Technical Note 35

Young Lives

Young Lives Rounds 1 to 4 Constructed files

Obiageri Bridget Azubuike and Kristine Briones

Young Lives Rounds 1 to 4 Constructed files

Obiageri Bridget Azubuike and Kristine Briones

First published by Young Lives in May 2016

Printed on FSC-certified paper from traceable and sustainable sources.

About Young Lives

Young Lives is an international study of childhood poverty, following the lives of 12,000 children in four countries (Ethiopia, India, Peru and Vietnam) over 15 years. **www.younglives.org.uk**

Young Lives is funded from 2001 to 2017 by UK aid from the Department for International Development (DFID), It is co-funded by the Netherlands Ministry of Foreign Affairs from 2010 to 2014, and by Irish Aid from 2014 to 2016.

Funded by

UKaid
from the British people

Ministry of Foreign Affairs of the Netherlands

Irish Aid
An Roinn Gnóthaí Eachtracha agus Trádála
Department of Foreign Affairs and Trade

Young Lives, Oxford Department of International Development (ODID), University of Oxford, Queen Elizabeth House, 3 Mansfield Road, Oxford OX1 3TB, UK

Tel: +44 (0)1865 281751 • E-mail: younglives@younglives.org.uk

Contents

Summary

This Technical Note accompanies the Constructed Files of Young Lives data which have been deposited with the UK Data Service to facilitate analysis of the household and child survey data across the four rounds of data collected to date. The constructed files are combined sub-sets of selected variables from Rounds 1 to 4 of the Young Lives survey, carried out in 2002 (Round 1), 2006 (Round 2), 2009 (Round 3) and 2013 (Round 4) when the Younger Cohort children were aged 1, 5, 8, and 12 years and the Older Cohort children were aged 8, 12, 15, and 19 years.

The files contain about 200 original and constructed variables, most of them comparable across the four rounds, presented in a panel format and classified in four broad groups: panel information, general characteristics, household characteristics, and child characteristics. This document is organised around the same groups.

The authors

Bridget Azubuike is a quantitative research assistant in the Young Lives team. Her role involves providing support to the research and education teams. She has an MSc in Economics from the University of Sussex and a BSc in Economics and Statistics from University of Benin, Nigeria. **Kristine Briones** is a quantitative research assistant in the Young Lives team and is currently providing support in the development of the Round 5 survey. She has an MSc in Impact Evaluation for International Development from the University of East Anglia, and a BSc in Statistics and MA in Economics from the University of the Philippines.

Acknowledgements

The dataset of constructed variables was compiled and checked by Bridget Azubuike and Kristine Briones, based on preliminary work carried out by Patricia Espinoza for the Round 1 to Round 3 constructed variables (see Young Lives Technical Note 30, 2014). Useful suggestions and comments were received from Patricia Espinoza, Abhijeet Singh, and Anne Solon.

1. Introduction

This Technical Note accompanies the Constructed Files of Young Lives data which have been deposited with the UK Data Service to facilitate analysis of the household and child survey data across four rounds of data collected. The constructed files are combined sub-sets of selected variables from Rounds 1 to 4 of the Young Lives survey carried out in 2002 (Round 1), 2006 (Round 2), 2009 (Round 3) and 2013 (Round 4) when the Younger Cohort children were aged 1, 5, 8, and 12 years and the Older Cohort children were aged 8, 12, 15, and 19 years.

The files contain about 200 original and constructed variables, most of them comparable across the four rounds, presented in a panel format and classified in four broad groups: panel information, general characteristics, household characteristics, and child characteristics. This document is organised around the same groups.

2. Panel information

Data: Four dummy variables showing whether the child has been found in each round, and an additional dummy signalling whether child was present in all rounds.

inr1	Child is present in Round 1
inr2	Child is present in Round 2
inr3	Child is present in Round 3
inr4	Child is present in Round 4
panel1234	Child is in all rounds

3. General characteristics

Data: Date of interview, and other basic geographical identifiers (area of residence, region of residence, sentinel site id, and community id). Note that the variable name for Peru's community ID is **placeid**.

dint	date of interview
commid	community id
clustid	sentinel site id
typesite	area of residence (urban/rural)
region	region of residence

4. Household characteristics

4.1. Household wealth

Data: Wealth index and constituent sub-indexes (housing quality, access to services, and consumer durables).

