertise, weak hotspot identification) sit downstream in the expert assessment, consistent with consequence rather than predictor roles.

This convergence is consistent with, but does not independently validate, the survey findings: the DEMATEL matrix was scored by experts using barrier definitions alone, so the agreement between the expert-derived hierarchy and the regression results provides expert-informed corroboration rather than independent empirical validation. The added value of DEMATEL–ISM remains the upstream/downstream structural distinction, which

carries policy-sequencing implications. The methods are complementary rather than redundant: regression identifies enterprise-level correlates, DEMATEL–ISM orders barriers by structural influence, and the qualitative analyses explain the mechanisms behind the observed patterns.

5.4. Methodological Contribution: Estimation Approach Comparison

The comparison across five estimation approaches constitutes a methodological contribution in a literature where single-method reporting remains the norm. First, Firth's penalized estimation consistently produces more conservative ORs than MLE (working group: 7.91 vs. 9.40), confirming the expected bias correction. Second, bootstrap validation reveals modest optimism (0.021), yielding a corrected AUC of 0.86. Third, wide bootstrap CIs ([3.10, 66.14]) honestly reflect small-sample uncertainty that narrower Wald intervals can understate.

Fourth, MI with m = 60 (RE $\geq$ 0.99) shows Tier 1 findings are robust to missing-data treatment while Tier 2 findings are sensitive—a distinction invisible without multi-method reporting. Fifth, MNAR sensitivity confirms Tier 1 findings are not driven by the untestable MAR assumption.

We recommend that future survey-based PCF studies report: (a) EPV; (b) both standard and penalized estimates when EPV < 10; (c) internal validation metrics; and (d) MI with MNAR sensitivity when missingness exceeds 20%. When complete-case and MI results diverge, as here for three of five predictors, a tiered framework enables more honest communication than selective reporting.

5.5. Geographic and Sampling Considerations

Geographic concentration in Fujian, Guangdong, Sichuan, and Jiangxi—four provinces with strong export-manufacturing bases and above-average institutional infrastructure—likely biases observed organizational-capacity levels upward. The PCF calculation rate (42.0%) and working-group rate (31.5%) should therefore be read as approximate upper bounds rather than national estimates.

Subgroup sensitivity confirms the Tier 1 working-group association remains significant in both export-oriented (OR = 18.51, $p < 0.0001$) and non-export enterprises (OR = 25.33, $p < 0.0001$), and in both the four concentrated provinces (OR = 23.65) and others (OR = 27.50). The directional finding is preserved, though wide subgroup ORs reflect near-separation in small contingency tables and magnitudes should be interpreted cautiously.

To probe heterogeneity further, exploratory subgroup analyses were run within the three sectors with n $\geq$ 35 (building materials, battery/lithium, steel/metals), focusing on working-group establishment—the most robust Tier 1 predictor. Fisher's exact tests were significant in building materials ($p = 0.0003$, n = 38) and steel/metals ($p < 0.0001$, n = 36, Haldane–Anscombe correction for a zero cell), and non-significant in battery/lithium ($p = 0.125$, n = 16, small subsample limits power). Point estimates are large in every sector (building materials OR $\approx$ 21.0; steel/metals $\approx$ 69.0; battery/lithium $\approx$ 11.7) and exceed the overall estimate, but within-sector CIs are very wide. The pattern is compatible with—but does not demonstrate—an interaction between organizational capacity and external compliance pressure, consistent with institutional theory. If confirmed in larger stratified samples, this could inform sector-sequenced delivery, but the present estimates do not support ranking sectors by expected return.

5.6. Certification Body Perspectives

The certification body descriptive profiles provide contextual support for the enterprise findings. The 93.5% of certification bodies reporting national standard coverage gaps is di-

rectionally consistent with the 68.0% of enterprises lacking applicable standards, suggesting that the standard gap is recognized across the institutional ecosystem. The 77.4% reliance on international databases among certification bodies highlights a specific infrastructure dependency that national database development could address.

