GHANA LEAP 1000 IMPACT EVALUATION BASIC INFORMATION AND DATA USE INSTRUCTIONS

Recommended citation for this document:
Transfer Project. 2020. "Ghana LEAP 1000 Impact Evaluation Basic Information Document". Carolina Population Center, Chapel Hill, NC, USA.

OVERVIEW

This document provides information for using the LEAP1000 impact evaluation data, a two-wave panel dataset that was created to analyze the impact of a pilot phase extending Ghana's Livelihood Empowerment Against Poverty (LEAP) cash transfer programme to a new category of pregnant women and women with children under one year old. In addition to explaining the data structure and steps for merging files, it provides brief information about the programme and the evaluation.

This dataset is released by The Transfer Project, housed at the Carolina Population Center at the University of North Carolina – Chapel Hill. Additional information about the project not found here or without a direct link can be found on The Transfer Project's Website: https://transfer.cpc.unc.edu/.

The data package contains 59 primary datasets (from two waves of individual/household surveys). The survey interviewed households at two points in time, in 2015 and 2017. The household survey was complemented by a health facility survey and price survey at baseline, and a community survey and price survey at endline.

BACKGROUND

The Livelihood Empowerment Against Poverty (LEAP) Programme and LEAP 1000 pilot

The LEAP (Livelihood Empowerment Against Poverty) Programme began in 2008 and is Ghana's flagship poverty alleviation programme. It is implemented by the LEAP Management Secretariat (LMS) and the Department of Social Welfare (DSW) under the guidance of the Ministry of Gender, Children and Social Protection (MoGCSP). Designed to fight poverty among extremely vulnerable populations, the LEAP Programme provides bimonthly cash payments to extremely poor households with orphans and vulnerable children, the elderly with no productive capacity, persons with acute disability, and, in 2015, a pilot called 'LEAP 1000' was launched to include a new category – pregnant women and children under the age of 12 months. Since then, the 'LEAP 1000' category has been mainstreamed into the larger LEAP Programme. As of December 2018, LEAP reaches more than 213,000 poor families in all 216 districts of Ghana

The LEAP 1000 pilot initially targeted a total of ten districts in Northern Ghana (three districts in Upper East region and seven districts in Northern region). These districts were selected by applying criteria based on the high proportion of poor people within a district, combined with a high incidence of poor nutrition. However, the LEAP 1000 category was quickly mainstreamed into the larger LEAP Programme and since the end of 2015 has been included alongside the other categories in the nationwide expansion

of LEAP. This demographic group now makes up about 8% of all LEAP beneficiaries. The description below refers to the initial roll-out in ten districts.

Communities within the first ten LEAP 1000 districts were targeted using official poverty rankings established at district level. Poverty rankings of communities in Ghana are based on a validation of census data by district assembly members (District Social Welfare Officers, District Health Officials, and District Chief Executives). Once the poorest communities were identified using the district ranking, priority was given to the poorest communities which were not already covered by mainstream LEAP.

Targeting of beneficiaries occurred between March 2015 and July 2015 using a demand-driven approach. In the ten selected districts, mobile units were deployed to advertise the programme and encourage potentially eligible women to apply to enter the programme. To be eligible to apply, pregnant women and households with infants under 15 months had to present proof of either: (a) antenatal cards, if pregnant; or (b) birth certificates and weighing cards, if they have an infant below 15 months. Women unable to present either document during the targeting process were advised that if selected, the necessary documentation should be provided during enrolment. All those who applied were then administered the standard LEAP proxy means test (PMT) and assigned a score to ensure they met the poverty criterion. Those that met the poverty criterion (households with a PMT score below the designated threshold), were enrolled into the programme from August 2015 onwards, receiving their first payment during the September 2015 payment cycle of LEAP. In total, LEAP 1000 enrolled 6,124 poor households with pregnant women and infants in 2015.

