



# **Impact Evaluation of the Graduation of the Ultra-Poor Pilot in the Philippines**

**Report - Endline Survey**

Prepared by Innovations for Poverty Action Philippines

Manila, April 1, 2022

## Contents

Executive summary .....	5
1. Background .....	6
The graduation approach.....	6
Description of the Philippines graduation pilot .....	7
Research design .....	10
Research questions and outcome measures .....	12
Survey instrument design and programming .....	13
Sample selection and randomization .....	14
2. Data Collection .....	18
Endline data collection .....	18
Survey attrition .....	19
Differential attrition .....	20
Balance on baseline characteristics.....	21
3. Program implementation and participation.....	23
Implementation challenges .....	24
Participation rates.....	25
4. Empirical specification.....	27
5. Endline results .....	28
Key welfare outcomes .....	29
Household consumption.....	30
Food security .....	31
Productive asset ownership.....	32
Household income.....	33
Subjective well-being .....	37
Secondary outcomes .....	39
Sub-treatment analysis.....	45
6. Cost-effectiveness .....	46
7. Conclusions and policy implications.....	50
8. References .....	53
Appendix .....	54
1. Summary statistics on key outcome components .....	54
2. Endline survey preparatory activities .....	60
3. Field data collection .....	62

4. Data Collection Relaunch Plan (December 2021).....	63
5. Endline survey preparatory activities .....	65

## List of tables

Table 1: Treatment and control group assignment, by barangay .....	16
Table 2: Number of households from baseline survey in endline sample .....	19
Table 3: Status of unreached households.....	19
Table 4: Endline sample completion rates disaggregated by barangay by treatment arm .....	20
Table 5: Means of baseline characteristics for endline sample (standard deviations in parentheses).....	21
Table 6: Program livelihood asset chosen by arm .....	23
Table 7: Reported reasons for opting out of program .....	26
Table 8: Overall coaching session attendance rates by treatment arm .....	27
Table 9: Program livelihood by treatment group .....	28
Table 10: Treatment effects on key welfare outcomes .....	29
Table 11: Consumption per capita per month (in pesos).....	30
Table 12: Food security.....	31
Table 13: Value of productive assets (in pesos).....	32
Table 14: Household income per month (in pesos) .....	34
Table 15: Household income components per month (in pesos) .....	34
Table 16: Income sources.....	35
Table 17: Business operations.....	36
Table 18: Labor supply in past 30 days.....	36
Table 19: Subjective well-being.....	37
Table 20: Kessler measure components 30 days (higher score = less distress).....	38
Table 21: Cantril's ladder components (higher score = better outcome).....	39
Table 22: Savings .....	40
Table 23: Loans .....	40
Table 24: Financial health.....	41
Table 25: Child welfare .....	42
Table 26: Social networks.....	44
Table 27: Sub-treatment analysis by livelihood choice.....	45
Table 28: Total program costs (in USD) .....	47
Table 29: Estimated costs by treatment arm (in USD) .....	48
Table 30: Return on investment under different cost and benefit assumptions .....	49
Table 31: 7-day consumption items in pesos .....	54
Table 32: 30-day consumption items in pesos .....	55
Table 33: 12-month consumption items in pesos.....	56
Table 34: Value of household assets.....	57
Table 35: 7-day food consumption score .....	57
Table 36: 4-week household food insecurity access scale .....	58
Table 37: 6-item Kessler psychological distress scale.....	58
Table 38: Cantril's ladder relative and expected life satisfaction .....	59
Table 39: Log of monthly consumption and income and productive asset values.....	59

## List of figures

Figure 1: Project timeline.....	10
Figure 2: Treatment and control group assignment.....	11
Figure 3: Survey rollout.....	18
Figure 4: Timeline of endline survey activities .....	18
Figure 5: Participation share and asset delivery over time, differential across arms.....	25
Figure 6: Monthly conditional coaching session attendance rates .....	27
Figure 7: Program activity disruptions timeline.....	65

## Terms and abbreviations

4Ps	Pantawid Pamilyang Pilipino Program - DSWD's conditional cash transfer program
ADB	Asian Development Bank
Barangay	A local government administrative unit, akin to a micro-district
BRAC	Building Resources Across Communities
COVID-19	Coronavirus disease
DOLE	Department of Labor and Employment of the Philippines
DSWD	Department of Social Welfare and Development of the Philippines
GCF	Graduation Community Facilitator
IPA	Innovations for Poverty Action
<i>Kabuhayan</i>	DOLE's asset transfer and training program, implemented for individuals and groups
PII	Personally Identifiable Information
RCT	Randomized Controlled Trial
SAP	DSWD's COVID-19 Social Amelioration Program
T1	Treatment group 1, which received group coaching and grouped livelihood
T2	Treatment group 2, which received group coaching and individual livelihood
T3	Treatment group 3, which received individual coaching and individual livelihood

## Evaluation partners

### ADB

The Asian Development Bank (ADB) is a multilateral development finance institution, providing loans, technical assistance, and grants to governments in the region. It also provides direct assistance to private enterprises of developing member countries through equity investments and loans, facilitates policy dialogues, provides advisory services, and mobilizes financial resources through co-financing operations that tap official, commercial, and export credit sources.<sup>1</sup> The ADB piloted the graduation approach to sustainable livelihoods in the Philippines with technical assistance grants from several

---

<sup>1</sup> Source: <https://www.adb.org/about/our-work>

sources, including the regional technical assistance *Unlocking Innovation for Development* (TA 9017-REG).

### **BRAC**

BRAC is a development organization that for nearly fifty years has concentrated on developing, testing, iterating, and refining entrepreneurial solutions to reducing global poverty. BRAC developed the Graduation Approach in 2002 and has since scaled it through direct implementation to reach more than 1.8 million households in Bangladesh, where the organization is founded, and across the 11 countries in Asia and Africa where BRAC operates. Currently, BRAC's Ultra Poor Graduation Initiative is partnering to provide technical support to governments in Kenya, Lesotho, Tanzania, Rwanda, and the Philippines as well as other partners in Bangladesh, India, and Malawi, among others.

### **DOLE**

The Philippine government's Department of Labor and Employment (DOLE) plays a key role in the Graduation Pilot. DOLE's Integrated Livelihood and Emergency Employment Program (DILEEP) targets poor and vulnerable workers and forms the basis of the asset transfer component of the Graduation Pilot in the Philippines. Under the *Kabuhayan* program, one component of DILEEP, beneficiaries receive a one-time asset transfer worth an average of about 15,000 pesos (\$300) per recipient plus a few days of training in business planning, organization, and accounting practices. *Kabuhayan* is implemented for individuals and groups. Beneficiaries of the individual asset transfer can receive a productive asset of their choosing worth up to P20,000 (\$400) and training. Beneficiaries of the group-based asset transfer collectively decide on an asset worth up to P1 million (\$50,000) (for groups with 100 members or more) in addition to training.

### **IPA**

Innovations for Poverty Action (IPA) is a specialized organization whose core competence is the successful design and implementation of rigorous impact evaluations of development interventions using randomized controlled trials. IPA's global network of country programs offers supportive infrastructure and established relationships with key stakeholders including expert researchers. IPA has previously worked with DOLE on the evaluation of KASAMA, a livelihood program that focuses on households with a high incidence of child labor.

## Executive summary

Between September and December 2021, IPA conducted an endline survey of households participating in the Graduation of the Ultra Poor pilot study in Negros Occidental, Philippines. The goal of this study was to measure the pilot's impact one year after the end of all program activities, or approximately two years after the bulk of livelihood asset transfers. The survey measured household consumption, food security, asset ownership, income, subjective well-being, and other outcomes.

The target sample included 2,339 households that were interviewed during the baseline survey and randomly assigned to one of three treatment arms or a control group. By collecting updated household contact information prior to data collection and using GPS data collected at baseline, IPA's enumerators surveyed 98% of the baseline sample (2,288 households), with roughly equal completion rates across experimental groups.

The study was designed to test the relative performance and cost-effectiveness of three program variations. Treatment arm "T1" households were coached in groups and received a livelihoods asset transfer as a group; treatment arm "T2" households were coached in groups and received individual asset transfers; treatment arm "T3" households were coached individually and received individual asset transfers.

While the study was designed to isolate the relative importance of the specific program elements that were explicitly varied, differences in key implementation details across treatment arms affect the interpretation of outcomes across arms. Specifically, compared to T2 and T3, T1 had a lower asset delivery rate (62% vs 73% in T2 and 78% in T3), a higher program attrition rate and members received their assets later. In addition, while T2 and T3 were designed to differ only in mode of coaching delivery, T2 and T3 also received markedly different types of assets, driven in large part by differences in BRAC's livelihood choice process for the two groups. As a result of these differences in key implementation details, interpreting the differences in impact estimates requires accounting for these implementation differences.

The main findings on key outcomes of interest are as follows:

- Among those who received any form of the graduation program, average monthly consumption per capita is between 7.8% (T1) and 8.6% (T2), or 311 to 342 pesos greater than the average among control-group households, in which consumption per capita averaged 3,992 pesos.
- Food security also increased among all graduation groups, with households scoring on average 0.20 (T1) to 0.28 (T3) standard deviations higher than control households on a composite food security index.

- Assignment to any graduation treatment arm increased the average value of productive assets, with treatment households owning productive assets worth 18.5% (T1) to 42.5% (T2) or 1,954 to 4,484 pesos more, on average, than control households whose productive asset value averaged 10,546 pesos.
- There is no evidence of an increase in monthly income for any individual treatment arm or jointly, in part because income is measured noisily. The upper ends of the confidence intervals do not rule out income increases of 25–51%, which encompass the estimated percent increase in consumption.
- In terms of subjective well-being, only households who received the individual livelihood have a statistically significant improvement of 0.08 (T3) and 0.13 (T2) standard deviations in a composite index of mental health and life satisfaction question.
- Members of all three treatment groups were more likely to have savings than the control group by 6 to 15 percentage points. Treatment households also had significantly more savings (between 224 and 443 pesos) than the control groups and were more likely to keep their savings in microfinance institutions by 2 to 4 percentage points.

The preliminary results from the endline survey demonstrate that program beneficiaries of the graduation pilot fared better than the households assigned to the control group in key economic outcomes such as consumption, food security, and accumulation of productive assets. These findings are in line with similar studies that were rigorously evaluated in other contexts. A multi-site randomized control trial (RCT) of graduation programs conducted by Banerjee et al. (2015) found that the programs increased consumption by 0.12 standard deviations, food security by 0.11 standard deviations, and productive asset values by 0.26 standard deviations in the pooled sample, all significant at the 1-percent level. The findings of the current study are particularly important given multiple implementation challenges between 2019 to 2020 along with the COVID-19 pandemic.

Combining these results with estimated costs demonstrates that all variations of the program demonstrate high benefits-to-cost ratios relative to studies in other contexts (Banerjee et al. 2015). While T1 and T2 have very similar ratios of 1159–1162%, assuming 100% persistence, the substantially higher cost of individual-level coaching means that T3 is less cost-effective, with a ratio of 758%. Consequently, even with a very conservative persistence rate of 40%, implying that benefits attrit by 60% annually, the total net benefits of T1 and T2 still exceed the program costs.

## 1. Background

### **The graduation approach**

The graduation approach aims to graduate ultra-poor households living on less than \$1.90 (PPP) a day into more sustainable and stable livelihoods. Pioneered by BRAC in 2002, the program follows a theory of change based on the premise that people in extreme poverty are trapped in an interrelated set of

challenges that include a lack of capital, skills, work opportunities, financial education, nutrition, and confidence. Thus, graduation programs are designed to offer a holistic package of interventions aimed at supporting participants in their short-run needs (i.e. with conditional cash transfers) while promoting sustainable and independent livelihoods in the long run (with the transfer of a productive asset and customized training to make profits out of this small business). The centerpiece of the program is the transfer of a productive asset and the training to manage it, but the program also includes consumption support (like regular cash transfers), life coaching, skills training, access to health information, and components of financial inclusion.

The graduation approach has been adapted to a different context and has been rigorously evaluated in countries including Bangladesh, Ethiopia, Ghana, Honduras, India, and Peru. Rigorous impact evaluations covering a total of 21,000 households have shown that one year after the program ended (three years after the initial asset transfer), participant households on average:

- » Had significantly more assets and savings.
- » Spent more time working.
- » Increased food security and went hungry on fewer days.
- » Experienced lower levels of stress, and
- » Had improved physical health compared to those who did not receive the program.

The program also proved cost-effective: the return on investment (ROI) ranged from 133 percent in Ghana to 433 percent in India. For every dollar spent on the program in the top-performing country, ultra-poor households received \$4.33 in longer-term benefits (Banerjee et al., 2015).<sup>2</sup> Scale-ups of the graduation program are being implemented in countries covering millions of households. The pilot program in the Philippines also proved to be very cost-effective in terms of ROI, with group coaching arms reaching rates of 1126% to 1159% and the individual coaching arm reaching 798%.

### Description of the Philippines graduation pilot

ADB engaged BRAC USA and Innovations for Poverty Action (IPA)<sup>3</sup> to implement and evaluate the graduation pilot in the Philippines in partnership with the Philippines government's Department of Labor and Employment (DOLE). The pilot builds upon the existing DOLE *Kabuhayan* (Livelihood) program, which targets poor and vulnerable households and has been implemented with pre-formed groups. The pilot's beneficiaries were selected among households recently enrolled in the Philippine Department of Social Welfare and Development's (DSWD) *Pantawid Pamilyang Pilipino Program*

---

<sup>2</sup> Banerjee, A. et al (2015). A multifaceted program causes lasting progress for the very poor: Evidence from six countries, Science 15 May 2015, v. 348, issue 6236.

<sup>3</sup> ADB engaged BRAC USA under TA-9017 REG to guide program design and implementation, and IPA under TA 8332-REG to undertake a baseline survey.

(4Ps), which provides recipients with bi-monthly conditional cash transfers (CCT). These transfers range from 1,100 pesos to up to 2,600 pesos per month, depending on family composition.

For the current study, DSWD provided the list of households to be included in the pilot sample and the consumption support those members of the treatment and control groups received.

Under normal circumstances, the Kabuhayan program is available only to those not currently receiving 4Ps cash transfers. To accommodate this study, DOLE issued a special memo to allow 4Ps recipients enrolled in the study to concurrently participate in the livelihood program.

Under the existing DOLE *Kabuhayan* program, beneficiaries of the group-based asset transfer form an association and receive a seed capital fund of a maximum of 1,000,000 pesos depending on the group size<sup>4</sup>. Besides the working capital for selected projects, qualified beneficiaries of selected project proposals also receive training on how to set up, start and operate the livelihood project and technical and business advisory services to support the sustainability of the business. The process involves the election of an executive committee with roles such as president, treasurer, operations manager, and so on. Since the funding disbursed by DOLE for such group livelihoods is a multiple of a typical individual micro-enterprise grant, the range of businesses a group can create is wider and its potential market is larger. Even if not all members are able or willing to contribute labor to a group business, successful enterprises can issue dividend payments to regular members, effectively treating them as shareholders. A group enterprise can also function as a social safety net for members, with funds they can borrow from in case of an emergency, for example.

### **Group vs individual implementation**

Typically, graduation programs fund and foster individual-level small business creation and development. The graduation pilot in the Philippines is the first to measure differences in outcomes between individual and group-based variations of two key components of BRAC's graduation program: i) asset transfers and ii) coaching. These variations align with DOLE's interest in making effective policy decisions about future *Kabuhayan* designs. Additionally, group-based variations may permit efficient scaling of such programs by reducing budget requirements, especially in the costly individual coaching component of the graduation model. The study tests three out of four potential iterations: group livelihoods and group coaching, individual livelihoods and group coaching, and individual livelihoods and individual coaching.

The graduation model as developed by BRAC is normally implemented with individual livelihood and individual coaching components. The logic behind testing group variations of these components

---

<sup>4</sup> Associations of 15-25 members can obtain a maximum financial assistance of 250,000 pesos (Micro-Livelihood), while organizations with 26-50 members can receive a maximum of 500,000 pesos (Small-Livelihood), depending on the project requirements. Associations with more than 50 members are considered Medium-Livelihoods and can benefit of a maximum of 1,000,000 pesos depending on the project requirements.

comes from the possibility of cost savings by way of reduced working hours per coach, social cohesion built from regular meetings, information sharing between members, peer encouragement, group economic activity, and accountability through peer pressure. The cons of group variations however may stem from less individual attention per beneficiary from coaching and potential negative group dynamics and inequality within groups.

Specifically, group-based livelihoods have several potential advantages: they may enable more capital-intensive livelihoods, allow participants to achieve economies of scale, create accountability mechanisms within groups, and increase production volumes. Additionally, the logistics of implementation are streamlined, reducing administrative costs. Conversely, group livelihoods may face challenges due to group coordination, different preferences, different capabilities, and standard moral hazard problems that prevent the transition from informal livelihoods to small firms. Implementing individual livelihoods as a separate treatment arm serves as a proof-of-concept of the graduation program in the Philippines and a benchmark against which to compare group livelihoods. The research design intended to allow further investigation into whether and how individual livelihoods and group livelihoods impact households differently in terms of profitability of businesses, longevity, and sustainability of the enterprise, and effective business management and growth. Further, the nature of group livelihoods suggests that they, relative to individual livelihoods, may change social capital for those in the groups. Depending on how the livelihoods prosper, social capital could be helped or hurt.

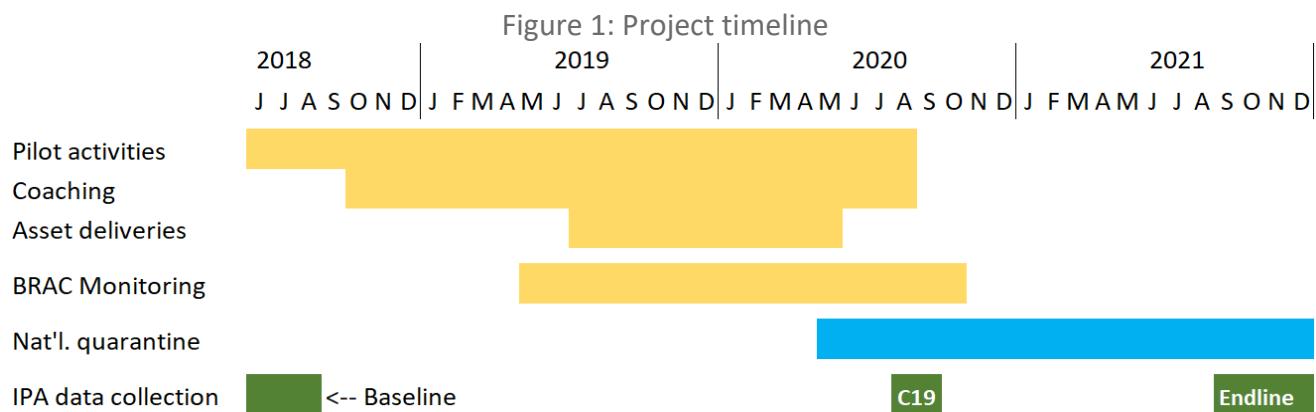
The second variation in the design of the graduation program is whether coaching is provided individually or in groups. Coaching is an integral part of BRAC's graduation approach, and in turn, comprises a large share of the program budget (see Section 0). Coaching accounts for roughly 30 percent of total program costs, and BRAC estimated that group-level coaching sessions can reduce the coaching component cost by 30-40 percent through increased caseloads without decreasing the frequency of each coach's community visits. The coaching sessions aim to foster encouragement, education, skills and confidence, and better access to resources. Through regular sessions, coaches act as trainers, mentors, and progress monitors for participant households. Coaches visit participants at home and provide personal support and business advice throughout the implementation period. The frequency and individualized nature of home visits mean that the caseload per coach needs to be relatively low. To cost-effectively scale up the program, the caseload per coach likely needs to be reasonably high.

