

Slow escape or trapped in poverty? Evidence from a randomized control trial in Bangladesh

PRE-ANALYSIS PLAN

1. RESEARCH QUESTION

Do social protection interventions provide the accelerant that allows households to permanently escape poverty and does this depend on design features of the intervention (type of modality, inclusion of complementary programming)?

We formalize this question as an adjudication between two hypotheses:

H1 (poverty trap): households below a critical threshold are structurally trapped; a sufficiently large transfer (a “big push”) enables threshold-crossing and permanently altered dynamics. Observable features of a poverty trap include treatment effects concentrated among households above or near a threshold in baseline consumption or assets; a detectable kink or interior maximum in the effect-by-baseline profile; treatment–control gaps in poverty measures that persist or widen over time; and convergence rates that differ between treated and untreated households.

H2 (slow convergence): poor households are on a slow convergent path; transfers produce level shifts (a durable “head start”) without altering the underlying dynamics. Observable signatures: smooth, weakly pro-poor treatment-effect profiles with no threshold; treatment–control gaps that shrink or remain flat over time; and convergence rates that are equal across arms.

2. STUDY DESIGN AND SAMPLE

The Transfer Modality Research Initiative study operated for 24 months, May 2012 to April 2014. It consisted of two randomized control trials implemented in two regions of Bangladesh: Rangpur division (province) in the northwest (hereafter “North”) and Barisal and Khulna divisions in the south (hereafter “South”). In the narrative below, we focus on the North as this was the region where we have data for survey baseline (2012), midline (2013), endline (2014), 2018, 2022, and 2026.

In the RCT fielded in the North, there were four treatment arms in addition to the control group: a cash transfer; a food ration; a half cash payment and half food ration; and a cash transfer plus nutrition BCC. Note, however, that for budgetary reasons, we did not include the half cash – half food treatment arm in the 2018 or 2022 survey rounds. Across all arms, the target beneficiary was the mother of an “index child” aged 0-24 months in March 2012, residing in a poor rural household.

The cash treatment (“Cash”) consisted of a monthly payment of 1,500 Taka (approximately \$19 US) per household. Payments were made to mothers via mobile money. The amount was about 25 percent of the average monthly household consumption expenditures of poor rural households in Bangladesh as of 2012.

The food treatment arm (“Food”) consisted of a monthly food ration of 30 kilograms (kg) of rice, 2 kg of *mosoor* pulse (a type of lentil), and 2 liters of micronutrient-fortified cooking oil. Several considerations guided the choice of these commodities. First, they were all foods that were familiar to beneficiaries. Second, they provided energy (through the rice and cooking oil), protein (via the pulses), and fats (the cooking oil). Third, they could all be obtained via existing World Food Programme (the intervention implementer) procurement channels. Fourth, for logistical and budget reasons, as food transfers could only be made once a month, and virtually none of the beneficiaries had access to refrigeration, it was necessary to ensure that the foods provided were shelf-stable (ie non-perishable). The quantities were chosen so that – based on prevailing prices at the time the intervention commenced - the value of the food ration was equal to the value of the cash provided in treatment arms that provided cash.

The BCC component consisted of a suite of activities led by a Community Nutrition Worker (CNW). CNWs, all women, were from local communities and came from the same villages as TMRI participants. The core BCC activity was a weekly, one-hour group session in each village with a CNW. The BCC session materials were derived from material developed for Alive & Thrive (A&T) in Bangladesh, a comprehensive program aimed at improving breastfeeding and complementary feeding practices and reducing stunting among young children. This curriculum is widely used throughout Bangladesh and follows World Health Organization and UNICEF guidelines for infant and young child nutrition (IYCN). TMRI also followed A&T in terms of the content of the BCC sessions and an approach that included community engagement, group BCC sessions, and home visits. Because many factors underlying the IYCN recommendations – including local food availability and seasonality – were comparable in the north and south of Bangladesh, TMRI used the same approach to the BCC across the two regions.

