

Group versus Individual Coaching for Rural Social Protection Programs: Evidence from Uganda, Philippines, and Bangladesh

[Emily A. Beam](#)

University of Vermont

[Lasse Brune](#)

Northwestern University

[Narayan Das](#)

BRAC University

[Stefan Dercon](#)

University of Oxford

[Nathanael Goldberg](#)

Innovations for Poverty Action

[Rozina Haque](#)

BRAC University

[Dean Karlan](#)

Northwestern University and IPR

Maliha Khan

BRAC University

Doug Parkerson

Innovations for Poverty Action

Ashley Pople

The World Bank

Yasuyuki Sawada

The University of Tokyo

Christopher Udry

Northwestern University and IPR

Rocco Zizzamia

University of Oxford

Version: March 13, 2026

DRAFT

Please do not quote or distribute without permission.

Abstract

Multifaceted social protection programs in low-income countries often include both capital grants and informational and behavioral support on the premise that households face simultaneous and multiple frictions. To tackle informational and behavioral constraints, programs typically deploy either individual or group coaching visits from field agents. The relative efficacy of individual versus group coaching could provide insights into the underlying mechanism through which information and behavioral support change household decisions. However, in three similar randomized evaluations in Uganda, the Philippines, and Bangladesh, the authors find no differences in efficacy. Given its 15–20% lower costs, group coaching is more cost-effective.

1 Introduction

Graduation-style social protection programs aim to lift people out of extreme poverty under the premise that multiple supports are needed to address the multiple constraints households face. These programs demonstrate marked improvements in household well-being in both the short and long run (Banerjee et al., 2015, 2021; Barker et al., 2024). Graduation programs typically provide a comprehensive package of assistance, including regular consumption support, training and coaching, a one-time asset transfer to support livelihood development, and improved access to savings. The combination of reinforcing supports, especially the provision of regular coaching and mentoring, is argued to be important to shift households into higher-return economic activities (Blattman et al., 2016; Balboni et al., 2022; Sedlmayr et al., 2020; Banerjee et al., 2022; Roelen and Devereux, 2019).

We focus on the training and coaching aspect of the program. As with any program, quality of implementation—not merely program design—may drive impact. Training and coaching pose obvious challenges for implementing programs with fidelity, more so than other key program components such as cash transfers and improved access to savings. Different modalities of training and coaching may also vary in efficacy, shedding light on the underlying frictions. Specifically in our case, if group-based coaching works better than individual-based coaching, this may indicate that social norms and complementarities across households are barriers best overcome through group discussion. On the other hand, if individual coaching works better, this may indicate highly heterogeneous information challenges across households, such that one-on-one engagement is more effective. For policy design, the considerably lower costs of group coaching mean that if both generate similar effects, group coaching would be most cost-effective.

The promising results of graduation programs have motivated rapid expansion by non-governmental organizations and governments, which in turn has stimulated demand for understanding the most cost-effective implementation strategy (Andrews et al., 2021; Banerjee et al., 2024). Versions of the graduation model have been implemented in at least 75 countries, reaching more than 92 million people, with governments playing an increasingly central role (Andrews et al., 2021). However, this push to scale requires an understanding of which program components are essential and which can be adapted to improve effectiveness or reduce cost. In practice, as in education and active labor market programs more generally, there is no clear consensus on whether individual- or group-based instruction modalities are more effective (McKenzie et al., 2023; Brooks et al., 2018; Ross and Begeny, 2011; Mühlberger and Traut-Mattausch, 2015; Miles et al., 2022; Roelen and Devereux, 2019).

Coaching, a standard feature of these graduation programs, relies on a well-trained labor

force and the organizational capacity to deliver complex, tailored support to participants. This element is costly, and the challenges of identifying highly skilled coaches grow with program size. One way to reduce program costs and logistical complexity is to shift away from individual-based coaching, usually delivered through home visits, to group-based coaching, in which a single coach works simultaneously with a group of participants. However, this decision presents a potential trade-off: individual coaching may deliver more tailored and confidential advice and problem-solving, but group coaching may facilitate peer learning and social capital building.

We experimentally measure the differential impact of group versus individual coaching within graduation programs in three sites: Uganda, the Philippines, and Bangladesh. We also test the impact of a lower intensity individual coaching arm in Bangladesh. Each site uses a cluster-randomized design, assigning communities to a control group that receives no program, to the Graduation program with individual coaching, or to the program with group coaching. Coaching sessions—whether individual or group—are typically fortnightly. While we vary the modality of coaching delivery, the syllabus and discussion topics remain identical across modalities. We measure short-run effects approximately one year after program conclusion in Uganda and the Philippines and four-year effects in Bangladesh.

Overall treatment effects are strong and consistent with prior research ([Banerjee et al., 2015, 2021](#)). Relative to the control group, graduation increased household consumption (0.17–0.36 standard deviations (s.d.)); reduced food insecurity (0.31–0.80 s.d. in Uganda and the Philippines, not statistically significant in Bangladesh), increased value of productive assets (0.22–0.50 s.d.), increased monthly household income (0.22–0.44 s.d. in Uganda and Bangladesh; not significant statistically in the Philippines) and improved subjective well-being (0.06–0.67 s.d.).

In terms of the relative effectiveness of group vs. individual coaching, we find no evidence of differences in treatment effects between group and individual coaching across a range of economic outcomes when examining sites separately or pooling across all three sites. Additionally, when considering the ratio of impact to cost, we find that group-coaching is more cost-effective in raising our aggregate outcomes index and key welfare outcomes—consumption, food security and subjective well-being—and we reject the equality of these group vs. individual ratios at the 5% level.

We also examine impacts on social connections and trust, with the hypothesis that group coaching should affect those more. We find suggestive evidence that group coaching promotes social connection and trust, but impacts are modest. We further test for heterogeneity, noting that while the two approaches perform similarly on average, this could mask underlying—and offsetting—heterogeneity. Following the generic machine learning approach in [Chernozhukov](#)

et al. (2020), we find no evidence of heterogeneous treatment effects across a broad set of baseline measures.

While common experimental features across these three sites ensure comparability, the diverse study contexts support the external validity of our findings and strengthen the case for scaling group coaching. Specifically, we provide evidence over shorter (Philippines and Uganda) and longer time horizons (Bangladesh); for graduation programs delivered at various scales, from a scoping study in the Philippines to a scaled intervention in Bangladesh; for refugee (Uganda) and non-refugee populations (Uganda, Philippines and Bangladesh); for graduation programs with large lump-sum cash transfers (Uganda) and for those with in-kind productive asset transfers (Philippines and Bangladesh); and across different implementation partners.

A key limitation of our study is the absence of a treatment arm in which recipients received the graduation program without the coaching component. This arm was not included for three reasons. First, budget constraints did not allow for an additional treatment arm trading off statistical power. Second, the program without coaching is proximate to a lump-sum cash transfer (albeit with improved access to savings), and there is a large literature on the short- and long-run impacts of lump-sum cash transfers (Crosta et al., 2024). Third, the implementers held a strong conviction that coaching was an essential component for success. Supporting this conviction, evidence suggests that coaching is, in fact, an important determinant of program impact: when coaching has been experimentally removed in similar interventions, program impacts have tended to dissipate (Burchi and Strupat, 2018; Bossuoy et al., 2022; Blattman et al., 2016; Banerjee et al., 2022; Sedlmayr et al., 2020; Roelen and Devereux, 2019). For these reasons, while we study the effects of the intensity and modality of coaching, we cannot examine whether receiving *any* coaching has a marginal impact on welfare outcomes. Thus, our contribution lies in studying how to, rather than whether to, deliver coaching within graduation programs.

Thus, we contribute to the literature on improving the effectiveness of training and education programs by comparing the relative costs and benefits of group versus dyadic training modalities. There is some experimental evidence from behavioral sciences on the effects of group versus dyadic coaching modalities on helping subjects reduce procrastination (Mühlberger and Traut-Mattausch, 2015), and their respective pedagogic effectiveness on reading from the education literature (Miles et al., 2022; Ross and Begeny, 2011). However, these studies are limited to high-income economies, have small sample sizes, and yield inconsistent results. In the context of development programs, the literature on firms indicates that peer interactions, fostered through group meetings, may have positive effects on entrepreneurial performance (Cai and Szeidl, 2018). In contrast, an RCT comparing mentors

(dyadic) to classroom training (group) by [Brooks et al. \(2018\)](#) shows better outcomes for entrepreneurs who receive one-on-one mentoring, although the treatment arms vary both the modality and instructional content.

We also advance our understanding of how best to deliver multifaceted economic inclusion programs to maximize impact and cost-effectiveness, a critical question as governments worldwide look to build and scale similar programs. Coaching, being both expensive and reliant on scarce human resources, represents a major obstacle to scaling ([Al-Ubaydli et al., 2017](#)). Motivated by this challenge, a growing body of evidence unpacks the role of particular program elements or modifies components to maximize the potential to scale. One-on-one personalized coaching and mentoring is a cornerstone of the early graduation model, and its value has been demonstrated across a range of contexts. For example, [Banerjee et al. \(2022\)](#) finds that asset provision alone is insufficient to generate the long-lasting impacts realized by pairing assets with additional coaching support and training. Similarly, [Leight et al. \(2023\)](#) finds that a lighter-touch graduation program without coaching in comparison with a control group yields only modest impacts on assets and income, and it has no detectable impacts on consumption or food security. Related work by [Sedlmayr et al. \(2020\)](#) and [Burchi and Strupat \(2018\)](#) arrive at similar conclusions. None of these studies experimentally tests variations in coaching modalities or intensities. While our paper does not evaluate the impact of coaching itself, we demonstrate that adjusting the mode and intensity of coaching delivery can substantially reduce costs without compromising effectiveness.

2 Context and interventions

2.1 Context

The Uganda program was implemented by the AVSI Foundation as part of a USAID Resilience Food Security Activity (RFSA). The 2019 program targeted households in the Rwamwanja refugee settlement and surrounding host communities. Eligible households were identified through community consultations and household assessments by the AVSI Foundation.

The Philippines program was developed and implemented by BRAC and the Philippine Department of Labor and Employment (DOLE). DOLE provided financial support for group livelihoods as an extension of its existing DOLE Integrated Livelihood and Emergency Program (DILEEP). BRAC recruited households from a list of those who had begun receiving the government’s conditional cash transfer program (“4Ps”) in 2015–2017. It then organized participants into groups, procured asset packages, conducted livelihood training, and

provided regular coaching. Assets were transferred in 2019.

The Bangladesh program was also developed and implemented by BRAC. The 2016 BRAC Ultra-Poor Graduation Program (UPGP) targeted the poorest households in selected rural upazilas (subdistricts). Eligibility was determined through a community poverty ranking exercise and predefined eligibility criteria.

2.2 Interventions

Program components are similar across all three sites, comprising an asset transfer valued at US\$300-500 (PPP 720 in the Philippines, PPP 850 in Uganda, and PPP 1180 in Bangladesh), bi-monthly regular coaching sessions, skills training and livelihood support, and savings and financial inclusion activities.¹ The Uganda and Bangladesh programs also included monthly consumption support. Appendix C provides full intervention details across all three sites.

Households in each site were assigned with equal probability to one of two (or three) treatment arms or a control arm:

T1 Graduation package, group-level coaching

T2 Graduation package, individual-level coaching

T3 Graduation package, light individual-level coaching (*Bangladesh only*)

C Control group with no program

Aside from coaching modality, households in T1–T3 received identical program elements, as described above. The coaching curriculum was identical between the individual and group versions of the program. In group coaching sessions, participants from the same implementation group met at a central location for each session. In the individual coaching version, the coach instead visited each participant at their homes. However, some program elements, such as initial program sensitization and livelihood trainings, were delivered in groups of 10–25 participants for both coaching modalities.

In T1, participants in Uganda and the Philippines received all coaching in fortnightly group sessions. In Bangladesh, the group-level coaching package used a hybrid approach in which participants alternated between group and individual coaching sessions every week,

¹We report all monetary values in 2024 PPP, converting from local currency units to PPP dollars as of the year of endline outcome measurement: 2021 for Uganda and the Philippines and 2022 for Bangladesh (Bank, 2025), and then adjusting to 2024 using the US CPI (U.S. Bureau of Labor Statistics, 2025). One 2021 PPP dollar equals 1315.27 Ugandan shillings, 20.75 Philippine Pesos, and 0.89 2024 PPP dollars. One 2022 PPP dollar equals 29.13 Bangladesh taka and 0.96 2024 PPP dollars.

balancing cost savings with tailored advice and attention. In T2, participants in Uganda and the Philippines received fortnightly individual coaching sessions.² Participants in Bangladesh received weekly individual coaching sessions. Thus, the intensity of coaching remained constant between T1 and T2, and only the modality differed. In T3, implemented only in Bangladesh, participants received individual coaching similar to T2, but with lower (fortnightly) intensity.

Coaching delivery was similar across contexts, ensuring results are comparable across sites. Coaching was provided by trained staff, and recruitment methods were the same for both coaching modalities. In Uganda, coaches specialized in either group or individual coaching, and in the Philippines and Bangladesh, the same coaches conducted both group and individual sessions. Each implementer had a fixed curriculum that coaches would work through over the sessions, and the curriculum did not differ between modalities. The content of the curriculum was broadly similar across the three studies and included topics like business development, management of the transferred asset, financial literacy, and non-business topics such as health, education, and sanitation.

In the Philippines, we observe a key difference between group and individual coaching arms regarding their choice of livelihoods, as detailed in Appendix Table C.1. BRAC coaches highlighted different livelihood options through the workshops where these selections were made. Since participants in the individual coaching arm (T2) made their selections just after those in the group coaching arm (T1), their choices differed slightly. T1 members were less likely than T2 members to select swine fattening (37% vs. 77%), with a larger share participating in a prepared food stand (21% vs. 9%) and free-range chicken raising (22% vs. 0%). As a result, the comparison between group and individual coaching in the Philippines reflects not only differences in coaching modalities, but also differences in livelihood selection.

Beyond the coaching modality, it is worth noting a few key differences in the other components of the Graduation package across the three study locations, although these do not affect the comparability of treatment arms within each country: (1) in the Philippines and Bangladesh, the asset transfer was delivered as in-kind support, while Ugandan households received cash; (2) consumption support formed part of the intervention in Uganda and Bangladesh, while in the Philippines all study participants—treatment and control households—received consumption support through a pre-existing government program; (3) the Uganda study organized participating households into Village Savings and Loan Associations (VSLAs) to promote greater financial inclusion; and (4) in Bangladesh all treatment groups received some individual coaching—in the group coaching arm, fortnightly individual

²In Uganda, individual coaching sessions were held weekly for the first four months, but the implementer switched to once every two weeks beginning after the first six months.

coaching was delivered alongside fortnightly group-based coaching. Figure 1 summarizes the key characteristics of the three programs.

		Consumption support	Asset transfer	Coaching	Coaching frequency
Uganda	Control				
	Individual	X	Cash, PPP 850	Individual	2x/month
	Group	X	Cash, PPP 850	Group	2x/month
Philippines	Control	X			
	Individual	X	In-kind, PPP 720	Individual	2x/month
	Group	X	In-kind, PPP 720	Group	2x/month
Bangladesh	Control				
	Individual	X	In-kind, PPP 1180	Individual	Weekly
	Group	X	In-kind, PPP 1180	Ind. + Group	Weekly
	Individual Light	X	In-kind, PPP 1180	Individual	2x/month

Figure 1: INTERVENTION COMPONENTS

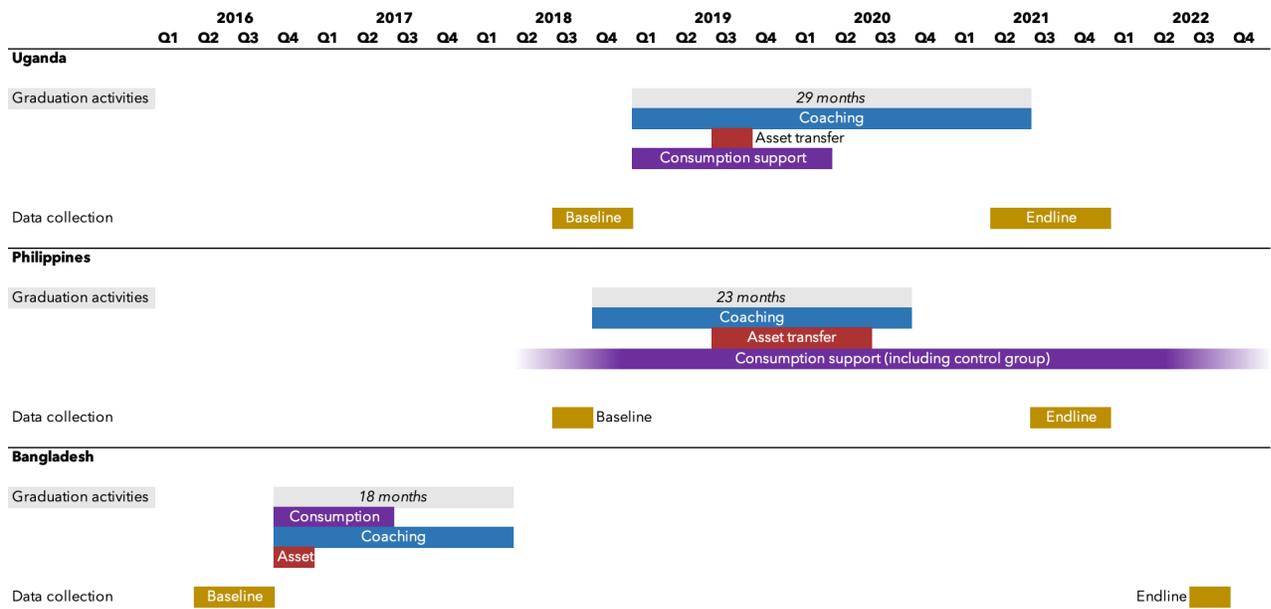


Figure 2: PROJECT AND DATA COLLECTION TIMELINE

The coaching duration varied: 29 months in Uganda, 23 months in the Philippines, and 18 months in Bangladesh. Asset procurement delays in the Philippines extended the transfer period substantially, with some groups receiving transfers one year after the first groups received theirs. Figure 2 shows the duration and timing of key program elements across the three studies.

