



YOUNG LIVES TECHNICAL NOTE NO. 1  
March 2008

# **An Assessment of the Young Lives Sampling Approach in Ethiopia**



**Ingo Outes-Leon**  
**Alan Sanchez**



# Contents

Executive summary	2
<b>1. Introduction</b>	<b>4</b>
<b>2. Overview of Ethiopia</b>	<b>5</b>
<b>3. Young Lives sampling procedure</b>	<b>6</b>
3.1 Sampling methodology	6
<b>4. Methodological issues</b>	<b>7</b>
4.1 Methodological tools	7
4.1.1 <i>Wealth index plots</i>	7
4.1.2 <i>Tools for comparison of indicators</i>	7
4.2 Ensuring comparability	8
4.2.1 <i>Demographic comparability</i>	8
4.2.2 <i>Geographic comparability</i>	9
4.3 Other concerns	10
<b>5. Comparison samples</b>	<b>10</b>
5.1 Demographic and Health Survey	10
5.2 Welfare Monitoring Survey	11
<b>6. Analysis of the wealth index</b>	<b>12</b>
<b>7. Comparison of living standard indicators</b>	<b>15</b>
7.1 Comparison with the Demographic and Health Survey	15
7.2 Comparison with the Welfare Monitoring Survey	18
<b>8. Assessing secular trends</b>	<b>21</b>
<b>9. Conclusion</b>	<b>22</b>
References	23
Appendix A: Comparison of the Demographic and Health Survey with Young Lives	24
Appendix B: Comparison of the Welfare Monitoring Survey with Young Lives	30
Appendix C: Non-parametric tests for comparison of distributions	32
Appendix D: Constructing a wealth index	34

# Executive summary

Young Lives is a longitudinal research project investigating the changing nature of childhood poverty. The study is tracking the development of 12,000 children in Ethiopia, Peru, India (Andhra Pradesh) and Vietnam through qualitative and quantitative research over a 15-year period. Since 2002, the study has been following two cohorts in each study country. The younger cohort consists of 2,000 children per study country aged between 6 and 18 months in 2002. The older cohort consists of 1,000 children per country aged between 7.5 and 8.5 in 2002. The key objectives of Young Lives are: (i) to improve the understanding of causes and consequences of childhood poverty, (ii) to inform the development and implementation of future policies and practices that will reduce childhood poverty.

The sampling methodology adopted by Young Lives is known as a sentinel site surveillance system. In Ethiopia, the Young Lives team used multi-stage, purposive and random sampling to select the two cohorts of children. This methodology randomised households within a study site while the sites themselves were chosen on the basis of predetermined criteria, informed by the Young Lives objectives. To ensure the sustainability of the study, and for resurveying purposes, a number of well-defined sites was chosen. The sites were selected with a pro-poor bias and to ensure a balanced representation of the Ethiopian regional diversity as well as rural/urban differences.

This paper assesses the sampling methodology by comparing the Young Lives sample with larger, nationally representative samples. In doing this, the Ethiopia team sought to:

- analyse how the Young Lives children and households compare with other children in Ethiopia in terms of their living standards and other characteristics
- examine whether this may affect inferences between the data
- establish to what extent the Young Lives sample is a relatively poorer or richer sub-population in Ethiopia
- determine whether different levels of living standards are represented within the dataset.

We used two nationally representative comparison samples, the Demographic and Health Survey 2000 (DHS) and the Welfare Monitoring Survey 2000 (WMS), and applied two different methodologies to assess the Young Lives sample. We first compared wealth index scores for the Young Lives households with those for DHS households. This provided a graphical illustration of the relative wealth of the Young Lives sample relative to the population of Ethiopia. We went on to use standard t-tests to test for statistical significance of the differences in several living standard indicators between Young Lives, the DHS, and the WMS samples. Finally, we investigated trends over time by comparing the Young Lives sample with the DHS 2000 and the DHS 2005. We compared variables that are common in the three surveys - area of residence, access to electricity and access to drinking water. In order to ensure comparability of the different samples we imposed constraints on the comparison samples to accommodate the fact that the Young Lives sample only includes households with at least one child aged between 6 and 18 months or aged 8. Second, the Young Lives sampling procedure selected sentinel sites to ensure a balanced representation of the Ethiopian regional diversity as well as rural/urban differences. This means that both rural and urban, and regional weights implied by the YL sample are not necessarily consistent with the true population weights.

We carried out our analysis at different levels of geographic aggregation. In particular, we focus our analysis on three different sets of tests. First, we compare Young Lives indicators against comparison sample indicators for the 'five Young Lives regions' only. Second, we carry out the tests for the rural and urban sub-samples separately. And finally, we carry out the tests for five different regional sub-samples, namely 'Addis Ababa urban', 'Amhara rural', 'Oromiya rural', 'Tigray rural' and 'SNNPR rural'. In the context of the Young Lives sampling procedure, we believe that the last set of regional tests provide the most accurate framework for comparison between the Young Lives sample and nationally representative samples.

Given the Young Lives sampling procedure, we nevertheless expected to find significant biases in the Young Lives sample. On the one hand, the study's aim to document childhood poverty implied that sentinel sites were sampled over-proportionally from food-deficient areas. On the other hand, budgeting constraints combined with the challenging topography of rural Ethiopia might have led to the selection of rural sentinel sites with relatively better transport communications and services. This is likely to have resulted in rural households being located in wealthier sites than the typical Ethiopian rural household.

Furthermore, the different survey years between the Young Lives sample and the DHS and WMS samples, implies that our analysis might be affected by improvements in the living standard across Ethiopia between year 2000 and 2002. We address this further concern, by carrying out comparison tests between the YL 2002 and the DHS 2005 sample.

We found that households in the Young Lives sample were slightly wealthier than households in the DHS sample. Further analysis revealed that households in rural areas and in urban areas, except Young Lives households in Addis Ababa, were wealthier. Young Lives households in Addis Ababa were poorer than households in the DHS sample. A similar picture emerged when we use t-tests to compare the means for a range of living standard indicators between the Young Lives and the DHS samples. Young Lives households in rural areas had better access to public services such as drinking water and electricity supply, while households in Addis Ababa had less access to basic services. These findings were supported by the comparison of common variables in Young Lives and the WMS. However, Young Lives households were less likely to own land or a house, and had smaller livestock holdings than WMS households.

To assess trends over time we compared the Young Lives sample with the DHS 2005 sample. Some of the differences, which we observed in the comparison of Young Lives with the DHS were reduced which indicates some improvements in living standards between 2000 and 2005.

The analyses show that households in the Young Lives sample were slightly better-off and had better access to basic services than the average household in Ethiopia, as measured by the nationally representative DHS and the WMS. However, our detailed analysis reveals that, while Young Lives households are located at sites with better access to services and utilities, they hold less land, less livestock. And are less likely to own their own house than the average Ethiopia household. This evidence is consistent with the sampling methodology applied with the Young Lives samples in Ethiopia.

Despite these biases, it is shown that the Ethiopian Young Lives sample covers the diversity of children in the country. Therefore, while not suited for simple monitoring of child outcome indicators, the Young Lives sample will be an appropriate and valuable instrument for analysing causal relations, modelling child welfare, and its longitudinal dynamics in Ethiopia.

# 1. Introduction

Young Lives is a longitudinal research project investigating the changing nature of childhood poverty. The study is tracking the development of 12,000 children in Ethiopia, Peru, India (Andhra Pradesh) and Vietnam through qualitative and quantitative research over a 15-year period. Since 2002, the study is following two cohorts in each study country. The younger cohort or 1-year-old cohort consists of 2,000 children per study country aged between 6 and 18 months in 2002. The older cohort or 8-year-old cohort consists of 1,000 children per country aged between 7.5 and 8.5 years in 2002. The key objectives of Young Lives are: (i) to improve the understanding of causes and consequences of childhood poverty, (ii) to inform the development and implementation of future policies and practices that will reduce childhood poverty.

The sampling methodology adopted for Ethiopia is known as sentinel site surveillance system. It consists of a multi-stage sampling procedure, whereby households within a sentinel site were selected randomly, while sentinel sites were chosen on basis of a number of predetermined criteria, informed by the objectives of the study. Specifically, in the Ethiopian context sentinel sites have been selected so as to ensure that (i) the cultural and geographic diversity of the country is reflected in the sample, (ii) the urban and rural differences are captured, and (iii) the pro-poor bias of the project is fulfilled.

The aim of this report is to assess the sampling methodology by comparing the samples with two larger, nationally representative samples. This analysis serves two main purposes. First, to analyse how the Young Lives children and households compare with children in Ethiopia, in terms of their living standards and other characteristics. Second, to examine whether and how this may affect inferences between the data. We will establish to what extent the Young Lives sample is a relatively poorer or richer Ethiopian sub-population, and whether different levels of living standards are represented in the dataset. The findings can provide guidance for interpreting research that uses Young Lives data.

Our analysis will compare a number of living standard indicators from the Young Lives sample with equivalent indicators from two different nationally representative samples: the Demographic and Health Survey 2000 (DHS) (CSA 2001a) and the 2000 Welfare Monitoring Survey (WMS) (CSA 2001b).

Based on the Young Lives sampling procedure we expect to find significant biases in the Young Lives sample. The aim of Young Lives is to document child poverty. Therefore, over-sampling of poor sites mainly in food-deficient areas took place. However, budgeting constraints and concerns regarding the long-term sustainability of the study meant that the rural sites selected were located in relatively better accessible areas. This is likely to have resulted in Young Lives rural households being located in wealthier sites than the typical Ethiopian rural household.

The report is structured as follows. In Section 2, we provide a brief overview of Ethiopia. Section 3, describes the Young Lives sampling approach. Section 4, presents the methodology used in our analysis. In Section 5, we briefly describe the two comparison samples. Section 6 and 7 discuss the main results from our comparison exercise. Section 8 analyses the effects of secular trends using information from the DHS 2005 sample. Finally, Section 9 concludes.

## 2. Overview of Ethiopia

Ethiopia has an estimated population of 74.3 million (2007), with about 86 per cent living in rural areas (CSA 2007). The country is characterised by its ethnic and geographic diversity. About 70 ethnic groups of highly variable sizes are spread all over the country. The two largest ethnic groups are the Oromo (32 per cent) and the Amhara (30 per cent), followed by the Tigray (6 per cent) and the Somali (6 per cent). The country is divided into nine ethnic-based regions and two cities, from which Oromiya, Amhara and SNNP (Southern Nations, Nationalities and People's region) account for more than 80 per cent of the total population (Table 1). The degree of ethnicity concentration in each region varies from very high (in Oromiya and Amhara) to low (in SNNP and Addis Ababa).

**Table 1.** *Education and population statistics (total and urban) of Ethiopia, by region*

Region	Primary school attendance (%)	% Urban	Population	
			Total (in mill.)	% of total
Tigray	51	19	4.3	6
Affar	15	9	1.4	2
Amhara	50	11	19.1	26
Oromiya	43	13	26.6	36
Somali	14	17	4.3	6
Benishangul-Gumuz	48	10	0.6	1
SNNP	35	9	14.1	19
Gambela	42	19	0.2	0
Harari	54	62	0.2	0
Addis Ababa	81	100	3.0	4
Dire Dawa	55	74	0.4	1
Total	42	14	74.3	100

Source: CSA 2007 and CSA 2006

The central area of Ethiopia is on a high plateau, with elevations that vary from 1,800 to 3,000 metres above sea level and some mountains reaching 4,620 meters (Mol 1960). Due to its topography, Ethiopia has a variety of agro-ecological conditions between and within regions and this, in turn, conditions the economic activities, from farming in the highlands to herding in the lowlands.

Agriculture is the main economic activity, accounting for 54 per cent of the Gross Domestic Product (GDP) and employing about 80 per cent of the population (CSA 2006). Ethiopia is one of the least developed countries in the world with a Human Development Index (HDI) rank of 170 (out of 177 countries) and with a Gross National Income (GNI) per capita of US\$110 in 2004 (World Bank 2006).

## 3. Young Lives sampling procedure

### 3.1 Sampling methodology

In 2002, the research team in Ethiopia selected a cohort of 2,000 children aged between 6 and 18 months and a control cohort of 1,000 children aged between 7.5 and 8.5 years. Using a methodology known as sentinel site surveillance system and according to the main guidelines of the project, the team first selected 20 sentinel sites across the country. Then they randomly selected 100 households with a 1-year-old child and 50 households with 8-year-old child in each site (Alemu et al. 2003). The selection of the sentinel sites followed a purposive sampling strategy, whereas the household selection within each sentinel site was done using simple random sampling.<sup>1</sup>

The methodology in the first stage was purposive because the sentinel sites were chosen such that: (i) poor areas - and, in particular, areas with food deficiency - were over-sampled; (ii) selected sites captured Ethiopia's diversity across regions and ethnicities, in both urban and rural areas; (iii) the costs of sampling were manageable, which reduced the probability of selecting sentinel sites in remote areas.

