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.

A Multi-LLM Consensus Weighting (MCDM) Approach to Redefining Microfinance Through Blockchain: Enriching Social Capital Outreach in Malaysia and the Maldives

Zahari Md Rodzi
Universiti Teknologi MARA (UiTM) Negeri Sembilan Branch
Nur Aima Shafie
aimashafie@uitm.edu.my

Universiti Teknologi MARA
Zuraidah Md Sanusi
Universiti Teknologi MARA
Aizatul Waznah Ghazali
Universiti Kebangsaan Malaysia (UKM)
Mariyam Niyaf Mohamed
Zikura International College

Research Article

Keywords: microfinance outreach, blockchain, Islamic microfinance, financial inclusion, social capital, multi-criteria decision making

Posted Date: July 16th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10234134/v1

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

Additional Declarations: No competing interests reported.

A Multi-LLM Consensus Weighting (MCDM) Approach to Redefining Microfinance Through Blockchain: Enriching Social Capital Outreach in Malaysia and the Maldives

Zahari Md Rodzi[1,2], Nur Aima Shafie[3*], Zuraidah Md Sanusi[4], Aizatul Waznah Ghazali[5], Mariyam Niyaf Mohamed[6]

[1]Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA (UiTM) Negeri Sembilan Branch, Seremban, Malaysia.
[2]Faculty of Science and Technology, Universiti Kebangsaan Malaysia (UKM), Bangi, Selangor, Malaysia.
[3]Faculty of Accountancy, Universiti Teknologi MARA (UiTM), Malaysia.
[4]Accounting Research Institute, Universiti Teknologi MARA (UiTM), Malaysia.
[5]Faculty of Economics and Management, Universiti Kebangsaan Malaysia (UKM), Bangi, Selangor, Malaysia.
[6]Research and Innovation Department, Zikura International College, Malé, Maldives.

*Corresponding author(s). E-mail(s): nur_aima@uitm.edu.my;

Abstract

Microfinance is valued, above all, for its *outreach*, its capacity to extend financial services to households and microenterprises that formal finance leaves behind, and that outreach rests on the *social capital* (trust, transparency, financial inclusion, connectivity) of the communities it serves. Blockchain is increasingly advanced as a means of redefining this outreach: tamper-evident ledgers, programmable shariah-compliant contracts, and low-cost digital onboarding can enrich the trust and accountability on which microfinance depends. This study measures the social-capital conditions for blockchain-enabled microfinance outreach in *Malaysia and the Maldives*, two collaborating economies at contrasting stages of Islamic-finance and digital-infrastructure development, and benchmarks them within a panel of Organisation of Islamic Cooperation (OIC) economies

so that the comparison is statistically meaningful. Measurement uses a *parsimonious indicator framework* (indicators with complete international coverage, high reliability, cross-country comparability, and full reproducibility) weighted by a supporting methodological innovation, *Multi-LLM Consensus Weighting* (MLCW), which aggregates the judgements of several large language models through reliability weighting and a disagreement penalty and fuses them with objective entropy weights. On real, fully traceable secondary data, Malaysia ranks 2nd and the Maldives 9th among 53 OIC economies; the ranking is robust to bootstrap resampling, indicator addition (Spearman **0.95**), and the choice of weighting method (Equal, Entropy, CRITIC, and MEREC agree, pairwise Spearman $\approx$ **1.0**). The study contributes an auditable, comparative measure of the social-capital outreach conditions for blockchain-enabled microfinance and derives concrete policy implications for transparency, financial inclusion, digital readiness, and governance in Malaysia, the Maldives, and the wider OIC.

Keywords: microfinance outreach, blockchain, Islamic microfinance, financial inclusion, social capital, multi-criteria decision making

JEL Classification: G21 , O16 , C43 , C44 , O53

1 Introduction

Microfinance is judged, above all, by its *outreach*: the breadth and depth with which it extends financial services to households and microenterprises that formal finance leaves behind. In Malaysia, institutions such as Amanah Ikhtiar Malaysia (AIM), the country's flagship microcredit provider, modelled on the Grameen approach, have for decades pursued this mission, and Islamic microfinance has become a central instrument of inclusive development across the Muslim world. Crucially, outreach is not a purely financial phenomenon: it rests on *social capital*, the trust, reciprocity, transparency, and participation that allow group lending, peer monitoring, and reputational collateral to substitute for physical collateral [1–4]. In Islamic microfinance this dependence is sharpened by instruments built on risk-sharing and solidarity (*qard al-hasan*, *murabahah*, *mudarabah*) and by the ethical prohibition of *riba* [6]. The social-capital base of a society therefore conditions how far, and how deeply, microfinance can reach.

Blockchain has emerged as a technology that could redefine this outreach. Tamper-evident records of contributions and repayments can reinforce reputational collateral; programmable, shariah-compliant smart contracts can lower agency and monitoring costs; transparent, auditable fund flows bear directly on transparency and accountability; and low-cost digital onboarding broadens reach to remote and underserved populations [7–13]. The promise, in short, is that blockchain can *enrich the social-capital conditions* for microfinance outreach. Yet this promise is largely asserted rather than measured: there is no auditable, comparable way to assess the social-capital conditions for blockchain-enabled microfinance outreach across countries, and recent multi-criteria work on evaluating blockchain-enabled microfinance [39] has focused on

ranking alternatives rather than on the prior, and equally decisive, problem of how the underlying criteria should be *weighted*.

That weighting problem is the crux. A credible composite measure of social capital requires criterion weights, and the weighting step is the most contested in index construction [14, 16, 17]: equal weighting embeds untested assumptions, subjective methods (Delphi, AHP, BWM) are costly and non-reproducible [18, 19], and objective methods (entropy, CRITIC, MEREC) are meaning-blind [20, 21, 26, 40]. A two-country comparison, moreover, cannot on its own support data-driven weighting or normalisation. This study addresses both the substantive and the methodological problem for *Malaysia and the Maldives*, two collaborating OIC economies at contrasting stages of Islamic-finance and digital-infrastructure development, and the natural focal cases for a bilateral research agenda on blockchain-enabled microfinance. We (i) operationalise the social-capital conditions for microfinance outreach through a deliberately small *parsimonious indicator framework* retained for complete coverage, reliability, comparability, and reproducibility; (ii) weight it with a supporting methodological innovation, *Multi-LLM Consensus Weighting* (MLCW), which treats several large language models as a reproducible panel of semantic assessors [35–37]; and (iii) *benchmark* the two focal economies within a panel of 53 OIC economies, so that their standing is read against meaningful peers and the data-driven weighting is statistically valid.

The contribution is therefore primarily applied: an auditable, comparative measurement of the social-capital outreach conditions for blockchain-enabled microfinance in Malaysia and the Maldives, benchmarked across the OIC, with substantive policy implications, supported by a methodological enhancement (MLCW) that makes the measurement simultaneously meaning-aware and reproducible. In doing so the study extends an ongoing programme of multi-criteria research on blockchain-enabled microfinance and objective weighting [39, 40, 42] from alternative *ranking* toward criterion *weighting* for composite social indicators. The remainder of the paper reviews the relevant literatures (Section 2), develops the methodology (Section 3), reports the outreach standing of the two focal economies within their OIC benchmark together with method comparison, robustness, and policy implications (Section 4), and concludes (Section 5).

