Development Impact Lab Data Task

Professor Michael Kremer

Alejandro Ortiz

Table of contents

Disclaimer: The responses and content presented in this document are solely my own work. They reflect my personal understanding, analysis, and perspective. All material contained herein has been created independently by me.

Question 1

Are schools comparable across treatment groups? Are student cohorts?

To assess the comparability of the groups, I conducted four statistical tests: a simple joint significance F-test, a Wald test for robustness (given the F-test's application to a heteroskedastic Linear Probability Model [LPM]), individual difference-of-means tests for the continuous variables, and individual chi-squared tests for the discrete variables.

The results are as follows:

Schools

For schools I calculated the following model on the baseline data:

$$Trearment_i = \beta_0 + \gamma_{d,i} + \beta X_i + \varepsilon_i$$

Where $\gamma_{d,i}$ represents district fixed effects, and X_i are the school-level covariates including the number of: teachers, female teachers, female studens, male students, number of schools in a 2 km radius, and number of latrines as well as:

- Urban dummy

- Female head teacher dummy

- Average teacher age

- The average student score

the F-statistic for this model has a value of 1.005 on 11 and 143 degrees of freedom. The corresponding p-value of this statistic is 0.445. However, because this is a Linear Probability Model (LPM), it is heteroskedastic and can yield out-of-bounds predictions. The heteroskedasticity is problematic because it can bias the standard errors, thereby rendering the F-statistic unreliable and typically increasing the probability of a Type I error. To account for this, the model is estimated using heteroskedasticity-robust standard errors, and a second test is performed. This is the Wald test. Even though this test is asymptotically equivalent to the F-test, the small sample size produces different results and serves as a robustness check on the joint significance of the model. The Wald test has a p-value of 0.17.

To test for individual significance, I conduct a series of T-tests on continuous variables

Table 1: Difference of means T-test results

variable	control_mean	treatment_mean	p.value
n_teachers	14.536585	14.180723	0.5788467
n_teachers_fem	7.146342	7.204819	0.9259432
n_students_fem	254.963415	231.674699	0.1283670
n_students_male	253.573171	232.951807	0.1713085
n_schools_2km	1.951220	2.012048	0.8315288
av_teacher_age	39.612828	39.961537	0.4866534
av_student_score	249.271605	255.225000	0.1880104
n_latrines	10.948718	11.626506	0.4334080

For discrete variables I do Chi-squared tests. For students I repeat this process. In all these tests p-values are never statistically significant at the 10% level. Given that all tests indicate balance between treatment and control groups at baseline I conclude that randomization was successful. Even is some p-values are a bit low there is no strong evidence of imbalance at the school or student level.

Note: More analysis in the script file, but I am short on time.

Question 2

What were the effects of the intervention on the outcomes of interest after 3 and 5 years?

Figure 1 shows the point estimates and 95% confidence intervals for the treatment effects on the main outcomes of interest after 3 and 5 years. The results indicate that the intervention did not have a statistically significant effect on school evasion rates, pregnancy rates or marriage rates after 3 years. After 5 years it did have a statistically significant effect on dropout and marriage rates, but not on pregnancy rates.

