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.

What a voluntary health pool needs from its members: a formation bound, two actuarial floors and a pre-registered persistence test in 141 Indian self- help groups

David M. Dror

daviddror@socialre.org

Principal Adviser, Health Investment & Financing Corp

Research Article

Keywords: Community-based health insurance, voluntary health insurance, informal sector, sustainability, persistence, trust, consensus, dual benefit, India

Posted Date: November 25th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11403134/v1

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

Additional Declarations: No competing interests reported.

Abstract

Background

Ghana made health insurance compulsory in 2012. By 2026, 44% of residents were not members, in part because the law cannot reach people without a payslip. For them, coverage depends on a voluntary pool. Research yields lists, not thresholds. This paper measures two actuarial floors and the lowest members' benefit at which pools formed, and tests whether that benefit predicted persistence.

Methods

Analysis of scheme records, a three-round household panel and two pricing spreadsheets (RECORD). The members' score is the product of trust, agreement and benefit. Only benefit is measured here, from households' baseline illness burden; the score is bounded, not computed. Two floors are reproduced from the spreadsheets (India 2010; Bangladesh 2011); a formation bound is read from 141 Indian self-help groups. A pre-registered test, written after the survival of 42 groups was known, asks whether baseline benefit predicted later participation in the 86 groups that formed a pool.

Results

The Indian stand-alone scheme needed 271.93 households to cover costs, 19 times the largest group served. The Bangladesh loading of 7.9858% is reproduced at 10% ruin probability. Among the 86, the smallest benefit level was 0.056, with 0.090 next; because trust and agreement cannot exceed 1, the minimum score at formation is at most 0.056. Fifty of the 86 kept a member enrolled in both later rounds; 20% of the cohort observed at endline was enrolled. Benefit did not detectably predict participation (within-partner correlation 0.11, one-sided permutation p = 0.15; 80% power at 0.27). Persistence rose across benefit terciles (43%, 56%, 76%), but the terciles align with partner; a post hoc within-partner trend is not significant. No baseline member of the 55 groups without a pool at baseline was enrolled later.

Conclusions

The condition is necessary, not sufficient; the bound is an upper limit on the minimum. Pools formed at benefit as low as 0.056, and 36 of the 86 did not persist. A scheme can compute the floors from a costing and the bound from an enrolment register before money moves, and see which limit binds.

Background

In 2012 Ghana made health insurance compulsory for every resident [1]. By 2026, fourteen years and many registration drives, identity-card linkages and premium exemptions later, 56% of residents were

active members and 44% were not [2, 3]. The law reaches people with a payslip, because their contribution can be deducted at source. The law does not reach farmers, traders and home workers, because there is no payslip to deduct from, and no penalty follows if they stay out [4]. For these households, being covered does not follow from the law. Coverage depends on whether a voluntary insurance pool exists where they live and keeps going from one year to the next. This paper is about what such a pool needs from the members in order to form and to keep going, and how a scheme can measure whether the pool has enough.

Governments and funders usually try to make a pool survive by putting more people into the pool. They make membership compulsory where enforcement is possible, subsidise premiums where enforcement is not, and set premiums so that what the pool collects covers what the pool pays out. None of these asks the question that decides the outcome: will a household pay, pay again next year, and receive something worth what the household paid? Researchers have asked that question, and their answer is a list. A systematic review of 51 studies from 22 countries found that trust in the insurer, a sense of ownership, the perceived value of the benefit and the size of the membership all matter. The review concluded that the field needs validated tools [5]. A second review, of 27 studies of renewal, found that people leave when they get no benefit, do not trust the management, or feel excluded. That review named a trade-off between equity and efficiency, without offering a way to measure where a scheme stands on the trade-off [6]. The one numerical threshold either review reports is a rule imposed by a scheme. In Uganda, 60% of a community had to enrol before services began, and both managers and members saw the rule as a threat to the scheme's survival [5]. A list of influences does not say how much of anything is enough. A threshold set by a designer is a barrier, not a measurement.

This paper measures two things a voluntary health pool needs, from records a scheme already holds, and tests a third. The first is two actuarial floors: how many households must join for premiums to cover costs, and what safety loading keeps the probability of ruin at a tolerated level. The second is a formation bound: the lowest level of members' benefit at which groups formed a pool, which bounds from above the minimum score a pool needs to form. The third is a pre-registered test of whether members' benefit at formation predicted that the pool persisted two survey rounds later. The evidence comes from India: 141 women's self-help groups in Bihar and Uttar Pradesh that were offered health insurance and were followed through three survey rounds, and two pricing spreadsheets recovered from scheme designers. A companion study of 169 community health insurance schemes in 38 countries, now published, relates survival over ten years to the three conditions [7]. The three together explain 39% of the differences in survival between countries, against 13% for how the schemes were financed. Within the three, agreement and benefit carry the explained variance; adding trust changes the fit by an amount not distinguishable from zero. That study finds that two of the conditions are associated with survival over years, and leaves trust open. The present paper asks how much of the benefit condition was present where pools formed, and whether the level of benefit at formation told the pool's fate two rounds on.

Methods

Design and definitions

This is an analysis of records that schemes already keep, a household panel collected under a research programme, and two pricing instruments. No new data were collected. Reporting follows the RECORD extension of the STROBE statement [8] (Additional file 1). The group records and the pricing instruments are records kept in the course of operations, which is what RECORD covers; the household panel is reported under the STROBE items the checklist carries.

We use the word sustainability in one sense only: a scheme is sustainable over a period of years if the pool survives every one of them. A year is the natural period, because membership is renewed yearly. A pool survives a year if four things are true of the pool, and all four can be read from the scheme's own records. Cover is in force and members can renew. Contributions are collected. Claims are paid out of the common fund. And the fund is still the members' shared property at the end of the year: not exhausted, not split into individual accounts, and not taken over by an outside body. The panel used below records the first of the four conditions, cover in force and renewed by members, at two survey rounds after formation. We call what the panel measures persistence, and we reserve sustainability for survival over years with all four conditions holding. The relation between the two is the subject of a separate theoretical paper. A pool that clears the floor in every year survives the period, so the yearly condition is sufficient for survival when no entrants and no reserve carry a bad year. The condition is not necessary where a reserve does carry one.

Two further points follow from the definition. First, the number of members at one moment is not sustainability. A published costing of a prototype plan shows why [9, Table 5]. At a premium of US$4.48 per person per year with three or fewer insured persons per household, no number of members makes the pool viable. At US$7.47 with four insured per household, 1,600 households are enough. Second, the group records observe formation once. They show that a group formed a pool, which is the first of the four conditions. The panel then shows whether the group's members were still enrolled at the two later rounds. Every statement made below says which of the two the statement rests on.