These sessions covered the following topics: (1) the importance of nutrition and diet diversity for health; (2) how handwashing and hygiene improve health; (3) diet diversity and micronutrients; (4) breastfeeding; (5) complementary foods for children 6-24 months; and (6) maternal nutrition. A variety of methods were used to deliver this information including presentations, questions and answers, interactive call-and-answer songs and chants, practical demonstrations, and role-playing. One of these sessions per month, with only beneficiaries participating, occurred on the day of the transfer distribution. For the remaining group BCC trainings each month, other household members – particularly mothers-in-law, husbands, and other pregnant or lactating women – were invited to attend along with beneficiaries, with the intention of creating a supportive household atmosphere and behavior change at the household level. These combined sessions served to facilitate women’s ability to participate in the BCC, as household members could see what women were participating in and reduce restrictions on attendance, and to increase uptake of BCC messages as husbands and mothers-in-law are also key decision makers on food purchases, IYCF, and child-rearing in the household. CNWs also made home visits to beneficiaries twice a month to follow up on topics discussed during the group sessions and to discuss specific concerns that mothers might have. While attendance at these BCC sessions was a condition for receipt of transfers, this was a “soft” condition. When a mother missed a session, CNWs followed up with a home visit to ascertain why the session had been missed, and there were no cases where a beneficiary was dropped from the study for failing to attend sessions. In addition, CNWs and project staff conducted community meetings and met with influential members (village leaders, imams, elders) of the villages in which the BCC took place to explain the purposes of the nutrition training and to provide them with the information being conveyed to study participants. CNWs received training prior to the start of the intervention, with

refresher training undertaken three and 12 months after the intervention began. In localities where the same payment point was used for both the Cash arm and the Cash+BCC arm, Cash beneficiaries were paid in the morning while Cash+BCC beneficiaries were paid in the afternoon, to minimize the likelihood of information from the BCC activities spilling over to the Cash treatment arm. BCC activities cost approximately \$50 per year per beneficiary.

Based on power calculations, in the North, five upazilas (sub-districts) were selected using simple random sampling from a list of upazilas where in 2010 the proportion of households living below Bangladesh's lower poverty line was 25% or higher. All villages within these five upazilas were listed. Villages classified as urban or villages with fewer than 125 households were dropped. Using a random number generator, each village was assigned a random number. Villages were then sorted in ascending numerical order with the first 275 retained. The first 50 villages were assigned to treatment group 1, the second 50 to treatment group 2, the third 50 villages to treatment group 3, the fourth 50 villages to treatment group 4, and the fifth 50 villages to the control group. The remaining 25 villages were held as a reserve. A complete village census was carried out in each of the 250 selected villages, collecting information on household demographics, a set of poverty indicators, and whether households participate in safety nets and other targeted interventions. Using these data, a list was compiled of households that: (1) were considered poor (i.e. based on the poverty indicators collected, they were estimated to have consumption below Bangladesh's lower poverty line); (2) would have at least one child aged 0-24 months when the intervention began; and (3) were not receiving benefits from other safety net interventions. These households were eligible to participate in the study. Using simple random sampling, 10 eligible households were selected from each village. The total sample in the North in the 2012-14 study thus included 250 clusters and 2,500 households; note, however, that for this work we only include households that were in the cash, the food, the cash and BCC, and control arms (200 clusters and 2,000 households at baseline).

Prior data and pre-registration boundary. The 2012, 2014, 2018 and 2022 survey rounds have been analyzed by the study team. No member of the analysis team has accessed any 2026 outcome data; the 2026 data files (including household tracking files) are sequestered from the analysis environment until this plan is registered.

3. TREATMENT AND CONTROLS

Treatment households are households who, between 2012 and 2014, were assigned to receive:

- (1) A monthly cash payment
- Or
- (2) A monthly food ration
- Or
- (3) A monthly cash payment plus participation in the nutrition BCC activities described above

Control households were households not assigned to any treatment group.

4. OUTCOMES

We have four primary outcomes: (1) log real per capita consumption; (2) headcount poverty (P0); (3) the poverty gap (P1); and (4) the squared poverty gap (P2). P1 and P2 are co-primary adjudication outcomes: because households can cross the poverty line without their distance from the line changing differentially across arms, P0 alone cannot discriminate between the trap and convergence hypotheses.

Our secondary outcomes include: poverty status measured at a poverty line 20 percent above the extreme poverty line — pre-specified because, given the secular decline in extreme poverty in rural Bangladesh, effects at the extreme line may be attenuated by a floor effect, with the policy-relevant variation moving up the consumption distribution.

