TD Econometrics of causality
Exercise 4 -- RCT

For this exercise, we will be replicating and exploring some of the results of Lorenzo Casaburi and Tristan Reed "Using individual-level randomized treatment to learn about market structure" (2021). This paper uses an RCT to measure the causal impact of the allocation of subsidies on market structure (price, quantities purchases, and payment in advance in particular)

The market that is studied is the cocoa market of Sierra Leone. In this market, cocoa is produced by a large number of farmers who sell to some traders who themselves sell to a few wholesalers.

In our data, we have 1818 farmers, 80 traders (40 treatment and 40 control traders), and 5 wholesalers. The treatment of this experiment is a subsidy of 5% on the price that wholesalers pay to traders.

To run our analysis, we will use 4 databases :

- In the first one, each row is an observation, which corresponds to a farmer.
- In the second one, each row is an observation, which corresponds to a shipment delivered by a trader to a wholesaler
- In the third one, a row is an observation, which corresponds to a week. There are 80 traders and 11 weeks of intervention. In this database we have 880 observations: one for each week for each trader
- In the fourth one, a row is an observation, which corresponds to a traders

Key variables :

If the variable end by _B it's a baseline variable

Variable name	Label	Note
pair	2 traders that have been pair randomized together will have the same "pair numb	Baseline / Endline
treat	Treatment Trader	Baseline / Endline
after	=1 if date is after 15 October 2011 / start of treatment	
week	Number of the week	
N_traders	How many traders are in the village ?	Baseline
N_treat	How many treated traders in the village ?	Baseline
N_treat_oth	How many other traders are treated in this village	Baseline
N_traders_oth	Number of other study traders in the village	Baseline
N_suppliers_village	Number of farmers in this village	Baseline
Luawa	=1 if the village is in Luawa	Baseline / Endline
Mandu	=1 if the village is in Mandu	Baseline / Endline
Njaluahun	=1 if the village is in Njaluahun	Baseline / Endline

Nongoawa	=1 if the village is in Nongoawa	Baseline / Endline
Upper_Bambara	=1 if the village is in Upper Bambara	Baseline / Endline
milesnearest	Distance of the village to the nearest town (miles)	Baseline / Endline
lshare_village	Share of farmers given credit since March for each village	Baseline
age_trader_B	Age, years of the trader	Baseline
years_experience_trading_B	Years trading cocoa	Baseline
years_selling_wholesaler_B	Years selling to study wholesalers	Baseline
nice_floor_B	Cement floor in traders house	Baseline
mobile_trader_B	Are traders mobile phone owner ?	Baseline
storage_trader_B	Does trader have access to a storage facility ?	Baseline
N_villages_B	Number of villages traders are operating in	Baseline
N_farmers	Number of farmer trader is buying from	Baseline
number_bags_solde_grade_A_B	Grade A bags sold to wholesaler in 2011	Baseline
weight_cacao_sold_gradeA_B	Baseline values of pounds of cocoa sold Grade A	Baseline
lshare_trader_B	Share of farmers given credit since March for each trader(baseline)	Baseline
lend	=1 if the farmer received an advanced payment during the experiment period	Endline
farmers_trader_village	Mean of number of farmers per village (for each trader)	Baseline
price_farmer	Price paid to farmers	Data collected during the 11 week of the experiment
total_weight_week	Quantities of cacao purchased by traders to farmers (pound)	Data collected during the 11 week of the experiment

Question 1 : Using the database data_traders, data_traders_week and data_shipment, check whether the treatment and control groups were similar at baseline (use the same variables as those reported by the authors in Table 1 of the paper). For each of the baseline variable, report the mean value in the treatment group and in the control group, along with the difference between the groups and a test for the statistical significance of this difference.

a) First, you need to determine the mean value of the variables for the treatment and the control group. Be careful, you have to focus on the baseline value.
b) Then you have to find if the difference between those mean are significant.

Interpret your results: did the randomization generate similar treatment and control groups? If not, what should you keep in mind for the rest of your analysis?

Question 2. What is the impact of the intervention?

Now you will replicate some regression from the article. To do that, look at the different tables and read the text that lies below each one.

Note 1: For each regression, you have to add pair fixed effect. To add them to the regression you have to control for the variable pair.

Note 2: Week FE means week fixed effect. To add them to the regression you have to control for the variable week.

a. Using data_shipment, replicates the results from the regressions of table 2, interpret the results (for this question, you need to only keep the observation where after = 1 (use filter)).
b. Using data_farmers, replicates the results from the regressions of table 3, interpret the results.
c. Using data_traders_week, replicates the results from the regressions of table 4, interpret the results (for this question, you need to only keep the observation where after = 1 (use filter)).

According to this empirical analysis, what is the overall impact of the subvention on the market structure ?