The graduation pilot design implemented in the Philippines added group-based and individual variations of the coaching component, which may point to potential differences in impact at the household level. This approach aligns with the interest in DOLE to make appropriate policy decisions for its future *Kabuhayan* design. Group coaching can prove to be as effective as individual coaching if

knowledge retention and behavior change are shown to be equal among treatment groups. Group coaching, however, runs the risk of being less effective if it leads to lower attendance rates.

An accurate account of the details of the intervention differences is key to interpreting any treatment differences (or a lack thereof), both from a research perspective and for the implementing partners. The results of the study are also of high interest to policymakers since it will enable DOLE to determine: (i) the optimal coaching model for its graduation program, and (ii) whether cost efficiencies in group coaching and/or livelihoods can generate sufficient impact for scaling up the approach nationally. BRAC started the implementation of the key graduation-style additions to DOLE's *Kabuhayan* program including regular individual or group-based coaching sessions in October 2018.

Altogether, these interventions aim to lead to sustainable livelihoods, income diversification, asset accumulation, improved income management, confidence in their abilities, and increased savings. These outcomes should give participants not only a better presence but also leave them more prepared for potential shocks in the future.



## Research design

To rigorously estimate program effects, the graduation pilot was designed as a randomized controlled trial among 2,339 program-eligible households interviewed at baseline. Through random assignment after the baseline survey, each household in the sample had the same chance of being assigned to any one of the experimental groups. This ensures that households in the treatment and control groups will, on average, be statistically identical. Any changes which are observed between the groups at the end of the intervention can be attributed to differences in the program rather than to other external or unobserved factors.

Households were grouped into quadrants within each of the 29 barangays and randomly assigned to one of three treatment arms or a control group after conducting a baseline survey, as follows:

1. T1: Households received group assets and group coaching (587 households)
2. T2: Households received an individual asset transfer and group coaching (583 households)
3. T3: Households received an individual asset transfer and individual coaching (583 households)
4. C: Control group (no intervention; 586 households)

Figure 2: Treatment and control group assignment

Intervention	Control (N=586)	T1: Group livelihoods and group coaching (N=587)	T2: Individual Livelihoods and group coaching (N=583)	T3: Individual livelihoods and individual coaching (N=583)
4Ps	✓	✓	✓	✓
In-kind asset transfer	X	\$1,500 per group of 5	\$300 per individual	\$300 per individual
% received transfer	X	62%	73%	78%
Coaching (two-weekly)	X	Group	Group	Individual
Skills training	X	✓	✓	✓
Savings facilitation	X	✓	✓	✓
Community mobilization	X	✓	✓	✓

The sample was drawn from recipient households of the DSWD 4Ps conditional cash transfer program. To remain eligible for the transfers, households are expected to attend monthly Family Development Sessions, pregnant women must avail of pre- and post-natal care, and children up to 18 years of age must regularly attend school and receive preventative health checkups and vaccines.

Just prior to program implementation, IPA conducted a baseline study from June to August 2018 with 2,418 households from the recipient list. Some initially surveyed households were ineligible to participate in the program due to their participation in DSWD's *Sustainable Livelihood Program*, so BRAC and IPA agreed to exclude Barangay IV from the sample and interviewed additional households from barangay VI-A (Victorias City) and barangay Alegria (municipality of Murcia), and assigned them to treatment and control groups in December 2018. However, BRAC did not implement the program with the additional households surveyed. These two barangays were also excluded from this study's analysis. The final sample was 2,339 households.

From August to September 2020, IPA conducted a phone survey with households of the graduation pilot sample to (i) collect information on how these vulnerable households were faring during the COVID-19 pandemic and the community quarantines, and (ii) get timely measures of the graduation pilot on households' income, livelihoods, and well-being that can inform the endline and policymakers who implement livelihood programs in the Philippines. The research team attempted to reach 84.3% of baseline respondents (phone survey sample = 1,972) who had provided at least one phone contact

number.<sup>5</sup> By collecting updated contact information of program participants from BRAC plus reaching out to barangay officials and previously surveyed respondents, IPA interviewed 1,243 (63.0%) of target respondents. See IPA's report, "COVID-19 socio-economic impact of the poor households in Negros Occidental," for additional information on the results of this survey (IPA 2020).

From September to December 2021, IPA conducted an endline survey to assess the impact of the graduation pilot on participants. The survey covered key outcomes such as household consumption, food security, the value of productive assets, household income, and subjective well-being. The survey also covered outcomes in financial health, savings and loan behavior, social networking, child welfare, and women's empowerment. The research team attempted to reach all households interviewed at baseline (2,339 households) and were able to reach 2,288 households, 97.8% of the target sample.

### **Research questions and outcome measures**

The study aims to answer the following research questions:

- Does the classic graduation model (individual livelihood + individual coaching) improve welfare for 4Ps households as much as found in other country contexts?
- Are individual livelihoods more profitable, sustainable, and resilient compared to the control group than group livelihoods?
- Does providing individual livelihood plus individual coaching yield greater impacts in building sustainable livelihoods and reducing poverty compared to the control group than group livelihoods and group coaching, in the context of this DOLE-implemented graduation model?
- Does receiving group vs. individual coaching affect intermediate outcomes such as program engagement, business formation, and social capital?
- What are the potential trade-offs between cost-effectiveness and impact through group coaching and/or livelihoods?
- Are there complementarities between group livelihoods and group coaching that are not present when implementing the program through individual livelihoods or individual coaching?

The impact evaluation of the graduation pilot in the Philippines uses the following primary outcomes in its analysis:

- 1) Household consumption. The primary measure of living standards used in this study is consumption-based. This will be measured by per capita consumption of the following variables:
  - a. Seven-day food consumption
  - b. 30-day non-food consumables
  - c. 12-month non-food consumables
  - d. Durable asset consumption

---

<sup>5</sup> The target sample is equally composed of the control and treatment groups at approximately 25% each.

- 2) Food security index (whether all members of the household get enough food to eat every day), composed of the following measures:
  - a. Food Consumption Score
  - b. Household Food Insecurity Access Scale
  - c. Rice and meat consumption index
  
- 3) Household income, measured as the sum of the following income streams:
  - a. Paid work
  - b. Business revenues
  - c. Agricultural revenues
  - d. Livestock revenues
  - e. Remittances
  - f. Social welfare payments
  
- 4) Total productive asset value or asset holdings, calculated as the sum of the following values:
  - a. Value of currently owned livestock
  - b. Value of livestock structures
  - c. Value of inventory
  - d. Durable asset value used for business
  
- 5) Subjective well-being and life satisfaction, measured in the following ways:
  - a. Kessler Psychological Distress Scale
  - b. Cantril's Ladder like satisfaction scale
  
- 6) Other outcomes of interest, such as:
  - a. Financial health
  - b. Total savings balance and loans
  - c. Child welfare

The five key outcomes of interest were defined and pre-registered prior to examining the endline data.

### **Survey instrument design and programming**

The endline instrument was designed to capture follow-up socio-economic and psychological measures taken from baseline with additional modules on women's empowerment and child welfare. The survey, which took on average 80 minutes to administer, included the following modules:

- Contact information
- Household roster
- Food security
- Assets
- Consumption
- Business

- Livestock
- Plot agriculture
- DOLE livelihoods
- Paid work
- Financial health
- Savings and loans
- Social safety net
- Social networks
- Child welfare
- Women's empowerment
- Subjective well-being
- Remittances
- Shocks

The questionnaire was translated to and delivered in Hiligaynon, which is the predominant language among pilot beneficiaries. IPA programmed and administered the instrument through the mobile data collection platform, SurveyCTO, using Android-based tablets. All data collected through SurveyCTO were housed in a secure server and encrypted through Boxcryptor once transferred to IPA's file-sharing system. Only persons with Institutional Review Board approval to handle data with Personally Identifying Information (PII) were granted the permissions necessary to view PII.

## **Sample selection and randomization**

### **Original sample selection**

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.

### **Randomization**

Random assignment was used to avoid selection and program placement bias. In each of the 30 barangays, about 80 eligible households surveyed by IPA were assigned to one of three treatment groups or the control group by computer. Randomization was conducted by dividing the barangay into 4 smaller geographic regions and assigning one treatment to each. This was done to simplify the logistics of implementation. In the Philippines, rural barangays often cover extensive areas. Had such households from these barangays been assigned to treatment groups completely at random, some participants would have had to travel far for trainings, meetings, and coaching sessions, which would likely affect attendance. The geographical separation of different treatments and the control group also reduces the likelihood of spillovers. 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 was 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<sup>6</sup> across the 30 sample barangays were filled.

---

<sup>6</sup> A minority of barangays have fewer than 80 sample households; some have more than 80.

2. Once all households had been assigned to clusters, randomized assignment to treatment clusters and control clusters began. For treatment groups to be comparable and the control group to be a valid counterfactual, all groups need to have, on average, statistically identical characteristics. The four clusters in all barangays were repeatedly assigned to different treatment groups or a control group. Statistical tests were run on each configuration, comparing every treatment to the control and the treatments to one another based on four key covariates. Finally, the statistically most balanced configuration was chosen.

Table 1 below shows the final treatment and control group assignment by barangay.<sup>7</sup> Households assigned to treatment groups were eligible to receive treatment as soon as the randomization was complete. IPA and BRAC mitigated the risk of differential attrition by minimizing the link between the research and program implementation. This was done through carefully monitoring implementation to ensure assignment to treatment and control groups was maintained.

Table 1: Treatment and control group assignment, by barangay

Barangay	Control	T1 (GrpLH/GrpC)	T2 (IndLH/GrpC)	T3 (IndLH/IndC)	Total
BAGTIC	20	20	20	20	80
BARANGAY IX	20	20	19	19	78
BARANGAY VI POB. (HAWAIIAN)	20	20	19	19	78
BARANGAY VIII (POB.)	20	19	20	19	78
BARANGAY XX	20	20	19	20	79
BARANGAY XXI	20	21	19	19	79
BLUMENTRITT	21	19	19	20	79
CABATANGAN	19	18	19	18	74
CANLANDOG	20	19	20	20	79
CONCEPCION	21	21	20	20	82
CONSING	21	20	20	20	81
DOS HERMANAS	20	19	19	22	80
EFIGENIO LIZARES	20	19	21	20	80
EUSTAQUIO LOPEZ	19	19	20	21	79
GUIMBALA-ON	20	20	20	20	80
GUINHALARAN	20	20	20	20	80
KAPITAN RAMON	19	20	20	20	79
KATILINGBAN	22	22	24	22	90
LANTAD	21	20	21	20	82
MAMBULAC	20	20	20	20	80
MATAB-ANG	20	20	20	20	80
MINOYAN	20	20	20	20	80
RIZAL	20	22	19	19	80

<sup>7</sup> 34 baseline households were assigned to a cluster after confirming their GPS coordinates. They are excluded from Table 1.

SALVACION	20	19	21	20	80
SAN FERNANDO	23	23	23	23	92
SAN ISIDRO	20	20	20	20	80
SANTA ROSA	19	21	19	20	79
ZONE 15 (POB.)	21	26	22	22	91
ZONE 3 (POB.)	20	20	20	20	80
<b>Total</b>	<b>586</b>	<b>587</b>	<b>583</b>	<b>583</b>	<b>2,339</b>

Just after the randomization was completed, 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). The master list of respondents IPA received from the Government of the Philippines should not have included households participating in SLP, as some variations of this program also include a livelihood asset transfer component.

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 drop it, leaving 29 participating barangays.<sup>8</sup>

To avoid the potential bias that could be introduced by dropping treatment households deemed ineligible but not doing the same in the control group across the remaining barangays, IPA kept all original households in the sample during the phone survey and endline surveys.

### Statistical Power

During baseline, using a sample size of 2,415 households, IPA calculated an intra-cluster correlation coefficient (ICC) on key variables of interest (total income, durable asset value, 30-day consumption) ranging from less than 0.001 to 0.0015 within the subdivided barangays. Conservatively using 0.0015, the study was estimated to have 80% power to detect a minimum detectable effect size (MDE) of 0.16 standard deviations between any treatment group and the control group when randomized by sub-barangay clusters, assuming full compliance and no attrition.

---

<sup>8</sup> IPA conducted additional surveys in barangays Alegria, in the municipality of Murcia, and in Barangay VI-A in the City of Victorias, to account for other baseline respondents across the remainder of the sample who were found to be ineligible through BRAC's validation. These barangays were originally dropped from the master list (see subsection

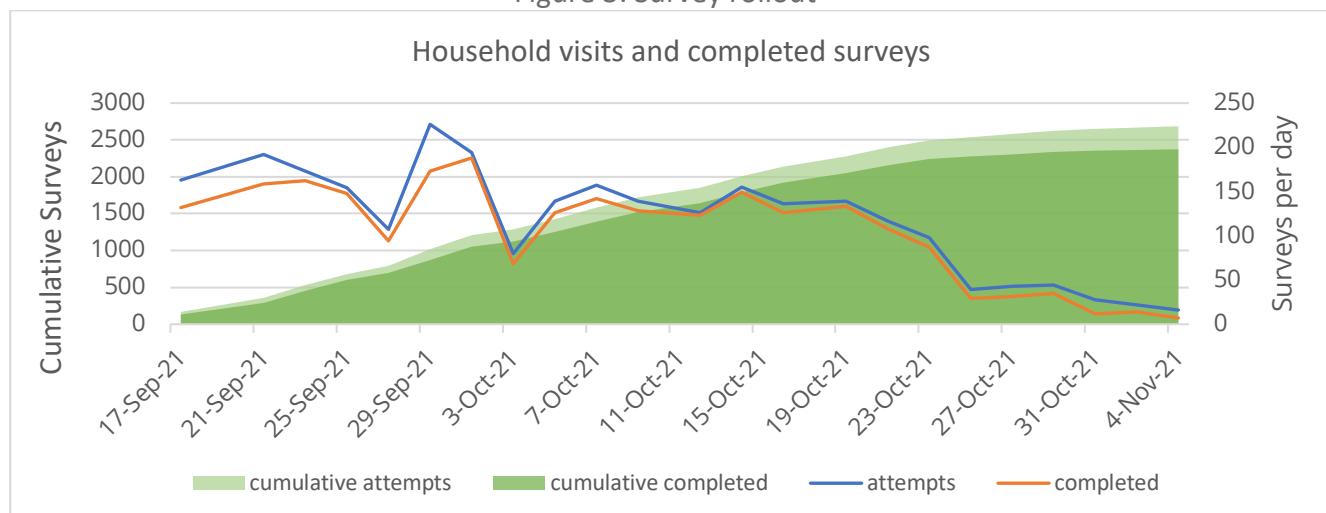
Original sample **selection** above) for having the fewest eligible respondents. However, BRAC did not include these barangays in the treatments, so they are excluded from the endline survey.

## 2. Data Collection

### Endline data collection

The target endline sample comprised households surveyed at baseline. The research team attempted to reach 2,339 households across the original 29 barangays in 5 municipalities of Negros Occidental from mid-September to early November 2022. By November 4, 2,288 households, or 97.8% of the baseline sample, were interviewed. Five teams with a total of 25 enumerators made 2,689 attempts and averaged 2.3 completed surveys per day over 41 workdays. Productivity levels peaked around late September and sharply dropped in late October as teams began sweeping households that were difficult to reach during the first attempt at each barangay. Productivity levels were also affected by the ongoing COVID-19 pandemic, insurgent activities in the City of Victorias and E. B. Magalona in early to mid-October, and inclement weather due to tropical storm Lannie in October.

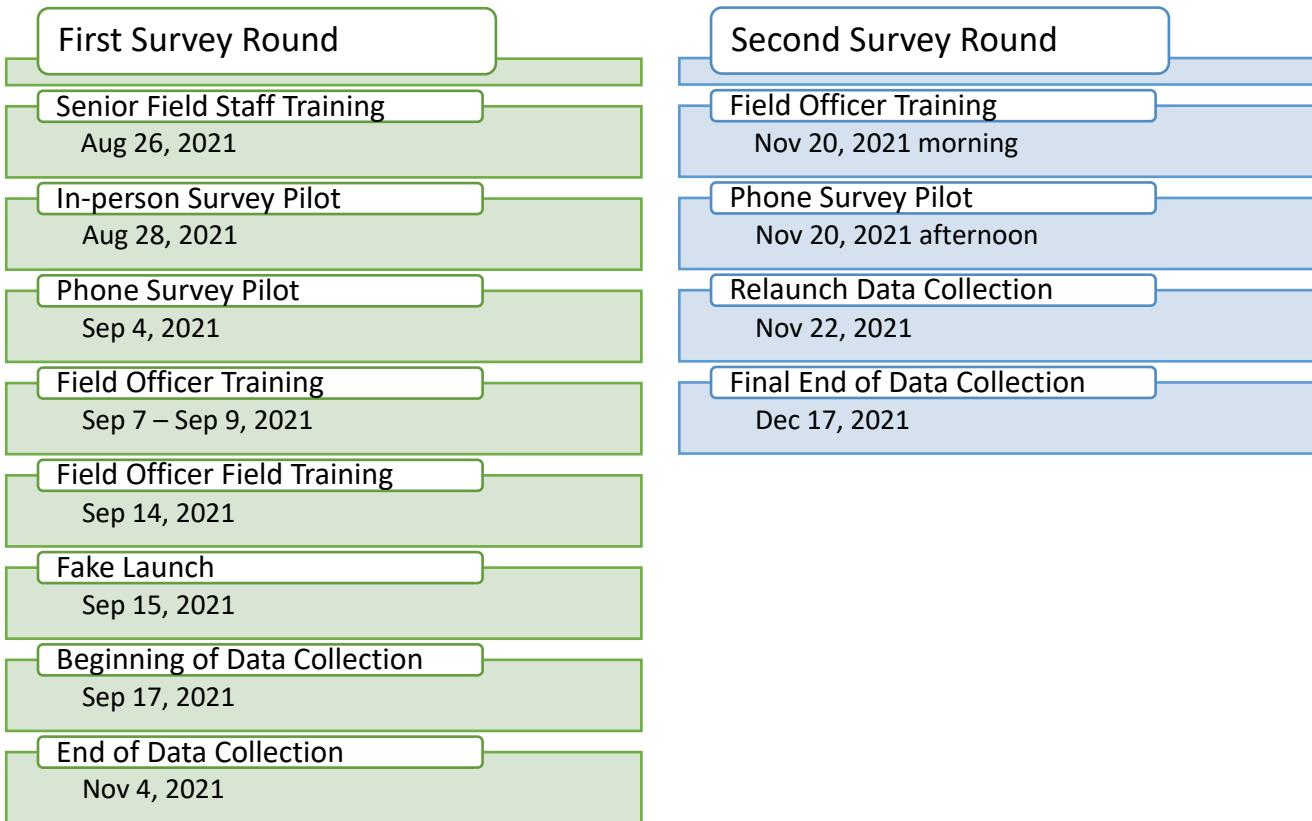
Figure 3: Survey rollout



Note: The survey rollout data in this figure includes interviews from Barangay VI Poblacion in the City of Victorias, which was later dropped in the analysis.

