Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

One boundary in two traces: concept vocabulary and cited references in base- of-the-pyramid, social enterprise and social entrepreneurship research, 2000 to 2025

Faria Zahan Sarna

faria.sarna@metrouni.edu.bd

Metropolitan University

Research Article

Keywords: Bibliometric analysis, Base of the pyramid, Social enterprise, Social entrepreneurship, Bibliographic coupling, Inclusive business

Posted Date: November 6th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11366380/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Research on market-based approaches to poverty centres on a small set of labels, including base of the pyramid, social enterprise, social entrepreneurship and inclusive business. Whether these labels index distinct bodies of work has consequences for how evidence is assembled and used. This study examines 554 journal articles published between 2000 and 2025, of which 461 carry at least one label, and compares the labels on two traces of scholarly organisation: the indexed concept vocabulary attached to each paper, and the prior work each paper cites. Both traces are tested against the same permutation null, generated by reassigning the corpus's own labels at random. The same boundary appears on both. Base-of-the-pyramid research is separated from social enterprise and social entrepreneurship research on vocabulary and on cited references alike, and in both cases the separation exceeds chance after correction across the six-pair family. The boundary is deeper on citations: overlap across it falls to 44 per cent of the within-pair value for cited references, against 61 per cent for vocabulary. Social enterprise and social entrepreneurship are not separated on either trace, which is a non-detection and does not establish their equivalence. A bibliographic coupling network is substantially modular under the same separation, but does not identify a unique placement for inclusive business, whose pools are too small for comparable inference. Findings describe indexing and citation behaviour in the analysed corpus, not the conceptual content of the constructs.

JEL Classification I32; L26; L31; M14; O35

1 Introduction

Market-based approaches to poverty reduction are discussed through a small vocabulary. Programmes are designed, funded and evaluated under labels such as base of the pyramid, social enterprise, social entrepreneurship and inclusive business. These labels organise more than terminology. They determine which journals a researcher reads, which prior results are treated as relevant evidence, and which assumptions about mechanism and measurement travel into a new study or a new programme design. Whether the labels index different bodies of evidence is therefore a practical question and not only a semantic one. If research published under two labels reads and cites largely the same prior work, then the labels are alternative names for one conversation, and evidence accumulates across them. If the reference bases diverge, then choosing a label selects an evidence base, and findings established under one label may not be reaching researchers working under another. The second situation is harder to detect from inside either literature, because the absence of a citation is not visible to the person who does not make it.

Existing reviews have mapped each of these literatures, and several have noted definitional instability within them (Chliova et al. 2020; Dembek et al. 2020; Doherty et al. 2014; Kolk et al. 2014; Saebi et al. 2019; Short et al. 2009). What has received less attention is the relationship between the labels, measured rather than asserted. Reviews tend to be organised around one label, which makes cross-label comparison a by-product rather than an object of analysis.

This study treats that relationship as the object. It assembles 554 journal articles published between 2000 and 2025, of which 461 carry at least one of the four labels, and compares the labelled groups on two measures: the conceptual vocabulary attached to each paper, and the prior work each paper cites. Observed overlap is compared against a null distribution generated by reassigning the corpus's own labels at random, so that the question is not whether overlap is large but whether it departs from what arbitrary grouping would produce.

Five research questions are stated in Section 3.2 and answered in Sections 4 to 7. The study's contribution is stated plainly, in line with the journal's fifth publication criterion. It examines the relationship between conceptual vocabulary and citation patterns across research on base of the pyramid, social enterprise and social entrepreneurship. It finds that the same boundary separates base-of-the-pyramid research from the social-economy pair on both traces, that the separation is deeper on cited references than on indexed vocabulary, and that the available evidence neither supports a unique placement of inclusive business nor establishes equivalence between social enterprise and social entrepreneurship.

This differs from the expectation that motivated the design. The study was built to test whether a shared vocabulary could coexist with divergent citation bases, which would have indicated communities talking alike while reading differently. The measurement does not support that reading. Vocabulary and citations point the same way, and the contribution is therefore the convergence of two traces on one boundary, together with the difference in its depth, rather than a contrast between them.

Three boundaries are stated here rather than deferred to the limitations, because they shape how every later result should be read. First, the analysis measures labels and citations. Citation behaviour is treated as an observable trace of which prior work an author treats as relevant (Boyack and Klavans 2010; Crane 1972; Small 1973). It is not a measure of conceptual content, so finding that two labelled literatures cite different work does not establish that the underlying constructs are distinct. Second, the analysis does not detect a separation between social enterprise and social entrepreneurship, and a non-detection is not evidence of equivalence (Altman and Bland 1995; Greenland et al. 2016; Lakens 2017). Third, inclusive business is represented by ten labelled papers and contributes one connected node to the displayed network. Its pairwise comparisons are reported as exploratory rather than confirmatory, because its vocabulary and reference pools are too small to support comparable inference, and no unique placement is claimed for it. Exploratory here means analysed and reported, including inside the correction family, but not used as the basis for a confirmatory claim.

Section 2 sets out the conceptual background and the interpretive lenses used. Section 3 describes the data and methods, including the research questions, the inferential strategy and the use of software tools. Section 4 profiles the corpus. Sections 5 and 6 report the vocabulary, citation and network results. Section 7 reports robustness. Sections 8 to 11 discuss the findings, draw implications, state limitations and conclude.

2 Conceptual background

2.1 The four labels

The four labels have separate origins, and their histories help explain why a measured comparison is worth making.

Base of the pyramid emerged from strategy and international business as a proposition about unserved markets, arguing that low-income populations constitute a commercially viable market and that serving it requires departures from established corporate practice (Hart and Christensen 2002; London and Hart 2004; Prahalad 2004; Prahalad and Hart 2002). It attracted early and sustained criticism on both evidential and ethical grounds (Karnani 2007). Reviews of the field a decade on documented a shift from selling to the poor toward working with them, and noted persistent difficulty in specifying what the concept excludes (Dembek et al. 2020; Kolk et al. 2014).

Social enterprise and social entrepreneurship developed largely within entrepreneurship, management and nonprofit scholarship, with distinct European and North American framings (Defourny and Nyssens 2010). Work here has concentrated on organisational form, hybridity and the management of competing institutional demands (Battilana and Lee 2014; Doherty et al. 2014; Santos 2012), on the boundaries of the entrepreneurial phenomenon (Austin et al. 2006; Mair and Marti 2006; Zahra et al. 2009), and on whether the field requires theory of its own (Dacin et al. 2010). Several contributions treat the categories themselves as unstable, and one documents category ambiguity as a persistent rather than transitional condition (Chliova et al. 2020; Saebi et al. 2019).

Inclusive business is the most recent and the least settled of the four. It is used in development policy and practice as well as in scholarship, often to describe business models that incorporate low-income populations as suppliers, distributors, employees or consumers (George et al. 2012; Hahn 2012; Likoko and Kini 2017). Its use in the indexed journal literature is correspondingly thin, which is reflected directly in the corpus assembled here and constrains what can be concluded about it.

These genealogies suggest two possibilities that observation can distinguish. The labels may have converged on a shared literature as the problems they address overlapped, or they may have retained separate reference bases while adopting a common vocabulary. The analysis that follows tests which pattern the citation record shows.

2.2 Interpretive lenses

The study does not test a theory, and no hypotheses are derived from one. It measures citation structure and interprets it. Three established lenses are used to interpret the patterns observed, and they are named here so that their role stays clear.

Construct clarity concerns whether a concept's definition, scope conditions and relationships to adjacent concepts are specified well enough for cumulative work (Suddaby 2010). It provides the vocabulary for

discussing what undifferentiated conceptual overlap between labels might mean.

The invisible college describes a community of researchers who read, cite and respond to one another, forming a bounded conversation that may not coincide with disciplinary or terminological boundaries (Crane 1972). It is the natural frame for interpreting a modular citation network, because modularity is a property of who cites whom.

Boundary objects are artefacts or concepts that hold enough shared meaning to coordinate work across communities while retaining different local interpretations within each (Star and Griesemer 1989). This offers a reading of the central pattern reported below, in which a vocabulary is shared while reference bases are not.

These lenses are applied after the measurement and are used to interpret it. They are not pre-registered predictions, and no result below should be read as confirming or disconfirming any of them.

3 Data and methods

3.1 Research design and the inference it licenses

The study asks whether four labels used at the business and development interface correspond to different patterns of scholarly citation. It assembles a corpus of journal articles carrying those labels, measures how far the labelled groups share a conceptual vocabulary, measures how far they draw on the same cited works, and tests whether any observed separation exceeds what reassigning the labels at random would produce.

The design measures labels and citations. It does not read conceptual content and does not observe how researchers understand or apply a construct. The distinction governs every claim that follows. Finding that two labelled literatures cite different work establishes that they read different work. It does not establish that the underlying constructs are distinct, that their definitions are stable, or that practitioners use them differently.

Three boundaries are fixed at the outset. First, the population is the retrieved, screened, OpenAlex-matched journal literature, not the whole of scholarship associated with these concepts. Books, chapters, practitioner reports and policy publications fall outside it, which matters particularly for inclusive business, much of whose output is practitioner and policy work. Second, classification precision was audited but recall was not, so the corpus is not described as complete. Third, the archived records contain no affiliation data, so no claim is made about countries, institutions or collaboration.

3.2 Research questions

RQ1. How has research using the four labels developed between 2000 and 2025, where is it published, and what body of prior work does it draw on?

RQ2. Do the four labels share a common conceptual vocabulary?

RQ3. Do the four labels draw on a common body of cited work, and is any observed separation greater than would arise from chance grouping?

RQ4. Does the structure observed across pooled reference sets also hold at the level of individual papers, and does the network identify a unique grouping of the four labels?

RQ5. Is the observed structure stable across time and robust to corpus composition?

3.3 Corpus construction and screening

Records were retrieved from Crossref, Scopus and Google Scholar under the protocol set out in Appendix A and Online Resource 1, Section 1. Retrieval operated under per-source result ceilings and returned 1,400 records, comprising 1,000 from Crossref, 200 from Google Scholar and 200 from Scopus. Deduplication on normalised title and on DOI reduced these to 739 unique records, contributed as 380 from Crossref, 186 from Google Scholar and 173 from Scopus. Of these, 614 matched an OpenAlex record by DOI, which is the condition for retrieving structured reference and concept data (Priem et al. 2022). Removing non-journal documents, principally working papers, conference proceedings, edited volumes and handbooks, left a final corpus of 554 journal articles published between 2000 and 2025 across 290 journals. Figure 1 reports the chain.

Three distinct source-level counts appear in this study and are not interchangeable. The ceilings above describe raw returns. The 739 unique records decompose as 380, 186 and 173 by contributing source. The final 554 articles decompose as Crossref 324, Scopus 170 and Google Scholar 60. Each is reported where it belongs and none is used as a proxy for another.