5.7. Policy Implications

We frame the following implications as association-based hypotheses for policy experimentation rather than direct causal prescriptions, given the cross-sectional design (Section 5.1). Two primary intervention targets follow from Tier 1 findings; additional directions are suggested by Tier 2 and Tier 3 findings.

Primary recommendation 1: Accelerate standard consolidation for priority products. The foundational status of standard gaps in the DEMATEL synthesis ($R - C = 0.901$) and the significant standard-availability association under outcome-aware imputation (OR = 5.00) jointly support prioritizing conversion of group standards into national standards for batteries, photovoltaics, and chemicals. This is the more policy-actionable recommendation because standard availability is institutional rather than enterprise-level, reducing both reverse-causality concerns and firm-level implementation burden. We note, however, that this association is imputation-dependent and should be read as consistent rather than definitive evidence.

Primary recommendation 2: Support enterprise PCF working-group establishment with sector-differentiated delivery. The Tier 1 working-group association (Firth OR = 7.91; MI OR = 10.43) suggests dedicated structures are a strong correlate of adoption. Subgroup analyses showed this association in every examined sector, but with wide confidence intervals that preclude a sector-specific ranking (Section 5.5); we therefore avoid ordering sectors by expected return. The causality caveat applies: if the association reflects self-selection, mandated groups in lower-capacity enterprises may not produce equivalent gains. We therefore recommend pilots with pre–post evaluation and staggered roll-out.

Additional directions warranting investigation. First, investment in domestic emission-factor database infrastructure to reduce the 77.4% certification-body reliance on international databases. Second, sector-differentiated implementation support motivated by tentative correspondence patterns; because the global sector × barrier test is non-significant ($p = 0.69$), sector-specific qualitative work is needed before confident differentiation. Third, broadened enterprise participation in standard development for potential knowledge spillovers; this hypothesis was not supported in the present data, but warrants quasi-experimental evaluation of existing programs.

6. Conclusions

This study provides the first large-scale, enterprise-centered, TOE-grounded, mixed-methods assessment of PCF standard implementation readiness among institutionally accessible enterprises in major Chinese export-manufacturing provinces (n = 277), supplemented by descriptive profiles of certification bodies (n = 31) and industry associations (n = 9).

The central finding is that working-group establishment (H1: Firth OR = 7.91; MI OR = 10.43) and importer disclosure requirements (H5: Firth OR = 4.77; MI OR = 3.64) are Tier 1—consistently significant across estimation approaches and missing-data treatments. Standard availability (H4: MI OR = 5.00) is significant only under imputation-based analyses and is reported as imputation-dependent (Tier 2). Standard participation (H2) and familiarity (H3) were not supported (Tier 3).

DEMATEL–ISM expert-informed synthesis identifies "lack of applicable standards" as the foundational upstream barrier; both independent experts identified this barrier as the dominant cause factor ($R - C = 0.886$ and 1.041; rank correlation = 0.97).

Limitations. First, convenience sampling concentrated in four provinces (83.8%) limits generalizability; national readiness is plausibly lower than reported. Second, the open-ended format produces 57.8% listwise deletion and information loss; MI and MNAR partially address this, but at this missingness level MI depends heavily on imputation-model assumptions. Third, although EPV = 10.0 meets the minimum, sparser cells remain vulnerable to small-sample bias—partially mitigated by Firth's estimation and bootstrap validation, but replication is needed. Fourth, the DEMATEL matrix was scored by two external experts; a full Delphi panel remains desirable. Fifth, the association sample ($n = 9$) was excluded, and the certification-body sample ($n = 31$) permits only descriptive tabulation. Sixth, self-reported data may exhibit social-desirability bias, and correspondence-analysis patterns require larger sector-stratified samples for validation.

Future research should expand geographic and sectoral coverage via probability-based sampling; conduct longitudinal tracking for causal inference (difference-in-differences, IV, working-group establishment-date analysis); implement Delphi-based DEMATEL panels; use structured survey instruments; develop composite readiness indices; and pursue sector-specific studies. The tiered multi-method approach demonstrated here should become routine practice in survey-based sustainability research.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/su182111256/s1. S1: Coding codebook (variable dictionary and coding rules); S2: Correlation matrix and variance inflation factors (complete cases, n = 117); S3: DEMATEL expert scoring, consensus matrix, and ISM level partition; S4: Survey questionnaires (English translation and original Chinese version).