After initial analysis of the endline data, the research team determined that additional data needed to be collected on sources of household income and to address or confirm low reported rates of paid labor. IPA conducted a second round of surveying from November 22 to December 23, 2021, with the initial endline sample. Using updated contact information collected during the first endline survey, 1,802 households were reached by phone while 454 were interviewed in person. The second survey's data collection activities were severely disrupted by Typhoon Rai, halting all interviews from December 16<sup>th</sup> through the 20<sup>th</sup>. After resuming data collection on December 20<sup>th</sup>, it became increasingly difficult to reach the remaining target households due to damage from the typhoon. Data collection for the second survey ended on December 23<sup>rd</sup>, after reaching 2,256 households (98.6% of the 2,288 households interviewed during the first endline survey).

Figure 4: Timeline of endline survey activities



### Survey attrition

Although at least one attempt was made to interview each of the 2,339 baseline households, enumerators were not able to reach 49 of the target households (2.1%). Table 2 shows that the final sample is spread evenly across all treatment and control arms ranging between 97.4% and 98.1% for each group. Respondents moving outside of the survey area with no available contact information was the most cited reason for a household remaining unreached, accounting for about half of all unreached households. In cases when respondents had moved to a location within the survey area, or when updated contact information was obtained, enumerators were able to interview respondents either at their new location or by phone. A total of 73 households (3.2% of endline interviews), were surveyed by phone.

Table 2: Number of households from baseline survey in endline sample

	Control	T1 (GrpLH/GrpC)	T2 (IndLH/GrpC)	T3 (IndLH/IndC)	Total
Baseline respondents	586	587	583	583	2,339
Endline respondents	575	574	571	568	2,288
Endline as % of baseline	98.1%	97.8%	97.9%	97.4%	97.8%

Table 3: Status of unreached households

	Control	T1 (GrpLH/GrpC)	T2 (IndLH/GrpC)	T3 (IndLH/IndC)	Total
Refused to continue survey after consent	1	2	2	0	5
Refusal to consent to survey	1	0	2	1	4
Respondent not reached	3	2	1	3	9
Deceased	0	2	1	3	6
Moved away	6	6	5	8	25
Total	11 (22%)	12 (25%)	11 (22%)	15 (30%)	49 (100%)

### Differential attrition

Breaking this sample further down at the barangay-treatment, or “quadrant” level (the unit of randomization), Table 4 below shows that the final sample is also spread fairly evenly among the 29 barangays and treatment groups of the baseline households. 100% of the target sample were interviewed in about a third of all barangay and more than 93% of respondents in nearly all other barangays. The only exception is barangay Canlandog, where only 86% of households were interviewed due to a large share of respondents reportedly moving outside of the survey area with no available contact information. Table 4 presents a breakdown of respondents by barangay and treatment arm.

Table 4: Endline sample completion rates disaggregated by barangay by treatment arm

Barangay	Control	T1 (GrpLH/GrpC)	T2 (IndLH/GrpC)	T3 (IndLH/IndC)	Total and % of baseline target
BAGTIC	20 (25%)	20 (25%)	20 (25%)	20 (25%)	80 (100%)
BARANGAY IX (DAAN BANWA)	20 (26%)	20 (26%)	19 (24%)	18 (23%)	77 (99%)
BARANGAY VI POB. (HAWAIIAN)	20 (26%)	20 (26%)	19 (24%)	18 (23%)	77 (99%)
BARANGAY VIII (POB.)	20 (26%)	18 (23%)	19 (24%)	18 (23%)	75 (96%)
BARANGAY XX	19 (24%)	20 (25%)	19 (24%)	18 (23%)	76 (96%)
BARANGAY XXI	20 (25%)	20 (25%)	19 (24%)	19 (24%)	78 (99%)
BLUMENTRITT	21 (27%)	18 (23%)	19 (24%)	19 (24%)	77 (98%)
CABATANGAN	19 (26%)	17 (23%)	19 (26%)	18 (24%)	73 (99%)
CANLANDOG	14 (18%)	19 (24%)	17 (22%)	18 (23%)	68 (86%)
CONCEPCION	21 (26%)	20 (24%)	20 (24%)	20 (24%)	81 (99%)
CONSING	21 (26%)	20 (25%)	20 (25%)	20 (25%)	81 (100%)
DOS HERMANAS	20 (25%)	19 (24%)	19 (24%)	22 (28%)	80 (100%)
EFIGENIO LIZARES	20 (25%)	19 (24%)	21 (26%)	20 (25%)	80 (100%)
EUSTAQUIO LOPEZ	18 (23%)	18 (23%)	20 (25%)	21 (27%)	77 (98%)
GUIMBALA-ON	20 (25%)	19 (24%)	17 (21%)	19 (24%)	75 (94%)
GUINHALARAN	19 (24%)	18 (23%)	20 (25%)	19 (24%)	76 (95%)
KAPITAN RAMON	18 (23%)	20 (25%)	20 (25%)	18 (23%)	76 (96%)

KATILINGBAN	22 (24%)	22 (24%)	24 (27%)	22 (24%)	90 (100%)
LANTAD	21 (26%)	20 (24%)	21 (26%)	20 (24%)	82 (100%)
MAMBULAC	20 (25%)	19 (24%)	20 (25%)	19 (24%)	78 (98%)
MATAB-ANG	20 (25%)	20 (25%)	19 (24%)	19 (24%)	78 (98%)
MINOYAN	20 (25%)	19 (24%)	18 (23%)	19 (24%)	76 (95%)
RIZAL	20 (25%)	21 (26%)	19 (24%)	19 (24%)	79 (99%)
SALVACION	20 (25%)	19 (24%)	21 (26%)	20 (25%)	80 (100%)
SAN FERNANDO	23 (25%)	23 (25%)	23 (25%)	23 (25%)	92 (100%)
SAN ISIDRO	20 (25%)	20 (25%)	19 (24%)	20 (25%)	79 (99%)
SANTA ROSA	19 (24%)	21 (27%)	19 (24%)	20 (25%)	79 (100%)
ZONE 15 (POB.)	20 (22%)	26 (29%)	22 (24%)	22 (24%)	90 (99%)
ZONE 3 (POB.)	20 (25%)	19 (24%)	19 (24%)	20 (25%)	78 (98%)
Total (Average)	575 (25%)	574 (25%)	571 (24%)	568 (24%)	2288 (98%)

### Balance on baseline characteristics

Table 5 shows the mean values of baseline outcomes for the sample interviewed in the endline survey. balance tests conducted confirm that there are no statistical differences on observable characteristics, which is a necessary condition to guarantee internally valid estimates for the impact evaluation.

Table 5: Means of baseline characteristics for endline sample (standard deviations in parentheses)

Control	T1 (GrpLH/ GrpC)	T2 (IndLH/ GrpC)	T3 (IndLH/ IndC)	C v. T1	C v. T2	C v. T3	T1 v. T2	T1 v. T3	T2 v. T3	C v. T123
	mean (sd)			p-value						
Household Size	5.92 (1.96)	5.95 (1.95)	5.84 (1.95)	5.97 (1.99)	0.91	0.32	0.75	0.29	0.84	0.22
Respondent gender, =1 if female	0.85 (0.36)	0.86 (0.35)	0.84 (0.37)	0.86 (0.35)	0.67	0.58	0.73	0.32	0.99	0.42
Age of respondent	43.25 (10.83)	43.24 (10.96)	42.52 (11.07)	43.99 (10.43)	0.31	0.78	0.02	0.17	0.20	0.01
Years of education of respondent	8.56 (4.33)	8.69 (4.24)	8.4 (4.29)	8.60 (4.47)	0.24	0.98	0.44	0.18	0.79	0.39
Gender of hh head, =1 if female	0.21 (0.40)	0.21 (0.41)	0.19 (0.39)	0.22 (0.41)	0.67	0.20	0.96	0.31	0.66	0.22
Number of adults	2.89 (1.32)	2.89 (1.32)	2.85 (1.28)	2.92 (1.36)	0.86	0.39	0.86	0.50	0.74	0.34
Number of children	3.02 (1.46)	3.04 (1.49)	2.99 (1.43)	3.04 (1.46)	0.80	0.72	0.74	0.57	0.97	0.50
Total income per person, incl. transfers (30d, pesos 100)	18.87 (50.38)	17.93 (38.09)	19.59 (58.17)	19.08 (52.85)	0.90	0.45	0.65	0.39	0.58	0.82
Total household consumption per person, incl. assets (30d, pesos 100)	21.24 (10.42)	21.14 (9.84)	21.46 (10.30)	21.12 (11.11)	0.73	0.85	0.79	0.60	0.96	0.67
Total value of durable assets per person (pesos 100)	229.46 (349.62)	229.09 (345.82)	225.56 (363.09)	233.74 (340.09)	1.00	0.88	0.86	0.85	0.82	0.66
										0.98

Total value of livestock assets per person (pesos 100)	75.58	70.1	72.02	84.69	0.88	0.96	0.23	0.83	0.18	0.22	0.52
	(182.96)	(163.66)	(189.58)	(194.28)							
Total received from transfers & remittances (30d, pesos 100)	27.43	27.77	26.74	27.8	0.51	0.30	0.60	0.69	0.99	0.74	0.77
	(59.54)	(52.50)	(40.51)	(79.17)							
Household spending (Expenditure module, 30d scale pesos 100)	114.68	116.24	114.94	112.84	0.47	0.76	0.73	0.64	0.29	0.52	0.75
	(52.11)	(53.97)	(53.44)	(48.75)							
Spending (food, 30d pesos 100)	86.82	89.22	86.45	84.76	0.39	0.76	0.36	0.24	0.10	0.50	0.42
	(41.12)	(43.99)	(40.88)	(38.21)							
Spending (non-food, 30d pesos 100)	27.55	26.67	28.2	27.77	1.00	0.18	0.31	0.14	0.26	0.71	0.35
	(19.30)	(16.78)	(21.37)	(19.48)							
Spending (festivals, 30d pesos 100)	0.31	0.35	0.28	0.3	0.06	0.71	0.42	0.13	0.29	0.66	0.27
	(0.90)	(1.04)	(0.69)	(0.93)							
Spending on recent assets (Asset module, 30d pesos 100)	3.77	3.58	3.19	4.54	0.79	0.32	0.22	0.51	0.16	0.03	0.19
	(10.67)	(9.68)	(7.90)	(13.63)							
Total household spending (incl. recent assets, 30d pesos 100)	118.45	119.83	118.13	117.38	0.53	0.93	0.94	0.57	0.49	0.87	0.89
	(54.75)	(56.77)	(54.96)	(52.48)							
Received income from paid work	0.79	0.79	0.8	0.79	0.35	0.43	0.40	0.80	0.86	0.93	0.74
	(0.41)	(0.41)	(0.40)	(0.40)							
Income: self-employment (Y/N)	0.17	0.18	0.16	0.18	0.17	0.77	0.22	0.39	0.96	0.41	0.44
	(0.38)	(0.38)	(0.37)	(0.38)							
Income: livestock (Y/N)	0.11	0.09	0.12	0.11	0.90	0.17	0.38	0.15	0.38	0.60	0.41
	(0.31)	(0.29)	(0.32)	(0.31)							
Income: agriculture (Y/N)	0.09	0.09	0.09	0.09	0.78	0.77	0.98	0.98	0.79	0.79	0.98
	(0.29)	(0.29)	(0.29)	(0.29)							
Income: labor per person (Y/N)	13.73	12.7	14.62	13.86	1.00	0.31	0.56	0.29	0.56	0.73	0.68
	(49.19)	(36.38)	(57.81)	(51.06)							
Household consumption per person (30d, pesos 100)	20.52	20.52	20.81	20.22	0.79	0.78	0.53	0.59	0.69	0.38	0.84
	(9.61)	(9.37)	(9.65)	(9.81)							
Total income from labor, livestock, and agriculture (pesos 100)	105.1	98.52	109.02	107.81	0.66	0.50	0.66	0.24	0.44	0.93	0.66
	(254.05)	(162.05)	(286.45)	(292.89)							
Income: paid work (pesos 100)	77.67	70.75	82.29	80.01	0.79	0.33	0.55	0.19	0.41	0.86	0.55
	(247.03)	(152.67)	(284.18)	(281.84)							
Income: self-employment (pesos 100)	6.37	6.4	6.94	5.78	0.92	0.53	0.67	0.61	0.61	0.35	0.81
	(22.41)	(20.36)	(23.64)	(23.15)							
Income: livestock (pesos 100)	2.24	0.75	-0.01	6	0.62	0.99	0.20	0.66	0.26	0.21	0.62
	(90.81)	(18.23)	(7.87)	(156.46)							
Income: agriculture (pesos 100)	13.44	6.65	19.13	14.59	0.92	0.22	0.54	0.18	0.50	0.73	0.54
	(224.23)	(139.75)	(280.18)	(230.40)							
N	568-575	568-575	568-575	568-575							

### 3. Program implementation and participation

At the beginning of the program, BRAC GCFs used a market assessment tool to determine which livelihoods are appropriate for beneficiaries. This tool has 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 do not have. Most beneficiaries (54%) chose swine fattening across all treatment groups. This was followed by business cart or NegoKart<sup>9</sup> (15%), meat processing (13%), and free-range chicken (8%).

Table 6: Program livelihood asset chosen by arm

	Number of households (total across arms)	T1 (GrpLH/GrpC)	T2 (IndLH/GrpC)	T3 (IndLH/IndC)
	Share of choices in each arm			
Swine Fattening	666	48%	77%	37%
NegoKart Business	190	15%	9%	21%
<i>NegoKart Fried Snacks</i>	121	15%	4%	11%
<i>NegoKart Fried Chicken</i>	67	0%	5%	10%
<i>NegoKart Fried Peanuts</i>	2	0%	0%	0%
Meat Processing	155	14%	9%	15%
Free Range Chicken	99	0%	0%	22%
Working Carabao	52	14%	0%	0%
Chicken Egg Production	32	9%	0%	0%
Backyard Vegetable Farming	24	0%	2%	3%
Salted Eggs Processing	16	0%	2%	2%
Cosmetology/ Massage	4	0%	0%	1%
Fish Vending	1	0%	0%	0%
Total	1239	100%	100%	100%

Note: Data taken from BRAC's administrative records

The research team anticipate differences in livelihood selection between the group and individual livelihood arms; for example, carabao and chicken egg production were only offered to grouped livelihoods (T1) as they were too expensive to provide to individual livelihoods (T2 and T3). However, important differences were also observed in the distribution of livelihood selection between the two individual livelihood treatment arms (T2 and T3). During the initial asset selection meetings, BRAC coaches observed that a very high share of participants preferred swine fattening, although the market assessment had revealed it to be among the less profitable choices. Conversations with BRAC indicated that this preference reflected many stay-at-home mothers who preferred home-based livelihood options that allowed them to manage their livelihoods while attending to other household duties, despite lower estimated net profits. Additionally, many respondents had some familiarity with swine fattening. Coaches began emphasizing other livelihoods as the meetings progressed, and BRAC

<sup>9</sup> This is an abbreviation of “Negosyo sa Kariton” which translates to business cart.

added a free-range chicken option. Because T3 asset meetings generally took place after T2 asset meetings, this led to 77% of T2 participants selecting swine fattening, compared with 37% of T3 members.

### Implementation challenges

BRAC faced multiple challenges that delayed livelihood asset delivery, and as a result, a substantial share of respondents dropped out of the program. Attrition was particularly high among members of T1, the group-livelihood and group-coaching treatment arm, who were last to receive their livelihood assets, by March 2020 (just before the start of the COVID-19 pandemic and the national lockdown). These delays were largely attributed to administrative hurdles in securing approvals for proposals, procurement, and delivery. Some of these hurdles identified by BRAC were inconsistencies in the review of proposals, lack of staff capacity at the regional office level, and variation in requirements and forms across partner organizations. BRAC also faced procurement challenges due to price fluctuations, inflation, and difficulty reaching rural households. Figure 7 shows in more detail how these challenges aligned with the study timeline.

Program activities began in October 2018 when individual and group life-skills coaching was launched. BRAC received approval from DOLE to proceed with procuring livelihood assets for T2 and T3 in June 2019 and later received approval for T1 in December 2019. This meant that they started deliveries of individual livelihoods nine months after program implementation was launched with the start of the life-skills coaching component, while the group livelihood deliveries were delayed fifteen months after the start of the program implementation.

In terms of asset selection, a market assessment tool was used to determine which livelihoods were appropriate for certain beneficiaries. This tool contained a list of requirements for the operation of each livelihood to help BRAC's GCFs make recommendations. During monitoring interviews conducted by IPA in August 2019, GCFs mentioned skills, experience, interests, and location as their most important considerations when suggesting livelihood assets towards or away from beneficiaries. The manner in which GCFs made livelihood recommendations to different treatment groups may explain some of the differences seen in the take-up of certain livelihoods across these groups. According to BRAC's monitoring reports, an overwhelming majority of households chose swine fattening by November 2018 (73% of individual livelihoods chosen at that time). This led BRAC to introduce additional packages to encourage other livelihood options, such as free-range chickens and massage therapy. Timing may have also played a role since funding for T3 households was not secured until October 2018, which delayed project activities in terms of asset selection for this group. By December 2018, while T1 and T2 households were registering their livelihoods with DOLE, T3 households were still undergoing targeting verification surveying and holding barangay assemblies.

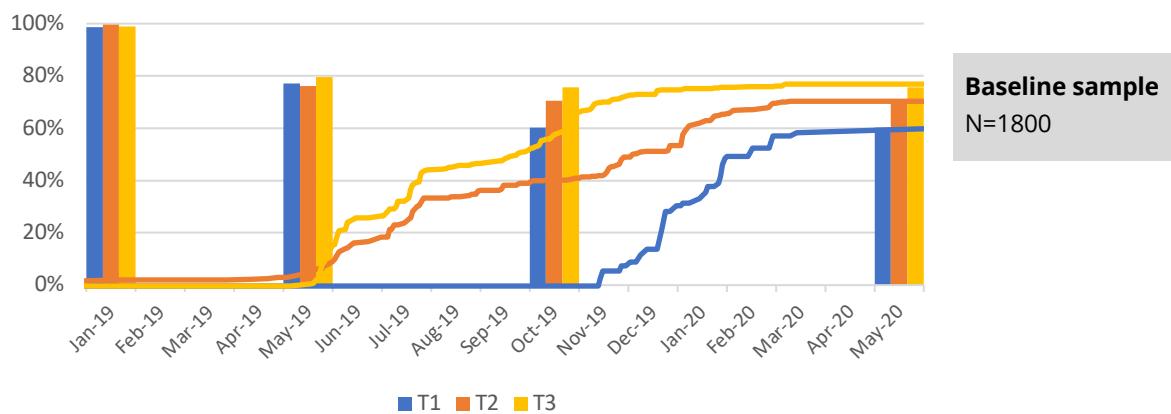
Unfortunately, the project experienced a major shock due to the COVID-19 pandemic: in March 2020 the national government imposed strict quarantine measures, halting BRAC's life-skills coaching sessions and program participants' normal operations. Coaching activities resumed in June 2020, and implementation of the graduation pilot was completed in Negros Occidental in September 2020.