Table 5
The pre-registered persistence test, 86 groups that formed a pool, baseline to endline 2013

Quantity	Value	Note
Pre-registration	28 September 2026	Analysis plan on partly known data: the survival outcome of 42 of the 86 groups had been read in the 43-group test; benefit had not been related to any post-baseline outcome. Files fixed by checksum; seed 20260928
Sample	86 groups	At least one household enrolled at baseline
Persisted; lapsed	50; 36	Persisted: at least one cohort household enrolled at midline and at endline. 31 groups had two or more at endline; 9 had six or more
Cohort enrolled at endline, 86 groups	147 of 749 observed (20%)	525 enrolled at baseline
Persisted by field partner (mean benefit)	BAIF 28 of 36 (0.263); NIDAN 16 of 35 (0.184); Shramik Bharti 6 of 15 (0.206)	
H1, primary: correlation of the endline participation share with benefit centred on the partner mean	0.11	One-sided permutation p = 0.15 (10,000 permutations, seed 20260928). Fisher z interval about − 0.10 to + 0.31 and 80% power at r = 0.27, both post hoc. Not supported
H1b, secondary: same, outcome persisted	0.10	p = 0.16. Not supported
H2, persistence limit: smallest benefit among persisting pools	0.0556	Group 3052, NIDAN; next 0.090, 0.090; holds for lines 1 to 6; 0.16 at 7; 0.27 at 8; none at 9 or 10. Formation bound 0.0556
H3: persisted by benefit tercile (low, middle, high)	13 of 30; 15 of 27; 22 of 29	43%, 56%, 76%. Mean endline participation share 0.18, 0.14, 0.27. Cochran-Armitage trend (post hoc, one-sided): unstratified z = 2.53, p = 0.006; stratified by partner z = 1.40, p = 0.08
Terciles by partner: persisted of groups (low; middle; high)	BAIF 2 of 5; 8 of 11; 18 of 20. NIDAN 9 of 19; 5 of 11; 2 of 5. Shramik Bharti 2 of 6; 2 of 5; 2 of 4	Tercile cut points 0.186 and 0.246 on the pooled 86
D1: groups enrolling nobody at baseline with a cohort household enrolled later	0 of 55	No late formers on the cohort definition, so the pre-registered smallest benefit among them is undefined; smallest benefit among the 55 non-formers 0.075. Three of the 55 had enrolled entrant households at a later round (1 and 1 at midline, 2 at endline in one of them; 5 at midline)
Sensitivity (a) to (g), declared and not decisive	(a), (b), (c), (e), (g): r 0.09 to 0.16; p 0.07 to 0.20	(e) unstratified r = 0.16, p = 0.07; (f) regression coefficient on benefit 0.39, SE 0.39, p = 0.16; (d)

Quantity	Value	Note
		midline trust vs endline participation, 69 groups, Spearman − 0.01

The three conditions and the score

The paper's unit of account is the members' score, built from three conditions. Trust: whether they trust the body that will pay their claims. Agreement: whether the group has settled what the pool is for, measured from the record of the group's deliberation. Benefit: whether the pool is worth joining to each household and to the group at once, measured from the illness households actually face. We combine the three into one number, which we call the members' score, by multiplying them. Each lies between 0 and 1, so the score does too. None of the three uses survival, enrolment or renewal in the definition. That is deliberate. A measure that included the outcome could not fail to predict the outcome, and a measure that can fail is the only kind worth reporting. Multiplication, rather than addition, expresses the claim that the three conditions are all needed. Three shortfalls of 20% each leave a score of $0.8 \times 0.8 \times 0.8 = 0.51$, about half of the maximum, even though each condition taken alone would pass an audit. The reasons for the multiplicative form are set out in a companion paper (Dror and Duru, unpublished; deposit [10]); nothing reported here depends on that derivation. Where the year is clear from the context we write the score simply as S. In this paper only benefit is measured, for all 141 groups. Trust is reported descriptively from one panel item, and agreement is not used; the reasons are given below. The score is bounded from above and never computed, except in the 43-group test, where two of the three factors are constants.

Two quantities complete the definition. The first is the pool's cost floor: the share of households in the catchment that must join for the premiums collected to cover the pool's costs. The floor comes from the scheme's pricing workbook. The second is the minimum score, written S*: the lowest members' score at which enough households join to reach that floor. Both are properties of how the pool is designed, and neither enters the measurement of trust, agreement or benefit. The claim the paper works with is a necessary condition. A pool whose members' score is below the minimum erodes. A pool whose score is at or above the minimum can form and can persist; the condition does not promise that the pool will. These data cannot refute the claim, and the paper does not say otherwise. The score is never computed here, only bounded from above by benefit, so no group can be shown to sit below the minimum. What the data give is an upper bound on the minimum, read from the groups that formed, and a test of whether the one measured condition predicted persistence among them. Pools that fail above the minimum do not refute a necessary condition. The rate at which they fail is reported here as a fact for the companion theory to account for.

Two results borrowed from a companion paper

Two results from a companion paper are used here and are restated so that the argument can be followed without the companion (Dror and Duru, unpublished; deposit [11]). Neither is used to compute any number in this paper. The first result says that the minimum score is a tipping point rather than a grade. Above the minimum, benefits paid build trust, trust brings members in, and their contributions fund more benefits, so membership is self-sustaining. Below the minimum the same loop runs backwards. Stated that way, the first result is a claim about sufficiency, and the 43-group sample reported in Results bears on that reading. Every group there had a score of at least 0.571 under that test's benefit construct, far above any minimum of at most 0.056, and 12 of the 43 pools failed. Under this paper's benefit measure the scores of those groups are unknown and at most 0.46. The companion theory has to say what rate of failure above the minimum the loop allows. This paper reports the rate and uses the first result in no computation. The second result concerns the kind of pool. If the minimum score were the same for every kind, one could not observe a stand-alone scheme needing most of the catchment and a self-help group's pool inside a shared scheme forming at a very low score. Both are observed, so the minimum depends on how the pool is organised.

Which constraint binds

Every scheme faces three limits, and each can be expressed as a share of the population the scheme is meant to cover. How many people can the law compel to join? How many must join before the design covers costs? How many join, and keep paying, by choice? Each limit has a data source of that limit's own, so a scheme's position against each can be measured rather than assumed. The three apply in turn. The legal limit divides the population into the part a mandate reaches and the part the mandate does not. The cost floor sets the share that both parts together must reach. The behavioural limit decides whether the voluntary part reaches that share. The law and the sponsor's money only work while they are maintained: enforcement must be funded, a subsidy must be renewed. The members' choice needs no upkeep, and that choice is the limit that binds wherever the other two stop. Table 1 sets out the three limits, what each measures, how close a scheme is to breaching each, and where the data come from.

Table 1
The three bounds, their slacks and their instruments

Bound	What the bound measures	The slack: distance to violation	Instrument	In this paper
Legal	The participation the state can compel in practice	The enforcement gap: the share of the covered population actually compellable	Compliance and coverage data	Not measured here
Actuarial	The participation or funding the pool's economics require	The distance of participation from the floor	The scheme's own cost structure	Results: two floors
Behavioural	The participation people choose	The voluntary margin: chosen participation against the same floor	Records of voluntary pools that formed and persisted	Results: formation bound; persistence test

This paper is about pools whose members join by choice, make their own rules, and can leave at any renewal. Their common resource is a fund that members' contributions build up and members' claims draw down. Such a pool does not need to make a profit; the pool needs the fund to exceed what the fund owes. The prototype costed in [9] is viable when running costs stay at or below 20% of premium income and claims settle at about 70%. Two rules place a real scheme in this class. Is there a pool, and is joining a choice? And is the law enforced in practice? A scheme that is compulsory on paper but cannot collect from people without payslips is voluntary in fact. Ghana is the documented case [4].

