The Dignity Gap: Redefining Poverty as a Multidimensional Computational Target and Its Implications for AI-Assisted Welfare Governance

Anonymous authors
Paper under double-blind review

Abstract

Official poverty measurement in India classifies 11.28% of the population as poor (NITI Aayog MPI 2022-23 round). We argue this figure reflects a definitional choice calibrated to administrative manageability rather than human reality, and formalise it as objective misspecification (Krakovna et al. 2020) at the institutional level: the official target is a truncated, binarised functional of the welfare objective the state is mandated to pursue. Applying a three-tier taxonomy – Survivability Failure, Dignity Poverty, and Aspirational Exclusion – grounded in published national survey data (GHI 2025; FAO-SOFI 2025; ASER 2023; NHA 2021-22; NITI Aayog 2024) and a dependency-free bound on inter-survey overlap, we estimate a corrected headcount of 58–69% of India's 1.4 billion population, an undercount of 5.1–6.1×. We introduce the Dignity-Adjusted Poverty Index (DAPI), extending Alkire-Foster (2011) with continuous deprivation scoring over six independently verifiable dimensions weighted by an entropy-based dispersion measure (Zeleny 1982) over NFHS-5 district frequencies, robust to ±20% perturbation. We then give three results on welfare routing: under additive separability, scheme-centric allocation is strictly Pareto-dominated by a person-centric formulation on the same budget; computing the person-centric optimum is NP-hard by reduction from the Generalised Assignment Problem; and a polynomial-time cost-benefit greedy router attains a $(1 - 1/e)$ approximation under a submodularity condition we verify for DAPI. On a synthetic population of 50,000 citizens calibrated to NFHS-5 marginals, the router more than doubles deprivation reduction per rupee and, with a non-exclusion constraint, eliminates false exclusion at negligible welfare cost. Critically, the status quo's false-exclusion rate is near-invariant across an eightfold budget increase: exclusion is architectural, not budgetary. We conclude with four minimum computational requirements for AI-assisted welfare coordination at scale.

1 Introduction

In 2022-23, India's official Multidimensional Poverty Index (MPI) recorded a poverty rate of 11.28%, representing 160 million people deprived in at least one-third of weighted welfare indicators (NITI Aayog 2024). This figure is accurate: 248 million people exited this classification between 2013-14 and 2022-23, one of the largest documented reductions in MPI headcount within a decade. It is also, we argue, the wrong number.

In the same measurement period: (a) 42.9% of India's population (≈600 million people) could not afford a nutritionally adequate diet (FAO-SOFI 2025); (b) 18.7% of children under five suffered wasting, the second highest rate globally (GHI 2025); (c) 26% of rural teenagers aged 14–18 could not read a Class 2 textbook in their own regional language despite years of school enrolment (ASER 2023); and (d) 47–48% of all health expenditure was paid out-of-pocket (NHA 2021-22), making health shocks the leading driver of households falling into survival poverty. These conditions characterise a population predominantly *above* the official poverty threshold.

The coexistence of an 11.28% poverty rate with these conditions is not a statistical contradiction: it is the expected consequence of an objective function designed to detect a narrow form of deprivation, toward which India's welfare system has increasingly optimised rather than toward the broader welfare goal that objective was meant to proxy. When a measure becomes a target, it ceases to be a good measure (Goodhart 1975; Strathern 1997). This is an institutional design failure rather than a measurement failure, analogous to the reward hacking documented by Krakovna et al. (2020) but occurring in the definitional architecture of the welfare state rather than in an algorithm. Sen's (1999) capabilities approach grounds the critique: poverty is the absence of substantive freedoms, not of income, and by that standard the official line measures only the most extreme capabilities failure.

This paper makes six contributions. First, we formalise the targeting failure as objective misspecification in the sense used in the AI alignment literature (§1.1). Second, we formalise a

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

three-tier poverty taxonomy – Survivability Failure (SF), Dignity Poverty (DP), and Aspirational Exclusion (AE) – with a dependency-free methodology for bounding tier estimates drawn from different national surveys (§3.1). Third, we introduce the Dignity-Adjusted Poverty Index (DAPI), extending Alkire-Foster (2011) with continuous scoring and entropy-derived weights from NFHS-5 (§3.2). Fourth, we give three results on welfare routing: a strict optimality gap between the scheme-centric and person-centric formulations (Theorem 1), NP-hardness of the person-centric problem by reduction from the Generalised Assignment Problem (Theorem 2), and a $(1 - 1/e)$ approximation guarantee for a polynomial-time greedy router under a submodularity condition we verify for DAPI (Theorem 3) (§3.3). Fifth, we validate the router on a synthetic population of 50,000 citizens calibrated to NFHS-5 marginals, finding that person-centric routing more than doubles deprivation reduction per rupee and that the status quo's false-exclusion rate is invariant to budget (§4.3). Sixth, we derive four minimum computational requirements any AI welfare system must satisfy (§4).

1.1 Objective Misspecification in Welfare Targeting

We state the failure in the vocabulary of reward misspecification. Let P be the population distribution, $d_i \in [0,1]^{|J|}$ person i's deprivation vector over dimensions J, and π an allocation policy. The *true latent welfare objective* is the expected capability shortfall the state is mandated to reduce, $W^*(\pi) = \mathbb{E}_{i\sim P}[U(d_i(\pi))]$ with $U(d) = -\sum_{j\in J} w_j\, d_j$ the (negated) DAPI aggregator of §3.2. What the system actually optimises is the *administrative proxy*

$$\widetilde{W}(\pi) \;=\; -\sum_{i} \mathbf{1}\,[\,\mathrm{MPI}(d_i(\pi)) \geq \tau\,], \qquad \mathrm{MPI}(d) \;=\; \sum_{j\in J_0} v_j\;\mathbf{1}[\,d_j > 0\,], \quad J_0 \subsetneq J.$$

Two operations separate $\widetilde{W}$ from W^*: *truncation* of the support ($J_0 \subsetneq J$ omits nutritional adequacy, functional literacy, and income security entirely) and *binarisation* (the indicators discard both deprivation intensity and all variation below τ). Both are many-to-one, so $\widetilde{W}$ is constant on level sets over which W^* varies, and $\arg\max_\pi \widetilde{W} \neq \arg\max_\pi W^*$ in general: moving a person from just below to just above τ earns full credit under $\widetilde{W}$ while contributing arbitrarily little to W^*, and relieving severe deprivation in $J \setminus J_0$ earns none. Successive rounds of scheme design and budget reallocation constitute ascent on $\widetilde{W}$; the 248-million-person MPI reduction is the return on that ascent, §3.1 estimates the resulting gap in W^*, and §4.3 shows quantitatively that the induced allocation is far from W^*-optimal.