The socioeconomic impact of the pandemic in the Philippines has been devastating. Across the Philippines, joblessness reached a record high of 45.5%,<sup>10</sup> and more than 164,000 workers were reported displaced in September 2020.<sup>11</sup> As of February 2022, Negros Occidental recorded 39,874 confirmed cases of COVID-19 and 1,786 related deaths.<sup>12</sup> The province, like many others in the country, has been under community quarantines with varying restrictions since March 2020.

## Participation rates

The number of participants remained relatively equal across treatment arms in May 2019 but dropped at different rates after five months. As of October 2019, the number of participants for T1, T2, and T3 declined by 40%, 30%, and 24%, respectively, despite BRAC's best efforts to prevent further attrition. The GCFs attribute this attrition to the extensive differential delay in livelihood asset deliveries, especially for T1 households, who did not begin to receive their livelihood assets until November 2019, about 6 months after the beginning of asset deliveries for the other two treatment arms, as depicted in Figure 5.

Figure 5: Participation share and asset delivery over time, differential across arms



Note: The graph is based on the baseline sample (N=1800). Participation share is represented by the bar graph while the share of respondents that received the livelihood asset is represented by the line graph. The data is based on BRAC's administrative records.

<sup>10</sup> Social Weather Stations (2020, August 16). SWS July 3-6, 2020 National Mobile Phone Survey – Report No. 16: Adult joblessness rises to record-high 45.5%.

<sup>11</sup> Aquino, L. (2020, September 07). DOLE monitoring report shows an increase in number of displaced workers.

<sup>12</sup> Provincial Government of Negros Occidental (2022, February 02). Negros Occidental COVID-19 Update.

In the phone survey conducted in September 2020, respondents were asked if they were ever contacted by BRAC before or during program implementation (see Table 7). 13% of respondents in the treatment groups reported that they were contacted by BRAC but later opted out of the program. Of these respondents, their reported reasons for leaving were having no time (33%), disliking the treatment group assignment (18%), disliking the livelihood options (7%), being a beneficiary of a similar program (5%), and miscommunication about the livelihood (5%).

Table 7: Reported reasons for opting out of program

		Mean	N
HH was contacted by BRAC but opted out of the program because...	No time	33%	150
	Did not like treatment groups	18%	150
	Other	15%	150
	Cannot raise livelihood in area	9%	150
	Did not like livelihood options	7%	150
	Beneficiary of other programs	5%	150
	Miscommunication	5%	150
	Relocated	4%	150
	Opted for other work opportunities	3%	150
	Improved economic status	2%	150

### Monitoring activities and data analysis

BRAC conducted household monitoring activities from April 2019 until September 2020. Twice per month, GCFs collected welfare and livelihood data using KoBo Toolbox, an open-source data collection platform through Android-based tablets. IPA assisted in setting up these monitoring tools. These data were analyzed monthly and used to make adjustments to the coaching curriculum, such as providing increased support for livestock management. After the start of the pandemic, BRAC began conducting remote monitoring and reduced the frequency of in-person coaching as well as the size of group coaching attendants.

IPA also hired a field manager to maintain a regular field presence and strong communication with BRAC. The field manager attended recruitment and livelihood selection meetings as well as selected coaching sessions and trainings in order to share qualitative information with the research team. She also communicated the challenges GCFs initially experienced with the monitoring tools so that revisions could be made. Finally, she implemented a brief survey of the GCFs in order to understand how livelihood options were presented and received by program participants. Further analysis of the monitoring data from the project's implementation phase shows that coaching activities ran from May 11, 2018, to February 20, 2020, with the bulk of sessions conducted between June and October of 2019. On average, each GCF served 43 beneficiaries through individual coaching sessions per month and 230 beneficiaries through group coaching sessions per month.

Data on attendance rates for coaching sessions show that rates differed substantially between treatment groups. Unconditional attendance rates (not accounting for attrition) for T1 households are estimated to be just 32% while conditional attendance rates (accounting for attrition) sat at just 52%. T3 had the highest attendance rates among treatment groups, with a 65% unconditional attendance rate and 84% conditional attendance rate. This may also be a consequence of delays in asset deliveries amongst treatment groups, especially in the substantial delay of group livelihoods (T1) deliveries.

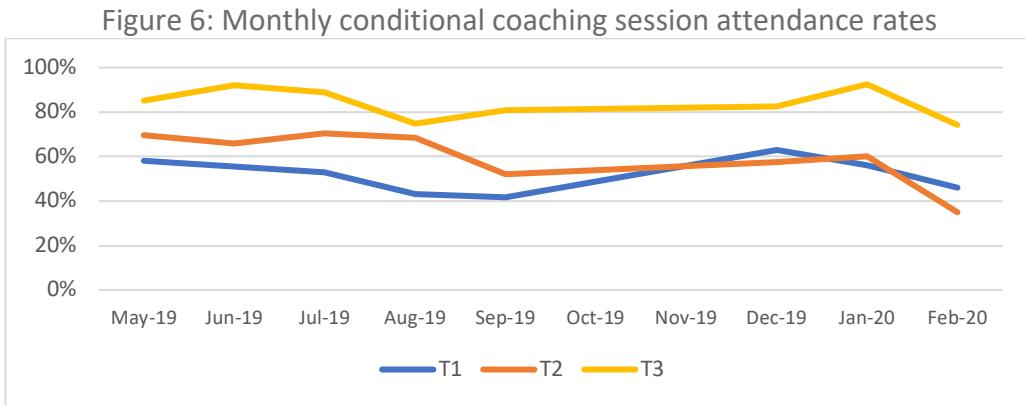


Table 8: Overall coaching session attendance rates by treatment arm

	T1 (GrpLH/GrpC)	T2 (IndLH/GrpC)	T3 (IndLH/IndC)
Unconditional attendance rates	32%	43%	65%
Conditional attendance rates	52%	60%	84%

IPA also conducted monitoring activities, administering two qualitative surveys on livelihood selection and suppliers' impression of government procurement. From April to June 2019, IPA interviewed 18 GCFs regarding topics such as methods of guiding beneficiaries and considerations when suggesting livelihoods. IPA also conducted interviews with suppliers to better understand the significant delays in the procurement of livelihood assets.

## 4. Empirical specification

Intention-to-treat (ITT) effects of each graduation treatment arm were estimated using the following specification:

$$y_{ibq} = \beta_0 + \beta_1 treat1_{bq} + \beta_2 treat2_{bq} + \beta_3 treat3_{bq} + \mathbf{X}'_i \boldsymbol{\gamma} + f_b + \epsilon_{ibq}$$

where  $y_{ibq}$  is the outcome of interest for individual  $i$  living in quadrant  $q$  of barangay  $b$ . The treatment indicators  $treat1_{bq}$ ,  $treat2_{bq}$ , and  $treat3_{bq}$  correspond to quadrant-level assignment to

the group livelihood-group coach treatment (treatment 1), the individual livelihood-group coaching treatment (treatment 2, and the individual livelihood-individual coaching treatment. A vector of individual-level covariates was included,  $X_i$  used in the re-randomization procedure which includes durable assets, value of livestock, number of household members, and number of adults, along with barangay fixed effects,  $f_b$ . Standard errors are clustered at the barangay-quadrant level, the unit of randomization.

Estimates of  $\beta_1$ ,  $\beta_2$ , and  $\beta_3$  were the focus of analysis, which reflect the causal impact of each assignment to each treatment arm versus the control group, along with the joint test of their significance. The impact of assignment to group livelihoods, holding constant group coaching, was measured by testing whether  $\beta_1 = \beta_2$ , and the impact of assignment to individual coaching, holding constant individual livelihood assignment, was measured by testing whether  $\beta_2 = \beta_3$ .

Two additional adjustments were made for key outcomes in Table 10, which were pre-specified in the registered analysis. The baseline measure of the outcome variable, or the closest approximation, was included in each regression. Additionally, standard p-values are reported alongside Anderson sharpened q-values, which hold constant the false discovery rate across the 30 primary hypothesis tests (6 tests per outcome, and 5 outcomes total).

## 5. Endline results

Among the 1,713 households assigned to a graduation treatment arm interviewed at endline, 60% of respondents said that their household has ever managed a BRAC livelihood, and 29% said that they currently have a livelihood set up by BRAC. In comparison, administrative data from BRAC indicates that 62–78% of households received a livelihood. Overall, the means for currently or ever having a BRAC livelihood are higher for treatment group 3 than for the other two treatment groups, with treatment group 1 recording the lowest means. Table 9 presents the means for each treatment group on the abovementioned variables.

Table 9: Program livelihood by treatment group

	Control	T1 (GrpLH/GrpC)	T2 (IndLH/GrpC)	T3 (IndLH/IndC)	Any T
	Mean	Mean	Mean	Mean	Mean
Any household member ever attended livelihood trainings conducted through BRAC	0.04	0.68	0.76	0.80	0.75
Any household member ever managed a BRAC livelihood	0.01	0.50	0.62	0.68	0.60
Currently has a livelihood set up by BRAC	0.00	0.24	0.29	0.33	0.29
N	575	574	571	568	1713

## Key welfare outcomes

Table 10 shows that each version of the graduation intervention improves monthly household consumption, food security, and the total value of productive assets, which comprise three of the five key welfare outcomes, and these effects are robust to adjustments for multiple hypothesis testing. The analysis reveals no general increase in monthly income and finds that only individual livelihoods have a statistically significant impact on participant well-being.<sup>13</sup>

Table 10: Treatment effects on key welfare outcomes

	(1) Monthly consumption per capita (pesos)	(2) Food security index	(3) Value of productive assets (pesos)	(4) Monthly income (pesos)	(5) Subjective well- being index
T1: Grp LH / Grp C	311** (119) [0.01]	0.20*** (0.05) [<0.0 1]	1954* (1054) [0.09]	4 (2072) [0.67]	0.05 (0.05) [0.28]
T2: Ind LH / Grp C	342*** (96) [<0.01]	0.26*** (0.05) [<0.01]	4200*** (1296) [<0.01]	1143 (2005) [0.45]	0.13*** (0.04) [<0.01]
T3: Ind LH / Ind C	334*** (114) [0.01]	0.28*** (0.05) [<0.01]	4484*** (1044) [<0.01]	2893 (2785) [0.26]	0.08* (0.04) [0.09]
Mean in Control group	3992	-0.00	10546	16451	-0.00
SD in Control group	1993	1.00	19429	31893	1.00
N	2287	2288	2288	2288	2288
R2	0.13	0.16	0.21	0.04	0.10
P-values of H0:					
Any Treatment =					
Control	0.00 [<0.01]	0.00 [<0.01]	0.00 [<0.01]	0.68 [0.46]	0.03 [0.04]
T1 = T2	0.79 [0.54]	0.35 [0.28]	0.08 [0.09]	0.52 [0.42]	0.07 [0.09]
T2 = T3	0.94 [0.64]	0.66 [0.46]	0.82 [0.55]	0.49 [0.41]	0.19 [0.17]

Notes: Standard errors are clustered at the quadrant level, which was the level of random assignment to experimental groups. Brackets indicate Anderson sharpened q-values. See text for details. All specifications include barangay fixed effects along with individual-level baseline controls used for re-randomization (durable assets, value of livestock, number of household members, and number of adults) during random assignment. Each column also includes the baseline measure of that outcome variable: per-capita consumption, days gone to bed hungry in past 7 days, baseline assets, baseline income, and mental-health index. Statistical significance denoted by \*\*\* for p<0.01, \*\* for p<0.05, and \* for p<0.1

In terms of consumption, control households spent 3,992 pesos on average each month on food, non-food consumables, services, and celebrations for each household member. Assignment to any graduation treatment arm increased average monthly per-capita consumption by 311 to 342 pesos (7.8% to 8.6%) compared to the control group. Food security also increased among all treatment groups, with households scoring 0.20 to 0.28 standard deviations higher than control households on

<sup>13</sup> Table 39 reports impacts on consumption, value of productive assets, and monthly income (columns 1, 3, and 4) in logged terms.

this index. The last key outcome with significant findings for all treatment groups is the value of productive assets, with treatment households owning additional productive assets worth 1954 to 4484 pesos (18.5% to 42.5%) more than control group households, whose total productive assets average 10,546 pesos. The impacts of the program on income, however, are noisy and not statistically significant.<sup>14</sup> The largest impact is among those in T3, who report a 17% increase in income relative to the control group (2,893 pesos), but this is imprecisely estimated ( $p = 0.299$ ). Graduation increases respondents' self-reported subjective well-being by 0.05 (T1) to 0.13 (T2) standard deviations, but impacts are only statistically significant among the individual livelihood treatment arms (T2 and T3).

## Household consumption

Household consumption was calculated as the sum of food, non-food consumables (including services and gifts), durable household assets (such as furniture and appliances), and festival expenditures (such as spending on birthdays and weddings). Data on food consumption was recorded using a 7-day recall period. Non-food consumables were divided into recall periods of 30 days and 12 months. Data on festival expenditure also used a recall period of 12 months. All consumption spending was adjusted to reflect a 30-day period and divided by the sum of adult equivalents in each household.

Table 11: Consumption per capita per month (in pesos)

	(1) Total	(2) Food	(3) Non-food consumables	(4) Durables
T1: Grp LH / Grp C	311*** (119)	92 (57)	210** (91)	9** (3)
T2: Ind LH / Grp C	343*** (96)	71 (65)	259*** (66)	12*** (3)
T3: Ind LH / Ind C	334*** (114)	149** (67)	175** (70)	10*** (3)
Mean in Control group	3992	2611	1334	48
SD in Control group	1993	1222	1157	66
N	2287	2287	2287	2287
R2	0.13	0.14	0.06	0.13
P-values of H0:				
Any Treatment = Control	0.00	0.17	0.00	0.00
T1 = T2	0.79	0.71	0.61	0.28
T2 = T3	0.94	0.25	0.27	0.46

Notes: Food consumption based on 7-day recall, non-food consumption (including festival consumption) based on 30-day and 12-month recall. All items rescaled to monthly consumption. Durables are defined as appliances and furnishings within the household; per-month durable consumption is computed as 1/12 of 10% of the value of durable household asset holdings. Monthly consumption reflects sums of all items. All values in Philippine Pesos. All specifications include barangay fixed effects along with individual-level baseline controls used for re-randomization (durable assets, value of livestock, number of household members, and number of adults) during random assignment. Standard errors are clustered at the quadrant level, which was the level of random assignment to experimental groups. See text for details. Statistical significance denoted by \*\*\* for  $p < 0.01$ , \*\* for  $p < 0.05$ , and \* for  $p < 0.1$ .

<sup>14</sup> Measures of household income are traditionally noisy due to large variation in income, particularly among those engaged in self-employment (de Mel, McKenzie, and Woodruff, 2009) and agriculture (Gollin and Udry, 2021), along with difficulty in pricing items and activities.

Table 11 presents the impact of each graduation program variation on per-capita monthly household consumption, disaggregating consumption into food, non-food consumables, and durables. Although there are general positive impacts across all three dimensions, the overall increase in total consumption is driven in large part by non-food consumables (such as toiletries, home goods, and services), which makeup 68–76 percent of total increases and is statistically significant for all treatment groups. Increases in per-capita monthly food consumption are between 71–149 pesos, although this increase is only significant for households in T3. Small but statistically significant increases of 9–10 pesos in the value of durable assets can also be seen for all treatment groups.

## Food security

The food security measure is an index composed of three different scores and indices. The first is the Food Consumption Score (FCS), which is the sum of weighted frequencies of a household's consumption of eight food groups over a 7-day recall period.<sup>15</sup> The second is the Household Food Insecurity Access Scale (HFIAS), which is the sum of the frequencies of nine different food access conditions over a four-week recall period. The HFIAS is normally calculated on a scale of zero to three for each condition, with zero signifying that the household has not experienced a given condition over the 4-week recall period.<sup>16</sup> For consistency with the other measures, this analysis reverses the scoring system, making three the most favorable score and zero the least favorable. The last measure included in the food security index is a rice and meat consumption index derived from questions pertaining to the quantity and frequency of meat and rice consumption over a 7-day recall period.

Table 12: Food security

	(1) Food security index	(2) Food Consumption Score	(3) Reversed HFIAS Score	(4) Rice and meat consumption index
T1: Grp LH / Grp C	0.20*** (0.05)	3.74*** (1.03)	0.85*** (0.29)	0.03 (0.04)
T2: Ind LH / Grp C	0.25*** (0.05)	5.13*** (1.04)	1.05*** (0.25)	0.02 (0.04)
T3: Ind LH / Ind C	0.28*** (0.05)	5.51*** (0.99)	0.83*** (0.26)	0.08** (0.04)
Mean in Control group	0.00	62.01	20.50	0.00
SD in Control group	1.00	18.56	5.22	1.00
N	2288	2288	2288	2288
R2	0.16	0.09	0.07	0.14
P-values of H0:				
Any Treatment = Control	0.00	0.00	0.00	0.13
T1 = T2	0.35	0.19	0.48	0.80
T2 = T3	0.66	0.71	0.39	0.16

Notes: The Column 1 food security index is an equally weighted average of the three food security measures in Columns 2–4, normalized to the control group. Column 2, Food Consumption Score (FCS),

<sup>15</sup> International Dietary Data Expansion Project (2021, January 27). Food Consumption Score (FCS).

<sup>16</sup> International Dietary Data Expansion Project (2021, January 27). Household Food Insecurity Access Scale (HFIAS)

groups consumed over the last 7 days. Column 3, Reversed HFIAS Score, is calculated using the frequency of occurrence of specific food insecurity events over the last 4 weeks. The score was reversed so that a higher score signals more positive outcomes. Column 4 is the aggregate score of three food security measures used during the COVID-19 phone survey conducted in 2020, namely how much rice and meat the household had consumed over the past 7 days. All specifications include barangay fixed effects along with individual-level baseline controls used for re-randomization (durable assets, value of livestock, number of household members, and number of adults) during random assignment. Standard errors are clustered at the quadrant level, which was the level of random assignment to experimental groups. Statistical significance denoted by \*\*\* for p<0.01, \*\* for p<0.05, and \* for p<0.1

Table 12 presents the results of the analysis of impacts on food security. All treatment groups reported significantly higher Food Consumption Scores as compared to the control group (3.74 to 5.51 points higher), mostly driven by increased consumption of pulses, vegetables, fruit, milk, and sugar. The same is true for the reversed HFIAS score, where the frequency of food access conditions is reduced by 0.83 to 1.05 points in any of the treatment groups. Only T3 saw a significant impact on rice and meat consumption, with an increase of 0.08 standard deviations above the control group, significant at the 5% level.

### Productive asset ownership

Productive asset ownership is calculated as the sum of the current value of livestock, livestock structures, business asset values, and business inventory values. Business assets are defined as any refrigerator or freezer, bicycle or pedicab, motor cab or motorcycle, sewing machine, washing machine, or any other unlisted asset the household may have used for business. Inventory is defined as any raw or final products the household maybe have, ready for sale or use to make a final product.