The indexes were estimated consistently across rounds. For this purpose, only variables common to the four rounds were included. The wealth index is composed of three sub-indexes: (i) housing quality index (hq), (ii) access to services index (sv), and (iii) consumer durables index (cd); all of which have equal weights in the estimation of the wealth index.

Therefore, the **wealth index** of household i will be defined as:

$$wi_i = \frac{hq_i + sv_i + cd_i}{3}$$

The **Housing quality index** (hq) is the simple average of the following indicators:

- Crowding (scaled sleeping rooms per person)
- Main material of walls – dummy variable that takes the value of 1 if main material of walls satisfied basic norms of quality.
- Main material of roof
- Main material of floor.

The **Access to services index** (sv) is the simple average of the following indicators:

- Access to electricity (elecq)
- Access to safe drinking water (drwaterq)
- Access to sanitation (toilet)
- Access to adequate fuels for cooking (cookingq).

The **Consumer durables index** (cd) is the simple average of a set of dummy variables which take the value of 1 if a household member owns at least one of each consumer durable. To ensure comparability across rounds, only those consumer durables common to all three rounds were included. The following tables show the lists of common consumer durables used in each country:

Ethiopia: 10 common items across all rounds	India: 9 common items across all rounds	Peru: 12 common items across all rounds	Vietnam: 9 common items across all rounds
Radio	Radio	Radio	Radio
Television	Television	Television	Television
Bicycle	Bicycle	Bicycle	Bicycle
Motorbike	Motorbike	Motorbike	Motorbike
Automobile	Automobile	Automobile	Automobile
Landline phone	Landline phone	Landline phone	Landline phone
Mobile phone	Mobile phone	Mobile phone	Mobile phone
Table and chair	Refrigerator	Refrigerator	Refrigerator
Sofa	Fan	Stove	Fan
Bedstead		Blender	
		Iron	
		Record player	

4.2. Consumption aggregates

Data: Total per capita expenditure, per capita food consumption, and per capital non-food expenditure, all in both nominal and real terms.

The construction of the consumption aggregates involved adding together a number of items grouped in two main classes: (i) food items, and (ii) non-food items. It should be noted that while a core set of items is similar in all four countries, other food and non-food items that are specific to each country were added in the design of the questionnaire.

Food consumption: Aggregation of all food items consumed in the last 2 weeks from different sources: (i) food purchased; (ii) food home-produced (from own harvest) or from stock; (iii) food items received as gifts or transfers; and (iv) food received from employers as payment in-kind for services rendered. In the case of Peru, there is an additional question related to all food that was left over. Therefore, this amount is subtracted from the final aggregate.

Each source is converted to monthly terms (by multiplying it by 2 because the recall period is 2 weeks) and finally aggregated.

Non-food consumption: Aggregation of all non-food items, which are classed together in 4 big groups: (i) expenditure on education; (ii) expenditure on health; (iii) expenditure on clothing and footwear; and (iv) expenditure on other non-food items.

The selection of the items was based on a comparability criteria (i.e. all those items that were included consistently in all rounds). Since this information was collected for different reference periods, they all were converted to months before aggregating them.

(i) ***Expenditure on education:*** Includes all money spent on school uniform for boys and girls, payments for tuition, fees or donations to school, books and stationary, and transport to school.

(ii) ***Expenditure on health:*** Includes all money spent on medical consultations and treatment and other medical expenses.

(iii) ***Expenditure on clothing and footwear****:* Includes all money spent on clothing and footwear for adults and children.

(iv) ***Expenditure on other non-food items****:* Includes all money spent on other non-food items such as: rents, dwelling and vehicle maintenance, water supply, electricity rates, telephone and mobile phone rates, fees and paperwork, legal advice, bribes, one-off family events, festivals and celebrations, personal care items, entertainment, presents for children, and jewellery (although this item was excluded from the consumption aggregate of Peru and Vietnam because it caused too much distortion).

Once food and non-food items are aggregated, they are adjusted for cost of living differences and for household composition. Prices to deflate nominal consumption aggregates are collected from external sources for all countries, except for India, where the information comes from community questionnaire, as it seemed more appropriate for our research purposes. This information is now archived together with the main data in order to be able to reproduce the same consumption aggregates if needed.