3 Experimental Design

3.1 Sample selection

All three study sites targeted poor or ultra-poor households in primarily rural areas, using eligibility criteria developed by each implementing partner. Appendix D provides additional details on sampling and randomization.

3.1.1 Uganda

The Uganda site targeted both refugee and host communities, focusing on the entire Rwamwanja Refugee Settlement and four neighboring subcounties in Kamwenge District. AVSI conducted a community-led poverty mapping and household visits to determine eligibility in the study’s 114 village clusters (42 in the settlement and 72 in the surrounding host community). Roughly 70% of households met eligibility criteria: being “extremely poor” or “poor” with a woman or youth as household members who are—or could be—economically active.

3.1.2 Philippines

The site included 2,339 program-eligible households living in 29 *barangays* (the smallest local administrative units) in northern Negros Occidental. These households were enrolled in the Pantawid Pamilya Pilipino Program (4Ps), a government conditional cash transfer program, between 2015–2017. From the 4Ps list, we randomly selected 80 households per barangay for the survey, using a randomly ranked backup list of up to 30 additional households.

3.1.3 Bangladesh

The broader 2016 UPGP was implemented in 126 BRAC branches across 26 districts. For this site, we work only in those 11 districts where a previous version of the program had never been implemented. The sample comprised households of 8,468 female recipients who

completed baseline surveys, drawn from 88 branches. Households were selected as follows: First, the poorest rural *upazilas* (subdistricts) in which BRAC operates were identified using World Food Program poverty mapping tools (Ahmad et al., 2010). Within these *upazilas*, BRAC field staff identified communities with the highest concentration of poverty. They then worked with local communities to list the poorest households via participatory wealth ranking (PWR), which was validated against eligibility criteria using a BRAC survey.

3.2 Randomization

All studies use a cluster-randomized design. However, the Uganda study used a two-level design: we first randomized 115 village clusters to either receive any treatment or serve as pure control. We then randomized the type of graduation program at the household level within treatment village clusters. Randomization clusters in the Philippines are based on dividing each of 29 barangays into four geographically proximate quadrants, with 116 clusters in total. In Bangladesh, we randomized at the branch level (88 clusters), the smallest unit in BRAC’s administrative structure. We also used a fixed re-randomization procedure in Uganda and the Philippines to ensure balance along key baseline characteristics.

In Uganda, there were two additional treatment arms randomly assigned to households within treatment clusters: one that received the graduation package without an asset transfer and one that served as a spillover control group. We include these observations in our main results, using binary indicators for assignment to each of these arms. Control clusters (*not* the control spillover sample) serve as the control group for this site. In the Philippines, a third arm received group livelihood delivery along with group coaching. Substantial delays of these group assets confound the interpretation of this arm. As with Uganda, we include these observations, adding binary indicators for assignment to these arms. In Bangladesh, experimental variation was nested within the group coaching arm (T1): the community-level mobilization element of the graduation package was withheld for randomly selected villages. However, errors in record-keeping prevent us from being confident of the treatment status for all villages, and so we do not include it as a covariate. Appendix A.9 reports results after controlling for this compromised treatment assignment variable.

3.3 Primary outcomes

Our main pre-specified outcomes of interest are consumption, food security, subjective well-being, productive asset values, and household income.³ We report monthly consump-

³In Bangladesh, only consumption and productive asset values were pre-specified as primary outcome variables, and the others were specified as secondary outcome variables.

tion per adult equivalent using 2024 PPP-adjusted values based on 2020 conversion factors (World Bank, 2022). We use a food security index that averages the Food Consumption Score (FCS) and the Household Food Insecurity Scale (HFIAS); reverse-coded so that higher values indicate greater food security), normalized to the control group in each site. We build a subjective well-being index, normalized to the control group of each country. In Uganda and the Philippines, we normalize the average of a reverse-coded 6-item Kessler score and a Cantril’s ladder score from four questions about current and future life satisfaction. In Bangladesh, we employ the same methodology, except that the 10-item Center of Epidemiological Studies Depression Scale is used instead of the Kessler score. Productive asset values and total monthly household income are also reported in 2024 PPP dollars.⁴ Appendix Table B.14 describes each primary outcome in detail.

We also estimate impacts on an overall index that aggregates these five primary outcomes based on their standardized inverse-covariance weighted average (Anderson, 2008).

3.4 Data collection

We collected baseline data approximately six months before randomization and program roll-out in all three sites. Asset transfers took place approximately one year after the baseline in Uganda and the Philippines, and six months after baseline in Bangladesh. In Uganda and Bangladesh, asset deliveries were completed within two months, while in the Philippines asset deliveries were ongoing until 12 months after program commencement. In all settings coaching started within six months of the baseline.

In Uganda and the Philippines, the endline survey was conducted roughly two years after the asset transfer.⁵ In Bangladesh, the endline survey was conducted approximately six years after the asset transfer.⁶ Figure 2 shows the full study timeline.

3.5 Program take-up and participation

In Uganda, nearly all households assigned to treatment (95%) participated in the program, with balance across treatment arms (see Appendix Figure A.1). Initial participation rates were lower in the Philippines, at 80%, but were balanced across treatment arms.⁷ By

⁴Asset and income sample sizes in the Uganda sample are smaller because these modules were administered to a randomly selected subset of respondents

⁵In Uganda, the endline survey was stalled for two months mid-implementation due to COVID restrictions.

⁶In Bangladesh, both baseline and endline surveys were completed before disruptions associated with the seasonal monsoons.

⁷This lower rate reflects weaker initial targeting, as 11% of initially screened households were later deemed ineligible by BRAC because they were receiving another livelihood program, their economic status

the end of the program, however, the difference in participation rates between individual and group coaching arms is statistically significant, at 75% for individual arms and 70% for group arms. In Bangladesh, 91% of selected beneficiaries received assistance from BRAC, while 6% of control households also reported receiving an asset from BRAC. Participation rates are balanced across treatment arms.

3.6 Balance and attrition

Appendix Tables A.1, A.2, and A.3 show no systematic imbalance across treatment arms for nearly all covariates. In all cases except for the group treatment arm in the Philippines, we fail to reject the null hypothesis of equality. To ensure robustness, we show that our results remain consistent when controlling for covariates using a post-double-selection lasso procedure in Appendix Table A.7.

At endline, we surveyed 98% of households in the Philippines, 94% of households in Uganda, and 87% (after six years) in Bangladesh.⁸ Appendix Table A.6 demonstrates no evidence of differential attrition by treatment arm when comparing individual to group coaching arms or when comparing treatment arms to the control group.

3.7 Empirical specification

We estimate intent-to-treat effects using the following estimation equation:

$$y_{ij} = \alpha + \beta_0 + \beta_1 \text{GroupCoach}_{ij} + \beta_2 \text{IndCoach}_{ij} + \beta_3 \text{IndLightCoach}_{ij} + \rho_j + X' \gamma + T' \phi + \epsilon_{ij} \quad (1)$$

where y_{ij} is outcome y measured for household i in cluster j . Our main pre-specified outcomes are consumption, food security, productive asset values, income, and well-being. The *IndLightCoach* indicator is only included in the Bangladesh specification. We control for stratification-cell fixed effects, ρ_j . X is a vector of re-randomization covariates used in Uganda and the Philippines. T is a vector of assignment to other treatment arms: in Uganda, this is assignment to the no-asset and the spillover control arms within treatment

had improved beyond the target range, or they were now located outside the target area. Among those who did not participate, roughly half (9% overall) declined the initial invitation, most citing lack of time. Figures are based on BRAC reports.

⁸For Bangladesh, we impose the most conservative definition of continuing sample members, excluding recipients whose households split to form two or more new households. A 13% attrition rate is comparable to other TUP studies in Bangladesh, such as Balboni et al. (2022); Bandiera et al. (2017), which use similar definitions of continuing sample members.

villages, such that we compare treated households to households in pure control villages. In the Philippines, this is assignment to the group livelihood arm.

In Uganda, we cluster standard errors at the “village cluster” level, as we randomly selected village clusters to receive any intervention or to be part of a pure control group. When we test for differences in impacts between group and individual coaching, we do not cluster our standard errors, reflecting household randomization to coaching modality *within* treatment village clusters.⁹ In the Philippines, we cluster standard errors at the “quadrant,” or “sub-barangay” level. And in Bangladesh, we cluster standard errors at the branch office level. In Appendix Table A.8 we show that our results are robust to multiple hypothesis test corrections, reporting Anderson’s sharpened q-values in brackets below the estimated standard errors (Anderson, 2008).

4 Results

4.1 Average effects

Graduation leads to substantial improvements across nearly all primary outcome indicators relative to the control group, and we find no evidence of differential impacts by the type of coaching received in all three sites. Null results on the hypothesis of no difference are precisely estimated.

Figure 3 presents the point estimates on each coefficient, along with a 95% confidence interval, relative to a standardized control group mean across the key pre-specified outcome variables. Table 1 presents the coefficient estimates for the aggregate index (Column 1) and each subsequent outcome, along with the p-value from testing whether the group and individual coaching estimates are statistically different.¹⁰ The first three panels of Table 1 does this separately for the three sites. Panel D presents results from a fixed-effects meta analysis, pooling results across the three sites. We weight the estimate from each site by the inverse of the sampling variance of its estimated treatment effect.

⁹Appendix Table A.13 demonstrates that our results are unaffected if we instead exclude pure control villages.

¹⁰Appendix Table A.7 shows that our results are robust to estimation controlling for baseline covariates selected using double lasso regression (Urminsky et al., 2016).

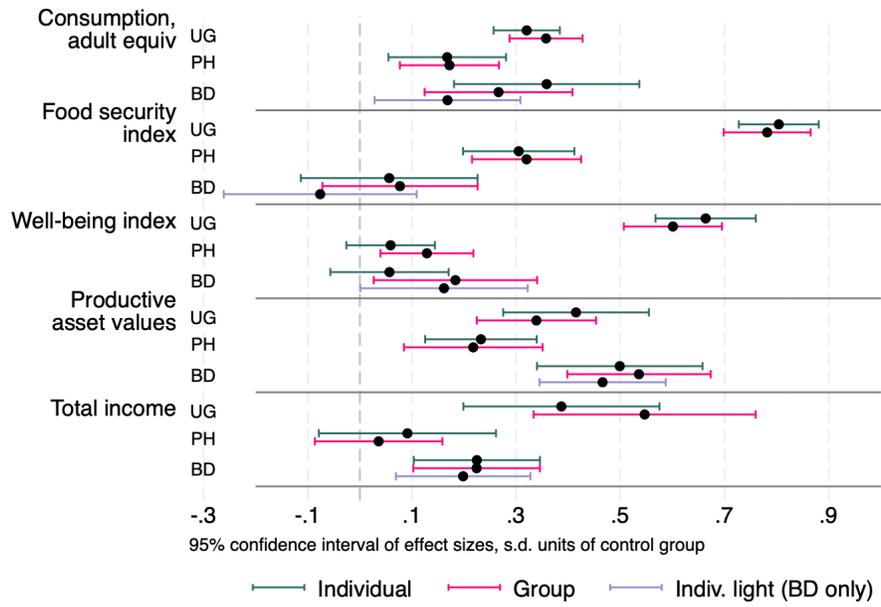


Figure 3: STANDARDIZED PROGRAM IMPACTS ON PRIMARY HOUSEHOLD OUTCOMES

All three programs yield substantial improvements in household economic well-being, particularly when considering our aggregate outcome index as well as household consumption. These results are statistically significant at the 1% level after adjusting for multiple hypothesis testing. (See Appendix Table A.8 for results alongside Anderson q -values.) Additionally, food security improves by 0.79–0.80 s.d. in Uganda and by 0.31–0.32 s.d. in the Philippines, also statistically significant at the MHT-adjusted 1% level. Reflecting the additional cash-transfer component, impacts on consumption and food security are higher in Uganda than in the Philippines, increasing by 28% in Uganda and 9% in the Philippines. In Bangladesh, for which we measure longer-run impacts, consumption rises by 8% (light-intensity individual coaching), 13% (group coaching), and 17% (individual coaching), while there is no detectable impact on food security. These more muted effects are possibly due to general, significant improvements in consumption between baseline and endline, including for the control group, which experienced an inflation-adjusted 25% increase in the value of household consumption.

Impacts on asset holdings are much greater in Bangladesh, at around 0.5 standard deviations relative to the control group. The absolute value of productive assets increases by PPP 543–675 in Uganda, by PPP 228–243 in the Philippines, and by PPP 425–489 in Bangladesh. Monthly household income increases for program participants in all sites, but while we estimate statistically significant average impacts of PPP 66–68 per month in Uganda and PPP 54–61 in Bangladesh, the estimated effect in the Philippines of PPP 62–157 per month is imprecisely estimated and not statistically significant. Additionally, well-being increases in all sites, with the largest effects in Uganda (0.61–0.67 s.d. increase) versus more modest increases in the Philippines (0.06–0.13 s.d.) and Bangladesh (0.06–0.18 s.d.).

Figure 3 shows that the impacts from individual and group coaching are close in magnitude with overlapping confidence intervals. The bottom rows of each panel in Table 1 show that no tests for equality between group and individual arms are statistically significant at the five-percent level. Impacts on consumption are 12% and 3% higher among group coaching recipients in Uganda and the Philippines, respectively, and 26% lower in Bangladesh. However, these differences are not statistically significant ($p = 0.29$ in Uganda, $p = 0.94$ in the Philippines, $p = 0.34$ in Bangladesh). The only exception is the well-being index, for which group coaching yields larger improvements in the Philippines (0.13 s.d. vs 0.06 s.d., $p = 0.137$) and Bangladesh (0.18 s.d. vs 0.06 s.d., $p = 0.10$). However, in Uganda, individual coaching generates greater well-being improvements, though this is noisily estimated (0.67 vs. 0.61 s.d., $p = 0.12$).

At the bottom of each panel of Table 1, we report the group impact per PPP 1,000 spent and the difference between group impact per PPP 1,000 and individual impact per PPP

1,000. We find that group coaching is more cost effective across nearly all outcomes and settings, reflecting that the impact of the group model is usually very close to or even higher than the individual model, while the cost is much less. Of course each treatment effect is estimated with noise. To consider what improvement in cost effectiveness the individual model could yield, we also report the 95% confidence interval of this difference in the ratios of each treatment effect to its cost and report this in the bottom row. The lower bound thus reports the threshold below which we can rule out (at the 5% significance level) the individual implementation model leading to an improvement relative to the group approach. We find that group coaching is more cost-effective in our aggregate index when looking separately by site or in our pooled analysis. Additionally, it is more cost effective in improving consumption, food security, and subjective well-being when we analyze all three sites together, although we cannot always reject at the 5% level when considering these outcomes at the site level (Panels A–C). Together, these results indicate that group coaching is more cost effective on our key welfare outcomes—consumption, food security, and well-being—across all three study sites.

We also test whether halving individual coaching intensity affects program impacts in Bangladesh. Similar to the group-versus-individual coaching comparisons, we find no statistically significant differences in effects when comparing weekly individual coaching to fortnightly individual coaching, apart from a difference in consumption impacts (PPP 20 for weekly coaching vs. PPP 10 for fortnightly coaching, $p = 0.05$). In general, we do not find evidence that light individual coaching is more cost-effective, as the difference in impacts per \$1,000 spent are negative for consumption and food security but positive for the other outcomes, and we cannot reject a difference of zero for any individual or aggregate measures.

Appendix Table A.11 reports graduation impacts on a broader range of economic outcomes, including land ownership, land use, and household labor supply. The Graduation program leads to increased land use and value, particularly in Uganda and Bangladesh, as well as increased household hours worked, particularly in livestock. However, we again find no evidence that coaching modality affects any of these outcomes.

The final rows of Table 1 present the estimated impact per \$1,000 (in 2024 PPP) spent. Compared to individual coaching, group-coaching is more cost-effective, yielding 0.01–0.07 additional standard deviations in our aggregate index per \$1,000 across sites, which is statistically significant at the 5% level.¹¹ Using the pooled sample, we estimate a 95% confidence interval of the group vs. individual difference in cost-effectiveness of [0.005, 0.023]. We can also reject a null of equal cost-effectiveness in the pooled analysis (Panel D) for consumption, food security, and well-being.

¹¹For Bangladesh (Panel C), we also compare regular vs. lighter group-coaching intensity.