2 Related Work

2.1 Poverty Measurement: The Composite Index Critique

The Alkire-Foster (2011) MPI and its adoption by NITI Aayog are a genuine advance over unidimensional income lines. Ravallion (2011) argues that multidimensional indices conflate measurement with normative weighting in ways that obscure policy choices, and Atkinson (2003) shows counting approaches are sensitive to cutoff selection in ways aggregate indices hide. Decancq and Lugo (2013) survey eight weighting approaches as data-driven, normative, or hybrid; DAPI's entropy weighting (§3.2.2) is data-driven, a positioning we make explicit. We take Ravallion's critique seriously by deriving the weights through a documented procedure and subjecting them to sensitivity analysis. DAPI extends the decomposable measures of Foster, Greer, and Thorbecke (1984) with continuous scoring and generalises the continuous poverty gap of Watts (1968) to multiple dimensions. Deaton (2010) shows poverty-line construction choices have large effects on measured headcounts, and Banerjee and Duflo (2007) document the co-occurring household deprivations motivating §1. Conceptually, DAPI operationalises Sen's (1999) capabilities approach and Nussbaum's (2001) capabilities list – poverty as the absence of substantive freedoms rather than of income – within the Alkire-Foster computational framework, anchoring thresholds to measurable national survey indicators.

2.2 Satellite-Based Poverty Prediction

Machine-learning poverty prediction achieves high accuracy at recovering official poverty labels, but thereby inherits whatever misspecification those labels carry. Jean et al. (2016) predict consumption and asset wealth in Sub-Saharan Africa from satellite imagery; Yeh et al. (2020) extend this over time; Chi et al. (2022) produce wealth microestimates for all low- and middle-income countries. All accept the official definition as the prediction target. In the notation of §1.1, they estimate $\widetilde{W}$ with decreasing error while leaving the $\widetilde{W}$-W^* gap untouched: optimising the wrong objective more precisely.

2.3 Algorithmic Welfare Systems and Specification Gaming

Eubanks (2018) documents how automated decision systems amplify welfare exclusions rather than reducing them. Krakovna et al. (2020) name specification gaming: optimising a measurable proxy in ways that undermine the goal it was meant to serve. Saxena et al. (2020) find US child-welfare risk-assessment systems built on variables selected for administrative availability rather than predictive validity, and Australia's Robodebt scheme (Chowdhury 2024) is the canonical case of welfare automation optimising a legally incorrect objective at scale. Our contribution is to locate the failure one level earlier: specification gaming in India's welfare system occurs not at the algorithmic level – there is no welfare-routing algorithm to misspecify – but at the definitional level, and all allocation downstream of that target inherits and amplifies the misspecification.

3 Formal Problem Definition

3.1 Three-Tier Poverty Taxonomy

Let person i be indexed in a population of size $N = 1.4$ billion. We define three mutually exclusive, exhaustive tiers, ordered by severity. **Tier 1, Survivability Failure (SF)**: deprivation in at least one of caloric access (<2,100 kcal/day), physical shelter, or emergency healthcare – broadly the MPI-poor population. **Tier 2, Dignity Poverty (DP)**: not in SF, but deprived in at least one of nutritional adequacy, functional literacy, healthcare without financial catastrophe, or dignified shelter. **Tier 3, Aspirational Exclusion (AE)**: in neither SF nor DP, but deprived in at least one of savings buffer, retirement security, formal employment, or digital participation. Full definitions and per-dimension thresholds are in Appendix C.

The tiers are mutually exclusive by construction, but combining tier-level headcounts requires care: the underlying surveys (NITI Aayog MPI, FAO-SOFI, ASER, NHA, ILO, NSAP, CMIE) use different sampling frames and reference periods, with no individual-level linkage. We therefore replace the unweighted summation used in earlier drafts with an explicit bounding methodology. DP membership is a logical OR across four dimensions whose joint dependency structure is unknown, so we bound it with Fréchet inequalities rather than assuming independence. Under the cited rates – nutritional adequacy 42.9% (FAO-SOFI 2025), healthcare catastrophe 47.5% (NHA 2021-22), shelter insecurity 18.5% (SECC 2011 proxy), and a corrected functional-literacy estimate of ≈21% (below) – the Boole upper bound exceeds 100% and is uninformative, while the Fréchet lower bound is the largest single rate, healthcare catastrophe. Netting out Tier 1 at its upper-bound estimate of 210 million gives a DP floor of $(0.475 \times 1.4\text{B}) - 210\text{M} \approx 455$ million. This floor is *dependency-free*: it holds however the four dimensions co-occur. We retain 480 million as a working upper *anchor* rather than a bound; we do not derive it from independence, since deprivations cluster within poor households (Alkire and Foster 2011), making the independence estimate (≈970 million) implausibly high. Formulating the bound over published statutory aggregates rather than linked microdata lets any oversight body recompute and contest every step from public sources; the price is that the upper anchor is not itself dependency-free.