Data

Group records. The main dataset is 141 women's self-help groups in Bihar and Uttar Pradesh, India, offered community health insurance under a European Commission research programme (grant HEALTH-F2-2009-223518). The Micro Insurance Academy collected the records between 2010 and 2013, and the records are deposited with the companion paper [11]. All 141 have the benefit measure on which the formation bound rests. Fifty-five enrolled nobody at the offer. Eighty-six enrolled at least one household, and these are the groups that formed a pool. The files are data/pillar1_EU_FP7/group_table_v1_2.csv and household_baseline_extract.csv in the deposit, version 1.2.0, where every group-level number used below is rebuilt from the baseline round of the household panel. Column S of the deposited group table is the baseline enrolment share, enrolled divided by size, and not the members' score. The pools sat inside one scheme per site. Under the programme's protocol each implementing partner set up one community-based health insurance scheme among the self-help groups in the partner's area. Groups affiliated to that scheme. Five levels of scheme staff, among them a claims committee, were elected from group members and partner staff. A scheme office at a focal point in the area ran premium collection and claims processing [12]. A group's pool is therefore a unit inside a scheme with shared administration. That is what the Discussion means by a federated micro-pool.

Household panel. The same programme surveyed the member households of the 141 groups in three rounds: baseline, a midline round in 2012 and an endline round in 2013. Each round records, for each household, the household's group code and whether the household was enrolled. The baseline members of the 141 groups number 1,335. Of these, 1,070 were observed at both later rounds, and at least one baseline member of every group was observed at endline. At midline 133 baseline members, and at endline 106, carried a group code different from the baseline code; 177 households at midline and 152 at endline appear in a group only after baseline. These four counts come from the panel file and are recorded by group in the inventory group table and report in the replication package; the deposited group table of the test carries only the enrolled subsets of them. The panel file is hh_master.sav (MD5 93d831f9fd239b2e7d40b825c0815c02); the baseline extract and the group table carry MD5 b4a6236a6a540764bb42c0f295149a60 and ba3e1f84f7aead599addce3b3707bc93.

Pricing instruments. Two spreadsheets built by scheme designers, recovered and traced cell by cell in the deposit. The first was built in 2010 for a stand-alone scheme in Pratapgarh district, Uttar Pradesh. The cells record a fixed cost of 108,500, a premium of 1,380 per household per year, an administration share of 30% chosen by the designers, and a variable cost of 15 per household. The 50 BAIF self-help groups the scheme served had at most 14 member households each; the largest group of any partner has 17 (Table 2). The second spreadsheet is a 2011 premium calculator from Bangladesh. The calculator records expected claims of 300.01 per person per year, a variance of 349,484, a community of 1,000 people, and a safety loading of 7.9858% on the premium. Both spreadsheets were built before the definitions above existed.

Table 2
Group records behind the formation bound

Groups	All groups	Enrolled no one	Formed a pool (enrolled ≥ 1)
	141	**55**	**86**
BAIF, Uttar Pradesh	50	14	36
NIDAN, Bihar	49	14	35
Shramik Bharti, Uttar Pradesh	42	27	15
Group size, households (min, median, max)	1, 10, 17	–	–
Share enrolled among groups that formed a pool (min, median, max)	–	–	0.071, 0.667, 1.000
Groups enrolling every household	–	–	7
Trust item at midline, groups with at least five respondents (min, max; 69 groups)	–	–	0.500, 1.000

Every group has the benefit measure on which the formation bound rests. The trust item is the share of baseline members who agree, midline round.

Survival sample. Ninety-six of the self-help groups chose their benefit package through a structured group deliberation, Choosing Healthplans All Together (CHAT) [13]. These groups were matched to the household survey by group code. Forty-three met every condition set out in advance for a one-year survival test, and their survival from the 2012 round to the 2013 round is recorded. The test on this sample was written down and deposited, with checksums, before any data file was opened. The result is reported here from the report of record of 9 September 2026. The procedures, the report and the group table of that test are in this paper's replication package and in deposit [11]. Two features of the sample were fixed before the data were opened. Agreement is 1 in every group, because each deliberation ends with one package chosen. And the benefit measure of that test is not the episode-burden measure used everywhere else in this paper. The companion procedures (SStar_Measurement_Procedures_v1_0.docx, in the scripts folder of deposit [11]) scored benefit from scheme documents: five institutional features, each scored 0 or 1, averaged with equal weights, once per scheme. The five are member share in rule-setting, transparency of grievance and benefit delivery, frequency of member assemblies, accommodation of heterogeneity, and non-extracted surplus. Both sites scored 1 on all five, so benefit was 1.000 at both sites and constant by construction. The test could therefore only move along the trust dimension. The 43 groups are among the 141. Forty-two of them are among the 86 that formed a pool at baseline, and for those 42 the survival outcome read in that test is the persistence outcome of the test below.

How each quantity is measured

Trust. The household panel carries one trust item, asked from the midline round: "I trust the ability of us people to arrange insurance protection by ours[elves]", answered agree, disagree or don't know. A group's value is the share of the group's baseline members who agree, among those who answered agree or disagree. The item measures confidence in the community's own capacity rather than trust in a named paying body, and the item does not enter the formation bound or the persistence test. The channel-specific question and the validation of that question are in the deposited package of the companion trust study (Dror, unpublished; [14]).

Agreement. Read from the record of the group's deliberation. Agreement is not used in computing the formation bound or as a predictor in the persistence test, and the reason is a finding in that reason's own right. The usual statistical measure of agreement, when applied to the decision to enrol, gives the same reading for unanimous rejection as for unanimous acceptance. In this dataset the measure scores all 55 groups that enrolled nobody at exactly 1.000. A bound that only requires agreement to be at most 1 is unaffected by this defect. A test that used agreement as a predictor of enrolment would predict enrolment by construction. The defect is analysed in a companion paper (Dror, unpublished; deposit [15]).

Benefit. The burden of illness episodes a household carries at baseline, summed over the household's members, capped at 10 episodes and scaled to lie between 0 and 1. A group's value is the mean over the group's baseline members. The measure is built from household circumstances and from no outcome. A

second version, built from borrowing for health and reported in earlier versions of the deposit, is withdrawn because that version could not be reproduced from the panel.

The cost floor. The number of households at which premiums cover costs is the fixed cost divided by each household's contribution to overhead. In symbols, n* = F/(πθ − v), where F is the fixed cost, π the premium, θ the administration share and v the variable cost per household. Dividing n* by the catchment gives the floor as a share. **The risk floor.** The safety loading that keeps the probability of ruin at a tolerated level is λ(n) = z·d/(π√n) [16]. Here d is the standard deviation of claims, n the pool size and z the normal quantile for the tolerated probability. Solving for n at the loading the design will bear gives the pool size below which the risk margin alone breaks the premium. Both formulas are reproduced from the spreadsheets' own cells, and the reproduction is reported in Results.

The formation bound. Every group that formed a pool must have had a score at least equal to the minimum. Since trust and agreement are each at most 1, a group's benefit is at least the group's score. Write L_form for the smallest benefit among the groups that formed a pool. Then every group that formed had a score at most L_form at the lowest, so the minimum score needed to form is at most L_form. The bound needs no measurement of trust. Where the line between forming a pool and not forming one should be drawn is debatable, so the bound is reported for every line from one enrolled household to ten.