Table 13: Value of productive assets (in pesos)

	(1)	(2)	(3)	(4)	(5)
	Value of productive assets	Livestock value	Livestock structure value	Business asset value	Business inventory value
T1: Grp LH / Grp C	1954*	868	133	881***	72
	1054	946	117	315	84
T2: Ind LH / Grp C	4200***	3060**	289**	597*	254***
	1296	1218	126	327	94
T3: Ind LH / Ind C	4484***	2497**	402***	1258***	327***
	1044	993	131	334	101
Mean in Control group	10546	8947	594	712	292
SD in Control group	19429	18452	1927	3230	1591
N	2288	2288	2288	2288	2288
R2	0.21	0.23	0.07	0.05	0.05
P-values of H0:					
Any Treatment = Control	0.00	0.02	0.01	0.00	0.00
T1 = T2	0.08	0.07	0.19	0.42	0.04
T2 = T3	0.82	0.65	0.39	0.07	0.49

Notes: Column 2, total value of all livestock currently owned. Column 3, total value of all livestock structures currently owned. Column 4, total value of all assets currently used for business activities. Column 5, total value of all finished and

unfinished goods and inventory. All values in Philippine Pesos. All specifications include barangay fixed effects along with individual-level baseline controls used for re-randomization (durable assets, value of livestock, number of household members, and number of adults) during random assignment. Standard errors are clustered at the quadrant level, which was the level of random assignment to experimental groups. Statistical significance denoted by \*\*\* for  $p < 0.01$ , \*\* for  $p < 0.05$ , and \* for  $p < 0.1$

Table 13 presents the impact of the treatment arms on productive asset ownership. For T1, the increased value of productive assets (1,954 pesos) is driven roughly equally by increases in the value of livestock (868 pesos) and the value of business assets (881 pesos), although only the impact on business assets is statistically significant ( $p < 0.01$ ). Members of the other two treatment groups show a greater increase in the value of livestock and livestock structures, both in relative and absolute terms, though only the increase in T3 livestock value is statistically significantly different from the T1 increase ( $p = 0.07$ ). Additionally, members of both groups see substantial increases in business inventories. Altogether, members of T2 and T3 see a 40–43% increase (4,200–4,484 pesos) in the total value of productive assets relative to the control group.

### **Household income**

Household income is calculated as the sum of different income streams the household may have had in the form of paid work, business revenue, revenue from agriculture and livestock, remittances, and social assistance (including 4Ps transfers). The analysis of these different income sources revealed no significant differences between the control group and any of the three treatment groups. Although there was a 2,893 peso increase in T3 households' monthly income over the control group, primarily driven by the 1,751 peso increase in net agricultural revenue, this finding is not statistically significant at any level ( $p = 0.299$ ) (see Table 14).

Although observed changes in income are not statistically significant, the average percentage increases in income across all three treatment groups relative to the control mean are broadly in line with the percentage increases relative to the control mean for consumption. These increases are well within the confidence intervals for these estimates. However, the standard errors for these estimates are large, making it difficult to detect statistically significant impacts, as the minimum detectable effect size equals 2.8 times the standard error.

Two potential explanations for the lack of meaningful impacts of the program on net revenue from agriculture and livestock (with the exception of the large but imprecise impacts on T3 agriculture) are either there was little additional economic activity in those domains, or revenue increased but was largely offset by higher costs. To investigate this further, Table 15 further disaggregates the net agricultural and livestock income components in Table 14. Overall, T1 and T2 led to little increase in even gross agricultural revenue. In contrast, gross livestock revenue rose across all treatment groups, although it was not individually statistically significant for T1; however, these are largely negated by

nearly equal increases in livestock expenses. All groups also experienced net losses in the value of livestock owned.

Table 14: Household income per month (in pesos)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Total (sum of columns 2-7)	Paid work	Business profits	Net revenue from agriculture	Net revenue from livestock	Remittances	Social assistance
T1: Grp LH / Grp C	4 (2072)	244 (293)	32 (168)	-296 (1999)	87 (128)	-60 (418)	35 (122)
T2: Ind LH / Grp C	1124 (2005)	396 (250)	131 (148)	-28 (1845)	25 (108)	633 (491)	-14 (145)
T3: Ind LH / Ind C	2893 (2785)	54 (242)	318* (169)	1751 (2600)	167 (114)	566 (473)	42 (161)
Mean in Control group	16451	5553	761	5483	-153	1502	3326
SD in Control group	31893	4797	2868	30368	1121	7861	3611
N	2288	2288	2288	2288	2288	2288	2267
R2	0.04	0.05	0.04	0.04	0.02	0.01	0.04
P-values of H0:							
Any Treatment = Control	0.68	0.37	0.26	0.86	0.51	0.33	0.98
T1 = T2	0.53	0.60	0.54	0.87	0.65	0.15	0.72
T2 = T3	0.49	0.15	0.25	0.44	0.27	0.90	0.74

Notes: Column 2, total value of all livestock currently owned. Column 3, total value of all livestock structures currently owned. Column 4, total value of all assets currently used for business activities. Column 5, total value of all finished and unfinished goods and inventory. All values in Philippine Pesos. All specifications include barangay fixed effects along with individual-level baseline controls used for re-randomization (durable assets, value of livestock, number of household members, and number of adults) during random assignment. Standard errors are clustered at the quadrant level, which was the level of random assignment to experimental groups. See text for details. Statistical significance denoted by \*\*\* for p<0.01, \*\* for p<0.05, and \* for p<0.1.

Table 15: Household income components per month (in pesos)

	(1)	(2)	(3)	(4)	(5)
	Agriculture gross revenue	Agriculture expenses	Livestock gross revenue	Livestock expenses and purchases	Change in livestock value
T1: Grp LH / Grp C	-264 (2012)	16 (23)	192 (130)	105* (58)	-93 (67)
T2: Ind LH / Grp C	38 (1856)	45** (23)	259** (106)	234*** (53)	-193*** (58)
T3: Ind LH / Ind C	1713 (2608)	27 (21)	344*** (123)	178*** (54)	-159** (61)
Mean in Control group	5491	52	149	301	-199
SD in Control group	30616	317	1098	762	885
N	2288	2288	2288	2288	2288
R2	0.04	0.06	0.04	0.08	0.06
P-values of H0:					
Any Treatment = Control	0.88	0.24	0.02	0.00	0.01
T1 = T2	0.85	0.25	0.60	0.02	0.11
T2 = T3	0.47	0.45	0.49	0.28	0.54

All specifications include barangay fixed effects along with individual-level baseline controls used for re-randomization (durable assets, value of livestock, number of household members, and number of adults) during random assignment. Standard errors are clustered at the quadrant level, which was the level of random assignment to experimental groups. See text for details. Statistical significance denoted by \*\*\* for p<0.01, \*\* for p<0.05, and \* for p<0.1

Table 16 reports the impact of the program on the number of income sources households have. Overall, control-group households averaged 1.6 income sources over the past three months. Households in T3 reported having 0.12 more income sources, significant at the 1% level, while T2 households reported having 0.06 more income sources, significant at the 10% level. Roughly one-third of households had to stop business operations and one-third had to stop workplace operations due to quarantines over the past year, and the graduation program did not affect the likelihood of disruptions in either domain.

In the absence of the graduation program, business ownership is relatively common, as roughly one-fourth of control group households own a business (22%) (see Table 17). Assignment to the individual-livelihood graduation arms has a modest impact on business ownership, increasing it by 4 percentage points (T2) to 9 percentage points (T3). Nearly all control group business owners (88% of those who owned a business, or 17% overall) operated it in the past 30 days, and the likelihood is 4 percentage points higher for T2 and 6 percentage points for T3. Total business expenses were also found to be significantly higher for households in T1 and T3 over the control group (737 and 494 pesos more), while total monthly sales were found to be higher only for T3 (by 709 pesos). Households in T2 were also 6 percentage points more likely to have a business operating within the past 30 days.

Table 16: Income sources

	(1)	(2)	(3)	(4)	(5)	(6)
	Number of income sources 3 months	Household had at least one income source in past 3m	Household had at least one business income source in past 3mo	Unexpected sale of business asset in past 12m	Stop business operations due to quarantine in past 12m	Stop workplace operations due to quarantine in past 12m
T1: Grp LH / Grp C	0.03 (0.04)	-0.02 (0.02)	0.02 (0.02)	-0.01 (0.01)	-0.01 (0.02)	-0.02 (0.02)
T2: Ind LH / Grp C	0.06* (0.04)	0.00 (0.01)	0.02 (0.02)	-0.01 (0.01)	-0.01 (0.02)	-0.00 (0.02)
T3: Ind LH / Ind C	0.12*** (0.04)	-0.01 (0.01)	0.05** (0.02)	0.00 (0.01)	0.03 (0.02)	0.02 (0.02)
Mean in Control group	1.61	0.94	0.26	0.04	0.35	0.32
SD in Control group	0.76	0.23	0.44	0.20	0.48	0.47
N	2288	2288	2288	2288	2288	2288
R2	0.19	0.10	0.07	0.05	0.12	0.13
P-values of H0:						
Any Treatment = Control	0.01	0.36	0.11	0.32	0.27	0.33
T1 = T2	0.52	0.08	0.81	0.37	0.91	0.46
T2 = T3	0.13	0.40	0.14	0.40	0.07	0.25

Notes: All specifications include barangay fixed effects along with individual-level baseline controls used for re-randomization (durable assets, value of livestock, number of household members, and number of adults) during random assignment. Standard errors are clustered at the quadrant level, which was the level of random assignment to experimental groups. Statistical significance denoted by \*\*\* for p<0.01, \*\* for p<0.05, and \* for p<0.1

Table 17: Business operations

	(1) Household owns business	(2) Business has operated in the past 30 days	(3) Monthly total sales (pesos)	(4) Monthly total expenses (pesos)
T1: Grp LH / Grp C	0.03 (0.02)	0.01 (0.02)	651 (466)	737** (332)
T2: Ind LH / Grp C	0.04** (0.02)	0.04* (0.02)	201 (375)	187 (242)
T3: Ind LH / Ind C	0.09*** (0.02)	0.06*** (0.02)	709* (401)	494** (237)
Mean in Control group	0.22	0.17	1818	1042
SD in Control group	0.41	0.38	6850	4374
N	2288	2288	2288	2288
R2	0.10	0.07	0.04	0.33
P-values of H0:				
Any Treatment = Control	0.00	0.01	0.26	0.07
T1 = T2	0.51	0.19	0.32	0.09
T2 = T3	0.02	0.16	0.18	0.18

Notes: All specifications include barangay fixed effects along with individual-level baseline controls used for re-randomization (durable assets, value of livestock, number of household members, and number of adults) during random assignment. Standard errors are clustered at the quadrant level, which was the level of random assignment to experimental groups. See text for details. Statistical significance denoted by \*\*\* for p<0.01, \*\* for p<0.05, and \* for p<0.1

Table 18: Labor supply in past 30 days

	(1) Hours worked by household members (total of columns 2 3 4)	(2) Hours worked in household business (total of all hh members)	(3) Hours worked in livestock activities (total of all hh members)	(4) Hours worked in paid work (total of all hh members)
T1: Grp LH / Grp C	-5.33 (10.96)	3.81 (5.14)	1.98 (2.64)	-11.12 (9.89)
T2: Ind LH / Grp C	6.04 (11.68)	-1.28 (4.68)	8.23*** (2.60)	-0.91 (10.87)
T3: Ind LH / Ind C	14.98 (11.61)	14.04** (5.68)	3.35 (2.37)	-2.40 (9.90)
Mean in Control group	267.45	34.46	33.32	199.67
SD in Control group	197.69	103.85	47.81	170.57
N	2288	2288	2288	2288
R2	0.07	0.06	0.13	0.05
P-values of H0:				
Any Treatment = Control	0.27	0.04	0.02	0.64
T1 = T2	0.29	0.30	0.03	0.32
T2 = T3	0.43	0.01	0.06	0.88

Notes: All specifications include barangay fixed effects along with individual-level baseline controls used for re-randomization (durable assets, value of livestock, number of household members, and number of adults) during random assignment. Standard errors are clustered at the quadrant level, which was the level of random assignment to experimental groups. See text for details. Statistical significance denoted by \*\*\* for p<0.01, \*\* for p<0.05, and \* for p<0.1

When examining labor supply (see Table 18), though there is no significant difference between treatment and control groups in the total number of hours supplied, there is a significant difference in how labor hours were supplied across specific income-generating activities. On average, T2 households contributed 8 more hours per month than the control group in livestock activities, while households in T3 contributed 14 more hours per month in household business activities.

### Subjective well-being

The final key outcome for analysis is the subjective well-being of each household, which is calculated as the standardized sum of two separate measures. The first measure is a six-item Kessler Psychological Distress Scale, which asks about the frequency of psychological distress over a four-week recall period. The typical Kessler scale scores the frequency of distress in increasing order, meaning higher scores reflect higher levels of psychological distress in individuals. The version used in this study reverses this scoring system so that higher scores reflect lower levels of psychological distress. The second score is the average of a four-item measure of current and future relative life satisfaction, using Cantril's ladder as a scale, with a score of 10 reflecting the highest level of life satisfaction and a score of 1 reflecting the lowest level of life satisfaction.

Table 19: Subjective well-being

	(1) Subjective well-being index	(2) Reversed Kessler mental health score (out of 5)	(3) Average Cantril's ladder life (out of 10)
T1: Grp LH / Grp C	0.05 (0.05)	0.03 (0.19)	0.06 (0.08)
T2: Ind LH / Grp C	0.13*** (0.04)	0.30 (0.19)	0.19** (0.07)
T3: Ind LH / Ind C	0.08* (0.04)	-0.02 (0.20)	0.14* (0.07)
Mean in Control group	-0.00	24.93	6.13
SD in Control group	1.00	3.78	1.57
N	2288	2288	2288
R2	0.10	0.09	0.10
P-values of H0:			
Any Treatment = Control	0.03	0.37	0.06
T1 = T2	0.07	0.19	0.11
T2 = T3	0.19	0.14	0.48

Notes: Column 1, standardized aggregate score of Kessler and Cantril's ladder life measures. Column 2, aggregate measure of the Kessler Psychological Distress Scale. The score was reversed so that a higher score signals more positive outcomes. Kessler scales range from 10–50, and the scores were reversed so that a higher number reflects lower psychological distress. Column 3, averaged aggregate measure of relative life satisfaction and future relative life satisfaction using Cantril's ladder as the scale, with 10 being the best outcome and 1 being the worst. All specifications include barangay fixed

effects along with individual-level baseline controls used for re-randomization (durable assets, value of livestock, number of household members, and number of adults) during random assignment. Standard errors are clustered at the quadrant level, which was the level of random assignment to experimental groups. See text for details. Statistical significance denoted by \*\*\* for p<0.01, \*\* for p<0.05, and \* for p<0.1

Table 19 shows that households in T2 and T3 reported significant increases in subjective well-being over the control group by 0.13 and 0.08 standard deviations, respectively. This is driven by significantly higher levels of life satisfaction, with T2 and T3 households scoring 0.19 and 0.14 points higher on the averaged Cantril's ladder life satisfaction measure than the control group.

Examining the components of the Kessler scale (see Table 20) shows that there are modest differences in the distribution of impacts on specific items. For example, T1 and T3 respondents feel significantly more restless than the control group by 0.04 to 0.06 points, while T2 and T3 respondents feel significantly less worthless than the control group by 0.01 to 0.02 points. Respondent from T2 households also feel significantly less depressed than the control group, by 0.03 points.

The Cantril's ladder life satisfaction measure used in Table 19 averages responses to four different questions about current and future life satisfaction, which Table 21 disaggregates. The impacts of each program version are relatively homogenous across life satisfaction and future life satisfaction (Columns 1 and 2), although only T3 is statistically significant (at the 10-percent level). Only T2 members report an increase in relative economic status, reporting a 0.20-point increase compared to a control-group average of 4.83, which is statistically significant at the 1-percent level. Future relative economic status is more uniform across treatment arms, though increases are greater for T2 (0.21 points) and T3 (0.16 points) relative to T1 (0.07 points).

Table 20: Kessler measure components 30 days (higher score = less distress)

	(1) Feel nervous	(2) Feel hopeless	(3) Feel restless	(4) Feel depressed	(5) Feel difficult	(6) Feel worthless
T1: Grp LH / Grp C	-0.00 (0.02)	0.00 (0.01)	-0.04*** (0.01)	0.01 (0.01)	-0.01 (0.02)	0.01 (0.01)
T2: Ind LH / Grp C	0.02 (0.02)	0.01 (0.01)	-0.02 (0.02)	0.03* (0.01)	0.01 (0.02)	0.01** (0.01)
T3: Ind LH / Ind C	0.00 (0.02)	0.01 (0.01)	-0.06*** (0.02)	-0.01 (0.01)	-0.01 (0.02)	0.02** (0.01)
Mean in Control group	0.90	0.97	0.93	0.93	0.70	0.97
SD in Control group	0.30	0.18	0.26	0.26	0.46	0.16
N	2288	2288	2288	2288	2288	2288
R2	0.02	0.02	0.04	0.04	0.14	0.03
P-values of H0:						
Any Treatment = Control	0.61	0.57	0.00	0.06	0.61	0.08
T1 = T2	0.22	0.47	0.15	0.18	0.26	0.15
T2 = T3	0.34	0.71	0.03	0.01	0.25	0.86

Notes: All specifications include barangay fixed effects along with individual-level baseline controls used for re-randomization (durable assets, value of livestock, number of household members, and number of adults) during random assignment. Standard errors are clustered at the quadrant level, which was the level of random assignment to experimental groups. See text for details. Statistical significance denoted by \*\*\* for p<0.01, \*\* for p<0.05, and \* for p<0.1

Table 21: Cantril's ladder components (higher score = better outcome)

	(1) Life satisfaction (out of 10)	(2) Future life satisfaction (out of 10)	(3) Relative economic status (out of 10)	(4) Future relative economic (out of 10)
T1: Grp LH / Grp C	0.12 (0.12)	0.12 (0.12)	-0.05 (0.07)	0.07 (0.10)
T2: Ind LH / Grp C	0.18 (0.11)	0.16 (0.11)	0.20*** (0.07)	0.21** (0.08)
T3: Ind LH / Ind C	0.22* (0.12)	0.19* (0.11)	-0.02 (0.07)	0.16** (0.08)
Mean in Control group	5.62	7.21	4.83	6.86
SD in Control group	2.27	2.06	1.83	1.92
N	2288	2287	2288	2285
R2	0.06	0.05	0.09	0.09
P-values of H0:				
Any Treatment = Control	0.25	0.33	0.00	0.06
T1 = T2	0.61	0.74	0.00	0.12
T2 = T3	0.68	0.72	0.00	0.50

Notes: All specifications include barangay fixed effects along with individual-level baseline controls used for re-randomization (durable assets, value of livestock, number of household members, and number of adults) during random assignment. Standard errors are clustered at the quadrant level, which was the level of random assignment to experimental groups. See text for details. Statistical significance denoted by \*\*\* for p<0.01, \*\* for p<0.05, and \* for p<0.1

## Secondary outcomes

In addition to findings on the five key welfare outcomes above, there are statistically significant differences between the three treatment groups and the control group in saving and loan behavior, financial health, and children's labor supply outcomes, which the following sub-sections discuss.