One input required correction. The ASER (2023) figure of 26% applies to *rural* respondents aged 14–18 and was extrapolated nationally without adjustment in earlier drafts. Reweighting by rural population share and an illustrative urban rate of half the rural rate gives a national estimate of ≈21% (≈294 million) rather than 26% (derivation in Appendix C). We flag this as illustrative pending an ASER urban supplement or microdata cross-tabulated against the Class-2 reading standard (§5.2). It does not alter the DP floor, which healthcare governs, so the central empirical conclusion is robust to the literacy-proxy bias identified in review. Tier 3 is reported using the

200–270 million range (ILO India 2024; NSAP; CMIE); the cited sources give only the bundled estimate, so a dependency-free AE bound cannot be constructed from published figures (§5.2).

Table 1. Revised tier headcounts using the dependency-free bounding methodology (§3.1).

Tier	Range (M)	% of 1.4B	Basis
Tier 1 SF	160–210	11–15%	NITI Aayog MPI 2024; GHI 2025
Tier 2 DP (add'l)	455–480	32–34%	Fréchet lower bound + working upper anchor
Tier 3 AE (add'l)	200–270	14–19%	ILO 2024; NSAP; CMIE
TOTAL	815–960	58–69%	Dependency-free bound

Note: Earlier drafts reported 65–75%, revised to 54–69% using unweighted tier summation, and again here to 58–69% using the dependency-free Fréchet lower bound for the DP tier with a corrected literacy estimate. Even at the lower bound this is an undercount of 5.1× relative to the official 11.28% figure; at the upper bound, 6.1×.

3.2 The Dignity-Adjusted Poverty Index (DAPI)

3.2.1 Formal Definition

Subscript i indexes persons, j deprivation dimensions, and d the $|D| = 707$ districts covered by NFHS-5 (IIPS 2021). We extend Alkire-Foster (2011) with continuous deprivation scoring. Let person i have indicators $d_{ij} \in [0,1]$ across $j \in J = \{\text{caloric, nutritional, literacy, healthcare, shelter, income_security}\}$ with weights w_j summing to 1, giving the score $c_i = \sum_j w_j \, d_{ij}$.

Here $d_{ij} = 0$ if person i meets or exceeds the dimension-j threshold; $d_{ij} = (\text{threshold}_j - \text{observed}_j)/\text{threshold}_j$ for continuous dimensions (caloric: kcal gap; healthcare: out-of-pocket excess above 10% of consumption); and $d_{ij} = 1$ for binary dimensions (literacy: cannot read at grade level; shelter: kutcha housing; income security: no pension *and* informal employment *and* no savings buffer). The cutoff is $k = 0.20$, lower than the Alkire-Foster $1/3$, because continuous intensity scoring means partial deprivations aggregate meaningfully and because the near-poverty population is the primary early-intervention target; person i is DAPI-poor if $c_i \geq k$. We report no k-sensitivity table analogous to Appendix B's weight analysis, since the population distribution of c_i it would require is not recoverable from published aggregates; §4.3 instead reports k-robustness on the synthetic population, where the joint distribution is known by construction.

3.2.2 Dimension Weights: Entropy-Based Dispersion Derivation

Weights w_j are derived from an entropy-based dispersion measure adapted from Zeleny (1982) and applied to NFHS-5 deprivation frequencies p_{jd} across the $|D| = 707$ districts. Following common practice when extending entropy weighting to deprivation rates rather than normalised shares (cf. Decancq and Lugo 2013), we apply the formula directly to the p_{jd}, which need not sum to 1 across districts; $E_j = -(\ln |D|)^{-1} \sum_d p_{jd} \ln p_{jd}$ (with $0 \ln 0 = 0$) then measures spatial homogeneity rather than a normalised Shannon entropy, but retains the qualifying property that more spatially uniform dimensions receive lower weight, uniform deprivation being less informative for geographic targeting. The divergence is $\delta_j = 1 - E_j$ and the normalised weight $w_j = \delta_j / \sum_{j'} \delta_{j'}$ (Table 2; full computation in Appendix A).

Table 2. Entropy-derived DAPI dimension weights from NFHS-5 district-level deprivation frequencies ($|D| = 707$). Thresholds are defined in §3.2.1. *Corrected national estimate (§3.1); raw ASER rural 14–18 figure is 26%. **Informal employment rate; full criterion in §3.2.1.

Dimension j	Freq.	E_j	δ_j	w_j
Caloric adequacy	12.0%	0.862	0.138	0.148
Nutritional adequacy	42.9%	0.812	0.188	0.202

Functional literacy	≈21%*	0.813	0.187	0.200
Healthcare	47.8%	0.821	0.179	0.192
Shelter security	18.5%	0.871	0.129	0.138
Income security	85%**	0.888	0.112	0.120
Total	n/a	n/a	0.933	1.000

The derived weights are broadly consistent with the manually-set weights of earlier drafts (within the ±15% perturbation range). Sensitivity analysis (Appendix B; Figure 2) shows the DAPI headcount is robust to weight perturbation: varying each w_j by ±20% changes the total by at most a few percentage points, and no tested perturbation approaches the official 11.28% figure. Because the derivation consumes only published district frequencies, any third party can reproduce and contest the weights from the public NFHS-5 fact sheets; the cost is that we report no bootstrap standard errors, which would require the district microdata release (§5.2).