Based on these criteria, the selection procedure was as follow (Alemu et al. 2003):

#### First stage: selection of sites

- Selection of five regions out of a total of nine. The main criterion for selection was national coverage. The five selected regions (Addis Ababa, Amhara, Oromiya, SNNP, and Tigray) account for 96 per cent of the national population.
- Selection of three to five districts (*weredas*) in each region (20 districts in total) with a balanced representation of rural-poor, urban-poor, and relatively less poor rural and urban households. Due to lack of official statistics, the classification and selection was made through consultation with local officials in each district.
- Selection of at least one *peasant association* (in rural areas) or *kebele* (the lowest level of administration in urban areas) in each district. Districts themselves are too wide in terms of population and extension to be considered as sentinel sites. A *peasant association* or *kebele* was considered as a sentinel site when it was possible to find at least 100 households with a 1-year-old child and 50 households with an 8-year-old child. If there were not enough households to fulfil the criteria the *peasant association* or the *kebele* was considered as centre point around which the sentinel site was established. This happened in five cases (Young Lives 2007).

#### Second stage: selection of households

- A village within each sentinel site was randomly selected and all the households on the periphery were interviewed until 150 eligible households were located.

## 4. Methodological issues

In this section, we describe the methodology used in this report to compare the Young Lives samples with the alternative samples. We aim to assess the samples used in the collection of the Young Lives dataset, and thereby determine how the living standard of a Young Lives household compares with the living standard of a representative Ethiopian household.

We compare a number of living standard indicators from the Young Lives sample with comparable indicators from two nationally representative Ethiopian samples, namely the Demographic and Health Survey (DHS) and the Welfare Monitoring Sample (WMS), both collected in 2000.

We use two complementary methodologies in our analysis. First, we plot the wealth index for Young Lives households against the wealth index of DHS households. This analysis provides a graphical illustration of the socioeconomic status of Young Lives households relative to a representative Ethiopian household. Second, we use appropriate statistical tools to test whether Young Lives living standard indicators are consistent with the indicators computed from nationally representative samples.

Finally, we restrict our statistical analysis to the 1-year-old cohort only. The 8-year-old cohort is a smaller sample, which would make the tests unreliable when applied to regional sub-samples.

### 4.1 Methodological tools

#### 4.1.1 *Wealth index plots*

The Wealth index is the primary instrument used in the Young Lives survey to measure the socioeconomic status of households. It produces values between 0 and 1, whereby a higher wealth index indicates a higher socioeconomic status. It is computed as the simple average of three individual indexes that also range between 0 and 1: housing quality, consumer durables and access to services.<sup>2</sup>

We calculate a comparable version of the wealth index for Young Lives and the DHS surveys (DHS 2000 and DHS 2005)<sup>3</sup> and compare the cumulative distribution functions (cdf) graphically. This is not a standard comparison test but it provides us with an initial analysis of the differences between the surveys.

The logic of this procedure is that, for a pair of samples A and B selected from the same population if the wealth index cdf of sample A is always above the cdf of sample B (first order dominance), it would indicate that sample B consists of a wealthier population than sample A.

#### 4.1.2 *Tools for comparison of indicators*

We want to investigate whether Young Lives indicators are consistent with indicators from nationally representative samples. We do this by testing for similarity of the means of a set of individual variables and other moments of the distribution from the Young Lives and the comparison samples. The statistical test chosen depends on characteristics of the variable (whether it is discrete or continuous and, if discrete, whether it is dichotomic or non-

---

<sup>2</sup> See Appendix D for a detailed description of the calculation of the Wealth index.

<sup>3</sup> The DHS 2000 does not include some of the household characteristics used to define the wealth index in Young Lives Round 1 survey. Therefore, the wealth index calculated for our comparison includes different variables. See Appendix D for details.

dichotomic), on the moments of the distributions that are compared, and on the objectives of the comparison.

For binary variables<sup>4</sup> standard t-tests can be used to test whether the means of the Young Lives sample and the comparison samples are statistically different from each other. In this case, a t-test also tests whether the Young Lives and the comparison samples have the same distribution.<sup>5</sup> For discrete variables<sup>6</sup> we use the Wilcoxon-Mann-Whitney test for ordered discrete variables and the Chi-Square test for independence for non-ordered discrete variables. For instance, for an ordered variable such as mothers' years of education, we use the Wilcoxon-Mann-Whitney test to investigate whether mothers sampled in Young Lives have on average more years of education than mothers from the comparison samples. For a non-ordered variable such as type of cooking fuel, we use the Chi-Square test for independence to test whether the use of different types of cooking fuel is statistically similar between the two samples.

Finally, in the case of continuous variables<sup>7</sup> we use the Kolmogorov-Smirnov test to test whether two distributions differ.<sup>8</sup>

Although we use all statistical tests previously mentioned in our analysis, we rely heavily on the results provided by standard t-tests as most of the comparable variables are either binary or can be easily transformed into binary variables. Transforming a categorical variable into a set of binary variables provides a more intuitive set of tests, and allows us to identify the source of the difference in cases when the null hypothesis of the Chi-Square test is rejected.

## 4.2 Ensuring comparability

To test for equality of two distributions, it is crucial to ensure that the samples used are comparable. Lack of comparability will lead to invalid tests. In the context of the Young Lives sample, two issues are of particular concern: (i) the demographic characteristics of the Young Lives sample; (ii) the geographic distribution of Young Lives sentinel sites.

### 4.2.1 Demographic comparability

The Young Lives sample only includes households with a child aged between 6 and 18 months or between 7.5 and 8.5 years in 2002. This selection imposed a demographic restriction on the sample.

For comparability purposes, we restrict the DHS and the WMS samples to include only households that fulfil the demographic restriction of Young Lives. For complete comparability we narrow the sample down to only include women that are the biological mother of at least one child aged between 6 and 18 months in the DHS and WHS. Moreover, we restrict the Young Lives sample to include only caregivers that are the biological mother of the child.<sup>9</sup>

---

<sup>4</sup> A binary variable is a variable that has only two possible values, e.g. whether a household has electricity supply is either yes or no.

<sup>5</sup> This is because a binary variable can be understood as a realisation of a Bernoulli distribution (discrete probability distribution) whose only parameter, the probability of success, is equal to the mean.

<sup>6</sup> A discrete variable is a variable that can only take a limited number of discrete values that are usually whole numbers, e.g. number of children in the household.

<sup>7</sup> A continuous variable is a variable that can take values on a continuous scale, e.g. Wealth index.

<sup>8</sup> See Appendix C for a technical description of each test.

<sup>9</sup> Since we do not carry out tests for the 8-year-old cohort, we do not restrict our sample for this cohort.

#### 4.2.2 *Geographic comparability*

The sentinel sites were selected to capture regional diversity as well as urban-rural differences. However, the regional and urban-rural sampling weights implied by Young Lives will inevitably differ from the nationally representative weights. A comparison of the samples would therefore be invalid since any uncovered discrepancies might be due to the artificial Young Lives sampling weights.

We address this by carrying out our analysis at different levels of geographic aggregation. In particular, we define the following levels of aggregation:

- **Level A; national level with sampling weights.** We compare the living standard indicators for the different samples at the national level. The nationally representative samples are weighted with the national sampling weights. This methodology is akin to assume that Young Lives regional and urban-rural weights are consistent with the nationally representative weights.<sup>10</sup>
- **Level B; national level restricted to Young Lives regions. The** DHS and WMS samples are restricted to include only the five Ethiopian regions that were included in the Young Lives sample. Non- Young Lives regions are excluded.
- **Level C; urban-rural level.** We compare the Young Lives sample with the DHS and WMS samples for rural and urban sites separately. This methodology will abandon the artificial Young Lives urban-rural weights, but might still suffer from regional biases.
- **Level D; urban-rural and regional level.** The samples are compared for the following geographic areas: rural Amhara, rural Oromiya, rural Tigray, rural SNNP and urban Addis Ababa. We do not do a similar exercise for the urban sub-samples of Amhara, Oromiya, Tigray, and SNNP because of the small sample size of these groups (less than 100 households each).

The disaggregation into region-rural areas enables us to avoid the use of the artificial Young Lives sampling weights. We believe that this methodology provides us with a meaningful comparison between the Young Lives living standard indicators and the comparison samples. However, this approach comes at an additional cost. Since neither the DHS nor the WMS samples are representative at such level of disaggregation, we can no longer speak in terms of comparisons between the Young Lives sample and the Ethiopian average.<sup>11</sup> Instead, our analysis only allows us to make statements regarding comparisons between the Young Lives means and the average of the respective samples. This raises the concern that the comparison samples are biased. In this context, the availability of a number of common indicators for both the DHS and WMS samples is helpful. It allows us to investigate whether tests from the DHS and WMS samples are consistent with each other. Finally, comparisons at a higher level of aggregation, while not accurate, are useful in uncovering the potential biases affecting the Young Lives sample at the national level.

---

<sup>10</sup> Or to assume that regional and rural/urban differences are random. However, if different regions and rural/urban areas are considered to have non-random properties, this comparison will be invalid.

<sup>11</sup> See Section 5 for a detailed description of the DHS and WMS samples and their levels of representativeness.

### 4.3 Other concerns

The methodology described suffers from a number of limitations that could render our statistical tests invalid. First, the disaggregating methodology could undermine the statistical power of our estimators. Dividing the data into smaller sub-samples increases the imprecision of the sample estimates, and therefore raises the probability of failing to reject the null hypothesis. We found this to be a major problem when using the much smaller 8-year-old cohort. Statistical tests became very unreliable, in particular, when testing at the rural/regional level of disaggregation. This was a major reason for restricting our analysis to the younger cohort.

Second, both comparison samples were collected in 2000 while the Young Lives sample was selected in 2002. General improvements in living standards over the two-year interim could have created biases that create the artificial impression that Young Lives households are wealthier than they really are. We address this concern in Section 8, where we present tests using the 2005 DHS sample.

Finally, Young Lives over-sampled poor sentinel sites in the study site selection. However, budgeting constraints and concerns regarding the long-term sustainability of the study meant that the selected rural sites were located in relatively more easily accessible areas.

## 5. Comparison samples

Theoretically, we could have compared Young Lives households with the census information. However, the last available census data are from 1994. Due to migratory movements and changes in the configuration of administrative units, this comparison would not have been adequate. Instead, we chose two nationally representative surveys that were as close to the Young Lives sample selection as possible: the Demographic and Health Survey 2000 (DHS) and the Welfare Monitoring Survey 2000 (WMS).

The reason why we use two surveys is that both present complementary information on different dimensions of well-being, which allows us to understand possible biases in the Young Lives sample better. For instance, the DHS provides information on childcare practices and the WMS delivers detailed information on socioeconomic characteristics of households.

Table 2 provides a list of variables used from each sample in the comparison exercise. There several indicators for which information are available from both the DHS and the WMS samples. These variables also allow us to check for consistencies as well as inconsistencies between the two comparable samples.

### 5.1 Demographic and Health Survey

The DHS is a nationally, regionally, urban-rural representative survey that provides data on maternal and child health, nutrition, fertility and mortality. 14,643 households with at least one woman aged between 15 and 49 years was interviewed (15,367 women in total).

For demographic comparability, we narrow the DHS sample to include only women that are the biological mother of at least one child aged between 6 and 18 months and reduce the Young Lives sample to consider only caregivers that are the biological mother of the child. For complete comparability, these sub-samples include only residents (persons who slept in the house but did not live there were excluded). The comparable sub-sample consists of 3,254 women at the national level and 2,286 women for the five regions sub-sample.

## 5.2 Welfare Monitoring Survey

The WMS is a nationally and regionally representative survey designed to monitor the social impacts of economic reforms. The survey is carried out on a regular basis and gathers information on household assets, access to facilities, education, health, nutritional status of children and female fertility. The 2000 WMS covers all eleven regional states of Ethiopia, and included 25,928 households, of which 66 per cent were located in rural areas.

For comparability with the 1-year-old cohort, we restrict the WMS sample to include only households with at least one child aged between 6 and 18 months whose biological mother is member of the household. After applying these demographic restrictions, the WMS sample is reduced to 3,856 households. The WMS does not collect information on caregivers of the children. Our analysis therefore assumes that the biological mother is the primary caregivers.