2 Literature Review

2.1 Microfinance outreach, social capital, and blockchain

Social capital is conventionally decomposed into structural elements (networks, participation, connectivity) and cognitive elements (trust, norms, reciprocity), following the foundational accounts of Bourdieu, Coleman, and Putnam [1–3], with a bridging-versus-bonding distinction emphasised in development policy [4, 5]. In microfinance the construct is constitutive rather than peripheral: group lending relies on peer selection, peer monitoring, and joint liability, so the social-capital endowment of a community shapes both the breadth of access and the depth of repayment that define outreach. Islamic microfinance sharpens this dependence by foregrounding risk-sharing, solidarity, and the prohibition of *riba* [6], and the fintech and digital-finance literatures document how technology reshapes inclusion and its distribution [7–10]. A growing

strand argues that blockchain (through immutable records, smart-contract shariah governance, transparent fund flows, and cheaper onboarding) can strengthen the trust, transparency, accountability, and inclusion dimensions of social capital [11–13], and recent multi-criteria research has begun to evaluate blockchain-enabled microfinance directly [39]. This strand remains, however, predominantly conceptual or alternative-ranking in focus; the *social-capital conditions* that determine whether such outreach can succeed are seldom measured comparably across countries.

The notion of *outreach* itself has two faces that a social-capital lens helps to reconcile: *breadth* (how many clients, and how far into underserved and remote populations, a programme reaches) and *depth* (how poor those clients are and how durable the relationship becomes). Group-based Islamic microfinance mobilises bonding social capital (within-group trust and monitoring) for breadth and bridging social capital (links to institutions and markets) for depth. The four blockchain mechanisms above map onto these faces directly: immutable repayment histories build the reputational capital that deepens relationships, while low-cost onboarding and interoperable rails extend breadth. A measurement framework for the social-capital conditions of outreach must therefore be multidimensional and comparable across very different national contexts, precisely the requirement that motivates a carefully weighted composite rather than a single headline indicator. The microfinance literature has long distinguished these aspects of outreach and warned that breadth and depth can trade off against one another, so that a programme may serve many clients shallowly or few clients deeply; a credible measure must therefore avoid collapsing the construct onto a single axis. Equally, the social-capital tradition cautions that trust and participation are latent and only partially observable, which is why proxies (and the weights attached to them) carry so much of the measurement burden and must be chosen and justified transparently.

2.2 Weighting methods, composite indicators, and the research gap

Composite indicators are the standard instrument for summarising latent, multidimensional constructs such as social capital for cross-country comparison [14–16]; their construction embeds choices at every step, and weighting is both the most consequential and the most contested, which is why uncertainty and sensitivity analysis are regarded as essential [17]. Weighting methods divide into subjective and objective families. Subjective weighting (Delphi, AHP [18], BWM [19]) elicits importance from human experts and thus encodes domain meaning, but is costly, slow, sensitive to anchoring, and difficult to reproduce. Objective weighting (entropy [25, 26], CRITIC [20], MEREC [21]) derives importance from statistical structure and is reproducible but meaning-blind; recent extensions embed these ideas in fuzzy and neutrosophic settings to handle uncertainty in expert judgement, including Pythagorean-neutrosophic MEREC for objective weighting [40], generalised neutrosophic aggregation operators [41, 45], neutrosophic DEMATEL for criterion interdependence [44], and adaptive utility-ranking with reproducible decision-support tooling [42]. Broad reviews map the proliferation of these techniques [22–24]. In parallel, a fast-moving literature shows that large language models can serve as structured evaluators [35, 38], that single

models are biased judges [37], and that aggregating a panel of diverse models improves reliability [36], a development not yet connected to composite-indicator weighting.

It is worth being concrete about how these methods differ, because that difference is the crux of our contribution. AHP infers weights from the principal eigenvector of a pairwise-comparison matrix and checks a consistency ratio [18]; BWM reduces the elicitation to best-to-others and others-to-worst vectors [19]; both require a human panel. Entropy assigns larger weight to indicators whose values are more dispersed across units [25, 26]; CRITIC augments dispersion with inter-criterion correlation [20]; MEREC infers weight from the change in aggregate performance when each criterion is removed [21]; all three are data-driven and meaning-blind. To handle the vagueness and hesitation inherent in expert judgement, a substantial recent literature embeds these ideas in Pythagorean-neutrosophic settings, PNMEREC for objective weighting [40], generalised Hamacher–Heronian and orthopair aggregation operators [41, 45], neutrosophic DEMATEL for criterion interdependence [44], and adaptive utility-ranking with reproducible decision-support tooling [39, 42, 43]. MLCW is complementary: rather than refine the algebra of a single expert's judgement, it makes the *semantic* half of weighting reproducible by replacing the human committee with a panel of language models, while retaining an objective entropy anchor.

Four gaps follow. *First*, most social-capital and development indices rely on equal or conventional ad hoc weights, embedding untested assumptions about relative importance. *Second*, the principled alternatives are polarised between subjective methods that are costly and non-reproducible and objective methods that discard domain meaning. *Third*, despite rapid progress in LLM-based judgement, almost no work integrates AI-based *semantic* assessment into social-capital measurement, where the meaning of indicators is central. *Fourth*, and most specifically, no known framework combines multiple LLMs, consensus weighting, disagreement attenuation, and entropy-based adjustment to assess the social-capital conditions for blockchain-enabled microfinance outreach in OIC economies. This study is positioned to close these gaps: it is, at its core, an applied measurement of social-capital outreach for blockchain-enabled microfinance in Malaysia and the Maldives, benchmarked across the OIC, with MLCW as the supporting methodological enhancement that makes the measurement simultaneously meaning-aware and reproducible.

Table 1 Positioning of MLCW relative to established weighting methods.

Method	Source	Repro.	Meaning	Disagreement	Tunable
AHP	subjective	—	✓	consistency only	—
BWM	subjective	—	✓	—	—
Entropy	objective	✓	—	n/a	—
CRITIC	objective	✓	—	n/a	—
MEREC	objective	✓	—	n/a	—
MLCW	semantic + objective	✓	✓	✓ (penalty)	✓ (α, λ)

3 Methodology

3.1 Measuring social-capital outreach: a parsimonious indicator framework

We treat the social-capital conditions for microfinance outreach as a latent, multi-dimensional construct and measure it through a *parsimonious indicator framework*. Rather than maximise the number of dimensions, we retain only indicators that satisfy four criteria jointly: *complete international coverage* of the unit set, *data reliability* (collection by a recognised international body), *cross-country comparability* of definition and source, and *reproducibility* (public, versioned, citable with an indicator code). This is a principled design choice, not a data limitation: enlarging the set with partially observed or non-comparable indicators would force imputation and silently inject assumptions, undermining the auditability the index is meant to provide. The retained indicators are, moreover, exactly those most relevant to blockchain-enabled Islamic microfinance: *transparency* (the dimension most directly affected by auditable, tamper-evident ledgers) and *financial inclusion* (the dimension the technology most directly broadens), with *network connectivity* added as a digital-readiness indicator for robustness. The choice of these three dimensions is theoretically motivated, not merely pragmatic. *Transparency* proxies the cognitive social capital, institutional trust and accountability, without which neither peer monitoring nor shariah governance can function, and it is the dimension that auditable, tamper-evident ledgers most directly strengthen. *Financial inclusion* proxies the structural reach of the financial system and is the most direct national-level correlate of microfinance *breadth*. *Digital readiness* proxies the connective infrastructure that conditions whether blockchain-based delivery is feasible at all. Other dimensions of social capital (participation, generalised trust, associational density) are conceptually relevant but lack complete, comparable, and reproducible OIC-wide series, and are therefore excluded by the parsimony criteria rather than by assumption, an exclusion revisited as a limitation in Section 4.5. Figure 1 locates the framework conceptually: blockchain-enabled Islamic microfinance acts on social-capital dimensions, which a weighting method aggregates into a Social Capital Index that, in turn, yields policy insight.



Fig. 1 Conceptual framework of the study. Blockchain-enabled microfinance acts on the social-capital dimensions that underpin outreach; these are aggregated by the MLCW weighting method into a Social Capital Outreach Index for Malaysia and the Maldives (benchmarked across the OIC), which informs policy.

3.2 Multi-LLM Consensus Weighting (MLCW)

MLCW supplies the weights that turn the indicator framework into an index. It couples a *semantic* branch (a panel of M large language models that judge indicator importance) with a *statistical* branch, entropy, which measures how strongly each indicator discriminates among units, and reconciles them through a single fusion parameter. Figure 2 shows the workflow.