3.3 Welfare Routing as Constrained Optimisation

3.3.1 Problem Formulation

Let B denote the total welfare budget (approximately INR 13–15 lakh crore/year across central and state government, NGO, and CSR sources); $S = \{s_1, \ldots, s_m\}$ the set of $m \geq 805$ welfare schemes; P the population of $n \approx 1.4$ billion persons; d_i the DAPI deprivation vector for person i; $b_{ij} \geq 0$ the gross budget allocation to person i under scheme j; $l_{ij} \in [0,1]$ the leakage rate, so that the net benefit received is $(1 - l_{ij})b_{ij}$; and $\Delta_i(\cdot)$ the total deprivation reduction for person i as a function of the net benefits received across all schemes. The person-centric (correct) welfare optimisation problem is:

$$\max_{b_{ij}} \quad \sum_i \Delta_i\left(\{(1 - l_{ij})\, b_{ij}\}_j\right) \quad \text{s.t.} \quad \sum_{i,j} b_{ij} \leq B, \quad b_{ij} \geq 0, \quad \text{eligibility}_j(i) = 1 \quad \forall\, j \in \text{route}(i).$$

The scheme-centric (current) formulation solved by India's welfare system is instead $\max_{b_j} \sum_j E_j(b_j)$ subject to $\sum_j b_j \leq B$, where $E_j(b_j)$ is the expenditure achievement for scheme j. This maximises expenditure per scheme independently, with no cross-scheme coordination and no deprivation-reduction objective. Both formulations measure cost identically, as gross budget allocated, so the budget constraint needs no leakage adjustment; leakage enters through Δ_i, which depends on the net benefit $(1 - l_{ij})b_{ij}$ received rather than the gross amount.

3.3.2 Routing Optimality Gap Theorem

Theorem 1. $\text{OPT}_{\text{PC}} > \text{OPT}_{\text{SC}}$ *whenever conditions (a) multidimensional deprivation served by uncoordinated schemes, or (b) differential leakage rates across overlapping schemes, hold for at least one person.*

Proof (summary). The scheme-centric feasible set is a strict subset of the person-centric one: any scheme-centric allocation is reproducible person-centrically, whereas person-centric routing additionally admits allocations serving a person's multidimensionally distributed deprivations across uncoordinated schemes (condition a) and shifting budget from higher- to lower-leakage schemes (condition b), neither representable scheme-centrically. Optimising over this strict superset yields $\text{OPT}_{\text{PC}} \geq \text{OPT}_{\text{SC}}$, strictly whenever the additional allocations carry positive marginal welfare, which both conditions guarantee. Appendix E gives the four-step proof; the general optimality gap (Step 4) assumes additive separability of Δ_i, while the leakage-reallocation argument (Step 3) is constructive and holds for any non-decreasing Δ_i. □

Both conditions plausibly hold in India. A malnourished child is simultaneously deprivation-eligible for PM POSHAN, ICDS, and Ayushman Bharat, yet no scheme triggers the others under current practice (a); and PDS grain leakage of 28% (ICRIER/IFPRI 2024) against near-zero leakage for PMJDY direct transfers is a 28-percentage-point differential across schemes serving overlapping populations (b).

3.3.3 Computational Complexity of Person-Centric Routing

Theorem 1 says nothing about whether the person-centric optimum can be attained, and a reader is entitled to ask why the state could not simply reorganise its way to $\mathrm{OPT}_{\mathrm{PC}}$ administratively. It cannot: the problem is intractable at population scale, which is what makes algorithmic coordination necessary rather than merely convenient. In deployment, benefits are indivisible (a household either receives the ration, the housing unit, or the procedure) and each scheme has a per-cycle servable capacity C_j, turning the programme above into an integer program over $x_{ij} \in \{0,1\}$ maximising $\sum_{i,j} w_j\, d_{ij} x_{ij}$ subject to the budget and the per-scheme capacities, where x_{ij} indicates that person i's deprivation in the dimension served by scheme j is closed at cost $\kappa_j d_{ij}/(1 - l_{ij})$ (Appendix G).

Theorem 2 (Hardness). *The person-centric welfare routing problem is* NP*-hard, and remains so with a single scheme per dimension and uniform leakage. Deciding whether a routing of value at least V exists is* NP*-complete.*

Proof (sketch). Reduction from the Generalised Assignment Problem: given a GAP instance with bins of capacity C_j and items of profit p_{ij} and size s_{ij}, set $w_j d_{ij} = p_{ij}$, $\kappa_j d_{ij}/(1 - l_{ij}) = s_{ij}$, and $B = \sum_j C_j$. Feasible routings correspond exactly to feasible GAP assignments of equal value, so an optimal router solves GAP; restricting to one scheme collapses this to 0-1 Knapsack. Membership in NP follows by verifying both constraint families in $O(nm)$. Appendix G gives the full reduction. □

With $n \approx 1.4$ billion and $m \geq 805$, exact optimisation is not merely administratively awkward but computationally infeasible, so the relevant question is what a polynomial-time router can guarantee. The DAPI objective turns out to be well behaved.

Theorem 3 (Approximation). *If each person's welfare function Δ_i is monotone and concave in the net benefit received per dimension, then the routing objective $F(S) = \sum_i \Delta_i$ is monotone submodular, and the cost-benefit greedy router with partial enumeration runs in polynomial time and returns S with $F(S) \geq (1 - 1/e)\ \mathrm{OPT}_{\mathrm{PC}}$.*

Proof (sketch). F is a sum over persons of monotone concave functions of the net benefit delivered to disjoint (person, dimension) slots, and a non-negative weighted sum of monotone concave functions of modular set functions is monotone submodular; the λ term of §3.4 adds a weighted coverage function, also monotone submodular, and submodularity is closed under non-negative addition. Maximisation under the knapsack constraint then admits the $(1 - 1/e)$ bound of Sviridenko (2004) via density-ordered greedy with enumeration over size-3 seed sets, and the per-scheme capacities form a partition matroid that preserves it. Appendix H verifies submodularity for the DAPI aggregator. □

The concavity premise is an assumption about welfare, not a technical convenience, and is where the guarantee is most contestable: it asserts diminishing returns within a dimension, standard in welfare economics but not universally true, since a partially-funded housing unit may deliver no welfare until complete. Where a dimension is genuinely all-or-nothing, Δ_i is a step function, submodularity fails, and the guarantee lapses to the setting of Theorem 2. §4.3 reports the empirical ratio, which far exceeds the worst case.