Table 1: IMPACT OF GRADUATION PROGRAMS ON PRIMARY OUTCOMES

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Uganda						
	Index of all primary outcomes	Monthly adult equiv consumption (2024 PPP)	Food security index	Wellbeing index	Productive asset values (2024 PPP)	Monthly HH income (2024 PPP)
	(1)	(2)	(3)	(4)	(5)	(6)
Group coaching	0.693*** (0.050)	30.339*** (3.068)	0.785*** (0.044)	0.611*** (0.048)	543.055*** (80.096)	65.630*** (10.243)
Individual coaching	0.710*** (0.048)	27.061*** (2.743)	0.804*** (0.043)	0.673*** (0.051)	675.301*** (103.364)	67.645*** (9.574)
Control mean	-0.00	100.84	-0.00	0.00	513.68	126.08
Control s.d.	1.00	84.76	1.00	1.00	1584.25	154.06
Observations	10263	10509	10509	10453	10430	10337
p-value, group = ind.	0.880	0.287	0.761	0.115	0.324	0.938
Group impact per 1k cost	0.09	3.75	0.10	0.08	67.05	8.10
Group impact per 1k cost - Ind impact per 1k cost	0.01	0.77	0.01	0.00	-7.33	0.65
CI: Group impact per 1k cost - Ind impact per 1k cost	[0.00, 0.02]	[0.18, 1.35]	[0.00, 0.02]	[-0.01, 0.01]	[-29.27, 14.62]	[-1.82, 3.13]
Panel B. Philippines						
Group coaching	0.315*** (0.054)	18.574*** (5.188)	0.320*** (0.053)	0.129*** (0.045)	227.630*** (70.263)	61.780 (106.056)
Individual coaching	0.292*** (0.056)	18.104*** (6.158)	0.305*** (0.054)	0.059 (0.043)	243.047*** (56.563)	156.678 (147.208)
Control mean	0.00	216.39	-0.00	0.00	560.88	715.37
Control s.d.	1.00	108.02	1.00	1.00	1046.01	1715.74
Observations	2288	2287	2288	2288	2288	2288
p-value, group = ind.	0.685	0.939	0.774	0.137	0.824	0.478
Group impact per 1k cost	0.23	13.34	0.23	0.09	163.53	44.38
Group impact per 1k cost - Ind impact per 1k cost	0.07	3.82	0.07	0.06	35.67	-38.04
CI: Group impact per 1k cost - Ind impact per 1k cost	[0.00, 0.15]	[-3.58, 11.22]	[0.00, 0.14]	[0.00, 0.12]	[-56.72, 128.07]	[-192.31, 116.24]
Panel C. Bangladesh						
Group coaching	0.363*** (0.073)	15.174*** (4.070)	0.074 (0.072)	0.181** (0.078)	488.939*** (63.158)	61.221*** (16.698)
Individual coaching	0.299*** (0.074)	20.437*** (5.101)	0.054 (0.082)	0.056 (0.056)	455.296*** (72.933)	61.342*** (16.646)
Individual coaching, light	0.223*** (0.078)	9.588** (4.014)	-0.074 (0.090)	0.159** (0.080)	425.304*** (55.603)	54.135*** (17.742)
Control mean	0.00	117.57	0.01	-0.00	458.70	224.74
Control s.d.	1.00	57.02	0.97	0.99	912.84	273.18
Observations	7565	7534	7534	7534	7447	7445
p-value, group = ind.	0.437	0.337	0.797	0.099	0.619	0.994
Group impact per 1k cost	0.12	4.82	0.02	0.06	155.27	19.44
Group impact per 1k cost - Ind impact per 1k cost	0.04	-0.45	0.01	0.04	37.92	3.63
CI: Group impact per 1k cost - Ind impact per 1k cost	[-0.01, 0.09]	[-3.49, 2.60]	[-0.03, 0.05]	[0.00, 0.09]	[-0.05, 75.90]	[-6.46, 13.72]
p-value, light = ind.	0.376	0.046	0.182	0.169	0.629	0.691
Light impact per 1k cost	0.08	3.51	-0.03	0.06	155.73	19.82
Light impact per 1k cost - Ind impact per 1k cost	0.00	-1.76	-0.04	0.04	38.39	4.01
CI: Light impact per 1k cost - Ind impact per 1k cost	[-0.05, 0.06]	[-4.97, 1.46]	[-0.10, 0.02]	[-0.01, 0.10]	[1.96, 74.81]	[-7.86, 15.88]
Panel D. Pooled, meta-analysis						
Group coaching	0.429*** (0.030)	23.706*** (2.208)	0.486*** (0.031)	0.293*** (0.030)	402.762*** (40.094)	64.356*** (8.697)
Individual coaching	0.427*** (0.029)	23.035*** (2.178)	0.510*** (0.031)	0.243*** (0.028)	382.357*** (40.929)	66.334*** (8.286)
Observations	3	3	3	3	3	3
p-value, group = ind.	0.935	0.747	0.360	0.086	0.629	0.826
Group impact per 1k cost	0.082	4.382	0.099	0.068	105.120	9.592
Group impact per 1k cost - Ind impact per 1k cost	0.014	1.143	0.013	0.015	0.591	1.052
CI: Group impact per 1k cost - Ind impact per 1k cost	[0.005, 0.023]	[0.457, 1.829]	[0.003, 0.023]	[0.002, 0.028]	[-21.537, 22.720]	[-1.434, 3.539]

All dollar values and costs reported in 2024 PPP. Aggregate index based on inverse-covariance weighted average of all five normalized primary outcomes. Monthly consumption includes food, non-durable and durable goods, and 10% of value of household durable assets, reported per adult equivalent. Food security index is normalized average of FCS and negative HFIAS. Subjective well-being index is normalized average of negative Kessler score and a Cantril's ladder score based on averaging four questions about current and future life satisfaction. All indices normalized to endline control group mean. UG asset and income based on imputation of sub-components for respondents randomly selected to answer abridged questionnaires (see Online Appendix for more details). Panel D meta-analysis weights sites by their inverse sampling variance. Robust standard errors are shown in parentheses and clustered at the unit of randomization. UG: village-level randomization for comparisons with control group, with individual-level randomization when testing the equality of group vs. individual coaching impacts. PH: quadrant (sub-barangay)-level randomization. BD: branch-level randomization * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

4.2 Channels of impact

Individual- and group-based coaching may differentially impact program effectiveness through several channels. Individual coaching, delivered at the recipient’s home, may increase engagement and accountability through reduced travel frictions, greater personalization, and direct observation of participants’ living conditions and the state of the transferred asset. Conversely, group coaching may strengthen social networks and facilitate peer learning.

Administrative and self-reported survey data provide little evidence of systematic differences in these mechanisms between coaching modalities, based on a set of proxy measures. Attendance data from Uganda and the Philippines, shown in Appendix Table A.4, reveal no clear pattern between coaching modality and attendance. The small and insignificant differences we observe are insufficient to result in meaningful changes in economic outcomes, as discussed in the section above.

Regarding personalization of content, there are only modest differences in whether participants asked about or discussed a range of topics in Uganda (Appendix Figure A.3). Similarly, in Bangladesh, confidence in performing skills taught during coaching did not differ by modality (Appendix Table A.5).

In terms of fostering stronger social connections, we examine the impact of coaching modality on a social capital index reflecting households’ perceptions about giving and receiving social assistance and trust indices (Appendix Table A.10).¹² We find suggestive evidence that group coaching did modestly improve social capital and trust relative to individual coaching, though the difference is only statistically significant in the Philippines (an increase of 0.12 vs. 0.03 standard deviations, $p = 0.03$), and marginally significant in Bangladesh (a reduction of 0.14 s.d. instead of 0.27 s.d., $p = 0.09$) In terms of trust, differences are modest and statistically insignificant.

4.3 Heterogeneous effects

While we find no differential impacts of group versus individual coaching on average, group coaching could have had heterogeneous and opposing effects relative to individual coaching for different groups of participants, effectively canceling each other out. To examine this possibility, we apply the generic machine learning approach detailed by Chernozhukov et al. (2020) to estimate group average treatment effects by terciles for each of our five primary outcomes, separately for each project site. We do not find evidence of treatment

¹²We only measure the trust index in Uganda and Bangladesh. Definitions of each index are provided in Appendix B.2.1.

heterogeneity across the broad set of covariates we consider. Appendix E contains additional details on our approach.

Overall, we fail to reject the null hypothesis of equal treatment effects between the top and bottom terciles across outcomes at a conservative 10% significance level. Results are shown in Appendix Figures A.4, A.5, and A.6.¹³ Similarly, we see no evidence of heterogeneous impacts of treatment intensity variation in Bangladesh (see Appendix Figure A.7). One exception is that we find evidence of differential impacts of group coaching on household income, shown in Appendix Figure A.6e. Classification analysis indicates that households with larger assets and higher income at baseline see greater improvements in income with group coaching than individual coaching, while individual coaching leads to larger income improvements among poorer households.

While this analysis cannot entirely rule out the possibility of heterogeneity along other unobserved characteristics, such heterogeneity would have to be based on characteristics uncorrelated with the broad range of potential predictors we consider. Additionally, it indicates that among subgroups that could be readily identified and targeted by policymakers—such as those based on household demographics, initial household economic activity, or asset holdings—here is no evidence of disproportionate harm or benefit from shifting from individual to group coaching.

4.4 Benefit-cost analysis

Having found that group coaching is generally more cost-effective than individual coaching across most measures, we now estimate the overall benefits relative to total program costs. We focus on consumption as an indicator of well-being, also allowing comparability with Banerjee et al. (2015).

Overall, group coaching yields per-participant cost savings of 21% (PPP 523) in the Philippines and 11% in Uganda (\$980) (see Appendix Table F.1). The hybrid group/individual coaching arm in Bangladesh saves 19%, or PPP 731 compared to the individual coaching arm. The lighter intensity individual coaching arm saves 30% (PPP 1149) relative to the standard individual coaching arm.

We calculate program implementation costs by summing reported expenditures on all treatments and staffing along with administrative costs, described in Appendix F, and estimate the cost per household *offered* the program, reflecting that program impacts are measured as intent-to-treat (ITT) effects.¹⁴ We calculate program benefits as the total

¹³Further disaggregation to quartiles or quintiles demonstrates a similar absence of detectable treatment heterogeneity.

¹⁴Because take-up is near universal in Uganda, we use cost per participant.

increase in annual household consumption, following [Banerjee et al. \(2015\)](#) and using the point estimates from [Table 1](#). We assume an annual discount rate of 5%, and we report the benefit-cost ratio at varying persistence rates, as we lack longer-run evidence on impacts. As a benchmark, [Bandiera et al. \(2017\)](#) and [Banerjee et al. \(2021\)](#) find that impacts increased over time, while [Barker et al. \(2024\)](#) finds that effects of a graduation program in Ethiopia fade but remain positive by year 7, with effects on consumption and asset holdings roughly 70% and 45% of the year 3 impacts, respectively. Our results in Bangladesh show, at a minimum, that effects persist for six years.

All three programs yield positive benefit-cost ratios under reasonable assumptions about persistence. In Uganda and Bangladesh, the benefit-cost ratio exceeds one so long as consumption impacts decay by at most 20% per year. In the Philippines, the total benefits exceed costs within the time frame of the program itself. Reflecting equivalent consumption impacts but substantial cost savings, group coaching is consistently more cost-effective than individual coaching in Uganda and the Philippines. As also shown in [Table 1](#), in Bangladesh, the individual coaching arm is marginally more cost-effective than the hybrid group coaching arm or lighter-intensity individual arm due to the larger coefficient estimates. However, we urge caution in over-interpreting this difference, as the estimated consumption benefits for Bangladesh are noisy, and the difference in benefits between treatment arms is not statistically significant (see [Table 1](#)).

Table 2: SITE-LEVEL BENEFIT-COST ANALYSIS

	Uganda		Philippines		Bangladesh		Ind. Light
	Individual	Group	Individual	Group	Individual	Group	
<i>Panel A: Estimated Program Costs</i>							
Cost per household at year 1	\$9,079	\$8,099	\$1,901	\$1,392	\$3,531	\$2,866	\$2,486
Cost per household discounted (5% per year)	\$10,060	\$8,974	\$2,107	\$1,543	\$5,014	\$4,070	\$3,530
<i>Panel B: Estimated Program Benefits</i>							
Monthly consumption per adult equiv	\$27 [\$22, \$32]	\$30 [\$24, \$36]	\$18 [\$6, \$30]	\$19 [\$8, \$29]	\$20 [\$10, \$30]	\$15 [\$7, \$23]	\$10 [\$2, \$17]
Annual total household consumption	\$1,348 [\$1,080, \$1,616]	\$1,512 [\$1,212, \$1,811]	\$861 [\$287, \$1,435]	\$884 [\$400, \$1,367]	\$620 [\$317, \$924]	\$461 [\$218, \$703]	\$291 [\$52, \$530]
Total benefits at persistence of ...							
100%	\$24,336	\$27,284	\$15,544	\$15,948	\$11,199	\$8,315	\$5,254
90%	\$7,553	\$8,468	\$4,824	\$4,949	\$3,476	\$2,581	\$1,631
80%	\$4,056	\$4,547	\$2,591	\$2,658	\$1,867	\$1,386	\$876
<i>Panel C: Cost Benefit Ratios</i>							
Benefit-cost ratio at persistence of ...							
100%	2.9 [2.3, 3.4]	3.6 [2.9, 4.3]	7.4 [2.5, 12.4]	10.4 [4.7, 16.1]	3.1 [1.6, 4.6]	2.8 [1.3, 4.3]	2.1 [0.4, 3.7]
90%	1.2 [0.9, 1.4]	1.5 [1.2, 1.8]	3.1 [0.9, 4.7]	4.4 [1.8, 6.1]	1.5 [0.8, 2.3]	1.4 [0.7, 2.2]	1.0 [0.2, 1.9]
80%	0.8 [0.7, 1.0]	1.0 [0.8, 1.3]	2.1 [0.6, 3.1]	2.9 [1.2, 4.0]	1.2 [0.6, 1.8]	1.1 [0.5, 1.7]	0.8 [0.1, 1.5]

Notes: All values converted to 2024 PPP. Calculations based on a social discount rate of 5%. Year 1 and 2 (and in Bangladesh's case also 3, 4, and 5) benefits are extrapolated based on Year 3 (in Bangladesh's case, Year 6) estimates, with no Year 1 benefit assumed in the Philippines because participants received the asset transfer and primary program supports in late Year 1 or early Year 2. Costs are inflated to year 3 in Uganda and the Philippines, and to year 6 in Bangladesh. 95% confidence intervals for program impacts in square brackets.

5 Conclusion

We conduct randomized controlled trials in Uganda, the Philippines, and Bangladesh to measure the impact of providing group versus individual coaching as part of multifaceted social protection (“graduation”) programs. All three programs show large improvements in key economic outcomes: consumption, food security, subjective well-being, and asset holdings. We find that benefits exceed program costs under reasonable assumptions about persistence. In the Philippines, where costs are the lowest of the three sites, the benefits accrued in years 2 and 3 alone already exceeded implementation costs.

Additionally, we find that delivering coaching in groups yields substantial cost reductions (11–21%) without any negative impacts on program effectiveness on average. Given the reduced costs of implementing group coaching, we find that the benefit per dollar spent is greater for group than individual coaching in aggregate and specifically for increasing consumption, food security, and subjective well-being. Reducing coaching intensity, which we test in Bangladesh, reduces costs by 30% but shows mixed effects on cost-effectiveness: confidence intervals for the difference in benefit per \$1,000 spent are wide and generally include zero, although point estimates, particularly for consumption, tend to favor regular intensity coaching. Furthermore, we do not find any evidence of differences in impacts among most- or least-benefited groups across any primary outcome, indicating that shifting from individual to group coaching is unlikely to harm, for example, the poorest or least socially connected households.

As governments look to existing evidence to develop and scale similar programs, our results indicate that programs can improve cost-effectiveness by shifting to group coaching modality, which in turn may also reduce logistical constraints, such as reducing the total number of coaches that need to be found, trained, and monitored.

References

- Ahmad, N., F. Ahmed, D. M. Jolliffe, M. A. Khan, M. A. Mahbub, I. Sharif, V. Swaroop, N. Yoshida, S. Zaidi, and J. Zutt (2010). Poverty maps of Bangladesh. (90487).
- Al-Ubaydli, O., J. A. List, and D. L. Suskind (2017). What can we learn from experiments? Understanding the threats to the scalability of experimental results. *American Economic Review* 107(5), 282–286.
- Anderson, M. L. (2008). Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects. *Journal of the American Statistical Association* 108(484), 1481–1495.
- Andrews, C., A. de Montesquiou, I. Arevalo Sanchez, P. V. Dutta, B. V. Paul, S. Samaranyake, J. Heisey, T. Clay, and S. Chaudhary (2021, January). *The State of Economic Inclusion Report 2021: The Potential to Scale*. The World Bank.
- Athey, S. and S. Wager (2019). Estimating treatment effects with causal forests: An application. *Observational Studies* 5(2), 37–51.
- Balboni, C., O. Bandiera, R. Burgess, M. Ghatak, and A. Heil (2022). Why do people stay poor? *The Quarterly Journal of Economics* 137(2), 785–844.
- Bandiera, O., R. Burgess, N. Das, S. Gulesci, I. Rasul, and M. Sulaiman (2017, May). Labor markets and poverty in village economies. *The Quarterly Journal of Economics* 132(2), 811–870.
- Banerjee, A., E. Duflo, N. Goldberg, D. Karlan, R. Osei, W. Parienté, J. Shapiro, B. Thuysbaert, and C. Udry (2015, May). A multifaceted program causes lasting progress for the very poor: Evidence from six countries. *Science* 348(6236), 1260799.
- Banerjee, A., E. Duflo, N. Goldberg, D. Karlan, R. Osei, W. Parienté, J. Shapiro, B. Thuysbaert, and C. Udry (2015). A multifaceted program causes lasting progress for the very poor: Evidence from six countries. *Science* 348(6236), 1260799.
- Banerjee, A., E. Duflo, and G. Sharma (2021). Long-term effects of the targeting the ultra poor program. *American Economic Review: Insights* 3(4), 471–86.
- Banerjee, A., R. Hanna, B. A. Olken, and D. S. Lisker (2024). Social protection in the developing world. *Journal of Economic Literature* 62(4), 1349–1421.
- Banerjee, A., D. Karlan, R. Osei, H. Trachtman, and C. Udry (2022). Unpacking a multifaceted program to build sustainable income for the very poor. *Journal of Development Economics* 155, 102781.
- Bank, T. W. (2025). PPP conversion factor, private consumption (LCU per international \$). <https://data.worldbank.org/indicator/PA.NUS.PRVT.PP?locations=UG-PH-BD>. Accessed on 12 March, 2025.