Each stage removes records, and the removals are not random. The 125 records without an OpenAlex match skew recent and social-enterprise-leaning, which is reported in Section 10. Retrieval operated under source-specific result ceilings, so the corpus is a large sample of the indexed journal literature rather than a census of it. Reporting follows the structure recommended for transparent review protocols (Page et al. 2021), adapted to a bibliometric design.

3.4 Label assignment and audit

A paper is assigned a label when a regular expression for that label matches its title or abstract. The expressions allow for plural forms, hyphenation variants and the common abbreviation of base of the pyramid, and are listed in full in Appendix A. Papers may carry more than one label, and no label is forced.

Of the 554 articles, 461 carry at least one label and 93, or 16.8 per cent, carry none. The unlabelled group was retrieved by the same strategy but names none of the four constructs in its title or abstract. It is retained as a contextual corpus rather than discarded, and it is excluded from all label comparisons. Group sizes are social enterprise 105, social entrepreneurship 121, base of the pyramid 297 and

inclusive business 10. Counting papers carrying exactly one label gives 56, 77, 259 and 4 respectively, with a further 65 papers carrying two or more labels, of which 33 span what the analysis later treats as different domains. Table 1 reports composition by label and period.

Table 1
Corpus composition by period and label

Period	Articles	Labelled	SE	SEship	BoP	IB
2000–2006	27	23	6	16	3	0
2007–2012	129	106	22	38	54	1
2013–2017	179	148	36	39	97	3
2018–2025	219	184	41	28	143	6
Total	**554**	**461**	**105**	**121**	**297**	**10**

SE social enterprise, SEship social entrepreneurship, BoP base of the pyramid, IB inclusive business. Label counts exceed the labelled total because 65 articles carry more than one label. The 93 articles carrying no label are the difference between the Articles and Labelled columns.

A manual audit read 25 randomly sampled labelled records in full and judged each assignment against the paper's substantive focus. Twenty-four of the 25 were consistent, a precision of 96 per cent, with a Wilson 95 per cent interval from 80.5 to 99.3 per cent. The sample is small and was pooled across labels, so it supports no label-specific claim. Recall was not estimated. Establishing it would require hand-coding a sample of unlabelled records, which was not undertaken. Its absence is treated as a limit on representativeness rather than a minor caveat, and it is the reason no claim in this paper is made about the field as a whole.

Abstracts are available for 418 of the 554 articles, or 75.5 per cent. Because the labelling rule reads title and abstract together, records without an abstract have fewer opportunities to match. The labelling rate is 86.1 per cent among records with an abstract and 74.3 per cent among those without, a gap of 11.8 percentage points. The gap is not evenly distributed across labels, and its consequences are tested directly in Section 7.1.

3.5 Analytical lenses

Four lenses are applied to the labelled corpus.

The vocabulary lens compares the sets of OpenAlex concepts attached to each labelled group using the Jaccard coefficient, which expresses the share of the combined vocabulary two groups hold in common.

The pooled citation lens compares the union of works cited by each group, again by Jaccard, with the cosine and overlap coefficients reported alongside because Jaccard is sensitive to differences in set

size (Salton and McGill 1983). Containment, the share of one group's references lying inside another's, is reported where set sizes differ sharply.

The paper-level lens computes similarity between individual papers rather than between pooled sets, which guards against pooled overlap being an artefact of group size.

The network lens builds a bibliographic coupling graph in which two papers are joined when they share at least two references and their Salton-normalised coupling reaches 0.05 (Kessler 1963). The graph is restricted to single-label papers so that each node carries an unambiguous label. Of these, 349 carry retrievable references, and the resulting graph has 5,272 edges over 296 connected nodes, leaving 53 isolates. A co-citation network over frequently cited works is reported as a secondary representation in Online Resource 1, Section 3 (Small 1973).

Reference availability is reported on two populations, which must not be conflated. Of the 461 labelled papers, 405, or 87.9 per cent, carry retrievable references, and those 405 are the papers entering the pooled citation comparisons in Section 5. Across the full screened corpus, 482 of 554 articles, or 87.0 per cent, carry retrievable references, and that larger population is the basis for the reference-base descriptives reported in Section 4. Analyses were implemented in Python, and the code, the banked results and a redistributable reproduction dataset are deposited; the full code and banked results are provided in Online Resource 2; comparable workflows are available in established bibliometric software (Aria and Cuccurullo 2017; van Eck and Waltman 2010), and the measurement choices follow standard guidance for the field (Donthu et al. 2021; Zupic and Čater 2015).

3.6 Inferential strategy

Observed similarity values are compared against a null generated by reassigning the corpus's own label sets at random across the 461 labelled papers, preserving group sizes and multi-label structure, and recomputing each statistic (Good 2005). The reassignment is confined to labelled papers. Including unlabelled papers in the pool would allow label sets to land on papers that never carried one, which alters every pooled reference set and inflates the null.

The confirmatory test of pooled reference overlap uses 300,000 draws in a single run. Empirical p is reported under the add-one convention, (b + 1) / (n + 1) (Phipson and Smyth 2010). Six pairwise comparisons are made, all six form the Holm family, and the adjustment is computed on the same run that produced the raw values (Holm 1979).

The draw count is set at 300,000 for a specific reason. Two earlier 10,000-draw runs of this test returned 1,177 and 1,267 draws at or below the observed value for the social-economy pair, p = .118 and p = .127. At 300,000 draws the Monte Carlo standard error for that pair is 0.0006 and the estimate is .121, so both earlier values lie within 1.8 standard errors of it and the difference between them was Monte Carlo noise rather than a substantive disagreement. The larger run also lifts the two cross-boundary pairs off the reporting floor, giving point estimates instead of the bound that 10,000 draws can supply. The robustness specifications in Section 7 retain 10,000 draws, which is adequate for sensitivity checks, and

the one consequence of that choice is stated in Section 7.2. Effect sizes are reported as the distance of each observed value from its null in null standard deviations, because significance alone conveys nothing about magnitude.

For the network, modularity is computed for a specified partition and compared against nulls in which labels are reassigned across nodes while the graph is held fixed (Newman 2006; Newman and Girvan 2004). Because the two sides differ substantially in connectivity, the primary null reassigns labels only within degree quartiles, preventing the degree-label association from inflating the statistic. Modularity here is Newman's measure, in which the null term is summed over every same-community pair rather than only over connected pairs. The implementation was validated against direct evaluation of the definition on seven test graphs spanning weighted edges, negative values, three communities, disconnected components and isolates, and that validation is reported in Online Resource 1, Section 2. Assortativity follows Newman (2003), label-free community detection uses label propagation (Raghavan et al. 2007), and agreement between partitions is measured by the adjusted Rand index and normalised mutual information (Danon et al. 2005; Hubert and Arabie 1985).

Where several candidate partitions are compared, two further procedures apply. The difference between two partitions is tested by paired randomisation, computing both statistics on the same reassignment so that their dependence is preserved. The best of several searched partitions is tested against the distribution of the best partition under the null, correcting for having searched at all. Paper-level intervals come from a bootstrap that resamples papers rather than pairs, respecting the dependence between pairs sharing a paper (Efron and Tibshirani 1993).

3.7 Use of artificial intelligence tools

A large language model (ChatGPT) was used to check and correct Python syntax in the analysis scripts. It did not select methods, specify models, choose parameters, interpret output or generate any numerical result reported in this paper. All analytical decisions, all specifications and all interpretation are the authors'. Every reported statistic was produced by executing the deposited code, and the code and its output are available in Online Resource 2 so that any reader can reproduce the values independently. The authors are accountable for the final text and for the correctness of the analysis.

4 The literature in profile

Annual output rose from a mean of 3.9 papers a year in 2000 to 2006 to 27.4 a year in 2018 to 2025, a 7.1-fold increase. Counts by period are 27, 129, 179 and 219 (Figure 2; Table 1). Shares by label across the four periods are shown in Figure 3.

The literature is dispersed. Its 554 articles appear in 290 journals. The most frequent outlet is the *Journal of Business Ethics* with 33 papers, or 6.0 per cent of the corpus. The ten most frequent carry 24.7 per cent between them, the fifty most frequent carry 50.2 per cent, and 204 journals, or 70.3 per cent, publish

exactly one paper each. A Bradford partition gives zones of 19, 89 and 182 journals (Bradford 1934). No venue organises this field.

Received citations are highly concentrated. The corpus has accumulated 112,208 citations, with a median of 34 and a maximum of 3,907, and 62 articles are uncited. The Gini coefficient of the citation distribution is 0.734, the top decile of articles holds 52.9 per cent of all citations, and the corpus h-index is 182 (Hirsch 2005). Table 2 reports the descriptive profile.

Table 2 Descriptive profile of the corpus

Dimension	Measure	Value
Growth	Mean annual output, 2000–2006	3.9
	Mean annual output, 2018–2025	27.4
	Growth multiple	7.1
Outlets	Distinct journals	290
	Most frequent outlet	*Journal of Business Ethics*, 33 (6.0%)
	Share in 10 most frequent outlets	24.7%
	Share in 50 most frequent outlets	50.2%
	Journals publishing one article	204 (70.3%)
	Bradford zones	19 / 89 / 182
Received citations	Total	112,208
	Median per article	34
	Maximum	3,907
	Uncited articles	62
	Gini coefficient	0.734
	Share held by top decile	52.9%
	Corpus h-index	182
Reference base	Labelled articles with references	405 of 461 (87.9%)
	Citation links	34,307
	Distinct cited works	21,302
	Mean references per article	71.2
	Cited once only	81.3%
	Cited ten or more times	287
	Of 100 most co-cited, inside corpus	40

Citation counts are global and reflect all citing literature, not only this corpus.

The shared intellectual base is thin. Across the 482 screened articles that carry retrievable references, 34,307 citation links reach 21,302 distinct works, a mean of 71.2 references per article, and 81.3 per cent of those works are cited by only one paper in the corpus. Just 287 works are cited ten or more times. Of the hundred most co-cited works, 40 are themselves inside the corpus and 60 lie outside it.

The most heavily co-cited works already fall into two groups. Within the corpus the leading items are London and Hart (2004) with 92 co-citations, Mair and Marti (2006) with 88, Karnani (2007) with 79, Austin et al. (2006) with 70 and Seelos and Mair (2007) with 60. The first, third and fifth belong to the base-of-the-pyramid conversation and the second and fourth to the social-economy one, which anticipates the pattern reported in Section 5 without establishing it.

Received citations are also distributed unevenly across labels, and the comparison depends on how publication age is handled. Raw median citation counts for single-label papers are 430 for social entrepreneurship, 379 for inclusive business, 248 for social enterprise and 17 for base of the pyramid. Dividing by years since publication, with 2026 as the reference year, gives medians of 31.8, 31.1, 17.7 and 2.0 per year. Both measures are reported because they are different quantities. The per-year figures divide by elapsed time and nothing else, so they do not account for the shape of citation accumulation over a paper's life, and base-of-the-pyramid papers are on average the more recent. Neither measure should be read as an age-adjusted estimate in any stronger sense.