The persistence test

The persistence test was pre-registered on 28 September 2026. The pre-registration states what was already known, and the paper repeats that statement here. Baseline enrolment of every group was published. The 43-group test of 9 September 2026 had read midline and endline enrolment for 43 groups, reported 12 failures and summarised renewal by trust level. Forty-two of those groups are among the 86 tested here, so the persistence outcome of 42 of the 86 groups (31 persisted, 11 lapsed) had been read before the test was written. Renewal patterns in this panel were also published in 2016 [17]. Two things had not been read: the relation of benefit to any post-baseline outcome, and the endline participation of the other 44 formed pools. The pre-registration is of an analysis plan on partly known data, not a blind test, and the word pre-registered is used in this paper in that sense. The pre-registration fixes the data files by checksum, the sample, the outcome, the predictor, the strata, the statistic, the decision rule and the sensitivity analyses. The analysis script runs only against the pre-registration file as saved and stops on any other checksum. The pre-registration, the script, the report of record and the group table without names are in the paper's replication package.

Sample. The 86 groups with at least one household enrolled at baseline. **Cohort.** The baseline member households of a group, followed by household number whatever group code they carry later. Households that enter a group after baseline are excluded from the primary analysis. **Primary outcome.** The endline participation share: cohort households enrolled at endline, divided by cohort households observed at endline. **Secondary outcome.** Whether the pool persisted: at least one cohort household enrolled at

midline and at least one at endline; otherwise the pool lapsed. This is the one-year survival rule applied to every formed pool. **Predictor.** Benefit at baseline, the episode-burden measure of the deposit, built from household circumstances and from no outcome. **Strata.** Field partner: BAIF, NIDAN, Shramik Bharti.

Hypotheses and rules. H1, primary: endline participation rises with benefit within field partner. The outcome is permuted within partner 10,000 times (seed 20260928); the statistic is the correlation of the raw outcome with benefit centred on the partner mean; the test is one-sided. The outcome is not centred. The permutation p does not depend on that choice, because only the within-partner covariance varies under permutation. H1 is supported if $p < 0.05$ and the correlation is positive; otherwise H1 is not supported, and that is reported as a failure of the paper's persistence prediction. H1b, secondary: the same test on whether the pool persisted, declared not testable if fewer than 8 pools persisted or fewer than 8 lapsed. H2, the persistence limit: trust and agreement are each at most 1, so a pool's score is at most the pool's benefit, and the score needed to persist is at most the smallest benefit among the pools that persisted. The limit is reported with the group that sets the limit, at each line k from 1 to 10 cohort households enrolled at endline, beside the formation bound. H3 and D1, descriptive. H3: persistence and endline participation by tercile of benefit among the formed pools. D1: among the 55 groups that enrolled nobody at baseline, the number with a cohort household enrolled at midline or endline, with the smallest benefit among any such late formers. Seven sensitivity analyses were declared and are not decisive. (a) Groups with at least half the cohort observed at both later rounds. (b) Only cohort households with an unchanged group code. (c) Including households that entered after baseline: enrolled entrants are added to the numerator and the denominator stays the cohort observed at endline, so this share is not bounded by 1. (d) The midline trust item, in groups with at least five respondents, against endline participation. (e) The test without partner strata. (f) A linear regression of endline participation on benefit with partner effects. (g) Endline participation net of the baseline enrolment share and partner.

What the test can and cannot show. Enrolment in a round shows that a household took cover in that round. The panel therefore measures the first of the four conditions of survival, cover in force and renewed, group by group, over two rounds after baseline. Claims paid out of the common fund and the fund's standing at year end are scheme-level facts and are not in the panel. Persistence here means that a group's members kept taking cover in the scheme's pool. The design is observational. An association between benefit and persistence is not a causal estimate. With 86 groups the one-sided test has 80% power to detect a correlation of about 0.27 (Fisher z approximation); a smaller true association is detected with lower probability. The interval on the correlation reported in Results is a Fisher z interval on the observed value. Both figures were computed after the test, outside the pre-registration, and both are approximations for a stratified permutation design.

What the design can and cannot show

The design supports bounds read from records, and associations. The two floors are arithmetic on a scheme's own cells and support nothing more. The formation bound is a necessary condition, and the

bound depends on the multiplicative form. If a pool must have a score of at least S* to form, then no pool that formed sits below S*, and the smallest of them bounds S* from above. The bound does not predict which groups will form a pool. The persistence limit has the same logic applied to the pools that persisted, and the persisting pools are a subset of the pools that formed, so the persistence limit cannot sit below the formation bound. The design does not support three further things. The design gives no causal estimate of how much any one condition raises persistence. The design does not generalise beyond one population, because the floors come from one scheme and one calculator, and the bound and the test from one project in two Indian states. And the 43-group survival test, as Results explains, could not vary two of the three conditions and so speaks to the trust question rather than to the score.

Software

Software tools, including a large-language-model assistant, were used to check the code and the internal consistency of the tables and to edit language. All analytic decisions, interpretations and conclusions are the author's, who takes full responsibility for the content. No other person contributed to the analysis or the text.

Results

The cost floor

The Pratapgarh spreadsheet sets the spreadsheet's own floor. Each household contributes 1,380 × 0.30 − 15 = 399 towards overhead. Dividing the fixed cost of 108,500 by 399 gives 271.93 households, which is the number the spreadsheet reports in the cell for households to be insured. The 50 BAIF self-help groups the scheme served had at most 14 member households each, so a pool confined to the largest of them would need 19.4 times that group's membership (Table 3). No participation rate between 0 and 1 reaches that floor in any single group; the design needs a pool drawn across many groups.

Table 3
The two floors and the formation bound, each in its own units

Bound	Value	Axis and source
Cost floor, Pratapgarh 2010	n* = 271.93 households	Households. Workbook cells: fixed cost 108,500; premium 1,380; administration share 0.30; variable cost 15
Required participation, largest BAIF self-help group served (14 member households)	19.4	Share of membership. No participation rate in [0, 1] reaches the floor in any single group
Risk floor, Bangladesh 2011	Loading 7.9858% at n = 1,000, ruin probability 10%	Loading on premium. Calculator cells: expected claims 300.01; variance 349,484; community 1,000
Formation bound, L_form = min benefit over groups that formed	0.0556	Score on [0, 1]. Group 3052, NIDAN, 7 of 9 enrolled. Holds for constitution lines 1 to 7; 0.090 at 8; 0.094 at 9 and 10
Next-lowest values among groups that formed a pool	0.090, 0.090	Score on [0, 1]

The risk floor

The Bangladesh calculator sets a floor of a different kind, the reserve against ruin [16]. The cells give expected claims of 300.01 per person per year, a variance of 349,484 and a community of 1,000. The safety loading that keeps the probability of ruin at 10% is 1.28155 × 591.172 / (300.01 × 31.623) = 0.07986, or 7.986%, which reproduces the calculator's 7.9858% to four significant figures. That is the figure the calculator records. The authors of the calculator did not write down the formula; the formula recovers their number and, with the number, the level of risk they accepted without naming the level. Solving the same expression for the pool size at the loading a design will bear gives the number of members below which the risk margin alone makes the premium unaffordable.

