<div align="center">

ENV ECON 118 / IAS 118 - Introductory Applied Econometrics

Assignment 2

Due <u>Tuesday September 29</u> at the beginning of class

</div>

Exercise 1: Do Microfinance Loans Affect Expenditure?

Most of the world's poor have limited access to formal credit. Traditionally, they have had to resort to their social networks (family, friends) or local moneylenders that charge exorbitant interest rates (upwards of 100%). This often prevents individuals from making investments in potential businesses or productive assets. In the last 15 years, microfinance institutions (MFI) have emerged all across the developing world to address this problem. The basic model of a microfinance institution (such as the Grameen Bank) is to provide small loans to a group of potential borrowers at much lower interest rates. The number of very poor families with a microloan has grown exponentially: from 7.6 million in 1997 to 137.5 million in 2010. Micro-credit has been heralded as a major advance in the reduction of global poverty. However, in recent years, critiques have emerged accusing micro-finance institutions of acting irresponsibly by holding the poor to very strict repayment schedules and charging unreasonably high interest rates.

So where does this leave us? Does micro-credit help or hinder? In the January 2015 issue of the *American Economic Journal: Applied Economics*, six papers evaluating the merits of micro-credit were released. We will explore the results from one of these papers by Augsburg et al., which evaluates a microfinance institution in Bosnia and Herzegovina.

Guidelines

This assignment should be completed using STATA, and you are encouraged to use a .do file to write your code for the exercise. To write notes in the .do file that STATA will not read as commands, type a * at the beginning of each line in which you have written a comment. This will help you keep track of which question you are trying to answer, the purpose of each command, etc. To get started:

- The first thing you should do in your .do file is to change directories. If you came to one of the Stata tutorials, somewhere on your computer you should have a folder called "EEP 118". Inside that folder, make a subfolder called PS2.

- When you download the data from the course website, save it in this folder. To change the directory to this folder, you will use a command that is something like the following (but your file path will vary):

$$\text{cd} \quad \text{``} C: \backslash EEP\_118 \backslash PS2\text{''}$$

- Then turn on a log file:
<div align="center">log using PS_2.txt, text replace</div>

- Display your present working directory to make sure the cd command was successful ("pwd")

- This time you are opening a Stata-formatted ".dta" file, so you will use the " use" command instead of insheet/import command that was applied to a .csv spreadsheet in the last assignment. Your command should look like this:
<div align="center">use "PS2_MFI.dta", clear</div>

- For Exercise 1 you need to submit your log file from your Stata work in addition to your written answers.

Stata Tips

- To count observations, use the command "count".

- To create a new variable named var1, use the command "generate var1". Open the Data Editor (Browse) to see and check what you have done.

- To create a scatter plot of variables y on x, use "scatter y x" and to create a histogram of a variable x, use the command "histogram x".

- The command "tabulate" lists all values a variable takes in the sample and the number of times it takes each value.

- To summarize data for a specified subset of the observations, you can use "summarize" along with an if statement.

Data Description

The data for this exercise comes from a study conducted in Bosnia and Herzegovina investigating the effects of a small loan on access to liquidity; self-employment and income; labor supply; expenditure and savings. These were individual-liability loans with monthly repayments and an interest rate of 22%. The sample consists of potential borrowers (who were just marginally eligible for loans). Approximately half the sample was randomly selected to receive the loan (the treatment group), while the other half did not receive anything (control group). You have the entire sample of individuals (both in the treatment and control group) that the researchers used for their analysis. The respondent (= the loan applicant) answered questions about the household they belonged to as well (no two respondents are from the same household). The "PS2_MFI.dta" file includes the following variables:

- *treatment* : dummy equal to 1 if the respondent is in the treatment group (which received a loan)
- *resp_female*: dummy equal to 1 if the respondent is female
- *resp_age*: the respondent's age
- *hhmem*: number of household members
- *hhmem_adults*: number of adults in household (≥ 14)
- *hhmem_children*: number of children in household (<14)
- *hhmem_elderly*: number of elderly in household (>64)
- *total_exp*: total annual household expenditure in Bosnian Convertible Marka (BAM)
- *food_exp*: annual household expenditure on food in BAM
- *nondur_exp*: annual household expenditure on nondurables (rent, fuel, transport, clothes, insurance...) in BAM
- *dur_exp*: annual household expenditure on durables (education, furniture, vehicle...) in BAM
- *temp_exp*: annual household expenditure on temptation goods (cigarettes, tobacco, alcohol) in BAM

The prevailing exchange rate at the time of the study was approximately US\$1 = 1.63 BAM.

Question 1   First, write a short paragraph that describes your data. In particular:

(a) How many respondents are in your data set? How many respondents are females? What is the age-range among the respondents in the sample? What is the median and mean age of respondents in your sample?

(b) Construct a variable *total_exp_pc* equal to total expenditures *per capita* in BAM. Plot a histogram of this constructed variable. What is the range of household total expenditures per capita? (You may want to refer to US\$ in the discussion, so as to make sense of the income level of these MFI clients).