Fig. 2 Methodological workflow of the MLCW-based index.

Formulation.

Let there be N indicators ($i = 1, \dots, N$) and M assessor models ($j = 1, \dots, M$); each model returns a weight vector $w^{(j)} \in \Delta^{N-1}$ on the simplex $\Delta^{N-1} = \{x \in \mathbb{R}^N : x_i \geq 0, \sum_i x_i = 1\}$, collected in $W = [w_i^{(j)}]$. With reliability weights $r \in \Delta^{M-1}$, the reliability-weighted consensus and its dispersion are

$$\bar{w}_i = \sum_{j=1}^{M} r_j w_i^{(j)}, \qquad \sigma_i = \Big(\sum_{j=1}^{M} r_j (w_i^{(j)} - \bar{w}_i)^2 \Big)^{1/2}. \tag{1}$$

A disagreement penalty discounts indicators on which the panel disagrees, and the result is renormalised to the *semantic* weights:

$$\widetilde{w}_i = \bar{w}_i e^{-\lambda \sigma_i}, \qquad w_i^{\text{sem}} = \widetilde{w}_i / \textstyle\sum_k \widetilde{w}_k, \tag{2}$$

with $\lambda \geq 0$ the penalty strength. From the normalised data matrix $X = [x_{ik}] \in [0,1]^{N \times K}$ over K units, the objective *entropy* weights are

$$E_i = -\frac{1}{\ln K} \sum_{k=1}^{K} p_{ik} \ln p_{ik}, \quad p_{ik} = \frac{x_{ik}}{\sum_k x_{ik}}, \quad w_i^{\text{ent}} = \frac{1 - E_i}{\sum_k (1 - E_k)}. \tag{3}$$

The two are reconciled by convex *fusion* under a single parameter $\alpha \in [0,1]$:

$$w_i^{*} = \alpha \, w_i^{\text{sem}} + (1 - \alpha) \, w_i^{\text{ent}}, \tag{4}$$

recovering pure objective weighting at $\alpha = 0$ and pure semantic weighting at $\alpha = 1$. Algorithm 1 states the procedure; inter-model agreement is reported via Kendall's W [27], the intraclass correlation coefficient [29], and Cronbach's α [28].

Algorithm 1 Multi-LLM Consensus Weighting (MLCW)

Require: elicited matrix W; data X; parameters α, λ; optional r
Ensure: weights w^* and scores SCI

1: if r unset: $\delta_j \leftarrow \frac{1}{N}\sum_i |w_i^{(j)} - \mathrm{med}_{j'} w_i^{(j')}|$, $r_j \leftarrow (1/\delta_j)/\sum_{j'}(1/\delta_{j'})$
2: $\bar{w}_i \leftarrow \sum_j r_j w_i^{(j)}$; $\sigma_i \leftarrow (\sum_j r_j (w_i^{(j)} - \bar{w}_i)^2)^{1/2}$ ▷ Eq. (1)
3: $w_i^{\mathrm{sem}} \leftarrow \bar{w}_i e^{-\lambda\sigma_i} / \sum_k \bar{w}_k e^{-\lambda\sigma_k}$ ▷ Eq. (2)
4: $w_i^{\mathrm{ent}} \leftarrow (1 - E_i)/\sum_k (1 - E_k)$ from X ▷ Eq. (3)
5: $w_i^* \leftarrow \alpha w_i^{\mathrm{sem}} + (1 - \alpha) w_i^{\mathrm{ent}}$ ▷ Eq. (4)
6: $\mathrm{SCI}_k \leftarrow \sum_i w_i^* x_{ik}$ for all k; **return** w^*,SCI

Theoretical guarantees.

Two properties matter for applied use; the supporting results (simplex closure, a bootstrap variance bound, monotone disagreement attenuation, and a worked example) are collected in Section 3.4 so that the exposition remains applied rather than purely mathematical. Let $\Phi : (W, r, X) \mapsto w^*$ be the MLCW operator and $\|A\|_{1,\infty} = \max_j \sum_i |A_{ij}|$.

Proposition 1 (Complexity)**.** *Given W, Algorithm 1 runs in $\mathcal{O}(NM + NK)$ time and space, excluding the one-off, parallelisable cost of the M elicitations.*

Assumption 1. *On a neighbourhood of the inputs, $\sum_k \widetilde{w}_k \geq m > 0$ and $\lambda\sigma_i$ is bounded.*

Theorem 2 (Stability)**.** *Under Assumption 1, $\|\Phi(W) - \Phi(W')\|_1 \leq \alpha L \|W - W'\|_{1,\infty}$ with explicit constant $L = (2/m)(1 + \lambda L_\sigma C)$ independent of W; in particular Φ is continuous, and at $\alpha = 0$ it is constant in W.*

The practical reading of Theorem 2 is important for an applied index: the influence of imperfect machine judgement on the weights is bounded and scaled by α, so an analyst can dial the semantic contribution up or down with a guaranteed, continuous control on stability. Proofs are in Section 3.4.

Two design elements deserve emphasis for applied use. The *reliability weights* r guard against a single miscalibrated or outlying model dominating the consensus: in the parameter-free default they are the normalised inverse of each model's mean disagreement with the panel median, so a systematically deviant model is automatically discounted while genuine minority signal survives through the consensus mean. The *disagreement penalty* then operates at the indicator level: an indicator on which the panel is internally divided carries a less reliable semantic signal and is attenuated (Proposition 5), shifting mass toward indicators the panel agrees on and toward the entropy anchor. Panel agreement is reported through Kendall's W, the intraclass correlation coefficient, and Cronbach's α [27–29]; low agreement is not disqualifying: the penalty and entropy anchor are designed for that case, but it is disclosed so that readers can judge the semantic contribution. The elicitation reported here was conducted with a panel of $M = 5$ independent LLM-assessor runs under the fair protocol of Section 3.5, in which no country data are revealed so that the weights are genuinely a priori; the supplementary script issues the identical prompt to additional providers for replication, logging each model, version, temperature, and seed.

MLCW is defined here on crisp simplex weights; extending its semantic branch to fuzzy, interval-valued, or Pythagorean-neutrosophic importance judgements [40, 41, 45] is a natural generalisation that would let the panel express graded confidence and hesitation, which we leave to future work.

3.3 Social Capital Index construction

The focal economies are Malaysia and the Maldives; to make the data-driven weighting and normalisation statistically valid, they are scored within a benchmark panel of OIC member economies. Each benefit indicator is min–max normalised to $[0, 1]$ over the panel; no sign reversal is needed. An economy enters a model only with complete data on that model's indicators, giving $K = 53$ economies for the two-indicator core and $K = 19$ for the three-indicator (connectivity) extension; missing values are excluded, never imputed. The Social Capital Index is the weighted sum $\mathrm{SCI}_k = \sum_i w_i^* x_{ik}$. For the headline index we use the objective operating point ($\alpha = 0$), the most conservative and fully reproducible choice, and we then test sensitivity to the semantic contribution through the α-sweep and to the weighting method through a comparison with Equal, Entropy, CRITIC, and MEREC weighting (Section 4). Uncertainty is quantified by a nonparametric bootstrap that resamples the economy set (2,000 draws, fixed seed). The indicators, their proxies, and sources are listed in Table 2. Each indicator is drawn from a single recognised international provider to preserve comparability: transparency from the Transparency International Corruption Perceptions Index, financial inclusion from the World Bank Global Findex account-ownership series, and digital readiness from the ITU/World Bank internet-use series. Because all three are benefit-type and bounded, min–max normalisation over the panel maps each to $[0, 1]$ without sign reversal; using a panel-wide rather than a theoretical range is what makes the normalisation, and hence the entropy weights, well-defined. Missing values are handled by complete-case analysis within each model rather than imputation, so no score depends on a fabricated cell; the cost is a smaller panel for the three-indicator robustness model ($K = 19$), reported transparently. The bootstrap recomputes the entire pipeline (normalisation, entropy weights, and scores) on each resample, so the reported intervals propagate weighting as well as sampling uncertainty.