The formation bound

The two floors and the bound are measured in different units. The floors are shares of households; the bound is a score between 0 and 1. They bound the same threshold from opposite sides and are reported separately, not combined.

All 141 groups have the benefit measure. Fifty-five enrolled nobody. Eighty-six enrolled at least one household and so formed a pool (Table 2). Among these 86, the share of the group that joined runs from 7% to 100%, with a median of 67%. Seven groups enrolled every household. Among these 86 the groups have between 5 and 17 households, with a median of 10; across all 141 the range is 1 to 17 (Table 2). The trust item runs from 0.500 to 1.000 across the 69 of these groups in which at least five baseline members answered the item.

The smallest benefit level among the 86 is 0.0556, set by a NIDAN group (code 3052) that enrolled seven of the group's nine households. The group's nine households reported five illness episodes at baseline,

one each in five households and none in the other four; five episodes over a possible 90 give 0.0556. So the minimum score needed to form a pool cannot be higher than about 0.056: a group formed a pool with a benefit level no higher than that. The bound is not set by a marginal group. The next-lowest values among the 86 are 0.090 and 0.090. Each group's benefit is a mean over a few households measured with survey error, so the minimum of 86 such means tends to fall below the minimum of the true group values. The first three order statistics, 0.056, 0.090 and 0.090, are the reading a cautious reader should carry: a bound of about 0.06 to 0.09, with 0.056 the literal minimum. Where exactly a pool begins is debatable, so the bound is reported for every line from one enrolled household to ten. The bound stays at 0.0556 for every line from one to seven, rises to 0.090 at eight, and to 0.094 at nine and ten. The order of magnitude does not change with the line.

The 55 groups that enrolled nobody were not below the bound: all 55 have a benefit level at or above 0.0556, and the lowest among them is 0.075. The bound neither predicts nor forbids this. The bound is an upper limit by construction: the bound says what every group that formed a pool had at least, and no enrolment pattern among the 55 could contradict that. The bound does not say that a group above the bound will form one.

Persistence

Of the 86 groups that formed a pool, 50 had at least one cohort household enrolled at midline and at least one at endline, and 36 did not (Table 5). The rule is lenient, and the table shows how lenient. At endline 147 of the 749 cohort households observed in the 86 groups were enrolled, 20%, against 525 enrolled at baseline. Thirty-one of the 86 groups had two or more cohort households enrolled at endline, and nine had six or more. Persistence differed by field partner: 28 of 36 BAIF groups, 16 of 35 NIDAN groups and 6 of 15 Shramik Bharti groups, with mean benefit 0.263, 0.184 and 0.206. The primary hypothesis was not supported. The correlation of the endline participation share with benefit centred on the partner mean was 0.11; the one-sided permutation p was 0.15 (10,000 permutations, seed 20260928). The 95% interval on the correlation runs from about − 0.10 to + 0.31. On the secondary outcome, whether the pool persisted, the correlation was 0.10 and p was 0.16. Five sensitivity analyses that re-run the benefit correlation, (a), (b), (c), (e) and (g), give 0.09 to 0.16 with one-sided p from 0.07 to 0.20; the unstratified test (e) comes nearest to the threshold (0.16, p = 0.07). The regression (f) gives a benefit coefficient of 0.39 (standard error 0.39, one-sided p 0.16). The trust analysis (d) gives Spearman − 0.01 across 69 groups. None overturns the verdict, and none was pre-registered to do so.

The descriptive results are consistent with a gradient the test could not resolve, and with no more than that. Persistence by tercile of benefit runs 13 of 30 (43%) in the lowest tercile, 15 of 27 (56%) in the middle and 22 of 29 (76%) in the highest. The mean endline participation share by tercile is not monotone: 0.18, 0.14 and 0.27. A Cochran-Armitage trend test across the three terciles, without strata, gives z = 2.53, one-sided p = 0.006. The terciles align with field partner: 20 of the 29 high-tercile groups are BAIF, whose pools persisted at 78%, and 19 of the 30 low-tercile groups are NIDAN, whose pools persisted at 46%. The same trend test stratified by partner gives z = 1.40, one-sided p = 0.08. Within BAIF the terciles run 2 of 5, 8 of 11 and 18 of 20. Within NIDAN the pattern reverses: 9 of 19, 5 of 11 and 2 of

5. Within Shramik Bharti the three terciles give 2 of 6, 2 of 5 and 2 of 4. Composition of the three field partners is therefore a candidate explanation, and the test cannot separate the two. Neither trend test was pre-registered.

The persistence limit, the smallest benefit among the 50 pools that persisted, is 0.0556, from the same group 3052 that sets the formation bound. The limit holds for every line from one to six cohort households enrolled at endline, rises to 0.16 at seven and to 0.27 at eight; no group had nine. The persisting pools are a subset of the formed pools, so the persistence limit cannot sit below the formation bound; equality tells us one thing, that group 3052 persisted. The next values among persisting pools are 0.090 and 0.090. The limit therefore rests on one group and is not independent evidence about the score needed to persist.

Of the 55 groups that enrolled nobody at baseline, none had a cohort household enrolled at midline or at endline. Three of the 55 had households that entered the group after baseline enrolled at a later round. In two BAIF groups one entrant was enrolled at midline, and in one of these two entrants were enrolled at endline. In one Shramik Bharti group five entrants were enrolled at midline. On the cohort definition the decision not to form a pool was never reversed; counting entrants, three groups later held a pool.

The 43-group survival test: uninformative about the score

Ninety-six self-help groups chose their benefit package through the CHAT deliberation. Of these, 95 were observed in the 2013 round and 85 had households interviewed in the 2012 round. Fifty-five had a pool in place in 2012, and 60 had at least five respondents to the trust question. Forty-three met every condition set in advance (Table 4). Thirty-one of them still had a pool in 2013 and 12 did not.

Table 4
The pre-registered one-year survival test, 43 CHAT (Choosing Healthplans All Together) groups, midline 2012 to endline 2013

Step or quantity	Value	Note
CHAT groups	96	All matched to the household panel by group code
With endline observation	95	
With midline households	85	
Enrolled at midline	55	A pool-period exists
Five or more trust respondents at midline	60	
Analysis sample	43	Survived 31, failed 12
Rule	T·C·DB ≥ 0.056	C = 1 in all groups; DB = 1.000 at both sites (scored from scheme documents, five features each 0 or 1, not the episode-burden measure); product reduces to T
Predicted viable	43 of 43	Observed: 31 survived, 12 failed; the prediction failed for 12. Sensitivity 1.000, specificity 0.000; MCC undefined
Verdict under pre-registered rule	Degenerate; uninformative about the score	Unchanged for any threshold below 0.571, with don't-know as disagree, and at ≥ 3 respondents (n = 50: 37 survived, 13 failed)
Trust item, direct (min, median, max)	0.571, 1.000, 1.000	31 of 43 groups at 1.000
Survival by trust	T < 1: 10 of 12; T = 1: 21 of 31	Spearman ρ(T, survival) = − 0.17
Groups with T ≤ 0.10	0	The second falsifier could not arise

The rule set in advance predicted that a pool would survive if the members' score was at least 0.056; the threshold was the formation bound read from these same 141 groups. In this sample agreement is 1 in every group, because each deliberation ended on one package, and the benefit measure of that test, scored from scheme documents as Methods describes, is 1.000 at both sites. The score therefore equals trust alone, and every one of the 43 groups was above 0.056. The rule predicted 43 survivors and there were 31. The pre-registered prediction, survival for every group above 0.056, failed for 12 of the 43. The rule identified every survivor and none of the 12 failures: sensitivity 1.000, specificity 0.000. The usual summary statistic for such a table cannot be computed when one predicted class is empty, and the pre-registered verdict for that case is that the test is degenerate. The result is the same for any threshold below 0.571, the lowest group value, and with don't-know answers counted as distrust.