**Table 2.** *Comparison indicators, DHS and WMS*

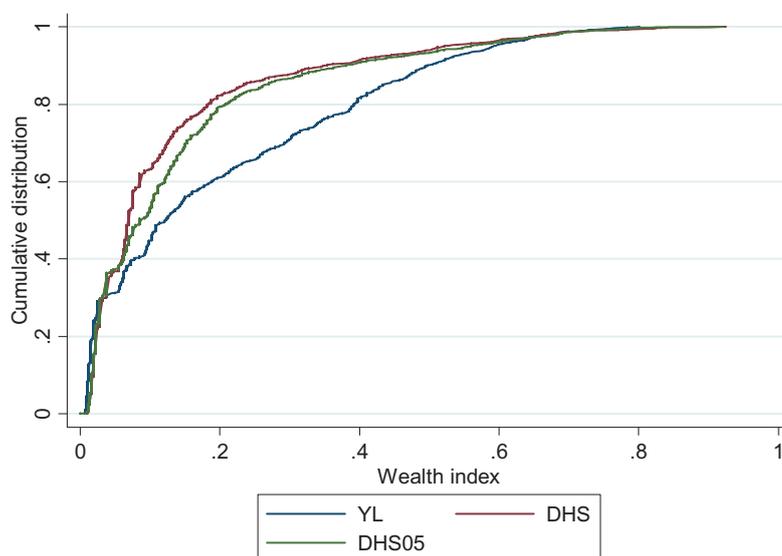
<b>Common Indicators (DHS and WMS)</b>	<b>WMS only</b>
<i>Household characteristics</i>	<i>Other household characteristics</i>
Household size	Land ownership
Electricity supply	Household ownership
Source of drinking water	Livestock ownership
Type of toilet facility	Number of cattle
Radio ownership	Number of rooms in the house
TV ownership	
Type of cooking fuel	<i>Child characteristics</i>
	Height-for-Age z-scores
<i>Characteristics of the head of household</i>	Weight-for-Height z-scores
Age	Measles and TB vaccination
Sex	
Level of education	
	<b>DHS only</b>
<i>Mother's background</i>	<i>Child care practices</i>
Literacy rate	Tetanus injections during pregnancy
Level of education	Antenatal care
Ethnic background	Place of delivery
Religion	Breastfeeding

## 6. Analysis of the wealth index

A graphical inspection of the wealth index cumulative distribution function (cdf) is useful to determine how much the socioeconomic status of households varies across a given sample. We can also use it to detect differences among our three samples: the Young Lives sample, the DHS 2000 sample and, for inter-temporal comparison, the DHS 2005 sample (Figure 1).<sup>12</sup> The Kolmogorov-Smirnov test shows that the wealth index cdf of Young Lives differs from the DHS 2000 one ( $p$ -value  $< 0.05$ ).<sup>13</sup> This result is shown in Figure 1. For the poorest 20 per cent of households, the wealth index has similar values for Young Lives and DHS 2000. However, when we consider the poorest 60 per cent of households, all Young Lives households report a wealth index below 0.2 compared with 0.1 in the DHS 2000. That is, the poorest 60 per cent of Young Lives households are richer than the same group of households in DHS 2000. The same is true when comparing Young Lives with the DHS 2005, but the difference is smaller due to improvements in living standards between 2000 and 2005 (CSA 2006). For further comparison with the DHS 2005, see Section 8.

Since there is no first order dominance of the DHS 2000 over Young Lives,<sup>14</sup> we cannot conclude that the Young Lives sample is richer than the DHS sample. However, the figure shows us that, if there is a difference, it is likely that Young Lives households are richer than DHS households and that this seems to be true for a large range of households.

**Figure 1.** *Wealth index cumulative distributions functions, Young Lives 2002, DHS 2000 and 2005, household-to-household*



<sup>12</sup> Both DHS 2000 and 2005 correspond to the national sample and included mothers with at least one child aged between 6 and 18 months at the time of the interview.

<sup>13</sup> The null hypothesis of equality of distributions is rejected. See Appendix A.

<sup>14</sup> Young Lives cdf is not always below the DHS cdf. For instance, both surveys capture a similar percentage of households with very low Wealth index values.

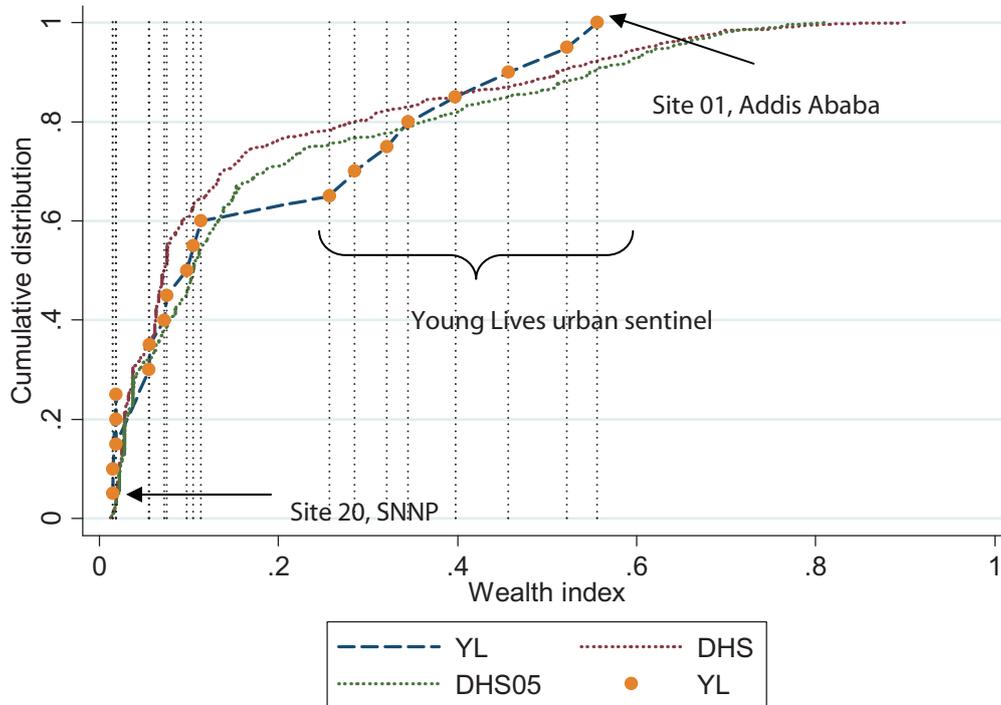
**Table 3.** *Wealth index by Young Lives sentinel site*

Sentinel Site Code	Region	Setting	Wealth Index Median
01	Addis Ababa	urban	0.427
02	SNNP	urban	0.392
03	Addis Ababa	urban	0.386
04	Tigray	urban	0.291
05	Amhara	urban	0.289
06	Addis Ababa	urban	0.288
07	Oromiya	urban	0.286
08	SNNP	rural	0.185
09	Tigray	rural	0.178
10	Tigray	rural	0.111
11	Tigray	rural	0.094
12	Oromiya	rural	0.092
13	SNNP	rural	0.053
14	Oromiya	rural	0.040
15	Oromiya	rural	0.037
16	Amhara	rural	0.028
17	Amhara	rural	0.014
18	Amhara	rural	0.014
19	SNNP	rural	0.014
20	SNNP	rural	0.011
<b>Total</b>			<b>0.120</b>

In the previous discussion, we abstracted from differences between sentinel sites, regions, urban and rural areas. Yet we know that the urban-rural composition of the Young Lives sample differs from the DHS sample (see Section 4.2.2). Low wealth index values are more likely to be reported for households in rural areas (Table 3 and Figure 2).<sup>15</sup> All rural sentinel sites have a median wealth index below 0.2 and all urban sites report a median wealth index above 0.2.

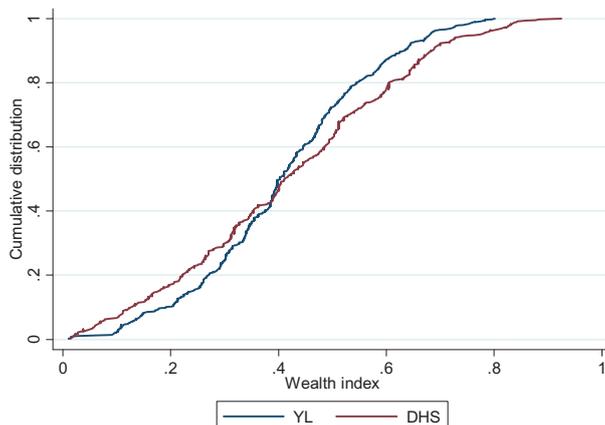
<sup>15</sup> Site-to-site means to compare the cdf of Young Lives sentinel sites against DHS sites (we grouped DHS households by census enumeration areas). Each observation represents a wealth index median in a site.

**Figure 2.** *Wealth index cumulative distribution functions, Young Lives 2002, DHS 2000 and 2005, site-to-site*

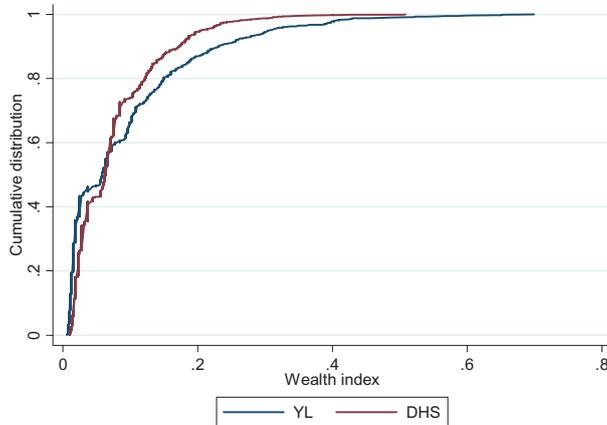


When reporting the cdf separately for urban and rural areas, differences between Young Lives and the DHS remain ( $p$ -value  $< 0.05$  for Kolmogorov-Smirnov test), but for different reasons. Figure 3 presents the wealth index cdf of urban sites for Young Lives and the DHS. It suggests that urban sentinel sites are poorer compared with the DHS sample. Figure 4 shows rural sites. We can see that the wealth indices of Young Lives households are more narrowly spread compared with DHS households. Analysis in section 7 shows that the graphical illustration masks substantial differences between Addis Ababa and other urban regions. In particular, we show that for Addis Ababa, Young Lives sampled poorer households compared with DHS 2000.

**Figure 3.** *Urban Ethiopia, Young Lives 2002 and DHS 2000, household-to-household*



**Figure 4.** Rural Ethiopia, Young Lives 2002 and DHS 2000, household-to-household



## 7. Comparison of living standard indicators

In order to assess the relative living standard of Young Lives households and individuals in more detail, we use a range of indicators available in the Young Lives sample and the comparison samples. The following groups of variables are used in our analysis: (i) household characteristics, (ii) characteristics of the household head and mother's background, (iii) child care practices and (iv) individual child characteristics. Table 2 provides the complete list of variables used in the analysis.

### 7.1 Comparison with the Demographic and Health Survey

Appendix A presents the results for all tests carried out between Young Lives and the DHS for the samples with 1-year-old children. It presents means and p-values for dichotomous variables and p-values for the Chi-Square, Wilcoxon-Mann-Whitney and the Kolmogorov-Smirnov test. It also presents the results for different aggregation levels as described in Section 4. For convenience, we present a summarised version with some key results in Table 4 (national level, total urban and total rural) and Table 5 (Addis Ababa and rural areas of the other four regions). Thereby, we rely on t-tests for both binary variables and categorical variables that were transformed into binary values. We present unweighted and weighted DHS results for the whole country.<sup>16</sup>

At national level, a key aspect is the difference in the proportion of urban households in the Young Lives sample and the DHS sample (34 per cent versus 16 and 10 per cent in the unweighted and weighted sample, respectively, Table 4, column 1). This difference is partially driven by the artificial composition of urban and rural areas in the Young Lives sample. The higher proportion of urban households induces a bias in the comparison of access to services and goods that are easier to obtain in urban areas. Thus, the Young Lives sample reports better access to electricity, protected drinking water and improved toilet facilities compared with the DHS sample. Mothers are also better educated and report better

<sup>16</sup> By including the sampling weights, we account for the DHS sampling design.

childcare practices. In all cases, differences are statistically significant at 5 per cent level. These results do not change when DHS sampling weights are included or the DHS sample is restrained to the five regions of Young Lives (Table 4, column 2).

To avoid the use of an artificial urban-rural composition we report results for urban Ethiopia (Column 3) and rural Ethiopia (Column 4) separately. In rural Ethiopia, we see that some differences remain. People in the rural Young Lives sample have better access to electricity, protected water and toilet facilities. For instance, 35 per cent of Young Lives households in rural Ethiopia have access to protected sources of drinking water compared with 16 per cent in rural DHS households. Mothers are also better educated in the Young Lives sample.

As for urban areas, results are less obvious. Compared with the DHS, people in the urban Young Lives sample have better access to electricity, toilet facilities and mothers report higher vaccination coverage. However, at the same time, urban Young Lives sentinel sites have worse access to drinking water and mothers are less educated. These results seem to contradict each other. It is possible that these results are driven by Young Lives artificial regional weighting (in particular, the weighting given to Addis Ababa compared with other urban areas).

**Table 4.** *DHS versus Young Lives, selected indicators*

Variables	National		Five regions		Total urban		Total rural		
	YL	DHS unw.	DHS w. YL	DHS	YL	DHS	YL	DHS	
<b>Household characteristics</b>									
Urban household	0.34	0.16 *	0.10 *	0.34	0.16 *	1.00	1.00	0.00	0.00
Electricity supply	0.34	0.13 *	0.07 *	0.34	0.14 *	0.89	0.77 *	0.06	0.01 *
Source of drinking water									
Piped into dwelling/yard	0.11	0.06 *	0.02 *	0.11	0.07 *	0.30	0.38 *	0.02	0.00 *
Protected public source	0.41	0.22 *	0.18	0.41	0.21 *	0.53	0.52	0.35	0.16 *
Unprotected source	0.48	0.72 *	0.80 *	0.48	0.72 *	0.17	0.11 *	0.63	0.84 *
<b>Mother's background</b>									
Literacy rate									
Easily	0.27	0.15 *	0.13 *	0.27	0.16 *	0.54	0.54	0.13	0.08 *
With difficulty	0.12	0.06 *	0.07 *	0.12	0.06 *	0.14	0.09 *	0.11	0.05 *
Not at all	0.61	0.79 *	0.80 *	0.61	0.77 *	0.32	0.38 *	0.76	0.87 *
<b>Child care practices</b>									
At least 2 injections of Tetanus during pregnancy	0.86	0.64 *	0.62 *	0.86	0.64 *	0.88	0.72 *	0.84	0.60 *
<b>Sample size</b>	1,924	3,113	3,112	1,924	2,286	647	518	1,277	2,736

Note: (\*) means the null hypothesis that the difference between both means is zero is rejected at the 5 % significance level.