This profile answers RQ1. It establishes the scope and maturity of the corpus and shows that its citation base is thin and already uneven. It motivates the questions that follow without answering them.

5 Vocabulary and cited references

5.1 Vocabulary separates the same groups that citations do

Concept vocabulary is measured from OpenAlex concepts at levels two and three with an assignment score of at least 0.3. This specification is stated because the result depends on it: concept pools hold 264 terms for social enterprise, 297 for social entrepreneurship, 506 for base of the pyramid and 43 for inclusive business.

Pairwise concept-vocabulary Jaccard is 0.399 between social enterprise and social entrepreneurship, against 0.250 and 0.234 across the boundary to base of the pyramid (Fig. 4a; Table 3). Subjected to the same permutation null used for cited references, both cross-boundary comparisons fall outside it, at $p = 0.00036$ and $p = 0.0000033$, surviving Holm correction over the six-pair family. The social-economy pair sits inside the null at $p = .412$. Vocabulary therefore separates the same groups that cited references separate, and fails to separate the same pair. This answers RQ2, and it answers it differently from the way the study was designed to expect.

Table 3
Concept-vocabulary overlap and its permutation inference

Pair	Status	Observed J	Null mean	Effect size (null SD)	p	p (Holm)
SE / SEship	confirmatory	0.3990	0.4050	-0.24	0.4118	1.0000
SE / BoP	confirmatory	0.2500	0.3032	-3.26	0.0004	0.0018
SEship / BoP	confirmatory	0.2335	0.3056	-4.84	3.3e-06	2.0e-05
SE / IB	exploratory	0.1004	0.1022	-0.11	0.4683	1.0000
SEship / IB	exploratory	0.0897	0.1023	-0.76	0.2250	0.9002
BoP / IB	exploratory	0.0786	0.0790	-0.04	0.5012	1.0000

Concept pools hold 264 terms for social enterprise, 297 for social entrepreneurship, 506 for base of the pyramid and 43 for inclusive business, from OpenAlex concepts at levels two and three with an assignment score of at least 0.3. The null and the correction family are identical to those used for cited references in Table 6, so the two traces are directly comparable. One run of 300,000 draws, seed 42, produced every value here. Inclusive business comparisons are exploratory, for the reasons in Section 5.3.

The separation is not produced by the label words themselves. Removing the four concepts that restate a label, namely social enterprise, social entrepreneurship, bottom of the pyramid and pyramid, leaves the ratio of within-pair to cross-boundary overlap at 1.66 against 1.65 with them retained. Removing instead every concept present in at least a fifth of the corpus gives 0.393, 0.245 and 0.229, which preserves both the ordering and the magnitudes. Neither filter changes the finding.

Two boundaries on the measure should be kept in view. OpenAlex concepts are assigned from title and abstract text, and the labels are assigned from the same text, so the two are not independent measurements of a paper. The label-term filter addresses the most direct form of that dependence but not subtler forms, such as topic words that co-occur with a label without naming it. And a concept pool is a property of how a corpus is indexed, not of how its authors define their terms.

5.2 Base of the pyramid has lower cited-reference overlap with the social-economy labels

Pooled cited-reference overlap shows a different pattern (Fig. 4b; Table 4). Social enterprise and social entrepreneurship share 0.233 of their combined reference pool. Each shares far less with base of the pyramid: 0.103 and 0.101 respectively. In absolute terms the social-economy pair shares 2,156 cited works, against 1,488 and 1,605 across the boundary, from pools of 4,929, 6,484 and 10,968 works. Among the fifty most cited works in each group, the social-economy pair shares 24, against 7 and 6 across the boundary.

Table 4
Pairwise similarity of pooled cited-reference sets

Pair	Jaccard	Cosine	Overlap coef.	Shared works	Union
SE / SEship	0.233	0.381	0.437	2,156	9,257
SE / BoP	0.103	0.202	0.302	1,488	14,409
SEship / BoP	0.101	0.190	0.248	1,605	15,847
SE / IB	0.025	0.079	0.224	137	5,405
SEship / IB	0.034	0.116	0.377	231	6,866
BoP / IB	0.031	0.133	0.564	346	11,235

Pool sizes: SE 4,929; SEship 6,484; BoP 10,968; IB 613. The rank order of the three non-IB pairs is identical under all three coefficients. The IB pairs reorder between Jaccard and the overlap coefficient, which is the expected consequence of a small pool largely contained within a larger one, and is the reason IB is discussed through containment rather than Jaccard.

Against 300,000 reassignments of the corpus's own labels, the social-economy pair sits inside the null, with 36,288 draws at or below the observed value, giving p = .121. Both cross-boundary pairs sit far outside it, with 3 and 2 draws respectively at or below, giving p = 0.0000133 and p = 0.0000100, and both survive Holm correction over the six-pair family at 0.0000667 and 0.0000600. Expressed as distance from the null, the cross-boundary values lie 3.92 and 4.07 null standard deviations below expectation, against 1.16 for the social-economy pair. Table 5 reports the inference. Together with Section 6 this answers RQ3.

Table 5
Permutation inference on pooled cited-reference overlap

Pair	Status	Observed J	Null mean	Effect size (null SD)	Draws at or below	p	p (Holm)
SE / SEship	confirmatory	0.2329	0.2598	−1.16	36,288	.1210	.4839
SE / BoP	confirmatory	0.1033	0.1498	−3.92	3	.000013	.000067
SEship / BoP	confirmatory	0.1013	0.1422	−4.07	2	.000010	.000060
SE / IB	exploratory	0.0253	0.0302	−0.59	86,343	.2878	.8634
SEship / IB	exploratory	0.0336	0.0354	−0.17	137,543	.4585	.8634
BoP / IB	exploratory	0.0308	0.0358	−0.52	94,317	.3144	.8634

Label sets are reassigned at random across the 461 labelled papers, preserving group sizes and multi-label structure. One run of 300,000 draws, seed 42, produced every value in this table, including the Holm adjustment, so no row combines figures from different Monte Carlo streams. Empirical p uses the add-one convention. The Holm family is all six pairs, which is why the three inclusive business comparisons appear: excluding them from the table while including them in the correction would misstate the adjustment. They are marked exploratory and no inference is drawn from them, for the reasons in Section 5.3. The Monte Carlo standard error for the social-economy pair is 0.0006, giving a 95 per cent interval of [.1198, .1221]. Two earlier 10,000-draw runs of this test returned p = .118 and p = .127, both within 1.8 standard errors of the value above.

The correction family is all six pairwise comparisons, including the three involving inclusive business. Those three are reported in Table 5 so that the adjustment is auditable, and they are marked exploratory because the group is too small to support inference. Including them makes the correction more conservative for the comparisons that do carry the argument, which is the defensible direction for the error to run.

The two traces can now be compared on the same footing, because the same null and the same correction family were applied to both. Cross-boundary overlap falls to 44 per cent of the within-pair value for cited references (0.103 and 0.101 against 0.233) and to 61 per cent for vocabulary (0.250 and 0.234 against 0.399). Expressed as a ratio of within-pair to cross-boundary overlap, the boundary is 2.28 deep on citations and 1.65 on vocabulary. The direction agrees, the magnitude does not. What the labels index and what their authors read both divide at the same place, and they divide more sharply in what is read.

Effect sizes do not order the two traces as cleanly as the ratios do. Against the null, the cross-boundary citation values lie 3.92 and 4.07 standard deviations below expectation and the vocabulary values 3.26 and 4.84. The vocabulary separation for social entrepreneurship and base of the pyramid is the largest of the four. The claim that the boundary is deeper on citations therefore rests on the proportional drop in overlap, which is a statement about magnitude, and not on the test statistics, which do not support an ordering.

The social-economy result requires care in statement. The test does not detect a separation between social enterprise and social entrepreneurship, and that is all it shows. It does not establish that the two literatures cite the same work, and it is not evidence of equivalence. A null result of this kind is consistent both with genuinely shared citation practice and with a real difference this corpus and this test lack the power to detect, and nothing in the design distinguishes those possibilities (Altman and Bland 1995; Greenland et al. 2016). Establishing equivalence would require a pre-specified margin of practical indifference and a test against it, which this design does not provide (Lakens 2017). The pair is never significant in any of the thirteen specifications in Section 7, which makes the non-detection stable, but stability of a null result is not evidence for the null.

The ordering is not an artefact of the similarity measure. Under the cosine coefficient the three values are 0.381, 0.202 and 0.190, and under the overlap coefficient 0.437, 0.302 and 0.248. The rank order of these three pairs is identical under all three measures. Nor does it rest on works cited once. Restricting each pool to works cited at least two, three or five times within its group preserves the ordering throughout, and the social-economy value rises rather than falls, from 0.233 to 0.277, 0.293 and 0.287 across the three thresholds.

Base of the pyramid also has the most self-contained base. Of the works it cites, 77.4 per cent are cited by no other group, against 53.1 per cent for social entrepreneurship, 41.9 per cent for inclusive business and 40.8 per cent for social enterprise.

5.3 Inclusive business

With ten labelled papers, a cited-reference pool of 613 works and a concept pool of 43 terms, inclusive business cannot support a confirmatory claim, and none is made. Its comparisons are reported as exploratory on both traces. Descriptively, 0.564 of its cited works lie inside the base-of-the-pyramid pool, measured by the overlap coefficient reported in Table 4. This containment is reported as a gradient rather than as evidence of nesting.

The asymmetry also explains why the inclusive business Jaccard values are low while its overlap with base of the pyramid is high. A small pool largely contained within a much larger one produces exactly that pattern. Jaccard divides by the combined pool and so is dominated by the larger set, whereas containment divides by the smaller. The two coefficients answer different questions rather than disagreeing, which is also why the rank ordering of the inclusive business pairs changes between metrics while the ordering of the other three pairs does not.

6 Network structure

6.1 The coupling network

The coupling graph displays 296 of the 349 single-label papers with references, in two components of which the larger holds 294 nodes. Displayed nodes comprise 191 base-of-the-pyramid papers, 66 social entrepreneurship papers, 38 social enterprise papers and one inclusive business paper (Fig. 5).

Of all edges, 93.6 per cent join papers whose labels fall on the same side of the division between the social-economy labels and base of the pyramid. Weighted assortativity by that division is 0.847, and by the four labels separately 0.686.

6.2 The specified division is strongly modular against a degree-stratified null

For the partition placing social enterprise and social entrepreneurship on one side and base of the pyramid and inclusive business on the other, modularity is Q = 0.290. Under a null reassigning labels

within degree quartiles, the expected value is 0.0049 with a standard deviation of 0.0067, giving z = 42.62 and 0 of 10,000 draws at or above the observed value. Under a free reassignment the null mean is − 0.003 and z = 54.71, so the degree-stratified specification is the conservative one and is used throughout. Table 6 reports network structure.