- Barker, N., D. Karlan, C. Udry, and K. Wright (2024). The fading treatment effects of a multifaceted asset-transfer program in Ethiopia. *American Economic Review: Insights* 6(2), 277–294.
- Beam, E., L. Brune, D. Karlan, Y. Sawada, and C. Udry (2020). Ultra poor graduation pilot in philippines. AEA RCT Registry.
- Blattman, C., E. P. Green, J. Jamison, M. C. Lehmann, and J. Annan (2016, April). The returns to microenterprise support among the ultrapoor: A field experiment in postwar Uganda. *American Economic Journal: Applied Economics* 8(2), 35–64.
- Bossuroy, T., M. Goldstein, B. Karimou, D. Karlan, H. Kazianga, W. Parienté, P. Premand, C. C. Thomas, C. Udry, J. Vaillant, et al. (2022). Tackling psychosocial and capital constraints to alleviate poverty. *Nature* 605(7909), 291–297.
- Brooks, W., K. Donovan, and T. R. Johnson (2018). Mentors or teachers? Microenterprise training in Kenya. *American Economic Journal: Applied Economics* 10(4), 196–221.
- Brune, L., N. Goldberg, D. Karlan, D. Parkerson, and C. Udry (2024). Randomized impact evaluation of a graduation program on livelihoods in refugee and host communities in uganda. AEA RCT Registry.
- Burchi, F. and C. Strupat (2018). Unbundling the impacts of economic empowerment programmes: Evidence from Malawi.
- Cai, J. and A. Szeidl (2018). Interfirm relationships and business performance. *The Quarterly Journal of Economics* 133(3), 1229–1282.
- Chernozhukov, V., M. Demirer, E. Duflo, and I. Fernández-Val (2020, December). Generic machine learning inference on heterogenous treatment effects in randomized experiments.
- Crosta, T., D. Karlan, F. Ong, J. Rüschenpöhler, and C. R. Udry (2024). Unconditional cash transfers: A bayesian meta-analysis of randomized evaluations in low and middle income countries. Technical report, National Bureau of Economic Research.
- Das, N., S. Dercon, I. Matin, A. Pople, N. Sulaiman, and R. Zizzamia (2022). The role of coaching within bundled ultra-poor graduation programmes: Experimental evidence from bangladesh. AEA RCT Registry.
- Leight, J., H. Alderman, D. Gilligan, M. Hidrobo, and M. Mulford (2023). Can a light-touch graduation model enhance livelihood outcomes? Evidence from Ethiopia. Technical report, International Food Policy Research Institute, Washington, DC.
- McKenzie, D., C. Woodruff, K. Bjorvatn, M. Bruhn, J. Cai, J. Gonzalez-Uribe, S. Quinn, T. Sonobe, and M. Valdivia (2023, September). Training entrepreneurs. *VoxDevLit* 1(3).
- Miles, K. P., K. E. McFadden, D. Colenbrander, and L. C. Ehri (2022). Maximising access to reading intervention: Comparing small group and one-to-one protocols of Reading Rescue. *Journal of Research in Reading* 45(3), 299–323.

- Mühlberger, M. D. and E. Traut-Mattausch (2015). Leading to effectiveness: Comparing dyadic coaching and group coaching. *The Journal of Applied Behavioral Science* 51(2), 198–230.
- Roelen, K. and S. Devereux (2019). Money and the message: The role of training and coaching in graduation programming. *The Journal of Development Studies* 55(6), 1121–1139.
- Ross, S. G. and J. C. Begeny (2011). Improving Latino, English language learners’ reading fluency: The effects of small-group and one-on-one intervention. *Psychology in the Schools* 48(6), 604–618.
- Sedlmayr, R., A. Shah, and M. Sulaiman (2020). Cash-plus: Poverty impacts of alternative transfer-based approaches. *Journal of Development Economics* 144, 102418.
- Urminsky, O., C. Hansen, and V. Chernozhukov (2016). Using double-lasso regression for principled variable selection. *Available at SSRN 2733374*.
- U.S. Bureau of Labor Statistics (2025). Consumer price index for all urban consumers (cpi-u). U.S. Bureau of Labor Statistics. Series Id: CUUR0000SA0, Not Seasonally Adjusted, All items in U.S. city average, all urban consumers, Base Period: 1982-84=100.
- Welz, M., A. Alfons, M. Demirer, and V. Chernozhukov (2022). *GenericML: Generic Machine Learning Inference*. R package version 0.2.3.

A Appendix Figures and Tables

Figure A.1: PROGRAM PARTICIPATION OVER TIME

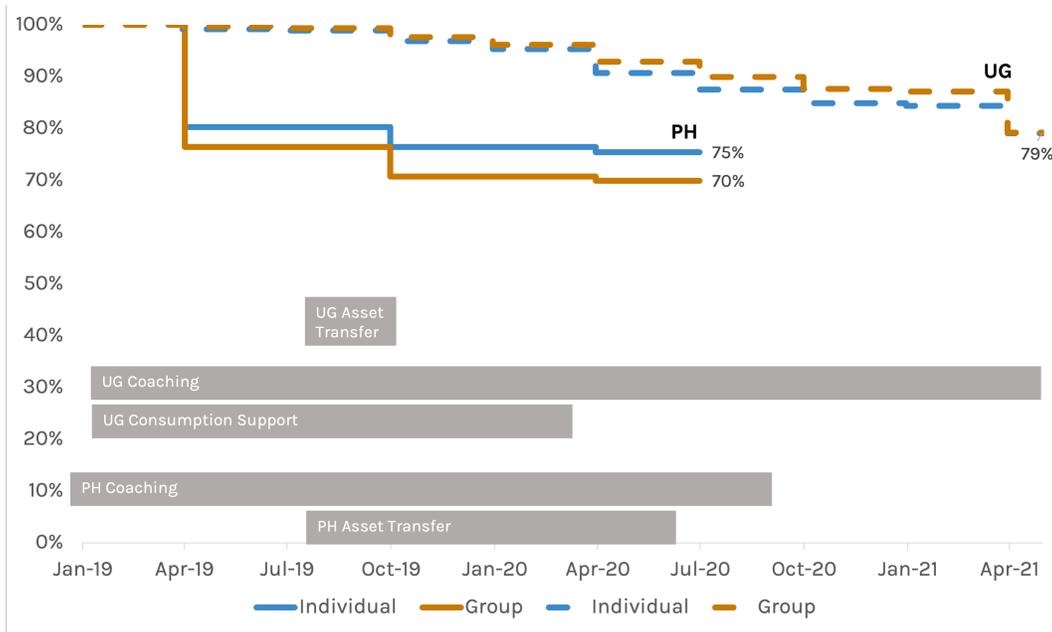
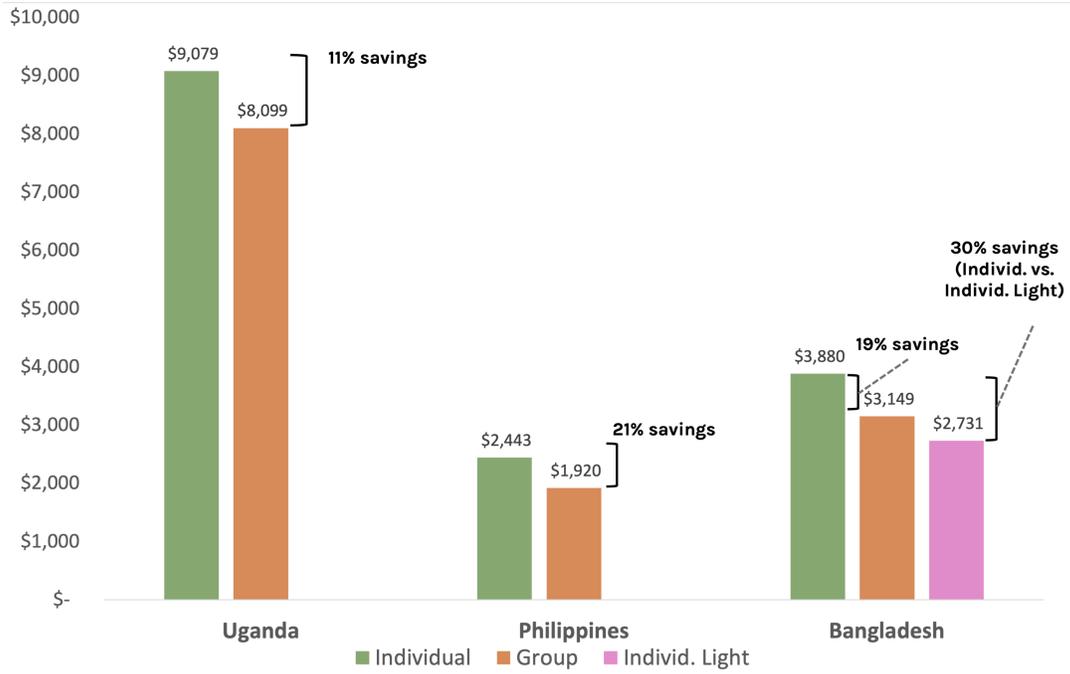


Figure A.2: PROGRAM COST PER PARTICIPANT



All values reported in 2024 PPP.

Figure A.3: TOPICS DISCUSSED AND ASKED ABOUT, UGANDA

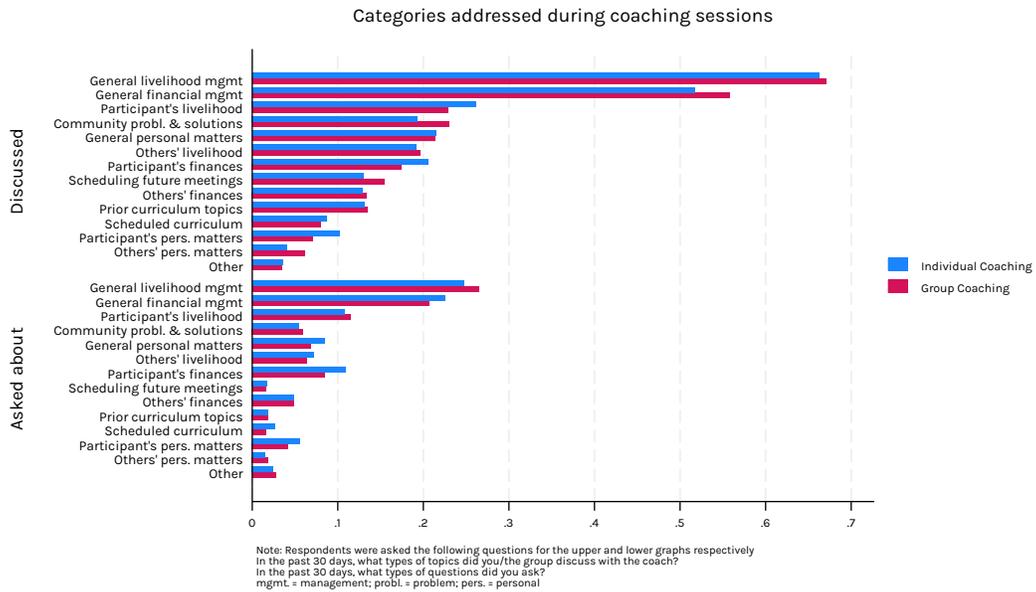
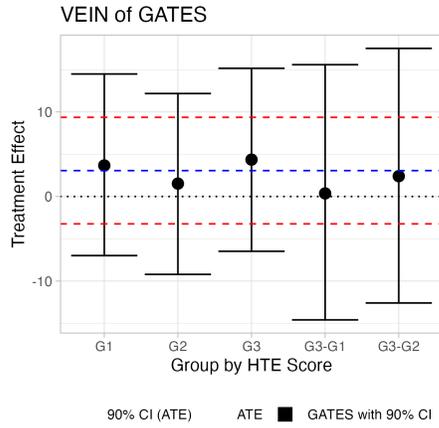
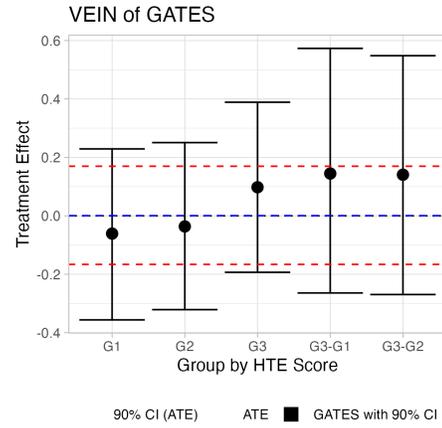


Figure A.4: HETEROGENEOUS TREATMENT EFFECTS BY TERCILE, PRIMARY OUTCOMES, UGANDA

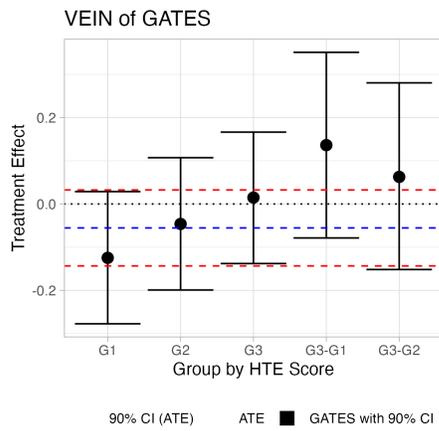
(a) MONTHLY ADULT EQUIV CONSUMPTION



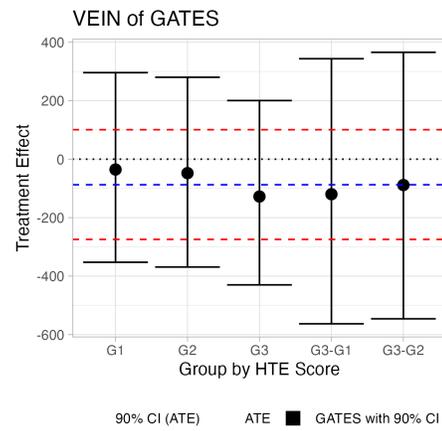
(b) FOOD SECURITY INDEX



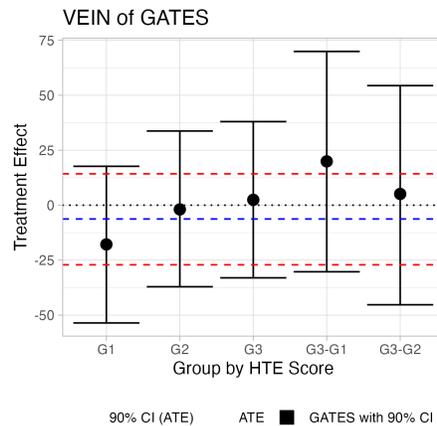
(c) WELL-BEING INDEX



(d) PRODUCTIVE ASSET VALUES



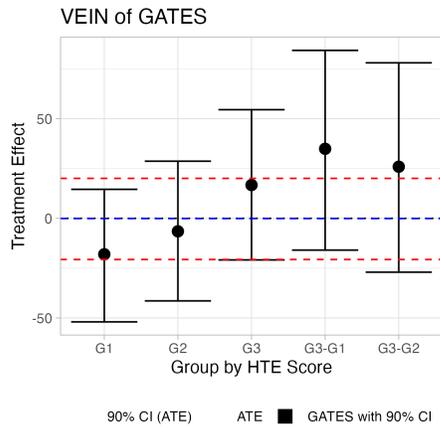
(e) MONTHLY HOUSEHOLD INCOME



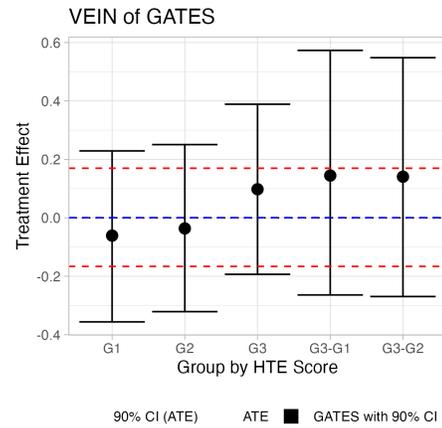
Note: All values reported in 2024 PPP.

Figure A.5: HETEROGENEOUS TREATMENT EFFECTS BY TERCILE, PRIMARY OUTCOMES, PHILIPPINES

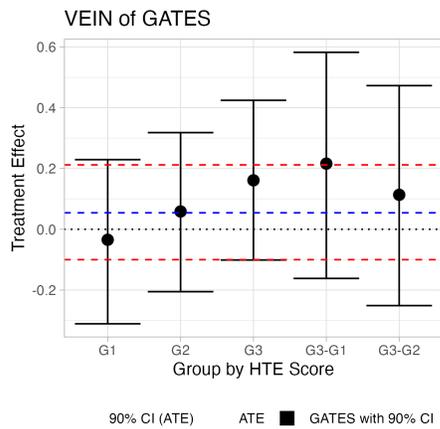
(a) MONTHLY ADULT EQUIV CONSUMPTION



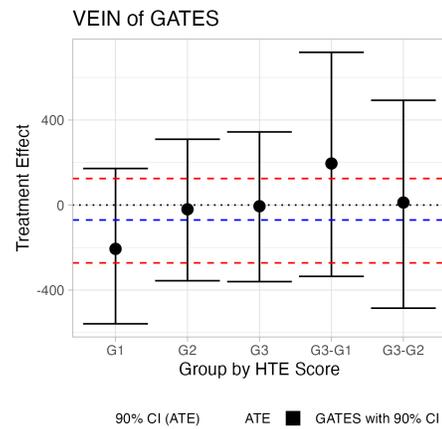
(b) FOOD SECURITY INDEX



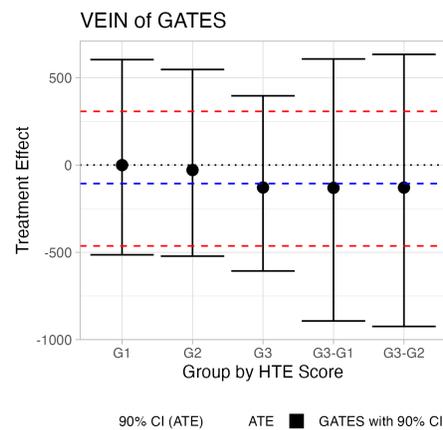
(c) WELL-BEING INDEX



(d) PRODUCTIVE ASSET VALUES



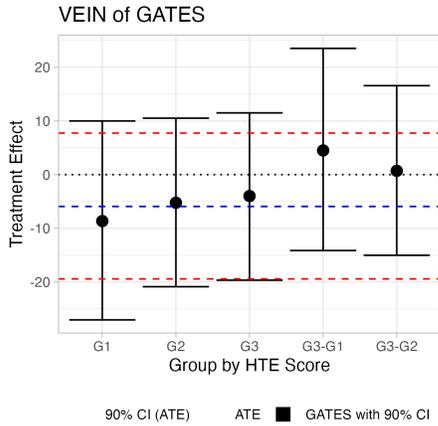
(e) MONTHLY HOUSEHOLD INCOME



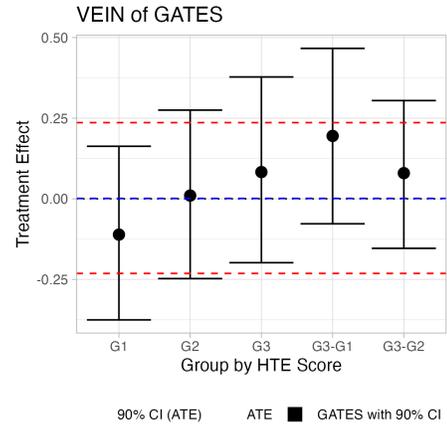
Note: All values reported in 2024 PPP.