**Table 5.** *DHS versus Young Lives, selected indicators by region*

Variables	Addis Ababa		Rural Amhara		Rural Oromiya		Rural SNNP		Rural Tigray	
	YL	DHS	YL	DHS	YL	DHS	YL	DHS	YL	DHS
<b>Household characteristics</b>										
Urban household	1.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Electricity supply	0.99	0.97	0.00	0.00	0.07	0.00 *	0.14	0.01 *	0.01	0.00
Source of drinking water										
Piped into dwelling/yard	0.34	0.61 *	0.01	0.00	0.00	0.00	0.05	0.00 *	0.00	0.00
Protected public source	0.51	0.00 *	0.49	0.09 *	0.40	0.13 *	0.15	0.14	0.44	0.28 *
Unprotected source	0.15	0.37 *	0.50	0.91 *	0.60	0.87 *	0.80	0.86 *	0.56	0.72 *
<b>Mother's background</b>										
Literacy rate										
Easily	0.66	0.64	0.10	0.0	0.16	0.09 *	0.17	0.10 *	0.09	0.03 *
With difficulty	0.14	0.14	0.09	0.06	0.16	0.07 *	0.12	0.07 *	0.04	0.02
Not at all	0.20	0.21	0.81	0.84	0.68	0.84 *	0.70	0.83 *	0.86	0.94 *
<b>Child care practices</b>										
At least 2 injections of Tetanus during pregnancy	0.95	0.75	0.29	0.12 *	0.39	0.25 *	0.37	0.26 *	0.52	0.30 *
<b>Sample size</b>	274	156	294	452	292	639	397	472	294	361

Note: (\*) means the null hypothesis that the difference between both means is zero is rejected at the 5 % significance level.

Differences due to Young Lives implicit regional weighting are also suggested by differences in the representation of major ethnic groups in Young Lives compared to the DHS sample. For instance, Oromo people are under-sampled in Young Lives according to the five-region sample (see Appendix A). Some of these differences remain after reporting urban and rural areas separately. To account for this, Appendix A presents disaggregated results for the following sub-groups: urban Addis Ababa, rural Amhara, rural Oromiya, rural SNNP and rural Tigray. This sub-sampling increases the comparability, but it also results in a loss of representativeness of the DHS sample at the urban-region and at the rural-region level. This means that we are no longer comparing the Young Lives sample with the Ethiopian average, but with the DHS average.

While Young Lives seems to under-sample poor households in rural Amhara, rural Oromiya, rural SNNP and rural Tigray, poor households are over-sampled in urban Addis Ababa compared with the DHS sample. For instance, in only 34 per cent of the Young Lives' households in Addis Ababa water is piped into the dwelling, compared with 61 per cent in DHS households. Furthermore, Young Lives in Addis Ababa reports worse toilet facilities, less ownership of a TV and radio, and higher use of wood in detriment of kerosene as cooking fuel than households in the DHS sample.

It is worth mentioning that there are also differences in the representation of religious groups between both samples at this level of disaggregation. In rural Amhara, rural Oromiya and rural Tigray, Young Lives over-sampled households with Orthodox mothers rather than Muslim mothers. The opposite happened in SNNP, where the Young Lives sample includes a higher proportion of Muslim mothers compared with the DHS sample. If differences in living standards are correlated with religious background, this might also help to explain part of the remaining differences between both samples.

## 7.2 Comparison with the Welfare Monitoring Survey

The WMS sample provides some common variables with the DHS that allow us to check the robustness of the results obtained by the DHS sample. The WMS sample also provides additional information not available in the DHS. In particular, we will use data on household assets and child anthropometrics as additional measures of wellbeing.

The complete comparisons exercise between the WMS and Young Lives is reported in Appendix B. We report results for t-tests for equality of means, as well as where appropriate Chi-Square, Wilcoxon-Mann-Whitney and the Kolmogorov-Smirnov tests. As before, we report our analysis at different levels of aggregation.

### Common indicators

Table 6 presents the results of the comparison tests between the WMS and the Young Lives sample indicators common to the DHS sample. The table reports test results for rural Tigray, rural Amhara, rural Oromiya and rural SNNP, and urban Addis Ababa. The results are similar to the Young Lives versus DHS comparisons reported in Table 5 and in Appendix A.<sup>17</sup> The results lend additional support to our analysis and dispel some concerns regarding the possibility that the comparison samples might include biases.<sup>18</sup>

From the comparison exercise, we find that Young Lives households in rural areas are better off than WMS households are in terms of access to electricity, protected water and toilet facilities. This indicates that Young Lives sites are located in areas that have better access to services. On the other hand, Young Lives households in urban Addis Ababa appear to be located in worse serviced areas than WMS households are. Similar differences across region/rural sub-samples are found in terms of education of the household head.

**Table 6.** *WMS versus Young Lives, selected common indicators by region*

Variables	Rural Tigray		Rural Amhara		Rural Oromiya		Rural SNNP		Addis Ababa	
	YL	DHS	YL	DHS	YL	DHS	YL	DHS	YL	DHS
<b>Household characteristics</b>										
Average household size <sup>a</sup>	5.88	5.75	5.22	5.45	5.66	5.86 *	5.94	5.80	5.74	6.27
Electricity supply	0.01	0.00 *	0.00	0.01	0.07	0.01 ***	0.14	0.01 ***	0.99	0.97
Source of drinking water <sup>b</sup>										
Piped into dwelling/yard	0.00	0.00	0.01	0.00	0.00	0.00	0.05	0.00 ***	0.34	0.56 ***
Protected public source	0.44	0.19 ***	0.49	0.09 ***	0.40	0.12 ***	0.15	0.14	0.51	0.43
Unprotected source	0.56	0.81 ***	0.50	0.91 ***	0.60	0.86 ***	0.80	0.85 **	0.14	0.01 ***
Toilet facility <sup>b</sup>										
At home	0.29	0.05 ***	0.04	0.01 **	0.09	0.08	0.16	0.16	0.17	0.35 ***
Outside home	0.06	0.00 ***	0.06	0.00 ***	0.01	0.02	0.04	0.03	0.70	0.52 ***
None	0.65	0.95 ***	0.90	0.98 ***	0.89	0.89	0.81	0.81	0.12	0.13

<sup>17</sup> In this section, we only report tests for the regional sub-samples. We refer to Appendix B for the results of the comparisons for the national and rural/urban samples.

<sup>18</sup> We are concerned with potential biases in the comparison samples because they are no longer representative when used at the rural/regional level of disaggregation. The fact that both samples provide similar results indicates that biases might be small. Another concern could be that the comparison samples are both biased in the same direction.

**Table 6.** *WMS versus Young Lives, selected common indicators by region continued*

Variables	Rural Tigray		Rural Amhara		Rural Oromiya		Rural SNNP		Addis Ababa	
	YL	DHS	YL	DHS	YL	DHS	YL	DHS	YL	DHS
<b>Household head characteristics</b>										
Sex (male)	0.87	0.77 ***	0.84	0.94 ***	0.93	0.88 ***	0.96	0.86 ***	0.78	0.73
Average age	38.37	41.26 ***	36.75	39.06 ***	36.28	36.33	35.81	35.84	38.50	40.82
Level of education <sup>a</sup>	0.27	0.14 **	0.24	0.17 **	0.72	0.43 ***	0.92	0.49 ***	1.60	1.69
None	0.80	0.87 **	0.82	0.88 **	0.49	0.70 ***	0.44	0.66 ***	0.21	0.28
Incomplete primary	0.17	0.12	0.13	0.08 **	0.33	0.18 ***	0.25	0.20 **	0.22	0.13 **
Complete primary	0.02	0.01	0.04	0.04	0.15	0.10 **	0.26	0.13 ***	0.39	0.31
Secondary	0.00	0.00	0.01	0.00	0.02	0.01	0.03	0.00 ***	0.12	0.18
Higher/Other	0.01	0.00 *	0.00	0.00	0.00	0.00	0.02	0.00 *	0.06	0.10
<b>Sample size</b>	294	203	296	513	295	834	397	968	287	89

Note: (\*) indicates that differences between Young Lives and the WMS are significant at 10%, (\*\*) at 5%, (\*\*\*) at 1% level.

a reported tests correspond to Wilcoxon-Mann-Witney test.

b reported tests correspond to Chi-Square test.

**Table 7.** *WMS versus Young Lives, household assets and child characteristics by region*

Variables	Rural Tigray		Rural Amhara		Rural Oromiya		Rural SNNP		Addis Ababa	
	YL	DHS	YL	DHS	YL	DHS	YL	DHS	YL	DHS
<b>Household assets</b>										
Household owns land	0.87	0.94 ***	0.90	0.98 ***	0.85	0.99 ***	0.80	0.99 ***	0.00	0.39 ***
Household owns house	0.71	0.88 ***	0.81	0.91 ***	0.88	0.92 *	0.90	0.93 **	0.12	0.31 ***
Household owns livestock	0.84	0.95 ***	0.80	0.95 ***	0.75	0.91 ***	0.76	0.89 ***	0.27	0.11 ***
Number of cattle owned <sup>a</sup>	1.49	3.81 ***	2.23	3.06 ***	1.79	3.74 ***	2.09	3.20 ***	0.03	0.10
Number of rooms <sup>a</sup>	1.40	1.62 ***	1.40	1.53 ***	1.26	2.01 ***	1.35	1.55 ***	1.52	2.94 ***
One room	0.68	0.57 **	0.73	0.63 ***	0.78	0.57 ***	0.75	0.65 ***	0.60	0.33 ***
Two rooms	0.25	0.26	0.17	0.26 ***	0.20	0.33 ***	0.16	0.27 ***	0.30	0.36
Three rooms	0.07	0.14 **	0.07	0.08	0.02	0.08 ***	0.08	0.05	0.08	0.16 *
Four rooms	0.00	0.02 **	0.03	0.03	0.01	0.01	0.01	0.02	0.01	0.06
<b>Child characteristics</b>										
Sex (male)	0.55	0.47 *	0.58	0.45 ***	0.53	0.49	0.48	0.51	0.54	0.53
Average age	11.45	11.40	11.82	11.34 *	12.13	11.49 ***	11.35	11.52	11.98	11.00 **
Height (cm)	70.93	66.94 ***	69.69	66.74 ***	70.07	68.37 ***	70.73	68.15 ***	72.25	68.49 ***
Weight (kg)	8.01	7.19 ***	7.63	7.27 ***	7.61	7.71 ***	7.63	7.67 ***	8.66	8.22 ***
Height-for-Age z-score	-1.08	-2.16 ***	-1.89	-2.48 ***	-1.74	-1.99 *	-1.22	-2.06 ***	-0.91	-1.40 **
Weight-for-Height z-score	-0.72	-0.60	-0.88	-0.50 ***	-1.03	-0.32 ***	-1.09	-0.37 ***	-0.17	-0.05
Measles vaccination	0.84	0.77 **	0.39	0.39	0.49	0.35 ***	0.49	0.38 ***	0.73	0.86 ***
TB vaccination	0.82	0.79	0.73	0.44 ***	0.61	0.42 **	0.60	0.45 ***	0.98	0.99
<b>Sample size</b>	294	203	296	513	295	834	397	968	287	89

Note: (\*) indicates that differences between Young Lives and the WMS are significant at 10%, (\*\*) at 5%, (\*\*\*) at 1% level.

a reported tests correspond to Wilcoxon-Mann-Witney test.

### **Additional indicators**

The WMS results presented support the conclusions from the analysis of the DHS sample that, in rural areas, richer household have been over-sampled. This appears to be consistent with our earlier concerns of potential biases due to selection of better accessible sentinel sites.

However, we should note that the indicators on which we base these conclusions, namely the wealth index, and access to services such as water, electricity and education, are likely to be highly correlated. They might therefore provide a misleading assessment of the living standard and the economic wellbeing of individual households.<sup>19</sup>

The WMS sample allows us to address these shortcomings when using information on individual household assets and child anthropometric outcomes. These groups of variables will provide a more accurate measure of the fortunes and misfortunes of individual households. Table 7 shows the results for the statistical tests on the additional indicators by region/rural sub-samples.

Analysis of the household asset variables provides very interesting results. In contrast to earlier indications, we find that Young Lives households, both in rural regions and urban Addis Ababa, tend to be poorer – in terms of their asset holdings – than the WMS average. Young Lives households are less likely to own a house, land or livestock, have less cattle and live in houses with fewer rooms. These differences are not only statistically significant, but can be large in magnitude. For example, the number of cattle held by Young Lives households is half the number held by households in the region/rural WMS sample.