Women enrolled in the programme receive bi-monthly payments of cash in alignment with the mainstream LEAP. During the period of the impact evaluation (2015-2017), the amount of the cash transfer, which depends on the number of eligible household members, was as follows:[1]

- One eligible household member: GH₵ 64
- Two eligible household members: GH₵ 76
- Three eligible household members: GH₵ 88
- Four or more eligible household members: GH₵ 106

Further, LEAP beneficiaries are entitled to free health insurance through the National Health Insurance Scheme (NHIS), giving them access to free out-patient and in-patient services, dental services, and maternal health services. This reflects a step towards better integration of social protection programming and is the result of a collaboration between the National Health Insurance Agency (NHIA) and the DSW starting in 2011 to enrol beneficiaries of LEAP into the NHIS. LEAP beneficiaries qualify for the NHIA “indigent” exemption which waives all fees for NHIS including card processing fees, premiums and renewals. All members of LEAP households are entitled to a complete waiver of NHIS enrolment fees and premiums In 2016, LEAP conducted a nation-wide exercise to enrol 97,536 LEAP beneficiaries on NHIS. However, it is important to note registration in NHIS must be formally renewed each year, and this national exercise is not a routine occurrence.

For more information on LEAP see the government of Ghana’s website, http://leap.gov.gh/

[1] Note that the demographic groups of the wider LEAP programme also count as eligible household members for this calculation. A pregnant woman and a caregiver with a child under 1 year both count for 2 beneficiaries, one mother/caregiver and one infant. The minimum amount a LEAP 1000 household receives is therefore by definition GH₵ 76.

The impact evaluation

The Ghana LEAP 1000 impact evaluation comprised quantitative surveys (community, health facility and household) and an embedded qualitative study with beneficiaries. This document only describes the quantitative evaluation design.

Since randomization was not an option for LEAP1000, the study uses a regression discontinuity design (RDD). The RDD works in situations where treatment is determined by whether a value on a continuous numerical score falls below or above a predetermined threshold or cutoff. The main idea of this approach is that households in the close vicinity of the eligibility threshold are 'as good as randomly' assigned to the treatment and comparison group. In the case of LEAP 1000, the numerical score is the PMT score, and the cutoff for eligibility was determined by the LEAP Management Secretariat (LMS), placing the score between the extreme poverty and poverty lines. This score would also be relatively close to the lowest wealth quintile of GLSS6, a group often used as a comparison for LEAP households.

The sample

The evaluation was conducted in five of the 10 districts in which LEAP 1000 operates: Yendi, Karaga and East Mamprusi in the Northern Region and Bongo and Garu Tempane in the Upper East Region.

Because the key idea behind RDD is that households just below and above the cutoff are highly similar, the sampling strategy sought to select those households that were closest to the cutoff. Out of the 8,058 households in 189 communities who applied for LEAP 1000, a sample of 1,250 households below the cutoff and 1,250 households above the cutoff were selected. This number was chosen based on power calculation around key outcomes (nutritional status). Based on their PMT score, households were sorted in ascending order for the comparison group and in descending order for the treatment group, and the first 1,250 top ranked households for each group were selected as the initial sample. Since it was deemed inefficient to visit communities with fewer than three selected households, the sample was restricted to communities in which at least three households were selected. An additional sample of 125 households on either side of the cutoff was added to serve as replacements in the case of refusals or inability to locate sampled households during fieldwork. At baseline, the number of successfully completed interviews was 2,497, of which 1,262 had a PMT score lower than the threshold (treatment) and 1,235 had a score above the threshold (comparison).

The baseline analysis found that the two groups were balanced at baseline, with fewer than five per cent of indicators (based on more than 500 statistical tests) showing a statistical difference between the two groups. At endline, 2,331 households were successfully re-interviewed and retained in the panel, indicating an attrition rate of 6.65 per cent.

Table 1: Sample for the evaluation

	Baseline (2015)	Endline (2017)
Treatment (LEAP 1000)	1,262	1,185
Comparison	1,235	1,146
TOTAL	2,497	2,331

All evaluation reports (baseline and endline) as well as the instruments used for the study can be found here: https://transfer.cpc.unc.edu/?page_id=1231

Contents of the data files

1. Household survey

The data package for the household survey consists of 59 separate data files, each one representing the answers to the various modules of the questionnaire at different waves. The contents of all the different files is described in Table 2 at the end of this manual.

2. Price survey

The price survey was conducted at baseline and endline in the main markets serving the households in the sample. At baseline, the survey was conducted at 32 markets, and at endline, at 33 markets. For convenience, all the prices are combined into one file by survey wave. Hence, there are 2 files, one for baseline and one for endline.

3. Health facility survey

At baseline, a survey of primary health facilities was conducted in the evaluation districts. A total of 142 facilities were surveyed. The data is provided separately for each section of the health facility survey (7 files).

