****

VERSION 1: 10 June 2021

****

Information relates to data collected in 2017 in Eastern Congo in a targeted 339 villages in the provinces: South Kivu, Haut Katanga and Tanganyika.

Additional, updated versions of the datasets may be posted later.

The PIs can only provide limited support to users.

****

Panel information:

- The 339 targeted villages were also (targeted to be) visited in 2015.
- The 2015 data can be found here: https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/ZFAG4S
- The unique village code (IDV) in 2015 was also used in 2017, and can thus be used to merge the data from both data collection rounds.
- We also aimed to re-interview the respondents from 2015.

****

Instruments:

- "**DRC2017_Instruments.pdf**" contains all instruments: surveys, protocols and the codebook related to the data.
- All variable names in the datasets (.dta format) have question numbers and a keyword that correspond to the instruments.
- All variables in the datasets are labeled, where the first word of the label indicates the survey model from where the data originated in the survey.

****

Datasets:

There are nine separate datasets in Stata format.

There are two identifying variables used throughout the datasets:

- IDV: Unique village ID
- IDS: Unique survey ID

The ID codes are nested: IDS = IDV*10 plus a number.

Generic value labels that are similar across the datasets:

- "-10": “Other”
- "-9": “Don't know”
- "-8": “Not applicable”
- "-7": “Refused to respond”

Villages were visited two times. The first time to distribute the leaflets and the reflection exercise. The second time for additional data collection.

DRC2017_Visit1_Distribution.dta

- This is the only dataset with information from visit 1.
- Distribution of leaflets and reflection exercise: conducted one week before the other surveys.
- At the end of the visit, the team leader filled out this document with information about how the leaflets and the reflection exercise was received
- A total of 339 villages were targeted

DRC2017_EC.dta

- Is the Chief survey (**E**nquete **C**hef in French)
- A total of 339 villages were targeted

DRC2017_EM.dta

- Is the household survey (**E**nquete **M**enage in French)
- The aim was to collect ten households per village, so 339 * 10 = 3,390 households were targeted

DRC2017_EE.dta

- This is the school survey (**E**nquete **E**cole in French)
- One school was targeted in each village. Where it existed, the school that participated in the T2+ program was visited. We thus targeted 339 schools.

DRC2017_ES.dta

- This is the hospital survey (**E**nquete **S**ante in French)
- One hospital was targeted in each village. Where it existed, the hospital that participated in the T2+ program was visited. We thus targeted 339 schools.

DRC2017_EU.dta

- This is the survey with the user committee
- The president of the user committee (or the secretary in the president’s absence) was interviewed.
- There are potentially two user committees to be interviewed: the COPA and the CODESA.
- We thus targeted 339 * 2 = 678 user committees.

DRC2017_ER.dta

- This is the enumerator survey that was filled out at the end of visit 2 to learn whether, in the opinion of the enumerator, the leaflet was spread, the reflection exercise was conducted, etc.

DRC2017_Village_Info.dta

- This file contains additional information from "the field"
- This file indicates the village's treatment status:
    - Treatment=A: are those villages without community nor government involvement
    - Treatment=B: are those villages with community but not government involvement
    - Treatment=C: are those villages with government but not community involvement
    - Treatment=D: are those villages with community and government involvement
- Survey teams were not always able to visit a village. The variable "Insecure" indicates whether a village was not visited for security reasons.

DRC2017_HQ_Coding.dta

- Hardcopy response forms (from the user committees and the service providers) were brought back to headquarters for independent coding of quality.
- This dataset contains more than 339 observations because we can have two forms per village.

****

Note that due to security and logistical reasons, etc. not all the targeted villages / households / schools / hospitals were visited.

****

We left out some data purposefully from these datasets:

- Anything that can identify the village, household or the individual (names, phone numbers, GPS locations, etc.). Identifiable data are available from the Chiefdom level and up.
- Treatment assignment. We hope to write academic papers based on these data.

****

For questions or suggestions, please send an email to petervanderwindt@nyu.edu

**** END ****