RESEARCH PAPER

Evaluating Efficiency Differences in Supply Chain Practices: Evidence from Indian SMEs Using the Mann–Whitney U Test

Dr. Karim Soliman[1*], Dr. Juhi Vinod Mehta[2], Dr. S. Jyothirmaye Reddy[3], Dr. Maqsood Husain Junnaidi[4], Dr. Rajinder Kumar[5], Dr. Ajeet Singh[6]

[1*]College of Business Administration, University of Business and Technology, Jeddah, Saudi Arabia; Arab Academy for Science, Technology, and Maritime Transport, Cairo, Egypt | ORCID: https://orcid.org/0000-0002-9216-6360 | Email: karim.mohamed@aast.edu

[2]Associate Professor, Department of Commerce, Jain Deemed-to-be University, Bengaluru, Karnataka, India | Email: juhimehta1314@gmail.com

[3]Associate Professor, Faculty of Management Studies, Marwadi University, Rajkot, Gujarat, India | Email: sathi.reddy@marwadieducation.edu.in

[4]Department of Supply Chain Management, College of Business, University of Jeddah, Jeddah, Saudi Arabia | Email: Mhjunnaidi@uj.edu.sa

[5]Associate Professor, Faculty of Computing, Guru Kashi University, Talwandi Sabo, Bathinda, Punjab, India | ORCID: https://orcid.org/0009-0001-4129-0388 | Email: drrajinder1983@gmail.com

[6]Assistant Professor, Department of Physics, School of Sciences, IFTM University, Moradabad, Uttar Pradesh, India

***Corresponding Author:** Dr. Karim Soliman, College of Business Administration, University of Business and Technology, Jeddah, Saudi Arabia | ORCID: https://orcid.org/0000-0002-9216-6360 | Email: karim.mohamed@aast.edu

Abstract

Small and medium enterprises (SMEs) supply the majority of India's manufacturing employment, yet the efficiency consequences of how they organise their supply chains remain thinly documented relative to large-firm evidence. This study compares supply chain efficiency between SMEs operating integrated supply chain practices and those operating traditional, fragmented practices, using a cross-sectional survey of 400 registered Indian SMEs and the Mann–Whitney U test. Efficiency is measured across six domains procurement and sourcing, inventory management, logistics and distribution, information integration, supplier relationships, and demand planning plus a composite index. Because the response distributions are ordinal, non-normal in every domain, and heteroscedastic across arms, a rank-based test is the appropriate inferential instrument; the analysis reports tie-corrected *z* statistics, Cliff's δ with bootstrap confidence intervals, and Hodges–Lehmann shift estimates rather than *p*-values alone. Integrated-practice firms rank significantly higher on the composite index (U = 24,416.0, z = 5.90, $p < .001$, δ = 0.360) and on five of the six domains after Holm correction. The advantage is largest for information integration (δ = 0.416) and demand planning (δ = 0.307), and statistically indistinguishable from zero for supplier relationships (δ = 0.063, p = .299). The composite result survives rank analysis of covariance on firm size, enterprise resource planning presence, export orientation and age, and an exactly matched 113-pair subsample. The pattern suggests that practice integration in Indian SMEs delivers its returns through informational and planning coordination rather than through relational quality, which appears to be achieved by traditional firms through other means. Implications for SME managers, cluster development agencies and non-parametric practice in operations research are discussed.

Keywords: supply chain practices; operational efficiency; small and medium enterprises; India; Mann–Whitney U test; Cliff's delta; non-parametric inference

How to cite this article: Soliman K, Mehta JV, Reddy SJ, Junnaidi MH, Kumar R, Singh A. Evaluating efficiency differences in supply chain practices: evidence from Indian SMEs using the Mann–Whitney U test. Int J Drug Deliv Technol. 2026;16(73s): 1282-1298. DOI: 10.25258/ijddt.16.73s.134

1. Introduction

Indian small and medium enterprises occupy an awkward position in supply chain research. They are simultaneously indispensable and under-theorised. They account for a very large share of the country's manufacturing units, industrial employment and export value, and they sit as first- and second-tier suppliers inside the domestic supply chains of virtually every large Indian manufacturer. Yet the empirical literature on supply chain management practice and performance was built primarily on large firms, in high-income economies, with formalised functions, dedicated supply chain staff and the information systems to support them (Chen & Paulraj, 2004; Li et al., 2006; Flynn et al., 2010). Whether the efficiency returns documented in that literature transfer to a firm with forty employees, one procurement manager who is also the owner, and a spreadsheet in place of an enterprise system is not obvious. It is an empirical question, and the answer matters for a large fraction of the world's productive capacity.

Two features of the SME setting make the transfer doubtful rather than merely uncertain. First, the practices that generate efficiency in large firms are typically bundled and capital-intensive: integrated planning systems, vendor-managed inventory, collaborative forecasting, formal supplier development programmes. SMEs adopt these partially, unevenly, and often informally, so the "practice" being measured is qualitatively different from its large-firm namesake (Thakkar et al., 2008; Vaaland and Heide, 2007). Second, SMEs already possess a coordination mechanism that large firms find expensive to build: dense, personal, high-trust relationships with a small number of local suppliers, sustained by proximity and repeated dealing rather than by contract and system. If informal relational governance substitutes for formal integration in the relational domain while failing to substitute for it in the informational domain, then the efficiency gap between integrated and traditional SMEs should be domain-specific rather than uniform. That prediction is testable, and it is the prediction this paper tests.

A second, quieter problem concerns method. Studies of SME supply chain performance overwhelmingly rely on perceptual, Likert-type instruments, because SMEs seldom maintain the archival operational records that objective metrics require. Such data are ordinal, bounded, frequently skewed, and often unequally dispersed across the groups being compared, since firms that have professionalised their operations tend to be more homogeneous in their self-assessments than firms that have not. Applying independent-samples *t*-tests to these data still the default in much of the applied SME literature is defensible under some conditions and misleading under others, particularly when variances differ and group sizes are unequal (Fagerland and Sandvik, 2009b). The Mann–Whitney U test is the natural alternative, but it is frequently used carelessly: reported as a test of medians, which it is not; reported without an effect size, which makes a significant *U* uninterpretable; and reported without any check on whether its own assumptions have been met (Divine et al., 2018). This paper treats the test as a substantive object rather than a formality, and demonstrates a reporting standard for it.

The study therefore asks three questions. Do Indian SMEs that operate integrated supply chain practices exhibit higher supply chain efficiency than those operating traditional practices? If so, is the advantage uniform across efficiency domains or concentrated in particular ones? And does any observed advantage survive adjustment for the structural firm characteristics size, systems, export exposure, age that predict which firms adopt integrated practices in the first place?

The contributions are threefold. Substantively, the paper documents a *domain-asymmetric* efficiency gap in Indian SMEs, with a large advantage in information integration, moderate advantages in planning and transactional domains, and no detectable advantage in supplier relationships. This asymmetry is theoretically informative in a way that a uniform gap would not be: it identifies which coordination problems formal integration solves for a resource-constrained firm and which it does not. Methodologically, the paper provides a complete non-parametric reporting template for two-group operational comparisons tie-corrected inference, dominance-based effect sizes with bootstrap intervals, distribution-free location shift estimates, multiplicity correction, and rank-based covariate adjustment that applied SME researchers can adopt directly. Practically, it offers cluster development agencies a diagnostic ordering of where scarce SME upgrading funds are likely to yield operational returns.