Table 6
Network structure

Representation	Quantity	Value
Bibliographic coupling	Single-label articles with references	349
	Edges at coupling 0.05	5,272
	Nodes with at least one edge	296
	Isolates	53
	Components	2
	Within-division edges	93.6%
	Modularity, two-way division	0.290
	Null mean, degree-stratified	0.0049 (SD 0.0067)
	z against degree-stratified null	42.62
	Null mean, free reassignment	−0.003 (SD 0.0054)
	z against free null	54.71
	Modularity, four labels	0.257
	Assortativity by division	0.847
	Assortativity by four labels	0.686
	Mean degree, BoP and IB side	42.7
	Mean degree, social-economy side	22.6
Label-free communities	Communities detected	5
	Adjusted Rand index vs division	0.815
	Normalised mutual information	0.702
	Same-community pairs sharing a side	93.1%
Expanded network	Nodes including single-side multi-label	377
	Edges	6,058
	Modularity	0.334 (z = 47.86)
Co-citation	Works cited ten or more times	287
	Modularity	0.345 (z = 76.25)
	Works lacking decisive majority	12.9%

All nulls use 10,000 draws. The degree-stratified null is the primary specification. The co-citation representation derives each cited work's side from the labels of the papers citing it, so it shares the same label input and is not independent corroboration.

The degree-stratified null is primary because connectivity is unevenly distributed. Displayed base-of-the-pyramid and inclusive business papers have a mean degree of 42.7 against 22.6 for displayed social-economy papers. A free reassignment would understate the null and overstate the result.

Running label propagation, which never sees the labels, recovers a comparable division. Five communities emerge, the two largest holding 189 and 101 nodes. Agreement with the two-way division gives an adjusted Rand index of 0.815 and normalised mutual information of 0.702, and 93.1 per cent of same-community pairs share a side. This agreement is independent of the label partition during community construction. It is not independent of the classification process or of the corpus, because the network population was restricted to single-label papers with retrievable references in the first place.

A co-citation network over the 287 works cited ten or more times is modular on the same division, with Q = 0.345 and z = 76.25 (Online Resource 1, Section 3). This is a second representation sharing the same citation records rather than independent corroboration, and 12.9 per cent of works lack a decisive two-thirds majority and concentrate at the boundary. Restricting to the 250 works with a decisive majority raises Q to 0.415, which is a consequence of removing ambiguous cases and not further evidence.

6.3 The network does not identify a unique grouping

Seven two-way divisions of the four labels are possible, and all were tested against the same degree-stratified null (Fig. 6; Table 7). The division above has the highest modularity at 0.290. The division separating base of the pyramid alone from the other three follows at 0.287. Six of the seven are individually significant at p = .0001, and only the division isolating inclusive business alone is not, at Q = − 0.0002 and p = .690. That six of seven clear the threshold is itself informative: individual significance does not discriminate between candidate groupings here, which is why the two further procedures below are necessary.

Table 7
Modularity of all seven candidate two-way divisions, degree-stratified nulls

Division	Q	Null mean	z	p
SE + SEship vs BoP + IB	0.2900	0.0049	42.62	.0001
BoP vs SE + SEship + IB	0.2868	0.0044	42.35	.0001
SEship vs SE + BoP + IB	0.1761	0.0002	40.34	.0001
SE + BoP vs SEship + IB	0.1749	−0.0001	39.94	.0001
SE vs SEship + BoP + IB	0.0518	0.0012	23.53	.0001
SE + IB vs SEship + BoP	0.0498	0.0008	21.60	.0001
IB vs SE + SEship + BoP	−0.0002	−0.0001	−0.19	.690

Six of the seven divisions are individually significant, so individual significance does not discriminate between candidate groupings. The paired randomisation test of the two leading divisions gives ΔQ = 0.00314, two-sided p = .060, a difference not detected at conventional levels. Correcting for having searched seven candidates, the best observed division gives a family-wise p of .0001 against a null whose best-division mean is 0.0064.

A paired randomisation test, computing both statistics on the same reassignment so their dependence is preserved, gives an observed difference of ΔQ = 0.00314 against a paired null with mean 0.0005 and standard deviation 0.0016, two-sided p = .060. A difference between these two divisions is therefore not detected at conventional levels. No equivalence claim is made, because no principled margin exists for what difference in modularity would be substantively negligible.

The two divisions differ in exactly one respect: whether inclusive business is grouped with base of the pyramid or with the social-economy labels. Inclusive business contributes one node to the displayed network, and three of its four single-label papers are isolates at the display threshold, an isolate rate of 75 per cent against 13 to 21 per cent for the other three labels. This is a further diagnostic from the same network rather than independent evidence, and it points the same way. The network identifies a separation of base-of-the-pyramid research from social enterprise and social entrepreneurship research. It does not identify where inclusive business belongs.

Correcting for having searched seven candidates, the best observed division still lies far outside the distribution of the best division under the null, whose mean is 0.0064 and whose 97.5th percentile is 0.0213, giving a family-wise p of .0001. The best observed division therefore has substantially greater modularity than the specified degree-stratified null produces, and that result is not an artefact of having searched. What the analysis does not do is identify which division describes the structure, because the comparison between the two leading candidates does not distinguish them at conventional levels.

6.4 The result holds at paper level and under both divisions

Pooled comparisons can manufacture structure through aggregation, so the comparison was repeated between individual papers. Mean pairwise Jaccard between papers on the same side is 0.01076, against 0.00223 for papers on opposite sides. The paper-level bootstrap puts the absolute difference at 0.00853, with a 95 per cent interval of [0.00701, 0.01022] and a ratio of 4.82 [3.91, 5.95].

Because the placement of inclusive business is unresolved, the same comparison was run under the rival division. The values barely move: 0.01078 against 0.00233, an absolute difference of 0.00845 [0.00693, 0.01008] and a ratio of 4.63 [3.72, 5.88]. The contrast is not sensitive to where inclusive business is placed. Sections 6.3 and 6.4 together answer RQ4.

The ratios are reported alongside the absolute differences deliberately. Individual papers share few references with each other under either division, and a ratio near 4.8 computed on similarities of 0.011 against 0.002 describes a consistent difference rather than a large one. The median between-side similarity is 0.000 against a within-side median of 0.0044, so the contrast is driven by how many pairs share any reference at all rather than by how much the typical pair shares.

7 Robustness

Thirteen specifications are summarised in Fig. 7 and Table 8, with complete tables for all six pairs in Online Resource 1, Section 2.

Table 8
Robustness of the pooled overlap result across thirteen specifications

Specification	SE / SEship	SE / BoP	SEship / BoP
Main (300,000 draws)	.1210	.000013	.000010
Title-only labelling	.572	.0004	.0001
Reference-count-stratified null	.0577	.0001	.0001
Minimum 5 references	.157	.0001	.0001
Minimum 10 references	.116	.0001	.0001
Minimum 20 references	.211	.0001	.0001
Period 2000–2012	.104	.0001	.0001
Period 2013–2025	.184	.0006	.0095
Excluding *J Bus Ethics*	.152	.0001	.0001
Excluding *J Bus Res*	.132	.0001	.0001
Excluding *Entrep Theory Pract*	.0514	.0001	.0002
Excluding *Bus Strategy Environ*	.127	.0001	.0001
Excluding *J Soc Entrep*	.111	.0001	.0001

Empirical p, add-one convention. The main row is the definitive 300,000-draw run reported in Table 5; every other row is a 10,000-draw sensitivity run, at which the reporting floor is 1/10,001 and the Monte Carlo standard error near p = .05 is about 0.002. The cross-boundary pairs remain outside the null in every specification. The social-economy pair is never significant, coming closest under the reference-count-stratified null and when *Entrepreneurship Theory and Practice* is excluded. Sample sizes are deliberately omitted from this table because the specifications are defined on different populations: title-only labelling yields 413 labelled articles, the minimum-reference rows are defined on screened articles meeting the reference threshold, and the outlet exclusions change the screened corpus. Each specification's population, group sizes, observed values, null parameters and effect sizes for all six pairs are reported in Online Resource 1, Section 2. ---

7.1 Abstract availability

Because abstracts are missing for a quarter of the corpus and the labelling rate differs by 11.8 percentage points between records with and without them, the core comparisons were repeated under a stricter rule reading titles only. This matters because the labels do not depend on abstracts equally. Under the title-only rule base of the pyramid retains 95.3 per cent of its papers, while social entrepreneurship retains 64.5 per cent and the other two about 70 per cent. Differential coverage could in principle produce the reported separation.

It does not. Under title-only labelling, with group sizes of 74, 78, 283 and 7, the two cross-boundary pairs remain outside the null at p = .0004 and p = .0001, surviving Holm correction at .002 and .0006, while the

social-economy pair remains inside it at p = .572. The separation is not an artefact of abstract availability.

This test addresses differential sensitivity to abstracts. It does not address differential sensitivity to titles across the four labels, which remains unexamined and is listed in Section 10.

7.2 Corpus composition

Restricting the corpus to papers with at least five, ten and twenty retrievable references leaves both cross-boundary pairs at p = .0001 throughout. Reassigning labels only within reference-count quartiles, so that the null preserves how many references each paper contributes, leaves them at p = .0001 with effect sizes of 6.97 and 8.65 null standard deviations. The social-economy pair moves to p = .0577 under that null, closer to the conventional threshold than in the main specification but still not detected.

Dropping each of the five largest journals in turn leaves the cross-boundary pairs between p = .0001 and p = .0002. The social-economy pair comes closest to the threshold when *Entrepreneurship Theory and Practice* is excluded, at p = .0514.

Two cautions apply to that value and to the p = .0577 above. Both come from 10,000-draw runs, at which the Monte Carlo standard error near p = .05 is about 0.002, so neither is distinguishable from .05 on the evidence of a single run. Neither is qualitatively different from .049. Both are reported in the main text rather than a supplement for that reason, and readers who treat .05 as decisive should regard the social-economy comparison as unresolved rather than settled in either direction.

7.3 Time

Splitting the corpus into two disjoint halves and analysing each independently preserves the pattern. In 2000 to 2012, with 156 papers, both cross-boundary pairs sit at p = .0001. In 2013 to 2025, with 398 papers, they sit at p = .0006 and p = .0095. The social-economy pair is not detected in either half, at p = .104 and p = .184. These are disjoint subsets of one corpus and one classification pipeline, so they demonstrate temporal stability rather than independent replication. The weakening of the cross-boundary effect in the later period is visible and is discussed in Section 8.

7.4 Excluded papers and network thresholds

Restricting the network to single-label papers leaves out those carrying more than one label. Adding back the 28 multi-label papers whose labels fall on one side of the division gives 377 nodes and 6,058 edges, with Q = 0.334 against a null mean of 0.0046, z = 47.86. The 28 papers whose labels span both sides cannot be assigned and remain excluded, and they are the boundary-crossing cases rather than noise. Their exclusion is a limit on the network analysis and is restated in Section 10.

The 0.05 coupling threshold is one choice among many. Across thresholds of 0.02, 0.05, 0.10 and 0.15, modularity rises from 0.265 to 0.362 while displayed nodes fall from 321 to 149. The structure is

therefore not created by the threshold, though higher thresholds reach it by discarding papers, which is why the lower and more inclusive value is used for the reported result.