Admitting groups with only three respondents brings the sample to 50, of which 37 survived and 13 did not, all predicted to survive.

The sample shows why the test is uninformative about the score. The trust question hit the ceiling. In 31 of the 43 groups every respondent agreed with the trust item; the median is 1.000 and the lowest group value is 0.571. No group had trust at or below 0.10, so the other way the prediction could fail, a pool surviving with very low trust, could not arise. Among the 12 groups where trust was below 1, ten survived; among the 31 where trust was exactly 1, twenty-one did. The rank correlation between trust and survival is − 0.17. Read against the definitions in Methods, this was a test in which two of the three conditions could not vary and the third was measured with a question most people answered yes to. The test tells us about the question, not about the score. The test is reported in full because a paper that proposes a minimum owes readers every sample on which the minimum was watched.

Discussion

What the bound and the test add to the list of influences

The two reviews cited in Background list what raises and lowers uptake and renewal, and both stop at the list [5, 6]. The bound turns one item on the list into a quantity with a threshold. Trust, agreement and benefit are measured as proportions a scheme's own staff can collect. The floor comes from the scheme's own cost cells. The comparison between them is arithmetic. What the reviews call for, a validated tool and a way to see where a scheme stands between equity and efficiency, is what a threshold can provide once validated, and a list of influences cannot. This paper supplies the threshold's form, two floors, a bound, and one test of the benefit component, which returned a null at this sample size. The persistence test adds the second thing the reviews lack: a prediction written down before benefit was related to any post-baseline outcome, and the verdict the outcome gave. The survival outcome was already known for 42 of the 86 groups when the prediction was written.

The two kinds of number reported here count as evidence because they are independent. The floor is computed from cost cells; the bound is read from the enrolment register. They share no data, no method and no assumption, and nothing was adjusted to make them agree. They are also in different units, a share of households and a score. To read them together as a single bracket around the minimum would need two further assumptions: that a higher score brings more households in, and that this relation holds where the system settles. Those assumptions belong to the companion paper and this paper does not make them. What the two facts support without them is narrower and firm. Each applies in the domain of that fact: the bound wherever joining is voluntary, the floor in every pooled scheme, sponsor-funded ones included. And the stand-alone design was below the design's floor in every self-help group the design served before a single household chose.

A necessary condition, and the rate of failure above the bound

Fifty-five groups above the bound enrolled nobody, and 36 of the 86 pools that formed above the bound did not persist. Neither fact bears on the bound, which is an upper limit by construction: the bound says what every group that formed a pool had at least, and promises nothing above the bound. No observation in this paper can refute the minimum-score claim, because the score is not measured. The rate of failure above the bound is a fact the theory must account for: 42% of pools that formed lapsed within two rounds. A statistic that measures how well the score separates survivors from failures tests a sufficiency claim. The test of 9 September 2026 made that claim, with the formation bound from these same groups as the threshold, and the claim failed for 12 of 43 groups. This paper withdraws the sufficiency reading and keeps the necessary condition. The sample of that test could not vary the score, as Results explains.

The persistence test asks the question that can be asked of a necessary condition: among pools that cleared the bound, did more benefit go with more persistence? The answer at 86 groups is that the data cannot say. The point estimate is positive and the interval includes zero and includes 0.27, the association the test had 80% power to detect. The tercile pattern points the same way as the point estimate and is confounded with field partner. A reader who wants to know whether benefit predicts persistence needs more groups or more rounds. The paper reports the verdict the pre-registered rule gave, and the rule gave a negative. On trust the paper has two readings, and both are null or negative. In the 43 groups the rank correlation of trust with survival was − 0.17, with the item at the ceiling in 31 groups. Across the 69 groups whose item ranged from 0.500 to 1.000, the Spearman correlation of midline trust with endline participation was − 0.01. Nothing in this paper supports the trust condition on the margin of staying; the evidence for that condition is the companion household study of joining [14].

One result of the test is not in doubt and is easy to pass over. No baseline member of any group that declined the offer at baseline was enrolled in the two rounds that followed; in three of the 55 groups, households that joined the group later did enrol. On the cohort definition the formation decision was not reversed in these data. If the bound is a necessary condition for forming, the groups below the condition at the offer did not later rise above the condition, or did not act on rising above the condition. Either way, a scheme that wants a pool in a group has one chance at the offer with the group's members, and the offer is the moment the three conditions must be in place.

Where management and government act

Management is not a fourth condition. Management is what produces the three. Trust comes from reliability and proximity: claims paid promptly and visibly, open books, familiar local faces [18]; this paper's own trust readings are null, as the preceding section says, and the sentence rests on the cited work. Agreement comes from structured deliberation, and a documented procedure can produce agreement [13]. Benefit comes from design: packages matched to local illness, pooling rules the community owns [9]. Multiplying the three writes management's job description. Keep all three above the minimum at once rather than maximising a favourite. Put effort where the weakest one is. Treat the minimum as the definition of enough. A management failure never appears in a post-mortem under the

failure's own name. The failure appears as one of the three conditions at zero. Denied claims appear as collapsed trust. A package imposed from above appears as absent benefit. Meetings that never happened appear as agreement that never formed. "Bad management" is at once the universal diagnosis of dead schemes and an empty one, because the phrase says nothing about where to look. The three conditions say where. Ostrom's design principles for common resources read, in this light, as ways of maintaining each of them [19].

Government acts on the minimum, not on the score. Legal recognition, enforceable contracts, proportionate regulation and stable policy set the minimum low. Legal invisibility, taxing mutuals as if they were commercial insurers, or subsidising a competing state programme set the minimum high. The two kinds of action differ in reach by orders of magnitude. A manager tends one pool. A ministry moves the minimum under every pool in the country at once, and the companion paper's first result says the tipping point moves with the minimum. What government cannot do is supply the score. Compulsory enrolment and premium subsidy raise the number of members and leave the score untouched. On the companion theory's account, where the score is below the minimum, they postpone insolvency until the money or the enforcement is withdrawn; this paper observes no score below the minimum. Changing the design acts on the minimum. Federation and shared administration lower what each unit must reach alone, with neither subsidy nor compulsion [20]. The two Indian records fit this. The stand-alone design could not reach the design's floor in any single group; the federated micro-pools formed with benefit levels as low as 0.056.