In conclusion, Young Lives households might have better access to services; nevertheless, they are poorer than the average household in the WMS sample. These results are consistent with the sampling methodology applied by Young Lives, whereby sites were chosen from better accessible areas, but were also selected for their poor and food-deficient status.

We now turn to the analysis of child characteristics. Young Lives children are more likely to be vaccinated against Measles and TB than WMS children. This is consistent with results from the DHS sample. Tests on child anthropometric provide a more complex picture. Young Lives children are more likely to be stunted compared with children from the WMS sample, but, rather unexpectedly, we also find that Young Lives children suffer less wasting. While the former is consistent with previous results and the common interpretation of stunting as a medium and long-term measure of nutritional deficiencies, the wasting discrepancies are more challenging. Although we have not found a fully satisfactory explanation, we note that 2000 was characterised by widespread growth across Ethiopia. It is therefore possible that, unlike for the 2002 Young Lives data, weight-for-height figures for WMS children measured in 2000 might not fully include the improved nutritional intake resulting from the 2000 harvests.

## 8. Assessing secular trends

In this section, we address the question of whether differences uncovered between the Young Lives and the comparison samples might be due to differences in the collection year. We investigate whether these differences could be due to economic improvements between 2000 and 2002. We include the DHS 2005 in the comparison at national level (Table 8, Appendix A). According to the DHS, two aspects have significantly improved between 2000 and 2005: (i) the proportion of households with access to protected sources of water, and (ii) the availability of toilet facilities at home.

Considering only these two variables, the Young Lives sample has better living standards than the DHS 2000 sample, but worse than the DHS 2005 sample. This shows that differences between Young Lives and DHS households could be due to different times of collection. Also note that since these two variables are included in the wealth index, these differences in the wealth index cdfs between Young Lives and the DHS are less when compared with the DHS 2005 (see Figure 1 in Section 6).

Apart from these two indicators, the DHS 2000 and the DHS 2005 do not differ significantly.

**Table 8.** *DHS 2000 and 2005 versus Young Lives, selected indicators*

Variables	National Sample			
	YL	2000 DHS unw.	DHS w.	2005 DHS w.
<b>Household characteristics</b>				
Urban household	0.34	0.16 *	0.10 *	0.08 *
Average household size	5.68	5.92 *	5.93 *	5.97 *
Electricity supply	0.34	0.16 *	0.10 *	0.08 *
Source of drinking water				
Piped into dwelling/yard	0.11	0.06 *	0.02 *	0.03 *
Protected public source	0.41	0.22 *	0.18 *	0.53 *
Unprotected source	0.48	0.72 *	0.80 *	0.44 *
<b>Mother's background</b>				
Literacy rate				
Easily	0.27	0.15 *	0.13 *	0.10 *
With difficulty	0.12	0.06 *	0.07 *	0.07 *
Not at all	0.61	0.79 *	0.80 *	0.83 *
<b>Pregnancy, delivery, breastfeeding</b>				
At least 2 injections of Tetanus during pregnancy	0.86	0.64 *	0.62 *	0.77 *
Number of antenatal visits	0.51	0.30 *	0.26 *	0.28 *
Breastfeeding	0.98	0.97 *	0.97	0.97
<b>Sample size</b>	1994	3113	3113	3229

## 9. Conclusion

Our assessment of the 2002 Young Lives Ethiopian sample indicates that the sample includes a wide range of living standards akin to the variability found in the Ethiopian population. On average Young Lives households appear to have higher wealth index values than the average Ethiopian household as measured by the nationally representative 2000 DHS sample. These differences might be the result of a higher proportion of urban sites in the Young Lives sample and secular time trends.

Further, our statistical tests indicate that households in Young Lives sites have better access to services and utilities, but are poorer in assets than the average Ethiopian household as measured by the WMS sample. This is consistent with the sampling methodology applied, whereby Young Lives sites are chosen from better accessible but poor and food-deficient areas.

However, even if poor sites have been over-sampled, the Young Lives sample covers the diversity of the children in the country and a wide variety of attributes and experiences. Therefore while not suited for monitoring child outcome indicators, the Ethiopian Young Lives sample will be an appropriate and valuable instrument for the analysis of causal relations and modelling of child welfare and its longitudinal dynamics.

# References

- Alemu T, G. Asgedom, L. Liebenberg, A. Mekonnen, J. Seager, B. Tefera and T. Woldehanna (2003) *Young Lives Preliminary Country Report: Ethiopia*, London: Young Lives
- Central Statistical Authority Ethiopia (CSA) and ORC Macro (2001a) *Ethiopia Demographic and Health Survey 2000*, Addis Ababa, Ethiopia and Calverton, MA CSA and ORC Macro
- Central Statistical Authority Ethiopia (CSA) (2001b) *Report on the 1999/2000 Welfare Monitoring Survey*, Vol. I and Vol. II, Addis Ababa: CSA
- Central Statistical Authority Ethiopia (CSA) and ORC Macro (2006) *Ethiopia Demographic and Health Survey 2005*, Addis Ababa, Ethiopia and Calverton, Maryland, USA: Central Statistical Authority and ORC Macro
- Central Statistical Agency of Ethiopia (CSA) (2007) 'The 2007 National Statistics Population', Addis Ababa
- Kanji G.K. (1993) *100 Statistical Tests*, London: Sage Publication
- Ministry of Information (Mol) (1960) *Ethiopia: Facts and Figures*, Addis Ababa: Imperial Ethiopian Government
- Wilson I., S.R.A Huttly and B. Fenn (2006) 'A Case Study of Sample Design for Longitudinal Research: Young Lives', *Int. J. Social Research Methodology* (9)3: 351-65
- World Bank (2006) 'World Development Indicators', [http://devdata.worldbank.org/wdi2006/contents/Table1\\_1.htm](http://devdata.worldbank.org/wdi2006/contents/Table1_1.htm) (accessed 8 March 2008)
- Young Lives (2007) 'The Young Lives Sampling Procedure in Ethiopia', [http://www.younglives.org.uk/pdf/ethiopia\\_sampling.pdf](http://www.younglives.org.uk/pdf/ethiopia_sampling.pdf) (accessed 10 March 2008)

# Appendix A

## Comparison of the Demographic and Health Survey with Young Lives

	National Sample				chi2 2000 Unw. (prob)	5 regions sample 2000		chi2 (prob)
	YL	DHS Unw.	DHS W.	2005 DHS W.		YL	DHS	
<b>Household characteristics</b>								
Urban households	0.34	0.16 *	0.10 *	0.08 *		0.34	0.16 *	
Average Household Size	5.68	5.92 *	5.93 *	5.97 *		5.68	5.89 *	
Electricity supply	0.34	0.13 *	0.07 *	0.08 *		0.34	0.14 *	
Source of drinking water					0.00			0.00
Piped into dwelling/yard/plot	0.11	0.06 *	0.02 *	0.03 *		0.11	0.07 *	
Protected public source	0.41	0.22 *	0.18 *	0.53 *		0.41	0.21 *	
Unprotected source	0.48	0.72 *	0.80 *	0.44 *		0.48	0.72 *	
Toilet facility					0.00			0.00
At home	0.21	0.10 *	0.09 *	0.25 *		0.21	0.09 *	
Outside home	0.16	0.10 *	0.06 *	0.09 *		0.16	0.10 *	
None	0.63	0.80 *	0.85 *	0.66		0.63	0.81 *	
Own radio	0.38	0.25 *	0.21 *	0.33 *		0.38	0.24 *	
Own TV	0.05	0.03 *	0.01 *	0.03 *		0.05	0.04 *	
Cooking fuel					0.00			0.00
Wood	0.64	0.79 *	0.80 *	0.89 *		0.64	0.74 *	
Kerosene/paraffin	0.08	0.06	0.02 *	0.01 *		0.08	0.08	
Charcoal	0.04	0.02 *	0.01 *	0.02 *		0.04	0.02 *	
Gas/electricity	0.00	0.00	0.00	0.00		0.00	0.00	
Coal	0.00	0.00	0.00	0.00		0.00	0.00	
Cow dung	0.23	0.12 *	0.17 *	0.08 *		0.23	0.16 *	
None	0.00	0.00	0.00	0.00		0.00	0.00	
Other	0.00	0.00 *	0.00 *	0.00 *		0.00	0.00 *	
<b>Characteristics of the head of the household</b>								
Average age	37.04	38.01 *	38.03 *	37.36		37.04	38.49 *	
Sex (male)	0.87	0.87	0.89	0.89 *		0.87	0.87	
<b>Mother's background</b>								
Literacy rate					0.00			0.00
Easily	0.27	0.15 *	0.13 *	0.10 *		0.27	0.16 *	
With difficulty	0.12	0.06 *	0.07 *	0.07 *		0.12	0.06 *	
Not at all	0.61	0.79 *	0.80 *	0.83 *		0.61	0.77 *	
Highest year of education								
None	0.58	0.67 *	0.66 *	0.58		0.58	0.65 *	
Elementary	0.35	0.26 *	0.30 *	0.39 *		0.35	0.28 *	
Secondary	0.06	0.05 *	0.03 *	0.02 *		0.06	0.05 *	
Higher	0.01	0.02 *	0.01	0.01		0.01	0.02 *	
Ethnic background					0.00			0.00
Other	0.04	0.24 *	0.12 *	0.16 *		0.04	0.10 *	
Agew	0.00	0.01 *	0.01 *	0.00 *		0.00	0.01 *	
Amhara	0.28	0.24 *	0.29	0.27		0.28	0.27	
Gurage	0.08	0.04 *	0.04 *	0.04 *		0.08	0.05 *	
Hadiva	0.05	0.01 *	0.01 *	0.02 *		0.05	0.01 *	
Kambata	0.00	0.01 *	0.01 *	0.01 *		0.00	0.01 *	
Oromo	0.21	0.31 *	0.38 *	0.38 *		0.21	0.34 *	
Sidama	0.06	0.03 *	0.05	0.05		0.06	0.05	
Tigran	0.22	0.10 *	0.06 *	0.06 *		0.22	0.13 *	
Wolavta	0.06	0.02 *	0.02 *	0.02 *		0.06	0.02 *	
Religion					0.00			0.00
Muslim	0.16	0.39 *	0.30 *	0.35 *		0.16	0.28 *	
Catholic	0.00	0.01	0.01	0.01 *		0.00	0.01	
Protestant	0.11	0.14 *	0.17 *	0.19 *		0.11	0.15 *	
Orthodox	0.71	0.43 *	0.48 *	0.43 *		0.71	0.53 *	
Evangelist	0.00	0.00 *	0.00 *	0.00 *		0.00	0.00 *	
Other	0.01	0.04 *	0.04 *	0.02 *		0.01	0.03 *	
<b>Pregnancy, delivery and breastfeeding</b>								
At least 2 inj. of tetanus d. pregn.	0.86	0.64 *	0.62 *	0.77 *		0.86	0.64 *	
Receiving antenatal care	0.51	0.30 *	0.26 *	0.28 *		0.51	0.31 *	
N° of antenatal visits d. pregnancy	2.13	1.13 *	0.85 *	1.00 *		2.13	1.19 *	
Place of delivery					0.00			0.00
Home	0.83	0.89 *	0.95 *	0.94 *		0.83	0.91 *	
Hospital	0.11	0.06 *	0.02 *	0.04 *		0.11	0.05 *	
Other health facility	0.06	0.04 *	0.03 *	0.02 *		0.06	0.04 *	
Breastfeeding	0.98	0.97 *	0.97	0.97		0.98	0.97 *	
<b>Sample size</b>	1924	3113	3113	3229		1924	2286	

\* P-value at the 5% of significance for the t-test. Null hypothesis: equality of means.

Chi2: Null hypothesis: independence between samples and categories.

