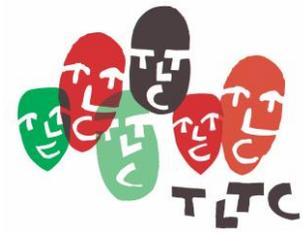# DOCUMENTATION GUIDE #14:
Constructed Dataset of All TLT-1 and TLT-2 Births

Dataset created by: Christie Sennott
Guide written by: Christie Sennott, Sara Yeatman and Jenny Trinitapoli
Last updated: February 8, 2019

Fertility was central to the goals of TLT and figured prominently in our data instruments. As with all longitudinal data sets, however, collecting data as the events are occurring and at multiple points of time creates challenges for analysts. Thus, based on our experience working with fertility in a dynamic fashion across waves and trying to identify child deaths, we created a full "Births Dataset" to assist future users.

The dataset pulls from all ten waves of TLT-1 and TLT-2. The dataset is a long dataset wherein each child ever born to a TLT respondent represents an observation. Some variables are specific to that child (e.g., date of birth, gender, vital status), while others refer to the parent's characteristics (e.g., children ever born, living children at a given wave) and are identical across siblings. Therefore, researchers interested in the characteristics of the TLT respondent at various waves may want to reshape the dataset to use those variables. And, researchers interested in children and their birth order or outcomes can use the data in its current, long format.

The variables described below come mainly from the child rosters[1] that were asked at each wave (see data key for details and small differences across waves). Additional information comes from the pregnancy section of the questionnaires, and particularly p17 (living children) and p19 (children ever born) that were asked at waves 1, 9 and 10. For women, g5 (have you had a new birth) was additionally useful to confirm the timing of some births.

Reminder: TLT-1 (waves 1-8), refresher sample (wave 9), TLT-2 (wave 10)

The dataset contains reports on children from all women and men in the sample who reported having children:

Core sample women:	1,171
Refresher sample women:	263
Core sample men:	225
Male partners entering at TLT1:	691
Male partners entering at TLT2:	230
Total:	**2,580**

VARIABLES
respid: Respondent ID
This variable assigns a unique number to each respondent in the dataset.

childnum: Child order
This variable assigns a unique number to each child of a respondent. Children are ordered by birth month/year with the first born child coded 1, the second born child coded 2, etc.

ch2: Child gender
This variable measures the child's gender and is coded 1 for male and 2 for female.
- 3 children who died and were never given names are missing (.d) on this variable

ch4: Child still alive
This variable indicates whether the child is still alive and is coded 1 for alive and 0 for dead.
- 8 cases are missing (.d) – these cases are also missing birth dates and death dates

ch3m: Child birth month
This variable indicates in which month the child was born (1-12).
- 288 missing values (.d)
  - 252 of these are births that occurred prior to the respondent joining TLT (see below)
  - 36 occur during TLT (2 reported by women, 34 reported by men)
  - 54 cases are missing both birth month and year (47 of these were pre-TLT)

ch3y: Child birth year
This variable indicates the year in which the child was born (1979-2015).
- 75 missing values
  - 57 cases are births that occurred prior to the respondent joining TLT (see below)
  - 18 cases occurred during TLT

ch_dob_cmc: Child DOB in CMC
This variable measures the child's date of birth (month and year) in Century Month Codes.
- 12(birth year−1900)+birth month
- http://demographicestimation.iussp.org/content/dhs-century-month-codes
- 309 children are missing month or year of birth and are missing on this variable (.d)

ch10m: Child death month
- 32 children who died are missing (.d) their month of death (24 from births before TLT)
  - Note that 8 of these cases are also missing on ch4, so we cannot know if they did in fact die
- If child is still alive, the variable is missing (.s)

ch10y: Child death year
- 14 children who died are missing (.d) their year of death (10 from births before TLT); all 14 were also missing the month of death