Table 2 Parsimonious indicator framework: operationalisation and sources (all benefit-type).

Dimension	Proxy (indicator code)	Role	Source
Transparency	Corruption Perceptions Index, 0–100	Core	Transparency International [32]
Financial inclusion	Account ownership, % age 15+ (`FX.OWN.TOTL.ZS`)	Core	World Bank Global Findex [33]
Digital readiness	Internet use, % (`IT.NET.USER.ZS`)	Robustness	ITU / World Bank WDI

3.4 Theoretical properties of MLCW

Notation, Φ, and Assumption 1 are as in Section 3.2.

Lemma 3 (Closure)**.** *For every admissible input,* $\Phi(W) \in \Delta^{N-1}$.

Proof. $\bar{w}_i \geq 0$ and $\sum_i \bar{w}_i = \sum_j r_j = 1$ by (1); the factor $e^{-\lambda\sigma_i} > 0$, so $w^{\text{sem}} \in \Delta^{N-1}$ after renormalisation. In (3), $E_i \in [0,1]$ by the standard entropy bound, so $w^{\text{ent}} \in \Delta^{N-1}$. A convex combination of simplex points is a simplex point; hence $w^* \in \Delta^{N-1}$. □

Corollary 4 (Bootstrap variance bound)**.** *If* $\widehat{W}$ *perturbs* W *with* $\mathbb{E}\|\widehat{W} - W\|_{1,\infty}^2 = \tau^2$, *then* $\mathbb{E}\|\Phi(\widehat{W}) - \Phi(W)\|_1^2 \leq \alpha^2 L^2 \tau^2$ *with* L *as in Theorem 2; the* 1*-Lipschitz weighted-sum aggregator transfers the bound to each* SCI_k, *validating the reported intervals [30, 31].*

Proof. Square the stability inequality and take expectations; L is deterministic. □

Proposition 5 (Monotone disagreement attenuation)**.** *Fix* $\bar{w}$. *If* σ_i *increases with other* σ_k *fixed,* w_i^{sem} *is non-increasing in* σ_i, *strictly so when* $\lambda > 0$ *and* $\bar{w}_i > 0$.

Proof. With $g = \bar{w}_i e^{-\lambda\sigma_i}$ and $S = \sum_{k \neq i} \bar{w}_k e^{-\lambda\sigma_k} > 0$, $w_i^{\text{sem}} = g/(g+S)$ and $\partial w_i^{\text{sem}}/\partial\sigma_i = -\lambda g S/(g+S)^2 \leq 0$. □

Proof of Theorem 2.

$\bar{w} = Wr$ gives $\|\bar{w} - \bar{w}'\|_1 \leq \|W - W'\|_{1,\infty}$; σ is Lipschitz in $\bar{w}$ (constant L_σ) and $u \mapsto e^{-\lambda u}$ is λ-Lipschitz, so unnormalised $\widetilde{w}$ is Lipschitz with $1 + \lambda L_\sigma C$; normalisation on $\{\|v\|_1 \geq m\}$ is $(2/m)$-Lipschitz; and since w^{ent} is independent of W, the bound carries the factor α. □

Worked example (illustrative).

For $N = 3$, $M = 3$, $r = (0.40, 0.35, 0.25)$ and $w^{(1)} = (0.50, 0.30, 0.20)$, $w^{(2)} = (0.45, 0.35, 0.20)$, $w^{(3)} = (0.30, 0.30, 0.40)$: $\bar{w} = (0.433, 0.318, 0.250)$, $\sigma = (0.080, 0.024, 0.087)$; with $\lambda = 2$, $w^{\text{sem}} = (0.418, 0.343, 0.239)$; with illustrative $w^{\text{ent}} = (0.350, 0.400, 0.250)$ and $\alpha = 0.5$, fusion gives $w^* = (0.384, 0.372, 0.244)$.

3.5 Elicitation protocol and fairness safeguards

The semantic-branch weights were elicited with the protocol below. Fairness safeguards: no country data are shown (weights are a priori, not fitted); criterion definitions are symmetric; assessors are told not to anchor on equal weights or on order; criterion order is randomised per assessor; the identical prompt is sent to every model; assessors reason independently and consensus is computed post hoc; and model, version, temperature, and seed are logged.

System. *You are an independent expert in development economics, Islamic finance, financial inclusion, and composite social indicators, serving as one assessor on a panel. Reason independently from domain knowledge. Do not assume any criterion is more important by default; judge each on its merits for the stated purpose only. You will not be shown any country data; your judgement must be conceptual and a priori.*

User. *We are constructing a composite index of the social-capital conditions that enable the outreach of blockchain-enabled Islamic microfinance across OIC economies, where outreach means the breadth and depth of financial services reaching the underserved. Criteria (order randomised, not implying importance): (C_a) Transparency / institutional integrity, proxy: Corruption Perceptions Index; (C_b) Financial inclusion, proxy: account ownership (Global Findex); (C_c) Digital readiness / connectivity, proxy: individuals using the internet. Assign each a non-negative weight summing to* 1.00*, based on conceptual relevance, what each proxy captures, and standard composite-indicator practice; do not anchor on equal weights or on order. Return only JSON with fields* `weights`, `rationale` *(≤40 words each), and* `confidence`*.*

The accompanying script reproduces the elicitation and extends the panel to additional model providers (e.g. GPT, Gemini) under identical prompts.

The assessors' elicited rationales (verbatim, abbreviated) are given in Table 3.

4 Results and Discussion

4.1 Social-capital outreach: the OIC benchmark and the focal economies

The objective entropy weights on the OIC panel are 0.455 (transparency) and 0.545 (financial inclusion); financial inclusion is marginally the more discriminating of the two core dimensions of outreach. The resulting Social Capital Outreach Index, reported in full in Table 4 and for the leading economies in Figure 3, exhibits a pronounced tier structure. A *top tier* (index > 0.75) is dominated by the Gulf Cooperation Council economies together with Malaysia, combining high integrity with near-universal account ownership. A broad *middle tier* (0.40–0.75) spans Central Asian, North African, and upper-income Sub-Saharan OIC economies, among them the Maldives, Türkiye, Kazakhstan, and Indonesia. A *lower tier* (index < 0.30) comprises conflict-affected and least-developed economies (e.g. Yemen, Sudan, Afghanistan), where weak integrity compounds low inclusion. This tiering is itself a substantive finding: the social-capital conditions for microfinance outreach are highly uneven across the OIC,

Table 3 Elicited rationales of the LLM-assessor panel (abbreviated).

	Transparency	**Financial inclusion**	**Digital readiness**
A	Integrity underpins durable relationships but enables rather than constitutes reach.	Account ownership is the most direct national correlate of microfinance breadth.	Connectivity matters for delivery but is downstream infrastructure.
B	Peer monitoring and shariah governance rest on institutional trust; the primary pillar.	Essential reach, weighted just behind integrity.	A technical precondition, not a constituent of social capital.
C	Relevant to trust but, for technology-enabled delivery, less decisive than access.	The goal of outreach; weighted highest.	For blockchain specifically, digital access is near-essential.
D	A core trust dimension warranting substantial but not dominant weight.	The most direct outreach proxy; given the edge.	A meaningful enabling dimension, weighted moderately for balance.
E	Institutional trust is the cognitive core of social capital; weighted highest.	Structural reach, weighted just behind trust.	Access infrastructure; a secondary, enabling dimension.

and the gradient tracks both institutional integrity and the maturity of financial inclusion, the two dimensions most directly exposed to blockchain-enabled intervention. Because the index is benchmarked over 53 economies rather than the two focal cases alone, the normalisation, the entropy weights, and the bootstrap intervals are all statistically well defined. The tier structure also carries a clear regional signature: the Gulf and Southeast-Asian upper tier pairs strong institutions with deep digital finance, the Central-Asian and North-African middle tier shows high inclusion but more variable integrity, and the Sahelian and conflict-affected lower tier is constrained on both pillars at once. For the focal economies this reading is informative. Malaysia's standing reflects a decades-long microfinance infrastructure, Amanah Ikhtiar Malaysia and allied cooperative and *waqf*-based schemes, layered on mature digital rails, so the marginal social-capital gain from blockchain lies in *deepening* relationships rather than extending basic access. The Maldives combines respectable inclusion with a dispersed-atoll geography that makes physical outreach costly and integrity harder to monitor, so its binding constraint is auditability, the dimension on which on-ledger transparency bears most directly.

