

The Effect of Improved Maize Seeds and Fertilizers on Crop Yields

Assignment: Identification and Causal Inference

Jonas Schmitt, Sina Streicher, Owen Cortner, Philipp Kronenberg

February 10, 2021

1 Introduction

We examine the effect of a package of improved maize seeds and fertilizers on the harvested amount of maize of small-scale agricultural households. Maize is an important source of carbohydrates in the imaginary country, which we examine. The result of the study should evaluate if the package is manageable and effective with a limited amount of extension service and if it can help to improve the food self-supply situation of those households.

The package of improved maize seeds and fertilizers can be picked up for free at several supply station throughout the research area. Agricultural households have been informed in person by a group of young students about the process and where they can pick up the package. The information was spread by considering a federal list of all agricultural households in the study area. The total number of agricultural households in the list was 10.000. 5.000 households were randomly picked and received a voucher to pick up the seed + fertilizer package.

However, we observe that only a subgroup of agricultural households takes the treatment by picking up the seed + fertilizer package. We guess that e.g. the distance to the supply stations of a household, the education (as read/write dummy) or number of very young children (< 10 years) influences the choice to pick up the package and that the selection of treatment is therefore not random.

Unfortunately, we have no information about the soil quality, which has an influence of the amount harvested even without any treatment. Therefore, we also face the issue of omitted variable bias.

a) Experiment

In the experiment, we would randomly distribute the package of improved seeds + fertilizers to a treatment group of agricultural households e.g. by an address-list and measure the causal effect by comparing the harvested amount between treated and untreated households. The treated households additionally receive a voucher for farm equipment after the harvest, if they apply the package to their best of knowledge and without sharing it with their neighbours or any family members. Thereby, we want to avoid any spill-over effects between treatment and control group.

b) Optimal Data Set

The optimal dataset would reveal all the relevant variables from an agricultural perspective (soil quality, topographic information, weather data on each farm etc.) and socioeconomic variables like income situation, education or family composition.

c) Threats to Identification

Omitted variable bias \Rightarrow control for all confounders, factors that determine assignment to treatment correlated with the outcome. The second challenge is the validity of the instrument (internal and external validity) \Rightarrow different weather effects during the experiment on the farms?

2 Data Generation and Model Specification

$$harv_i = \beta_0 + \beta_1 area_i + \beta_2 rain_i - \beta_4 child_i + \beta_5 exp_i + \Delta_i D_i + e_i. \quad (1)$$

Amount harvested (kg) $\sim U(30, 800)$

Cultivated area with maize (m²) $\sim U(100, 5000)$

Sun-hours during growing period (h) $\sim U(400, 600)$

total rainfall during growing period (mm) $\sim U(600, 850)$

(Here: if we need to adjust, we can also assume that the soil quality of the study area has no differences between the farms. However, I thought maybe it would be interesting for task 5 to consider a farm fixed effect to control

Number of children under 10 years $\sim U(1, 5) \Rightarrow$ Idea: more children influence the time, which can be spend on the field to look after the plants; can change in both direction in 5.)

Years of experiences $\sim U(1, 35) \Rightarrow$ increases in 5.)

Education $\sim U\{0,1\}$, where 1 is good and 0 bad reading and writing skills

3 Assignment into Treatment

$$D_i^* \sim U(-1, 1)$$

$$D_i = \begin{cases} 1 & \text{if } D_i^* \geq 0 \\ 0 & \text{if } D_i^* < 0 \end{cases}$$

$$D_i^* = \gamma_0 - \gamma_1 dist_i + \gamma_2 educ_i - \gamma_3 child_i + \gamma_4 exp_i + u_i. \quad (2)$$

$$D_i = \begin{cases} 1 & \text{if } D_i^* \geq 0 \\ 0 & \text{if } D_i^* < 0 \end{cases}$$

Education (can read/write) $\sim U(1, 0) \Rightarrow$ is supposed to have an influence on farmers evaluation regarding the effectiveness / advantage of improved seeds + fertilizer package.

4 Selection Bias

text...

5 Difference in Difference

Test citation:

bla bla Foroni et al. (2011) bla bla.

References

Foroni, C., Marcellino, M., and Schumacher, C. (2011). U-MIDAS: MIDAS regressions with unrestricted lag polynomials. Discussion Paper Series 1: Economic Studies 2011,35, Deutsche Bundesbank.