This folder contains the programs necessary for reproducing the results shown in the paper "Distributional Effects of Education on Health."

The analysis uses data from the UK Biobank. See http://www.ukbiobank.ac.uk/register-apply/ for how to gain access to the data.

The do-file "crdata.do" cleans the data. The do-file "analysis.do" produces the figures and tables shown in the paper.

There are special programs and datasets to run the hypothesis tests shown in Table 3. In particular, the p-values are generated using the ado-file "WAD_ROSLA_RF_gen_lin" in combination with the following datasets:

- Wad_stats_5yr_lin_t1957_body_size_full
- Wad_stats_5yr_lin_t1957_lung_function_full
- Wad_stats_5yr_lin_t1957_blood_pressure_full

- Wad_stats_5yr_lin_t1957_body_size_half
- Wad_stats_5yr_lin_t1957_lung_function_half
- Wad_stats_5yr_lin_t1957_blood_pressure_half

The datasets contain simulated statistics which were generated using the ado-file "WADdraw.ado"