In all countries, except Ethiopia, the results are adjusted by the household size (i.e. all members that live permanently in the household) and, thus, are reported in per capita terms. In Ethiopia, the results are reported in ‘per adult’ terms. This means that real expenditure is

divided by the household size adjusted for adult equivalence. Dercon and Krishnan (1998)[1] proposed the following equivalences based on nutritional (caloric) requirements of different ages and men and women.

Age (years)	Men	Women
0-1	0.33	0.33
1-2	0.46	0.46
2-3	0.54	0.54
3-5	0.62	0.62
5-7	0.74	0.70
7-10	0.84	0.72
10-12	0.88	0.78
12-14	0.96	0.84
14-16	1.06	0.86
16-18	1.14	0.86
18-30	1.04	0.80
30-60	1.00	0.82
60 plus	0.84	0.74

4.3. Livestock ownership

Data: Number of animals owned by the household.

In order to have comparability across rounds, animals were aggregated in four big groups: milk animals (*animilk*), draught animals (*anidrau*), small ruminants (*anirumi*), and animals that are specific to each country (*anispec*). For Rounds 2 to 4, the number of livestock for specific animals is reported also reported.

Note: While **anibull** refers to the number of bullocks owned by the household in Ethiopia and Vietnam, the variable refers to the number bullocks AND oxen owned by the household in India.

4.4. Land ownership

Data: access to land (owned, borrowed, rented, etc.) in the last year (*accland*), total area of land the household has had access to (in hectares) in the last year (*totland*), total area of land owned by the household (in hectares) in the last year (*ownland*).

In India, land area was collected consistently in acres. Therefore, for the panel datasets, the information was transformed to hectares using the conversion factor 1 acre= 0.4047 hectares. In Vietnam, the unit in which land extension was collected was squared metres. For this reason, the conversion factor used was 0.0001 (to hectares). Note that in Vietnam information about total area owned refers to having a long-term use certificate for the land.

In Ethiopia, while in Round 1 this information was recorded in squared metres, in the succeeding rounds, it was recollected using (most commonly used) local units. Therefore, all the information was harmonised by converting it to hectares. This was done by using the following conversion factors:

[1] Stefan Dercon and Pramila Krishnan (1998) 'Changes in Poverty in Rural Ethiopia 1989-1995: Measurement, Robustness Tests and Decomposition', CSAE Working Paper 1998-07, Centre for the Study of African Economies, University of Oxford.

1 unit of	In hectares
Gasha	40.00
Timad, Gemed, Kert, Kedema	0.25
Massa	0.67
Kufaro	0.03
Zhir	0.0003
Boy	0.005
Tinto	0.06
Ermija, squared metres	0.0001
Dearo	0.045
Gezem	0.1666
Kend	0.000025
Fer	0.25

Similarly in Peru, Rounds 2 to 4 collected information on land using local units. The conversion factors are listed in the following table:

1 unit of	In hectares	In locality (department)
Cuadra	0.70	Cajamarca, Piura, Tumbes
Cuadra	1.00	Huanuco
Metro cuadrado (m2)	0.0001	across the country
Tarea	0.13	Apurimac, Cusco
Tarea	0.10	Tumbes
Tarea aradura, tarea siege, Tarea carguio	0.14	Lambayeque
Tarea trasplante	0.07	Lambayeque
Tongo	0.08	Junin (Tarma)
Tongo	0.05	Junin
Topo	0.35	Arequipa, Maquegua, Tacna
Topo	0.33	Apurimac, Cusco, Puno
Topo	0.13	Ayacucho, Junin
Yugada	0.25	Amazonas, Ancash, Apurimac, Ayacucho, Cusco, Huancavelica, Huanuco, Puno
Yugada	0.33	Cusco, Huanuco, Junin
Yugada	0.30	Pasco
Yugada	0.20	Lima
Yugada chaquitaclla	0.05	Lima (Canta)

4.5. Characteristics of the household head

Data: basic characteristics of the household head such as the ID in the roster, age, sex, level of education (i.e. highest grade attained).