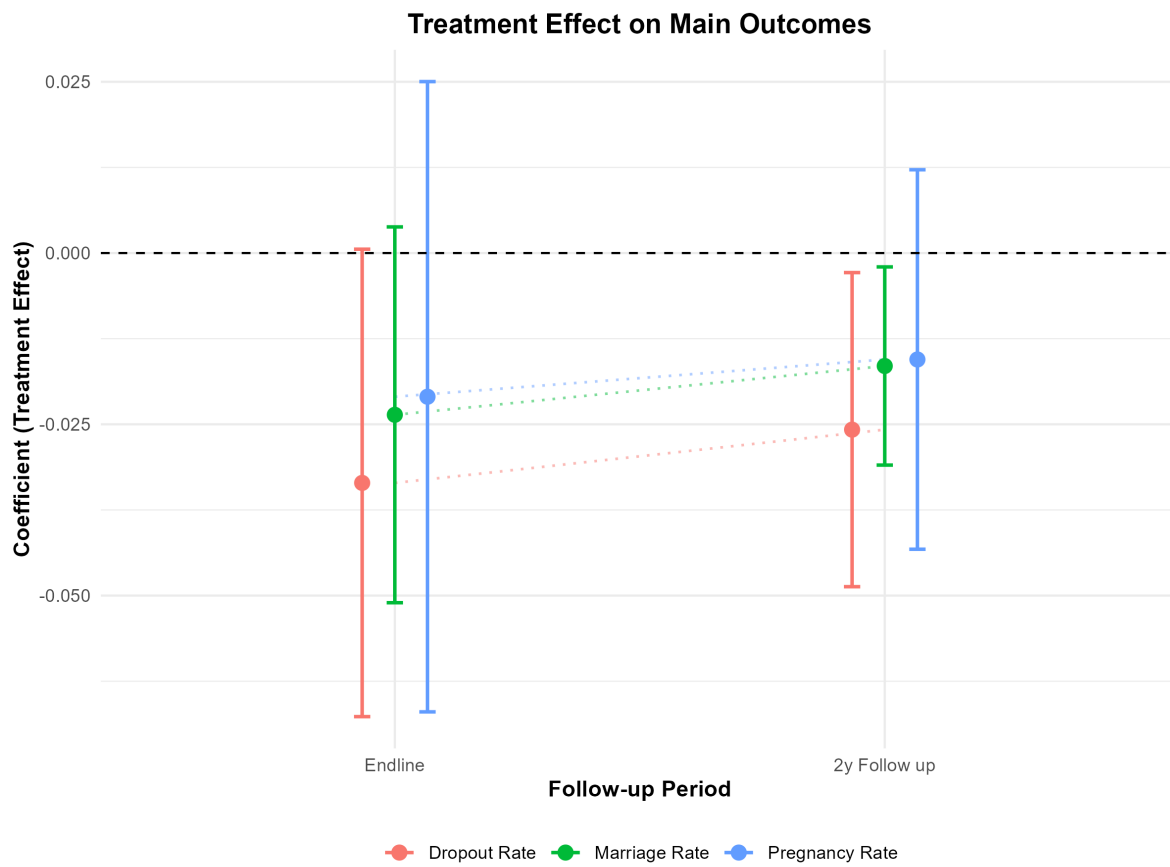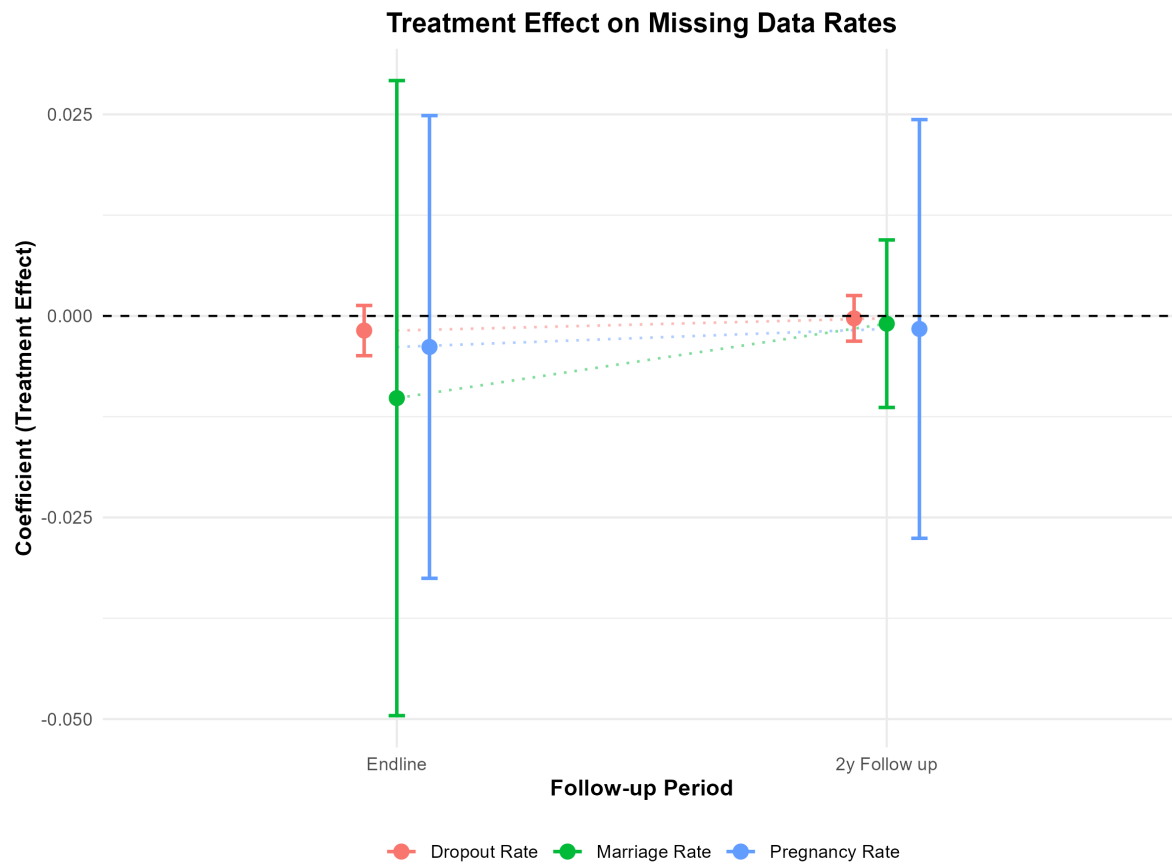Figure 1: Treatment Effect on Main Outcomes

Treatment Effect on Main Outcomes

Figure 2: Treatment Effect on Missing Data Rates

Question 3

Were the effects on school evasion different for girls and boys?

They were not. Please see results in the script file.

References