Limitations

These statements mark where the numbers stop. The formation bound is an upper limit by construction, and no observation in this paper can refute the minimum-score claim, because the score is bounded and not measured. The formation bound is taken over groups observed at the offer, so the bound concerns the score needed to form a pool. The persistence limit is a minimum over a subset of the same groups, rests on one group, and is not independent evidence. The floors come from one scheme and one calculator, and the bound and the test from one project in two Indian states; they are specific to those designs and must be recomputed elsewhere from that setting's own figures. The design is observational and supports bounds and associations, not causal estimates; members' decisions are the channel through which the other limits bite, not a variable independent of design. Under the premise that each group's benefit is measured without error, the minimum is a certain upper limit and needs no sampling interval; a bootstrap interval on a sample minimum would in any case be invalid [21]. Group benefit is a mean over a few households measured with survey error, and the minimum of such means tends to fall below the true minimum, so the literal 0.0556 may understate the bound. The first three order statistics, 0.056, 0.090 and 0.090, and the stability across the ten constitution lines are shown for that reason. The persistence test was written after the survival outcome of 42 of the 86 groups had been read in the 43-group test, so the test is pre-registered as an analysis plan, not blind. The persistence test had 80% power at a correlation of about 0.27, and the interval and power figures are post hoc approximations. The tercile gradient is confounded with field partner. The panel observes one of the four conditions of

survival over two rounds, not survival over years. The 43-group survival test could not vary two of the three conditions, so that test speaks to the trust question's ceiling rather than to the score. No equity-stratified result is reported, because the analysis plan of 28 September 2026 did not include one; a stratification of formation and persistence by baseline household characteristics is the next step, and would be labelled post hoc. And the three conditions are measured through proxies, so a bound computed from proxies bounds what the proxies measure.

Who is left out when the pool fails

This section is an interpretive frame built on the three limits, not a finding of the analysis; the paper carries no equity-stratified result. Three failures, the Ghanaian one described in Background and two from the literature, can each be named by the limit that bound. Ghana's law exists and does not reach everybody: the legal limit binds where the payslip ends, and the 44% outside are what is left. In Ethiopia members enrolled steadily and claims outran premiums [22, 23, 24]: the cost floor bound. The authors listed five causes, and put adverse selection and moral hazard among members first, beside stock-outs of medicines, late settlement of providers' claims and low premiums [22]. In India the offer stood and members who received nothing back stopped paying [17]: the behavioural limit bound. Three limits, three remedies, and the same word, sustainability, for all three. Knowing which limit binds tells a practitioner what to do; the word does not.

The three limits are not alike, and the difference is what an equity reader needs. Regulation works by obligation and pricing works by arithmetic. Both treat the member as something to be governed, either compelled or priced. But the member decides twice, whether to join and whether to use what she has joined, and each decision can undo one of the two governed limits. Where a mandate cannot reach, joining stays a choice, and the household that stays out is the one coverage was meant to reach. Where a scheme is priced on who will join and how they will claim, members supply both numbers. The instruments the field uses to set coverage and price measure someone other than the member. Renewal studies do ask whether members keep paying [6, 17]; what no instrument does is turn the answer into a threshold a scheme can measure before the money moves. Measurement follows agency: those who commission measurement measure what they can move, and a household's decision is the one thing no ministry or actuary can move. That is why remedies recycle. When members' behaviour defeats the pricing, the answer is more compulsion; when compulsion cannot reach, the answer is more registration, identity linkage, exemptions and waiting periods. Each works partway and then stops. Fourteen years after the mandate, Ghana's active membership stands at 56% [3]. A remedy aimed at the wrong limit cannot secure survival, whatever the remedy costs, and the households the remedy fails are the ones without a payslip.

Conclusions

This paper measures two things a voluntary health pool needs, from records a scheme already holds, and tests a third: two actuarial floors, a formation bound, and whether members' benefit at formation predicted persistence. The floors came from two designers' spreadsheets, the bound and the test from

one measured condition, benefit, in 141 groups. Trust and agreement were not measured here in the form the score requires.

What follows is a proposal for how a scheme would take all three measurements, and these data do not validate the proposal. A scheme can take the measurements with what the scheme already does. Trust is asked: one question to households, whether the body that pays the claims is trusted to pay, reported as the share who say yes. Agreement is observed in the room. Run the decision on the benefit package as a structured deliberation, record each member's choice before and the package agreed after, and take the share whose own choice matches the group's. Benefit is read from what members face and what the pool returns: the illness episodes a household meets in a year, set against the benefits the pool paid. Each is a proportion, and each can be collected by the scheme's own staff. The floor comes from the same desk: the administration workbook gives the households required, the catchment turns that into a share, and the enrolment register shows who renewed against who was eligible.

The rest is three questions asked in order. How many can the law reach? Compliance and coverage data answer. How many must pay to cover costs? The workbook answers. How many choose to stay? The enrolment register answers. Whichever leaves the least room is the one that will kill the scheme, and each has a remedy: enforcement, the way the pool is organised, and the design of trust, agreement and benefit. Repeat yearly, and the direction of change is the forecast. In the one Indian scheme whose workbook we recovered, the second question had no answer within any single self-help group: the floor was 19 times the membership of the largest group the scheme served. In 86 federated micro-pools the third question was answered at formation, and the lowest benefit at which a pool formed was 0.056, so the score needed to form is at most 0.056. Two rounds later 50 of the 86 still had at least one member enrolled, with 20% of the cohort observed at endline enrolled, and benefit at formation did not tell which, at the sample size in hand. No baseline member of a group that declined at the offer was enrolled afterwards. On the companion theory's account, where the score is below the minimum, compulsion and subsidy add members and leave the reason to stay untouched. The households the measurement serves are those outside the reach of a mandate, for whom coverage depends on a voluntary pool. The legal limit says whether they are reached. The cost floor says whether their pool can pay. The behavioural limit says whether they will keep paying. The measurement proposed here is taken before the money moves. The measurement tells a ministry, a funder or a scheme which of the three limits will bind, which is the one fact none of them holds today. What holds for the health pool is proposed for the wider class of voluntary collective organisations. A savings group, a water-user association and a burial society face the same three conditions against a minimum their own costs set. That is the next test.

Abbreviations

CHAT
Choosing Healthplans All Together
MCC
Matthews correlation coefficient

MD5
message-digest checksum
RECORD
REporting of studies Conducted using Observational Routinely-collected health Data
SE
standard error
STROBE
STrengthening the Reporting of OBservational studies in Epidemiology.

Declarations

Ethics approval and consent to participate

This paper is a secondary analysis of de-identified group-level and household-level records and administrative pricing instruments; no new data were collected. The Micro Insurance Academy, New Delhi, collected the Indian data in 2010 to 2013, and the author was the Academy's Chairman and Managing Director until the Academy's closure in 2017. The independent ethics committee of the University of Cologne approved all study tools in accordance with the Ethical Guidelines for Biomedical Research on Human Participants of the Indian Council of Medical Research (2006) [13]. Informed consent was obtained before each interview, as reported in the published study protocol [12].

Consent for publication

Not applicable.

Competing interests

The author has advised governments and non-governmental organisations on community-based health insurance and has published on microinsurance; six such publications are cited. He serves pro bono as Principal Adviser to Health Investment & Financing Corp., which has no involvement in the schemes studied.

Additional files

Additional file 1. RECORD checklist (DOCX).

Funding

This research received no external funding.