### Savings and loans

Members of all three treatment groups were 6 to 15 percentage points more likely to have savings than the control group, which 26% reported having some savings. All three treatment groups also had significantly more total savings than the control group, with an increase of 224–443 pesos (53– 104%) on average. All three treatment groups were also more likely to keep savings with a microfinance institution and other locations and were more likely to regularly save cash in general (see Table 22).

Borrowing is common among households in the study, with 61% of control-group households having taken out at least one loan in the past 12 months. Graduation reduced this likelihood by 2 (T1 and T2)

to 5 percentage points (T3), and the T3 decrease is statistically significant at the 5% level. The unconditional number of loans taken out by T3 was also modestly less (0.09 loans fewer), which is significant at the 10% level. However, the total remaining value of loans outstanding was 42 pesos higher on average for these households relative to the control group (see Table 23).

Table 22: Savings

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Household has savings	Total savings (pesos)	Saving location: home	Saving location: private bank	Saving location: microfinance institution	Saving location: other	Household regularly saves cash
T1: Grp LH / Grp C	0.06** (0.03)	224** (92)	-0.01 (0.02)	0.02* (0.01)	0.02** (0.01)	0.03*** (0.01)	0.04* (0.02)
T2: Ind LH / Grp C	0.11*** (0.02)	238*** (86)	0.05** (0.02)	0.01 (0.01)	0.04** (0.01)	0.02** (0.01)	0.09*** (0.02)
T3: Ind LH / Ind C	0.15*** (0.03)	443*** (101)	0.07*** (0.03)	0.01 (0.01)	0.04*** (0.01)	0.03*** (0.01)	0.13*** (0.02)
Mean in Control group	0.26	425	0.21	0.01	0.03	0.02	0.21
SD in Control group	0.44	1314	0.41	0.10	0.18	0.14	0.41
N	2288	2268	2288	2288	2288	2288	2288
R2	0.08	0.05	0.07	0.02	0.05	0.04	0.08
P-values of H0:							
Any Treatment = Control	0.00	0.00	0.00	0.21	0.01	0.01	0.00
T1 = T2	0.05	0.89	0.01	0.58	0.42	0.35	0.07
T2 = T3	0.10	0.06	0.46	0.91	0.98	0.20	0.12

Notes: All specifications include barangay fixed effects along with individual-level baseline used for re-randomization (durable assets, value of livestock, number of household members, and number of adults) during random assignment. Standard errors are clustered at the quadrant level, which was the level of random assignment to experimental groups. See text for details. Statistical significance denoted by \*\*\* for p<0.01, \*\* for p<0.05, and \* for p<0.1.

Table 23: Loans

	(1)	(2)	(3)	(4)	(5)
	Household has taken out loans 12 months	Number of loans taken 12 months	Total outstanding value of loans taken adjusted for 30 days (pesos)	Loans taken out up to date on payments	Loans given out up to date on payments
T1: Grp LH / Grp C	-0.02 (0.02)	-0.01 (0.04)	26 (20)	0.00 (0.00)	-0.00 (0.00)
T2: Ind LH / Grp C	-0.02 (0.02)	-0.06 (0.04)	-1 (19)	-0.00 (0.00)	0.00 (0.00)
T3: Ind LH / Ind C	-0.05** (0.02)	-0.09* (0.05)	42* (25)	-0.00 (0.00)	-0.00 (0.00)
Mean in Control group	0.61	0.88	148	0.07	0.01
SD in Control group	0.49	1.00	374	0.09	0.06
N	2288	2288	2288	2288	2288
R2	0.09	0.10	0.04	0.07	0.04
P-values of H0:					
Any Treatment = Control	0.11	0.19	0.20	0.80	0.77
T1 = T2	0.73	0.29	0.17	0.57	0.47

T2 = T3	0.08	0.51	0.08	0.62	0.31
---------	------	------	------	------	------

Notes: All specifications include barangay fixed effects along with individual-level baseline used for re-randomization (durable assets, value of livestock, number of household members, and number of adults) during random assignment. Standard errors are clustered at the quadrant level, which was the level of random assignment to experimental groups. See text for details. Statistical significance denoted by \*\*\* for p<0.01, \*\* for p<0.05, and \* for p<0.1

### Financial health

The analysis reveals several significant impacts in terms of financial security, particularly for households who received individual livelihoods. When asked what source they would use to come up with 5,000 pesos of emergency funds in 30 days, households in T2 and T3 were more likely to cite savings as a source of money (3 to 5 percentage points more likely than the control group). These two treatment groups were also significantly less likely to cite family or friends as a source of emergency money (6 to 9 percentage points less likely than the control group). When asked how difficult it would be to come up with 5,000 pesos in the next 30 days, households in T2 and T3 reported less difficulty in coming up with the money than control households. This was also true for T3 in terms of coming up with the same amount of cash within 7 days.

Table 24: Financial health

	(1)	(2) Source of 5000 pesos emergency money 30 days: savings	(3) Source of 5000 pesos emergency money 30 days: family or friends	(4) Source of 5000 pesos emergency money 30 days: salary	(5) Source of 5000 pesos emergency money 30 days: borrow	(6) Difficulty coming up with 5000 30 days (out of 3)	(7) Difficulty coming up with 5000 7 days (out of 3)	(8) Experienced income loss due to COVID restrictions
T1: Grp LH / Grp C	0.02 (0.02)	-0.04 (0.03)	-0.01 (0.02)	0.02 (0.02)	-0.00 (0.01)	-0.03 (0.02)	-0.02 (0.02)	-0.02 (0.03)
T2: Ind LH / Grp C	0.05*** (0.02)	-0.06** (0.03)	0.00 (0.02)	-0.01 (0.02)	0.00 (0.01)	-0.06*** (0.02)	-0.05*** (0.02)	0.01 (0.02)
T3: Ind LH / Ind C	0.03** (0.02)	-0.09*** (0.03)	0.04* (0.02)	-0.00 (0.02)	0.02 (0.01)	-0.05* (0.03)	-0.00 (0.01)	-0.02 (0.02)
Mean in Control group	0.09	0.47	0.17	0.17	0.07	2.62	2.95	0.66
SD in Control group	0.28	0.50	0.37	0.37	0.26	0.53	0.24	0.47
N	2288	2288	2288	2288	2288	2232	2232	2288
R2	0.03	0.06	0.05	0.07	0.03	0.04	0.02	0.10
P-values of H0:								
Any Treatment = Control	0.05	0.01	0.05	0.48	0.49	0.06	0.01	0.44
T1 = T2	0.17	0.52	0.51	0.14	0.79	0.17	0.13	0.23
T2 = T3	0.37	0.26	0.11	0.81	0.26	0.62	0.01	0.13

Notes: All specifications include barangay fixed effects along with individual-level baseline used for re-randomization (durable assets, value of livestock, number of household members, and number of adults) during random assignment. Standard errors are clustered at the quadrant level, which was the level of random assignment to experimental groups. See text for details. Statistical significance denoted by \*\*\* for p<0.01, \*\* for p<0.05, and \* for p<0.1

### Child welfare

Child labor in the context of this study uses a definition set forth by DOLE as “any work or economic activity performed by a child that subjects him/her to any form of exploitation or is harmful to his/her

health and safety or physical, mental or psychosocial development." The term "child" for this module is defined as household members aged 10 to 17 years old. A child may also be considered economically active under the UN System of National Accounts definition of economic activity, which is "all production that could be destined for the market, regardless of whether the decision is made to sell or retained for own use" (Edmonds and Theoharides 2020). The survey asks about whether children participated in any economic activity over the past 7 days across a range of categories and asks about two types of work that would be classified as child labor: whether children were carrying heavy loads and whether they were injured or unwell from work.

Table 25 shows that on average, households in all three treatment arms were more likely to have a child in the household participating in economic activities by 3 to 7 percentage points, and of children participating rose by 0.06 to 0.13 children relative to the control group. Children were also more likely to have worked in agriculture and business in the past seven days in all three treatment groups, and they were more likely to have worked in livestock in T1 and T2. Furthermore, children in T2 households were more likely to have worked in casual and domestic economic activity as well, while children in households in T3 households were less likely to work in fishing.

In terms of hazardous labor, there were no significant differences between treatment and control groups with children working becoming sick or injured from work, but children were 2 to 3 percentage points more likely to carry heavy loads at work in all three treatment groups. Children in T2 and T3 also spent 2.5 more hours per week taking care of siblings or doing household chores. It is also important to distinguish between "child labor," as defined by DOLE, and whether children are economically active. Like Edmonds and Theoharides (2020), this study finds an increase in economic activity among children, which may result from households needing to fill labor supply gaps to operate family businesses. However, the study does not find an increase in "child labor" along the reduced set of dimensions measured.

Table 25: Child welfare

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Household has members aged 4 to 17	Number of days children spend on school activities 7 days	Household has child that is working	Number of children working	Children working in agriculture 7 days	Children working in livestock 7 days	Children working in fishing 7 days
T1: Grp LH / Grp C	0.00 (0.01)	0.13 (0.13)	0.05** (0.02)	0.09** (0.04)	0.04*** (0.01)	0.04*** (0.02)	-0.00 (0.00)
T2: Ind LH / Grp C	0.03** (0.01)	0.22* (0.12)	0.07*** (0.02)	0.13*** (0.04)	0.03*** (0.01)	0.03** (0.02)	-0.00 (0.00)
T3: Ind LH / Ind C	0.00 (0.01)	0.15 (0.12)	0.03* (0.02)	0.06* (0.04)	0.03*** (0.01)	0.02 (0.02)	-0.01** (0.00)
Mean in Control group	0.89	4.75	0.12	0.24	0.03	0.07	0.01
SD in Control group	0.32	2.33	0.33	0.65	0.16	0.25	0.10
N	2288	2288	2288	2288	2288	2288	2288

R2	0.12	0.09	0.11	0.11	0.07	0.11	0.02
P-values of H0:							
Any Treatment = Control	0.06	0.34	0.01	0.01	0.00	0.05	0.10
T1 = T2	0.03	0.42	0.35	0.35	0.87	0.58	0.80
T2 = T3	0.03	0.50	0.10	0.10	0.52	0.42	0.09

Table 25: Child welfare (cont.)

	(8) Children working in business 7 days	(9) Children working in casual labor 7 days	(10) Children working in domestic labor 7 days	(11) Children carrying heavy load at work 7 days	(12) Children injured unwell from work	(13) Total amount children earn from labor 7 days (pesos)	(14) Number of hours children spend doing housework 7 days
T1: Grp LH / Grp C	0.02** (0.01)	0.00 (0.01)	0.01 (0.01)	0.03*** (0.01)	0.00 (0.01)	12* (6)	0.98 (0.90)
T2: Ind LH / Grp C	0.03*** (0.01)	0.03*** (0.01)	0.02** (0.01)	0.03*** (0.01)	0.01 (0.01)	18*** (7)	2.45** (1.11)
T3: Ind LH / Ind C	0.02** (0.01)	0.00 (0.01)	0.01 (0.01)	0.02* (0.01)	-0.00 (0.01)	8 (7)	2.56*** (0.94)
Mean in Control group	0.01	0.03	0.01	0.02	0.02	21	9.75
SD in Control group	0.10	0.17	0.12	0.14	0.12	116	18.74
N	2288	2288	2288	2288	2288	2288	2288
R2	0.03	0.03	0.02	0.05	0.01	0.03	0.15
P-values of H0:							
Any Treatment = Control	0.00	0.05	0.14	0.00	0.28	0.04	0.03
T1 = T2	0.15	0.03	0.19	0.94	0.95	0.44	0.13
T2 = T3	0.18	0.03	0.14	0.12	0.08	0.23	0.92

Notes: All specifications include barangay fixed effects along with individual-level baseline used for re-randomization (durable assets, value of livestock, number of household members, and number of adults) during random assignment. Standard errors are clustered at the quadrant level, which was the level of random assignment to experimental groups. See text for details. Statistical significance denoted by \*\*\* for p<0.01, \*\* for p<0.05, and \* for p<0.1

## Social networks

In examining program effects on households' social networks, there are increases in the likelihood that T1 and T3 arms have membership in any type of group by 6 and 4 percentage points. This increase for T1 households is driven by increased membership in cooperatives and finance groups, although this modest impact is surprising given that the control-group membership rates are low (2% and 7%, respectively) and participating a group livelihood would in theory be perceived as membership in a cooperative. This may have resulted from respondents' interpretation of the survey question<sup>17</sup> or their degree of involvement in their livelihood group. There is also no evidence of change in the

<sup>17</sup> Question wording: "Are you or members of your households part of any of the following types of groups or associations?" Answer choices included the following: 1. Civic organizations such as senior citizen's groups, women's associations, recreational groups (dancing, sports club) 2. Cooperatives, irrigator's associations, natural resource management groups 3. Credit/finance groups, such as credit cooperative or savings groups 4. Religious groups such as the Barangay Pastoral Council or church associations 5. Governmental group or institution or purok/barangay organization 6. Other(s)

strength of social networks or ties to the community, as the coefficients on each treatment arm are close to zero and not statistically significant.

Table 26: Social networks

	(1)	(2)	(3)	(4)	(5)	(6)
	Group membership in any group	Group membership: civic organization	Group membership: cooperative	Group membership: finance group	Group membership: religious group	Group membership: governmental group
T1: Grp LH / Grp C	0.06*** (0.02)	-0.01 (0.01)	0.03** (0.01)	0.03* (0.02)	-0.00 (0.00)	0.01 (0.01)
T2: Ind LH / Grp C	0.02 (0.02)	0.01 (0.01)	-0.00 (0.01)	-0.00 (0.01)	0.01 (0.01)	0.02* (0.01)
T3: Ind LH / Ind C	0.04** (0.02)	-0.00 (0.01)	-0.00 (0.01)	0.04** (0.02)	0.00 (0.00)	0.01 (0.01)
Mean in Control group	0.23	0.04	0.02	0.07	0.01	0.07
SD in Control group	0.42	0.20	0.15	0.25	0.10	0.26
N	2288	2288	2288	2288	2288	2288
R2	0.12	0.04	0.04	0.08	0.02	0.21
P-values of H0:						
Any Treatment =						
Control	0.03	0.55	0.11	0.05	0.41	0.33
T1 = T2	0.08	0.15	0.02	0.09	0.18	0.75
T2 = T3	0.23	0.48	0.96	0.02	0.81	0.45

Table 26: Social networks (cont.)

	(7)	(8)	(9)	(10)	(11)	(12)
	Group membership: other	Average frequency of social network group visits	Number of households respondent knows	Number of households respondent would go to for advice	Number of households respondent has given or take a loan from	Number of households respondent would consider investing with
T1: Grp LH / Grp C	0.01 (0.01)	-0.10 (0.16)	0.16 (0.34)	0.03 (0.16)	-0.02 (0.02)	-0.05 (0.17)
T2: Ind LH / Grp C	-0.01 (0.01)	-0.09 (0.16)	-0.01 (0.30)	0.02 (0.14)	0.00 (0.02)	-0.03 (0.17)
T3: Ind LH / Ind C	0.00 (0.01)	-0.01 (0.15)	0.32 (0.33)	-0.03 (0.15)	-0.03 (0.02)	-0.31* (0.16)
Mean in Control group	0.02	3.39	5.75	1.77	0.09	1.52
SD in Control group	0.13	1.64	6.26	3.42	0.45	3.62
N	2288	587	2288	2288	2288	2288
R2	0.04	0.26	0.24	0.07	0.02	0.05
P-values of H0:						
Any Treatment =						
Control	0.30	0.87	0.68	0.97	0.16	0.08
T1 = T2	0.06	0.98	0.57	0.90	0.09	0.91
T2 = T3	0.22	0.55	0.26	0.70	0.07	0.04

Notes: All specifications include barangay fixed effects along with individual-level baseline used for re-randomization (durable assets, value of livestock, number of household members, and number of adults) during random assignment. Standard errors are clustered at the quadrant level, which was the level of random assignment to experimental groups. See text for details. Statistical significance denoted by \*\*\* for p<0.01, \*\* for p<0.05, and \* for p<0.1

### Sub-treatment analysis

In order to explore the relationship between livelihood selection and individual livelihood implementation, Table 27 presents a sub-treatment analysis of livelihood outcomes based on treatment group assignment, considering the four most popular livelihoods: swine fattening, nego-cart business, meat processing, and free-range chicken, along with a separate category pooling all other selections. The top four livelihoods comprise 90% of all livelihoods received by participants across the three treatment arms (see Table 6).

Table 27: Sub-treatment analysis by livelihood choice

	Control	T1 (GrpLH/ GrpC)	T2 (IndLH/ GrpC)	T3 (IndLH/ IndC)
Panel A. Overall	575	574	571	568
Ever attended livelihood trainings conducted through BRAC	4%	68%	76%	80%
Ever managed a BRAC livelihood	1%	50%	62%	68%
Currently has a livelihood set up by BRAC	0%	24%	29%	33%
Panel B. By livelihood choice				
Swine fattening	0	206	346	175
Ever attended livelihood trainings conducted through BRAC		83%	90%	90%
Ever managed a BRAC livelihood		69%	81%	79%
Currently has a livelihood set up by BRAC		36%	38%	37%
Nego cart	0	88	56	101
Ever attended livelihood trainings conducted through BRAC		66%	79%	94%
Ever managed a BRAC livelihood		49%	54%	83%
Currently has a livelihood set up by BRAC		20%	16%	34%
Meat processing	0	76	60	69
Ever attended livelihood trainings conducted through BRAC		68%	65%	93%
Ever managed a BRAC livelihood		51%	47%	83%
Currently has a livelihood set up by BRAC		13%	18%	41%
Free range chicken	0	0	0	108
Ever attended livelihood trainings conducted through BRAC				86%
Ever managed a BRAC livelihood				80%
Currently has a livelihood set up by BRAC				45%
Others	0	98	19	28
Ever attended livelihood trainings conducted through BRAC		89%	89%	89%
Ever managed a BRAC livelihood		61%	79%	64%
Currently has a livelihood set up by BRAC		37%	58%	46%
Panel B Total	0	468	481	481

Note: Livelihood selection figures taken from BRAC's monitoring preload data and may exclude those who initially declined to participate in the program.

Across the board, T3-assigned participants have the highest rates of attending livelihood trainings, while households in T1 were the least likely to report “yes” on questions about training participation, livelihood management, and currently having their BRAC livelihood. In the case of swine fattening, training attendance was relatively equal across treatment groups, with only modestly lower training attendance rates among T1 households. However, there are large gaps between T3 and T1 or T2 for those who selected nego cart or meat processing. Specifically, 93–94% of T3 recipients attended training, 83% ever managed a livelihood, and 34–41% were still operating the livelihood. In contrast, 64–79% attended training, 47–53% ever managed the livelihood, and only 13–20% still had their nego cart or meat processing, livelihood among those assigned to T2 and T3.