- o Note that 8 of these cases are also missing on ch4, so we cannot know if they did in fact die
- If child is still alive, the variable is missing (.s)

ch_dod_cmc: Child DOD in CMC
This variable measures the child's date of death (month and year) in Century Month Codes
- 12(death year−1900)+death month
- http://demographicestimation.iussp.org/content/dhs-century-month-codes
- If child is still alive, the variable is missing (.s)
- 32 children who have died are missing month or year of death and are missing on this variable (.d)

p17_tlt2: Living kids – at TLT-2
This variable measures the number of living kids as reported by the respondent (men and women) at TLT-2.
- Respondents who were not interviewed at TLT-2 have missing values on this variable (.m)
- No respondents who were interviewed at TLT-2 are missing

p19_tlt2: Children ever born – at TLT-2
This variable measures the number of children that were ever born as reported by the respondent (women only) at TLT-2.
- Men who were interviewed at TLT-2 have missing values on this variable because they were not asked this item (.s)
- Women and men who were not interviewed at TLT-2 have missing values on this variable (.m)

pretltbirth: Pre-TLT birth
This variable indicates that a birth occurred before the respondent entered the study (coded 1). If the birth occurred while the respondent was in the study, the variable is coded 0. All core sample men and women entered the study at wave 1: (May-Aug 2009). Refresher sample women entered at wave 9 (March-April 2012). Men in the partners sample entered the study across all waves (1-10).
- 57% of all births occurred before the respondent entered the study
- When the child's date of birth was missing:
  - o The birth was coded as 1 if it was reported at the respondent's first interview.
  - o The birth was coded as 0 if it was first reported at TLT-2 and the respondent was interviewed during TLT-1.

pretltdeath: Pre-TLT death
This variable indicates that a child death occurred before the respondent entered the study (coded 1). If the death occurred while the respondent was in the study, the variable is coded 0.
- 41% of child deaths occurred before the respondent entered the study
- 4 cases are missing (.d); they are all from the same respondent (6142751) and are missing the month and year of death
- 8 cases are missing on this variable because they are missing on ch4 (.d)
- Children who are still alive are coded as missing (.s)

https://tsogololathanzi.uchicago.edu

cebtot: Total children ever born
This variable measures the total number of children as reported on the child roster across all waves.
- Note: this does not always align perfectly with p19_tlt2, the summary measure of children ever born asked at TLT-2

ceb_w1 (1-10): # children ever born w1
This set of variables measures the wave-specific number of children the respondent has ever reported.
- Coded at waves 1-10, where wave 9 is the refresher sample and wave 10 is TLT-2
- Coded based on the child's birth month and year, and interview month and year
- The respondent had to be interviewed at the wave in order for the count of children ever born to increase at that wave
- If the respondent was not interviewed at the wave, coded as missing (.m)
- ceb_w1 ceb_w2 ceb_w3 ceb_w4 ceb_w5, etc.

livingkidstot: Total living kids at last wave
This variable measures the total number of children that are still alive at the respondent's last interview, based on the child roster.
- Note: this does not always align perfectly with p17_tlt2, the summary measure of living kids asked at TLT-2

livingkids_w1 (1-10): # living kids w1
This set of variables measures the wave-specific number of living children for each respondent.
- Coded at waves 1-10, where wave 9 is the refresher sample and wave 10 is TLT-2
- Coded based on the child's birth month and year; death month and year; and interview month and year
- The respondent had to be interviewed at the wave in order for the count of living children to increase or decrease at that wave
- If the respondent was not interviewed at the wave, coded as missing (.m)
- livingkids_w1 livingkids_w2 livingkids_w3 livingkids_w4 livingkids_w5, etc.

CODING PROCEDURES
Because information on children was collected at every wave, there is room for discrepancies across waves. Birth date inconsistencies were very common, especially among men (10-15%). When these discrepancies arose, we used the following guiding principles to make reasoned decisions about the most likely situation.

Child birth date (month/year) inconsistencies across waves: general rules
- Take child birth date from TLT-2 for women interviewed at TLT-2 and men who only participated in TLT-2 (respid>7000000)
  - Women's child rosters were pre-populated at TLT-2, so any changes to the child roster at TLT-2 were due to the woman actively correcting the roster information. We prioritized these corrections as more likely to reflect the actual situation. If part of the child's birth date or death date was missing at TLT-2 (month or year), that information was pulled from the latest TLT-1 roster, as available.

https://tsogololathanzi.uchicago.edu

- o Men only interviewed at TLT-2 were only interviewed once, so there was no room for discrepancy and their TLT-2 child roster was considered unproblematic.
- Take child birth date from latest TLT-1 wave for all men who participated in TLT-1 (respid>5000000 & respid<7000000)
  - o Men's child rosters were not prepopulated at TLT-2. They were, however, prepopulated at TLT-1, so any changes made were made because men actively corrected erroneous information from earlier in the child roster. We thus, prioritize their latest TLT-1 child roster for births and deaths that occurred within or before TLT-1. Births between TLT-1 and TLT-2 came from these men's TLT-2 rosters.
- Take child birth date from latest TLT-1 wave for women not interviewed at TLT-2