AN ASSESSMENT OF THE YOUNG LIVES SAMPLING APPROACH IN ETHIOPIA

	Total Urban 2000		chi2 (prob)	Total Rural 2000		chi2 (prob)
	YL	DHS		YL	DHS	
<b>Household characteristics</b>						
Urban households	1.00	1.00		0.00	0.00	
Average Household Size	5.66	5.79		5.70	5.94 *	
Electricity supply	0.89	0.77 *		0.06	0.01 *	
Source of drinking water			0.00			0.00
Piped into dwelling/yard/plot	0.30	0.38 *		0.02	0.00 *	
Protected public source	0.53	0.52		0.35	0.16 *	
Unprotected source	0.17	0.11 *		0.63	0.84 *	
Toilet facility			0.03			0.00
At home	0.34	0.27 *		0.15	0.06 *	
Outside home	0.39	0.46 *		0.04	0.03	
None	0.27	0.26		0.81	0.90 *	
Own radio	0.64	0.69		0.24	0.16 *	
Own TV	0.14	0.21 *		0.00	0.00 *	
Cooking fuel			0.00			0.00
Wood	0.55	0.45 *		0.68	0.86 *	
Kerosene/paraffin	0.22	0.40 *		0.00	0.00	
Charcoal	0.12	0.12		0.00	0.00	
Gas/electricity	0.01	0.02		0.00	0.00	
Coal	0.00	0.00		0.00	0.00	
Cow dung	0.09	0.02 *		0.31	0.14 *	
None	0.00	0.00		0.00	0.00	
Other	0.01	0.00 *		0.00	0.00	
<b>Characteristics of the head of the household</b>						
Average age	37.74	37.33		36.69	38.14 *	
Sex (male)	0.79	0.77		0.91	0.88 *	
<b>Mother's background</b>						
Literacy rate			0.01			0.00
Easily	0.54	0.54		0.13	0.08 *	
With difficulty	0.14	0.09 *		0.11	0.05 *	
Not at all	0.32	0.38 *		0.76	0.87 *	
Highest year of education						
None	0.32	0.34		0.72	0.74	
Elementary	0.50	0.34 *		0.27	0.25	
Secondary	0.16	0.24 *		0.01	0.01	
Higher	0.02	0.08 *		0.00	0.01 *	
Ethnic background			0.00			0.00
Other	0.05	0.17 *		0.03	0.25 *	
Agew	0.00	0.00		0.00	0.01 *	
Amhara	0.31	0.40 *		0.27	0.21 *	
Gurage	0.08	0.09		0.08	0.03 *	
Hadiva	0.00	0.01		0.08	0.01 *	
Kambata	0.00	0.01		0.00	0.01 *	
Oromo	0.29	0.25		0.17	0.32 *	
Sidama	0.02	0.00 *		0.08	0.04 *	
Tigrrian	0.21	0.07 *		0.23	0.10 *	
Wolavta	0.04	0.00 *		0.07	0.02 *	
Religion			0.00			0.00
Muslim	0.10	0.32 *		0.20	0.40 *	
Catholic	0.01	0.00		0.00	0.01 *	
Protestant	0.06	0.09 *		0.13	0.15	
Orthodox	0.83	0.58 *		0.65	0.40 *	
Evangelist	0.00	0.00		0.00	0.00	
Other	0.00	0.01		0.01	0.04 *	
<b>Pregnancy, delivery and breastfeeding</b>						
At least 2 inj. of tetanus d. pregn.	0.88	0.72 *		0.84	0.60 *	
Receiving antenatal care	0.74	0.71		0.39	0.22 *	
N° of antenatal visits d. pregnancy	3.72	3.79		1.32	0.63 *	
Place of delivery			0.13			0.00
Home	0.55	0.50		0.97	0.97	
Hospital	0.28	0.34 *		0.02	0.01 *	
Other health facility	0.16	0.16		0.01	0.01	
Breastfeeding	0.97	0.97		0.99	0.97 *	
Sample size	647	518		1277	2736	

\* P-value at the 5% of significance for the t-test. Null hypothesis: equality of means

Chi2: Null hypothesis: independence between samples and categories.

	Addis Ababa			Rural Amhara		
	YL	DHS 2000	DHS 2005	YL	DHS 2000	DHS 2005
<b>Household characteristics</b>						
Urban households	1.00	1.00	1.00	0.00	0.00	0.00
Average Household Size	5.70	5.92	5.55	5.22	5.78 *	5.73 *
Electricity supply	0.99	0.97	0.96	0.00	0.00	0.00
Source of drinking water						
Piped into dwelling/yard/plot	0.34	0.61 *	0.67 *	0.01	0.00	0.00
Protected public source	0.51	0.00 *	0.32 *	0.49	0.09 *	0.56
Unprotected source	0.15	0.37 *	0.01 *	0.50	0.91 *	0.44
Toilet facility						
At home	0.17	0.04	0.27 *	0.04	0.01 *	0.21 *
Outside home	0.70	0.21 *	0.69	0.06	0.00 *	0.06
None	0.13	0.60	0.04 *	0.90	0.99 *	0.73 *
Own radio	0.71	0.83 *	0.89 *	0.10	0.05 *	0.24 *
Own TV	0.25	0.31	0.59 *	0.00	0.00	0.00
Cooking fuel						
Wood	0.24	0.09 *	0.12 *	0.66	0.64	0.79 *
Kerosene/paraffin	0.49	0.79 *	0.72 *	0.00	0.00	0.00
Charcoal	0.18	0.09 *	0.10 *	0.00	0.00	0.00
Gas/electricity	0.03	0.03	0.05	0.00	0.00	0.00
Coal	0.00	0.00	0.00	0.00	0.00	0.00
Cow dung	0.05	0.01 *	0.01 *	0.34	0.35	0.21 *
None	0.00	0.00	0.00	0.00	0.00	0.00
Other	0.02	0.00 *	0.00 *	0.00	0.00	0.00
<b>Characteristics of the head of the household</b>						
Average age	0.00	38.82	38.65	36.75	39.42 *	38.91 *
Sex (male)	0.79	0.79	0.74	0.84	0.94 *	0.91 *
<b>Mother's background</b>						
Literacy rate						
Easily	0.66	0.64	0.73	0.10	0.10	0.05 *
With difficulty	0.14	0.14	0.07 *	0.09	0.06	0.07
Not at all	0.20	0.21	0.20	0.81	0.84	0.87 *
Highest year of education						
None	0.19	0.26	0.19	0.83	0.86	0.79
Elementary	0.60	0.37 *	0.31 *	0.17	0.14	0.20
Secondary	0.19	0.27 *	0.24	0.00	0.00	0.02
Higher	0.02	0.11 *	0.26 *	0.00	0.00	0.00
Ethnic background						
Other	0.10	0.03 *	0.04 *	0.00	0.03 *	0.02 *
Agew	0.00	0.00	0.00	0.01	0.05 *	0.00
Amhara	0.28	0.47 *	0.55 *	0.99	0.88 *	0.94 *
Gurage	0.16	0.17	0.24	0.00	0.00	0.00
Hadiva	0.00	0.01	0.00	0.00	0.00	0.00
Kambata	0.00	0.01	0.00	0.00	0.00	0.00
Oromo	0.34	0.23 *	0.12 *	0.00	0.04 *	0.04 *
Sidama	0.00	0.00	0.00	0.00	0.00	0.00
Tigran	0.11	0.09	0.04 *	0.00	0.00	0.00
Wolavta	0.00	0.00	0.00	0.00	0.00	0.00
Religion						
Muslim	0.18	0.15	0.13	0.00	0.16 *	0.16 *
Catholic	0.01	0.01	0.00	0.00	0.00	0.00
Protestant	0.03	0.06	0.11 *	0.01	0.00	0.00
Orthodox	0.78	0.78	0.76	0.99	0.84 *	0.84 *
Evangelist	0.00	0.00	0.00	0.00	0.00	0.00
Other	0.01	0.01	0.00	0.00	0.00	0.00
<b>Pregnancy, delivery and breastfeeding</b>						
At least 2 inj. of tetanus d. pregn.	0.95	0.75 *	0.76 *	0.79	0.51 *	0.69
Receiving antenatal care	0.87	0.85	0.90	0.29	0.12 *	0.24
N° of antenatal visits d. pregnancy	5.17	5.90	6.45 *	0.85	0.26 *	0.61 *
Place of delivery						
Home	0.26	0.32	0.17 *	0.97	0.98	0.98
Hospital	0.46	0.35 *	0.57 *	0.01	0.01	0.00
Other health facility	0.28	0.33	0.25	0.02	0.01	0.02
Breastfeeding	0.97	0.97	0.96	1.00	0.98 *	0.98 *
Sample size	274	156	112	294	452	473

\* P-value at the 5% of significance for the t-test. Null hypothesis: equality of means

Chi2: Null hypothesis: independence between samples and categories.

	Rural Oromia			Rural SNNP			Rural Tigray		
	YL	DHS 2000	DHS 2005	YL	DHS 2000	DHS 2005	YL	DHS 2000	DHS 2005
<b>Household characteristics</b>									
Urban households	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Average Household Size	5.65	6.19 *	6.18 *	5.94	5.83	6.12	5.88	5.75	5.81
Electricity supply	0.07	0.00 *	0.02 *	0.14	0.01 *	0.04 *	0.01	0.00	0.02
Source of drinking water									
Piped into dwelling/yard/plot	0.00	0.00	0.00	0.05	0.00 *	0.00 *	0.00	0.00	0.00
Protected public source	0.40	0.13 *	0.55 *	0.15	0.14	0.52 *	0.44	0.28 *	0.65 *
Unprotected source	0.60	0.87 *	0.45 *	0.80	0.86 *	0.48 *	0.56	0.72 *	0.35 *
Toilet facility									
At home	0.09	0.08	0.13	0.16	0.11	0.56 *	0.29	0.01 *	0.08 *
Outside home	0.01	0.04 *	0.05 *	0.04	0.02	0.12 *	0.06	0.01 *	0.01 *
None	0.90	0.88	0.82 *	0.81	0.87 *	0.33 *	0.65	0.98 *	0.90 *
Own radio	0.29	0.22 *	0.33	0.26	0.14 *	0.30	0.30	0.17 *	0.31
Own TV	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.01
Cooking fuel									
Wood	0.70	0.83 *	0.93 *	0.99	1.00	1.00	0.29	0.77 *	0.91 *
Kerosene/paraffin	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00
Charcoal	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00
Gas/electricity	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Coal	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Cow dung	0.30	0.17 *	0.07 *	0.00	0.00	0.00	0.69	0.23 *	0.08 *
None	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Other	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
<b>Characteristics of the head of the household</b>									
Average age	36.16	37.53	36.73	35.81	38.67 *	37.54 *	38.34	39.83	39.31
Sex (male)	0.93	0.92	0.92	0.96	0.83 *	0.90 *	0.87	0.85	0.90
<b>Mother's background</b>									
Literacy rate									
Easily	0.16	0.09 *	0.06 *	0.17	0.10 *	0.07 *	0.09	0.03 *	0.05 *
With difficulty	0.16	0.07 *	0.07 *	0.12	0.07 *	0.08 *	0.04	0.02	0.07
Not at all	0.68	0.84 *	0.87 *	0.70	0.83 *	0.85 *	0.86	0.94 *	0.88
Highest year of education									
None	0.65	0.64	0.54 *	0.63	0.58	0.49 *	0.84	0.83	0.74 *
Elementary	0.34	0.33	0.45 *	0.36	0.40	0.49 *	0.16	0.16	0.25 *
Secondary	0.00	0.02 *	0.01	0.02	0.01	0.02	0.00	0.00	0.01
Higher	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00
Ethnic background									
Other	0.11	0.02 *	0.02 *	0.02	0.37 *	0.43 *	0.00	0.02 *	0.03 *
Agew	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Amhara	0.14	0.09 *	0.07 *	0.02	0.02	0.02	0.00	0.01 *	0.02 *
Gurage	0.00	0.01	0.00	0.26	0.18 *	0.12 *	0.00	0.00	0.00
Hadiva	0.01	0.00	0.01	0.23	0.05 *	0.07 *	0.00	0.00	0.00
Kambata	0.00	0.00	0.00	0.00	0.05 *	0.03 *	0.00	0.00	0.00
Oromo	0.73	0.88 *	0.90 *	0.00	0.02 *	0.02 *	0.00	0.23 *	0.18 *
Sidama	0.01	0.00	0.00	0.24	0.22	0.21	0.00	0.00	0.00
Tigrayan	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.74 *	0.76 *
Wolavta	0.00	0.00	0.00	0.22	0.10 *	0.10 *	0.00	0.00	0.00
Religion									
Muslim	0.23	0.49 *	0.57 *	0.46	0.14 *	0.13 *	0.00	0.25 *	0.22 *
Catholic	0.01	0.00	0.01	0.00	0.03 *	0.03 *	0.00	0.00	0.00
Protestant	0.05	0.11 *	0.14 *	0.39	0.53 *	0.58 *	0.00	0.00	0.00
Orthodox	0.71	0.32 *	0.26 *	0.13	0.25 *	0.21 *	0.96	0.75 *	0.78 *
Evangelist	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00
Other	0.00	0.08 *	0.03 *	0.02	0.06 *	0.06 *	0.04	0.00 *	0.00 *
<b>Pregnancy, delivery and breastfeeding</b>									
At least 2 inj. of tetanus d. pregn.	0.87	0.67 *	0.78 *	0.87	0.62 *	0.79 *	0.82	0.54 *	0.76
Receiving antenatal care	0.39	0.25 *	0.22 *	0.37	0.26 *	0.31	0.52	0.30 *	0.32 *
N° of antenatal visits d. pregnancy	1.33	0.79 *	0.70 *	1.21	0.81 *	1.12	1.94	0.68 *	0.98 *
Place of delivery									
Home	0.96	0.98	0.97	0.97	0.98	0.97	0.97	0.99 *	0.97
Hospital	0.03	0.00 *	0.02	0.02	0.00 *	0.01	0.02	0.00	0.02
Other health facility	0.01	0.01	0.01	0.01	0.01	0.02	0.01	0.00	0.01
Breastfeeding	0.97	0.97	0.96	0.99	0.97 *	0.98	1.00	0.98 *	0.98 *
Sample size	292	639	652	397	472	568	294	361	467

\* P-value at the 5% of significance for the t-test. Null hypothesis: equality of means

Chi2: Null hypothesis: independence between samples and categories.