There are two key takeaways from this table: first, it highlights that conditional on livelihood selection, livelihood participation rates differed substantially between treatment arms, with T3 recipients experiencing higher engagement rates in terms of training, managing, and currently owning a BRAC livelihood in two key livelihoods: nego cart and meat processing. However, these differences are not uniform, with nearly identical rates of currently implementing swine fattening across treatment arms. Second, conditional on treatment assignment, the rates of participation across livelihood vary substantially among those assigned to T1 and T2. Taken together, they indicate that livelihood selection alone is unlikely to solely drive T2 vs T3 differences, but it may remain a factor. In particular, there appear to be important interactions between livelihood selection and treatment type.

## 6. Cost-effectiveness

### Cost breakdowns

Table 28 shows total expenditures by expense type, using information provided by BRAC and ADB on the implementation cost of the graduation pilot in the Philippines. The implementation of program materials, including the delivery of assets and life-skills coaching material development, account for nearly half (40.8%) of total program expenses. Staff costs, including BRAC coaches' salaries and allowances, along with other program administration costs, represent 48.2% of total costs. The remaining 11% reflects trainings, staff transportation, communications, office operations, studies, surveys and reproduction, and monitoring equipment. Including all cost items, total cost per beneficiary is estimated to be USD 541. Although all participants also received support through 4Ps, these costs are excluded because they reflect “business as usual” expenditures. That is, control group members also received this support, and the measured program impacts reflect the benefits of the additional program elements only.

Table 28: Total program costs (in USD)

Project costs	USD	% of total cost
<b>Program administration and staff costs</b> (GCFs salaries and allowance, program coordination, remuneration, and out-of-pocket expenses for 8 municipal links)	457,247	48.2%
<b>Training</b> (seminars, workshops, other training)	40,755	4.3%
<b>Transportation</b> (international air travel national air travel, land transport and vehicle)	32,156	3.4%
<b>Office operations and communications</b>	5,550	0.6%
<b>Implementation and program materials</b> (assets, life-skill coaching material development, artist contracts, printing)	386,660	40.8%
<b>Studies, Surveys and Reproduction</b>	19,202	2.0%
<b>Equipment for monitoring and other activities</b>	6,355	0.7%
<b>Total program cost</b>	947,925	100%

<b>Number of beneficiaries</b>	<b>1,753</b>
<b>Cost per beneficiary</b>	<b>541</b>

Note: Program expenses in USD were calculated using the exchange rate used in the BRAC-ADB contract, which is 1 USD = 50.02 pesos. The total cost assumes an attrition rate of 30% for the purposes of computing asset costs, so only 1,239 households are assumed to receive a productive asset.

Breaking these figures down by treatment arm in Table 29 shows the relative cost-effectiveness of different program variations: individual vs. group coaching and individual vs. group livelihoods. Although IPA was not able to obtain program costs broken down by treatment type directly from BRAC, the research team was able to derive cost estimates using monitoring data on GCF productivity during the coaching period and cost information provided by ADB.

Before program implementation started, BRAC estimated that group-level coaching sessions could reduce total program costs by 30-40% through increased coaching caseloads without decreasing the number of community visits that each coach completes. Cost analysis by treatment arm after program implementation shows that group coaching yields cost savings of 27–32%. Between the two treatment arms that received group coaching, households that also received group livelihoods show a total cost saving of 7% compared to households that received individual livelihoods, although this reflects higher program attrition in the group livelihood arm leading to fewer delivered assets.

Table 29: Estimated costs by treatment arm (in USD)

	T1 (GrpLH/GrpC)	T2 (IndLH/GrpC)	T3 (IndLH/IndC)	Total
N	362	423	454	1239
ITT N	587	583	583	1753
Administrative costs	121,588	121,588	121,588	364,765
Assets	108,600	126,900	136,200	371,700
Coaching	38,798	38,533	134,129	211,460
Total sum	268,986	287,022	391,917	947,925
Cost per household	743	679	863	765
Cost savings vs. T3	-14%	-21%		
ITT cost per household	458	492	672	541
Cost savings vs. T3	-32%	-27%		

Table 29 estimates are based on calculating costs either (a) per household that remained in the program (N) (or (b) per household that was initially offered the program (ITT N). For the purposes of estimating cost-effectiveness, the ITT N is aligned with the estimated program benefits, which are calculated as an average benefit per household offered the graduation program. Administrative costs are divided equally between the 3 treatment groups, reflecting that they are not affected by program variation, and asset costs differ based only on the number of assets that were actually delivered.

The primary driver in differential costs per treatment arm is the cost of coaching. Estimated coach productivity is 28.8 beneficiaries served per day with group coaching and 5.3 beneficiaries served per day with individual coaching. These productivity estimates are based on average sessions held per coach per day, restricting to days in which coaches either only conducted group sessions (average of 1.4/day) or only conducted individual sessions (average of 5.3/day). This leads to a per-participant group cost of \$4.13/month, or \$52.78 over the estimated 16 months of coaching. Assuming that coaches hold 100% of intended individual sessions, the per-participant individual coaching cost is \$14.38, or \$230.07 over the estimated 16 months of coaching.

### Cost-benefit comparisons

The program's return on investment is the value of benefits incurred through the program relative to its costs. This estimation reflects the cost calculations above, using ITT cost per household, and the estimated ITT treatment effect on consumption as the program's benefits, reflecting Banerjee et al. (2015).<sup>18</sup> The return on investment reflects the average net present value of consumption benefits under certain assumptions about the continued impacts of consumption benefits over time. Scaling monthly consumption effects by the sample's average household size and multiplying by 12 leads to an estimate of yearly household consumption benefit. Assuming an annual social discount rate of  $d$ ,

<sup>18</sup> The estimated value of consumption differs slightly from Banerjee et al. (2015), as this study incorporates the value of durable assets as a continuous share of consumption rather than adding it as a one-time factor.

which captures future benefit loss due to cost of capital, opportunity cost, and uncertainty, the net present value of consumption can be calculated using the following formula, starting with  $B$  in year 1.

$$B = (1 - d)^2 * b * (1/(1 - k))$$

The second-year endline consumption treatment effect is denoted by  $b$ , and  $k$  is equal to  $p * (1 - d)$  where  $p$  is the annual rate of persistence of the consumption benefit. Table 30 presents return on investment estimates for each treatment arm under various social discount rates and levels of persistence.

For all treatment arms, return rates are very high using standard social discount rates of 5 or 7 percent at 100% persistence, ranging from 798%–1159% at a 5% social discount rate and 528%–780% at a 7% social discount rate. Reflecting substantially higher costs of individual coaching without greater consumption impacts, T1 and T2 are the most cost-effective, and the estimated ROI is relatively similar. Return rates decrease substantially with lower persistence levels, although they only approach zero for T1 and T2 at 40% persistence, at which point T3 has a benefits-to-cost ratio of –28%.

These results compare very favorably to impact evaluations of other graduation programs (Banerjee et al. 2015), which have an estimated benefits-to-cost ratio of 133%–433% (excluding Honduras, which had a –198% ratio) when assuming 100% persistence. These studies see a full dissipation of benefits, such that costs fully equal benefits, with persistence rates of 68.9–98.2%, much higher than in this study.

Table 30: Return on investment under different cost and benefit assumptions

	T1 (GrpLH/GrpC)	T2 (IndLH/GrpC)	T3 (IndLH/IndC)
Year 3 benefit estimate (USD)	296	326	317
Cost (USD)	458	492	672
<b>ROI</b>			
d=5%			
Persistence=			
100%	1126%	1159%	798%
80%	155%	162%	87%
40%	-1%	2%	-28%
d=7%			
Persistence=			
100%	757%	780%	528%
80%	134%	141%	72%
40%	-4%	-2%	-30%

## 7. Conclusions and policy implications

The graduation program is designed to assist ultra-poor households facing an interrelated set of challenges that keep them in a poverty trap, which includes a lack of cash for short-term consumption, lack of capital, skills, work opportunities, financial education, and confidence<sup>19</sup>. The centerpiece of the program is the transfer of a productive asset and the training to manage it, and the program also includes a holistic package of interventions including consumption support (like regular cash transfers), life coaching, skills training, access to health information, and components of financial inclusion.

The graduation pilot in the Philippines is the first RCT to measure differences in outcomes among individual and group-based variations of two key components of BRAC's graduation program: i) asset transfers and ii) coaching. These variations are aligned with DOLE's interests to maximize the effectiveness of the *Kabuhayan* through policy revisions. The graduation pilot started in October 2018 with the implementation of the individual and group life skills coaching; asset deliveries for individual livelihoods started in June 2019 and ended in January 2020; and asset deliveries for group livelihoods started in October 2019 and ended in March 2020.

Between September and December 2021, IPA conducted an endline data collection to survey the households of the graduation pilot. The pilot program faced many implementation challenges, such as delays in the transfer of productive assets to group livelihood participants and the COVID-19 pandemic. Despite these problems, the graduation approach had a significant positive impact on all treatment arms in terms of monthly consumption per capita, food security, the value of productive assets, and savings, relative to a control group that received only consumption support through the government conditional cash transfer program (4Ps). This outcome is in line with findings in several other graduation programs evaluated in a variety of settings. In particular, the endline survey shows the following results:

- Among those who received any form of the graduation program, monthly consumption per capita is between 311–342 pesos greater than households assigned to the control group.
- Food security also increased among all graduation groups, with households scoring 0.20–0.28 standard deviations higher than control households on this index.
- Assignment to any graduation treatment arm increased the average value of productive assets, with treatment households owning 1,954–4,484 pesos worth more productive assets than control households.

---

<sup>19</sup> In this context, not all beneficiaries in this pilot necessarily qualified as ultra-poor.

- In terms of subjective well-being, only households who received the individual livelihood have a statistically significant improvement (0.08–0.13 s.d.) in this index relative to the control group.
- There was no evidence that monthly income changed in any of the graduation arms compared to the control group, in part due to the high dispersion observed in income.
- Modest effects can be seen in the likelihood of group membership among households in treatment groups 1 and 3, although these effects are surprisingly low for T1 members given that they were offered a group livelihood and group coaching.
- There was a significant increase in the likelihood that treatment households had any savings over the control group by 6–15 percentage points as well as having significantly more savings (between 224 and 443 pesos). Treatment households were also significantly more likely to keep their savings in microfinance institutions by 2–4 percentage points.

These findings are largely in line with other programs that were rigorously evaluated in Banerjee et al. 2015, in which households saw significant increases in household consumption, food security, and the value of productive assets, and they saw modest increases in subjective well-being.

Because costs and administrative burdens across treatment arms differ substantially, understanding how impacts differ is important to offer policy recommendations. Among the key outcomes, the research team cannot reject a null hypothesis of identical treatment impacts across arms, except in the case of T2 outperforming T1 on the value of productive assets value and subjective well-being, for which the differences are statistically significant at the 10-percent level. However, these differences are in line with a general pattern that the individual livelihood/group coaching (T2) and individual livelihood/individual coaching (T3) treatments consistently produce better outcomes than the group livelihood/group coaching (T1) arm.

However, these results should be interpreted with some caution. The measured effectiveness of the group livelihood arm (T1), reflecting intention-to-treat effects, was hampered by relatively lower program take-up. In part, this could reflect hesitation among program participants (reflected in initially higher drop-out rates relative to T3 along with anecdotal reports), which might be inherent to any group livelihood program. However, the drop-out observed in this context was exacerbated by significant delays in the delivery of their productive assets, which were specific to this study. Despite these problems, it is notable that the group livelihood arm still yields overall positive effects relative to the control group.

In a similar vein, the substantial cost savings (27%–32%) associated with group coaching means that even modestly worse impacts relative to individual coaching (comparing T2 vs. T3) could still indicate a path to improved cost-effectiveness. However, analysis shows that the impacts of the group coaching arm are statistically indistinguishable from the individual coaching arm across all major

outcome variables, and group coaching leads to a larger increase in subjective well-being, though the difference is not statistically significant. The main secondary outcomes in which T3 performs better than T2 is in the likelihood of business ownership (9pp vs 4pp increase,  $p = 0.02$ ) and, consequently, in business asset values (1258 vs 597-peso increase,  $p=0.07$ ). This finding is particularly promising for the future of group coaching, but one caveat is that there are substantial differences in the nature of the livelihood selected between the two treatment arms, such that the net effect of the program reflects both differences.

Combining these results with estimated costs in Section 6 demonstrates that all variations of the program are cost-effective with standard or even more conservative estimates of persistence and social discount rates. Specifically, all versions are substantially more cost-effective than the studies included in Banerjee et al. 2015. Taking seriously the magnitudes of consumption estimates, the greater estimated impacts of T2 relative to T1 are accompanied by higher participation rates (and therefore higher asset costs), such that their cost-effectiveness is virtually equal. In comparison, the substantially lower costs of group-level coaching, which reduces total program costs by roughly 30%, means that T3 is substantially less cost-effective, although it still performs well compared to studies in other contexts.

To sum up the results of the study's research questions and outcome measures:

- Does the classic graduation model (individual livelihood + individual coaching) improve welfare for 4Ps households as much as found in other country contexts?
  - The graduation in the pilot in the Philippines has proved effective in improving the welfare of 4Ps households as much as has been found in other countries, with cost-effectiveness that appears to exceed studies conducted by Banerjee et al. (2015).
- Are individual livelihoods more profitable, sustainable, and resilient compared to the control group than group livelihoods?
  - There is evidence that both group and individual livelihoods are profitable, sustainable, and resilient compared to the control group. There is also suggestive evidence that individual livelihoods may be more effective than group livelihoods. However, these comparisons reflect potential differences in the effectiveness of the treatments as well as different rates of attrition and varied livelihood selection across treatment arms.
- Does providing individual livelihood plus individual coaching yield greater impacts in building sustainable livelihoods and reducing poverty compared to the control group and group livelihoods and group coaching, in the context of this DOLE-implemented graduation model?
  - The research team finds that individual livelihoods and individual coaching yield substantial gains in consumption, food security, and value of productive assets relative to the control group. The research team also finds suggestive evidence that they are more effective than group livelihood and group coaching, but these comparisons reflect not only potential

differences in the effectiveness of the treatments, but also different rates of attrition and varied livelihood selection across treatment arms.

- Does receiving group vs. individual coaching affect intermediate outcomes such as program engagement, business formation, and social capital?
  - There is no evidence that group versus individual coaching has any differential effect on program engagement, business formation, or social capital.
- What are the potential trade-offs between cost-effectiveness and impact through group coaching and/or livelihoods?
  - The research team cannot reject equal impacts between individual coaching and group coaching on nearly all dimensions, and on some dimensions, the group coaching arm outperforms the individual coaching arm. This comes alongside significant cost savings through group coaching, indicating that these cost savings are not harming program effectiveness. One caveat is that group-coaching attendance rates were lower overall.
  - Group livelihoods were ultimately cheaper to implement because of higher attrition. However, because of reduced impacts, they are, on average, as cost-effective in terms of consumption impacts as individual livelihoods, holding constant group coaching.
- Are there complementarities between group livelihoods and group coaching that are not present when implementing the program through individual livelihoods or individual coaching?
  - There is no evidence in this study of complementarities between group livelihoods and group coaching versus individual livelihoods and individual coaching.

## 8. References

- 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*, 103(484), 1481-1495.
- Banerjee, A. et al. (2015). A multifaceted program causes lasting progress for the very poor: Evidence from six countries. *Science*, 15 May 2015, v. 348, issue 6236. Available at: <http://science.sciencemag.org/content/348/6236/1260799>.
- de Mel, S., McKenzie, D. J., & Woodruff, C. (2009). Measuring microenterprise profits: Must we ask how the sausage is made? *Journal of Development Economics*, 88(1), 19-31.
- Edmonds, E. & Theoharides, C. (2020). The short-term impact of a productive asset transfer in families with child labor: Experimental evidence from the Philippines. *Journal of Development Economics*, September 2020, v. 146, Available at:

<https://www.sciencedirect.com/science/article/abs/pii/S0304387820300614> .

- Gollin, D., & Udry, C. (2021). Heterogeneity, measurement error, and misallocation: Evidence from African agriculture. *Journal of Political Economy*, 129(1), 1-80.
- Innovations for Poverty Action (2020). Philippines: COVID-19 socio-economic impact of the poor households in Negros Occidental. Report prepared for the Asian Development Bank.
- Innovations for Poverty Action (2018). Impact Evaluation of the Graduation of the Ultra-Poor Pilot in the Philippines. Report prepared for the Asian Development Bank.

## Appendix

### 1. Summary statistics on key outcome components

#### Consumption

Table 31: 7-day consumption items in pesos

How much in total did your household consume in the past week?  
 How much would the amount you consumed cost if purchased?

7 Day Consumption Food Items	Mean	Std. Dev.	N
Rice	553	256	2,288
Coconut	19	32	2,288
Banana	88	97	2,288
Meat excluding poultry and seafood	155	180	2,288
Poultry	183	132	2,288
Milk and formula	107	126	2,288
Other dairy such as cheese and yoghurt	11	30	2,288
Eggs	112	81	2,288
Fish and Seafood	335	249	2,288
Corn	12	30	2,288
Bread and cereals	75	69	2,288
Beans and nuts	8	21	2,288
Casava	11	26	2,288
Potatoes	13	28	2,288
Camote	15	29	2,288
Yams	3	16	2,288

Other Starches such as squash and carrots	29	39	2,288
Other Fruits	2	22	2,288
Vegetables	76	78	2,288
Coffee, tea and Milo	72	64	2,288
Toyo, vinegar, and other condiments and seasonings	67	56	2,288
Oils and fats	33	18	2,288
Sugar, jam and honey	32	21	2,288
Snacks, such as cookies, biscuits, chips, candy, ice cream, suman	79	92	2,288
Water	52	57	2,288
Other non-alcoholic beverages	53	65	2,288
Alcoholic beverages	37	75	2,288
Cigarettes	67	135	2,288
Meals outside the household	26	76	2,288
Any other food or drink not mentioned before	2	19	2,288

Table 32: 30-day consumption items in pesos

In the past 30 days, how much did the household spend on the following items?

30 Day Consumption Items	Mean	Std. Dev.	N
Soaps	129	91	2,288
Cosmetics	50	98	2,288
Hair products	125	90	2,288
Detergents	178	137	2,288
Toothpaste and toothbrushes	103	76	2,288
Phone credit	205	259	2,288
Postal services	35	248	2,288
Hairdressing salons and personal grooming establishments	6	26	2,288
Electric appliances for personal care	0	0	2,288
Other appliances, articles and products for personal care	1	10	2,288
Articles for babies	61	176	2,288
Expenses for the preparation of personal transport equipment	204	513	2,288
Electricity, gas, and other fuels for the home	571	548	2,288
Water supply and miscellaneous services used at home	97	174	2,288
Maintenance and repair of the dwelling	1074	4661	2,288
Medicines and Other Medical Products	374	634	2,288
Outpatient medical services	96	359	2,288
Hospital services	30	322	2,288
Game and entertainment	2	24	2,288
Accommodation Services	0	0	2,288
Transport Services	386	552	2,288

Audio-visual, photography and information processing equipment	0	0	2,288
Newspaper, books, and stationery	79	118	2,288
Do you own the current dwelling in which you reside?	1	0	2,288
How much do you pay in rent to reside at your current dwelling?	44	254	2,288
How much would you earn if you were able to rent out your current dwelling?	1000	983	2,288

Table 33: 12-month consumption items in pesos

In the past 12 months, how much did the household spend on the following items?

