Let’s prepare a short report summarizing your findings about the impact of a hypothetical policy initiative! A description of the study design is included below. You can use any statistical software of your choice to process and analyze the data for the study :)

By the end of the quarter, please submit:

1. All the code used to process and analyze the data. We will run and read all the codes in detail, so make sure it is reproducible, well-commented, and organized and that you include all files necessary to run it.
2. Any derived data sets created to analyze the data.
3. A short report (.pdf) written in LaTeX, with a regression table, describing the data and the methods you used to analyze it and discussing your findings. Write this report with an audience of policymakers with training on impact evaluation methods in mind. Its contents should be enough for them to understand the intervention, the methods used to evaluate it, and the evaluation results.

**Study Description**

This study evaluates an education subsidy program that aims to reduce financial barriers to schooling by providing in-kind transfers—non-cash educational materials and supplies—to students in public schools. The program was implemented over a three-year period (2010–2012), targeting all students who were enrolled in grade 6 at the beginning of 2010. By offsetting the cost of education, the intervention sought to encourage students to remain in school through the critical transition from primary to lower-secondary education, a stage at which drop-out rates tend to rise, especially among low-income households.

Each year, participating students in the treatment group received the subsidy package at the start of the school year. These in-kind transfers typically included school uniforms, textbooks, notebooks, and other essential materials required for learning. The intention was to reduce the indirect costs of education that often drive school evasion—such as the expense of supplies or the need for children to contribute to household income. By improving access and easing financial pressure, the program aimed to keep children in school longer and delay early entry into the labor force or family formation.

To rigorously measure the program’s impact, researchers designed a randomized controlled trial (RCT) at the school level. Schools within each participating district were pairwise-matched based on similar baseline characteristics (such as size, location, and socioeconomic context). Within each pair, one school was randomly assigned to the treatment group (receiving the subsidy) and the other to the control group (not receiving it). This design ensured that any systematic differences in outcomes between the two groups could be attributed to the intervention rather than to pre-existing conditions.

The primary outcome of interest was school evasion (drop-out)—whether students left school before completing the expected years of study. The study also examined secondary outcomes reflecting broader social consequences of continued schooling, including teen pregnancy and teen marriage, which are often correlated with early school departure and socioeconomic vulnerability.

To assess these outcomes, data collection occurred in three phases:

1. Baseline survey (2010): Conducted before the start of the intervention, it gathered detailed demographic, educational, and socioeconomic information about both students and schools.
2. First follow-up (2013): Conducted immediately after the final year of subsidy provision to measure short-term program effects on enrollment and early behavioral outcomes.
3. Second follow-up (2015): Conducted two years later to capture long-term effects, including whether initial gains in school participation persisted and whether secondary social outcomes (such as early marriage or pregnancy) were affected.

During each follow-up, researchers tracked all students originally enrolled in grade 6 in 2010, recording whether they were still in school, had dropped out, were married, had children, or were pregnant. This longitudinal panel structure allowed for an evaluation of both immediate and lasting impacts of the subsidy program.

Overall, this experimental design provides robust evidence on how in-kind education subsidies influence school retention and early life outcomes, offering insights for policymakers seeking cost-effective interventions to improve educational attainment and reduce social risks among adolescents.

**Report content**

Think of this report as a policy brief to be shared with policymakers who have training on impact evaluation methods and have no previous knowledge of the intervention or the study. The report must contain at least one table and one visualization, but you may include more exhibits. You should also feel free to describe any additional data or information you would like to use for the study and further research questions you think are relevant or interesting.

Your report should try to answer the following question.

* What were the effects of the intervention on the outcomes of interest after 3 and 5 years?
* Describe and explain your choice of regression specification, unit of analysis, controls variables, and post-estimation tests. Making use of graphs and tables to illustrate your answers will be appreciated. ##

**Data description**

The materials for this assessment include 4 data tables and a data dictionary describing the fields contained in each of them.

* **schools.csv**: school-level baseline and treatment assignment data
* **school\_visits\_log.csv**: A log of the dates when each school was visited at each round of follow-up
* **student\_baseline.csv**: student-level data with demographic information about all students in the study cohort, observed at baseline
* **student\_follow\_ups.csv**: student panel measuring outcome values after 3 and 5 years of the start of the program The data can be downloaded by clicking on this link: [Raw dataLinks to an external site.]