The 53 isolates are distributed roughly in proportion to group size, at 30 base-of-the-pyramid, 10 social enterprise, 10 social entrepreneurship and 3 inclusive business papers. They carry fewer references than connected nodes, a median of 51 against 66. This describes isolation at the display threshold only and says nothing about reference-retrieval bias or corpus representativeness, which are separate limitations. Section 7 as a whole answers RQ5.

8 Discussion

The central pattern is a convergence rather than a divergence. Two traces of scholarly organisation were measured on the same corpus and tested against the same null: the concept vocabulary by which papers are indexed, and the prior work they cite. Both place a boundary between base-of-the-pyramid research and the social-economy pair, and neither separates social enterprise from social entrepreneurship. The traces differ in the depth of that boundary, not in its location. Cross-boundary overlap falls to 44 per cent of the within-pair value on citations and to 61 per cent on vocabulary.

This was not the expected result, and the way it departs from expectation is worth stating. A design that measures vocabulary and citations separately anticipates the possibility that they disagree, which would describe communities using common language while drawing on separate evidence. The data do not show that. They show one boundary, visible twice.

Three explanations are consistent with that pattern, and the design distinguishes between them only partly. The first is that these are genuinely separate scholarly communities, which would produce agreement on both traces. The second is a common cause: OpenAlex concepts are derived from title and abstract text, and the labels are derived from the same text, so some agreement between the two traces is built into the measurement rather than discovered by it. The label-term filter in Section 5.1 removes the most direct form of this dependence and leaves the result unchanged, but it cannot remove topic vocabulary that travels with a label without naming it. The third is that both traces reflect a difference in subject matter rather than in community: papers about multinational distribution in emerging markets and papers about hybrid organisational form would differ in indexed concepts and in cited work whatever the citation community behind them.

The second and third explanations cannot be separated from the first with this design, and the paper does not claim to have done so. What would separate them is evidence that does not derive from the same title and abstract text: author overlap between the labels, editorial or disciplinary venue structure, or a hand-coded reading of how each literature defines its object. None of those is available here, and the first is impossible in this corpus because the archived records carry no affiliation or disambiguated author data.

Read through the boundary-object lens, the shared vocabulary is doing coordination work. Terms such as poverty, development and business model carry enough common meaning for researchers under different labels to recognise each other's topics, while the reference bases show that they are reading different literatures to address them (Star and Griesemer 1989). Read through the invisible college lens, the coupling network describes bounded conversations that do not coincide with the shared vocabulary (Crane 1972). Both readings are interpretations of the measurement and not findings in themselves.

Three features of the result constrain how far it should be taken.

First, the separation is a matter of degree on both traces. Cross-boundary citation Jaccard is 0.103 and 0.101 against 0.233 within the social-economy pair, and cross-boundary vocabulary Jaccard is 0.250 and 0.234 against 0.399, so these are literatures with partially overlapping reference bases and substantially overlapping vocabularies, not disconnected ones. At paper level the absolute similarities are small under either division, 0.011 against 0.002, and the median between-side pair shares nothing at all. The honest characterisation is a consistent and measurable difference in reading, not a wall.

Second, the social enterprise and social entrepreneurship comparison is unresolved rather than settled. The test does not detect a separation in any of thirteen specifications, and the closest approach is p = .0514. That stability is informative, but it cannot be converted into a claim that the two literatures are equivalent, and the paper does not make one.

Third, inclusive business cannot be placed. It is the label with the clearest policy and practitioner currency and the thinnest presence in the indexed journal literature, contributing ten labelled papers and one connected node. The two leading network divisions differ only in where it sits and are not distinguished by a paired test. This is a limit of the evidence rather than a finding about the concept, and it is plausibly a consequence of the corpus boundary itself, since much inclusive business output is practitioner and policy work that an indexed journal corpus does not reach.

The temporal split adds a qualification worth stating. The cross-boundary separation is present in both halves but weaker in 2013 to 2025 than in 2000 to 2012, with p moving from .0001 to .0006 and .0095 and cross-boundary Jaccard rising from 0.027 and 0.061 to 0.111 and 0.098. This is consistent with gradual convergence of the reference bases over time, which would be a substantively interesting development, but two periods cannot establish a trend and the later period is larger, which affects the comparison. The observation is offered as a direction for future work rather than as a finding.

9 Implications

9.1 For research

The practical implication for researchers is a search problem, though not the one the design anticipated. Because the boundary is visible on vocabulary as well as citations, a terminology-led search will not reliably bridge it either: the indexed concepts of base-of-the-pyramid work differ from those of the social-

economy literature in the same direction as their reference bases. A reviewer assembling evidence on, say, distribution models serving low-income consumers is therefore liable to stay within one side whichever strategy is used, by keyword or by citation chaining. The boundary is shallower on vocabulary, so a terminology-led search crosses it more readily than citation chaining does, but it does not dissolve it.

This suggests that cross-label searching should be deliberate rather than incidental, and that reviews organised around a single label should state that boundary rather than present themselves as covering the phenomenon. It also suggests a concrete empirical opportunity. The 33 papers in this corpus whose labels span the boundary, and the 12.9 per cent of frequently cited works without a decisive majority, are candidate sites for studying how evidence moves across the labels. They are identified here but not analysed, and they are the most direct extension of this work.

For future bibliometric work, two methodological points from this study generalise. Reporting significance for a single preferred partition is insufficient when several groupings are plausible, because six of the seven candidates here are individually significant while the two leaders are not distinguished from each other. And pooled set-overlap measures should be accompanied by a paper-level check, because pooled sets can manufacture apparent structure through aggregation.

9.2 For policy and practice

For funders, programme designers and evaluators, the implication is that label choice is not presentational. Framing an intervention as inclusive business, as a social enterprise or as a base-of-the-pyramid venture will tend to orient its designers toward a particular evidence base and toward the evaluation criteria that base takes for granted. The base-of-the-pyramid literature reaches the question through market viability and firm strategy. The social enterprise and social entrepreneurship literatures reach it through organisational form, hybridity and the management of competing demands. Both bear on whether a venture serving low-income populations will work and for whom, and a design informed by only one imports its assumptions silently.

This orientation argument is an inference from the citation evidence, not an observed effect on behaviour. What the study measures is which prior work researchers cite under each label. That designers and funders inherit those evidence bases when they adopt a label is a plausible consequence, and it is offered as such rather than as a finding.

The practical suggestion is correspondingly modest: where an intervention is framed under one of these labels, the evidence review behind it should be checked explicitly against the others, since topical similarity will not surface the gap. This paper provides no evidence about programme outcomes and makes no claim about which framing produces better results. That question requires outcome data this design does not contain.

10 Limitations

The limitations below are consequences of the design rather than incidental gaps, and several constrain the conclusions directly.

The design measures labels and citations, not concepts. Nothing here establishes that the underlying constructs are conceptually distinct, that their definitions are stable or that practitioners use them differently. A reader who wants claims about construct content will not find them supported by this evidence.

Classification recall is unmeasured. Precision was audited at 96 per cent, but on 25 records, giving a Wilson 95 per cent interval from 80.5 to 99.3 per cent, and the audit was pooled across labels. Estimating recall would require hand-coding unlabelled records, which was not done, and precision carries no information about recall. Conclusions therefore apply to the retrieved, screened, OpenAlex-matched journal corpus and not necessarily to the entire literature.

The corpus boundary excludes books, chapters, practitioner reports and policy publications. This bears unevenly on the four labels and most heavily on inclusive business, which is used extensively outside the indexed journal literature. The thin presence of that label here is partly a property of the corpus and not only of the field.

Retrieval operated under source-specific ceilings, and 125 of 739 unique records had no OpenAlex match. The archived logs preserve the sources, phrase variants, ceilings and per-source counts at each screening stage, but not the verbatim query strings or the retrieval date for each source. The screening chain and every derived statistic can therefore be reproduced from the deposited identifiers and code, while the original retrieval cannot be reproduced exactly. This is stated rather than reconstructed, because a reconstructed query string could not be verified against what was run. Those without matches skew recent and social-enterprise-leaning. Among labelled papers, 56 of 461 lack retrievable references, and these skew later, toward base of the pyramid and toward Crossref-sourced records.

The title-only replication addresses differential sensitivity to abstract availability. It does not address differential sensitivity to titles across labels, which is not tested here.

The network analysis excludes the 28 cross-boundary multi-label papers with references, which are precisely the boundary-spanning cases. The co-citation representation derives the domain of a cited work from the labels of the papers citing it, so it shares the same label input and is not independent corroboration.

No affiliation data exists in the archived records, so no analysis of countries, institutions, authors or collaboration is possible. Author-keyword analysis was also not possible, as that field was unpopulated.

Received citation counts are global and reflect all citing literature, not only this corpus, so they measure attention from the wider field rather than within-corpus influence.

Finally, the social enterprise and social entrepreneurship comparison and the placement of inclusive business are both unresolved. They are reported as open rather than closed, and a reader should not treat either as settled by this study.

11 Conclusions

This study compared four labels used in research on market-based approaches to poverty across two traces of scholarly organisation, applied to the same corpus of 554 journal articles published between 2000 and 2025 and tested against the same permutation null. The indexed concept vocabulary and the cited reference base place a boundary in the same position. Base-of-the-pyramid research is separated from social enterprise and social entrepreneurship research on both, and in both cases the separation survives correction across the six-pair family. Social enterprise and social entrepreneurship are not separated on either trace. A bibliographic coupling network is substantially modular under the same separation, a result that holds after correcting for having searched seven candidate groupings.

The contribution is the convergence of two traces on one boundary, together with a difference in its depth: overlap across the boundary falls to 44 per cent of the within-pair value on cited references and to 61 per cent on vocabulary. The study set out to test whether a shared vocabulary could coexist with divergent citation bases, and reports that it does not in this corpus, which is a correction to the expectation rather than a confirmation of it.

Three boundaries on the result are reported rather than minimised. The analysis does not detect a separation between social enterprise and social entrepreneurship and does not establish their equivalence. It does not identify where inclusive business belongs, because that label contributes one connected node and its vocabulary and reference pools are too small for comparable inference. And because OpenAlex concepts and the labels are both derived from title and abstract text, part of the agreement between the two traces may be a property of the measurement rather than of the field, which the label-term filter constrains but cannot eliminate.

The most direct extension is to test the boundary against evidence that does not derive from the same text, such as author overlap or venue structure, and to examine the boundary-spanning cases this corpus identifies but does not analyse.

Declarations

Author Contribution

F.Z.S. conceived and designed the study, carried out the literature retrieval, screening and data curation, developed the label-assignment procedure and analysis code, performed the statistical and network analyses, prepared all figures and tables, interpreted the results, and wrote the main manuscript text. F.Z.S. reviewed and approved the final manuscript.