Author Contribution

D.M.D. conceived the study, recovered and verified the two pricing instruments, wrote the pre-registration, performed all analyses, built the replication package, and wrote the manuscript. D.M.D. read and approved the final manuscript.

Acknowledgements

The author thanks the members of the self-help groups whose records this study analyses, and the field teams of the Micro Insurance Academy, New Delhi, who collected the Indian data in 2010 to 2013.

Data Availability

The replication package for this study is openly available in Zenodo under the concept DOI (https://doi.org/10.5281/zenodo.22131769), which always resolves to the latest version; the version accompanying this manuscript is 2.2.0 (https://doi.org/10.5281/zenodo.23165609). It contains the group-level records, the two recovered pricing instruments with cell-level provenance, the pre-registration of the persistence test with its checksum, the analysis and inventory scripts, the reports of record, the group table of the persistence test without names, a per-group table of household episode counts, and the procedures, report of record and group table of the 43-group test. A single script, REPLICATE_ALL.R, reproduces every value in Tables 2 and 3 and checks each against the printed value.The group-level and household-level records analysed here are in the replication deposit of a companion manuscript, version 1.2.0, (https://doi.org/10.5281/zenodo.22987235). The participant-level panel file hh_master.sav is not published, following the standard restriction on individual records collected under ethics approval. The script SUST_persistence_test_v1.R rebuilds the deposited group table of the persistence test from that file, so every statistic in Table 5 can be recomputed by anyone holding it.

References

1. National Health Insurance Act. 2012 (Act 852), § 27(1). Republic of Ghana; assented October 31, 2012 (repealing Act 650 of 2003).
2. van der Wielen N, Falkingham J, Channon AA. Determinants of National Health Insurance enrolment in Ghana across the life course: are the results consistent between surveys? Int J Equity Health. 2018;17:49. https://doi.org/10.1186/s12939-018-0760-x.
3. National Health Insurance Authority. NHIS active membership hits 56 percent coverage. Accra: National Health Insurance Authority. 2026. Accessed August 23, 2026. https://www.nhis.gov.gh/News/nhis-active-membership-hits-56-percent-coverage--5966
4. Kotoh AM, Van der Geest S. Why are the poor less covered in Ghana's national health insurance? A critical analysis of policy and practice. Int J Equity Health. 2016;15(1):34.

https://doi.org/10.1186/s12939-016-0320-1.

5. Fadlallah R, El-Jardali F, Hemadi N, et al. Barriers and facilitators to implementation, uptake and sustainability of community-based health insurance schemes in low- and middle-income countries: a systematic review. Int J Equity Health. 2018;17:13. https://doi.org/10.1186/s12939-018-0721-4.
6. Hussien M, Azage M. Barriers and facilitators of community-based health insurance policy renewal in low- and middle-income countries: a systematic review. ClinicoEcon Outcomes Res. 2021;13:359–375. https://doi.org/10.2147/CEOR.S306855.
7. Dror DM. Measuring what matters for the viability of inclusive insurance: behavioural diagnostics from 169 community-based health insurance schemes across 38 countries. Geneva Pap Risk Insur Issues Pract. 2026. https://doi.org/10.1057/s41288-026-00400-8.
8. Benchimol EI, Smeeth L, Guttmann A, Harron K, Moher D, Petersen I, Sørensen HT, von Elm E, Langan SM, RECORD Working Committee. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) statement. PLoS Med. 2015;12(10):e1001885. https://doi.org/10.1371/journal.pmed.1001885.
9. Dror DM, Majumdar A, Jangle N. Estimating capital requirements to scale health microinsurance serving rural poor populations. Geneva Pap Risk Insur Issues Pract. 2019;44:410–44. https://doi.org/10.1057/s41288-019-00126-w.
10. Dror DM, Duru G. Replication materials for An exact characterization of the product on the closed cube, without continuity, monotonicity or associativity [software]. Version 1.1.0. Zenodo; 2026. https://doi.org/10.5281/zenodo.20021890
11. Dror DM. Replication data and code for: The boundary below which trust cannot sustain cooperation in a voluntary collective organization [dataset and software]. Version 1.2.0. Zenodo; 2026. https://doi.org/10.5281/zenodo.22987235
12. Doyle C, Panda P, Van de Poel E, Radermacher R, Dror DM. Reconciling research and implementation in micro health insurance experiments in India: study protocol for a randomized controlled trial. Trials. 2011;12:224. https://doi.org/10.1186/1745-6215-12-224.
13. Dror DM, Panda P, May C, Majumdar A, Koren R. One for all and all for one: consensus-building within communities in rural India on their health microinsurance package. Risk Manag Healthc Policy. 2014;7:139–53. https://doi.org/10.2147/RMHP.S66011.
14. Dror DM. Replication package for: Trust in whom? Only trust in the delivering entity predicts collective action [dataset]. Version 3.0.1. Zenodo; 2026. https://doi.org/10.5281/zenodo.20501661
15. Dror DM. Replication package for: When disagreement looks like agreement: a measurement-theoretic foundation for consensus in collective-action research [software]. Version 2.0.2. Zenodo; 2026. https://doi.org/10.5281/zenodo.20079778
16. Asmussen S, Albrecher H. Ruin Probabilities. 2nd ed. Singapore: World Scientific; 2010. https://doi.org/10.1142/7431.
17. Panda P, Chakraborty A, Raza WA, Bedi AS. Renewing membership in three community-based health insurance schemes in rural India. Health Policy Plan. 2016;31(10):1433–44.

https://doi.org/10.1093/heapol/czw090.

18. Dror DM, Jacquier C. Micro-insurance: extending health insurance to the excluded. Int Soc Secur Rev. 1999;52(1):71–97. https://doi.org/10.1111/1468-246X.00034.
19. Ostrom E. Governing the Commons: The Evolution of Institutions for Collective Action. Cambridge: Cambridge University Press; 1990. https://doi.org/10.1017/CBO9780511807763.
20. Bonnevay S, Dror DM, Duru G, Lamure M. A model of microinsurance and reinsurance. In: Dror DM, Preker AS, eds. Social Reinsurance: A New Approach to Sustainable Community Health Financing. Washington, DC: World Bank; Geneva: International Labour Office; 2002:153–186. https://doi.org/10.1596/0-8213-5041-2
21. Bickel PJ, Freedman DA. Some asymptotic theory for the bootstrap. Ann Stat. 1981;9(6):1196–217. https://doi.org/10.1214/aos/1176345637.
22. Hussien M, Azage M, Bayou NB. Financial viability of a community-based health insurance scheme in two districts of northeast Ethiopia: a mixed methods study. BMC Health Serv Res. 2022;22:1072. https://doi.org/10.1186/s12913-022-08439-8.
23. Hussien M, Azage M, Bayou NB. Continued adherence to community-based health insurance scheme in two districts of northeast Ethiopia: application of accelerated failure time shared frailty models. Int J Equity Health. 2022;21:16. https://doi.org/10.1186/s12939-022-01620-9.
24. Hussien M, Azage M, Bayou NB. A mixed methods study of community-based health insurance enrolment trends and underlying challenges in two districts of northeast Ethiopia: a proxy for its sustainability. PLoS ONE. 2022;17(8):e0266583. https://doi.org/10.1371/journal.pone.0266583.