Figure A.6: HETEROGENEOUS GROUP VS. INDIVIDUAL TREATMENT EFFECTS BY TERCILE, PRIMARY OUTCOMES, BANGLADESH,

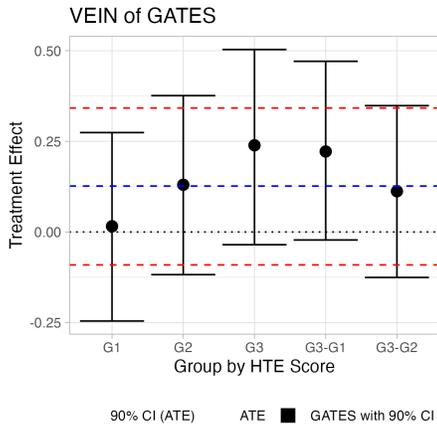
(a) MONTHLY ADULT EQUIV CONSUMPTION



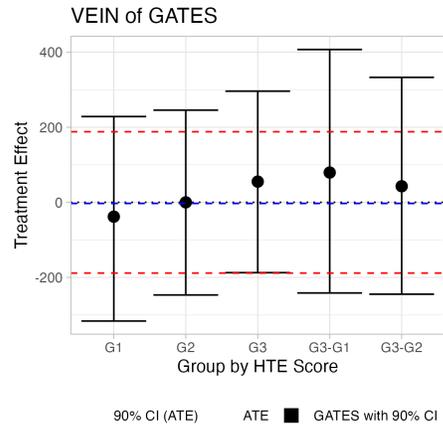
(b) FOOD SECURITY INDEX



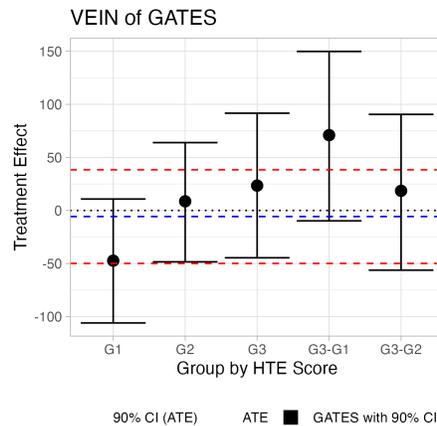
(c) WELL-BEING INDEX



(d) PRODUCTIVE ASSET VALUES



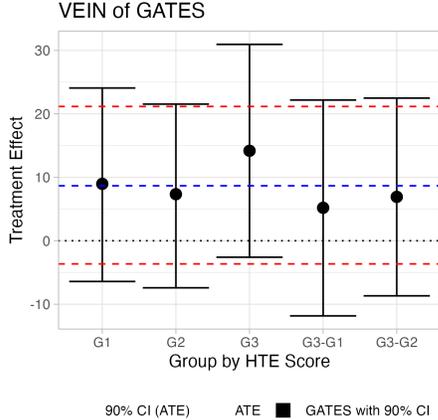
(e) MONTHLY HOUSEHOLD INCOME



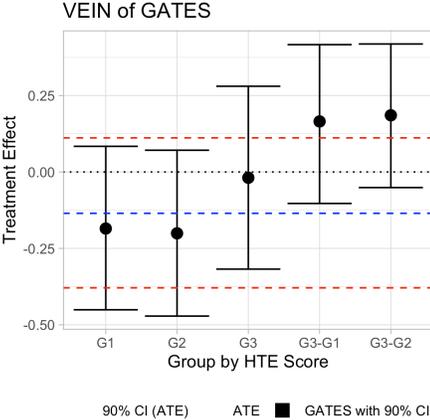
Note: All values reported in 2024 PPP.

Figure A.7: HETEROGENEOUS INDIVIDUAL VS. INDIVIDUAL LIGHT TREATMENT EFFECTS BY TERCILE, PRIMARY OUTCOMES, BANGLADESH,

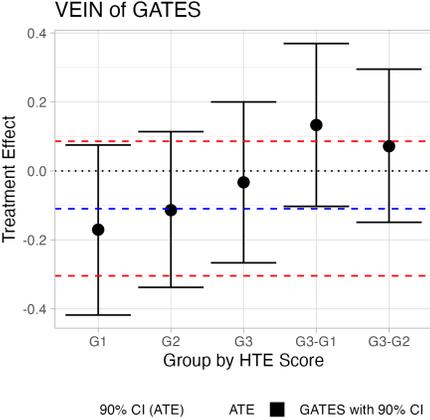
(a) MONTHLY ADULT EQUIV CONSUMPTION



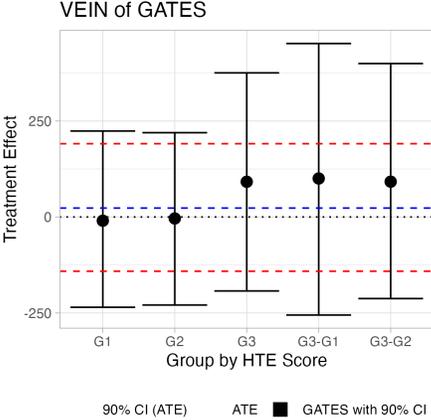
(b) FOOD SECURITY INDEX



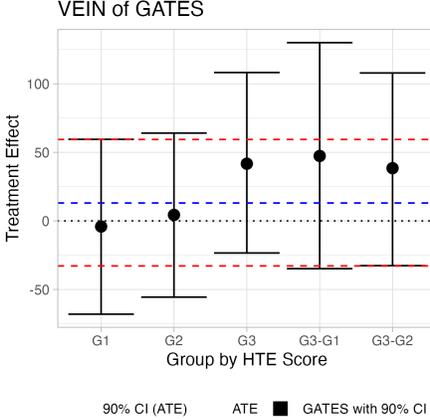
(c) WELL-BEING INDEX



(d) PRODUCTIVE ASSET VALUES



(e) MONTHLY HOUSEHOLD INCOME



Note: All values reported in 2024 PPP.

Table A.1: BALANCE TABLE – UGANDA

Variable	(1) Control Mean/(SE)	(2) Group Mean/(SE)	(3) Individual Mean/(SE)	F-test for balance across all groups F-stat/P-value
HH members	5.756 (0.069)	5.713 (0.059)	5.744 (0.064)	0.085 0.919
Age HH head	39.489 (0.658)	39.190 (0.552)	39.785 (0.611)	0.795 0.454
Female head	0.547 (0.014)	0.553 (0.011)	0.544 (0.012)	0.302 0.740
HH head completed primary	0.175 (0.022)	0.167 (0.015)	0.179 (0.010)	0.609 0.546
Children under 18 in HH	3.513 (0.059)	3.489 (0.047)	3.504 (0.046)	0.028 0.973
Any income selling crops	0.867 (0.040)	0.899 (0.016)	0.897 (0.012)	0.523 0.594
Skipped any meals	0.669 (0.019)	0.651 (0.022)	0.662 (0.016)	0.658 0.520
F-test of joint significance (F-stat)		0.912	0.984	
Number of observations	2172	2113	2079	6364
Number of clusters	57	57	57	114

Table A.2: BALANCE TABLE – PHILIPPINES

Variable	(1) Control Mean/(SE)	(2) Group Mean/(SE)	(3) Individual Mean/(SE)	F-test for balance across all groups F-stat/P-value
HH members	5.939 (0.084)	5.842 (0.101)	5.974 (0.104)	1.023 0.364
Age HH head	45.739 (0.446)	45.404 (0.433)	45.901 (0.501)	0.540 0.584
Female head	0.217 (0.019)	0.189 (0.018)	0.218 (0.022)	1.154 0.320
HH head completed primary	0.661 (0.028)	0.674 (0.022)	0.621 (0.028)	2.288 0.108
Children under 18 in HH	3.019 (0.059)	2.989 (0.073)	3.044 (0.067)	0.286 0.752
Any income selling crops	0.125 (0.022)	0.142 (0.028)	0.143 (0.028)	0.401 0.671
Skipped any meals	0.063 (0.008)	0.093 (0.012)	0.093 (0.013)	3.572** 0.032
F-test of joint significance (F-stat)		4.720***	1.367	
Number of observations	575	571	568	1714
Number of clusters	29	29	29	87

Table A.3: BALANCE TABLE – BANGLADESH

Variable	(1) Control Mean/(SE)	(2) Group Mean/(SE)	(3) Individual Mean/(SE)	(4) Individual Intense Mean/(SE)	F-test for balance across all groups F-stat/P-value
HH members	3.128 (0.158)	3.297 (0.205)	3.259 (0.183)	3.102 (0.175)	3.844** 0.012
Age HH head	47.304 (1.253)	46.899 (1.153)	46.507 (0.912)	47.638 (0.980)	1.029 0.384
Female head	0.548 (0.030)	0.498 (0.037)	0.529 (0.037)	0.519 (0.037)	2.080 0.109
HH head completed primary	0.044 (0.009)	0.048 (0.007)	0.048 (0.008)	0.060 (0.013)	0.969 0.411
Children under 18 in HH	0.708 (0.082)	0.801 (0.113)	0.781 (0.085)	0.675 (0.074)	3.851** 0.012
Skipped any meals	0.632 (0.059)	0.583 (0.077)	0.610 (0.069)	0.578 (0.060)	0.303 0.823
F-test of joint significance (F-stat)		1.387	1.405	1.913	
Number of observations	1723	2328	2389	2115	8555
Number of clusters	22	22	22	22	88

Notes: *p < .1; ** p < .05; *** p < .01.

Table A.4: COACHING ATTENDANCE BY MODALITY, UGANDA AND PHILIPPINES

	Uganda			Philippines		
	Group	Individual	Difference	Group	Individual	Difference
	<i>Last three months</i>			<i>Out of nine sessions</i>		
Any attendance	0.87	0.83	0.04***	0.73	0.79	-0.05**
Avg. attendance	2.74	2.29	0.45***	3.52	5.13	-1.61***
Observations	2,002	1,959	3,961	583	583	1,166

Uganda: Attendance based on self-reports over the past 3 months, while implementation was ongoing. Philippines: Attendance based on administrative data collected by graduation coaching facilitators across 9 coaching sessions and merged with survey data.

Table A.5: CONFIDENCE IN ABILITY TO PERFORM VARIOUS TASKS, BANGLADESH ONLY

	(1) Collecting debts	(2) Educating children	(3) Growing herd	(4) Accessing medical care	(5) Starting business	(6) Protecting business	(7) Resolving dispute	(8) Running business
Group coaching	-0.027 (0.075)	-0.010 (0.074)	0.039 (0.054)	0.013 (0.079)	-0.055 (0.059)	-0.001 (0.072)	-0.054 (0.078)	-0.077 (0.063)
Individual coaching	0.010 (0.079)	0.006 (0.075)	0.048 (0.061)	0.041 (0.084)	0.001 (0.064)	0.055 (0.074)	-0.008 (0.088)	-0.016 (0.066)
Individual coaching, light	0.060 (0.074)	0.092 (0.075)	0.090* (0.054)	0.095 (0.078)	0.034 (0.066)	0.071 (0.074)	0.069 (0.081)	0.009 (0.067)
Control mean at endline	0.58	0.52	0.70	0.56	0.36	0.54	0.51	0.46
P-val, group = ind.	0.61	0.81	0.87	0.71	0.29	0.39	0.56	0.27
P-val, group = ind. light	0.19	0.13	0.28	0.24	0.11	0.27	0.08	0.13
P-val, ind. = ind. light	0.47	0.20	0.45	0.47	0.59	0.81	0.34	0.67
Adjusted R -squared	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
Number of observations	7534	7534	7534	7534	7534	7534	7534	7534

Notes: Outcome variables are dummy variables indicating whether a respondent considers herself capable of performing a particular task. From left to right, these are: Collecting debts, furthering the education of children in one's household, increasing the number of cows one's household owns, accessing medical care, finding new business opportunities, protecting businesses from harm/loss by someone else, resolving an interpersonal dispute, and managing one's own business. Respondents could answer "No, cannot," "Maybe," or "Definitely." We recoded variables so that "No" = 0, and "Maybe" or "Definitely" = 1. Robust standard errors are shown in parentheses and clustered at the unit of randomization, which in Bangladesh was the branch level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.6: RESPONSE RATES BY TREATMENT ASSIGNMENT

	Uganda		Philippines		Bangladesh	
	(1)	(2)	(3)	(4)	(5)	(6)
	No covariates	Covariates	No covariates	Covariates	No covariates	Covariates
Group coaching	0.008 (0.006)	0.009 (0.006)	-0.002 (0.009)	-0.003 (0.008)	-0.000 (0.014)	-0.001 (0.014)
Individual coaching	0.006 (0.006)	0.006 (0.006)	-0.007 (0.008)	-0.009 (0.008)	0.012 (0.013)	0.012 (0.013)
Individual coaching, light					0.017 (0.012)	0.018 (0.012)
Observations	11145	11145	2339	2285	8555	8555
Response rate, control						
P-val, group = ind	0.760	0.750	0.449	0.350	0.271	0.260
P-val, group = ind. light					0.096	0.081
P-val, ind. = ind. light					0.566	0.516
P-val, joint signif.	0.628	0.562	0.811	0.699	0.271	0.245

Notes: Sample includes all randomized baseline respondents. Standard errors are clustered at the unit of randomization. Stratification-cell fixed effects are included in all regressions. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.7: IMPACT OF GRADUATION PROGRAMS ON PRIMARY OUTCOMES, POST DOUBLE SELECTION LASSO

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Uganda						
	Index of all primary outcomes	Monthly adult equiv consumption (2024 PPP)	Food security index	Wellbeing index	Productive asset values (2024 PPP)	Monthly HH income (2024 PPP)
Group coaching	0.681*** (0.049)	29.970*** (3.096)	0.777*** (0.043)	0.600*** (0.045)	551.616*** (71.089)	64.219*** (10.328)
Individual coaching	0.696*** (0.043)	27.170*** (2.581)	0.789*** (0.042)	0.657*** (0.048)	655.435*** (86.933)	64.967*** (8.668)
Control mean	-0.00	100.84	-0.00	0.00	513.68	126.08
Control s.d.	1.00	84.76	1.00	1.00	1584.25	154.06
Observations	10263	10509	10509	10453	10430	10337
p-value, group = ind.	0.880	0.287	0.761	0.115	0.324	0.938
Group impact per 1k cost	0.08	3.70	0.10	0.07	68.11	7.93
Group impact per 1k cost - Ind impact per 1k cost	0.01	0.71	0.01	0.00	-4.08	0.77
CI: Group impact per 1k cost - Ind impact per 1k cost	[0.00, 0.02]	[0.11, 1.31]	[0.00, 0.02]	[-0.01, 0.01]	[-24.80, 16.63]	[-1.71, 3.26]
Panel B. Philippines						
Group coaching	0.318*** (0.058)	20.011*** (5.372)	0.319*** (0.053)	0.125*** (0.046)	230.619*** (74.455)	56.803 (111.690)
Individual coaching	0.305*** (0.061)	17.737*** (6.610)	0.308*** (0.054)	0.059 (0.043)	280.213*** (67.934)	168.208 (152.487)
Control mean	0.02	216.39	0.00	0.00	569.28	715.66
Control s.d.	1.00	108.02	1.00	1.00	1052.52	1717.22
Observations	2287	2287	2287	2287	2287	2287
p-value, group = ind.	0.841	0.734	0.834	0.170	0.520	0.412
Group impact per 1k cost	0.23	14.38	0.23	0.09	165.67	40.81
Group impact per 1k cost - Ind impact per 1k cost	0.07	5.05	0.07	0.06	18.27	-47.68
CI: Group impact per 1k cost - Ind impact per 1k cost	[-0.01, 0.15]	[-2.99, 13.08]	[0.00, 0.13]	[0.00, 0.12]	[-81.29, 117.83]	[-204.70, 109.34]
Panel C. Bangladesh						
Group coaching	0.362*** (0.073)	15.174*** (4.070)	0.074 (0.072)	0.181** (0.078)	489.180*** (63.471)	61.221*** (16.698)
Individual coaching	0.302*** (0.073)	20.437*** (5.101)	0.054 (0.082)	0.056 (0.056)	459.517*** (73.293)	61.342*** (16.646)
Individual coaching, light	0.224*** (0.077)	9.588** (4.014)	-0.074 (0.090)	0.159** (0.080)	426.961*** (55.841)	54.135*** (17.742)
Control mean	0.00	117.57	0.01	-0.00	460.24	224.74
Control s.d.	1.00	57.02	0.97	0.99	913.99	273.18
Observations	7534	7534	7534	7534	7416	7445
p-value, group = ind.	0.470	0.337	0.797	0.099	0.662	0.994
Group impact per 1k cost	0.12	4.82	0.02	0.06	155.34	19.44
Group impact per 1k cost - Ind impact per 1k cost	0.04	-0.45	0.01	0.04	36.91	3.63
CI: Group impact per 1k cost - Ind impact per 1k cost	[-0.01, 0.08]	[-3.49, 2.60]	[-0.03, 0.05]	[0.00, 0.09]	[-1.19, 75.02]	[-6.46, 13.72]
p-value, light = ind.	0.356	0.046	0.182	0.169	0.601	0.691
Light impact per 1k cost	0.08	3.51	-0.03	0.06	156.34	19.82
Light impact per 1k cost - Ind impact per 1k cost	0.00	-1.76	-0.04	0.04	37.91	4.01
CI: Light impact per 1k cost - Ind impact per 1k cost	[-0.05, 0.06]	[-4.97, 1.46]	[-0.10, 0.02]	[-0.01, 0.10]	[1.41, 74.40]	[-7.86, 15.88]
Panel D. Pooled, meta-analysis						
Group coaching	0.420*** (0.029)	21.882*** (2.123)	0.489*** (0.030)	0.305*** (0.029)	400.099*** (37.202)	55.767*** (8.370)
Individual coaching	0.423*** (0.028)	22.340*** (2.000)	0.518*** (0.030)	0.254*** (0.027)	405.616*** (38.113)	62.211*** (7.536)
Observations	3	3	3	3	3	3
p-value, group = ind.	0.896	0.821	0.267	0.067	0.888	0.460
Group impact per 1k cost	0.082	4.192	0.101	0.070	102.238	8.708
Group impact per 1k cost - Ind impact per 1k cost	0.016	0.991	0.014	0.015	0.002	0.742
CI: Group impact per 1k cost - Ind impact per 1k cost	[0.007, 0.025]	[0.293, 1.689]	[0.005, 0.024]	[0.002, 0.027]	[-19.513, 19.518]	[-1.757, 3.240]

