EXERCISE SHEET 7: LA FERRARA ET AL. (2012): TELEVISION AND FERTILITY

Main learning goals: 1. Understanding individual panel data structure

2. Including year fixed effects

3. Evaluating LP models and their goodness of fit

4. Introducing interaction effects

5. Producing nicely formatted tables with `outreg2`

6. Regional fixed effects in individual data vs. regional panel

Hand-in: As usual: do-file and the corresponding log-file (saved in .log-format) containing the following steps (numbered in the do-file).

1. Open the dataset `LCD_ind.dta` and generate the main explanatory variable: `globo`, an indicator denoting Globo coverage being present one year before. Also generate an indicator for regional AMC areas being above the 95th percentile of area distribution and make sure to exclude them from the analysis. Save the new dataset under `LCD_new.dta`.

2. Re-estimate columns (1) to (6) of table 2 by a linear probability model, using sampling weights. Include year fixed effects and use `areg` to eliminate regional fixed effects. Store your estimation results.

3. Estimate column (3) "by hand", meaning that you should implement the fixed effect regression by de-meaning the variables.

4. Investigate how many predicted values of the linear probability model of model (6) are outside of the unit interval. Next, evaluate the goodness of fit for the linear probability model by implementing the following strategy: Predict the outcome variable for column (6) and assign values of 0 or 1, depending on whether the predicted values are below or above the threshold of 0.5. Then, compute a share of correct predictions, given that the true outcome variable is 0 or 1. Briefly discuss the results.

5. Re-estimate columns (1) to (3) of table 3. Interpret the economic meaning of the interaction effects in your do-file.

6. Export the estimated coefficients of columns (4) to (6) from exercise 2 in a nicely formatted and labeled table using `outreg2`.

7. Use `collapse` to generate a regional level panel dataset, containing the average values of the dependent and explanatory variables and population size (remember the sampling weights); save it under a different name.
   Use `xtset` to define the resulting dataset as panel data and `xtdes` to inspect it.

Further comments: Submit 3 files (.do, .log, .pdf(with an amazingly beautiful table)) to Ilias. Usual procedures apply.