5. ESTIMATION STRATEGY

We will conduct intent-to-treat (ITT) analysis using analysis of covariance (ANCOVA), controlling for the baseline value of the outcome. We will estimate the following equation:

$$Y_{int} = \beta_{0N} + \beta_{FN} * Food_{iN} + \beta_{CN} * Cash_{iN} + \beta_{BN} * Cash_{BCC_{iN}} + \delta_N * Y_{iN1} + \varepsilon_{int}$$

where Y_{int} denotes the outcome for household i , in the North, at time t (endline, 4yPP, 8yPP, and 12yPP). Y_{iN1} denotes baseline outcomes, β_{0N} is the constant term, and ε_{int} is the term for unobservables, which we assume to be normally distributed. $Food_{iN}$, $Cash_{iN}$, and $Cash_{BCC_{iN}}$ are dummy variables for a household in the North being assigned to Food, Cash, or Cash+BCC respectively; corresponding coefficients reflect treatment impacts of these arms relative to the region's control group.

We estimate all outcomes using ordinary least square regressions.

In addition to the ANCOVA specification above, we pre-specify five families of tests that jointly adjudicate between H1 and H2 (Section 1):

Test 1 — Poverty trajectories and gap dynamics. We estimate the ANCOVA ITT for P0, P1 and P2 at each post-intervention round. An important statistic is the change in each treatment–control gap between every survey round.

Test 2 — Kink and threshold detection in consumption space. We interact treatment with baseline log consumption (linear and quadratic terms), estimate quintile-specific treatment effects with a test of equality across quintiles, and conduct a grid search for a treatment-effect threshold over baseline-consumption percentile ranks 10–90. Inference on the threshold statistic is by village-level randomization inference. Under H1, effects concentrate above a threshold; under H2, the profile is smooth and, if anything, declining in baseline position.

Test 3 — β -convergence with correction for regression to the mean. We regress consumption growth from 2012 to 2026 on baseline log consumption, by arm, instrumenting baseline consumption with the inverse hyperbolic sine of baseline productive assets to purge Galton (measurement-error) bias, and we test equality of the convergence coefficient across arms. Equal convergence across arms combined with level shifts is H2's functional form; arm-specific

convergence — in particular, absent or slower convergence among control households below a candidate threshold — is H1's.

Test 4 — Heterogeneity in asset space. We repeat the kink and threshold tests of Test 2 using the household's baseline productive-asset (excluding land) percentile rank as the running variable. Because a non-trivial share of households holds exactly zero of some asset categories, the running variable is rank-based, and all asset-denominated outcomes are reported in levels per capita, inverse hyperbolic sine, and ranks side by side, so that no conclusion depends on a logarithmic transformation of variables with zeros.

Test 5 — Control-group falsification. Using control households only, we test whether consumption transitions between consecutive rounds (including 2022–2026) are nonlinear, using polynomial terms in lagged consumption, with and without village fixed effects. Nonlinear untreated dynamics are expected in this environment and will not be interpreted as evidence of a poverty trap; the purpose of this test is to demonstrate that dynamics-based (S-curve) estimation can detect threshold-like structure in the absence of any trap mechanism, which disciplines the interpretation of Tests 2 and 4.

6. INFERENCE

We cluster our standard errors at the unit of randomization, the baseline randomization village; households that have relocated since 2012 remain assigned to their baseline village cluster. We will report significance levels at the 10%, 5% and 1% levels. Across the family of primary hypotheses (P0, P1 and P2 at each post-endline round), we will report Romano–Wolf stepdown p-values, with wild cluster bootstrap p-values as an agreement check on the headline effects. For all threshold and kink statistics (Section 5, Tests 2 and 4), inference is by randomization inference: we re-randomize treatment assignment at the village level (at least 500 permutations), holding the design fixed, and report permutation p-values for the maximal threshold statistic.

7. STATISTICAL POWER AND MINIMUM DETECTABLE EFFECTS

Using realized control-group means, standard deviations and intracluster correlations from the 2022 round, together with the realized attrition trajectory, we will compute minimum detectable effects (MDEs) for the 2026 ITT on each primary and secondary outcome.

8. ATTRITION AND MISSING DATA

Our data spans a period of 14 years. Not surprisingly, we expect there to be attrition in our sample. We will describe attrition in the following ways:

First, we will document the prevalence of attrition in these data and its primary causes, distinguishing between households who attrit because they migrated, they could not be traced or were traced but declined to be interviewed.

Second, we will assess, by survey round, whether attrition is random or non-random. We will compare mean values of household baseline characteristics and treatment status. We will estimate models where the outcome variables are whether the subject was successfully in each survey round as a function of the baseline characteristics.

Third, we will test whether attrition is differential with respect to the interaction of treatment arm and baseline consumption (and baseline poverty status). If attrition differs across arms by more than five percentage points, or if any arm-by-baseline-consumption interaction is significant at the 10 percent level, we will compute Lee (2009) bounds on the primary treatment effects.