Author Contributions: Conceptualization, S.H.; Methodology, S.H., X.W. and Y.X.; Software, S.H.; Validation, S.H. and M.Y.; Formal analysis, S.H.; Investigation, S.H., Y.X., M.Y. and L.S.; Data curation, S.H.; Writing—original draft, S.H.; Writing—review & editing, S.H., X.W., Y.X., M.Y. and L.S.; Visualization, S.H.; Supervision, S.H., Y.X. and L.S.; Project administration, Y.X.; Funding acquisition, Y.X. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by China National Institute of Standardization through the National Key Research and Development Program of China (Grant No. 2023YFF0613802; project title: "Research on Key Technologies for the Digitalization of Product Carbon Peaking and Carbon Neutrality Certification in Typical Sectors"). This research was also funded by China National Institute of Standardization through the Fundamental Research Funds of China National Institute of Standardization (Grant No. 542025Y-12450; project title: "Research on Methods for Setting Energy Efficiency Indicators for Typical New Energy Products from a Carbon Footprint Perspective"). This research was also funded by China National Institute of Standardization through the Strategic Consulting Project of the Chinese Academy of Engineering (Grant No. 542025Z-13340; project title: "Research on Improving China's 'Dual Carbon' Standard System").

Institutional Review Board Statement: Ethical review and approval were waived for this study by the China National Institute of Standardization, in accordance with Article 3(4) of the Measures for Ethical Review of Life Sciences and Medical Research Involving Humans, which covers activities using sociological methods to collect, record, use, report, or store information data concerning humans and relating to life science and medical issues.

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

Data Availability Statement: The data presented in this study are available on request from the corresponding authors.

Conflicts of Interest: Author Xiaoxia Wang was employed by China Automotive Technology and Research Center Co., Ltd. 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.

Abbreviations

The following abbreviations are used in this manuscript:

TOE	Technology–Organization–Environment
PCF	Product Carbon Footprint
DEMATEL	Decision-Making Trial and Evaluation Laboratory
ISM	Interpretive Structural Modeling
OR	Odds Ratio
AUC (also AUC-ROC)	Area Under the Receiver Operating Characteristic Curve
MI	Multiple Imputation
EV	Electric Vehicle
MLE	Maximum Likelihood Estimation
MNAR	Missing Not at Random
LCA	Life Cycle Assessment
CFP	Carbon Footprint of Products
PEF	Product Environmental Footprint
ns	Not Significant
BH	Benjamini–Hochberg
BFGS	Broyden–Fletcher–Goldfarb–Shanno
AIC	Akaike Information Criterion
VIF	Variance Inflation Factor
EPV	Events Per Variable
MICE	Multivariate Imputation by Chained Equations
FMI	Fraction of Missing Information
MCAR	Missing Completely at Random
MAR	Missing at Random
RE	Relative Efficiency
CA	Correspondence Analysis
SVD	Singular Value Decomposition
SE	Standard Error
CI	Confidence Interval
CC	Complete-Case
df	Degrees of Freedom