All dollar values and costs reported in 2024 PPP. Aggregate index based on inverse-covariance weighted average of all five normalized primary outcomes. Monthly consumption includes food, non-durable and durable goods, and 10% of value of household durable assets, reported per adult equivalent. Food security index is normalized average of FCS and negative HFIAS. Subjective well-being index is normalized average of negative Kessler score and a Cantril's ladder score based on averaging four questions about current and future life satisfaction. All indices normalized to endline control group mean. UG asset and income based on imputation of sub-components for respondents randomly selected to answer abridged questionnaires (see Online Appendix for more details). Panel D meta-analysis weights sites by their inverse sampling variance. Robust standard errors are shown in parentheses and clustered at the unit of randomization. UG: village-level randomization for comparisons with control group, with individual-level randomization when testing the equality of group vs. individual coaching impacts. PH: quadrant (sub-barangay)-level randomization. BD: branch-level randomization * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.8: IMPACT OF GRADUATION PROGRAMS ON PRIMARY OUTCOMES,
MULTIPLE HYPOTHESES ADJUSTMENTS

	(1)	(2)	(3)	(4)	(5)
Panel A. Uganda					
	Monthly adult equiv consumption (2024 PPP)	Food security index	Wellbeing index	Productive asset values (2024 PPP)	Monthly household income (2024 PPP)
Group coaching	30.934*** (2.873) [0.001]	0.783*** (0.042) [0.001]	0.605*** (0.046) [0.001]	513.829*** (99.158) [0.001]	73.974*** (15.336) [0.001]
Individual coaching	28.091*** (2.519) [0.001]	0.803*** (0.039) [0.001]	0.669*** (0.048) [0.001]	657.474*** (116.403) [0.001]	67.658*** (13.820) [0.001]
Control mean	100.84	-0.00	0.00	755.50	127.40
Control s.d.	84.76	1.00	1.00	2077.63	165.90
Observations	10509	10509	10453	5259	2217
p-value, joint test	0.00	0.00	0.00	0.00	0.00
p-value, group = ind.	0.29	0.76	0.12	0.39	0.71
95% CI, Group-Ind	[-7.2, 2.1]	[-0.05, 0.07]	[-0.01, 0.11]	[-131.5, 335.3]	[-41.4, 28.4]
Panel B. Philippines					
Group coaching	18.574*** (5.188) [0.001]	0.320*** (0.053) [0.001]	0.129*** (0.045) [0.004]	227.630*** (70.263) [0.002]	61.780 (106.056) [0.203]
Individual coaching	18.104*** (6.158) [0.004]	0.305*** (0.054) [0.001]	0.059 (0.043) [0.069]	243.047*** (56.563) [0.001]	156.678 (147.208) [0.107]
Control mean	216.39	-0.00	0.00	560.88	715.37
Control s.d.	108.02	1.00	1.00	1046.01	1715.74
Observations	2287	2288	2288	2288	2288
p-value, joint test	0.00	0.00	0.02	0.00	0.57
p-value, group = ind.	0.94	0.77	0.14	0.82	0.48
95% CI, Group-Ind	[-12.6, 11.6]	[-0.12, 0.09]	[-0.16, 0.02]	[-121.9, 152.8]	[-168.9, 358.7]
Panel C. Bangladesh					
Group coaching	15.721*** (4.343) [0.001]	0.083 (0.076) [0.101]	0.184** (0.079) [0.016]	528.022*** (72.410) [0.001]	67.457*** (17.750) [0.001]
Individual coaching	22.021*** (5.425) [0.001]	0.060 (0.086) [0.151]	0.058 (0.058) [0.108]	500.362*** (80.258) [0.001]	68.463*** (18.147) [0.001]
Individual coaching, light	11.006** (4.549) [0.014]	-0.065 (0.093) [0.151]	0.165** (0.082) [0.031]	442.112*** (57.614) [0.001]	76.689*** (24.508) [0.003]
Control mean	118.04	-0.00	-0.00	461.98	226.61
Control s.d.	61.68	1.00	1.00	936.24	287.07
Observations	7534	7534	7534	7447	7445
p-value, joint test	0.00	0.34	0.07	0.00	0.00
p-value, group = ind.	0.28	0.76	0.11	0.74	0.96
p-value, ind. = ind. light	0.07	0.20	0.17	0.42	0.74

Robust standard errors are shown in parentheses and clustered at the unit of randomization. UG: village-level randomization for comparisons with control group, with individual-level randomization when testing the equality of group vs. individual coaching impacts. PH: quadrant (sub-barangay)-level randomization. BD: branch-level randomization. We control for the false discovery rate using Anderson sharpened q-values, which are reported in brackets below the standard errors. All dollar values and costs reported in 2024 PPP. Aggregate index based on inverse-covariance weighted average of all five normalized primary outcomes. Monthly consumption includes food, non-durable and durable goods, and 10% of value of household durable assets, reported per adult equivalent. Food security index is normalized average of FCS and negative HFIAS. Subjective well-being index is normalized average of negative Kessler score and a Cantril's ladder score based on averaging four questions about current and future life satisfaction. All indices normalized to endline control group mean. Panel D meta-analysis weights sites by their inverse sampling variance. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.9: IMPACT OF GRADUATION PROGRAMS IN BANGLADESH ON PRIMARY OUTCOMES, CONTROLLING FOR VILLAGE ASSISTANCE COMMITTEE RANDOMIZATION

	(1)	(2)	(3)	(4)	(5)	(6)
Bangladesh						
	Monthly adult equiv consumption (2024 PPP)	Food security index	Productive asset values (2024 PPP)	Monthly household income (2024 PPP)	Wellbeing index	
Group coaching	0.275*** (0.076)	11.456** (4.878)	-0.015 (0.095)	0.169** (0.073)	406.112*** (79.931)	57.359*** (18.833)
Individual coaching	0.299*** (0.072)	20.439*** (5.086)	0.054 (0.082)	0.056 (0.056)	455.259*** (72.250)	61.343*** (16.588)
Individual coaching, light	0.224*** (0.078)	9.620** (3.997)	-0.073 (0.090)	0.159** (0.080)	426.067*** (56.098)	54.166*** (17.750)
Control mean	0.00	117.57	0.01	-0.00	458.70	224.74
Control s.d.	1.00	57.02	0.97	0.99	912.84	273.18
Observations	7565	7534	7534	7534	7447	7445
p-value, group = ind.	0.775	0.144	0.477	0.097	0.558	0.830
Group impact per 1k cost	0.09	3.64	-0.00	0.05	128.97	18.21
Group impact per 1k cost - Ind impact per 1k cost	0.01	-1.63	-0.02	0.04	11.63	2.40
CI: Group impact per 1k cost - Ind impact per 1k cost	[-0.04, 0.06]	[-5.12, 1.86]	[-0.08, 0.04]	[0.00, 0.08]	[-37.54, 60.80]	[-8.55, 13.36]
p-value, light = ind.	0.381	0.046	0.184	0.169	0.638	0.693
Light impact per 1k cost	0.08	3.52	-0.03	0.06	156.01	19.83
Light impact per 1k cost - Ind impact per 1k cost	0.01	-1.75	-0.04	0.04	38.68	4.02
CI: Light impact per 1k cost - Ind impact per 1k cost	[-0.05, 0.06]	[-4.95, 1.46]	[-0.10, 0.02]	[-0.01, 0.10]	[2.00, 75.35]	[-7.86, 15.90]

All dollar values and costs reported in 2024 PPP. Aggregate index based on inverse-covariance weighted average of all five normalized primary outcomes. Monthly consumption includes food, non-durable and durable goods, and 10% of value of household durable assets, reported per adult equivalent. Food security index is normalized average of FCS and negative HFIAS. Subjective well-being index is normalized average of negative Kessler score and a Cantril's ladder score based on averaging four questions about current and future life satisfaction. All indices normalized to endline control group mean. Robust standard errors are shown in parentheses and clustered at the unit of randomization. BD: branch-level randomization * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.10: IMPACT OF GRADUATION ON SOCIAL CAPITAL AND TRUST

	(1)	(2)
Panel A. Uganda		
	Social capital index	Trust index
Group coaching	0.090* (0.051)	0.101*** (0.037)
Individual coaching	0.032 (0.054)	0.075* (0.039)
Control mean	-0.00	0.00
Control s.d.	1.00	1.00
Observations	5256	10509
p-value, joint test	0.19	0.02
p-value, group = ind.	0.15	0.32
Panel B. Philippines		
	Social capital index	
Group coaching	0.120** (0.051)	
Individual coaching	0.026 (0.052)	
Control mean	0.00	
Control s.d.	1.00	
Observations	2288	
p-value, joint test	0.03	
p-value, group = ind.	0.03	
Panel C. Bangladesh		
	Social capital index	Trust index
Group coaching	-0.142* (0.083)	-0.067 (0.096)
Individual coaching	-0.266*** (0.088)	-0.165** (0.074)
Individual coaching, light	-0.212** (0.098)	-0.131 (0.081)
Control mean	0.00	0.00
Control s.d.	1.00	1.00
Observations	7534	7534
p-value, joint test	0.03	0.14
p-value, group = ind.	0.09	0.24
p-value, ind. = ind. light	0.54	0.62

Robust standard errors are shown in parentheses and clustered at the unit of randomization. UG: village-level randomization for comparisons with control group, with individual-level randomization when testing the equality of group vs. individual coaching impacts. PH: quadrant (sub-barangay)-level randomization. BD: branch-level randomization * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.11: IMPACT OF GRADUATION ON SECONDARY ECONOMIC OUTCOMES

	(1)	(2)	(3)	(4)	(5)
Panel A. Uganda					
	Own/use land	Plot value	Plot size	Hours HH worked, past 7 days	Hours HH worked in livestock, past 7 days
Group coaching	0.030*** (0.010)	1625.067*** (208.478)	1456.945*** (211.565)	6.014* (3.353)	5.340*** (1.030)
Individual coaching	0.028*** (0.010)	1568.185*** (229.898)	1378.453*** (219.206)	6.503** (3.152)	4.538*** (1.034)
Control mean	0.90	3455.55	3830.07	41.89	9.36
Control s.d.	0.30	4447.67	5604.58	43.69	20.14
Observations	10095	10509	10505	2046	5252
p-value, joint test	0.01	0.00	0.00	0.09	0.00
p-value, group = ind.	0.67	0.32	0.31	0.82	0.44
Panel B. Philippines					
	Use land		Plot size	Hours HH worked, past 7 days	Hours HH worked in livestock, past 7 days
Group coaching	0.056** (0.027)		1623.659 (1098.676)	1.404 (2.717)	1.913*** (0.604)
Individual coaching	0.036 (0.029)		1171.421 (1301.189)	3.484 (2.699)	0.778 (0.552)
Control mean	0.31		2272.10	62.20	7.75
Control s.d.	0.46		18593.20	45.97	11.12
Observations	2288		2178	2288	2288
p-value, joint test	0.11		0.31	0.43	0.01
p-value, group = ind.	0.50		0.74	0.43	0.06
Panel C. Bangladesh					
	Own/use land	Plot value		Hours HH worked, past 7 days	Hours HH worked in livestock, past 7 days
Group coaching	0.089** (0.037)	3376.793*** (942.162)		4.703*** (1.552)	3.702*** (0.582)
Individual coaching	0.120*** (0.036)	2718.066** (1065.083)		3.772** (1.657)	2.898*** (0.473)
Individual coaching, light	0.091*** (0.035)	2024.724** (850.443)		3.550** (1.348)	3.445*** (0.541)
Control mean	0.57	6355.50		33.70	4.87
Control s.d.	0.50	11577.81		19.03	7.07
Observations	7534	7416		7445	7445
p-value, joint test	0.01	0.00		0.01	0.00
p-value, group = ind.	0.33	0.54		0.60	0.18
p-value, ind. = ind. light	0.34	0.49		0.89	0.33

All dollar values and costs reported in 2024 PPP. Robust standard errors are shown in parentheses and clustered at the unit of randomization. UG: village-level randomization for comparisons with control group, with individual-level randomization when testing the equality of group vs. individual coaching impacts. PH: quadrant (sub-barangay)-level randomization. BD: branch-level randomization * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.12: IMPACT OF GRADUATION ON FINANCIAL INCLUSION

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Uganda						
	Any savings	Savings value (2024 PPP)	Any borrowing	Outstanding loans (2024 PPP)		
Group coaching	0.090*** (0.008)	192.974*** (24.596)	0.024 (0.018)	43497.813*** (10422.595)		
Individual coaching	0.095*** (0.008)	216.700*** (23.379)	0.038** (0.017)	38799.579*** (10435.654)		
Control mean	0.89	153.61	0.71	162020.85		
Control s.d.	0.32	370.69	0.45	273837.29		
Observations	10509	10407	10509	10509		
p-value, joint test	0.00	0.00	0.09	0.00		
p-value, group = ind.	0.58	0.42	0.39	0.29		
Panel B. Philippines						
	Any savings	Savings value (2024 PPP)	Any borrowing	Outstanding loans (2024 PPP)	Any loans given	Outstanding loans given (2024 PPP)
Group coaching	0.113*** (0.024)	12.882*** (4.650)	-0.017 (0.024)	10.980 (25.880)	0.036*** (0.012)	5.165 (5.454)
Individual coaching	0.152*** (0.027)	23.999*** (5.472)	-0.055** (0.023)	48.901 (30.969)	0.014 (0.014)	20.722* (10.731)
Control mean	0.26	23.05	0.61	250.25	0.06	7.06
Control s.d.	0.44	71.21	0.49	515.58	0.25	55.44
Observations	2288	2268	2288	2288	2286	2288
p-value, joint test	0.00	0.00	0.05	0.28	0.01	0.13
p-value, group = ind.	0.10	0.06	0.08	0.23	0.13	0.16
Panel C. Bangladesh						
	Any savings	Savings value (2024 PPP)	Any borrowing	Outstanding loans (2024 PPP)	Any loans given	Outstanding loans given (2024 PPP)
Group coaching	0.146*** (0.034)	67.709*** (19.087)	0.028** (0.014)	13.710 (8.327)	0.062*** (0.023)	213.031*** (79.660)
Individual coaching	0.066** (0.032)	44.763** (18.401)	0.022 (0.014)	4.032 (7.204)	0.014 (0.023)	150.489** (69.567)
Individual coaching, light	0.055* (0.029)	19.980 (15.718)	0.007 (0.015)	5.396 (8.858)	0.045* (0.023)	160.049** (77.502)
Control mean	0.36	96.84	0.06	24.12	0.57	502.62
Control s.d.	0.48	283.56	0.24	144.73	0.50	1212.06
Observations	7421	8529	7534	8529	7534	8529
p-value, joint test	0.00	0.00	0.18	0.43	0.03	0.05
p-value, group = ind.	0.02	0.26	0.65	0.22	0.03	0.39
p-value, ind. = ind. light	0.72	0.14	0.33	0.87	0.15	0.89

All dollar values and costs reported in 2024 PPP. PH: Whether given or received loans in the past 12 months, with current outstanding balances reported. UG: Savings based on any savings in past 3 months. BD: Savings and loans based on current balances. Robust standard errors are shown in parentheses and clustered at the unit of randomization. UG: village-level randomization for comparisons with control group, with individual-level randomization when testing the equality of group vs. individual coaching impacts. PH: quadrant (sub-barangay)-level randomization. BD: branch-level randomization * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.13: IMPACT OF GROUP VS. INDIVIDUAL COACHING IN UGANDA, FULL SAMPLE AND COACHING ONLY

	All treatment arms		Group and individual coaching only	
	Group coaching (1)	Obs. (2)	Group coaching (3)	Obs. (4)
Monthly adult equiv consumption (2024 PPP)	3.3 (2.5)	10509	-2.5 (2.4)	8337
Food security index	0.01 (0.03)	10509	0.01 (0.03)	8337
Wellbeing index	0.05 (0.03)	10453	0.05 (0.03)	8295
	-132.2 (96.9)	10430	79.8 (80.9)	8270
	-2.0 (10.6)	10337	-0.6 (8.2)	8191

B Variable Definitions

B.1 Primary outcomes

Table B.14: DEFINITION OF PRIMARY OUTCOMES

Primary Outcome	Description	Calculation
Consumption	Monthly adult equiv consumption (PPP 2024)	Per capita household consumption is calculated by summing the following: (1) food consumption over the past seven days (including consumption out of purchases, consumption out of home produce, and consumption received in-kind/free.); (2) non-durable good consumption and frequently purchased services during the last 7 days and 30 days; (3) and semi-durable goods and durable goods and services during the last 12 months; and (4) 10% of the value of household durable assets. Adjusted to adult equivalents and scaled to monthly values (excluding durable assets). Adjusted to 2024 PPP dollars.
Food security	Food security index	In Uganda and the Philippines, normalized average of FCS and negative HFIAS standardized to the control group. In Bangladesh, a 6-item scale with substantial overlap with the HFIAS is used, also as a normalized average standardized to the control group.
Subjective-wellbeing	Subjective well-being index (Cantril and Kessler)	Z-score index of Negative Kessler (K6) score (in Uganda and the Philippines) or the Center of Epidemiologic Studies Depression Scale (CES-D 10) (in Bangladesh) and Cantril's Ladder (in all three countries), standardized to the mean of control villages. Negative Kessler Score: Kessler Psychological Distress Scale (K6) sums responses to six mental health questions and ranges from 0 to 24. Reverse-coded so that higher scores indicate less psychological distress and normalized to the control group. Center of Epidemiologic Studies Depression Scale: The CES-D 10 scale is a 10-item scale designed to measure depressive symptomatology in the general population. Scores range from 0-30 and are reverse-coded where appropriate. Cantril's Ladder: Average of four (three in the case of Bangladesh) questions, each ranging from 1 (low) to 10 (high), normalized to the control group.
Assets	Value of productive assets (PPP 2024)	Sum of (1) value of non-fixed durable productive assets, (2) value of livestock fixed assets, (3) total value of livestock owned, (4) total value of business inventory. Excludes imputed values in Uganda. Data on business inventories were not collected in Bangladesh. Adjusted to 2024 PPP dollars.
Income	Total monthly household income (PPP 2024)	Sum of paid labor income, business profit, net income from livestock, and value of crop production (net of costs) across all household members. Data on crop production was not collected in Bangladesh. Adjusted to 2024 PPP dollars

B.2 Secondary outcomes

B.2.1 Social outcomes

- Social capital index

– Uganda:

Table B.15: SECONDARY ECONOMIC OUTCOMES

Outcome	Description
Own/use land	Whether own land
Use land	Whether use land
Plot value	Value of land owned or rented, in PPP 2024 dollars
Plot size	Size of land cultivated, in square meters
Hours HH worked, past 7 days	Total hours in past 7 days dedicated by the respondent or any other household member to all remunerative work, including casual hours, livestock hours, and other hours
Hours HH worked in livestock, past 7 days	Total hours in past 7 days dedicated by the respondent or any other household member to livestock rearing and to other forms of agricultural self-employment.

