Documentation file for Datasets corresponding to paper titled:

“**Turning a Shove into a Nudge? A "Labeled Cash Transfer" for Education**”

*by Najy Benhassine, Florencia Devoto, Esther Duflo, Pascaline Dupas and Victor Pouliquen*

July 1, 2014

**Data Sets**

There are many databases coming from 5 different data collection activities:

* Administrative databases from Tayssir administration:
  + “**cct\_tayssir\_admin\_data\_an.dta**”: Dataset with information on students enrolled in the Tayssir program, including number of days of absence by month and amount of transfer received by month.
* Household level data:
  + “**cct\_baseline\_an.dta**”: baseline household survey data
  + “**cct\_endline\_an.dta**”: endline household survey data
  + “**cct\_monitoring\_endline\_an**”: Monitoring data with reason for survey non-response during endline household survey.
  + “**cct\_hh\_weights\_an.dta**”: Database with sampling weights for household surveys.
* School level data:
  + “**cct\_preliminary\_survey\_an**”: data from the school-level survey conducted in preparation for the study in 2008.
  + “**cct\_school\_visits\_an.dta**”: Data from the 7 school visits done between 2008 and 2010.
  + “**cct\_correction\_dropout\_date\_an.dta**”: database created to input corrections to the school visit data about dropout dates. When dropout dates were not consistent between two visits, data were manually checked to determine which data was the most accurate.
  + “**cct\_stratum\_an.dta**”: database at the school sector level with the stratum used for the randomization.
* ASER data:
  + “**cct\_aser\_an.dta**”: data from tests in mathematics administered to one child per household during the endline household survey.
* Awareness surveys:
  + “**cct\_knowledge\_households\_year1\_an.dta**”: data from the awareness survey administered to a subset of households in 2009.
  + “**cct\_knowledge\_teachers\_year1\_an.dta**”: data from awareness survey administered to a subset of teachers in 2009.
  + “**cct\_knowledge\_teachers\_year2\_an.dta**”: data from awareness survey administered to subset of teachers in 2010.

All surveys were administered in French. All databases are labeled with the exact English translation of the questions.

All survey instruments used for the research are included in the “Survey\_instruments” folder.

**Codebooks:**

For all databases, an Excel file describes all variables included in the datasets. Codebooks files are available in the “Codebook folder”. They are named with the prefix “codebook” followed by the name of the database it describes.