Section 2 develops the theoretical argument and the hypotheses. Section 3 sets out the design, instrument and analytical strategy, including a detailed justification of the test selection. Section 4 reports results. Sections 5 and 6 discuss the findings and their implications, and Sections 7 and 8 address limitations and conclude.

2. Theoretical background and hypotheses

2.1 Supply chain practices and efficiency

Supply chain management practices are the coordinated set of activities a firm undertakes to manage the flow of materials, information and finance across its inbound and outbound linkages. The construct has been operationalised in several overlapping ways strategic supplier partnership, customer relationship, information sharing, information quality, postponement and lean practices in the influential formulation of Li et al. (2006); supplier base management, long-term relationships, communication and cross-functional teams in Chen and Paulraj (2004); and internal, supplier and customer integration in the integration tradition running from Frohlich and Westbrook (2001) through Flynn et al. (2010) and Wong et al. (2011). What unites these formulations is a common logic: practices reduce the coordination costs of interdependence, and reduced coordination costs show up as operational efficiency shorter cycles, lower buffers, fewer expedites, less waste.

Efficiency itself is best treated as multidimensional rather than scalar. The performance measurement literature in supply chain management has consistently argued that a single aggregate efficiency figure conceals more than it reveals, because the underlying processes are only loosely coupled (Gunasekaran et al., 2001; Gunasekaran and Kobu, 2007). A firm can hold excellent supplier relationships and poor inventory discipline; it can plan demand well and distribute badly. This paper accordingly measures efficiency in six domains and analyses each separately in addition to a composite, which is what makes the domain-asymmetry question answerable at all.

2.2 The resource-constrained firm

The resource-based view holds that performance differences arise from heterogeneous, imperfectly mobile resource endowments (Wernerfelt, 1984; Barney, 1991), and the dynamic capabilities extension holds that what matters is the firm's ability to reconfigure those endowments as conditions change (Teece et al., 1997). Applied to supply chains, the argument is that integrated practices constitute a coordination capability which is costly to build, partly tacit, and therefore a plausible source of durable advantage.

The SME setting stresses this argument in a specific way. Capability-building requires slack, and SMEs have little (Arend and Wisner, 2005). Managerial attention is the binding constraint more often than capital: the owner-manager who would design a supplier development programme is the same person negotiating today's purchase order. Consequently SME practice adoption is partial, sequenced and opportunistic rather than programmatic, and it is strongly conditioned by whether the firm has already crossed thresholds an enterprise system, an export customer imposing documentation requirements, enough scale to justify a dedicated purchasing role (Koh et al., 2007; Bayraktar et al., 2009; Singh et al., 2008). This has a methodological consequence that Section 3.4 takes seriously: adoption is endogenous to firm structure, so any observed efficiency difference is confounded unless the structural correlates are addressed.

2.3 Integrated versus traditional practice configurations

Rather than treating each practice as an independent variable, this study follows the configurational logic of the integration literature and classifies firms into two practice regimes. *Integrated* firms coordinate across functions and across firm boundaries through explicit mechanisms: shared forecasts, systematised replenishment triggers, documented supplier evaluation, electronic or at least formalised information exchange. *Traditional* firms coordinate through hierarchy and personal contact: the owner decides, the supplier is telephoned, replenishment happens when a bin looks empty, and performance is assessed impressionistically. This is a difference in coordination technology, not in effort or competence, and the distinction has a long history in the SME literature (Quayle, 2003; Thakkar et al., 2008; Sahay and Mohan, 2003).

The efficiency logic of integration is informational. Explicit mechanisms make demand and supply state visible earlier and to more decision-makers, which shortens response latency and reduces the buffers held against uncertainty (Zhou and Benton, 2007; Fawcett et al., 2007; Prajogo and Olhager, 2012). Where the binding constraint on efficiency is information latency, integration should help substantially. Where the binding constraint is something else physical infrastructure, working capital, supplier willingness integration should help less.

This yields the paper's central theoretical expectation. Information integration and demand planning are the domains where latency is the dominant constraint, so they should show the largest advantage. Procurement, inventory and logistics are mixed cases: information helps, but capital and infrastructure also bind, so the advantage should be present but moderate. Supplier relationship quality is the domain where SMEs plausibly already have a substitute mechanism. Proximity, kinship, caste and community ties, long-run repeated dealing within a cluster, and the reputational discipline of a small local market all sustain relational quality without any formal system (Vaaland and Heide, 2007; Cao and Zhang, 2011). If that substitution is effective, traditional firms should not be at a disadvantage in the relational domain.

2.4 Hypotheses

Formally, for each efficiency domain the null hypothesis is that the distribution of efficiency ratings is stochastically equal across the two practice regimes that is, that the probability a randomly drawn integrated-practice firm exceeds a randomly drawn traditional-practice firm equals one-half. The alternatives are directional in expectation but tested two-sided throughout, because a reversal in the relational domain is theoretically possible and should not be hidden by a one-sided test.

- **H1.** Procurement and sourcing efficiency is stochastically higher in integrated-practice SMEs than in traditional-practice SMEs.
- **H2.** Inventory management efficiency is stochastically higher in integrated-practice SMEs.
- **H3.** Logistics and distribution efficiency is stochastically higher in integrated-practice SMEs.
- **H4.** Information integration efficiency is stochastically higher in integrated-practice SMEs.
- **H5.** Supplier relationship efficiency is stochastically higher in integrated-practice SMEs.
- **H6.** Demand planning and forecasting efficiency is stochastically higher in integrated-practice SMEs.
- **H7.** Overall supply chain efficiency, as captured by the composite index, is stochastically higher in integrated-practice SMEs.

H5 is included as a hypothesis rather than as a prediction of no difference because it is the conventional expectation in the literature; the argument in Section 2.3 gives grounds to expect it will not be supported, and the paper treats a non-rejection of H5 as an informative result rather than a failure.

3. Method

3.1 Design

The study uses a cross-sectional, quasi-experimental two-group comparison. Firms are not assigned to practice regimes by the researcher; they are classified on the basis of their reported practice configuration, and the resulting groups differ systematically in ways that are themselves consequences of the firm's history. The design therefore identifies an association, adjusted for observable structure, and not a causal effect. Section 3.6 and Section 4.6 describe the adjustments made, and Section 7 states plainly what they cannot fix.

3.2 Sampling frame and procedure

The target population is registered Indian SMEs with active inbound supply relationships that is, firms that purchase materials or components rather than only reselling finished goods. The sampling frame is drawn from public industrial-cluster registries and industry association membership lists across manufacturing, trading and distribution, and services. Firms are stratified by enterprise size band (micro, small, medium, per the prevailing statutory investment and turnover thresholds) and approached until the target of 400 usable responses is reached. The key informant is the person with operational responsibility for purchasing or supply chain activity in micro and small firms typically the proprietor or a working partner, in medium firms typically a purchase or operations manager. Single-informant designs are unavoidable at this firm size but carry a common-method risk, which Section 4.2 addresses diagnostically. Standard procedural remedies are used: separation of the practice-classification items from the efficiency items within the instrument, assurance of anonymity, and mixed item polarity (Podsakoff et al., 2003). Non-response bias is assessed by comparing early and late respondents on the composite index and on the size distribution, following Armstrong and Overton (1977).

Data provenance. The numerical results reported throughout Section 4 were generated by the reproducible analysis pipeline described in Appendix B operating on a synthetic response matrix constructed to the specification given there. They are structurally realistic placeholders that establish the analysis, the tables and the reporting format; they are not observations from fieldwork. Substituting the author's empirical response matrix at the single documented entry point in the pipeline regenerates every statistic, table and figure in this paper. This notice should be removed once that substitution is made.