- * The indicator is calculated as the mean of bonding and bridging social capital, each ranging from 0–6. Respondents choose whether they could turn to or help each of the following: relatives, non-relatives in my ethnic group/clan, and non-relatives in other ethnic group/clan.

* **Question wording:**

- Bridging Social Capital (0–3): “If your household had a problem and needed help urgently (e.g., food, money, labor, transport, etc.), who OUTSIDE THIS VILLAGE could you turn to for help?”
- Bridging Social Capital (0–3): “Who OUTSIDE THIS VILLAGE would you help if they needed help urgently (e.g., food, money, labor, transport, etc.)?”
- Bonding Social Capital Index (0-3): “If your household had a problem and needed help urgently (e.g., food, money, labor, transport, etc.), who IN THIS VILLAGE could you turn to for help?”
- Bonding Social Capital Index (0-3): “Who INSIDE THIS VILLAGE would you help if they needed help urgently (e.g., food, money, labor, transport, etc.)?”

– **Philippines:**

- * Sum of binary responses to four questions.

* **Question wording:**

- “Has your household provided social assistance in the past 12 months?”
- “Did you receive assistance from community groups in the past 12 months?”
- “Did you receive assistance from family outside your household in the past 12 months?”

- “Did you receive assistance from neighbors and friends in the past 12 months?”
- **Bangladesh:**
 - * Sum of binary responses to four questions about involvement in the shalish, or village courts.
 - * **Question Wording:**
 - “Do you think you will get help from shalish?”
 - “Can you participate in shalish willingly?”
 - “Can you give an opinion in shalish?”
 - “Did you vote in the last national election?”
 - “Are you treated as equal by villagers richer than you?”
- **Trust index**
 - **Uganda:**
 - * Sum of 5 variables (5-item Likert scale).
 - * **Question Wording:**
 - “People in this community trust one another to borrow and loan items.”
 - REVERSE CODE: “I am jealous of others in my community.”
 - REVERSE CODE: “Others in the community are jealous of me.”
 - “If I have a problem, there is always someone to help me.”
 - “Most people in this community are willing to help if you need it.”
 - **Bangladesh:**
 - * Sum of 3 binary variables.
 - * **Question Wording:**
 - “Do you think other villagers will share/lease/mortgage land with you?”
 - “Do you think most people can be trusted?”
 - “If you lost Tk 15,000 and a person around you found it, how likely is it to be returned?”

C Program details

C.1 Uganda

The graduation model in Uganda comprised six components: consumption support, an asset transfer, coaching, core training and skills, savings, and business linkages. In addition to the two program versions studied here—group coaching and individual coaching arms—another set of randomly selected households in treatment villages received a package identical to the individual coaching arm but *without* the asset transfer.

1. **Asset transfer** Participant households received an asset transfer beginning in July 2019, which was delivered on a staggered basis. Participant households received 1,100,000 UGX to use as start-up capital to invest in their recently chosen livelihood.
2. **Coaching**
 - (a) **Individual Coaching:** Individual coaches, who most closely resemble social workers in their job responsibilities, met one-on-one with individual participant households. These coaches guided participant households on AVSI's graduation pathway, helping them to identify, manage, and reach their personal goals. Each individual coach was assigned 23–27 households, meeting biweekly for nearly 30 months. Each meeting lasted, on average, 60 minutes. Topics of discussion include nutrition, water, health and sanitation, savings, business, and preventive health strategies.
 - (b) **Group Coaching:** Group coaches, drawn from the same pool of coaches as those providing individual coaching, met with participant households in groups ranging between 23–27 households. Rather than one-on-one meetings, group coaches delivered their program in group settings. Each group coach managed three groups in their portfolio. Group coaches did not conduct home visits except in exceptional cases. Group coaches meet with their groups weekly for 60 minutes.
3. **Business Linkages:** In May 2019, AVSI, in partnership with NGOs and the local government, held a three-day workshop to connect beneficiary households with private sector vendors and market-linked NGOs. During the first two days, the organizers allowed only Graduating to Resilience beneficiary households to attend the event. During the third and final day, organizers opened the event to all surrounding households. The organizers invited seven main categories of exhibitors to attend the event: (1) big agribusiness firms, (2) agricultural tool and equipment suppliers, (3) agro-processing,

post-harvest handling and value addition technology suppliers, (4) green energy solution service providers, (5) produce buyers, (6) providers of market information and extension, and (7) financial service providers. Overall, this market event served as the beginning of AVSI's work to connect private-sector agricultural input suppliers with participant households. The objective of this component was to help participant households access quality seeds and agricultural inputs that AVSI determined were not readily available in the local market context. Over the course of the program, AVSI worked to engage with these external service providers.

4. **Consumption support:** Households received a monthly cash stipend for the first 12 months of the program. The stipend was valued at 18,000 UGX per household member per month for refugee households, and 15,000 UGX per household member per month for host households.
5. **Core Training and Skills:** Throughout the program lifecycle, AVSI delivered a series of trainings improve household livelihood management, separately from coaching. These include selection, planning and management training designed to help households form sustainable businesses, agronomic training to improve household farming practices by introducing new farming techniques (e.g. use of organic manure, fertilizer), and business coaching to help participant households manage their businesses.
6. **Savings:** AVSI randomly allocated participant households to Village Savings and Loan Associations (VSLAs). A VSLA is a member-run financing mechanism that helps participant households save weekly.

C.2 The Philippines

BRAC USA implemented the graduation pilot in partnership with the Philippine government's Department of Labor and Employment (DOLE), with financial support from the Asian Development Bank. Prior to program launch, BRAC's conducted a market analysis to identify potential livelihoods that were likely to be feasible and profitable in this setting. They used the results of this assessment to construct livelihood packages from which participants could choose.

At the beginning of the program, BRAC Graduation Coaching Facilitators (GCFs) used a market assessment tool to determine which livelihoods would be appropriate for each community. This tool had a list of requirements for each livelihood to help GCFs cross-check. For example, livelihoods like swine, working carabao, and chicken egg-laying require abundant space, which beneficiaries in populated lowland areas did not have. As Appendix

Table C.1 shows, most beneficiaries (54%) chose swine fattening across all treatment groups. This was followed by business cart or NegoKart (15%), meat processing (13%), and free-range chicken (8%).

The graduation program was built upon the existing DOLE Kabuhayan (Livelihood) program, which targets poor and vulnerable households and had been previously implemented with “pre-formed” applicants, who applied as a group for seed capital. In this study, individuals formed groups within their barangay quadrant and applied for group or individual livelihoods, depending on their quadrant assignment.

Besides the working capital for selected projects, beneficiaries also received training on setting up and running the livelihood project as well as technical and business advisory services to support the sustainability of the business.

In addition to the livelihood training and asset transfer, participants received regular coaching from BRAC GCFs delivered fortnightly in individuals or groups. After facilitating the asset transfer and technical and business management training for households, GCFs provided life-skills training on topics ranging from household financial management to water, sanitation, and hygiene, health, domestic violence, and child nutrition, tailored to the most prominent issues among the poor in the community. Additionally, coaches provided access to savings and financial inclusion mechanisms as well as troubleshooting business development concerns for households to grow and improve upon their enterprises. Finally, coaches promoted social integration through supporting community-led activities and linking households to community associations or committees, including local government officials and initiatives.

C.3 Bangladesh

C.3.1 Program eligibility

The program has five inclusion and two exclusion criteria. To be eligible, a household must fulfill at least three of the inclusion criteria and none of the exclusion criteria. *The inclusion criteria:* 1. The household depends on female day labor for income; 2. It has less than or equal to ten decimal land¹⁵; 3. It has no productive assets; 4. The household has children of school-going age (6-14 years) who have to work; and 5. The household has no adult male members. *The exclusion criteria:* 1. The household has no female members capable of working; and 2. The household has one or multiple members that are participants in other government/non-government development projects.

¹⁵A unit of measurement used in rural Bangladesh, a decimal is one hundredth of an acre of land, or 40.47 square meters.

Table C.1: PHILIPPINES: PROGRAM LIVELIHOOD ASSET CHOSEN BY ARM

Livelihood	Total house- holds	Group	Individ- ual
Swine Fattening	545	77%	37%
NegoKart Business	170	9%	21%
Meat Processing	133	9%	15%
Free Range Chicken	99	0%	22%
Working Carabao	46	0%	0%
Chicken Egg Production	32	0%	0%
Backyard Vegetable Farming	24	2%	3%
Salted Eggs Processing	16	2%	2%
Cosmetology/ Massage	4	0%	1%
Fish Vending	1	0%	0%
Total	1134	100%	100%

C.3.2 Program components

We describe below the key components of the program. Where treatment is at the individual level, the recipient is always the main woman in the household. Unless explicitly stated otherwise, the treatment elements described below apply to the interventions received by all treatment groups in the study. On average, approximately five percent of households in treated villages received the intervention, though this proportion was somewhat variable.

1. *Asset transfers:*

At the start of the program, the main woman in the household is presented with a menu of assets, each of which can be used productively in an income-generating activity. These include, for instance, livestock, land, and capital equipment. Nearly all recipients (97%) chose a package with either cows and/or goats, and 76% chose a package containing cows. Fewer than 3% received a package that included land or non-agricultural capital goods.

Prior to the asset transfers, BRAC staff provide recipients with classroom training on how to manage the asset for income-generating activities. Assets were transferred within a month of this training session, and BRAC encouraged beneficiaries not to liquidate the transferred assets for at least two years after receipt, although there were no hard sanctions if the asset was liquidated prematurely.

Table C.2: BANGLADESH: PROGRAM LIVELIHOOD ASSET CHOSEN BY ARM

Asset package	Total households	Group	Individual	Individual Light
Cow	4255	79%	77%	77%
Goats	1030	18%	19%	20%
Other	178	3%	4%	3%
Total	5463	100%	100%	100%

Notes: The ‘Cow’ asset package is defined as an asset package that consists of one cow and either one or no goats. The ‘Goats’ asset package is defined as an asset package that consists of four or five goats and no cows. The ‘Other’ package is any other combination of assets. The average value of the ‘Cow’ asset package is \$500, the ‘Goats’ asset package is \$370 and the ‘Other’ package is \$240 (prices in 2020 \$PPP).

2. Coaching

All beneficiaries receive regular one-on-one coaching through home visits by BRAC program staff. As described above, BRAC introduced experimental variation in whether this coaching was received on a weekly or a fortnightly basis, as well as whether a group coaching element was included. For all treatment groups, these coaching sessions included life skills training, such as guidance on health and education-related issues, enterprise management skills, and guidance on how to care for the asset transferred through the program, among other things. Health topics included family planning, prevention and treatment of waterborne diseases, non-transmittable diseases and hookworm, vaccinations, and nutrition. Coverage of social issues included, for instance, guidance on how to further the education of children. The participants also received coaching meant to boost confidence, were prompted to set goals, were taught about their rights and responsibilities, and were encouraged to save and invest in simple businesses such as retail trading. They were also informed about different government and non-government services, such as access to credit and eligibility for support through various government rural development schemes.

An important part of the home visits was to follow up on the status of the asset. For instance, coaches observed whether the asset was being used for income generation and discussed challenges that the beneficiary may have experienced in putting the asset to productive use. If the asset was livestock, coaches also helped ensure that it was healthy

and receiving appropriate care. For instance, coaches typically provided training on animal husbandry best practices, such as insemination processes, vaccination, feeding, milking, and so on.

In summary, coaching comprised several elements serving different purposes and potentially operating through different mechanisms. Training was provided to transmit hard skills along with soft skill development, beneficiaries were exposed to mentors who might serve as role models, beneficiaries could access other social services through the intermediation of the coach, and the regular visits served as a compliance check, ensuring that beneficiaries were using the transferred assets as intended.

Coaching sessions lasted on average approximately 30 minutes and continued for 18 months after the initial asset transfer. Coaches, also called BRAC program officers, were young and well-educated: they were required to have a masters' degree and be under the age of 35. Each branch had on average three program officers serving between 80–120 households.

3. *Group meetings (Group coaching only):*

For the group meetings, a BRAC coach convened a gathering of between 3–8 beneficiaries in a village. These group meetings were typically held on the same week as the fortnightly home visit in one of the beneficiary's houses. Beneficiaries of the T1 group variant received the subsistence allowance stipend once every two weeks during these group meetings instead of during home visits. Group meetings were not set up as rotating savings and credit associations.

The topics covered during the group visits were the same in substance as those covered during the one-on-one coaching home visits. Group-based instruction and the moderation of these sessions was also undertaken by the same BRAC coaches who were responsible for the one-on-one sessions. However, while BRAC program officers also used the group meetings to provide instruction on various social and health-related matters, the group setting provided beneficiaries with a less structured forum in which to discuss concerns and raise questions. The group meeting was also a potential opportunity to create social networks between the beneficiaries. In addition, group-based coaching may be complementary to one-on-one house visits in other ways, for instance, representing a potentially more interactive and engaging approach to hard- and soft-skill acquisition, a more binding form of commitment device, and allowing recipients to learn from the experiences of other recipients.

4. *Subsistence Allowance:*

All beneficiaries receive regular consumption support in the form of small, unconditional cash grants. These cash grants were delivered in person by a BRAC officer during either weekly or fortnightly coaching sessions.¹⁶ Consumption support in the form of this subsistence allowance lasted for the first 40 weeks of the program. Recipients were encouraged to save a small amount of the transfer, with a target set at approximately 10 percent.

5. *Health and Miscellaneous Social Services:*

The beneficiaries were provided with access to medical support and encouraged to seek the assistance of a BRAC health worker if a family member was sick. BRAC health workers could provide basic medicines, give medical advice, and refer household members to local hospitals and health centers.

C.3.3 Nested experiment on community mobilization

In the 2016 cohort, BRAC also introduced experimental variation in community mobilization within the group coaching treatment arm (T1). Specifically, in all treatment villages apart from a subset within T1, informal institutions, called VACs or “Village Poverty Reduction Committees” were set up. VACs were *not* established in approximately half of those villages in T1, with randomization at the village level. Our records indicate that of 2328 households in T1, 1029 were in villages assigned to the establishment of VACs and 1299 were in villages in which VACs were not set up. However, errors in record-keeping prevent us from being confident of the treatment status for all villages, and so we do not include it as a covariate. Appendix A.9 reports results after controlling for this compromised treatment assignment variable.

These committees were created with the intention of inducing community mobilization, linking villagers across social strata, and creating mechanisms for better-off villagers to support the ultra-poor. The stated goal of these committee structures was to support the economic security of UPGP beneficiaries, create an enabling environment for the beneficiaries in the village, help recipients protect their assets, increase access to government services, and offer support in times of need by coordinating local community support. Committees consisted of 9–11 villagers and were comprised of UPGP recipients and community leaders—for instance wealthier land owners, teachers, imams,

¹⁶The total value received did not differ between beneficiaries who received weekly versus those who received fortnightly training—the latter simply received double the amount half as frequently. Since the amounts are small in absolute terms, we do not expect that the differences in the frequency of transfers would affect investment or have enduring impacts on consumption.

and local authorities. The VAC would meet monthly and discuss issues raised by the UPGP recipients. Committees were typically established within three months of the asset transfer.

D Experimental details

D.1 Uganda

D.1.1 Study area and listing

The study area comprises the entire refugee community of the Rwamwanja settlement and four subcounties contiguous to the settlement (Nkoma, Biguli, Bihanga, and Bwizi). Between July–September 2018, AVSI conducted an individual household poverty scorecard exercise and a community-based poverty ranking covering all households in the study area. AVSI reached 35,204 households and identified 25,104 households as eligible for program participation—households that were assessed as poor or extremely poor and included a woman or youth who could be economically active as part of the program. The program had an implementation target of 6,600 households evenly split by host and refugee communities. To accommodate implementation targets and the study design with four experimental groups in treatment clusters (three treatment arms and a spillover control) plus a “holdout” control for future implementation, each with a minimum target implementation group size of 25 members, four large villages were split and 14 small villages merged to create a total of 42 village clusters for the study in the refugee settlement and 72 village clusters in the host community. The study target sample consisted of a total of 11,000 households randomly sampled from the set of eligible households, split evenly across five experimental groups (three treatment groups plus a spillover control in treatment clusters and a pure control in control clusters), with a target of 2,200 households in each group evenly split between host and refugee communities. The final study sample consists of 11,145 program-eligible households interviewed for the baseline data collection, 10,514 of which were interviewed during the endline survey.

