EC 590

Data assignment 2

In this assignment, you must work with STATA or R to analyze the data from a pilot experiment carried out in Tanzania in 2011 to test the effect of family obligations on the ability to save.

The experiment worked as follows. Researchers randomly selected approximately 40 individuals in 4 rural villages to answer a questionnaire; respondents were either men or women, living in a household with a spouse. At the end of the interview, the respondent was given an envelope with 10,000 shillings ($7). The respondent was told that the research team was to return in three weeks for a follow-up visit. The researchers were going to verify how much was kept in the envelope, and provide the respondent with an additional 20% on whatever was left inside it (up to a maximum of 2,000 shillings). Respondents were allowed to keep all the money given to them.

The experiment had three relevant features.

- During the baseline interview, researchers played some behavioral games with the respondents to capture their time preferences (patience, impatience, time consistency) similar to Ashraf, Karlan and Yin.
- Before giving the envelope, researchers wrote down the individual bill code number. During the follow-up verification visit, they verified whether the bills in the envelope were put there by the researchers or placed there by the respondents. By checking code numbers, they could check whether the respondent (or someone in the family) took money out of the envelope.
- Respondents were randomly assigned to two treatments.
    1. The first treatment, which I call $PRIVATE$, corresponds to the case where the recipient was privately given the envelope, without the knowledge of relatives or family. The envelope was nondescript and had no clear markings.
    2. The second treatment, $PUBLIC$ involved handing over the envelope to the recipient in front of at least some members of the family. The envelope had a transparent window, through which one could see that there was money.
- Researchers randomized who in the household would get the envelope. Sometimes it was the wife, and other times it was the husband.

The objective of the exercise is to determine whether savings behavior is affected by other people's knowledge of their savings. The hypothesis is that it's hard to hold on to savings if people know you have money, and that it may be harder for women than for men.

The data is contained in the tex file *moneyexperiment.dta*. The treatment is written in the dummy *public*.

Keep in mind the following features of the data:

- The sample size is very small–this was a pilot. A lot of the results from the pilot will not be statistically significant. You can rely on point estimates for your analysis, and put less emphasis on significance.
- The database includes information collected at baseline (before the envelopes of cash are given out) and at endline. Information collected at endline includes the contents of the envelope and what happened to the money (Variables hasenvelope until pcsaved). The rest (starting with variable shock) is information collected at baseline. The variable labels explain what each variable is.

Questions to answer

1. The first task is to *verify randomization*, by showing whether $E(X^{Pu}) = E(X^{Pr})$.

   Create a summary table with summary statistics for relevant baseline characteristics. Make sure to include the average and standard deviation for each variable and each treatment group. Show standard deviations of the difference in means between the public and private treatment groups.

2. The rest of the analysis is for you to decide. You've got the dataset. You are now free to carry out whatever analysis you'd like! Work in teams to decide on an empirical strategy (regressions, outcome variables of interest etcetera). Make a short report with at least:
   - A justification for your analysis (what are you testing?)
   - At least one table containing regression results
   - A short writeup, summarizing the results and providing a short commentary.