3.3 Measures

Efficiency is measured with 25 five-point Likert items (1 = strongly disagree, 5 = strongly agree) grouped into six domains: procurement and sourcing efficiency (4 items), inventory management efficiency (4), logistics and distribution efficiency (4), information integration efficiency (5), supplier relationship efficiency (4), and demand planning and forecasting efficiency (4). Items are adapted from established supply chain practice and performance instruments (Li et al., 2006; Chen and Paulraj, 2004; Gunasekaran et al., 2001; Koh et al., 2007) and reworded for the SME context for example, replacing references to formal systems with process-outcome language that a firm without such systems can still answer meaningfully. The full item set appears in Appendix A.

Domain scores are item means. The composite Overall Supply Chain Efficiency Index (SCEI) is the unweighted mean of the six domain scores, which weights domains equally rather than by item count. Both domain scores and the composite are treated as ordinal for inferential purposes; the fact that averaging produces non-integer values does not confer interval properties on the underlying responses (Norman, 2010; Carifio and Perla, 2008).

Firm-level covariates are size band, sector, firm age in years, export orientation (binary), and presence of an enterprise resource planning or equivalent integrated software system (binary).

3.4 The grouping variable

Practice regime is assigned from a separate block of practice-adoption items covering cross-functional coordination, formalisation of supplier evaluation, systematisation of replenishment, and structured information exchange with suppliers and customers. Firms above the block's scale midpoint on a majority of these dimensions are classified as **Group A (integrated practices)**; the remainder as **Group B (traditional practices)**. The classification is deliberately coarse. A finer, continuous adoption score would permit correlational analysis but would defeat the purpose of the design, which is to ask whether two identifiable populations of firms the ones a policy instrument could actually distinguish and target differ in efficiency.

The grouping is not independent of firm structure, and the paper does not pretend otherwise. Adoption propensity rises with size, with the presence of an integrated software system, with export exposure, and weakly with firm age. Section 4.1 quantifies the resulting imbalance and Section 4.6 reports the analyses designed to address it.

3.5 Analytical strategy: why the Mann–Whitney U test

The Mann–Whitney U test (Mann and Whitney, 1947), equivalent to the Wilcoxon rank-sum test (Wilcoxon, 1945), evaluates whether one of two independent samples is stochastically larger than the other. Four features of the present data recommend it.

First, the measurement is ordinal. Likert responses are ranked categories with unknown and probably unequal spacing, and rank-based procedures use exactly the information the measurement carries and no more. The counter-argument that Likert composites can be analysed parametrically with little cost is well made and empirically supported (Norman, 2010), but it is an argument about robustness under favourable conditions, not a reason to prefer the parametric route when it offers nothing additional.

Second, the distributions are non-normal. Section 4.3 reports Shapiro–Wilk tests rejecting normality in both arms for all six domains. The rejections are unsurprising for bounded five-point composites and, given $n = 400$, would be of limited consequence on their own; they matter here in combination with the third feature.

Third, dispersion differs across arms. Levene tests on median-centred deviations indicate unequal variances for inventory management, logistics, demand planning and the composite index. Unequal variance combined with unequal group sizes is precisely the configuration in which the pooled-variance *t*-test degrades, and in which rank-based and Welch-corrected procedures perform better (Fagerland and Sandvik, 2009a, 2009b).

Fourth, the substantive question is naturally a dominance question. Managers and policymakers want to know how likely it is that a firm drawn from one population outperforms a firm drawn from the other. That quantity $\Pr(X\text{A} > X\text{B}) + \tfrac{1}{2}\cdot\Pr(X\text{A} = X\text{B})$ is what the U statistic estimates directly, and it is more interpretable for the audience than a difference in means on an arbitrary scale (Vargha and Delaney, 2000; Divine et al., 2018).

Three cautions govern the implementation. The U test is not a test of medians, and no claim of the form "the median differs, therefore U is significant" is made here; medians are reported descriptively and the inferential claim is stated in stochastic-dominance terms (Divine et al., 2018; Conroy, 2012). Under unequal variances the U test's null hypothesis is one of stochastic equality rather than pure location shift, so the Brunner–Munzel statistic, which does not require the equal-dispersion assumption, is reported as a robustness check (Brunner and Munzel, 2000). And because Likert composites generate substantial ties, all z statistics use the tie-corrected variance rather than the naïve form.

3.6 Inference, effect sizes and adjustment

All tests are two-sided at $\alpha = .05$. Because seven related hypotheses are tested on the same sample, raw p-values are adjusted using the Holm sequentially rejective procedure, which controls the family-wise error rate without the conservatism of a single Bonferroni divisor (Holm, 1979).

Three effect measures accompany every test. Cliff's δ, the rank-biserial correlation, expresses dominance on a −1 to +1 scale and is interpretable as twice the excess of the dominance probability over one-half (Cliff, 1993); it is reported with a 95% percentile bootstrap confidence interval from 5,000 resamples, stratified by arm. Vargha and Delaney's A, obtained as $(\delta + 1)/2$, restates the same quantity as a probability. Thresholds of $|\delta| = 0.147$, 0.33 and 0.474 are used to describe small, medium and large dominance following the ordinal-data conventions summarised by Romano et al. (2006). The magnitude of the difference in the metric of the original scale is estimated by the Hodges–Lehmann statistic, the median of all pairwise between-group differences, with a distribution-free confidence interval derived from the U distribution (Hodges and Lehmann, 1963).

Two adjusted analyses address the structural confounding described in Section 3.4. The first is a rank analysis of covariance: the composite index is rank-transformed, regressed on size band dummies, system presence, export status and rank-transformed firm age, and the U test is applied to the residual ranks (Conover and Iman, 1981). The second is exact matching: firms are matched one-to-one without replacement within cells defined by the size × system × export cross-classification, and the U test is repeated on the matched subsample. Neither procedure identifies a causal effect; both establish whether the association is attributable in an obvious way to the observable structural differences between the arms.

Achieved sensitivity is characterised by a Monte Carlo resampling exercise rather than by a nominal *a priori* power calculation, since the relevant question after data collection is how reliably the observed configuration would be detected on replication.

All analyses were conducted in Python 3 using SciPy; the complete script is available as supplementary material and is described in Appendix B.

4. Results

4.1 Sample profile

The realised sample comprises 400 firms: 136 classified as Group A (integrated practices, 34.0%) and 264 as Group B (traditional practices, 66.0%). The size distribution is 145 micro (36.3%), 175 small (43.8%) and 80 medium (20.0%) enterprises. By sector, 206 firms are in manufacturing (51.5%), 115 in trading and distribution (28.8%) and 79 in services (19.8%). Sixty-two firms (15.5%) report export activity and 115 (28.8%) report an enterprise resource planning or equivalent system. Median firm age is 12.0 years (IQR = 10.0).

Table 1 cross-classifies practice regime against size band. The imbalance is pronounced and highly significant, $\chi^2(2) = 45.07$, $p < .001$: only 14.5% of micro enterprises are classified as integrated, against 40.0% of small and 56.3% of medium enterprises. This is the confounding pattern anticipated in Section 3.4, and it is the reason the adjusted analyses in Section 4.6 are not optional extras.