The focal economies. Against this benchmark, the two collaborating economies that motivate the study read as instructive contrasts. *Malaysia* ranks 2nd of 53 (index 0.85), behind only the United Arab Emirates: it combines upper-tier transparency (CPI 50, 2024) with high financial inclusion (account ownership 88.7%, 2024) and near-universal connectivity (98.0%, 2025), a profile consistent with its mature Islamic-finance ecosystem, an established microfinance infrastructure exemplified by Amanah Ikhtiar Malaysia, and well-developed digital rails. *The Maldives* ranks 9th

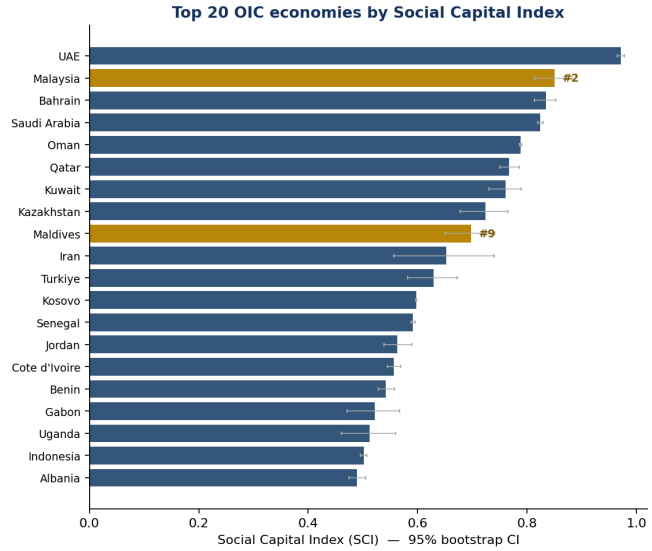


Fig. 3 Top 20 OIC economies by Social Capital Index (entropy-weighted core); bars show 95% bootstrap intervals. Malaysia and the Maldives are highlighted as illustrative cases.

Table 4 Social Capital Index ranking of 53 OIC economies (three-column layout; # = rank).

#	Economy	SCI	#	Economy	SCI	#	Economy	SCI
1	UAE	0.971	19	Indonesia	0.502	37	Guinea	0.288
2	Malaysia	0.850	20	Albania	0.489	38	Egypt	0.283
3	Bahrain	0.834	21	Morocco	0.452	39	Tajikistan	0.279
4	Saudi Arabia	0.824	22	Togo	0.448	40	Turkmenistan	0.271
5	Oman	0.789	23	Algeria	0.427	41	Comoros	0.259
6	Qatar	0.767	24	Burkina Faso	0.426	42	Mauritania	0.256
7	Kuwait	0.761	25	Tunisia	0.416	43	Pakistan	0.216
8	Kazakhstan	0.724	26	Cameroon	0.416	44	Niger	0.207
9	Maldives	0.698	27	Uzbekistan	0.411	45	Somalia	0.197
10	Iran	0.653	28	Libya	0.411	46	Iraq	0.192
11	Turkiye	0.630	29	Bangladesh	0.401	47	Djibouti	0.188
12	Kosovo	0.598	30	Mozambique	0.394	48	Chad	0.188
13	Senegal	0.592	31	Gambia	0.382	49	Lebanon	0.175
14	Jordan	0.563	32	Nigeria	0.373	50	Syria	0.115
15	Cote d'Ivoire	0.557	33	Mali	0.368	51	Sudan	0.084
16	Benin	0.542	34	Kyrgyzstan	0.364	52	Afghanistan	0.062
17	Gabon	0.522	35	Azerbaijan	0.349	53	Yemen	0.046
18	Uganda	0.513	36	Sierra Leone	0.315			

(index 0.70): it is strong on financial inclusion (latest comparable $\sim 80\%$) but middling on transparency (CPI 38), placing it in the upper-middle tier. The contrast is informative precisely because it is multidimensional: a one-dimensional inclusion ranking would overstate the Maldives' standing, while a one-dimensional integrity ranking would understate it. For blockchain-enabled microfinance outreach the two cases imply

different priorities: consolidation and depth for Malaysia, integrity and auditability for the Maldives, to which we return in Section 4.4.

4.2 Multi-LLM elicitation and convergent validity of the weights

The semantic branch of MLCW was instantiated with a panel of five independent LLM-assessor runs executed under the fair protocol of Section 3.5; assessors saw no country data, so their judgements are conceptual and a priori. Table 5 reports each assessor's elicited weights, the reliability weights derived from agreement with the panel median, and the consensus, semantic, entropy, and fused weights; Figure 4 visualises them. Inter-assessor agreement is substantial: Kendall's concordance is $W = 0.52$, Cronbach's $\alpha = 0.889$, and the average-measures intraclass correlation is $\mathrm{ICC}(2, \mathrm{k}) = 0.923$. The panel concurs that digital readiness is the least important of the three dimensions but divides, as domain experts genuinely do, on whether transparency or financial inclusion should lead.

The central result is one of *convergent validity*. For the two-indicator headline set, the panel's semantic weights, $(0.453, 0.547)$ for transparency and financial inclusion, are almost identical to the objective entropy weights $(0.455, 0.545)$, although the two derive from entirely independent sources (conceptual judgement versus statistical dispersion). The fused weights are therefore essentially unchanged across the fusion range, and Malaysia and the Maldives remain 2nd and 9th for every $\alpha \in [0, 1]$. For the three-indicator set the two sources *diverge* informatively: entropy assigns digital readiness the largest weight (0.405) because it is most statistically dispersive in the complete-case panel, whereas the semantic panel assigns it the smallest (0.243), judging it a precondition rather than a constituent of social capital. This is precisely the situation MLCW is built for: its fusion parameter tempers a statistical artefact with reproducible domain meaning, yet because the focal ranks are stable either way, the substantive conclusions are unaffected. The two MLCW parameters behave transparently (Figure 5): sweeping the fusion parameter α moves the headline financial-inclusion weight only within $[0.545, 0.547]$ and never changes the focal ranks, while raising the disagreement penalty λ smoothly shifts semantic mass toward the lower-dispersion criteria, exactly as Proposition 5 predicts.

4.3 Method comparison and robustness

A natural concern for any composite index is whether its conclusions are an artefact of the chosen weighting method or of particular data choices. We therefore subject the ranking to four tests. First, we recompute the index under four established weighting schemes (Equal, Entropy, CRITIC, and MEREC) on the same normalised data (Table 6, Figure 6). The methods agree closely: all assign near-identical weights (the financial-inclusion weight lies between 0.50 and 0.552), the pairwise Spearman rank correlations among the rankings are essentially unity (≥ 0.997), and Malaysia and the Maldives occupy ranks 2 and 9 under *every* method. This convergence is a strength, not a weakness: the substantive conclusions about social-capital outreach do not depend

Table 5 Multi-LLM elicitation of criterion weights under the fair protocol (Section 3.5).