4. Community survey

At endline only, an additional survey was conducted to collect information about the communities in the evaluation sample. For logistical and efficiency reasons, the community survey was conducted only in communities where 5 or more households were located. A total of 131 community surveys were completed, covering 2,215 household, or 88.7% of the total household sample. The data comes in 7 separate files.

Merging datasets

1. Household survey

Datasets can be merged using the household (hhid) and individual identifiers (pid) provided in the Table 2 at the end of this document.

2. Price survey

The baseline price survey can be matched to the household file using the auxiliary file '*LEAP1000_bl_hf_market.dta'*. This file includes for each household (hhid) the ID (qid_market) of the closest market, based on the GPS coordinates.

The endline price survey followed a different structure and was collected at the community level. Therefore, this file can be merged to the household data using the *'community'* variable in both datasets.

3. Health facility survey

The health facility survey can also be matched to the household file using the auxiliary file '*LEAP1000_bl_hf_market.dta'*. This file includes for each household (hhid) the ID of the closest health facility (qid_healthfac) and nearest health center (qid_healthctr), based on the GPS coordinates.

4. Community survey

Data from the community survey can be merged to the household data using the *'community'* variable in both datasets.

Supporting documentation

All supporting documents, such as questionnaires, codebooks and evaluation reports, are either included in this package or can be found on the Transfer Project Website, https://transfer.cpc.unc.edu/?page_id=1231.

Notes to the household data

1. Differences between the baseline and endline questionnaire and data

To preserve consistency across the two waves, the research team has attempted to preserve the questionnaire and data structure between baseline and endline. However, a number of changes, updates and additions were made to the endline questionnaire, which are listed below:

- Section A1 and A2 were added to track old household members and list new members
- Section 3: Structure of q10 and q12 was changed from multiple variables for each option to one single variable. Addition of question 12a (not the same as q12a at baseline), 13b, 13c, 17, 18 and 19 at endline. Q15 from baseline was dropped, and q15b from baseline is equivalent to q15 from endline
- Section 4A: q10a was added
- Section 4B: q2a-q2g, and q6-q7 were added.
- Section 4D: question 2 about the ownership of animals was revised and additional questions included about the purchase and sale of animals. Data on animal ownership are in wide format at enline and are provided in a separate datafile 'SEC4D – ANIMALS.dta'. It is possible to count the number of animals from this file to make the ownership data comparable to the baseline data. Note that due to this change, variables s4d_2a s4d_2b s4d_2c s4d_2d s4d_2e s4d_2f at endline are not equivalent to variables with the same names at baseline.
- Section 4D: q16, q16a, q17, q17a and q17b were added
- Section 5B at endline only records new births since July 2015. This file should be combined with the birth history file from baseline to get a complete birth history of the respondent
- Section 6 at endline only records information of newly born children since July 2015
- Section 11: q1 was removed from endline. Q9 (Life distress scale), q10 (vignettes), q11, q12 (group membership) and q13-q14 (social capital) were added.
- Section 16: the age limit for measuring anthropometrics was raised to 83 months for endline
- Section 17 Shocks and Coping Mechanisms was added
- Section 17A Positive Shocks was added
- Section 18 Operational Performance was added

2. Inconsistency between section 5A and 5B at baseline and endline

During the data cleaning stages at time of the endline data collection, a number of inconsistencies were found between Section 5A (Reproductive health of all women in the household) and 5B (Birth history of the main respondent). In theory, the total number of births and children alive for the main respondent reported at endline in section 5A should be equal to the number of births and children alive reported in section 5B at baseline and endline combined.

- For 849 cases, the number of total births reported in Section 5B (BL+EL) is different than in 5A (EL)

- For 616 cases, the number of living kids reported in Section 5B (BL+EL) is different than in 5A (EL)

For researchers interested in using this data for their research purpose, we recommend using the data from section 5B.

3. Inconsistencies between anthropometric measurements

The research team did everything in their power to get accurate anthropometric measurements from the children in the study, which resulted in high-quality anthropometric data. Despite all efforts, a number of cases were identified with inconsistencies between the baseline and endline measurements:

- Age progression between BL and EL is inconsistent (>24 or <22 months) and there is no birth date for these observations to verify their age (155 cases)
- Negative weight gain between BL and EL (50 cases)
- Weight gain beyond +/- 3SD of the mean weight gain (29 cases)
- Negative height gain between BL and EL (21 cases)
- Height gain beyond 3SD of mean height gain (36 cases)

When combined, there are 256 out of 3,298 (~8%) observations with any of these inconsistencies. We have not altered the raw data, and leave the researcher with the discretion to decide how to handle these cases.