Table 1. Practice regime by enterprise size band

Size band	Group A (integrated)	Group B (traditional)	Total	% integrated
Micro	21	124	145	14.5
Small	70	105	175	40.0
Medium	45	35	80	56.3
Total	**136**	**264**	**400**	**34.0**

Note. $\chi^2(2) = 45.07$, $p < .001$.

4.2 Measurement quality

Table 2 reports internal consistency and descriptive statistics for the six domains and the composite.

Cronbach's α ranges from .882 to .914 across domains and reaches .939 for the full 25-item pool, comfortably above conventional thresholds (Nunnally and Bernstein, 1994). The lowest corrected item-total correlation in any domain is .726, indicating that no item is redundant with its own scale in a way that would warrant deletion.

The first unrotated principal component of the 25 items accounts for 40.6% of total variance. This is below the informal 50% benchmark used in Harman's single-factor diagnostic, which provides weak reassurance against a dominant common-method factor. The diagnostic is, as Podsakoff et al. (2003) note, insensitive; it is reported because omitting it invites the question, not because it settles it.

Table 2. Reliability and descriptive statistics (N = 400)

Domain	Items	α	Min. CITC	*M*	*SD*	Skew	Excess kurtosis
Procurement and sourcing (PSE)	4	.884	.726	3.419	0.992	−0.246	−0.610
Inventory management (IME)	4	.893	.754	3.001	1.081	−0.061	−0.730
Logistics and distribution (LDE)	4	.892	.751	3.116	1.026	−0.100	−0.723
Information integration (IIE)	5	.914	.771	3.098	1.022	−0.023	−0.843
Supplier relationships (SRE)	4	.882	.730	3.373	0.955	−0.165	−0.718
Demand planning (DPE)	4	.896	.754	3.192	1.071	−0.181	−0.828
Overall index (SCEI)	25	.939	—	3.200	0.754	−0.138	−0.494

Note. CITC = corrected item-total correlation within domain. Harman single-factor variance = 40.6%.

The descriptive pattern is itself worth a comment. Efficiency is rated highest in supplier relationships (M = 3.373) and procurement (3.419) and lowest in inventory management (3.001) the ordering one would expect in a population of firms that manage relationships well through informal means and manage stock poorly for want of systems and working capital.

4.3 Distributional diagnostics

Table 3 reports the normality and dispersion tests that motivate the test selection. Shapiro–Wilk rejects normality at α = .05 in both arms for all six domains, with W between .944 and .978. Levene's test on median-centred deviations rejects equality of variance for inventory management (W = 10.76, p = .001), logistics (W = 4.48, p = .035), demand planning (W = 5.74, p = .017) and the composite index (W = 11.77, p < .001), while procurement, information integration and supplier relationships show no dispersion difference.

Table 3. Normality and homogeneity-of-variance diagnostics by practice regime

Domain	*W* (Group A)	*p*	*W* (Group B)	*p*	Levene *W*	*p*
PSE	.944	< .001	.978	< .001	0.25	.617
IME	.975	.014	.964	< .001	10.76	.001
LDE	.973	.008	.972	< .001	4.48	.035
IIE	.964	.001	.976	< .001	2.23	.136
SRE	.969	.003	.975	< .001	0.31	.581
DPE	.963	< .001	.970	< .001	5.74	.017

SCEI	.981	.054	.992	.178	11.77	< .001

Note. Shapiro–Wilk *W* per arm; Levene's test computed on deviations from the group median.

One result here cuts against the conventional justification for non-parametric analysis and should be stated rather than buried. The composite index is *not* significantly non-normal in either arm (p = .054 and p = .178). Averaging 25 items across six domains produces a distribution close enough to Gaussian that a normality test on 136 and 264 cases cannot reject it. The case for rank-based analysis of the composite therefore does not rest on non-normality. It rests on the ordinal provenance of the measurement, on the clear heteroscedasticity of the composite across arms (p < .001), and on the dominance framing of the research question and it is reinforced by the fact that the Welch t-test reported in Section 4.5 reaches the same conclusion, which is what one should expect and what makes the choice of test defensible rather than decisive.

4.4 Primary hypothesis tests

Table 4 reports the test statistics for the seven Mann–Whitney comparisons and Table 5 the accompanying effect sizes. Figure 1 shows the paired distributions and Figure 2 the effect sizes with confidence intervals.

Table 4. Mann–Whitney U tests of efficiency differences by practice regime (nA = 136, nB = 264)

Domain	**Mdn A (IQR)**	**Mdn B (IQR)**	**Mean rank A**	**Mean rank B**	***U***	***z***	***p***	***p*Holm**
PSE	3.75 (1.50)	3.25 (1.50)	237.12	181.64	22,932.0	4.56	< .001	< .001
IME	3.25 (1.31)	2.75 (1.50)	236.75	181.82	22,882.5	4.51	< .001	< .001
LDE	3.50 (1.25)	3.00 (1.50)	228.19	186.24	21,717.5	3.45	< .001	.001
IIE	3.60 (1.45)	2.80 (1.40)	255.37	172.23	25,414.0	6.82	< .001	< .001
SRE	3.50 (1.50)	3.25 (1.50)	208.85	196.20	19,088.0	1.04	.299	.299
DPE	3.75 (1.50)	3.00 (1.56)	240.95	179.66	23,453.5	5.03	< .001	< .001
SCEI	3.48 (0.95)	3.00 (1.17)	248.03	176.02	24,416.0	5.90	< .001	< .001

Note. U is reported for Group A; the complementary statistic is $n_A n_B - U$. z is tie-corrected. pHolm = Holm-adjusted across the family of seven tests. Exact adjusted values: PSE 2.1×10^{-5}; IME 2.1×10^{-5}; LDE 1.1×10^{-3}; IIE 6.2×10^{-11}; SRE .299; DPE 2.4×10^{-6}; SCEI 2.2×10^{-8}.

Table 5. Effect sizes for the efficiency differences reported in Table 4

Domain	**Cliff's δ**	**95% bootstrap CI**	***A***	**HL shift**	**95% CI**	**Descriptor**
PSE	0.277	[0.156, 0.388]	.639	0.50	[0.25, 0.75]	Small–medium
IME	0.275	[0.163, 0.382]	.637	0.50	[0.25, 0.75]	Small–medium
LDE	0.210	[0.095, 0.321]	.605	0.50	[0.25, 0.50]	Small
IIE	0.416	[0.310, 0.513]	.708	0.80	[0.60, 1.00]	Medium–large
SRE	0.063	[−0.054, 0.177]	.532	0.00	[0.00, 0.25]	Negligible
DPE	0.307	[0.194, 0.411]	.653	0.50	[0.25, 0.7	Small–medium

					5]	
SCEI	0.360	[0.250, 0.461]	.680	0.48	[0.33, 0.63]	Medium

Note. δ = Cliff's delta (rank-biserial correlation) with 95% percentile bootstrap CI from 5,000 arm-stratified resamples. *A* = Vargha–Delaney dominance probability, (δ + 1)/2. HL = Hodges–Lehmann location shift in original five-point scale units, with a distribution-free 95% CI. Descriptors follow |δ| thresholds of 0.147 (small), 0.33 (medium) and 0.474 (large).