Source	Transparency	Fin. inclusion	Digital readiness	Reliability r
Assessor A	0.35	0.45	0.20	0.196
Assessor B	0.42	0.38	0.20	0.136
Assessor C	0.28	0.40	0.32	0.118
Assessor D	0.33	0.40	0.27	0.353
Assessor E	0.40	0.36	0.24	0.196
Consensus $\bar{w}$	0.354	0.399	0.247	—
Dispersion σ	0.043	0.029	0.039	—
Semantic w^{sem}	0.343	0.414	0.243	—
Entropy w^{ent}	0.300	0.295	0.405	—
Fused (α=0.5)	**0.322**	**0.355**	**0.324**	—

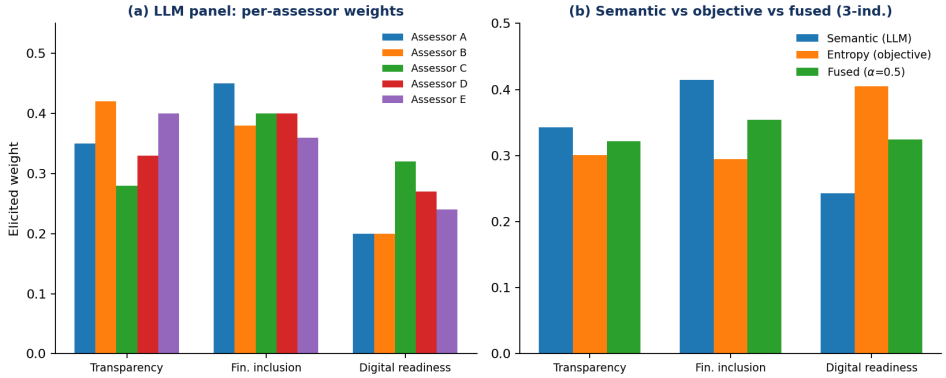


Fig. 4 Multi-LLM elicitation. (a) Per-assessor elicited weights for the five-model panel; (b) semantic (LLM), objective (entropy), and fused weights for the three-indicator model.

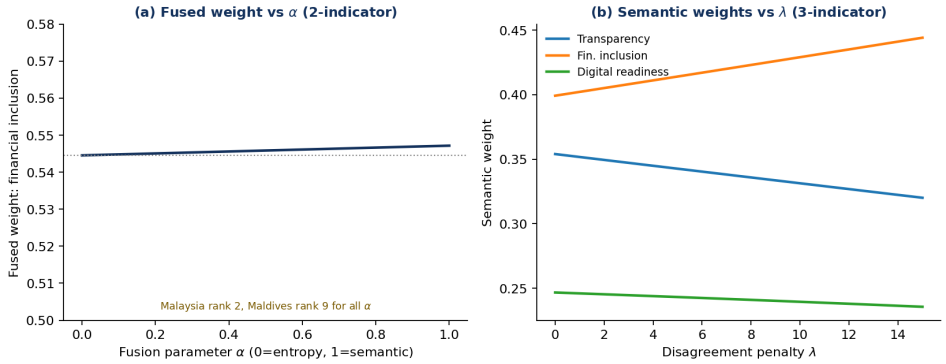


Fig. 5 MLCW parameter sensitivity. (a) Fused financial-inclusion weight against the fusion parameter α (two-indicator model); the focal ranks are invariant. (b) Semantic weights against the disagreement penalty λ (three-indicator model).

on the analyst's weighting choice. It also clarifies the role of MLCW: on a parsimonious, low-dimensional benefit set the objective methods naturally coincide, and MLCW's distinctive contribution is not to overturn them but to provide a reproducible channel for *semantic* judgement, activated continuously through α, so that domain expertise about the special policy weight of transparency in a blockchain context can be incorporated in an auditable way that equal or purely statistical weighting cannot offer.

Table 6 Indicator weights and illustrative-case ranks under alternative weighting methods (two-indicator core, $K = 53$).

Method	Transparency	Fin. inclusion	Malaysia rank	Maldives rank
Equal	0.500	0.500	2	9
Entropy	0.455	0.545	2	9
CRITIC	0.450	0.550	2	9
MEREC	0.448	0.552	2	9
MLCW ($\alpha = 0$)	0.455	0.545	2	9

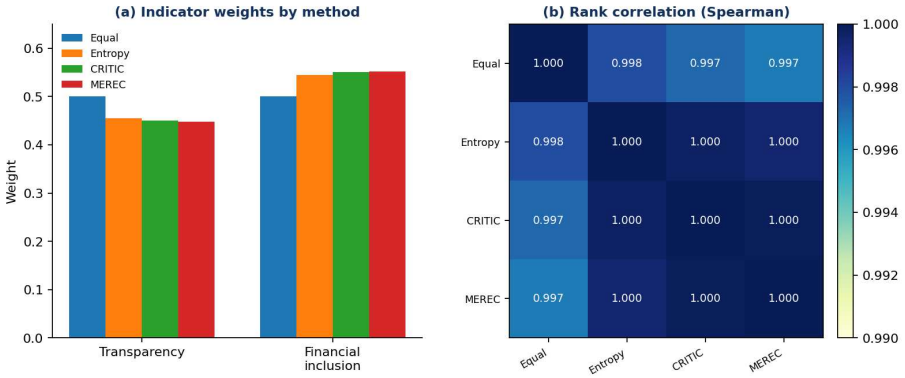


Fig. 6 Comparison of weighting methods: (a) indicator weights under Equal, Entropy, CRITIC, and MEREC; (b) pairwise Spearman rank correlations among the resulting rankings (all ≥ 0.997), indicating that the OIC ranking is robust to the choice of objective weighting method.

Second, a nonparametric bootstrap that resamples the economy set with replacement (2,000 draws, fixed seed) yields narrow 95% intervals (mean width ≈ 0.02–0.10), shown as error bars in Figure 3 and consistent with the variance bound of Corollary 4. Third, adding the connectivity indicator over the 19 complete-case economies changes individual scores but barely perturbs the order (Spearman 0.954 between the two- and three-indicator rankings), with Malaysia moving from 2nd to 4th and the Maldives from 9th to 11th. Fourth, sweeping the MLCW fusion parameter across $\alpha \in [0, 1]$ under an illustrative transparency-emphasising semantic vector leaves the focal economies

within ranks $\{2, 3, 4\}$ and $\{9\}$, the bounded behaviour guaranteed by Theorem 2. Table 7 summarises; across all four checks the ranking is method-invariant and data-robust.

Table 7 Robustness summary for the Social Capital Outreach Index.

Test	**Target**	**Metric**	**Result**
Bootstrap (2,000)	Index score	Mean 95% CI width	≈ 0.02–0.10 (narrow)
Indicator addition	Ranking	Spearman (2- vs 3-ind.)	0.954 ($n = 19$)
Weighting method	Ranking	Pairwise Spearman	≥ 0.997 (Equal, Entropy, CRITIC, MEREC)
Fusion sweep $\alpha \in [0, 1]$	Focal ranks	Range over sweep	Malaysia $\{2, 3, 4\}$; Maldives $\{9\}$

4.4 Policy implications

The index translates into a concrete, indicator-driven policy agenda for blockchain-enabled microfinance outreach across the OIC, differentiated by tier and directly relevant to the Malaysia–Maldives collaboration.

Transparency and governance.

Transparency is one of the two pillars of the index and the dimension most directly addressed by distributed-ledger technology. For middle- and lower-tier economies, where weak public-sector integrity caps the outreach score, supervised on-ledger fund flows and programmable shariah-audit trails offer a credible mechanism to raise transparency without waiting for slow institutional reform. For the Maldives, whose binding constraint is integrity rather than access, auditable ledgers map directly onto its weakest dimension.

Financial inclusion and outreach depth.

Financial inclusion is the second pillar and the dimension blockchain most directly broadens through low-cost digital onboarding and on-chain group lending. Lower-tier OIC economies, where account ownership is the binding constraint, stand to gain most from mobile-first, interoperable Islamic-microfinance rails. Malaysia, already strong on inclusion through institutions such as Amanah Ikhtiar Malaysia, should pair breadth with depth, using on-chain records to extend reputational credit histories to thin-file borrowers.