In terms of missing data, we do not undertake any imputation for missing outcome variables.

9. BALANCE

We will assess balance across selected baseline characteristics (subjects' age and sex; maternal and paternal age and schooling; and household wealth) through pairwise t tests and an omnibus F test (regressing these characteristics on assignment to treatment village) and note these findings.

10. HETEROGENEITY ANALYSIS

At baseline, more than 70 percent of households were poor. This, together with the fact that few households (less than five percent) were female headed has meant that in past analyses, typical heterogeneity analysis (disaggregating by baseline consumption or wealth, or female headship) did not produce meaningful differences across sub-groups. For this work, we pre-specify heterogeneity at three tiers. Tier 1 (pre-specified, confirmatory): quintile-specific treatment effects by baseline consumption and by baseline productive-asset rank, with tests of equality across quintiles (Section 5, Tests 2 and 4). Tier 2 (measurement discipline): because measured baseline consumption contains substantial measurement error, heterogeneity in baseline consumption is estimated both naively and using the asset-based instrument of Test 3 (or rank-based conditioning); conclusions that hold only in the naive specification will be flagged as fragile. Tier 3 (exploratory, hypothesis-generating): a data-driven search over baseline characteristics — household demographics and dependency structure, head and spouse occupation and education, female headship, credit and savings access, land, and village characteristics — using penalized (lasso-type) variable selection, to characterize which households escape poverty and which do not. Tier 3 results will be labeled as exploratory in the paper.

11. POVERTY DYNAMICS AND MOBILITY

With the 2026 round we observe poverty status in up to five rounds (2012, 2014, 2018, 2022, 2026). On the balanced poverty-status panel we pre-specify: (1) round-to-round poverty transition matrices, by arm; (2) the share of households poor in all observed rounds (“never escaping”) and the share never poor in any post-baseline round, by arm, with pairwise tests against control; (3) “permanent escape,” defined as being non-poor in 2018, 2022 and 2026, by arm; and (4) correlates of never escaping, conditional on treatment arm, using baseline consumption, baseline productive assets, land, household size, and the target woman's age and education.

Two quantities carry particular interpretive weight: (i) the share of households never escaping poverty, by arm; and (ii) whether, conditional on baseline consumption, baseline asset position predicts remaining persistently poor. If asset position does not predict persistent poverty conditional on consumption, this would indicate that persistent poverty reflects consumption distance from the poverty line rather than an asset threshold — a convergence fact rather than a trap fact — implying that targeting of future programs should key on the consumption-poorest rather than the asset-poorest households.

12. ROBUSTNESS AND SENSITIVITY ANALYSIS

We will undertake the following robustness and sensitivity analyses:

- (1) We will estimate our model using level (rather than log) real consumption per capita.
- (2) We will estimate using log consumption per adult equivalent.
- (3) Where the outcome variable is log or level real consumption, we will winsorize at the 2nd and 98th percentiles and re-estimate
- (4) We will assess whether our results are sensitive to the inclusion in the consumption aggregate of large infrequent lumpy expenditures such as ceremonies or religious expenditures.
- (5) We will estimate all models with and without baseline characteristics.
- (6) We will estimate P0 poverty using a probit and for poverty depth and severity, we will estimate censored regressions based on tobit models; for all three, we will then calculate marginal effects (dy/dx).
- (7) We will calculate Romano-Wolf p-values to assess whether our results are robust to multiple hypothesis concerns.
- (8) We will assess whether our results are sensitive to the calculation of the poverty line. Specifically, we will re-estimate all primary models at poverty lines 20 percent above and 20 percent below the extreme line. We note in advance that, given secular poverty decline, effects at the lower line may be attenuated by a floor effect; effects that are larger at the higher line are consistent with impacts having moved up the consumption distribution, as convergence implies.
- (9) We will estimate all primary models on the balanced panel (households observed in every round) as a robustness check; the round-specific (unbalanced) sample is the primary analysis sample.
- (10) Analyses of whether households were permanently poor or permanently escaped poverty are pre-specified in Section 11 (Poverty dynamics and mobility).

13. TRANSPARENCY

We end by noting that we regard this pre-analysis plan as a “serious statement of intent” but not a binding contract. More specifically, depending on the results we obtain from these analyses, there may be additional statistical work that is warranted and/or additional work suggested/required by journal reviewers, editors and others who read the paper and provide comments. In the paper we write, we will note that such analyses were not included in this pre-analysis plan.