(c) Calculate the proportion of household expenditures spent on durable goods (education, furniture, repairs, vehicle, appliances). You will need to create this new variable in Stata. What is the mean? What is the median? Compare the mean value of these proportions among treatment and control households; among households where the respondent (i.e. the loan applicant) is a female vs. a male.

Question 2 Now estimate the following model of what economists call "temptation" expenditures (cigarettes, alcohol, and tobacco):

$$temp\_exp = \beta_0 + \beta_1 total\_exp + \beta_2 treatment + u \tag{1}$$

(a) Interpret your $\hat{\beta}_1$ and $\hat{\beta}_2$ coefficients, remembering the triplet S(ign), S(ignificance), and S(ize), though you don't need to comment on significance in this problem set.

(b) Taken together, how well do total expenditures and treatment predict temptation good expenditures?

(c) What is the predicted level of temptation expenditures for a treated household with total annual expenditures of BAM 12000?

Question 3 Consider the following model of food and non-durable goods (rent, fuel, transport, clothes, recreation, magazines, insurance) expenditures

$$\ln(food\_exp) = \beta_0 + \beta_1 \ln(total\_exp) + \beta_2 treatment + u \tag{2}$$
$$\ln(nondur\_exp) = \beta_0 + \beta_1 \ln(total\_exp) + \beta_2 treatment + u \tag{3}$$

(a) Estimate the two models, and interpret β_1 and β_2 in equation (2) only.

(b) Comparing the estimates of $\hat{\beta}_1$ in equations (2) and (3), do your results seem reasonable? (Hint: What does it mean for an elasticity to be greater or less than 1?)

(c) Using the results from estimating equation (3), how would you expect food expenditure for a treated household to change if its total expenditure increases by 25%?

(d) How does treatment status (having a loan) affect food or non-durable consumption (interpret $\hat{\beta}_2$ in both cases i.e. equations 2) and 3))?

Question 4 We will now explore the role of household size in food consumption. Consider these two models:

$$\ln(food\_exp\_pc) = \beta_0 + \beta_1 \ln(total\_exp\_pc) + u \tag{4}$$
$$\ln(food\_exp\_pc) = \beta_0 + \beta_1 \ln(total\_exp\_pc) + \beta_2 ln(hhmem) + u \tag{5}$$

(a) Estimate equations (4) and (5).

(b) Interpret the estimated parameters of equation (5). The results you are finding on the role of household size may a priori seem suprising. Try to think about a scenario whereby two households of the same size, with the same total expenditures per capita decide to move in together. What does your estimation predict about food consumption per capita?

(c) How did your estimate $\hat{\beta}_1$ change between equation (4) and equation (5)? Without performing any calculations, what information does this give you about the correlation between total expenditure per capita and household size? (Explain your reasoning in no more than 4 sentences.)

(d) Predict the expected value of food expenditure of a household with 3 members and total annual expenditure per capita of BAM 3000 using your estimates from equation (5).

Department of Agricultural and Resource Economics   Elisabeth Sadoulet and Erin Kelley
University of California at Berkeley   Fall 2015

Question 5 A country's dependency ratio is the ratio of old and young dependents (dependents are those not in the labor force) to the working-age population. A similar measure could be constructed for the household:

$$hhdr = \frac{\text{hh members under 14 or over 64}}{\text{hh members aged } 14-64} \tag{6}$$

Equation (5) does not capture how the composition of a household, i.e. the characteristics of the members, is associated with food consumption per capita. You suspect that the structure of the family affects food expenditure per capita controlling for the log of household size and the log of total expenditure per capita (think about how children and older people might consume less food than adults; and how larger households might have more children). Specifically you hypothesize that a higher dependency ratio is associated with lower food expenditure per capita.

(a) Write an equation you could estimate that would test this hypothesis.
(b) Estimate the equation in part (a). What can you conclude about the hypothesis?

Exercise 2: Can Sweatshops in Bangladesh help Women?

It is well known that working conditions in Bangladesh's garment industry are bleak. Employees work long hours, for very little pay, and in very dangerous working conditions. In 2013, these conditions were brought to the forefront of most news outlets when a factory in Bangladesh collapsed, causing over 1000 casualties. This sparked a huge debate within the international community. Some called for a boycott of all companies sourcing their clothes from Bangladesh. Others argued this wasn't an appropriate solution, as the garment industry provides jobs to many women who would otherwise be unemployed. This begs the question: what is the impact of the ready-made garments export industry on employment outcomes, schooling, marriage, autonomy and fertility outcomes for women? We start by estimating:

$$fertility = \beta_0 + \beta_1 Garment\_village + u$$

Where $Garment\_village = 1$ if a women is living in a village with a factory

(a) What do you expect the sign of β_1 to be? Why? Can you think of a reason why it could have the opposite sign?
(b) List three other factors that could influence whether a woman has more children. Give an example of one factor that would not induce β_1 to be biased. Explain.
(c) Give an example of one factor that would induce β_1 to be biased. State the direction of the bias and how you determined that direction.