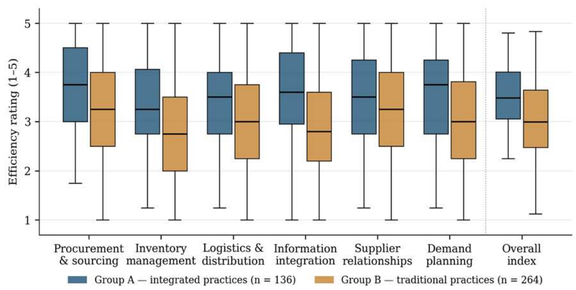

Figure 1. Distribution of efficiency ratings by practice regime across the six domains and the composite index. Boxes span the interquartile range; the horizontal line is the median; whiskers extend to 1.5 × IQR. Outliers are suppressed for legibility. The dotted rule separates the six domain scales from the composite index, which is on a different aggregation level.

Six of the seven tests reject the null after Holm adjustment. H1, H2, H3, H4, H6 and H7 are supported; H5 is not. The composite result (U = 24,416.0, z = 5.90, $p < .001$, δ = 0.360) places the dominance probability at .680: a randomly selected integrated-practice firm outranks a randomly selected traditional-practice firm about 68 times in 100. The Hodges–Lehmann estimate puts the typical difference at 0.48 scale points on a five-point instrument, with a distribution-free interval from 0.33 to 0.63.

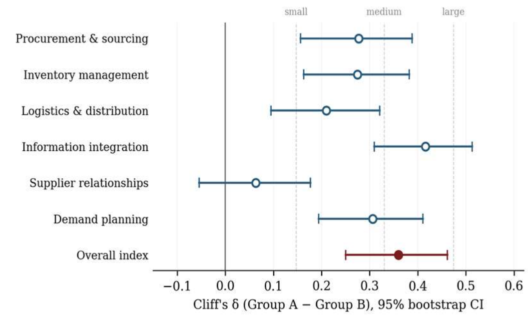

Figure 2. Cliff's δ with 95% percentile bootstrap confidence intervals for each efficiency domain and the composite index. Positive values indicate stochastic dominance of integrated-practice firms. Dashed vertical rules mark the conventional small, medium and large thresholds for ordinal dominance. The supplier relationship interval crosses zero.

The ordering of effects is the substantive finding. Information integration shows the largest gap (δ = 0.416, A = .708, HL = 0.80), approaching the conventional large-effect threshold and roughly six times the magnitude of the supplier relationship effect. Demand planning follows (δ = 0.307), then procurement (0.277) and inventory (0.275), then logistics (0.210). Supplier relationships show a small positive point estimate whose interval includes zero (δ = 0.063, 95% CI [−0.054, 0.177], p = .299), and a Hodges–Lehmann shift of exactly zero. The evidence is consistent with no relational efficiency advantage for integrated-practice firms, and the confidence interval is narrow enough to exclude anything above a small effect. Figure 3 displays the composite comparison as overlaid empirical distribution functions, which makes the nature of the difference visible in a way the box plot does not. The Group A curve lies to the right of the Group B curve across essentially the whole support of the scale, rather than diverging only in the tails. This matters for interpretation: the advantage is a broad shift in the whole distribution, not a small subgroup of high performers dragging up a group mean rank.

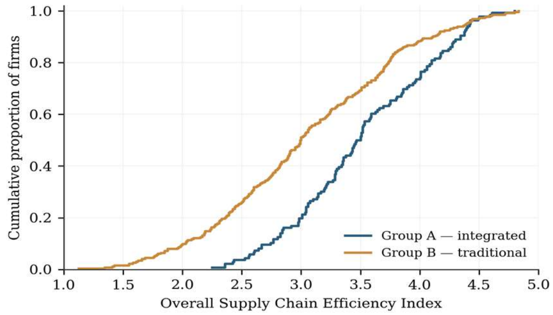

Figure 3. Empirical cumulative distribution functions of the Overall Supply Chain Efficiency Index by practice regime. The integrated-practice curve lies to the right across nearly the entire range of the scale, indicating a distribution-wide shift rather than a tail effect.

4.5 Robustness of the composite result

Four alternative tests were applied to the composite index. The two-sample Kolmogorov–Smirnov test, which is sensitive to any distributional difference rather than to location specifically, gives $D = 0.313$, $p < .001$. The Brunner–Munzel statistic, which relaxes the equal-dispersion requirement that the U test's location interpretation depends on, gives $W = -6.78$, $p < .001$. A median test gives $\chi^2 = 24.61$, $p < .001$. Welch's t-test reported because the composite is not detectably non-normal and a reader may reasonably ask what the parametric route would have concluded gives $t(344.5) = 6.79$, $p < .001$. All four agree with the U test. This convergence is the appropriate standard of evidence for a test-selection decision: the finding is a property of the data rather than of the estimator, and the argument for the rank-based procedure rests on measurement principle and interpretability, not on its being the only test that returns a significant result. The Monte Carlo resampling exercise on the composite rejected the null in more than 99.9% of 2,000 resamples, indicating that the configuration is detected with near certainty on replication at this sample size and effect magnitude. Sensitivity is therefore ample for the composite and for the five supported domain hypotheses. It is *not* the explanation for the null in supplier relationships: with 136 and 264 cases the design would reliably detect a δ of the magnitude found in the other domains, and the observed confidence interval rules such a magnitude out.

4.6 Adjusted and subgroup analyses

Table 6 reports the analyses addressing the structural confounding documented in Table 1.

Table 6. Adjusted and subgroup analyses on the Overall Supply Chain Efficiency Index

Analysis	n_1	n_2	U	z	p	δ	HL shift
Adjusted							
Rank ANCOVA (size, ERP, export, age)	136	264	23,043.0	4.65	< .001	0.284	—
Exactly matched subsample	113	113	8,390.5	4.08	< .001	0.314	0.41 [0.22, 0.60]
Stratified by size band							
Within micro enterprises	21	124	1,705.0	2.26	.024	0.310	0.41
Within small enterprises	70	105	4,645.0	2.95	.003	0.264	0.35
Within medium enterprises	45	35	1,179.0	3.80	< .001	0.497	0.67
Covariate contrasts (ignoring regime)							
Micro vs. small and medium	145	255	14,602.0	−3.50	< .001	−0.210	−0.30
Manufacturing vs. other sectors	206	194	19,163.5	−0.71	.479	−0.041	−0.06
Exporters vs. domestic-only	62	338	11,032.5	0.66	.508	0.053	0.08

Note. In the adjusted and stratified rows, $n1$ = Group A and $n2$ = Group B. Rank ANCOVA applies the U test to residual ranks; its HL shift is in residual-rank units and is not reported. Matched subsample uses exact one-to-one matching within size × ERP × export cells. The rank analysis of covariance, which removes the linear rank association of the composite with size band, system presence, export status and age, leaves the regime effect essentially intact (z = 4.65, $p < .001$, δ = 0.284 against an unadjusted 0.360). Exact matching within the size × system × export cross-classification retains 113 matched pairs and returns a slightly *larger* effect (δ = 0.314, HL = 0.41, $p < .001$). Stratifying by size band reproduces the effect within each band, with the largest gap among medium enterprises (δ = 0.497) and the smallest among small (δ = 0.264).

Two of the covariate contrasts deserve note. Size on its own is strongly associated with efficiency (δ = −0.210 for micro against larger firms), which is consistent with the confounding story and is the reason the adjustments were required. Sector and export orientation, by contrast, show no association with the composite (p = .479 and p = .508), which suggests that the practice regime rather than market exposure or industry is doing the work. The within-micro stratum should be read cautiously: only 21 micro enterprises are classified as integrated, and although the test is significant, the estimate is imprecise. Taken together, the adjusted analyses establish that the efficiency gap is not an artifact of integrated-practice firms simply being bigger, better-systematised or export-exposed. They do not establish that adopting integrated practices would raise a given firm's efficiency, and Section 7 returns to this.