D.1.2 Randomization

We randomized village clusters into treatment and control, separately by host and refugee communities. In host clusters, random assignment was stratified by parish. In refugee clusters, strata were based on distance to the economic center of the settlement and on whether a cluster was above or below the median poverty score in its respective geographic grouping. In addition to stratification, we used a re-randomization procedure based on listing data collected that included household size, poverty scores, and female headship.

Within treatment clusters, households were then individually randomized with equal probabilities via public lotteries into one of four experimental conditions, which include a spillover control and a no-asset arm in addition to the two arms that vary only by coaching

modality that are the focus of this paper.

D.1.3 Imputation

In Uganda, to limit the survey length, households were randomly assigned to complete a full ("long") or shorter versions of several survey modules. Randomization to short versus long survey modules was done separately by groups of modules, resulting in only a limited subset of households receiving all component modules needed to compute the productive assets and income aggregates. The depth of the information available in short modules varied across questions but typically included at least gateway questions for whether the household had any of a given activity/purchase/item category (e.g. any non-farm business, any purchase in each agricultural input category, any livestock structure for each livestock structure type) and sometimes included additional detail (e.g. business profits for the most important business; crop income from the most important crop, number of animals for all livestock types).

Appendix Table D.2 and D.1 shows the allocation to component modules for the asset and income aggregates.

Table D.1: SHORT SURVEY PATTERNS FOR INCOME MODULES

Pattern	Employment (E)	Business (B)	Livestock (L)	Crops (C)	Frequency
----	0	0	0	0	683
---C	0	0	0	1	1,375
-B--	0	1	0	0	1,609
-B-C	0	1	0	1	1,589
E-L-	1	0	1	0	1,616
E-LC	1	0	1	1	1,586
EBL-	1	1	1	0	1,364
EBLC	1	1	1	1	692

Table D.2: SHORT SURVEY PATTERNS FOR ASSET MODULES

Pattern	Asset (A)	Business (B)	Livestock (L)	Frequency
---	0	0	0	2,058
--L	0	0	1	3,202
AB-	1	1	0	3,198
ABL	1	1	1	2,056

In our analysis, we conduct a multiple imputation procedure for these missing values in order construct aggregates for the full sample. Importantly, the imputations are done separately by treatment arms (and refugee status). We use Multiple Imputations by Chained

Equations (MICE) to fill in missing values for each sub-component of the aggregates based on information collected in the short survey and the (imputed) value of the other sub-components. We specify a predictive mean matching procedure, which matches predictions from a linear model with the nearest neighbors in the distribution of the non-imputed values, to create 80 imputed data sets. The asset and income aggregates are recombined as the sum of the imputed components (where missing).

We use Stata’s `mi estimate` to fit the treatment-effect regressions on the imputed aggregates. `mi estimate` combines point estimates and variances across imputations using Rubin’s rules, so reported standard errors reflect between-imputation variability (i.e., imputation uncertainty). Since short survey modules were randomly assigned and we impute separately by experimental conditions, the imputed outcomes—and resulting treatment effect point estimates for the aggregates—match closely with the outcomes for observations where the full information is available.

For lasso selection of covariates, we use a randomly selected single-imputed dataset for our income and asset outcomes. We similarly use a single randomly selected imputed dataset when testing for treatment heterogeneity.

D.2 Philippines

D.2.1 Study area and listing

The original sample included poor households in 29 barangays across five municipalities in northern Negros Occidental who receive the government’s conditional cash transfer program and were added to the program during the same two-year period between 2015 and 2017. The following describes how the final list of households eligible for inclusion in the pilot program was determined.

During the inception workshop that included representatives of ADB, BRAC, DOLE, and IPA, it was agreed that the study sample would consist of 2,400 households across 30 barangays (80 households per barangay). In May 2018, IPA received a list of households that were all added to the government’s conditional cash transfer program (4Ps) during the same two-year period. The original list consisted of approximately 3,200 households from 32 barangays that had 80 or more eligible 4Ps recipients. The number of households per barangay eligible for inclusion in the study varied: some barangays had more than 200 names, others just over 80. IPA defined the final master list of households as follows:

1. The number of sample barangays was reduced to 30 by dropping the two barangays with the fewest number of eligible households: barangay Alegria in the municipality of

Murcia (81 eligible households) and barangay VI-A in the City of Victorias (81 eligible households).

2. IPA then randomly sampled 110 households per barangay from barangays with more than 110 eligible households. This reduced the likelihood that surveyors would skip respondents unavailable upon their first interview attempt, which has the potential of introducing selection bias, and left a sufficient number of “reserve” households in case the list proved outdated, or a significant fraction of respondents could not be reached during the 3-4 days IPA field staff spent in each barangay.

After random sampling, the master list numbered 3,098 households eligible for inclusion in the study.

3. The list of up to 110 names per barangay was then randomly sorted. For each barangay, surveyor teams were provided with a hard copy list of up to 110 respondents and instructed to begin with the names at the top and work their way down until the team reached 80 interviews. They were granted permission to skip a name only if a respondent could not be reached in the 3–4 days of fieldwork allotted per barangay.

D.2.2 Randomization

Households were randomized in two stages:

1. Each barangay was divided into four clusters based on GPS coordinates collected during surveying. Households were assigned points based on their distance to the nearest cluster’s center minus the distance to the farthest and ranked accordingly. Then each of the four clusters were filled with the highest-ranked households until 20 households were placed, at which point that cluster was taken out of the equation and the next cluster filled with the highest-ranking households.

This process was repeated until all 120 clusters of 20 households across the 30 sample barangays were filled. Once all households had been assigned to one of four quadrants in their barangay, we randomized assignment of these quadrants into one of three treatment arms or a control group. We re-randomized as described above in the Uganda experiment, aiming to ensure balance on household size, number of adult household members, total value of durable assets, and total value of livestock.

2. Following randomization, BRAC field staff conducted a validation survey of households in treatment areas. The team found several households that were recipients of DSWD’s Sustainable Livelihood Program (SLP). BRAC identified 44 households that had or

were participating in SLP. In one barangay (Barangay VI, Poblacion), more than half of study households were part of SLP, so BRAC and IPA agreed to exclude it, leaving 29 participating barangays.

D.3 Bangladesh

D.3.1 Study area and listing

This study draws a sample from the 2016 cohort of the BRAC UPGP intervention, a scaled version of BRAC’s graduation program delivered to 80,000 households. A subsample of those identified as eligible for the 2016 intervention were selected for the experimental evaluation.

Initial selection into the program proceeded in the following steps: First, the poorest rural *upazilas* (subdistricts) in which BRAC operates were identified using poverty mapping tools developed by the World Food Programme (Ahmad et al., 2010). Within these *upazilas*, communities with the highest concentration of poverty were identified based on the knowledge of existing BRAC field staff already operating in these areas (for instance, in BRAC’s microfinance, health, and education programs). Next, BRAC officers produced a list of the poorest households in these communities using participatory wealth ranking (PWR) exercises in which all households within a community are ranked in terms of material well-being by villagers in an open meeting¹⁷ The PWR serves to classify all households into categories such as very poor, poor, middle-class, and non-poor. Based on this classification, BRAC program staff then administer a questionnaire to those households in the village who were classified as “very poor” or “poor” to generate a preliminary list of ultra-poor households. This survey serves to verify whether particular households met the inclusion and exclusion criteria set by BRAC.¹⁸ The primary female adult in those households who are verified as

¹⁷The process starts with rapport building where a program staff walks around a village with a few villagers to decide whether there is a substantial concentration of poor households in the village. Once a village is chosen, the program staff engages with more villagers to identify a venue for the PWR. Villagers were invited for the PWR, which takes the form of a group discussion, usually on the following day. While the PWR may have differed from village to village, it usually included a group of about 40-50 villagers. A map of the village was drawn on the ground and villagers discussed the economic condition of each household as it was pointed out on the map. In the process of ascertaining economic conditions, the participants were asked to keep in mind issues such as the roofing material of the household, if any children were out of school, whether members had a steady income source, had any productive assets, et cetera. This was done for every household in the village.

¹⁸The program has five inclusion and two exclusion criteria. To be eligible, a household must fulfill three of the inclusion criteria and none of the exclusion criteria. *The inclusion criteria:* 1. The household depends on female day labor for income 2. It has less than or equal to ten decimal land 3. It has no productive assets 4. The household has children of school-going age (6-14 years) who have to work 5. The household has no adult male members. *The exclusion criteria:* 1. The household has no female members capable of working. 2. The household has one or multiple members that are participants in other government/non-government

eligible through this survey were then registered as beneficiaries.

The 2016 UPGP was implemented in 126 branches from 26 districts. Of these, 11 districts where the program had never been implemented before were selected for inclusion in the experimental evaluation. The program then selected eight branches from each of these selected districts. These branches were then randomly assigned to each of the four treatment or control groups. In total, 8,468 women and their households completed baseline surveys, and they constitute the study sample.

D.3.2 Randomization

A baseline survey was conducted between April and August 2016, prior to randomization or the implementation of the intervention. Baseline surveys were completed by 8,468 recipients and their households, drawing a sample of between 50 and 200 eligible ultra-poor households from each branch, which is the smallest unit in BRAC's administrative structure.

The study covers 88 branches from 11 districts. These 88 branches were randomly assigned to three treatment groups and a control group of 22 branches each. Randomization was stratified by district.

development projects.

E Heterogeneity

We follow the generic machine learning procedure outlined by [Chernozhukov et al. \(2020\)](#) to test for differential impacts by sub-groups, utilizing the `genericML` package by [Welz et al. \(2022\)](#) with the set of covariates that had above average influence for each outcome variable. Specifically, we do the following for each outcome, considering each program separately:

1. We restrict our sample to those receiving group or individual coaching, classifying group coaching as the “treatment group.”
2. We identify the list of key baseline covariates. Note that all covariates are continuous or binary.
3. For each outcome variable, we train a random pilot forest using the training sample and the potential covariates as a pilot. We restrict the set of potential covariates to those with above-average importance, following [Athey and Wager \(2019\)](#).
4. Using the `genericML` package, we use three potential learners: lasso, random forest (with 1,000 trees), and a support vector machine.
 - To ensure results are robust to how the data are split between a “main” and “auxiliary” sample, for each of 250 splits, we train the machine learning algorithm on the auxiliary sample and use those estimates to predict outcomes in the main sample. We take the median of these results over 250 random sample splits.
 - We account for clustering by quadrant in the Philippines and by village in Bangladesh. Conditional on village-level assignment to treatment, randomization of coaching modality in Uganda was implemented at the household level, so we do not cluster standard errors.
5. We then estimate group average treatment effects (GATES) based on terciles.

F Cost-effectiveness analysis

We calculate the net present value of all benefits based on the estimated year 3 (year 6 in Bangladesh) impact on consumption and assumed social discount and persistence rates. We use that same social discount rate to inflate the incurred costs to year 3 (year 6 in Bangladesh). We follow (Banerjee et al., 2015) and assume a social discount rate of 5%. For this example, we take a persistence rate of 100%, although Table 2 reports the estimated benefit-cost ratio at persistence rates of 90% and 80%, following the same strategy.

We report program costs and benefits in PPP 2024 dollars (Bank, 2025). We assume that all program expenses were incurred during the first year of implementation. Expense data was collected in USD, which we inflation adjust to the endline survey year (2021 in Uganda and the Philippine and 2022 in Bangladesh) before converting to local currency units and then into PPP dollars, ensuring the LCU to PPP conversion occurs in the same year for both costs and benefits. We then use the US CPI to inflation-adjust to 2024 PPP dollars.

F.1 Program costs

In the Philippines, we received cost estimates from BRAC, which we reconciled with funding requests from the Asian Development Bank, which funded coaching support and the asset transfer for the individual-coaching treatment arm. This funding covered the time of both field and supervisory staff within BRAC, as well as staff time at higher management levels that was specifically dedicated to this project. The livelihood project is integrated into the Department of Labor and Employment’s DILEEP (or Kabuhayan) program, and we account for the value of these assets and note that training was paid for by BRAC. However, we are not able to account for the value of staff time within DOLE to support the livelihood program, which includes reviewing proposals, approving and monitoring procurement, and liaising with BRAC. Consequently, we may underestimate the overall administrative program costs.

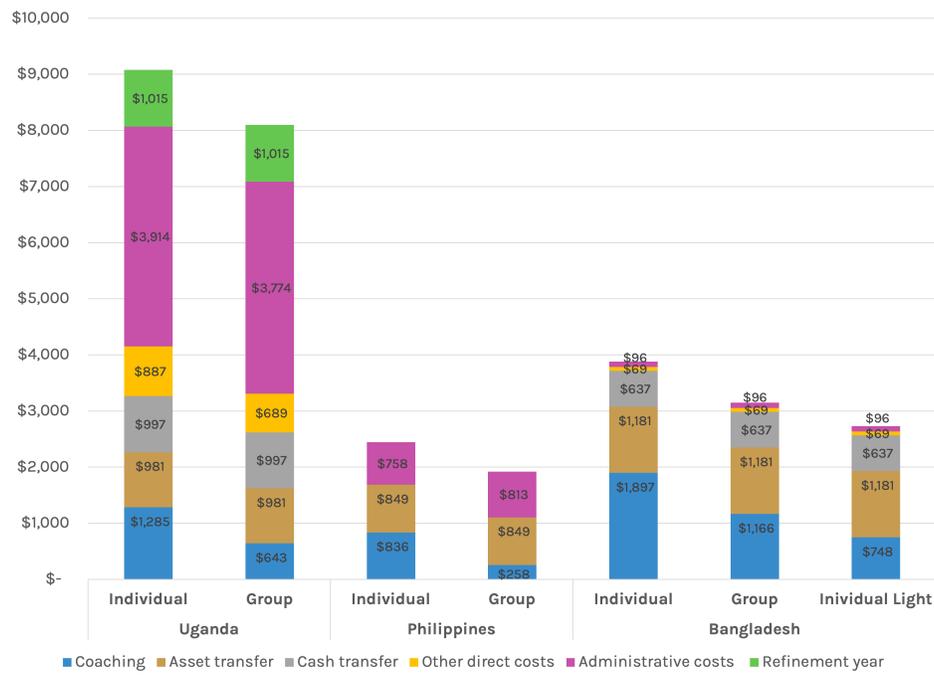
Program implementation costs in Uganda comprise the value of consumption support, asset transfers, coach and trainer salaries, other direct costs, and project management and monitoring and evaluation expenses. AVSI also spent one year refining the activity prior to launch. Because we anticipated that this refinement would benefit both this cohort and a second cohort, we ascribe half of the refinement cost as part of the overall program cost.

Cost estimates were harder to obtain in Bangladesh. BRAC either did not record or has been unable to provide detailed cost data disaggregated by treatment arm. They have only been able to supply total cost data for the entire 2016 cohort of 80,000 recipients—of which approximately 6,000 of which were included in the experiment. We therefore had to make

assumptions about the per-recipient cost and about the cost differential across treatment arms. We can break down the cost estimates provided between costs that are likely to remain fixed across treatment groups (i.e. the price of the asset, etc) and coaching-related costs that are likely to vary (coaching staff salaries, staff transport, etc). We estimate that 37% (\$262) of per-participant program costs would be sensitive to the amount of coaching delivered, while the remainder would be fixed across participants (\$496). We estimate that, compared to the fortnightly-plus-group coaching arm, BRAC saves 40% of the coaching costs in the fortnightly-only arm, bringing coaching-related costs down to \$187 for that arm. We then assume that doubling the amount of coaching delivered in the weekly treatment arm relative to the individual and group coaching arm increases coaching costs to BRAC by a factor of 1.7, to \$474. This yields the cost figures in Table F.1.

Across all three sites, have not included the cost of beneficiary time or any resources that were expended to participate in the program, such as transportation to program activities. However, reasonable estimates of the value would not substantially change impacts.

Figure F.1: COST PER PARTICIPATING HOUSEHOLD



F.2 Program benefits

We calculate the estimated increase in monthly consumption per adult equivalent in Table 1 for both group and individual coaching arms. We adjust consumption to an annual household total by multiplying by the number of adult equivalents and by 12.

Table F.1: COSTS PER PARTICIPATING HOUSEHOLD

	Uganda		Philippines		Bangladesh		
	Group	Individual	Group	Individual	Group	Individual	Individual light
Coaching	643	1,285	258	836	1,166	1,897	748
Asset transfer	981	981	849	849	1,181	1,181	1,181
Cash transfer	997	997	–	–	637	637	637
Other direct costs	689	887	–	–	69	69	69
Administrative costs	3,774	3,914	813	758	96	96	96
Refinement year	1,015	1,015	–	–	–	–	–
Total	8,099	9,079	1,920	2,443	3,149	3,880	2,731

Notes: The Bangladesh program was nested within a scaled intervention of 80,000 recipients: economies of scale in delivery help explain why administrative and other fixed costs are lower in this program than in the Philippines and Uganda programs.

Consumption support and asset transfers were delivered in Year 1 in Uganda and Bangladesh, while in the Philippines the asset transfers were not completed until Year 2. For this reason, we inflate the Year 3 (Year 6) benefit by the discount rate to estimate a benefit in Year 1 and Year 2 (Years 1 to 5) in Uganda (Bangladesh). In the Philippines, we inflate the Year 3 benefit only to Year 2, and we assume there is no impact on consumption in Year 1.

We calculate the net present value of benefits from year 4 (year 7 in Bangladesh) onward assuming a persistence rate of 100% and discount rate of 5%.

We then divide the total benefit by the total inflated costs to estimate a benefit-cost ratio for each persistence rate. A ratio greater than 1 indicates that the total realized benefits exceed the costs, although we note that these only account for the value of consumption and ignore any non-pecuniary benefits of the program.