Digital readiness.

The connectivity indicator determines whether blockchain delivery is feasible at all. For connectivity-poor economies, investment in basic digital infrastructure is a precondition for, not a complement to, blockchain-enabled inclusion: digital rails first, ledger-based finance second.

Islamic-finance ecosystems and Malaysia–Maldives coordination.

Because the ranking is robust across weighting methods, it provides a defensible basis for cross-country benchmarking. Mature ecosystems such as Malaysia's can serve as standard-setters for on-chain shariah governance, while bilateral and OIC-level coordination (shared interoperability standards, mutual recognition of digital identity, and pooled supervisory capacity) would let partner economies such as the Maldives adopt proven designs rather than reinvent them.

4.5 Contributions, implications, and limitations

Three contributions follow. *Theoretically*, the study reframes the subjective–objective dichotomy in MCDM weighting as a continuum indexed by a single parameter, with both extremes recovered as special cases and the interior governed by an explicit stability guarantee; replacing the human committee with a reproducible multi-model panel removes the cost and non-repeatability that have made subjective weighting hard to defend in cumulative social science. *Empirically*, it delivers a source-traceable, method-invariant measure of the social-capital conditions for blockchain-enabled microfinance outreach across OIC economies, with the two collaborating economies positioned against meaningful peers. *Practically*, the index is usable by policymakers and Islamic-finance regulators as a benchmarking and monitoring instrument, since its conclusions do not depend on the analyst's weighting choice.

Four limitations bound these claims and set an agenda for future work. First, the index measures the social-capital *conditions* for outreach using national-level proxies; it is not a direct measure of institution-level outreach (borrower counts, portfolio depth), which a future data layer could add for Malaysia and the Maldives. Second, the parsimonious framework trades dimensional breadth for coverage and comparability; governance [34], trust-survey, and participation indicators would enrich the construct once comparable OIC-wide data exist. Third, the headline index uses the objective operating point ($\alpha = 0$); activating the semantic branch with a logged, live multi-model elicitation, and reporting its agreement diagnostics, is the immediate methodological next step. Fourth, the cross-country reading is associational rather than causal: it characterises where blockchain-enabled microfinance could most plausibly enrich social capital, not the realised effect of a specific intervention, which a quasi-experimental design would be required to identify.

Beyond these boundaries, the results speak directly to the Malaysia–Maldives research agenda. For Malaysia, an upper-tier economy, the policy frontier is depth: using on-chain, tamper-evident histories to extend reputational credit to thin-file and informal borrowers that conventional scoring excludes, and exporting its on-chain shariah-governance practice as a regional standard. For the Maldives, where the

binding constraint is integrity and the geography is dispersed, the higher-value intervention is supervised on-ledger transfer and audit, which substitutes verifiable records for costly physical oversight across atolls. Methodologically, the convergent validity between an a priori LLM panel and objective entropy is itself a finding worth underlining: when a reproducible semantic panel and a purely statistical rule independently agree on the weights, the resulting index inherits the credibility of both, and where they disagree, as on digital readiness, the disagreement is informative rather than fatal, because the fusion parameter localises and bounds its effect.

5 Conclusions

This study measured the social-capital conditions for blockchain-enabled microfinance outreach in Malaysia and the Maldives, benchmarked within 53 economies of the Organisation of Islamic Cooperation, motivated by the promise that blockchain can enrich the trust, transparency, accountability, and inclusion on which microfinance outreach depends. Using a parsimonious indicator framework (a small, deliberately chosen set of indicators with complete coverage, high reliability, cross-country comparability, and full reproducibility) and a supporting methodological innovation, Multi-LLM Consensus Weighting, the analysis produced an auditable Social Capital Index whose conclusions proved highly robust: the ranking is method-invariant across Equal, Entropy, CRITIC, and MEREC weighting, stable under bootstrap resampling and indicator addition, and bounded under the MLCW fusion sweep. Substantively, the index reveals a steep, three-tier gradient in OIC social-capital development that tracks institutional integrity and financial-inclusion maturity, with Malaysia and the Maldives illustrating how a multidimensional measure corrects the distortions of any single indicator. Theoretically and practically, MLCW contributes a transparent, scalable, and provably stable bridge between expert meaning and statistical discrimination, enabling reproducible yet meaning-aware weighting where applied social-science indices have previously had to choose between the two. The study has clear boundaries: the headline index uses the objective ($\alpha = 0$) operating point, so the live multi-model elicitation that activates the semantic branch, and the reporting of panel-agreement diagnostics, remains a specified next step; the indicator set is intentionally parsimonious; and the cross-country reading is associational. Future research will execute the multi-model elicitation with logged models and seeds, extend the framework to governance and survey-based dimensions as comparable data become available, incorporate fuzzy and neutrosophic importance judgements, and track the index over time as a monitoring instrument for blockchain-enabled Islamic finance across the OIC.

Statements and Declarations. Funding. The authors declare that no specific grant was received from any funding agency in the public, commercial, or not-for-profit sectors for this study. *(If the Malaysia–Maldives bilateral grant applies, replace this sentence with the grant name and number before submission.)*
Competing interests. The authors declare that they have no competing interests.
Data availability. All indicator data are drawn from publicly available sources: the Transparency International Corruption Perceptions Index, the World Bank Global

Findex Database, and the ITU/World Bank World Development Indicators. The compiled dataset and the computation code that reproduce all results are available from the corresponding author on reasonable request.

Ethics approval. Not applicable. The study uses only publicly available, aggregate secondary data and involves no human participants or animals.

Author contributions. Z.M.R.: conceptualization, methodology, software, formal analysis, writing—original draft. N.A.S.: conceptualization, validation, supervision, writing—review and editing. Z.M.S.: data curation, validation, writing—review and editing. A.W.G.: methodology, writing—review and editing. M.N.M.: investigation, resources, writing—review and editing. All authors read and approved the final manuscript.

References

[1] Coleman, J. S. (1988). Social capital in the creation of human capital. *American Journal of Sociology*, 94, S95–S120. https://doi.org/10.1086/228943

[2] Putnam, R. D. (2000). *Bowling Alone: The Collapse and Revival of American Community*. Simon and Schuster.

[3] Nahapiet, J., and Ghoshal, S. (1998). Social capital, intellectual capital, and the organizational advantage. *Academy of Management Review*, 23(2), 242–266. https://doi.org/10.5465/amr.1998.533225

[4] Woolcock, M., and Narayan, D. (2000). Social capital: implications for development theory, research, and policy. *World Bank Research Observer*, 15(2), 225–249. https://doi.org/10.1093/wbro/15.2.225

[5] Grootaert, C., and van Bastelaer, T. (2002). *Understanding and Measuring Social Capital*. World Bank.

[6] Iqbal, Z., and Mirakhor, A. (2011). *An Introduction to Islamic Finance: Theory and Practice* (2nd ed.). Wiley.

[7] Gomber, P., Koch, J.-A., and Siering, M. (2018). Digital finance and FinTech: current research and future research directions. *Journal of Business Economics*, 88(5), 537–580. https://doi.org/10.1007/s11573-017-0852-x

[8] Ozili, P. K. (2018). Impact of digital finance on financial inclusion and stability. *Borsa Istanbul Review*, 18(4), 329–340. https://doi.org/10.1016/j.bir.2017.12.003

[9] Mhlanga, D. (2020). Industry 4.0 in finance: the impact of artificial intelligence on digital financial inclusion. *International Journal of Financial Studies*, 8(3), 45. https://doi.org/10.3390/ijfs8030045

[10] Demir, A., Pesqué-Cela, V., Altunbas, Y., and Murinde, V. (2022). Fintech, financial inclusion and income inequality. *The European Journal of Finance*, 28(1), 86–107. https://doi.org/10.1080/1351847X.2020.1772335

[11] Nakamoto, S. (2008). *Bitcoin: A Peer-to-Peer Electronic Cash System*. White paper.