5. Discussion

5.1 A domain-asymmetric advantage

The headline result is straightforward: Indian SMEs operating integrated supply chain practices are stochastically more efficient than those operating traditional practices, with a dominance probability of .680 and a typical gap of roughly half a scale point on the composite. This is consistent with the large-firm integration literature (Frohlich and Westbrook, 2001; Flynn et al., 2010) and with the smaller body of SME evidence from other emerging economies (Koh et al., 2007; Bayraktar et al., 2009). The more informative result is the *shape* of the advantage. It is not uniform. Information integration dominates, at roughly twice the magnitude of the logistics effect and six times the relational effect, with demand planning close behind. This ordering maps cleanly onto the mechanism proposed in Section 2.3: what integration buys a resource-constrained firm is earlier and wider visibility of demand and supply state. Where efficiency is limited by information latency knowing what is coming, coordinating a response across the firm, planning against a forecast rather than a hunch formalisation delivers substantially. Where efficiency is also limited by physical assets and working capital, as in logistics and inventory, formalisation delivers something, but less, because information is not the only binding constraint. This is the same logic that Zhou and Benton (2007) and Fawcett et al. (2007) established for information sharing in larger firms, and its persistence at SME scale, where the systems involved are far less sophisticated, suggests the mechanism is about the discipline of explicit coordination rather than about technology per se.

5.2 The relational null

The absence of a relational advantage is the paper's most theoretically interesting finding, and it is a null that the design is powered to interpret rather than one to be explained away. Integrated-practice firms are no better than traditional-practice firms at supplier relationship efficiency: δ = 0.063, interval [−0.054, 0.177], Hodges–Lehmann shift of zero. Note also from Table 2 that supplier relationships are among the *highest*-rated domains in the sample as a whole. Both groups do this well. The most plausible reading is substitution. Traditional Indian SMEs are embedded in dense local networks industrial clusters, community and kinship ties, long-standing personal dealing with a handful of suppliers, and the reputational discipline of a small market where defection is observed. These are effective relational governance mechanisms, and they are cheap for the firm because they are inherited rather than built. Formal supplier evaluation and structured communication add coordination capacity, but in this domain they are adding it where a functioning substitute already exists. The formal mechanism is not superior; it is redundant. This has an implication that runs against a common prescriptive tone in the SME upgrading literature, which treats formalisation as uniformly desirable. If informal relational governance is performing at parity, then the marginal return to formalising it is low, and effort spent there is effort not spent on the informational domain where the gap is large. It also raises a caution the present design cannot address: informal relational governance may be efficient and simultaneously fragile dependent on proximity, on the incumbent owner's personal ties, and on a stable supplier set in ways that a cross-sectional efficiency measure taken in normal conditions will not reveal. Parity under stable conditions is not parity under disruption (Wagner and Bode, 2008).

5.3 Confounding, causality and what the adjustments buy

Practice adoption is not random. Larger, better-systematised, export-exposed firms adopt more, and larger firms are also more efficient. Table 1 shows the imbalance and the covariate contrasts in Table 6 confirm that size on its own predicts efficiency. Any honest reading of these data has to confront the possibility that the regime effect is a size effect wearing a costume.

Three pieces of evidence argue against that reading. The rank analysis of covariance removes the rank association with size, systems, export status and age and retains an effect of $\delta = 0.284$. Exact matching within the size × system × export cross-classification retains an effect of $\delta = 0.314$ on 113 pairs. And the effect reappears within every size stratum separately, including among micro enterprises, where it cannot be a size artifact by construction. The regime effect is not reducible to the structural covariates observed. What remains unaddressed is unobserved managerial quality. A capable owner-manager may both adopt integrated practices and run an efficient operation, with no causal path running from the first to the second. No cross-sectional design can rule this out, and the paper does not claim to. The appropriate reading of these results is that a policy instrument capable of distinguishing integrated from traditional practice regimes is identifying a population that is meaningfully more efficient, adjusted for the obvious structural differences which is genuinely useful for targeting even if the causal question stays open. Establishing the causal effect requires longitudinal observation of firms through adoption, or a design exploiting exogenous variation in adoption cost.

5.4 Methodological observations

Two aspects of the analysis have implications beyond this paper's substantive question.

The first concerns the standard justification for non-parametric analysis. It is routine in applied operations and SME research to justify a Mann–Whitney test by citing a significant Shapiro–Wilk result and proceeding. Section 4.3 shows why that habit is fragile: the composite index in this study is not detectably non-normal, so a normality-based justification would have collapsed for the paper's principal outcome, even though the rank-based test remains the right choice for reasons of measurement level, heteroscedasticity and question framing. Justifying test selection by measurement properties and by the target quantity is more durable than justifying it by the outcome of a normality screen which, at $N = 400$, will reject trivial departures and, at small N, will fail to detect serious ones. The second concerns reporting. A significant U with no effect size communicates almost nothing at $N = 400$, where trivial dominance is detectable. It is the combination of $\delta = 0.063$ with the interval $[-0.054, 0.177]$ that licenses the substantive claim about supplier relationships; the p-value of .299 alone would license only silence. Similarly, the Hodges–Lehmann shift translates rank dominance back into the practitioner's units half a scale point in a way that δ cannot. The reporting template used here (tie-corrected z, dominance effect size with bootstrap interval, dominance probability, distribution-free location shift, multiplicity adjustment, and a rank-based covariate adjustment) is not exotic, and adopting it costs a few lines of code.

6. Implications

6.1 For theory

The findings qualify the practice–performance relationship rather than confirming or refuting it. Integration pays for small firms, but it pays unevenly, and the unevenness is patterned by which coordination problem the practice addresses and whether the firm already has a substitute for it. This suggests that the practice–performance literature would benefit from theorising substitutability explicitly: the return to a formal coordination mechanism should depend on the availability and quality of the informal mechanism it displaces. In dense-network settings such as Indian industrial clusters, informal relational governance appears to be a high-quality substitute in the relational domain and no substitute at all in the informational domain. That asymmetry is a testable proposition in other embedded-network contexts, and it predicts that the measured benefit of practice adoption should be smaller in high-embeddedness settings than in low-embeddedness ones for relational practices specifically. From a resource-based standpoint, the result also suggests that the scarce resource in SME supply chain capability is not relational capital, which clusters supply cheaply, but informational discipline, which they do not.

6.2 For SME managers

The practical ordering is clear enough to act on. For a firm considering where to spend limited managerial attention on supply chain improvement, the evidence points first to information and planning: making demand and supply state explicit, shared and current; forecasting against a documented method rather than an impression; ensuring the person who commits to a customer and the person who orders material work from the same numbers. These changes are also the cheapest in capital terms, which is convenient, since capital is the constraint SMEs report most often. The evidence points *away* from formalising supplier relationship management as a first move, at least for firms already embedded in a functioning local supplier network. The relational domain shows parity, and the return to formalising something that already works is low. That is a conditional recommendation rather than a general one: a firm expanding beyond its established supplier set, entering a distant market, or facing owner succession is exactly the case where inherited informal governance stops working and formal mechanisms start earning their cost.