Data Availability

All data supporting the findings are derived from publicly available bibliographic metadata retrieved from Crossref, Scopus and OpenAlex. The screened corpus identifiers, the derived concept and reference sets, and the complete analysis output in JSON form are provided in Online Resource 2. Redistribution of full third-party records is restricted by the source providers' terms, so the deposit contains record identifiers and derived measures sufficient to reproduce every reported statistic rather than the raw vendor records.

References

1. Altman DG, Bland JM (1995) Statistics notes: absence of evidence is not evidence of absence. BMJ 311(7003):485. https://doi.org/10.1136/bmj.311.7003.485
2. Aria M, Cuccurullo C (2017) bibliometrix: an R-tool for comprehensive science mapping analysis. J Informetr 11(4):959–975
3. Austin J, Stevenson H, Wei-Skillern J (2006) Social and commercial entrepreneurship: same, different, or both? Entrep Theory Pract 30(1):1–22. https://doi.org/10.1111/j.1540-6520.2006.00107.x
4. Battilana J, Lee M (2014) Advancing research on hybrid organizing: insights from the study of social enterprises. Acad Manag Ann 8(1):397–441. https://doi.org/10.5465/19416520.2014.893615
5. Boyack KW, Klavans R (2010) Co-citation analysis, bibliographic coupling, and direct citation: which citation approach represents the research front most accurately? J Am Soc Inf Sci Technol 61(12):2389–2404
6. Bradford SC (1934) Sources of information on specific subjects. Engineering 137:85–86
7. Chliova M, Mair J, Vernis A (2020) Persistent category ambiguity: the case of social entrepreneurship. Organ Stud 41(7):1019–1042
8. Crane D (1972) Invisible colleges: diffusion of knowledge in scientific communities. University of Chicago Press, Chicago
9. Dacin PA, Dacin MT, Matear M (2010) Social entrepreneurship: why we don't need a new theory and how we move forward from here. Acad Manag Perspect 24(3):37–57. https://doi.org/10.5465/amp.2010.52842950
10. Danon L, Díaz-Guilera A, Duch J, Arenas A (2005) Comparing community structure identification. J Stat Mech Theory Exp 2005(09):P09008. https://doi.org/10.1088/1742-5468/2005/09/P09008
11. Defourny J, Nyssens M (2010) Conceptions of social enterprise and social entrepreneurship in Europe and the United States: convergences and divergences. J Soc Entrep 1(1):32–53
12. Dembek K, Sivasubramaniam N, Chmielewski DA (2020) A systematic review of the bottom/base of the pyramid literature: cumulative evidence and future directions. J Bus Ethics 165(3):365–382
13. Doherty B, Haugh H, Lyon F (2014) Social enterprises as hybrid organizations: a review and research agenda. Int J Manag Rev 16(4):417–436

14. Donthu N, Kumar S, Mukherjee D, Pandey N, Lim WM (2021) How to conduct a bibliometric analysis: an overview and guidelines. J Bus Res 133:285–296
15. Efron B, Tibshirani RJ (1993) An introduction to the bootstrap. Chapman and Hall, New York
16. George G, McGahan AM, Prabhu J (2012) Innovation for inclusive growth: towards a theoretical framework and a research agenda. J Manag Stud 49(4):661–683
17. Good PI (2005) Permutation, parametric, and bootstrap tests of hypotheses, 3rd edn. Springer, New York
18. Greenland S, Senn SJ, Rothman KJ, Carlin JB, Poole C, Goodman SN, Altman DG (2016) Statistical tests, P values, confidence intervals, and power: a guide to misinterpretations. Eur J Epidemiol 31(4):337–350
19. Hahn R (2012) Inclusive business, human rights and the dignity of the poor: a glance beyond economic impacts of adapted business models. Bus Ethics Eur Rev 21(1):47–63
20. Hart SL, Christensen CM (2002) The great leap: driving innovation from the base of the pyramid. MIT Sloan Manag Rev 44(1):51–56
21. Hirsch JE (2005) An index to quantify an individual's scientific research output. Proc Natl Acad Sci USA 102(46):16569–16572
22. Holm S (1979) A simple sequentially rejective multiple test procedure. Scand J Stat 6(2):65–70
23. Hubert L, Arabie P (1985) Comparing partitions. J Classif 2(1):193–218
24. Karnani A (2007) The mirage of marketing to the bottom of the pyramid: how the private sector can help alleviate poverty. Calif Manag Rev 49(4):90–111
25. Kessler MM (1963) Bibliographic coupling between scientific papers. Am Doc 14(1):10–25. https://doi.org/10.1002/asi.5090140103
26. Kolk A, Rivera-Santos M, Rufín C (2014) Reviewing a decade of research on the base/bottom of the pyramid (BOP) concept. Bus Soc 53(3):338–377
27. Lakens D (2017) Equivalence tests: a practical primer for t tests, correlations, and meta-analyses. Soc Psychol Personal Sci 8(4):355–362
28. Likoko E, Kini J (2017) Inclusive business: a business approach to development. Curr Opin Environ Sustain 24:84–88
29. London T, Hart SL (2004) Reinventing strategies for emerging markets: beyond the transnational model. J Int Bus Stud 35(5):350–370
30. Mair J, Marti I (2006) Social entrepreneurship research: a source of explanation, prediction, and delight. J World Bus 41(1):36–44
31. Newman MEJ (2003) Mixing patterns in networks. Phys Rev E 67(2):026126
32. Newman MEJ (2006) Modularity and community structure in networks. Proc Natl Acad Sci USA 103(23):8577–8582
33. Newman MEJ, Girvan M (2004) Finding and evaluating community structure in networks. Phys Rev E 69(2):026113

34. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD et al (2021) The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 372:n71
35. Phipson B, Smyth GK (2010) Permutation P-values should never be zero: calculating exact P-values when permutations are randomly drawn. Stat Appl Genet Mol Biol 9(1):39
36. Prahalad CK (2004) The fortune at the bottom of the pyramid: eradicating poverty through profits. Wharton School Publishing, Upper Saddle River
37. Prahalad CK, Hart SL (2002) The fortune at the bottom of the pyramid. Strategy Bus 26:1–14
38. Priem J, Piwowar H, Orr R (2022) OpenAlex: a fully-open index of scholarly works, authors, venues, institutions, and concepts. arXiv:2205.01833
39. Raghavan UN, Albert R, Kumara S (2007) Near linear time algorithm to detect community structures in large-scale networks. Phys Rev E 76(3):036106
40. Saebi T, Foss NJ, Linder S (2019) Social entrepreneurship research: past achievements and future promises. J Manag 45(1):70–95
41. Salton G, McGill MJ (1983) Introduction to modern information retrieval. McGraw-Hill, New York
42. Santos FM (2012) A positive theory of social entrepreneurship. J Bus Ethics 111(3):335–351
43. Seelos C, Mair J (2007) Profitable business models and market creation in the context of deep poverty: a strategic view. Acad Manag Perspect 21(4):49–63
44. Short JC, Moss TW, Lumpkin GT (2009) Research in social entrepreneurship: past contributions and future opportunities. Strateg Entrep J 3(2):161–194
45. Small H (1973) Co-citation in the scientific literature: a new measure of the relationship between two documents. J Am Soc Inf Sci 24(4):265–269
46. Star SL, Griesemer JR (1989) Institutional ecology, translations and boundary objects: amateurs and professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39. Soc Stud Sci 19(3):387–420
47. Suddaby R (2010) Construct clarity in theories of management and organization. Acad Manag Rev 35(3):346–357
48. van Eck NJ, Waltman L (2010) Software survey: VOSviewer, a computer program for bibliometric mapping. Scientometrics 84(2):523–538
49. Zahra SA, Gedajlovic E, Neubaum DO, Shulman JM (2009) A typology of social entrepreneurs: motives, search processes and ethical challenges. J Bus Ventur 24(5):519–532
50. Zupic I, Čater T (2015) Bibliometric methods in management and organization. Organ Res Methods 18(3):429–472

Figures

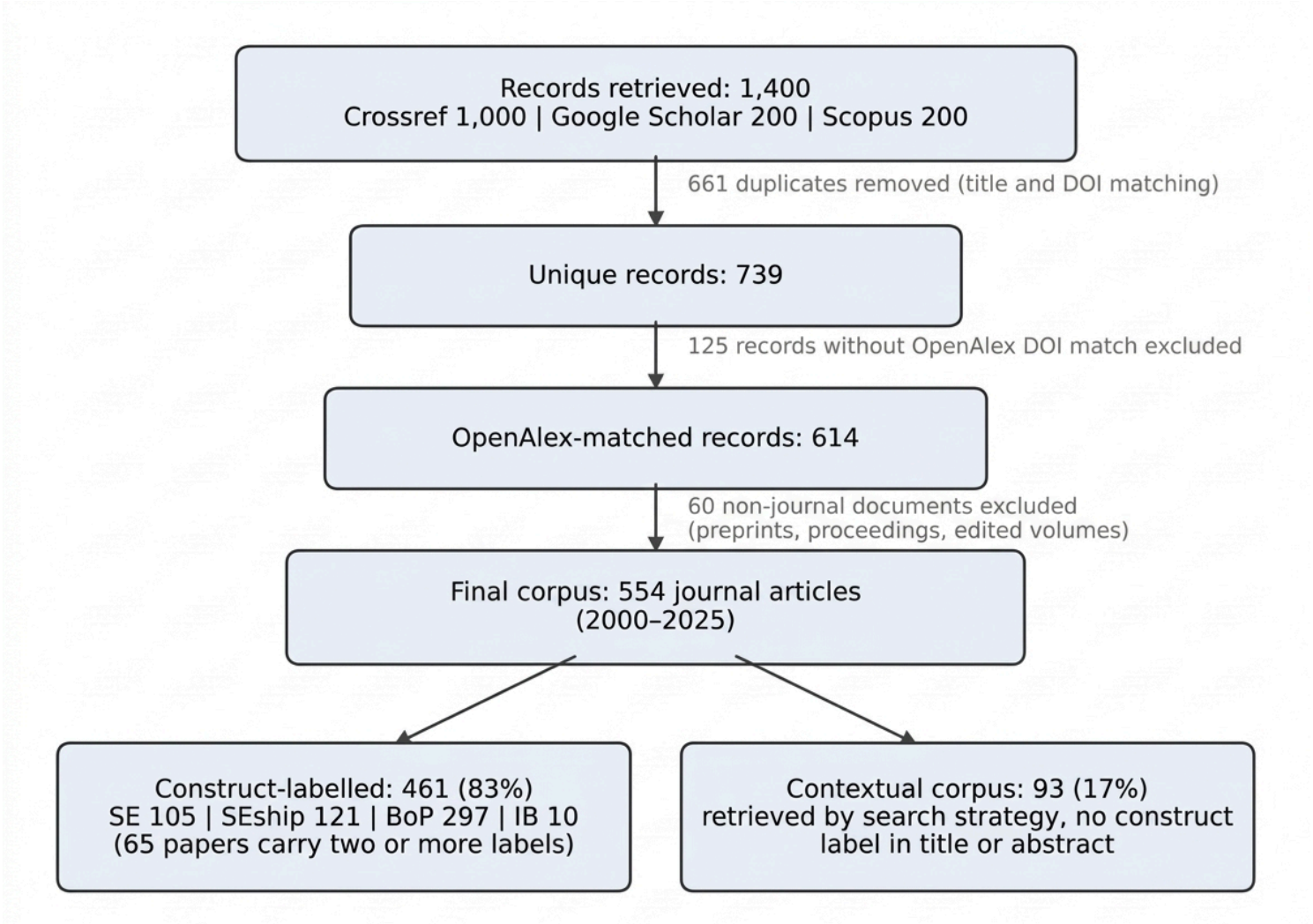


Figure 1

Identification, screening and classification of the corpus, from 1,400 retrieved records to 554 screened journal articles, of which 461 carry at least one label and 93 carry none