Because it is expected that household heads maintain their level of education across rounds, unless they were enrolled in formal schooling and had attained new grades as a consequence, the variable for the level of education has been harmonised for those households that have the same household heads throughout the three waves of interview or in at least two of them. In this way, we overcome reporting inconsistencies and we end up with only one level of education for a singular household head.

4.6. Characteristics of the primary caregiver

Data: basic characteristics of the primary caregiver such as the ID in the roster, age, sex, literacy, level of education (i.e. highest grade attained), relationship to household head, relationship to index child, and an additional variable related to caregiver's subjective well-being (in terms of the nine steps of a life-satisfaction ladder).

The variable of education was harmonised in the same way as for the household head (see 4.5).

careid	Caregiver's ID in roster
careage	Caregiver's age
caresex	Caregiver's gender
carehead	Caregiver's relationship to head of household
carerel	Caregiver's relationship to child
carelit	Caregiver is literate
caredu	Caregiver's level of education (harmonised variable)
ladder	Caregiver's subjective well-being (ladder)

Caregiver relationship to the index child (*carerel*) has been harmonised across rounds. It has 6 options: (1) biological parent, (2) partner of biological partner, (3) Grandparents, (4) Uncles/Aunts, (5) Siblings (which includes half-siblings), and (6) Other (including all other household or non-household members).

Information about the caregiver is available for both Younger and Older Cohorts in all countries in Rounds 1 to 3 except for the Older Cohort in Vietnam in Round 3. In this case, the respondent for Section 5 (Social Capital) Household Questionnaire was considered as the caregiver. In Round 4, no information about the caregiver was obtained for the older cohort in all countries.

4.7. Household composition

Data: household size (including the index child), and number of household members by sex and age groups (excluding index child).

hhsize	household size
male05	Number of males aged 0-5
male612	Number of males aged 6-12
male1317	Number of males aged 13-17
male1860	Number of males aged 18-60
male61	Number of males aged 61+
female05	Number of females aged 0-5
female612	Number of females aged 6-12
female1317	Number of females aged 13-17
female1860	Number of females aged 18-60
female61	Number of females aged 61+

4.8. Shocks and adverse events

Data: occurrence of events that have affected negatively the economic situation of the household.

All shock-related variables are binary, being 1: shock was reported, 0: shock was not reported. Note that answers are based on perceptions, this is, they do not show whether a negative event has occurred or not, rather they show whether the respondent considers the event has affected the welfare of the household negatively.

5. Child characteristics

5.1. General characteristics

Data: sex, age (in completed months), first language, ethnicity, and religion.

Given that sex, ethnicity, and religion are time-invariant variables, they were taken from Round 1.

Age in months is estimated by taking the age of the child in days (date of Interview-date of birth) and dividing this number by 365/12 (number of days per month). The final number is rounded up to one decimal point. In order to preserve anonymity, dates of birth cannot be publicly archived, therefore external users will not be able to estimate this variable, but details on how it was estimated can be found in the do-files.

Although child's first language is a time-invariant variable, it was taken from Round 2 because the information was not collected in Round 1. Therefore, most missing values in this variable are explained by attrition.

5.2. Child health and nutrition

Data: Anthropometric information, weight at birth, antenatal care, constructed measures for malnutrition, and self-reported health (i.e. in relation to other children of same age, serious injuries, and long-term health problems).

Anthropometric information includes health and weight, and z-scores for weight-for-height, height-for-age, and BMI-for age. The z-scores were estimated using WHO references tables and software (available for download at: http://www.who.int/childgrowth/en/). These measures were estimated using the age of children in days. To keep anonymity, the latter information cannot be publicly archived; therefore, the results we provide cannot be reproduced exactly. Age of child in months rounded up to one decimal, however, provides very close estimators.

Weight at birth and antenatal care information was only asked to the Younger Cohort in Round 1.