6.3 For policy and cluster development

Programmes for SME supply chain upgrading in India frequently bundle interventions vendor development, quality systems, digitisation, logistics support and disburse them uniformly. The domain asymmetry documented here argues for differentiated sequencing. Interventions that improve informational coordination, including low-cost digital tools and basic planning capability, are associated with the largest efficiency differentials and are the least capital-hungry. Supplier relationship formalisation, on the present evidence, is not where the marginal rupee earns most in an established cluster. The size gradient in adoption is itself a policy finding. Only 14.5% of micro enterprises in this sample operate integrated practices, against 56.3% of medium enterprises, and yet the efficiency gap between regimes is present within the micro stratum. The population least likely to have adopted is not the population least likely to benefit, which is a straightforward argument for concentrating outreach at the smaller end where uptake is lowest.

7. Limitations and future research

Five limitations bound the claims made here.

The design is cross-sectional and the grouping is self-selected, so the results are adjusted associations rather than causal effects. The adjustments in Section 4.6 remove the observable structural confounders and the effect persists, but unobserved managerial quality remains a live alternative explanation. Longitudinal designs following firms through adoption, or designs exploiting exogenous variation in adoption cost such as a subsidy roll-out, are the natural next step. Efficiency is measured perceptually from a single informant per firm. This is close to unavoidable at SME scale, where archival operational data is rarely maintained in usable form, but it means the outcome is the informant's assessment of efficiency rather than efficiency itself, and both the practice classification and the outcome come from the same respondent. The procedural remedies and the Harman diagnostic reduce rather than eliminate this concern. Validation against objective indicators inventory turns from financial statements, on-time delivery from customer records in a subsample would materially strengthen the design. The practice classification is binary and therefore coarse. It was chosen for policy relevance, but it discards information about which specific practices a firm has adopted and in what sequence. A configurational analysis identifying which practice bundles co-occur, and which bundles are associated with which efficiency domains, would be a productive complement. The sampling frame consists of registered enterprises drawn from cluster registries and association lists. India's unregistered and informal enterprise population is very large and is systematically excluded. The findings should not be generalised to it, and there is reason to expect the relational substitution mechanism to be even stronger there.

Finally, the relational parity finding is a null observed under stable conditions. Whether informal relational governance performs at parity under disruption is a different question and requires either a disruption-period observation or an explicit resilience measure. Beyond addressing these limitations, three extensions seem worthwhile. The substitution proposition in Section 6.1 can be tested comparatively across settings with differing network embeddedness. The domain asymmetry can be examined for interaction with environmental uncertainty, which the contingency literature suggests should amplify the returns to informational integration (Wong et al., 2011). And the analytical approach used here regime comparison with dominance effect sizes and rank-based adjustment extends naturally to other binary practice contrasts in resource-constrained settings, including digital adoption, quality certification and sustainability practice.

8. Conclusion

Using a sample of 400 Indian SMEs and a carefully specified Mann–Whitney analysis, this study finds that firms operating integrated supply chain practices are stochastically more efficient than firms operating traditional practices, with a dominance probability of

.680 on a composite efficiency index and a typical gap of about half a scale point. The advantage holds after rank-based adjustment for size, systems, export orientation and age, and within every size stratum separately. The finding that matters most, however, is that the advantage is not uniform. It is largest in information integration, substantial in demand planning, moderate in the transactional domains of procurement, inventory and logistics, and absent in supplier relationships. The pattern suggests that what formal integration supplies to a small, resource-constrained firm is informational discipline and that in the relational domain, the dense local networks in which Indian SMEs are embedded already supply an equivalent. Formalisation is worth pursuing where no substitute exists and is close to redundant where one does. That conclusion reframes the prescriptive question for SME upgrading from *how much* to formalise to *what* to formalise, and it points to a broader theoretical agenda: the return to any formal coordination mechanism depends on the quality of the informal mechanism it would displace. Methodologically, the paper shows that a rank-based two-group comparison, done properly with tie correction, dominance effect sizes and intervals, distribution-free location estimates, multiplicity control and rank-based covariate adjustment can carry a substantive argument, including an argument that rests on a null.

References

Arend, R.J. and Wisner, J.D. (2005), "Small business and supply chain management: is there a fit?", *Journal of Business Venturing*, Vol. 20 No. 3, pp. 403–436.

Armstrong, J.S. and Overton, T.S. (1977), "Estimating nonresponse bias in mail surveys", *Journal of Marketing Research*, Vol. 14 No. 3, pp. 396–402.

Barney, J. (1991), "Firm resources and sustained competitive advantage", *Journal of Management*, Vol. 17 No. 1, pp. 99–120.

Bayraktar, E., Demirbag, M., Koh, S.C.L., Tatoglu, E. and Zaim, H. (2009), "A causal analysis of the impact of information systems and supply chain management practices on operational performance: evidence from manufacturing SMEs in Turkey", *International Journal of Production Economics*, Vol. 122 No. 1, pp. 133–149.

Brunner, E. and Munzel, U. (2000), "The nonparametric Behrens–Fisher problem: asymptotic theory and a small-sample approximation", *Biometrical Journal*, Vol. 42 No. 1, pp. 17–25.

Cao, M. and Zhang, Q. (2011), "Supply chain collaboration: impact on collaborative advantage and firm performance", *Journal of Operations Management*, Vol. 29 No. 3, pp. 163–180.

Carifio, J. and Perla, R. (2008), "Resolving the 50-year debate around using and misusing Likert scales", *Medical Education*, Vol. 42 No. 12, pp. 1150–1152.

Chen, I.J. and Paulraj, A. (2004), "Towards a theory of supply chain management: the constructs and measurements", *Journal of Operations Management*, Vol. 22 No. 2, pp. 119–150.

Cliff, N. (1993), "Dominance statistics: ordinal analyses to answer ordinal questions", *Psychological Bulletin*, Vol. 114 No. 3, pp. 494–509.

Conover, W.J. and Iman, R.L. (1981), "Rank transformations as a bridge between parametric and nonparametric statistics", *The American Statistician*, Vol. 35 No. 3, pp. 124–129.

Conroy, R.M. (2012), "What hypotheses do 'nonparametric' two-group tests actually test?", *The Stata Journal*, Vol. 12 No. 2, pp. 182–190.

Divine, G.W., Norton, H.J., Barón, A.E. and Juarez-Colunga, E. (2018), "The Wilcoxon–Mann–Whitney procedure fails as a test of medians", *The American Statistician*, Vol. 72 No. 3, pp. 278–286.

Fagerland, M.W. and Sandvik, L. (2009a), "The Wilcoxon–Mann–Whitney test under scrutiny", *Statistics in Medicine*, Vol. 28 No. 10, pp. 1487–1497.

Fagerland, M.W. and Sandvik, L. (2009b), "Performance of five two-sample location tests for skewed distributions with unequal variances", *Contemporary Clinical Trials*, Vol. 30 No. 5, pp. 490–496.

Fawcett, S.E., Osterhuis, P. and McCarter, M.W. (2007), "Information sharing and supply chain performance: the role of connectivity and willingness", *Supply Chain Management: An International Journal*, Vol. 12 No. 5, pp. 358–368.

Field, A. (2018), *Discovering Statistics Using IBM SPSS Statistics*, 5th ed., Sage, London.

Flynn, B.B., Huo, B. and Zhao, X. (2010), "The impact of supply chain integration on performance: a contingency and configuration approach", *Journal of Operations Management*, Vol. 28 No. 1, pp. 58–71.

Frohlich, M.T. and Westbrook, R. (2001), "Arcs of integration: an international study of supply chain strategies", *Journal of Operations Management*, Vol. 19 No. 2, pp. 185–200.