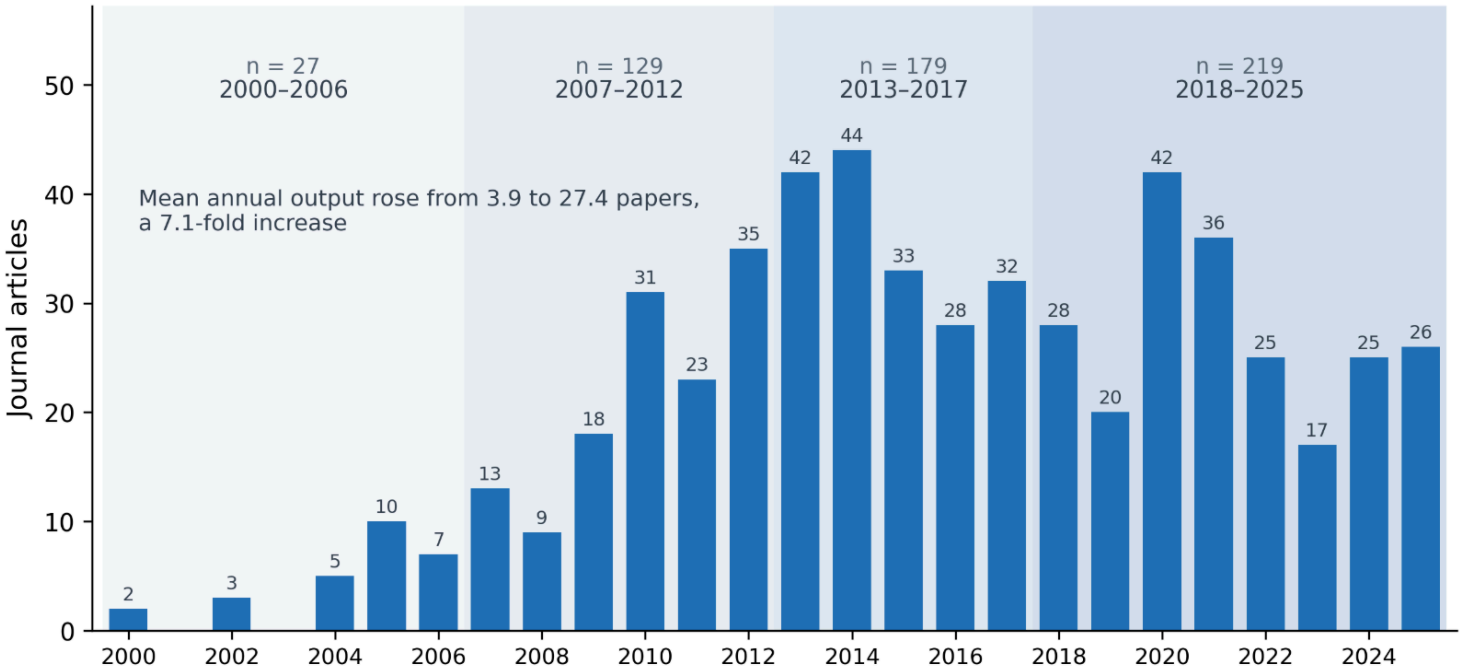


Figure 2

Annual output of labelled and contextual articles, 2000 to 2025, with the four periods shaded and their counts given as 27, 129, 179 and 219

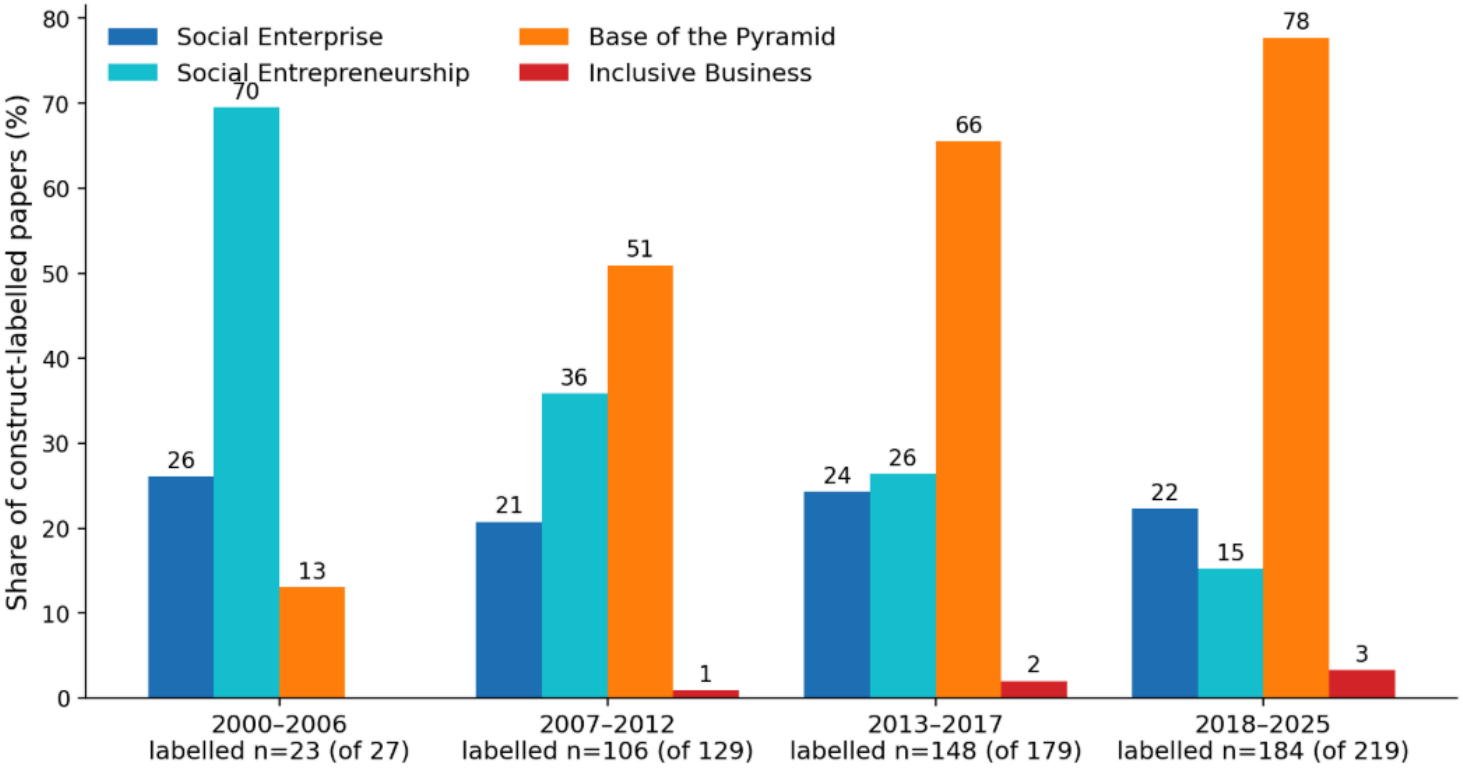


Figure 3

Share of labelled articles carried by each of the four labels across the four periods

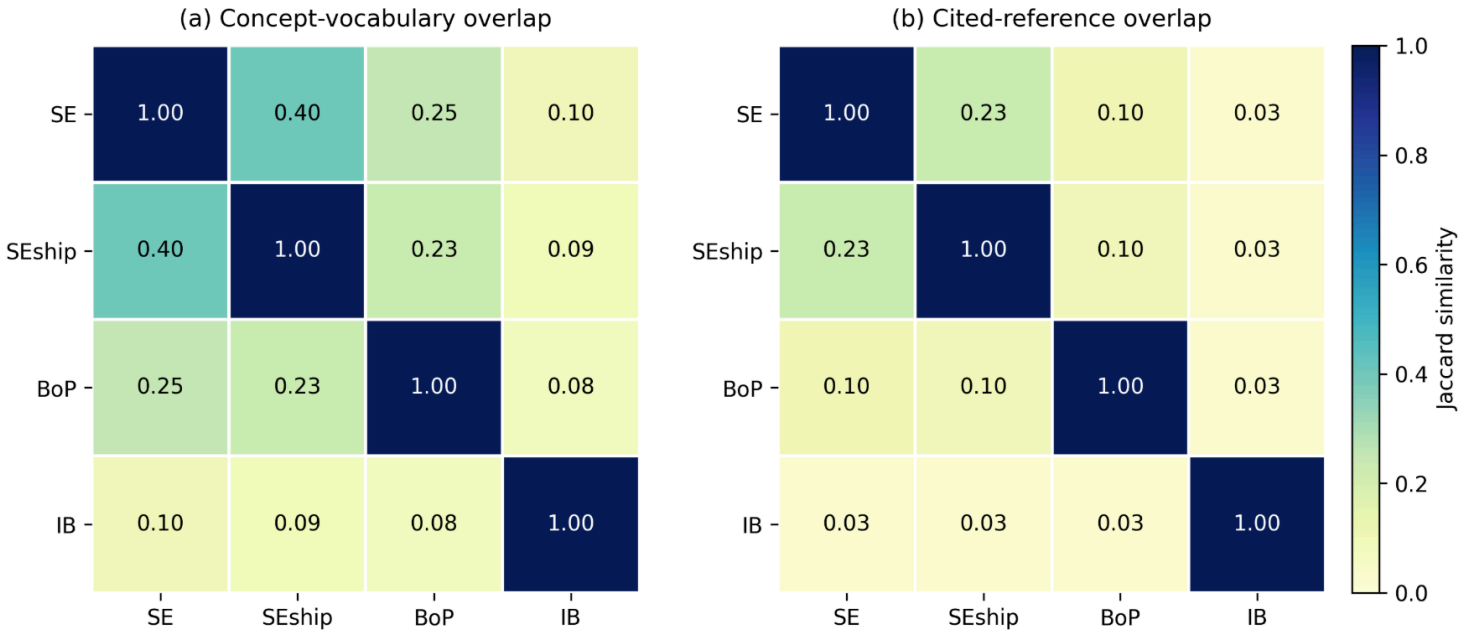


Figure 4

Pairwise overlap between the four labels, measured by the Jaccard coefficient, for **a** attached conceptual vocabulary and **b** pooled cited references