Malnutrition estimators were constructed on the basis of the z-scores. The six estimators that are included in the panel files are: stunting, severe stunting, thinness, severe thinness, underweight, and severe underweight. The following table provides the definition for each:

Malnutrition estimator	Variable name	Definition
Stunted	stunt	<-2 SD of height-for-age z-score
Severely stunted	sstunt	<-3 SD of height-for-age z-score
Underweight	underweight	<-2 SD of weight-for-age z-score
Severely underweight	sunderweight	<-3 SD of weight-for-age z-score
Thinness	thinness	<-2 SD of BMI-for-age z-score
Severe thinness	sthinness	<-3 SD of BMI-for-age z-score

Finally, three variables of self-reported health are included:

(1) Child's health relative to other children of the same age (*chhrel*)

Although this information exists for both cohorts in Rounds 1 to 3 in all countries and in Round 4 in Peru, the question was asked to different respondents (either the caregiver or child). For the Younger Cohort, the question was asked only to the caregiver. For the Older Cohort, the question was asked to both the caregiver and child in Round 2, and only to the child in Round 3 and Round 4.

In order to have a panel information, the variable 'chhrel' was created using the responses of the caregiver in all rounds for both the Younger Cohort and Older Cohort except in Ethiopia and Peru where the child's response is used for Rounds 2 to 4. See table below.

	Round 1 respondent	Round 2 respondent	Round 3 respondent	Round 4 respondent
Ethiopia				
Younger Cohort	Caregiver	Caregiver	Caregiver	(data not available)
Older Cohort	Caregiver	Child	Child	(data not available)
India				
Younger Cohort	Caregiver	Caregiver	(data not available)	(data not available)
Older Cohort	Caregiver	Caregiver	(data not available)	(data not available)
Peru				
Younger Cohort	Caregiver	Caregiver	Caregiver	Caregiver
Older Cohort	Caregiver	Child	Child	Child
Vietnam				
Younger Cohort	Caregiver	Caregiver	(data not available)	(data not available)
Older Cohort	Caregiver	Caregiver	(data not available)	(data not available)

Additionally, for the scales to be comparable across rounds, the scales of Round 3 and Round 4 were reduced from 5 to 3. See table below.

Constructed data	Rounds 1 and 2	Rounds 3 and 4
1 = Same	Same	Same
2 = Better	Better	Better, Much better
3 = Worse	Worse	Worse, Much worse

(2) Child's health in general (*chhealth*)

The overall health of the child was also asked in Round 3 and Round 4. The respondent for this question is the caregiver for the Younger Cohort and the child for the Older Cohort.

(3) Serious illnesses and injuries when caregiver thought the child might die (*chhilin*)

This information was collected consistently in Round 1 and Round 2. The specific question that was asked was: 'in the last [##] years, has the child had a serious illnesses or injuries,

when you REALLY thought he/she might die?'. The constructed variable is dichotomic, with 1 meaning that child has had a very serious illness or injury, and 0; otherwise.

(4) Serious injury (*chhinjury*)

In Round 3 and Round 4, the question was somewhat different compared to the previous question in (3) above, focusing only on serious injuries (instead of asking for serious injuries and illnesses as in Round 1 and Round 2).

(5) Long-term health problems (*chhprob*)

Similarly, information related to this variable was collected with some variation across rounds. In Round 1 and Round 2, the question was straightforward: 'Does child have any long-term health problem that affects how he/she attends to school or work?' However, in Round 3 and Round 4, this question was omitted and instead a list of long-term health problems was asked, in two steps: first, asking whether the child suffers from any of the listed long-term health problems, and second, asking if the problem affects his/her capacity to study.

For this reason, the panel variable corresponds to Round1 and Round 2 only; a dummy that takes the value of 1 if child has long term health problem, and 0 otherwise.

Additionally, a set of dummy variables is added for specific long-term health problems listed in Round 3 and Round 4. These are: poor vision, hearing problems, frequent headaches, and respiratory problems (*prvision, prhear, prhead, and prrest,* respectively).

5.3. Child education / formal skills

Data: enrolment, problems with reading and writing, literacy, school type, highest level attained, travel time to school, raw and standardised scores of PPVT, and number of correct answers in Ravens Test.

Enrolment (*enrol*) is constructed based on information on whether the child was attending formal school at the time of the interview. For Round 1, this information was collected for the Older Cohort only (the Younger Cohort was too young to be enrolled in school). In Round 2, when the Younger Cohort children were about 5 years old, enrolment includes pre-school enrolment. Following this, if the child was 'still attending pre-school,' at the time of the interview, he/she will be counted as enrolled in school.