12 Month Consumption Items	Mean	Std. Dev.	N
School costs such as tuition and school fees	275	1777	2,288
School costs such as allowance for family members studying away from home	50	351	2,288
Computer rental services and printing services	55	148	2,288
School uniform	3	38	2,288
All other educational expenses apart from tuition fees, and those already listed	134	463	2,288
Garments excluding uniform for adult household members 18 years old and older	593	851	2,288
Garments excluding uniform for children younger than 18 years old	692	894	2,288
Furnishings, carpets and other floor coverings	145	650	2,288
Household textiles	108	254	2,288
Repair of household appliances	26	120	2,288
Glassware, tableware and household utensils	179	331	2,288
Small tools and miscellaneous accessories	11	40	2,288
Goods and services for routine household maintenance	97	608	2,288
Footwear	817	916	2,288
Jewelry, clocks and watches	7	38	2,288
Handbag, travelling bag, wallet, purse, etc.	35	122	2,288
Articles for babies such as baby carriages, etc.	72	203	2,288
Umbrella, sunglasses, fan, etc.	101	217	2,288
Funerary articles such as coffins, gravestones, etc.	1292	6926	2,288
Other personal effects, n.e.c.	7	45	2,288
Insurance such as life and non-life insurance	148	581	2,288
Financial services such as charges from using banks and other financial providers like Western Union	14	138	2,288
Social protection such as Pension, PhilHealth, TWSP, ARCDP, CBEP, Disability Benefit, Scholarship and Student Financial Aid	226	1028	2,288
Church tithes and offerings	205	453	2,288
Taxes	9	91	2,288
Gifts and contributions to others	44	160	2,288

## Assets

Table 34: Value of household assets

If you were to sell all of your named asset how much would you receive in total?

30 Day Consumption Items	Mean	Std. Dev.	N
Refrigerator/Freezer	1,960	4,611	2,288
Bicycle or Pedicab	347	1,935	2,288
Motorcar or Motorcycle	10,302	21,192	2,288
Sewing machine	39	448	2,288
Washing machine	300	1,238	2,288
TV Set	2,347	3,787	2,288
VTR/VHS/VCD/DVD player	340	806	2,288
Radio/Transistor/Stereo	415	1,322	2,288
Electric fan	461	792	2,288
Telephone/Mobile phone	4,034	5,483	2,288
Sala set	723	2,136	2,288
Boat	401	4,626	2,288
Tractor or Power Tiller	0	0	2,288
Chair/Stool	613	1,050	2,288
Bed or cot	719	1,651	2,288
Table	479	948	2,288
Watch or Clock	170	594	2,288
Jewelry	94	1,346	2,288
Gas Stove	285	1,004	2,288
Farm tools and gardening tools (spade, shovel, sprayer, wheelbarrow, rake, hoe, pitchfork)	665	1,208	2,288
Nego-Kart/food cart	60	682	2,288
Storage for business (chest or drum)	16	245	2,288
Kitchen tools for business (meat grinder, frying pan, metal trays with cover, skimmer with handle, tongs, spoons, mixing bowls, measuring spoons, funnel)	118	880	2,288

## Food security

Table 35: 7-day food consumption score

Over the past 7 days, how many days has anyone in your household consumed any food from the following groups?

Food Consumption Score	Mean	Std. Dev.	N
Number of days consumed <b>staples</b> in the past 7 days	6.94	0.55	2,288
Number of days consumed <b>pulses</b> in the past 7 days	0.68	1.33	2,288
Number of days consumed <b>vegetables</b> in the past 7 days	2.93	2.30	2,288

Number of days consumed <b>fruit</b> in the past 7 days	3.59	2.46	2,288
Number of days consumed <b>meat and fish</b> in the past 7 days	4.53	2.32	2,288
Number of days consumed <b>milk</b> in the past 7 days	4.68	2.99	2,288
Number of days consumed <b>sugar</b> in the past 7 days	6.25	1.94	2,288
Number of days consumed <b>oil</b> in the past 7 days	6.24	1.67	2,288

Table 36: 4-week household food insecurity access scale

Household Food Insecurity Access Scale	Mean	Std. Dev.	N
In the past four weeks, did you worry that your household would not have enough food?	1.24	1.05	2,288
In the past four weeks, were you or any household member not able to eat the kinds of foods you preferred because of a lack of resources?	1.07	1.04	2,288
In the past four weeks, did you or any household member have to eat a limited variety of foods due to a lack of resources?	0.93	1.03	2,288
In the past four weeks, did you or any household member have to eat some foods that you really did not want to eat because of a lack of resources to obtain other types of food?	0.79	0.99	2,288
In the past four weeks, did you or any household member have to eat a smaller meal than you felt you needed because there was not enough food?	0.77	0.97	2,288
In the past four weeks, did you or any other household member have to eat fewer meals in a day because there was not enough food?	0.59	0.89	2,288
In the past four weeks, was there ever no food to eat of any kind in your household because of lack of resources to get food?	0.27	0.65	2,288
In the past four weeks, did you or any household member go to sleep at night hungry because there was not enough food?	0.11	0.44	2,288
In the past four weeks, did you or any household member go a whole day and night without eating anything because there was not enough food?	0.05	0.28	2,288

\*Binary variables with 1 = yes and 0 = no

## Mental health

Table 37: 6-item Kessler psychological distress scale

Kessler Psychological Distress Scale	Mean	Std. Dev.	N
About how often during the past 30 days did you feel nervous?	4.17	1.04	2,282
About how often during the past 30 days did you feel hopeless?	4.57	0.74	2,281
About how often during the past 30 days did you feel restless or fidgety?	4.00	1.05	2,286
About how often during the past 30 days did you feel so depressed that nothing could you cheer you up?	4.28	0.98	2,286

About how often during the past 30 days did you feel that everything was difficult?	3.25	1.38	2,283
About how often during the past 30 days did you feel worthless?	4.78	0.59	2,283

\*Based on a 5-point scale, 5 being the highest (most frequent) and 1 being the lowest (least frequent)

Table 38: Cantril's ladder relative and expected life satisfaction

Life Satisfaction	Mean	Std. Dev.	N
How would describe your satisfaction with life? The top rung of the ladder (10) represents very satisfied and the bottom of the ladder (1) represents very dissatisfied. On which step would you place yourself?	5.76	2.24	2,288
On the same ladder, which step do you believe you will be on in 5 years?	7.33	1.96	2,287
Now assume that the top rung of the ladder (10) represents the best-off members of your community and that the lowest rung (1) are the poorest individuals. Where would you place your household on the ladder in terms of economic status?	4.86	1.79	2,288
Where do you think you will be on this ladder 5 years from now in terms of your economic status?	6.97	1.90	2,285

\* Based on a 10-point scale with 10 being the highest score and 1 being the lowest

Table 39: Log of monthly consumption and income and productive asset values

	(1)	(2)	(3)
	Monthly consumption per capita (pesos)	Value of productive assets (pesos)	Monthly income (pesos)
T1: Grp LH / Grp C	0.07*** (0.02)	0.19* (0.10)	0.03 (0.05)
T2: Ind LH / Grp C	0.08*** (0.02)	0.38*** (0.08)	0.08 (0.05)
T3: Ind LH / Ind C	0.08*** (0.02)	0.37*** (0.10)	0.05 (0.05)
Mean in Control group	8.18	8.45	9.19
SD in Control group	0.47	1.68	0.94
N	2287	1691	2246
R2	0.18	0.16	0.05
P-values of H0:			
Any Treatment = Control	0.00	0.00	0.44
T1 = T2	0.86	0.03	0.29
T2 = T3	0.81	0.97	0.53

Notes: All specifications include barangay fixed effects along with individual-level baseline used for re-randomization (durable assets, value of livestock, number of household members, and number of adults) during random assignment. Standard errors are clustered at the quadrant

level, which was the level of random assignment to experimental groups. See text for details.  
 Statistical significance denoted by \*\*\* for  $p<0.01$ , \*\* for  $p<0.05$ , and \* for  $p<0.1$

## 2. Endline survey preparatory activities

### Questionnaire development

The PIs of the project, in close coordination with the Research Associate and the Research Manager, started developing the instrument to be implemented for the endline survey in early July. On July 26, 2021, IPA shared with ADB the first draft version of the questionnaire<sup>20</sup> for comments, feedback, and suggestions. The instrument included the following modules:

- Updated Household Roster
- Food security
- Household assets
- Consumption (7 days, 30 days, and 12 months)
- Household business
- DOLE livelihoods
- Agriculture
- Livestock
- Financial health
- Savings and lending
- Social safety nets
- Social networks
- Child welfare
- Women's empowerment
- Mental health and well-being

Given the average daily productivity observed per enumerator during the implementation of the baseline survey in 2018, as well as the timeframe and budget constraints (which included additional costs to follow safety protocols during the pandemic) to complete the survey, the instrument was designed to be implemented both in person and remotely via phone surveys for a subsample of the program participants, in case they couldn't be reached in person.

The PIs of the study reviewed carefully all the comments and suggestions received from ADB on July 27, 2021, as feedback to improve the questionnaire.

---

<sup>20</sup> This was the first deliverable of the contract signed between IPA and ADB for the endline data collection

## Pilot

The pilot activity was split into two sessions to cover both the in-person and phone survey versions of the instrument, with different versions programmed in SurveyCTO. The In-person survey pilot was conducted on August 28 by a senior field team of two field coordinators and four senior field officers, who interviewed a total of 13 households in Barangay VI-A in the City of Victorias (this barangay was dropped from the original pilot sample after completion of the Graduation Pilot's baseline survey in 2018). Each senior field staff completed at least two surveys and debriefed the research staff team in the afternoon.

On September 04, 2021, the same senior field team piloted the phone survey version of the instrument programmed in SurveyCTO for remote data collection. The sample chosen for the pilot was drawn from the barangay Alegria in the municipality of Murcia (this barangay was also dropped from the original sample after completion of the Graduation Pilot's baseline survey in 2018). The team faced difficulty completing their assigned interviews because some phone numbers from baseline were outdated and there often was poor connectivity quality in the respondents' area. After the activity, another debriefing session was held in the afternoon with the field team.

As part of the preparatory activities before starting the data collection, the senior field team was working in close coordination with barangay officials of target endline survey areas to collect updated phone numbers of the households included in the full sample.

## In-class training

The IPA research team traveled to the City of Talisay to conduct field officer training. The in-class training was held from Wednesday, September 7 through Friday, September 9, while the field practice and the “fake launch” were scheduled for September 14 and September 15.

IPA invited 33 field officer candidates as well as seven members of the senior field team to join the three-day in-class sessions. All attendees were required to submit a negative COVID-19 antigen test to participate. Annex 1 details the training agenda which focused heavily on familiarizing officers with i) the goals of the project and their role as enumerators, ii) the protocols for in-person and remote data collection, iii) the content and the substance of the survey instrument, iv) utilizing the SurveyCTO app in the tablets for data collection.

During the first day of the training, the participants were introduced to the work of IPA, the goals of the Graduation Pilot in the Philippines, and the endline data collection. Enumerators were given a thorough explanation of their roles, responsibilities (and expectations) as well as the protocols and procedures for data collection.

Day two of training focused on familiarizing candidates with the questionnaire itself as well as the protocols of the in-person survey. On day three the enumerators were trained on the protocols for conducting phone surveys as well as interview etiquette. Candidates were evaluated through multiple-choice quizzes at the end of days one and two to measure how well they absorbed information from training. At the end of the training, the research team convened with the senior field team to assess candidates for hiring eligibility.

### **Field training and fake launch**

The field training was held on September 14 in two of the barangays that were dropped from the Graduation baseline survey: Barangay VI-A (City of Victorias) and Barangay Alegria (Murcia). A total of 151 households were used as the training sample (not including the respondents who were interviewed during the pilot). The field training continued with 31 candidates (out of a total of 33) since two candidates dropped out due to health concerns. The results from the field training were assessed by the senior field team and only candidates that were eligible for hire based on their performance were invited to the fake launch.

The fake launch was conducted on September 15 in the same two barangays using the remaining un-interviewed households. Candidates were once again assessed based on their performance and a final hiring decision was made at the end of the fake launch day. The following day, another candidate dropped out for personal reasons, so only one field officer from the remaining candidates was not offered a contract of employment.

### **3. Field data collection**

Data collection was officially launched on September 17 after all field staff were tested for COVID-19 and received negative results. The Research Associate remained in Negros Occidental for the first two weeks of data collection to further help facilitate field activities, monitor the performance of the field team accompanying them in the conduction of the interviews, and assess data quality.

Field officers (FOs) were divided into groups of five, each headed by a senior field officer (SFO). The five teams were further divided into two groups each led by a field coordinator (FC). Data collection for the endline survey was scheduled to end on November 4, 2021. Working with 25 enumerators, data collection took approximately 7 weeks, reaching about 98% of the original 2,339 baseline surveyed households.

### **Data quality monitoring**

To ensure the high quality of the data collected, IPA monitored daily real-time information in a Google tracking sheet and ran high-frequency data checks and backcheck analyses over the course of data

collection. The RA alongside the field supervisors held debriefings every week to provide feedback based on the data quality monitoring.

### Tracking sheet

A Google sheet tracked the daily and overall progress of each field officer. This was linked to the SurveyCTO database, allowing real-time updates on the number of attempts and interview status of each respondent. A summary sheet was also set up to compare the level of effort of each enumerator.

### High-frequency checks

A reliability check was conducted every other day to identify surveyor factors in a variety of performance indicators. These checks involve validation and consistency checks, average survey time by module, value outliers, enumerators who are performing below or above average, and enumerator-specific patterns. The results of these checks are then used as input for the feedback that the enumerators receive during the regular team debriefing sessions.

### Back checks

18% of interviews were randomly backchecked by the field supervisors on a timely basis. The backcheck questionnaire consisted of a short 10-questions mini-survey to check if the interview occurred, and discrepancy rates in key questions such as business ownership, livestock ownership, and plot use.

## 4. Data Collection Relaunch Plan (December 2021)

Immediately after data collection ended during the first week of November, the research team started cleaning the dataset and continued working to identify outliers, missing values, and other data quality issues. After conducting a pre-analysis assessment of the data, the research team noticed some significant gaps between consumption outcomes and income outcomes. While the data showed significant increases in overall consumption in the overall sample, the collected data on income did not sufficiently explain this increase in consumption variables. After some consideration, the research team considered it necessary to relaunch field activities to collect additional data that may explain these differences. A second round of data collection began on November 22 and ended on December 23.

For this extension of the data collection, IPA retained 21 field officers divided into three teams (two teams of seven and eight enumerators, respectively, handling phone interviews and one team of six FOs conducting in-person interviews with households that provided no phone contact information). Enumerators were instructed to attempt all numbers provided for each household at least once and switch cases to in-person interviewing after six failed attempts. These households would then either be assigned to be interviewed by an in-person team, or phone teams would interview them later in

the survey period when all teams would transition to in-person interviewing. This transition occurred on December 8, once enumerators had exhausted all attempts at reaching households by phone.

Using updated contact information from the recent data collection and relying heavily on phone surveying<sup>21</sup> to reach most of the sample, by December 8 the field team was able to reach almost 80% of the first endline survey's sample. Households without phone numbers and those that were determined unreachable by phone were interviewed in person between December 8 and December 23, where enumerators were able to reach almost 99% of the first endline survey's sample.

The instrument employed was a short questionnaire that targets the assumed gaps in household income was administered. This includes the noisy sources of household income such as livestock slaughter information, remittances, paid work, and economic shocks the households may have experienced in the past year.

### **Slaughtered livestock**

A pre-analysis assessment of the data showed an overall negative change in the average quantity of livestock per household with fewer live animals being sold, having died, or being given away as gifts. Since this result might be due to a significant number of households slaughtering their livestock for their own consumption, the extension of the data collection has included key questions to capture income sourced by slaughtered livestock.

### **Sharp decline in paid work**

A second issue found during the pre-assessment of the data collected was an unusually sharp decline in the number of households reporting paid work income sources in the past year as compared to the baseline and phone surveys.

### **Remittances**

Remittances sent to households constitute a potential additional source of household income.

### **Negative shocks**

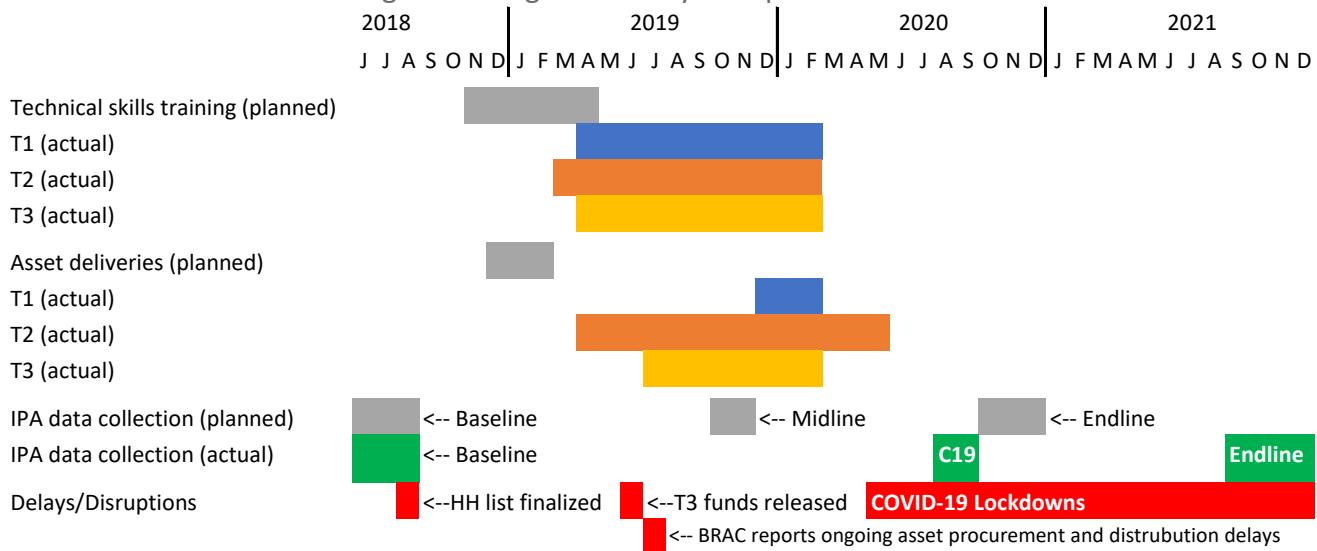
The research team included a set of short questions to understand whether households experience an inability to generate income in the past 12 months as a consequence of an unexpected negative shock, such as a climate event.

---

<sup>21</sup> During the first part of the data collection, only 74 households (out of 2,367) were interviewed by phone. Households were only interviewed by phone if it was extremely difficult to meet the respondent in person, the respondent or a household member appeared to have COVID-19 symptoms, or if the respondent themselves requested to be interviewed by phone.

## 5. Endline survey preparatory activities

Figure 7: Program activity disruptions timeline



Note: "Planned" timelines are based on a 2017 research design document available here:

<https://ipastorage.box.com/s/9xviwn03qqm0u778p0bjyet6e5pg9us3>