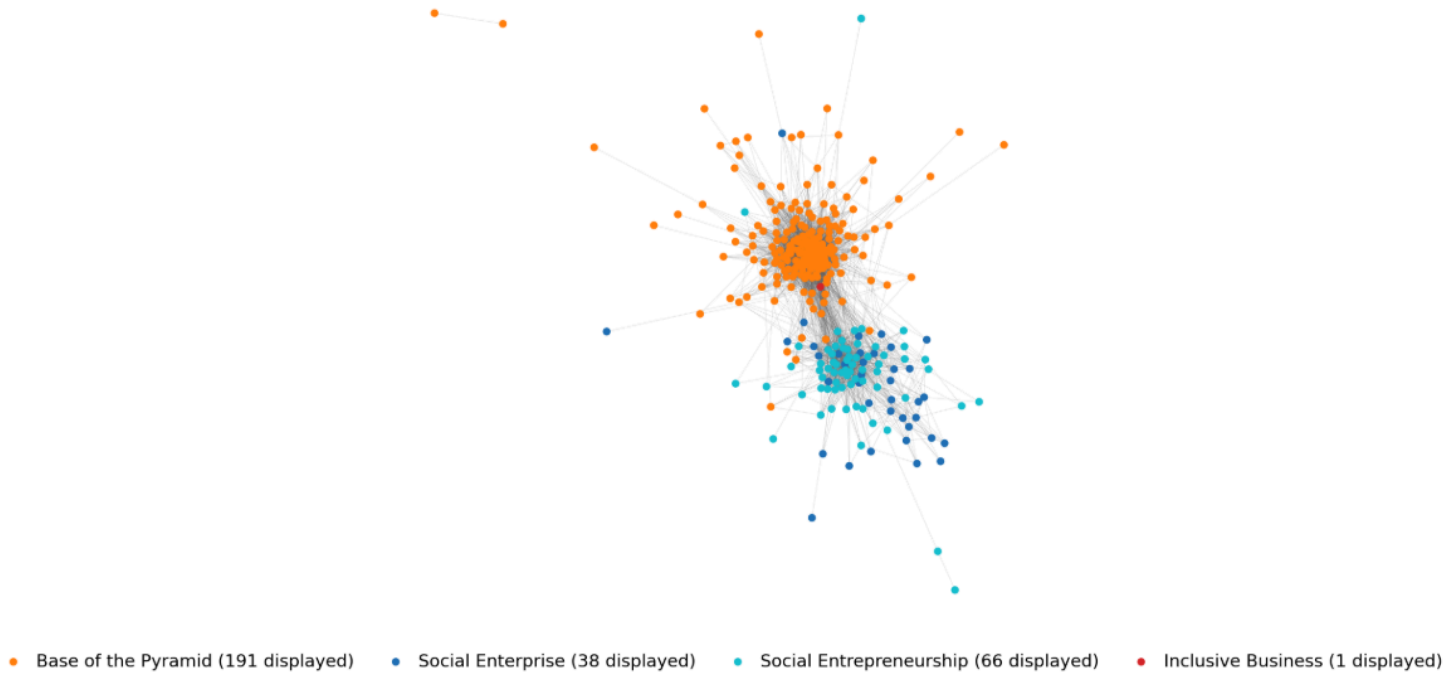


Figure 5

Bibliographic coupling network over single-label articles with retrievable references, showing the 296 nodes with at least one edge at a Salton-normalised coupling of 0.05, with node colour and marker shape indicating label

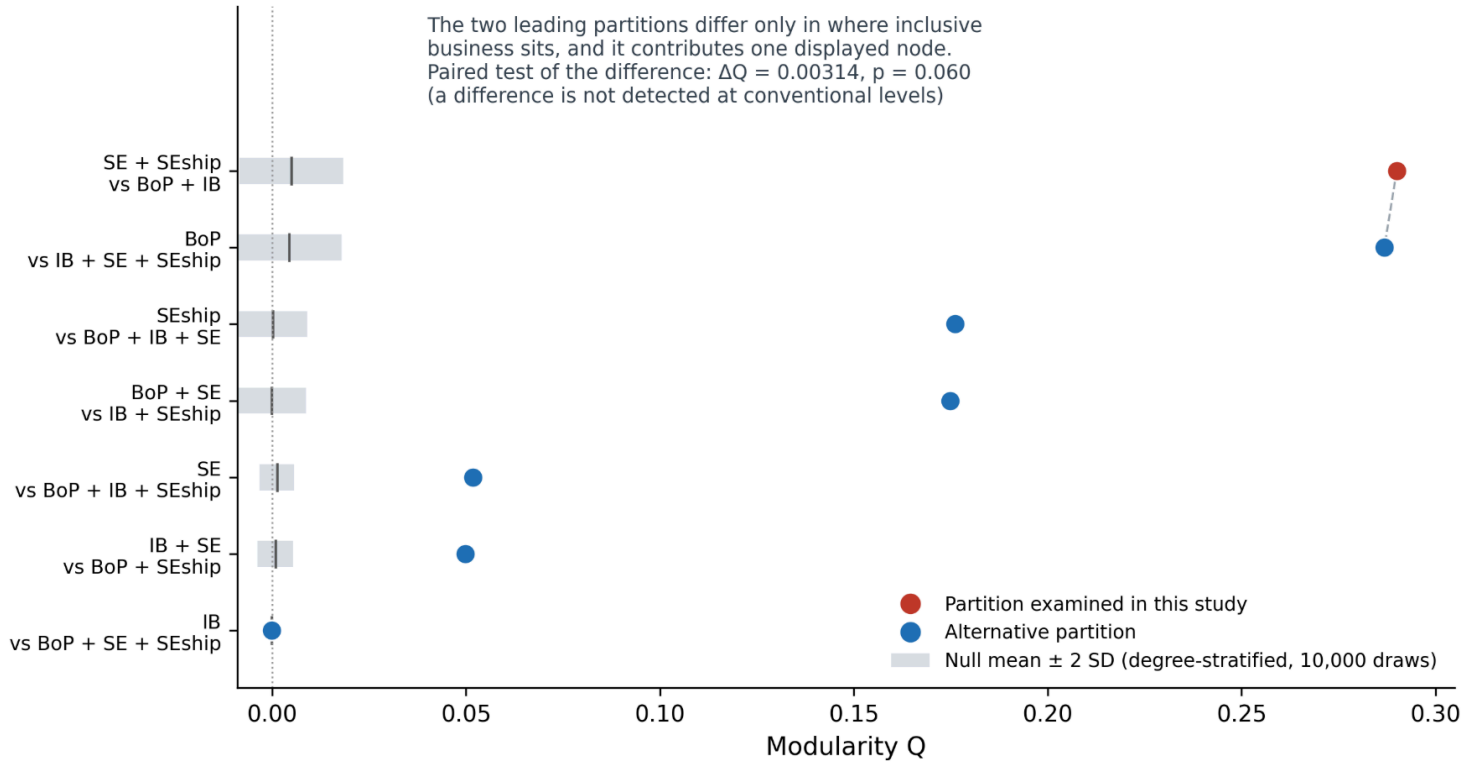


Figure 6

Modularity of all seven candidate two-way divisions of the four labels, each against its degree-stratified null band of mean plus or minus two standard deviations, with the two leading divisions joined and annotated with the paired randomisation test

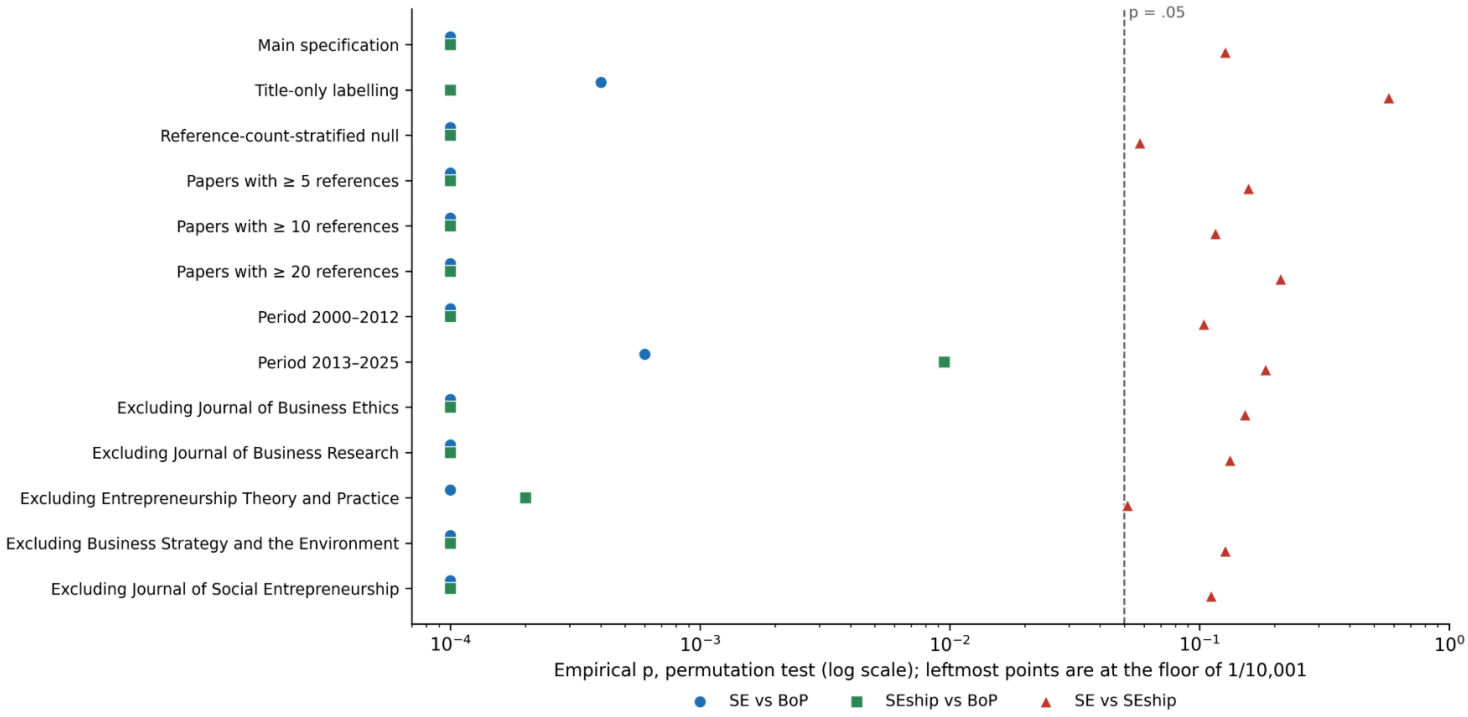


Figure 7

Empirical p for the three focal label pairs across thirteen specifications, on a logarithmic scale, with the conventional 0.05 threshold marked ---

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- ESM1.docx
- ESM2.zip
- AppendixA.docx