Problems with reading and writing are based on the reading and writing items of the achievement and development instruments. For each of the items, a test was administered to the child. After the finalisation of the tests, fieldworkers are asked to fill in some answers based on what/how the child responded, as follows:

Reading item
01 - Can't read anything
02 - Reads letters
03 - Reads words
04 - Reads sentences

Writing item
01 - No (can't write anything)
02 - Writes with difficulties or error
03 - Yes, without difficulties or error

Based on this information, problems with reading (*readprob*) accounts for all those cases where children are not able to read sentences (not equal to 04). Problems with writing (*writeprob*), in turn, will be defined as not being able to write without difficulties or errors (not equal to 03).

In close relation to the two previous variables, literacy (*literate*) is defined as the ability to read and write without problems.

The school type variable (*schtype*) is a scaled-down version of the original that includes a range of options (private, public, community, government funded, community, NGO-Charity, etc.). The variable in the panel file only shows the options public, private, and other, for all countries. For Round 2, the type of pre-school is included for all those children enrolled in pre-school at the time of the interview. In Round 3, the question related to school type was not asked. For the Older Cohort, however, yearly school information is included in the school history. For the purposes of the panel dataset, the type of school that the child was attending the year of the interview (2009) is included. The only exception is Vietnam where school history was not collected in Round 3.

Ethiopia Round 1: question only referred to public and private.

The grade of school (*chgrade*) refers to the highest grade completed by the child at the time of the interview. For children who never attended school, the variable will have the value of 0.

In terms of the Ravens Test (*ravens*), administered to the Older Cohort in Round 1 only, information was not collected uniformly across countries; while in Peru and India the test was administered to the whole cohort, in Vietnam it was administered only to a sample of 200 children living in urban areas. In Ethiopia, in turn, the test was not administered at all.

5.4. Time allocation

Data: number of hours that the child spends performing different activities on a typical day.

Time-allocation information was collected in Rounds 2 to 4 in all countries. For the Younger Cohort, the information was obtained from the caregiver or other well-informed adult in the household in all three rounds. For the Older Cohort, on the other hand, information was obtained from the caregiver in Round 2 and from the child in Round 3 and Round 4.

Since the information was collected for household members age 5 to 17, missing values in Round 2, Younger Cohort, correspond to children younger than age 5 at the time of the interview.

The activities around which time-use information is collected are:

- Sleeping (*hsleep)*
- Caring for others in the household (younger children or sick household members) (*hcare)*
- Running household chores (fetching water, cleaning, cooking, etc.) (*hchore)*
- Working on household tasks (farming, herding, etc.) (*htask)*
- Working outside household on paid activities (*hwork)*
- At school (including travel time) (*hschool)*
- Studying outside school (doing homework, extra classes, learning languages, etc.) (*hstudy)*
- Leisure time (including playing, seeing friends, using the internet, etc.) (*hplay).*

5.5. Subjective and psychosocial well-being

Data: Child's self-assessment of personal well-being in terms of a nine-step ladder.

The information displays the step number in which the child situates his/her personal well-being at the time of the interview, where 9 represents the 'best possible life' and 1 'the worst possible life'.

This information is available for the Older Cohort in Round 2 and for both Younger Cohort and Older Cohorts in Round 3 and Round 4.

5.6. Parental characteristics

Data: for both biological mother and father the information in the panel file is:

- ID in roster
- age
- literacy - taken from Round 2
- highest grade attained– harmonised across rounds
- presence in household (at the time of the interview).

Appendix: How to access the Young Lives data

The datasets from the Young Lives household and child surveys in 2002 (Round1), 2006 (Round 2), 2009 (Round 3), and 2013 (Round 4) are publicly archived and available to download from the **UK Data Service** (https://www.ukdataservice.ac.uk/) along with the documentation and questionnaires for each survey round. For users in our study countries, they are also available on CD-Rom, on request from the Principal Investigator.

The data archive also includes community data from Rounds 2 to 4, school survey data from the four study countries, and a dataset with constructed variables from across the four rounds of the household and child survey to facilitate longitudinal analysis. Data from our qualitative sub-sample research are not archived in the same way as the survey data because of concerns about confidentiality.

