**Exercise that Involve Quantitative Analyses**

This document has 2 pages. Please read everything carefully before starting to work on this exercise.

**1 Data Preparation**

Load the data set data\_exercise.csv that we provided to you. This data is based on a random sample of farm households in a district in Tanzania. The data set includes the following variables:

**hhNum** The number of the household in this data set

**Village** The village in which the household is located

**Sex** The sex of the household head

**Age The** age of the household head

**Educ** The number of years of education of the household head

**hhSize** The number of household members.

**landOwned** Land area owned by the household [acres]

**landCultivated** Land area cultivated by the household [acres]

**fo** Membership in a farmer organisation [1 = member, 0 = not member]

**Income** Household income [1,000 TZS]

The enumerators made a mistake in villages “D” and “G” and reported the **income** in TZS instead of in 1,000 TZS in these two villages. You need to correct these data errors in a transparent way.

**2 Graphical Comparisons**

Create a figure that compares the distribution of household incomes between members of farmer organization and non-members of farmer organisations. Interpret this figure.

**3 Proportions of Members in Farmer Organisations in Each Village**

Calculate the proportion of households in each village that are members in a farmer organisation and add this variable to the (household-level) data set. Create a figure that illustrates the relationship between a household’s membership in a farmer organisation and the proportion of households in each village that are members in a farmer organisation. Interpret this figure.

**4 Effect of Membership in Farmer Organisations on Household Income**

Use a statistical/econometric method to estimate the effect of membership in a farmer organisation on the household income. Briefly describe your model specification and method, present the regression results in a table, interpret the regression results, and briefly discuss to which extent the assumptions for obtaining unbiased estimates of the causal effect of membership in a farmer organisation on the household income are fulfilled.