4. Number of observations in Section 11-14 at endline

While the number of households at endline was 2,331, an additional number of 70 main respondents were lost at follow-up. Hence the number of observations in the section specifically for the main respondent is lower than the number of households at endline.

5. Imputed value of gifts in 15A

The food consumption module (Section 15A) collected detailed information on the amount and value of the food consumed by the household. Purchases, own production and gifts were counted as consumption. The value of the purchases, as well as the market value of produced food was directly elicited from respondents. Values of goods received (and for a minimal number of missing values) were imputed as follows and stored in a new variable, s15a_8 (457 values at baseline and 901 at endline)[2]:

1. If the household bought or produced the item as well, take the mean household expenditure on units bought and produced (40 imputations at baseline/58 at endline)
2. If the household has not bought or produced the item, take the mean expenditure by district/item/unit (411/806)
3. If the household has not bought or produced the item, and no other household in the district as well, take the mean expenditure by item/unit for full sample (5/6)
4. As a final resort, the data was inspected if another item/unit combination could provide information on the value of the gift, or the community price data was consulted for information about the value of the concerned goods (1/31)

[2] At baseline (endline), there were a total of 37,903 (35,837) nonzero observations for consumption goods, so the imputations represent a rather small part of the total consumption aggregates.

Notes to the price data

There were a few inconsistencies with the numbering of the items between the baseline and endline price survey. Therefore, a variable was added in the endline price data '*itemno_bl*' which contains the corresponding item number from the baseline survey.

Notes to the health facility data

A post-survey mop up was conducted in April 2016 to validate health facility data collected and to survey uncovered health facilities that were missed in the main baseline survey in September 2015. The mop up exercise found a total of 67 uncovered health facilities across the five districts. This include: 7 in East Mamprusi, 25 in Garu-Tempane, 23 in Bongo, 7 in Karaga, and 5 in Yendi. The date of interview is provided in the data which allows the distinction between the originally surveyd facilities and the facilities added during the mop-up exercise.

The team presented the list of health facilities surveyed in the main visit in 2015 for validation by the district health directorate. It turned out that some facilities which were covered in Bongo and Garu-Tempane should not have been surveyed because they were listed as feeding or rehabilitation centres and not health facilities in the records of the health directorate. Another category has subsequently been added to the "type of facility" variable (*facility_type2*) as "feeding or nutrition centre" in the data. They can also be expunged from the data since they are not regarded as health facilities as per the official checks in the district health directorates. This is left up to the researcher to decide based on their research needs.

It was also confirmed that 5 facilities that were surveyed should not have been covered since they were not located in the study districts. They are Bugri Health Centre and Kuka Clinic in Bawku District and not Garu-Tempane District; Pitanga CHPS compound and Dasobiligo CHPS compound were in the Nabdam district and not in Bongo District; and Zinindoo CHPS compound was in Saboba District and not Karaga district as covered in the data. These facilities are *not* dropped from the data as households may still cross district boundaries to access health services.

Additional information

In part C under kind of services rendered in the facility, and if any, the number of clients served in the previous month, that is March 2016, as mop up was conducted in April;

- ANC and PNC/CWC, we recorded only the number of new registrants in the month of March as against total attendance in March. The total attendance figure was normally very high in hundreds, which included new registrants in March and the "revisitors".
- Family Planning Services, only the figures of 'new or fresh acceptors' in March were recorded as against the figures of those who are in the continuation list. 'New acceptors' means clients who subscribed to the services of one or a mix of the available family planning services for the first time in that facility.
- In the case of malnourished children, the recorded figures reflect the number of new cases of malnourished children who were admitted to the CMAM Program in March. A zero (0) figure means the facility did not record a new case, and also has treated and discharged all clients before March.

- For mobile clinics, the figure refers to the number of times the facility embarked on outreach programs in March. The figures do not mean actual persons served, but reflect how many times the facility carried out clinical or preventive care services in the outreach services. The clients they served on outreach services are part of those reported under the services mentioned above.
- For Home Visits, under other services rendered, the facilities report the number of houses they visited, hence figures under this column reflects only the number of houses and not persons.