The UK Data Service has developed a **Study Guide for Young Lives** (http://discover.ukdataservice.ac.uk/series/?sn=2000060) which acts as an entry point for the data. The individual datasets have been assigned the following study numbers:

- Young Lives Round 1 (2002): study number 5307
- Young Lives Round 2 (2006): study number 6852
- Young Lives Round 3 (2009): study number 6853
- Young Lives Round 4 (2013): study number 7931
- Young Lives School Survey, India (2010-11): study number 7478
- Young Lives School Survey, Peru (2011): study number 7479
- Young Lives School Survey, Vietnam (2011-12): study number 7663
- Young Lives School Survey, Ethiopia (2012-13): study number 7823.

Documentation

The Archive contains complete documentation relating to the survey, including:

- The household, child and community questionnaires for each survey round;
- Fieldworker manuals;
- Justification documents that describe what questions were asked and how they were arrived at;
- A data dictionary that describes each variable, the relevant question, and gives the code values where appropriate;
- For calculated variables, the description includes the method of calculation.

File format

The datasets are deposited as SPSS data files. For each survey round, there is one file containing all of the household and child data, plus other files containing sub-tables (e.g. the household roster, which is a list of all family members). The household-level file also contains the key composite variables that were used in the original tabulation plans, including the wealth index.

Using our data

Users are required to register and apply for a password with the UK Data Service and sign a confidentiality agreement before they can access the data. We also ask that users inform the Archive and Young Lives of any analysis or publications resulting from their work with the dataset. This helps us maintain an overview of how the data is being used, and is also required in our reporting to our funders.

If you use the Young Lives data in any publication, we would be grateful if you include the following acknowledgement:

> 'The data used in this publication come from Young Lives, a 15-year study of the changing nature of childhood poverty in Ethiopia, India (Andhra Pradesh and Telangana), Peru and Vietnam (www.younglives.org.uk). Young Lives is funded by UK aid from the Department for International Development (DFID) and co-funded by Irish Aid. The views expressed here are those of the author(s). They are not necessarily those of Young Lives, the University of Oxford, DFID or other funders.'

Further information

www.younglives.org.uk

Young Lives Rounds 1 to 4 Constructed files

This Technical Note accompanies the Constructed Files of Young Lives data which have been deposited with the UK Data Service to facilitate analysis of the household and child survey data across the four rounds of data collected to date. The constructed files are combined sub-sets of selected variables from Rounds 1 to 4 of the Young Lives survey, carried out in 2002 (Round 1), 2006 (Round 2), 2009 (Round 3) and 2013 (Round 4) when the Younger Cohort children were aged 1, 5, 8, and 12 years and the Older Cohort children were aged 8, 12, 15, and 19 years.

The files contain about 200 original and constructed variables, most of them comparable across the four rounds, presented in a panel format and classified in four broad groups: panel information, general characteristics, household characteristics, and child characteristics. This document is organised around the same groups.

Young Lives

An International Study of Childhood Poverty

About Young Lives

Young Lives is an international study of childhood poverty, involving 12,000 children in 4 countries over 15 years. It is led by a team in the Department of International Development at the University of Oxford in association with research and policy partners in the 4 study countries: Ethiopia, India, Peru and Vietnam.

Through researching different aspects of children's lives, we seek to improve policies and programmes for children.

Young Lives Partners

Young Lives is coordinated by a small team based at the University of Oxford, led by Professor Jo Boyden.

- *Ethiopian Development Research Institute, Ethiopia*
- *Pankhurst Development Research and Consulting plc, Ethiopia*
- *Centre for Economic and Social Studies, Hyderabad, India*
- *Sri Padmavathi Mahila Visvavidyalayam (Women's University), Andhra Pradesh, India*
- *Grupo de Análisis para el Desarollo (GRADE), Peru*
- *Instituto de Investigación Nutricional (IIN), Peru*
- *Centre for Analysis and Forecasting, Vietnamese Academy of Social Sciences, Vietnam*
- *General Statistics Office, Vietnam*
- *Oxford Department of International Development, University of Oxford, UK*

Contact:
Young Lives
Oxford Department of
International Development,
University of Oxford,
Mansfield Road,
Oxford OX1 3TB, UK
Tel: +44 (0)1865 281751
Email: younglives@younglives.org.uk
Website: www.younglives.org.uk

Young Lives