## AN ASSESSMENT OF THE YOUNG LIVES SAMPLING APPROACH IN ETHIOPIA

	Rural Oromia			Rural SNNP			Rural Tigray		
	YL	DHS 2000	DHS 2005	YL	DHS 2000	DHS 2005	YL	DHS 2000	DHS 2005
<b>Household characteristics</b>									
Urban households	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Average Household Size	5.65	6.19 *	6.18 *	5.94	5.83	6.12	5.88	5.75	5.81
Electricity supply	0.07	0.00 *	0.02 *	0.14	0.01 *	0.04 *	0.01	0.00	0.02
Source of drinking water									
Piped into dwelling/yard/plot	0.00	0.00	0.00	0.05	0.00 *	0.00 *	0.00	0.00	0.00
Protected public source	0.40	0.13 *	0.55 *	0.15	0.14	0.52 *	0.44	0.28 *	0.65 *
Unprotected source	0.60	0.87 *	0.45 *	0.80	0.86 *	0.48 *	0.56	0.72 *	0.35 *
Toilet facility									
At home	0.09	0.08	0.13	0.16	0.11	0.56 *	0.29	0.01 *	0.08 *
Outside home	0.01	0.04 *	0.05 *	0.04	0.02	0.12 *	0.06	0.01 *	0.01 *
None	0.90	0.88	0.82 *	0.81	0.87 *	0.33 *	0.65	0.98 *	0.90 *
Own radio	0.29	0.22 *	0.33	0.26	0.14 *	0.30	0.30	0.17 *	0.31
Own TV	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.01
Cooking fuel									
Wood	0.70	0.83 *	0.93 *	0.99	1.00	1.00	0.29	0.77 *	0.91 *
Kerosene/paraffin	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00
Charcoal	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00
Gas/electricity	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Coal	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Cow dung	0.30	0.17 *	0.07 *	0.00	0.00	0.00	0.69	0.23 *	0.08 *
None	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Other	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
<b>Characteristics of the head of the household</b>									
Average age	36.16	37.53	36.73	35.81	38.67 *	37.54 *	38.34	39.83	39.31
Sex (male)	0.93	0.92	0.92	0.96	0.83 *	0.90 *	0.87	0.85	0.90
<b>Mother's background</b>									
Literacy rate									
Easily	0.16	0.09 *	0.06 *	0.17	0.10 *	0.07 *	0.09	0.03 *	0.05 *
With difficulty	0.16	0.07 *	0.07 *	0.12	0.07 *	0.08 *	0.04	0.02	0.07
Not at all	0.68	0.84 *	0.87 *	0.70	0.83 *	0.85 *	0.86	0.94 *	0.88
Highest year of education									
None	0.65	0.64	0.54 *	0.63	0.58	0.49 *	0.84	0.83	0.74 *
Elementary	0.34	0.33	0.45 *	0.36	0.40	0.49 *	0.16	0.16	0.25 *
Secondary	0.00	0.02 *	0.01	0.02	0.01	0.02	0.00	0.00	0.01
Higher	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00
Ethnic background									
Other	0.11	0.02 *	0.02 *	0.02	0.37 *	0.43 *	0.00	0.02 *	0.03 *
Agew	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Amhara	0.14	0.09 *	0.07 *	0.02	0.02	0.02	0.00	0.01 *	0.02 *
Gurage	0.00	0.01	0.00	0.26	0.18 *	0.12 *	0.00	0.00	0.00
Hadiva	0.01	0.00	0.01	0.23	0.05 *	0.07 *	0.00	0.00	0.00
Kambata	0.00	0.00	0.00	0.00	0.05 *	0.03 *	0.00	0.00	0.00
Oromo	0.73	0.88 *	0.90 *	0.00	0.02 *	0.02 *	0.00	0.23 *	0.18 *
Sidama	0.01	0.00	0.00	0.24	0.22	0.21	0.00	0.00	0.00
Tigrayan	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.74 *	0.76 *
Wolavta	0.00	0.00	0.00	0.22	0.10 *	0.10 *	0.00	0.00	0.00
Religion									
Muslim	0.23	0.49 *	0.57 *	0.46	0.14 *	0.13 *	0.00	0.25 *	0.22 *
Catholic	0.01	0.00	0.01	0.00	0.03 *	0.03 *	0.00	0.00	0.00
Protestant	0.05	0.11 *	0.14 *	0.39	0.53 *	0.58 *	0.00	0.00	0.00
Orthodox	0.71	0.32 *	0.26 *	0.13	0.25 *	0.21 *	0.96	0.75 *	0.78 *
Evangelist	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00
Other	0.00	0.08 *	0.03 *	0.02	0.06 *	0.06 *	0.04	0.00 *	0.00 *
<b>Pregnancy, delivery and breastfeeding</b>									
At least 2 inj. of tetanus d. pregn.	0.87	0.67 *	0.78 *	0.87	0.62 *	0.79 *	0.82	0.54 *	0.76
Receiving antenatal care	0.39	0.25 *	0.22 *	0.37	0.26 *	0.31	0.52	0.30 *	0.32 *
N° of antenatal visits d. pregnancy	1.33	0.79 *	0.70 *	1.21	0.81 *	1.12	1.94	0.68 *	0.98 *
Place of delivery									
Home	0.96	0.98	0.97	0.97	0.98	0.97	0.97	0.99 *	0.97
Hospital	0.03	0.00 *	0.02	0.02	0.00 *	0.01	0.02	0.00	0.02
Other health facility	0.01	0.01	0.01	0.01	0.01	0.02	0.01	0.00	0.01
Breastfeeding	0.97	0.97	0.96	0.99	0.97 *	0.98	1.00	0.98 *	0.98 *
Sample size	292	639	652	397	472	568	294	361	467

\* P-value at the 5% of significance for the t-test. Null hypothesis: equality of means

Chi2: Null hypothesis: independence between samples and categories.

**YL vs DHS 2000**

<b>Wilcoxon / Mann-Whitney test</b>	National Sample	5-region Sample	Total Urban	Total Rural	Addis Ababa	Rural Amhara	Rural Oromia	Rural SNNP	Rural Tigray
Variable	Prob > z	Prob > z	Prob > z	Prob > z	Prob > z	Prob > z	Prob > z	Prob > z	Prob > z
Number of antenatal visits	0.00	0.00	0.99	0.00	0.04	0.00	0.00	0.00	0.00
Household size	0.00	0.00	0.87	0.00	0.62	0.00	0.00	0.52	0.20
Years of schooling									
Head of the household	0.00	0.00	0.01	0.52	0.11	0.27	0.84	0.07	0.93
Mother	0.00	0.00	0.00	0.00	0.40	0.00	0.00	0.00	0.00

**YL vs DHS 2000**

<b>Kolmogorov-Smirnov test</b>	National Sample	5-region Sample	Total Urban	Total Rural	Addis Ababa	Rural Amhara	Rural Oromia	Rural SNNP	Rural Tigray
Variable	p-value	p-value	p-value	p-value	p-value	p-value	p-value	p-value	p-value
Wealth Index	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00

# Appendix B

## Comparison of the Welfare Monitoring Survey with Young Lives

	National Sample			YL Regions Only		Total Urban		Total Rural		
	YL	WMS Unw.	WMS W.	YL	WMS	YL	WMS	YL	WMS	
<b>Household characteristics</b>										
Average Household Size <sup>b</sup>	5.70	5.76	5.72	n.a	5.70	5.75	5.69	5.69	5.70	5.77
Electricity supply	0.35	0.16 ***	0.06 ***		0.35	0.14 ***	0.90	0.73 ***	0.06	0.01 ***
Source of drinking water <sup>a</sup>										
Piped into dwelling/yard/plot	0.12	0.07 ***	0.03 ***		0.12	0.06 ***	0.31	0.33	0.02	0.00 ***
Protected public source	0.41	0.23 ***	0.17 ***		0.41	0.20 ***	0.53	0.56	0.35	0.15 ***
Unprotected source	0.47	0.69 ***	0.80 ***		0.47	0.73 ***	0.16	0.11 ***	0.63	0.84 ***
Toilet facility <sup>a</sup>										
At home - toilet/private latrine	0.21	0.16 ***	0.11 ***		0.21	0.15 ***	0.34	0.44 ***	0.15	0.09 ***
Outside home	0.16	0.08 ***	0.04 ***		0.16	0.07 ***	0.39	0.31 ***	0.04	0.02 **
None	0.63	0.76 ***	0.86 ***		0.63	0.78 ***	0.27	0.25	0.81	0.88 ***
Own radio	0.38	0.26 ***	0.18 ***		0.38	0.23 ***	0.65	0.64	0.24	0.16 ***
Own TV	0.05	0.03 ***	0.01 ***		0.05	0.02 ***	0.15	0.14	0.00	0.00
Own Fridge	0.01	0.01	0.00		0.01	0.01	0.03	0.04	0.00	0.00
Cooking fuel <sup>a</sup>										
Wood	0.64	0.78 ***	0.78 ***		0.64	0.77 ***	0.55	0.55	0.68	0.83 ***
Gas/Electro/Coal/Kerosene	0.13	0.08 ***	0.02 ***		0.13	0.06 ***	0.36	0.39	0.01	0.00
Dung/Leaves/Other	0.24	0.14 ***	0.20 ***		0.24	0.17 ***	0.09	0.06 **	0.31	0.16 ***
Roof - Iron	0.44	0.28 ***	0.19 ***		0.44	0.26 ***	0.89	0.84 ***	0.21	0.14 ***
Roof - Straw/Thatch	0.40	0.60 ***	0.70 ***		0.40	0.63 ***	0.03	0.09 ***	0.60	0.72 ***
<b>Household Assets</b>										
HH Owns Land	0.60	0.89 ***	0.95 ***		0.60	0.90 ***	0.13	0.53 ***	0.85	0.98 ***
HH Owns House	0.66	0.83 ***	0.88 ***		0.66	0.84 ***	0.34	0.47 ***	0.83	0.92 ***
HH own Livestock	0.65	0.80 ***	0.88 ***		0.65	0.82 ***	0.38	0.34 *	0.78	0.91 ***
Nr Cattle Owned <sup>b</sup>	1.34	3.09 ***	3.23	n.a	1.34	2.95 ***	0.25	0.68 ***	1.92	3.70 ***
Nr Rooms in HH House <sup>b</sup>	1.49	1.78 ***	1.75	n.a	1.49	1.81 ***	1.74	2.12 ***	1.35	1.70 ***
One Room	0.66	0.57 ***	0.60 ***		0.66	0.56 ***	0.52	0.41 ***	0.74	0.61 ***
Two Rooms	0.23	0.31 ***	0.28 ***		0.23	0.30 ***	0.30	0.35 **	0.19	0.29 ***
Three Rooms	0.08	0.08	0.09		0.08	0.09	0.13	0.14	0.06	0.07 *
Four Rooms	0.02	0.02	0.02		0.02	0.03	0.03	0.05	0.01	0.01
<b>Household Head Characteristics</b>										
Sex (male)	0.86	0.86	0.86		0.86	0.85	0.78	0.78	0.91	0.88 ***
Average age	37.15	37.07	37.27		37.15	37.28	37.99	37.62	36.72	36.94
Level of Education <sup>b</sup>										
None	0.87	0.60 ***	0.47	n.a	0.87	0.57 ***	1.46	1.50	0.57	0.38 ***
Incomp. Primary	0.52	0.66 ***	0.70 ***		0.52	0.67 ***	0.32	0.34	0.62	0.74 ***
Comp. Primary	0.21	0.16 ***	0.17 ***		0.21	0.16 ***	0.18	0.15 *	0.22	0.16 ***
Secondary	0.18	0.13 ***	0.10 ***		0.18	0.13 ***	0.28	0.29	0.13	0.09 ***
Higher/Other	0.06	0.03 ***	0.02 ***		0.06	0.03 ***	0.14	0.13	0.02	0.01 ***
	0.03	0.02 *	0.01 ***		0.03	0.02 **	0.07	0.10	0.01	0.00 *
<b>Child Information</b>										
Sex (male)	0.53	0.50 **	0.48 ***		0.53	0.49 ***	0.53	0.52	0.53	0.49 **
Average age	11.68	11.47 **	11.47 *		11.68	11.44 **	11.71	11.40 *	11.66	11.49
Height (cm)	70.85	68.08 ***	67.94 ***		70.85	67.84 ***	71.70	68.38 ***	70.38	68.01 ***
Weight (kg)	7.97	7.71 ***	7.59 ***		7.97	7.67 ***	8.40	8.23 ***	7.72	7.59 ***
Height-for-Age scores	-1.32	-1.99 ***	-2.09 ***		-1.32	-2.07 ***	-1.06	-1.67 ***	-1.46	-2.07 ***
Weight-for-height scores	-0.73	-0.30 ***	-0.40 ***		-0.73	-0.31 ***	-0.36	0.14 ***	-0.95	-0.40 ***
Measles Vaccination	0.58	0.47 ***	0.41 ***		0.58	0.47 ***	0.64	0.81 ***	0.55	0.38 ***
TB Vaccination	0.74	0.54 ***	0.47 ***		0.74	0.54 ***	0.83	0.88 ***	0.69	0.46 ***
Sample size	1,953	3856	3856		1,953	3,082	671	770	1,282	3,086

Kolmogorov-Smirnov test	p-value	p-value	p-value	p-value	p-value
HH Head - Average age	0.00	n.a		0.01	0.88
Notes:					0.01

