Problem Set 2
Global Poverty and Economic Development

Lorenzo Casaburi & Manon Delvaux

NB: This problem set is due on the 15^{th} of November by 5.00pm. Please submit your work on OLAT in the section "Problem Set 2" with your name in the title of the document and also on the first page.

1 Theory Question: Intertemporal choices

An individual has to decide how much to consume and invest. We consider a 2-period model like the one we saw in class. The invididual maximizes her expected utility: $U = u(c_1) + \delta E[u(c_2)]$. Utility is logarithmic: $u(c) = \ln(c)$.
In the first period, she has an endowment Y and she can consume, save/borrow, invest in a risky input, x_r, or a hedging input x_h: $Y = c_1 + b + x_r + x_h$. The interest on saving/borrowing b is R (gross).
In the second period there are two states of the world, H, L, with probabilities $\pi_H = \frac{1}{2}$ and $\pi_L = 1 - \pi_H = \frac{1}{2}$, respectively. In state H, only the investment in x_r pays off, giving output $A_H f(x_r)$. In state L, only the investment in x_h pays off, giving output $A_L f(x_h)$. Productivity is higher in state G: $A_H > A_L$. We assume the production function is isoelastic: $f(x) = \frac{1}{\rho}x^\rho, 0 < \rho < 1$. For simplicity, we assume the prices of all the inputs and output are equal to 1.

Throughout the exercise, we assume perfect credit markets.

Let's consider first the case with *perfect insurance markets*, too: $c_2 = Rb + \pi_H A_H f(x_r) + (1 - \pi_H) A_L f(x_h)$

1. Derive the Euler Equation.
2. Find the optimal level of investment in x_r, x_h as a function of exogenous parameters only ($i.e. \pi_H, R, \rho$). Discuss the result.
3. Describe how to find b and c_1 (you do not need to do the actual algebra)

Second, let's consider the case with incomplete insurance markets. c_2 now depends on the state of the world: $c_{2H} = Rb + A_H f(x_r)$; $c_{2L} = Rb + A_L f(x_h)$;

5. Derive the Euler Equation

6. How does the investment level for x_r, x_h compare to the case of perfect markets? Derive the results and discuss.

2 Experiments

You have been hired to help design an experiment to study the impact of a new type of insecticide-diffuser on malaria rates in a region of Tanzania. The experimenters plan to conduct their research across 100 villages, studying the effect on households.

1. How would you design an experiment to test for this impact? (*Don't need more than half a page, focus for example on type of experiment, experimental strategy, outcome measures etc.*)

2. Given the nature of the treatment, which kind of spillovers would you expect? How would this affect your experimental design?

3. A reviewer of your proposal wants to know a little bit more about your estimation strategy.

   (a) When you get to the estimation part of your experiment, at which unit would you cluster your regressions?

   (b) There is a high chance that by the end of your experiment you would have had attrition in your sample. How might you test if this is a problem?

   (c) One of the reviewers wants you to add a lot of baseline control variables to your initial mean comparison regression. How would you expect this may affect your estimates? Why?

3 Data Exercise: Difference-in-Differences (DiD)

Use the dataset *data_did.dta* available on OLAT. It contains data on 500 towns over a 10-year period.

250 of these towns receive a treatment starting in period 6. The variable $after$ equals one for observations from period 6 to 10. The variable *treatment* equals one for those towns that are targeted by the treatment. The database also includes controls x_1 and x_2.

1. Produce a 2x2 table with the mean y by type of town (treatment/control) and time period group (before/after). What is the DiD estimator in this table?

2. Run a DiD regression of y on $after$,$treat$ and $after * treat$. How does the coefficient of interest compare to the estimate from the above table? (note: all regression should be clustered by "town".)

3. Are variables x_1 and x_2 different between treatment and control towns? Does this affect the validity of the DiD approach? How do the results of the regression in point 2 change if one controls for x_1 and x_2? Why?

4. Now run a Panel DiD with town and year fixed effects. Is the DiD coefficient similar/different from the one in the two points above? Why?

5. Run a full panel DiD where you estimate treatment coefficients in every year of the sample. For convenience, you may normalize the coefficient in the last year before the program (i.e t=5) to zero. Graph all the treatment coefficients and their standard errors. Inspect visually the graph. Does the assumption of parallel trends seem to hold?

6. One option to inspect pre-trends is to run a *placebo test*. Run a regression similar to the one in point 2, but where you include only years 1 to 5 and define the placebo $after$ variable as equal to one if year=4,5. What would you expect on the DiD coefficient in this placebo regression if pre-trends were parallel? What do you find in this case? Interpret.

7. What can you conclude about the initial result in points 1 and 2?