Gunasekaran, A. and Kobu, B. (2007), "Performance measures and metrics in logistics and supply chain management: a review of recent literature for research and applications", *International Journal of Production Research*, Vol. 45 No. 12, pp. 2819–2840.

Gunasekaran, A., Patel, C. and Tirtiroglu, E. (2001), "Performance measures and metrics in a supply chain environment", *International Journal of Operations & Production Management*, Vol. 21 No. 1/2, pp. 71–87.

Hodges, J.L. and Lehmann, E.L. (1963), "Estimates of location based on rank tests", *The Annals of Mathematical Statistics*, Vol. 34 No. 2, pp. 598–611.

Holm, S. (1979), "A simple sequentially rejective multiple test procedure", *Scandinavian Journal of Statistics*, Vol. 6 No. 2, pp. 65–70.

Koh, S.C.L., Demirbag, M., Bayraktar, E., Tatoglu, E. and Zaim, S. (2007), "The impact of supply chain management practices on performance of SMEs", *Industrial Management & Data Systems*, Vol. 107 No. 1, pp. 103–124.

Li, S., Ragu-Nathan, B., Ragu-Nathan, T.S. and Rao, S.S. (2006), "The impact of supply chain management practices on competitive advantage and organizational performance", *Omega*, Vol. 34 No. 2, pp. 107–124.

Mann, H.B. and Whitney, D.R. (1947), "On a test of whether one of two random variables is stochastically larger than the other", *The Annals of Mathematical Statistics*, Vol. 18 No. 1, pp. 50–60.

Norman, G. (2010), "Likert scales, levels of measurement and the 'laws' of statistics", *Advances in Health Sciences Education*, Vol. 15 No. 5, pp. 625–632.

Nunnally, J.C. and Bernstein, I.H. (1994), *Psychometric Theory*, 3rd ed., McGraw-Hill, New York, NY.

Podsakoff, P.M., MacKenzie, S.B., Lee, J.-Y. and Podsakoff, N.P. (2003), "Common method biases in behavioral research: a critical review of the literature and recommended remedies", *Journal of Applied Psychology*, Vol. 88 No. 5, pp. 879–903.

Prajogo, D. and Olhager, J. (2012), "Supply chain integration and performance: the effects of long-term relationships, information technology and sharing, and logistics integration", *International Journal of Production Economics*, Vol. 135 No. 1, pp. 514–522.

Quayle, M. (2003), "A study of supply chain management practice in UK industrial SMEs", *Supply Chain Management: An International Journal*, Vol. 8 No. 1, pp. 79–86.

Romano, J., Kromrey, J.D., Coraggio, J. and Skowronek, J. (2006), "Appropriate statistics for ordinal level data: should we really be using *t*-test and Cohen's *d* for evaluating group differences on the NSSE and other surveys?", paper presented at the Annual Meeting of the Florida Association of Institutional Research, Cocoa Beach, FL.

Sahay, B.S. and Mohan, R. (2003), "Supply chain management practices in Indian industry", *International Journal of Physical Distribution & Logistics Management*, Vol. 33 No. 7, pp. 582–606.

Singh, R.K., Garg, S.K. and Deshmukh, S.G. (2008), "Strategy development by SMEs for competitiveness: a review", *Benchmarking: An International Journal*, Vol. 15 No. 5, pp. 525–547.

Teece, D.J., Pisano, G. and Shuen, A. (1997), "Dynamic capabilities and strategic management", *Strategic Management Journal*, Vol. 18 No. 7, pp. 509–533.

Thakkar, J., Kanda, A. and Deshmukh, S.G. (2008), "Supply chain management in SMEs: development of constructs and propositions", *Asia Pacific Journal of Marketing and Logistics*, Vol. 20 No. 1, pp. 97–131.

Thakkar, J., Kanda, A. and Deshmukh, S.G. (2011), "Mapping of supply chain learning: a framework for SMEs", *The Learning Organization*, Vol. 18 No. 4, pp. 313–332.

Vaaland, T.I. and Heide, M. (2007), "Can the SME survive the supply chain challenges?", *Supply Chain Management: An International Journal*, Vol. 12 No. 1, pp. 20–31.

Vargha, A. and Delaney, H.D. (2000), "A critique and improvement of the CL common language effect size statistics of McGraw and Wong", *Journal of Educational and Behavioral Statistics*, Vol. 25 No. 2, pp. 101–132.

Wagner, S.M. and Bode, C. (2008), "An empirical examination of supply chain performance along several dimensions of risk", *Journal of Business Logistics*, Vol. 29 No. 1, pp. 307–325.

Wernerfelt, B. (1984), "A resource-based view of the firm", *Strategic Management Journal*, Vol. 5 No. 2, pp. 171–180.

Wilcoxon, F. (1945), "Individual comparisons by ranking methods", *Biometrics Bulletin*, Vol. 1 No. 6, pp. 80–83.

Wong, C.Y., Boon-itt, S. and Wong, C.W.Y. (2011), "The contingency effects of environmental uncertainty on the relationship between supply chain integration and operational performance", *Journal of Operations Management*, Vol. 29 No. 6, pp. 604–615.

Zhou, H. and Benton, W.C. (2007), "Supply chain practice and information sharing", *Journal of Operations Management*, Vol. 25 No. 6, pp. 1348–1365.

Zhu, Q. and Sarkis, J. (2004), "Relationships between operational practices and performance among early adopters of green supply chain management practices in Chinese manufacturing enterprises", *Journal of Operations Management*, Vol. 22 No. 3, pp. 265–289.

Appendix A. Measurement instrument

All efficiency items use a five-point agreement scale (1 = strongly disagree, 5 = strongly agree) with the stem "Considering the last twelve months of operations in your firm…".

Procurement and sourcing efficiency (PSE)

1. Purchase requirements are consolidated so that we avoid repeated small, urgent orders.
2. We obtain competitive prices from our suppliers relative to what comparable firms pay.
3. The time from identifying a material requirement to placing a confirmed order is short.
4. Emergency or off-plan purchasing is rare in our operations.

Inventory management efficiency (IME)

5. Our stock levels are matched to actual consumption rather than held as a general buffer.
6. Material shortages that halt or delay production are rare.
7. We seldom write off or discount material because it has become obsolete or damaged in storage.
8. We know our current stock position accurately without physically checking.

Logistics and distribution efficiency (LDE)

9. Our deliveries to customers arrive on the committed date.
10. Our transport and freight cost per unit dispatched is competitive.
11. Inbound material arrives when we expect it to arrive.
12. We rarely pay premium freight to recover lost time.

Information integration efficiency (IIE)

13. Everyone who needs demand information in our firm is working from the same figures.
14. Our suppliers receive our requirements early enough to respond without expediting.
15. Information about a change in a customer order reaches purchasing and production quickly.
16. We can answer a customer's status query without having to check with several people.
17. Order, stock and dispatch information is recorded in a form we can retrieve and use later.

Supplier relationship efficiency (SRE)

18. Our key suppliers accommodate changes in our requirements when we ask.
19. Problems with supplied material are resolved with the supplier without escalation or dispute.
20. Our key suppliers treat us as a priority customer relative to their other accounts.
21. We can rely on our key suppliers' commitments without formal follow-up.

Demand planning and forecasting efficiency (DPE)

22. We plan our material requirements against an explicit demand estimate rather than reacting to orders as they arrive.
23. Our estimate of next month's requirement is usually close to what actually happens.

24. Capacity and manpower are planned ahead of the period in which they are needed.
25. We are seldom caught with committed capacity and no orders, or orders and no capacity.