Notes:

(\*) - Indicates Differences between YL and WMS are significant at 10% level of confidence. (\*\*) and (\*\*\*) indicate significance at 5% and 1% respectively

a - Reported tests correspond to Chi-Square Tests

b - Reported tests correspond to Wilcoxon Ranksum Tests

## AN ASSESSMENT OF THE YOUNG LIVES SAMPLING APPROACH IN ETHIOPIA

	Rural Tigray		Rural Amhara		Rural Oromia		Rural SNNPR		Urban Addis Ababa	
	YL	WMS	YL	WMS	YL	WMS	YL	WMS	YL	WMS
<b>Household characteristics</b>										
Average Household Size <sup>b</sup>	5.88	5.75 *	5.22	5.45	5.66	5.86 *	5.94	5.80	5.74	6.27
Electricity supply	0.01	0.00	0.00	0.01	0.07	0.01 ***	0.14	0.01 ***	0.99	0.97
Source of drinking water <sup>a</sup>										
Piped into dwelling/yard/plot	0.00	0.00 ***	0.01	0.00 ***	0.00	0.00 ***	0.05	0.00 ***	0.34	0.56 ***
Protected public source	0.44	0.19 ***	0.49	0.09 ***	0.40	0.12 ***	0.15	0.14	0.51	0.43 ***
Unprotected source	0.56	0.81 ***	0.50	0.91 ***	0.60	0.86 ***	0.80	0.85 **	0.14	0.01 ***
Toilet facility <sup>a</sup>										
At home - toilet/private latrine	0.29	0.05 ***	0.04	0.01 **	0.09	0.08	0.16	0.16	0.17	0.35 ***
Outside home	0.06	0.00 ***	0.06	0.00 ***	0.01	0.02	0.04	0.03	0.70	0.52 ***
None	0.65	0.95 ***	0.90	0.98 ***	0.89	0.89	0.81	0.81	0.12	0.13
Own radio	0.30	0.16 ***	0.10	0.06 **	0.29	0.18 ***	0.26	0.15 ***	0.71	0.82 **
Own TV	0.00	0.00 ***	0.00	0.00 ***	0.00	0.00	0.01	0.00	0.26	0.28
Own Fridge	0.00	0.00 ***	0.00	0.00 ***	0.01	0.00	0.00	0.00	0.03	0.08
Cooking fuel <sup>a</sup>										
Wood	0.28	0.60 ***	0.66	0.69	0.69	0.78 ***	0.99	0.97 ***	0.24	0.10 ***
Gas/Electro/Coal/Kerosene	0.02	0.00 **	0.00	0.00	0.00	0.00 ***	0.01	0.00	0.70	0.83 ***
Dung/Leaves/Other	0.70	0.40 ***	0.34	0.30	0.31	0.21 ***	0.00	0.03 ***	0.06	0.07
Roof - Iron	0.13	0.06 **	0.18	0.19	0.29	0.13 ***	0.22	0.07 ***	0.95	0.99 **
Roof - Straw/Thatch	0.08	0.31 ***	0.82	0.77 *	0.69	0.80 ***	0.74	0.80 ***	0.00	0.01
<b>Household Assets</b>										
HH Owns Land	0.87	0.94 ***	0.90	0.98 ***	0.85	0.99 ***	0.80	0.99 ***	0.00	0.39 ***
HH Owns House	0.71	0.88 ***	0.81	0.91 ***	0.88	0.92 *	0.90	0.93 **	0.12	0.31 ***
HH own Livestock	0.84	0.95 ***	0.80	0.95 ***	0.75	0.91 ***	0.76	0.89 ***	0.27	0.11 ***
Nr Cattle Owned <sup>b</sup>	1.49	3.81 ***	2.23	3.06 ***	1.79	3.74 ***	2.09	3.20 ***	0.03	0.10
Nr Rooms in HH House <sup>b</sup>	1.40	1.62 ***	1.40	1.53 ***	1.26	2.01 ***	1.35	1.55 ***	1.52	2.94 ***
One Room	0.68	0.57 **	0.73	0.63 ***	0.78	0.57 ***	0.75	0.65 ***	0.60	0.33 ***
Two Rooms	0.25	0.26 ***	0.17	0.26 ***	0.20	0.33 ***	0.16	0.27 ***	0.30	0.36
Three Rooms	0.07	0.14 ***	0.07	0.08	0.02	0.08 ***	0.08	0.05	0.08	0.16 *
Four Rooms	0.00	0.02 **	0.03	0.03	0.01	0.01	0.01	0.02	0.01	0.06
<b>Household Head Characteristics</b>										
Sex (male)	0.87	0.77 ***	0.84	0.94 ***	0.93	0.88 ***	0.96	0.86 ***	0.78	0.73
Average age	38.37	41.26 ***	36.75	39.06 ***	36.28	36.33	35.81	35.84	38.50	40.82
Level of Education <sup>b</sup>										
None	0.27	0.14 **	0.24	0.17 **	0.72	0.43 ***	0.92	0.49 ***	1.60	1.69
Incomp. Primary	0.80	0.87 **	0.82	0.88 **	0.49	0.70 ***	0.44	0.66 ***	0.21	0.28
Comp. Primary	0.17	0.12	0.13	0.08 **	0.33	0.18 ***	0.25	0.20 **	0.22	0.13 **
Secondary	0.02	0.01	0.04	0.04	0.15	0.10 **	0.26	0.13 ***	0.39	0.31
Higher/Other	0.00	0.00	0.01	0.00	0.02	0.01	0.03	0.00 ***	0.12	0.18
Higher/Other	0.01	0.00 **	0.00	0.00	0.00	0.00	0.02	0.00 *	0.06	0.10
<b>Child Information</b>										
Sex (male)	0.55	0.47 *	0.58	0.45 ***	0.53	0.49	0.48	0.51	0.54	0.53
Average age	11.45	11.40	11.82	11.34 *	12.13	11.49 ***	11.35	11.52	11.98	11.00 **
Height (cm)	70.93	66.94 ***	69.69	66.74 ***	70.07	68.37 ***	70.73	68.15 ***	72.25	68.49 ***
Weight (kg)	8.01	7.19 ***	7.63	7.27 ***	7.61	7.71 ***	7.63	7.67 ***	8.66	8.22 ***
Height-for-Age scores	-1.08	-2.16 ***	-1.89	-2.48 ***	-1.74	-1.99 *	-1.22	-2.06 ***	-0.91	-1.40 **
Weight-for-height scores	-0.72	-0.60	-0.88	-0.50 ***	-1.03	-0.32 ***	-1.09	-0.37 ***	-0.17	-0.05
Measles Vaccination	0.84	0.77 **	0.39	0.39	0.49	0.35 ***	0.49	0.38 ***	0.73	0.86 ***
TB Vaccination	0.82	0.79	0.73	0.44 ***	0.61	0.42 ***	0.60	0.45 ***	0.98	0.99
Sample size	294	203	296	513	295	834	397	968	287	89

Kolmogorov-Smirnov test	p-value	p-value	p-value	p-value	p-value
HH Head - Average age	0.01	0.04	0.77	0.00	0.68

Notes:

(\*) - Indicates Differences between YL and WMS are significant at 10% level of confidence. (\*\*) and (\*\*\*) indicate significance at 5% and 1% respectively

a - Reported tests correspond to Chi-Square Tests

b - Reported tests correspond to Wilcoxon Ranksum Tests

# Appendix C

## Non-parametric tests for comparison of distributions<sup>20</sup>

### 1. Kolmogorov-Smirnov test

The Kolmogorov-Smirnov test is applied to test for the equality of distributions of continuous variables. It investigates the significance of the difference between two population distributions based on two sample distributions, being the null hypothesis that the population distributions are equal.

#### *Method*

Given samples 1 and 2 of size  $n_1$  and  $n_2$  from two populations, the cumulative distribution functions  $S_{n_1}(x)$  and  $S_{n_2}(x)$  can be determined and plotted. Based on this, the maximum value of the difference between the plots can be found and compared with a critical value. If the observed value exceeds the critical value, the null hypothesis that the two population distributions are identical is rejected.

### 2. Wilcoxon-Mann-Whitney test

The Wilcoxon-Mann-Whitney test is a non-parametric method used to test the null hypothesis that two random samples could have come from two populations with the same mean. When the shape of the distribution of both samples is not the same, it tests whether one population has larger values than the other.

Formally, the test assumes that the data are continuous, but it can be used for ordered categorical data as well (the test focuses on the rank order).

#### *Method*

Consider sample 1 and 2. The observations of these two samples are combined and arranged in order of increasing size and given a rank number. In cases where equal results occur, the mean of the available rank numbers is assigned. Next, the rank sum  $R$  of the smaller sample is calculated. Let  $N$  denote the size of the combined samples and  $n$  denote the size of the smaller sample. A second quantity  $R' = n(N+1) - R$  is calculated. The values  $R$  and  $R'$  are compared with critical values. If either  $R$  or  $R'$  are less than the critical value, the null hypothesis of the same mean would be rejected.

### 3. Chi-square test for consistency in a 2 x K table

This test is used to investigate the significance of the differences between two distributions based on two samples spread over K classes.

#### Method

Consider two samples 1 and 2 of size  $N_1$  and  $N_2$ , respectively, where  $n_{ij}$  represents the frequency of individuals in the  $i$ th sample in the  $j$ th class ( $i=1,2$  and  $j=1,\dots,K$ ) as described in the following table:

	1	2	...j....	K	Total
Sample 1	$n_{11}$	$n_{12}$	... $n_{1j}$ ...	$n_{1k}$	$N_1$
Sample 2	$n_{21}$	$n_{22}$	... $n_{2j}$ ...	$n_{2k}$	$N_2$
Total	$n_{.1}$	$n_{.2}$	... $n_{.j}$ ...	$n_{.k}$	$N=N_1+N_2$

Based on this information, another table of expected frequencies is calculated where the value in the  $i$ th row and  $j$ th column is:

$$e_{ij} = \frac{N_i * n_{.j}}{N_1 + N_2}$$

and the test statistic is:

$$\chi^2 = \sum_{j=1}^K \frac{(n_{1j} - e_{1j})^2}{e_{1j}} + \sum_{j=1}^K \frac{(n_{2j} - e_{2j})^2}{e_{2j}}$$

If this value exceeds the critical value obtained from  $\chi^2$  tables with  $(K-1)$  degrees of freedom, the null hypothesis that the two samples originate from two populations with the same distribution is rejected.

# Appendix D

## Constructing a wealth index

Different definitions and possible methods of calculation of a wealth index can be found in the literature. We have chosen this particular combination of indicators on basis of information available in both the Young Lives and the DHS samples. For this reason, our wealth index differs from the wealth index defined in the Young Lives Preliminary Country Report Ethiopia (Alemu et al. 2003).<sup>21</sup>

Our wealth index is the simple average of three different indexes: Housing Quality Index (HQ), Consumer Durables Index (CD) and Services Index (S). Each of these is simple means of variables that are in a 0-1 range. Hence, the wealth index can take a value between 0 and 1, whereby a higher wealth index value indicates a higher socioeconomic status.

### Housing quality index

- HQ1: Rooms per Person. Number of rooms divided by the number of household members. The HQ1 variable is set to take a maximum value of unity. Ratios higher than 1 are recoded accordingly.
- HQ2: Floor Quality. It takes the value of 1 if the floor is made of a finished material (cement, tile or laminated material); 0 otherwise.
- HQ3: Roof Quality. It takes the value of 1 if the roof is made of iron, concrete tiles or slates; 0 otherwise.

### Consumer durables index

- Constructed from simple means of the following dummy variables:<sup>22</sup> ownership of (i) radio, (ii) bicycle, (iii) TV, (iv) motorbike or scooter, (v) motorised vehicle or truck, (vi) landline telephone, and (vii) a modern bed or a table.

### Services index:

- S1: Electricity. Has value of 1 if the household has access to electricity; 0 otherwise.
- S2: Water. Has value of 1 if the household's source of drinking water is piped into dwelling or yard; 0 otherwise.
- S3: Toilet. Has value of 1 if the household has access to its own pit latrine or flush toilet; 0 otherwise.
- S4: Cooking fuel. Has value of 1 if the household uses electricity, gas or kerosene as cooking fuel; 0 otherwise.

---

<sup>21</sup> The Wealth index calculated in the country report includes the following additional variables: HQ4 – Wall quality (brick/plastered walls in dwelling), CD – ownership of (i) fridge, (ii) mobile phone, and (iii) sofa, CD – ownership of (i) modern bed, (ii) table or chair.

<sup>22</sup> A dummy variable is a variable that has the value zero or one for each observation, e.g. ownership of a radio yes or no.

---

## THE AUTHORS

**Ingo Outes-Leon** is a research assistant for Young Lives and a doctoral student in development economics affiliated to the Centre for the Study of African Economies at the University of Oxford. His research is focused on issues of poverty dynamics, health, nutrition and risk.

**Alan Sanchez** is a research assistant for Young Lives and DPhil student in Economics at the University of Oxford. His research is focused on human capital accumulation in the context of low-income families in developing countries. Currently, he is doing empirical research on the nutrition-learning nexus and on the impact of adverse climate events on nutrition in Peru.