Documentation regarding datasets and do-files Macours, Schady, Vakis, 2011.

Do-files:

random_selec_m.do : code for descriptive statistics of baseline characteristics and randomization results (table 1)

frequency_delays_m: code to calculate the frequency of delays compared to the international reference population (table 2)

main_analysis_m.do: code for the main analysis of impacts on test scores (table 3 and 4) and consumption (table 5)

main_analysis_mt13.do: code for the analysis comparing test scores of the lump sum payment package with the basic package (table 6)

econact_time_mom_m: code for comparison economic activities and care giving mothers with lump sum payment package compared to basic package (table 7)

analysis_inputs_m: code for analysis of impacts on intermediary inputs (table 8)

gradient_fig_m.do: produces the figure of socio-economic gradients (figure 1)

Data-sets:

macoursetal_main.dta: This is the main dataset for most of the analysis

macoursetal_t13.dta: This is the dataset used for table 6, comparing only the lump-sum package with the basic package. Compared to the main dataset, this dataset contains a handful fewer observations for some of the test scores, in order to be able to estimate the variance-covariance matrix of the SUR regression. This is due to unique values for some of the ages in this analysis that compared to the main analysis is on a reduced sample.

Note on variable names: when the year is not specified in the label, the last 2 numbers of the variable name indicate from which year of the survey the variable comes. _05 is data from the 2005 round (baseline); _06 is data from 2006 (first follow-up survey); and _08 is data from 2008 (second follow-up survey).