3.4 Non-Exclusion Bias Parameter

Requirement R4 (§4.2) specifies a loss parameter $\lambda = C_{\mathrm{FN}}/C_{\mathrm{FP}}$, the ratio of the social cost of falsely excluding a genuinely DAPI-poor person to that of falsely including a non-poor one. We derive it decision-theoretically rather than asserting it: estimating both costs from Indian data (DALY burden of nutritional deficiency, the WHO-CHOICE convention, and representative central-scheme supplement levels) gives $\lambda \approx 7.6$–12.6, and we adopt $\lambda = 10$. Appendix F gives all cost components, sources, and caveats. Coverage and false-exclusion rates are insensitive to λ across 7–13, so the specific value matters less than the decision to impose the constraint at all.

4 Implications for AI-Assisted Welfare System Design

4.1 Three Design Changes When DAPI Replaces MPI

Replacing MPI with DAPI changes three things. **The eligibility question** moves from 'Is this household poor?' (binary, per-decade survey) to 'Which dimensions is this person deprived in, to what degree, and how is that changing?' (continuous, per-month), requiring a system that updates deprivation estimates continuously rather than a classifier assigning a fixed label. **The routing logic** moves from assigning a household to a BPL list and delivering all applicable benefits, to matching scheme capabilities against per-person dimension deficits and activating complementary schemes simultaneously. **The success metric** moves from headcount-ratio reduction to aggregate deprivation-score reduction $\sum_i \Delta_i$: continuous, decomposable, and resistant to the threshold manipulation §1.1 formalises. Delivering these changes imposes four minimum requirements.

4.2 Minimum Computational Requirements

R1: Dynamic Identity Layer. A unified identity linking each person's DAPI profile across all 805+ schemes. Aadhaar provides biometric identity with a 6.5% failure rate (Policy Circle 2026; Public Accounts Committee 2025) and no cross-scheme entitlement linkage; what is needed is federated identity with mandatory biometric fallback and portable entitlement registration.

R2: Real-time Deprivation Estimation. A model updating d_i continuously from satellite imagery (nutritional proxy), community health worker inputs, school administrative data, and transaction patterns. The raw infrastructure exists via ISRO, ABHA, UDISE+, and PMJDY; integration requires API development, data-quality validation, and cross-system legal authorisation, none of which currently exists.

R3: Cross-Scheme Optimisation Engine. A component solving the person-centric problem of §3.3.1. We propose, as a reference architecture rather than a fielded system, a multi-agent formulation in which each scheme s_j is wrapped as an agent exposing $\text{eligibility}_j(i) \in \{0,1\}$, $\text{leakage}_j(i) = l_{ij}$, and $\text{capacity}_j(t)$, the last an operational primitive absent from the idealised programme giving scheme j's remaining servable budget at cycle t. A coordinating router reads each d_i and solves the person-centric problem over these interfaces each cycle, with an audit interface exposing the R4 constraint for post hoc verification. No scheme need surrender its internal logic, only expose the interface. §4.3 evaluates this router in simulation; Appendix D gives the schema and allocation-cycle protocol.

R4: Non-Exclusion Bias Constraint. A hard constraint on the eligibility classification loss, weighting false negatives λ times more heavily than false positives, with λ derived in §3.4. The two levels should be encoded differently: the principle – that any algorithmic welfare-eligibility classifier must weight false exclusion more heavily than false inclusion – belongs in law, durable across model updates and governments, while the numerical value of λ is an empirical estimate that should be revisable through ordinary rule-making as cost data improves.

4.3 Empirical Validation on a Synthetic Population

We cannot evaluate the router on live administrative data, but the central claims – that person-centric routing dominates, and that the λ constraint buys non-exclusion cheaply – are testable on a population whose joint deprivation structure we control. We generate $N = 50{,}000$ synthetic citizens by thresholding a Gaussian copula calibrated so that each dimension reproduces its NFHS-5-anchored marginal rate (Table 2) while deprivations co-occur with the positive dependence documented by Alkire and Foster (2011). Nine schemes with documented leakage rates spanning PDS in-kind grain (0.28) to PMJDY direct transfer (0.02) each serve one dimension under a fixed budget share. Critically, scheme eligibility is screened on the *MPI proxy* $\widetilde{W}$ of §1.1, not on DAPI, which makes this a test of objective misspecification rather than of optimisation quality alone. The population is 57.8% DAPI-poor at $k = 0.20$ but only 18.3% MPI-poor, and the DAPI headcount is robust to the cutoff (65.8% at $k = 0.15$, 49.6% at $k = 0.25$, 39.1% at $k = 0.30$). We compare three allocators on an identical budget B, fixed at 30% of the cost of closing

every deprivation. The *scheme-centric* baseline is not a straw man: each scheme spends its full programme budget greedily, maximising beneficiaries served, which is the actual bureaucratic incentive; it is efficient given its information, and handicapped only by MPI screening and the absence of cross-scheme communication. The *DAPI router* is the cost-benefit greedy of Theorem 3 over all (person, dimension) slots; the third adds the $\lambda = 10$ non-exclusion term. Table 3 reports means over 20 seeds, all standard deviations below 0.25 pp; Appendix D gives the full specification.

Table 3. Synthetic allocation benchmark, $N = 50{,}000$, 20 seeds, budget fixed at 30% of total need. Deprivation reduction is the fall in population DAPI mass $\sum_i \sum_j w_j \, d_{ij}$; false exclusion is the share of DAPI-poor persons receiving zero net benefit.