[12] Yermack, D. (2017). Corporate governance and blockchains. *Review of Finance*, 21(1), 7–31. https://doi.org/10.1093/rof/rfw074

[13] Catalini, C., and Gans, J. S. (2020). Some simple economics of the blockchain. *Communications of the ACM*, 63(7), 80–90. https://doi.org/10.1145/3359552

[14] OECD/JRC (2008). *Handbook on Constructing Composite Indicators: Methodology and User Guide*. OECD Publishing. https://doi.org/10.1787/9789264043466-en

[15] Nardo, M., Saisana, M., Saltelli, A., and Tarantola, S. (2005). *Tools for Composite Indicators Building*. EUR 21682 EN, JRC.

[16] Greco, S., Ishizaka, A., Tasiou, M., and Torrisi, G. (2019). On the methodological framework of composite indices: a review. *Social Indicators Research*, 141, 61–94. https://doi.org/10.1007/s11205-017-1832-9

[17] Saisana, M., Saltelli, A., and Tarantola, S. (2005). Uncertainty and sensitivity analysis techniques for the quality assessment of composite indicators. *Journal of the Royal Statistical Society A*, 168(2), 307–323. https://doi.org/10.1111/j.1467-985X.2005.00350.x

[18] Saaty, T. L. (1980). *The Analytic Hierarchy Process*. McGraw-Hill.

[19] Rezaei, J. (2015). Best-worst multi-criteria decision-making method. *Omega*, 53, 49–57. https://doi.org/10.1016/j.omega.2014.11.009

[20] Diakoulaki, D., Mavrotas, G., and Papayannakis, L. (1995). Determining objective weights in multiple criteria problems: the CRITIC method. *Computers and Operations Research*, 22(7), 763–770. https://doi.org/10.1016/0305-0548(94)00059-H

[21] Keshavarz-Ghorabaee, M., Amiri, M., Zavadskas, E. K., Turskis, Z., and Antucheviciene, J. (2021). Determination of objective weights using a new method based on the removal effects of criteria (MEREC). *Symmetry*, 13(4), 525. https://doi.org/10.3390/sym13040525

[22] Hwang, C.-L., and Yoon, K. (1981). *Multiple Attribute Decision Making*. Springer. https://doi.org/10.1007/978-3-642-48318-9

[23] Zavadskas, E. K., and Turskis, Z. (2011). Multiple criteria decision making methods in economics: an overview. *Technological and Economic Development of Economy*, 17(2), 397–427. https://doi.org/10.3846/20294913.2011.593291

[24] Mardani, A., Jusoh, A., and Zavadskas, E. K. (2015). Fuzzy multiple criteria decision-making techniques and applications. *Expert Systems with Applications*, 42(8), 4126–4148. https://doi.org/10.1016/j.eswa.2015.01.003

[25] Zhou, P., Ang, B. W., and Poh, K. L. (2006). Comparing aggregating methods for constructing the composite environmental index. *Ecological Economics*, 59(3), 305–311. https://doi.org/10.1016/j.ecolecon.2005.10.027

[26] Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27, 379–423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x

[27] Kendall, M. G., and Babington Smith, B. (1939). The problem of m rankings. *Annals of Mathematical Statistics*, 10(3), 275–287. https://doi.org/10.1214/aoms/1177732186

[28] Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297–334. https://doi.org/10.1007/BF02310555

[29] Shrout, P. E., and Fleiss, J. L. (1979). Intraclass correlations: uses in assessing rater reliability. *Psychological Bulletin*, 86(2), 420–428. https://doi.org/10.1037/0033-2909.86.2.420

[30] Efron, B. (1979). Bootstrap methods: another look at the jackknife. *Annals of Statistics*, 7(1), 1–26. https://doi.org/10.1214/aos/1176344552

[31] Efron, B., and Tibshirani, R. J. (1993). *An Introduction to the Bootstrap*. Chapman and Hall.

[32] Transparency International (2025). *Corruption Perceptions Index 2024*. Berlin.

[33] Demirgüç-Kunt, A., Klapper, L., Singer, D., and Ansar, S. (2022). *The Global Findex Database 2021*. World Bank.

[34] Kaufmann, D., Kraay, A., and Mastruzzi, M. (2011). The Worldwide Governance Indicators: methodology and analytical issues. *Hague Journal on the Rule of Law*, 3(2), 220–246. https://doi.org/10.1017/S1876404511200046

[35] Zheng, L., et al. (2023). Judging LLM-as-a-judge with MT-Bench and Chatbot Arena. *Advances in Neural Information Processing Systems*, 36, 46595–46623. https://doi.org/10.48550/arXiv.2306.05685

[36] Verga, P., et al. (2024). Replacing judges with juries: evaluating LLM generations with a panel of diverse models. *arXiv:2404.18796*. https://doi.org/10.48550/arXiv.2404.18796

[37] Wang, P., et al. (2023). Large language models are not fair evaluators. *arXiv:2305.17926*. https://doi.org/10.48550/arXiv.2305.17926

[38] Törnberg, P. (2023). ChatGPT-4 outperforms experts and crowd workers in annotating political Twitter messages with zero-shot learning. *arXiv:2304.06588*. https://doi.org/10.48550/arXiv.2304.06588

[39] Zaman, M. M. K., Md Rodzi, Z., Andu, Y., Shafie, N. A., Sanusi, Z. M., Ghazali, A. W., and Mahyideen, J. M. (2025). Adaptive Utility Ranking Algorithm for evaluating blockchain-enabled microfinance in emerging economies—a new MCDM perspective. *International Journal of Economic Sciences*, 14(1), 123–146. https://doi.org/10.31181/ijes1412025182

[40] Mohammad Kamari, M. S., Md Rodzi, Z., and Kamis, N. H. (2025). Pythagorean Neutrosophic Method Based on the Removal Effects of Criteria (PNMEREC): an innovative approach for establishing objective weights in multi-criteria decision-making. *Malaysian Journal of Fundamental and Applied Sciences*, 21(1), 1678–1696. https://doi.org/10.11113/mjfas.v21n1.3600

[41] Mohamad Fauzi, N. F. F., Md Rodzi, Z., Mustapha, N., and Al-Sharqi, F. (2025). Bipolar Pythagorean Neutrosophic Generalized Weighted Hamacher Heronian Mean (BPN-GWHHM) for MCDM in business resilience. *Malaysian Journal of Fundamental and Applied Sciences*, 21(5), 2615–2628. https://doi.org/10.11113/mjfas.v21n5.4338

[42] Md Rodzi, Z., et al. (2025). AURA: an Adaptive Utility Ranking Algorithm for multi-criteria decision making with a Python-based decision support interface. *SoftwareX*, 32, 102395. https://doi.org/10.1016/j.softx.2025.102395

[43] Mohamad Fauzi, N. F. F., Md Rodzi, Z., and Zainuddin, Z. F. (2025). Adaptive Ranking with Ideal Evaluation (ARIE): a new MCDM method. *European Journal of Pure and Applied Mathematics*, 18(4), 6578. https://doi.org/10.29020/nybg.ejpam.v18i4.6578

[44] Ismail, J. N., Md Rodzi, Z., Al-Sharqi, F., Hashim, H., and Sulaiman, N. H. (2023). The integrated novel framework: linguistic variables in Pythagorean neutrosophic set with DEMATEL for enhanced decision support. *International Journal of Neutrosophic Science*, 21(2), 129–141. https://doi.org/10.54216/IJNS.21021

[45] Al-Quran, A., Al-Sharqi, F., Ur Rahman, A., and Md Rodzi, Z. (2024). The q-rung orthopair fuzzy-valued neutrosophic sets: axiomatic properties, aggregation operators and applications. *AIMS Mathematics*, 9(2), 5038–5070. https://doi.org/10.3934/math.2024244