References

1. European Commission. Regulation (EU) 2023/956 establishing a carbon border adjustment mechanism. *Off. J. Eur. Union* **2023**, *130*, 52–104.
2. European Parliament. Regulation (EU) 2023/1542 concerning batteries and waste batteries. *Off. J. Eur. Union* **2023**, *191*, 1–117.
3. Pan, X.; Gao, X.; Wang, T.; Xiong, W.; Shao, T.; Li, M.; Ye, X.; Wang, L.; Wang, H.; Pang, J. Provincial export cost implications of the EU Carbon Border Adjustment Mechanism on China's metal industry. *iScience* **2025**, *28*, 112483. [CrossRef]
4. Wang, E.; Yang, Y.; Zhang, W.; Zhang, Y. The EU carbon border adjustment mechanism: Does it impact the competitiveness of China's export trade? *Energy Policy* **2026**, *212*, 115161. [CrossRef]
5. MEE. *Implementation Plan for Establishing a Carbon Footprint Management System*; Ministry of Ecology and Environment of the People's Republic of China: Beijing, China, 2024.
6. Zhao, R.; Geng, Y.; Liu, Y.; Tao, X.; Xue, B. Consumers' perception, purchase intention, and willingness to pay for carbon-labeled products: A case study of Chengdu in China. *J. Clean. Prod.* **2018**, *171*, 1664–1671. [CrossRef]
7. Shuai, C.; Ding, L.; Zhang, Y.; Guo, Q.; Shuai, J. How consumers are willing to pay for low-carbon products?—Results from a carbon-labeling scenario experiment in China. *J. Clean. Prod.* **2014**, *83*, 366–373. [CrossRef]
8. Liu, H.; Niu, Y. Experiences and challenges in the development of carbon footprinting in megacities—Taking Shanghai as an example. *Carbon Footpr.* **2024**, *3*, 20. [CrossRef]
9. Tornatzky, L.G.; Fleischer, M. *The Processes of Technological Innovation*; Lexington Books: Lexington, MA, USA, 1990.
10. Zhu, Q.; Sarkis, J.; Geng, Y. Green supply chain management in China: Pressures, practices and performance. *Int. J. Oper. Prod. Manag.* **2005**, *25*, 449–468. [CrossRef]