Allocator	Deprivation red.	False excl.	Overlap waste
Scheme-centric greedy (status quo)	19.1%	40.0%	25.7%
DAPI router (Theorem 3 greedy)	41.7%	9.9%	0.0%
DAPI router, $\lambda = 10$	41.6%	0.0%	0.0%
LP relaxation bound	41.7%	n/a	n/a

Three findings bear on the argument. **First**, person-centric routing more than doubles the deprivation reduction obtained from the same budget (41.7% against 19.1%, a 118% relative gain), and eliminates the 25.7% of baseline expenditure that uncoordinated schemes duplicate onto already-served slots. **Second**, the $\lambda = 10$ constraint drives false exclusion among the DAPI-poor from 9.9% to 0.0% at a cost of 0.08 percentage points of aggregate welfare: at this budget, universal non-exclusion is nearly free. **Third**, and most consequentially, the baseline's false-exclusion rate is essentially *invariant to budget* – 46.7%, 42.8%, 40.3%, 39.8%, and 39.6% at budgets of 5%, 10%, 20%, 30%, and 40% of total need (Figure 1). Because eligibility is screened on a proxy that cannot see the deprivations these persons have, they are not reached at any budget. Exclusion under the status quo is an artefact of objective misspecification, not of scarcity, and is not remediable by increased spending.

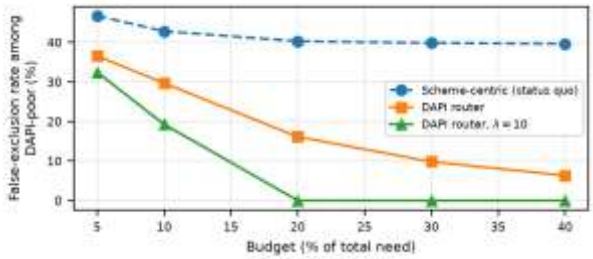

Figure 1. False-exclusion rate among the DAPI-poor against budget. The status-quo curve is near-flat: because eligibility is screened on the MPI proxy, the persons it cannot see are not reached at any budget. Appendix D adds the deprivation-reduction panel.

The greedy router attains 0.99999 of the LP relaxation bound (s.d. 4×10^{-6}), far exceeding the $1 - 1/e \approx 0.632$ worst-case guarantee, indicating that the hard instances driving the bound are not typical of deprivation structures resembling India's; we report this as evidence about this instance family, not a general claim.

5 Discussion

5.1 Comparison with Existing Poverty Measurement Approaches

DAPI differs from the Alkire-Foster MPI in four respects: continuous rather than binary scoring; entropy-derived rather than equal weights; person-level rather than household-level targeting; and a lower cutoff ($k = 0.20$ vs. $1/3$). It differs from income poverty lines (Government of India

2009; 2014) in capturing non-income deprivation and in resisting threshold manipulation. The trade-off against both is computational cost: continuous updating across 1.4 billion persons requires the infrastructure of R1–R4, which substantially exists in India's JAM Trinity and ISRO constellation, so the outstanding investment is in integration and governance rather than new data collection.

5.2 Limitations

First, DAPI weights remain dependent on dimension proxy choices, each of which happens to be the higher-deprivation-indicating NFHS-5 variable available (Appendix A); we have not systematically compared alternative proxies, and the 0.10 weight floor constrains weight allocation but not proxy selection. Relatedly, the literacy estimate in the DP tier (§3.1) is a scope-corrected illustrative figure, not a measured national rate, and requires validation against an ASER urban supplement or Census/NSS microdata before standing as a primary statistic.

Second, both routing results rest on structural assumptions about Δ_i: Theorem 1 assumes additive separability across dimensions and Theorem 3 per-dimension concavity, whereas in practice dimensions interact – a malnourished child performs worse in school – and some benefits are all-or-nothing. The general optimality-gap guarantee (Appendix E, Step 4) is restricted to the separable case, though the leakage-based Pareto-improvement argument (Step 3) is not; the $(1 - 1/e)$ guarantee lapses wherever concavity fails (§3.3.3). A model specifying Δ_i over joint dimension improvements is left to future work.

Third, DAPI is formulated over published district- and national-level aggregates rather than linked individual records. This is deliberate: a measure intended to hold the state accountable should be recomputable by any civil-society actor or oversight body from statutory public releases, without privileged data access or proprietary compute. The cost is real – the reported headcounts are dependency-free-bounded approximations rather than individually verified counts, we report no bootstrap standard errors on the weights, and tier boundaries remain fuzzy, since the confidence intervals on d_{ij} and the uncertainty band around $k = 0.20$ that an operational deployment would require cannot be constructed from aggregates. Validating DAPI against the NFHS-5 microdata release would address both and tighten the DP and AE anchors into genuine bounds.

Fourth, the benchmark of §4.3 is synthetic: its marginals are NFHS-5-anchored but its dependence structure, severity distribution, and per-scheme cost and capacity parameters are calibrated rather than observed. It supports the architectural claim that routing dominates screening, and the sharper claim that status-quo exclusion is budget-invariant under proxy-based eligibility, but not any forecast of rupees saved in a real programme.

5.3 The Political Economy of Definitional Choice

A candid account of DAPI's adoption barriers must acknowledge why the official poverty line sits where it does. Declaring DAPI the official standard would raise the headcount from 11.28% to roughly 58–69% overnight, and no elected government will voluntarily multiply its declared failure five- or six-fold. This is rational political economy, not hypocrisy. The workable response is to publish DAPI as a parallel shadow index maintained by an independent statutory body, creating accountability pressure without forcing a formal redefinition.

6 Conclusion

India cannot solve a problem it has not correctly stated. The official poverty figure of 11.28% is not wrong: it is a precise answer to a wrong question. The right question is what share of India's 1.4 billion people lack access to a dignified life, and the answer is between 58 and 69%, or 815–960 million people, derived using a dependency-free bounding methodology rather than unweighted survey summation. The gap between these two numbers is not a measurement error but an institutional design choice, perpetuated because accuracy is politically costly.