In part E:

- Figures under Classified Daily Employees are combinations of the number of Voluntary Health Extension Workers, Security Personnel, and Cleaners.
- Figures under Auxiliary Nurses combine Enrolled Nurses (they have certificates from the Health Assistants' Training Schools across the country); and Community Health Nurses (those who graduate from the Community Health Schools in the country).

Table 2: Data contents for Ghana LEAP 1000 data files

Dataset	Observation level	Unique ID	Baseline	Endline	Obs. BL	Obs. EL
SECA1 - HOUSEHOLD MEMBER TRACKING	Individual	hhid, pid		X		15,520
SECA2 - NEW HOUSEHOLD MEMBERS	Individual	hhid, pid		X		1,198
SEC0 - HH LEV.dta	Household	hhid	X	X	2,497	2,497
SEC1 - HH ROSTER.dta	Individual	hhid, pid	X	X	16,493	15,256
SEC2 - EDUCATION.dta	Individual	hhid, pid	X	X	13,738	15,262
SEC3 - HEALTH.dta	Individual	hhid, pid	X	X	12,301	15,262
SEC4A - HOUSING CONDITIONS AND WASH.dta	Household	hhid	X	X	2,497	2,331
SEC4B - FOOD SECURITY.dta	Household	hhid	X	X	2,497	2,331
SEC4C - TIME USE AND EMPLOYMENT.dta	Individual	hhid, pid	X	X	11,646	15,256
SEC4D - AGRIC INPUTS.dta	Item	hhid, itno	X	X	17,479	16,317
SEC4D - LOANS.dta	Loan	hhid, lid	X	X	1,027	1,030
SEC4D - PRODUCTIVE LIVELIHOODS.dta	Household	hhid	X	X	2,497	2,331
SEC4D - ANIMALS.dta	Animal	hhid, animalid		X		15,259
SEC4E - NON-FARM ENT.dta	Entreprise	hhid, s4e_id	X	X	583	720
SEC5A - REPRODUCTIVE HEALTH.dta	Individual (females)	hhid, pid	X	X	3,968	3,632
SEC5B - BIRTH HISTORY.dta	Birth	hhid, birthorder	X	X	8,964	780
SEC6 - MATERNAL AND NEWBORN HEALTH.dta	Child	hhid, pid	X	X	2,507	788
SEC7 - PREVENTIVE CARE AND CHILD HEALTH.dta	Child	hhid, pid	X	X	3,782	3,208
SEC8 - IMMUNIZATIONS.dta	Child	hhid, pid	X	X	3,782	3,209
SEC9 - CHILD NUTRITION AND FEEDING.dta	Child	hhid, pid	X	X	3,782	3,209
SEC10 - BIRTH REGISTRATION AND CHILD DEVELOPMENT.dta	Child	hhid, pid	X	X	3,782	3,209
SEC11 - CONTRACEPTION.dta	Main respondent	hhid	X	X	2,497	2,270
SEC12 - WOMEN'S EMPOWERMENT, STRESS AND PREFERENCES.dta	Main respondent	hhid	X	X	2,497	2,270
SEC12 - WOMEN'S EMPOWERMENT, STRESS AND PREFERENCES_II.dta	Group	hhid, groupid		X		18,648
SEC13 - NUTRITION AND FEEDING KNOWLEDGE.dta	Main respondent	hhid	X	X	2,497	2,270
SEC14 - DOMESTIC VIOLENCE.dta	Main respondent	hhid	X	X	2,497	2,270
SEC15A - FOOD CONSUMPTION.dta	Item	hhid, itemno	X	X	222,233	207,459
SEC15B - CONSUMPTION OF DURABLE GOODS (6M).dta	Item	hhid, itemno	X	X	122,353	102,564
SEC15C - CONSUMPTION OF DURABLE GOODS (12M).dta	Item	hhid, itemno	X	X	39,952	37,296
SEC16 - ANTHROPOMETRY.dta	Child	hhid, pid	X	X	3,742	4,477
SEC17 - SHOCKS	Shock	hhid, sid		X		34,965
SEC17A - POSITIVE SHOCKS	Shock	hhid, sid		X		9,324
SEC18 - OPERATIONS	Household	hhid		X		2,329