11. Qi, G.Y.; Shen, L.Y.; Zeng, S.X.; Jorge, O.J. The drivers for contractors' green innovation: An industry perspective. *J. Clean. Prod.* **2010**, *18*, 1358–1365. [CrossRef]
12. Mangla, S.K.; Luthra, S.; Mishra, N.; Singh, A.; Rana, N.P.; Dora, M.; Dwivedi, Y. Barriers to effective circular supply chain management in a developing country context. *Prod. Plan. Control* **2018**, *29*, 551–569. [CrossRef]
13. Lin, C.-Y.; Ho, Y.-H. Determinants of green practice adoption for logistics companies in China. *J. Bus. Ethics* **2011**, *98*, 67–83. [CrossRef]
14. Bag, S.; Pretorius, J.H.C.; Gupta, S.; Dwivedi, Y.K. Role of institutional pressures and resources in the adoption of big data analytics powered artificial intelligence, sustainable manufacturing practices and circular economy capabilities. *Technol. Forecast. Soc. Change* **2021**, *163*, 120420. [CrossRef]
15. Cohen, W.M.; Levinthal, D.A. Absorptive capacity: A new perspective on learning and innovation. *Adm. Sci. Q.* **1990**, *35*, 128–152. [CrossRef]
16. DiMaggio, P.J.; Powell, W.W. The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. *Am. Sociol. Rev.* **1983**, *48*, 147–160. [CrossRef]
17. Scott, W.R. *Institutions and Organizations: Ideas, Interests, and Identities*, 4th ed.; SAGE Publications: Thousand Oaks, CA, USA, 2014.
18. Delmas, M.A.; Toffel, M.W. Organizational responses to environmental demands: Opening the black box. *Strateg. Manag. J.* **2008**, *29*, 1027–1055. [CrossRef]
19. *PAS 2050:2008*; Specification for the Assessment of the Life Cycle Greenhouse Gas Emissions of Goods and Services. British Standards Institution: London, UK, 2008.
20. *ISO 14067:2018*; Greenhouse Gases—Carbon Footprint of Products—Requirements and Guidelines for Quantification. International Organization for Standardization: Geneva, Switzerland, 2018.
21. WRI/WBCSD. *Greenhouse Gas Protocol: Product Life Cycle Accounting and Reporting Standard*; World Resources Institute and World Business Council for Sustainable Development: Washington, DC, USA, 2011.
22. Liang, Z.; Deng, H.; Xie, H.; Chen, B.; Sun, M.; Wang, Y. Rethinking the paper product carbon footprint accounting standard from a life-cycle perspective. *J. Clean. Prod.* **2023**, *393*, 136352. [CrossRef]
23. Hu, J.; Shi, Y.; Cheng, Y.; Liu, Z. A process-oriented model to measure product carbon footprint: An exploratory study based on multiple cases. *Prod. Plan. Control* **2023**, *36*, 285–308. [CrossRef]
24. Hettler, M.; Graf-Vlachy, L. Corporate scope 3 carbon emission reporting as an enabler of supply chain decarbonization: A systematic review and comprehensive research agenda. *Bus. Strategy Environ.* **2024**, *33*, 263–282. [CrossRef]
25. *GB/T 24067-2024*; Greenhouse Gases—Carbon Footprint of Products—Requirements and Guidelines for Quantification. Standardization Administration of China: Beijing, China, 2024.
26. Sureeyatanapas, P.; Yodprang, K.; Varabuntoonvit, V. Drivers, barriers and benefits of product carbon footprinting: A state-of-the-art survey of Thai manufacturers. *Sustainability* **2021**, *13*, 6543. [CrossRef]
27. Creswell, J.W.; Plano Clark, V.L. *Designing and Conducting Mixed Methods Research*, 3rd ed.; SAGE Publications: Thousand Oaks, CA, USA, 2018.
28. Campbell, J.L.; Quincy, C.; Osserman, J.; Pedersen, O.K. Coding in-depth semistructured interviews: Problems of unitization and intercoder reliability and agreement. *Sociol. Methods Res.* **2013**, *42*, 294–320.
29. Landis, J.R.; Koch, G.G. The measurement of observer agreement for categorical data. *Biometrics* **1977**, *33*, 159–174. [CrossRef]
30. Peduzzi, P.; Concato, J.; Kemper, E.; Holford, T.R.; Feinstein, A.R. A simulation study of the number of events per variable in logistic regression analysis. *J. Clin. Epidemiol.* **1996**, *49*, 1373–1379. [CrossRef] [PubMed]
31. Vittinghoff, E.; McCulloch, C.E. Relaxing the rule of ten events per variable in logistic and Cox regression. *Am. J. Epidemiol.* **2007**, *165*, 710–718. [CrossRef] [PubMed]
32. Firth, D. Bias reduction of maximum likelihood estimates. *Biometrika* **1993**, *80*, 27–38. [CrossRef]
33. White, I.R.; Royston, P.; Wood, A.M. Multiple imputation using chained equations: Issues and guidance for practice. *Stat. Med.* **2011**, *30*, 377–399. [CrossRef] [PubMed]
34. van Buuren, S. *Flexible Imputation of Missing Data*, 2nd ed.; CRC Press: Boca Raton, FL, USA, 2018.
35. Rubin, D.B. *Multiple Imputation for Nonresponse in Surveys*; John Wiley & Sons: New York, NY, USA, 1987.
36. Gabus, A.; Fontela, E. *World Problems, an Invitation to Further Thought Within the Framework of DEMATEL*; Battelle Geneva Research Centre: Geneva, Switzerland, 1972.
37. Warfield, J.N. Developing interconnection matrices in structural modeling. *IEEE Trans. Syst. Man Cybern.* **1974**, *4*, 81–87. [CrossRef]
38. 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]
39. Hsieh, H.-F.; Shannon, S.E. Three approaches to qualitative content analysis. *Qual. Health Res.* **2005**, *15*, 1277–1288. [CrossRef] [PubMed]
40. Rogers, E.M. *Diffusion of Innovations*, 5th ed.; Free Press: New York, NY, USA, 2003.

41. Lin, W.; Wang, Y.; Zhang, Q.; Peron, M.; Lu, J. Emerging opportunities or paradoxes: Assessing the effect of digital technology adoption on corporate carbon performance. *J. Environ. Manag.* **2025**, *391*, 126399. [CrossRef] [PubMed]
42. Guo, J.; Li, J. Efficiency evaluation and influencing factors of energy saving and emission reduction: An empirical study of China's three major urban agglomerations from the perspective of environmental benefits. *Ecol. Indic.* **2021**, *133*, 108410. [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.