This paper provides five formal tools for addressing the misspecification. The proxy-versus-true-objective formulation (§1.1) names the failure precisely. The three-tier taxonomy and its

dependency-free bounding methodology establish that the undercount survives methodological scrutiny, and DAPI offers a computable, continuous, entropy-derived, sensitivity-tested alternative to MPI. Theorems 1–3 establish that person-centric routing strictly dominates, that computing it is NP-hard, and that a polynomial-time greedy router recovers a $(1 - 1/e)$ fraction of the optimum under a stated concavity condition; the synthetic benchmark shows the resulting gain is large and that status-quo exclusion does not shrink with budget. Correct measurement is necessary, though not sufficient, for AI-assisted welfare governance to target the right problem.

Ethics Statement

Algorithmic welfare targeting raises three ethical concerns that this paper addresses directly with self-contained mitigating principles. **First, potential for misuse.** DAPI scores could be used not only for welfare targeting but for surveillance, loan discrimination, or social credit scoring. The core mitigating principle is purpose limitation: welfare-deprivation data computed for DAPI should be legally prohibited from secondary use for credit, insurance, law-enforcement, or any non-welfare purpose, enforceable through specific statutory penalties rather than general data-protection principles alone.

Second, data privacy. DAPI computation at person level requires aggregating satellite, health, educational, and financial data, a privacy-sensitive aggregation regardless of downstream misuse. The core architectural mitigation is data minimisation through federation: individual-level source data should never be centralised in a single system; each contributing system (ISRO, ABHA, UDISE+, PMJDY) should compute and transmit only the minimum derived signal needed for its dimension of d_i, with cross-dimension aggregation into c_i performed using privacy-preserving computation rather than raw-data pooling. This is consistent with the data-minimisation principles of India's Digital Personal Data Protection Act (2023). A complete federated-systems specification is beyond this paper's scope, but the architectural constraint (no centralised raw-data pool) is binding on any implementation claiming to satisfy this framework.

Third, algorithmic discrimination. The entropy-based weighting method may assign low weight to dimensions where historically excluded populations (Dalit and Adivasi communities) show low variance because they are uniformly deprived, since the method of §3.2.2 weights by spatial dispersion. We address this through two mechanisms already specified in this paper: (a) the floor-weight constraint of §3.2.2 and Appendix A.3, guaranteeing $w_j \geq 0.10$ for every dimension regardless of its entropy-derived value; and (b) a requirement that DAPI coverage rates be audited for demographic parity across SC/ST/OBC/general categories within a pre-registered tolerance band (we suggest ±3 percentage points as a starting point for deliberation, subject to revision through the same regulatory process discussed for λ in §4.2), audited by a body independent of the welfare-disbursing agency. A fuller statutory and constitutional architecture for making these mitigations legally durable is pursued in dedicated future work; the principles stated here do not depend on that future work and are intended to stand as binding design requirements for any system built on this paper's framework.

Democratic accountability. Formalising welfare targeting as an optimisation problem does not remove human judgment: it makes human judgment explicit and auditable. The DAPI weights w_j are normative choices about the relative importance of nutritional adequacy versus functional literacy, and the non-exclusion parameter λ encodes a normative judgment about the relative harm of excluding a deserving person. These choices must be made through democratic deliberation, encoded in law, publicly visible, judicially challengeable, and democratically revisable. The algorithm implements the democratic choice; it does not substitute for it.

Reproducibility Statement

All DAPI calculations in this paper use publicly available aggregated data: NFHS-5 district-level tables (International Institute for Population Sciences); ASER 2023 district data (Pratham); FAO-SOFI 2025 country data; GHI 2025 country data; NHA 2021-22 state tables; NSAP programme

data. No individual-level data is used or released. The proofs of Theorems 1–3 (§3.3) require no data and are fully reproducible from the specifications given in Appendices E, G, and H.

The synthetic allocation benchmark of §4.3 and Appendix D.4 is generated by a single self-contained script of roughly 200 lines with no dependencies beyond NumPy and Matplotlib. It takes fixed integer seeds (0–19 for the main table, 0–4 for the budget sweep), reproduces every number reported in Table 3, Table 9, and Figure 1 exactly, and runs to completion in under a minute on a standard laptop CPU; no GPU, cluster, or proprietary solver is required. Appendix D.4 states the full generative specification – marginal rates, copula correlation blocks, severity transform, the nine scheme parameter tuples, and the three allocator definitions – in sufficient detail to reimplement the benchmark from the paper alone. That script, the DAPI computation code, the entropy weight derivation, and the sensitivity analysis (Appendix B) will be released under an MIT licence at an anonymised public code repository, following ICLR's anonymisation guidance for supplementary material. The district-level entropy weight computation (§3.2.2, Appendix A) is reproducible in principle from the cited NFHS-5 source but not from the aggregate figures in this paper alone, since district-level frequency values are not separately tabulated here (§5.2); we commit to releasing these alongside the code repository.

The Use of Large Language Models

The authors used AI writing assistance (Claude, Anthropic) to aid in drafting, structuring, and polishing sections of this manuscript. All claims, mathematical proofs, data citations, and intellectual contributions are the authors' own and were independently verified. The AI tool functioned analogously to a writing assistant. No AI-generated content was included without author review, verification, and explicit endorsement. The DAPI methodology, the proofs of Theorems 1–3, the entropy weight derivation, and the design, implementation, and analysis of the synthetic benchmark were conducted and verified by the authors independently of the AI tool.

Acknowledgements

Acknowledgements withheld for double-blind review.

References

Alkire, S., and Foster, J. 2011. Counting and multidimensional poverty measurement. Journal of Public Economics 95(7–8):476–487.

ASER Centre / Pratham. 2023. Annual Status of Education Report (Rural) 2023. New Delhi: Pratham Education Foundation. asercentre.org.

Atkinson, A. B. 2003. Multidimensional deprivation: Contrasting social welfare and counting approaches. Journal of Economic Inequality 1(1):51–65.

Banerjee, A., and Duflo, E. 2007. The economic lives of the poor. Journal of Economic Perspectives 21(1):141–167.