Estimating missing information on child birth and death dates
- In some cases, the birth/death occurred during TLT-1, but either the birth month or year was missing
- When possible, we estimated the month or year of birth/death using information about the wave at which the birth/death was reported, using other variables and sometimes partner reports to verify (n=18 cases)
- In these cases, the birth/death date was set roughly halfway between interviews based on the respondent's interview dates
- We did not estimate dates that occurred prior to the respondent's participation in the study
- 3 dates of death (month or year) were estimated; 15 dates of birth (month and year)

Procedure for estimating waves when birth/death fell in the same month/year of interview (in order listed below)
Sometimes a birth or death occurs in the same month as the interview (e.g., a woman's TLT-2 roster might indicate that a birth occurred in October 2010, the same month that she completed her wave 5 interview).
- coded at same wave if the woman (or man's partner) was not pregnant at that wave
- coded at next wave if the woman (or man's partner) was pregnant at that wave
- coded as same wave if g5=1 at that wave
- coded at next wave if g5=1 for next wave, but not for current wave
- coded wave based on the wave at which the birth was reported on child roster
- if birth/death month/year was missing, coded as wave at which the birth/death was reported

https://tsogololathanzi.uchicago.edu

Example of wave 3 child roster, which would have been prepopulated based on early waves.

SECTION CH: CHILDREN

> INTERVIEWER: Living children reported at prior wave have been pre-populated in this table. Ask about each child in turn. Make corrections where needed. Add children born since last wave (even if died after birth).

	CH1	CH2	CH3	CH4	CH5	CH6	CH7	CH8	CH9	CH10		CH13
								If ALIVE and not living with R		IF DEAD		
	List names of living children from Wave 1 found on coversheet Add children born between waves	Kodi [Dzina] ndi Wamkazi/ Wammuna? Is [NAME] male or female? 1=male 2=female	Kodi [Dzina] anabadwa mwezi ndi chaka chiti? In what month and year was [NAME] born? Probe: what is his/her birthday? If can't remember month, probe for best guess. If still no, enter "88".	Kodi [Dzina] Anakali Moyo? Is [NAME] still alive? 1=Yes 0=No If NO, SKIP TO CH10	Kodi [Dzina] Akudwala? Is __ ill? 1=Yes 0=No IF NO, SKIP TO CH7	Kodi [Dzina] Akudwala Kwambiri? Is__ seriously ill? 1=Yes 0=No	Ngati [Dzina] ali moyo, kodi amakhala ndi inu? Is [NAME] living with you? 1=Yes 0=No IF YES, SKIP TO CH13	Kodi[Dzina] akukhala ndi ndani? With whom is [NAME] living?	Chifukwa chani [Dzina] sakhala ndi inu? Why isn't [NAME] living with you? SKIP TO CH13	Kodi[Dzina] anamwalira liti? When did [NAME] he/she die? List MONTH and YEAR		IVER Verify Status Since Last Wave
1										Month _____ Year _____		
2										Month _____ Year _____		
3										Month _____ Year _____		
4										Month _____ Year _____		
5										Month _____ Year _____		
6										Month _____ Year _____		
7										Month _____ Year _____		
8										Month _____ Year _____		

VARIABLE CODES

WITH WHOM CHILD RESIDES (CH8):
1=father/Babo
2=R's parents/ makolo awoyankha
3=R's siblings/ mchimwene/ mchemwali wawoyankha
4=R's other relatives/ abale awoyankha
5=father's family/ banka la kwabambo ake
6=other (SPECIFY)/ china (tchulani)

REASON CHILD LIVES ELSEWHERE (CH9):
1=Divorce/ ukwati unatha
2=For school/ chifukwa cha sukulu
3=Remarriage/ ndinakwatiwanso
4=Can't take care of him/her/ sindingathe kumusamalira
5= Other (SPECIFY)/ china (tchulani)

IVER VERIFICATION SINCE LAST WAVE (CH13):
1=No Change
2=Information Change
3=Member Left
4=New Member
5=Dead

https://tsogololathanzi.uchicago.edu