Bertram, M. Y.; Lauer, J. A.; De Joncheere, K.; Edejer, T.; Hutubessy, R.; Kieny, M.-P.; and Hill, S. R. 2016. Cost–effectiveness thresholds: pros and cons. Bulletin of the World Health Organization 94(12):925–930.

Centre for Monitoring Indian Economy (CMIE). 2024. Consumer Pyramids Household Survey, employment and poverty indicators. Mumbai: CMIE. cmie.com.

Chi, G.; Fang, H.; Chatterjee, S.; and Blumenstock, J. E. 2022. Microestimates of wealth for all low- and middle-income countries. Proceedings of the National Academy of Sciences 119(3):e2113658119.

Chowdhury, S. 2024. Technology is never neutral: Robodebt and a human rights analysis of automated decision-making. Australian Journal of Human Rights 30(1):20–40.

Deaton, A. 2010. Price indexes, inequality, and the measurement of world poverty. American Economic Review 100(1):5–34.

Decancq, K., and Lugo, M. A. 2013. Weights in multidimensional indices of wellbeing: An overview. Econometric Reviews 32(1):7–34.

Digital Personal Data Protection Act 2023. Ministry of Law and Justice, Government of India.

Eubanks, V. 2018. Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor. New York: St. Martin's Press.

FAO/IFAD/UNICEF/WFP/WHO. 2025. The State of Food Security and Nutrition in the World 2025. Rome: FAO. doi.org/10.4060/cd1973en.

Fisher, M. L.; Jaikumar, R.; and Van Wassenhove, L. N. 1986. A multiplier adjustment method for the generalized assignment problem. Management Science 32(9):1095–1103.

Foster, J.; Greer, J.; and Thorbecke, E. 1984. A class of decomposable poverty measures. Econometrica 52(3):761–766.

GHI / Welthungerhilfe, Concern Worldwide, IFPRI. 2025. Global Hunger Index 2025. globalhungerindex.org. Accessed May 2026.

Global Burden of Disease Collaborative Network (IHME). 2021. Global Burden of Disease Study 2021: India. Institute for Health Metrics and Evaluation, Seattle.

Goodhart, C. A. E. 1975. Problems of monetary management: The UK experience. In Papers in Monetary Economics. Reserve Bank of Australia.

Government of India, Planning Commission. 2009. Report of the Expert Group to Review the Methodology for Estimation of Poverty [Tendulkar Committee Report]. New Delhi: Planning Commission.

Government of India, Planning Commission. 2014. Report of the Expert Group to Review the Methodology for Measurement of Poverty [Rangarajan Committee Report]. New Delhi: Planning Commission.

ICRIER/IFPRI. 2024. PDS grain leakage study, August 2022–July 2023. Cited in: The Wire, November 2024.

International Institute for Population Sciences (IIPS). 2021. National Family Health Survey (NFHS-5), 2019–21: District-Level Factsheets. Mumbai: IIPS. iipsindia.ac.in.

International Labour Organization. 2024. India Labour Market Update 2024. Geneva: ILO Regional Office for Asia and the Pacific.

Jean, N.; Burke, M.; Xie, M.; Davis, W. M.; Lobell, D. B.; and Ermon, S. 2016. Combining satellite imagery and machine learning to predict poverty. Science 353(6301):790–794.

Krakovna, V.; Uesato, J.; Mikulik, V.; Martic, M.; Leike, J.; Kohli, P.; and Legg, S. 2020. Avoiding side effects in complex environments. Advances in Neural Information Processing Systems (NeurIPS) 33.

Ministry of Statistics and Programme Implementation (MoSPI), Government of India. 2025. National Accounts Statistics 2025. New Delhi: MoSPI.

National Health Accounts Technical Secretariat. 2023. National Health Accounts Estimates for India 2021–22. New Delhi: MoHFW. nhsrcindia.org.

National Social Assistance Programme (NSAP). 2025. Programme data and pension disbursement rates. New Delhi: Ministry of Rural Development, Government of India. nsap.nic.in.

NITI Aayog. 2024. India's Multidimensional Poverty Index 2024: Accelerated Progress 2013–14 to 2022–23. Discussion Paper, January 2024. niti.gov.in.

Nussbaum, M. 2001. Women and Human Development: The Capabilities Approach. Cambridge: Cambridge University Press.

Policy Circle. 2026. Aadhaar authentication failures: 20.3 million per month at 6.5% failure rate. April 2026. policycircle.org.

Public Accounts Committee, Parliament of India. 2025. Report on Aadhaar Authentication Failures in Welfare Delivery. New Delhi: Lok Sabha Secretariat, July 2025.

Ravallion, M. 2011. Mashup indices of development. World Bank Research Observer 27(1):1–32.

Saxena, D.; Kawakami, J.; Weller, J.; Zhu, H.; and Ramchurn, S. D. 2020. Conducting ethical algorithmic decision-making in welfare services. In Proceedings of the AAAI-20 Workshop on AI for Social Impact.

Sen, A. 1999. Development as Freedom. New York: Anchor Books.

Socio-Economic and Caste Census (SECC) 2011. Ministry of Rural Development, Government of India. secc.gov.in.

Strathern, M. 1997. Improving ratings: Audit in the British university system. European Review 5(3):305–321.

Sviridenko, M. 2004. A note on maximizing a submodular set function subject to a knapsack constraint. Operations Research Letters 32(1):41–43.

Watts, H. W. 1968. An economic definition of poverty. In D. P. Moynihan, ed., On Understanding Poverty. New York: Basic Books.

Yeh, C.; Perez, A.; Driscoll, A.; Azzari, G.; Tang, Z.; Lobell, D.; Ermon, S.; and Burke, M. 2020. Using publicly available satellite imagery and deep learning to understand economic well-being in Africa. Nature Communications 11(1):2580.

Zeleny, M. 1982. Multiple Criteria Decision Making. New York: McGraw-Hill.