Population and development Field Exam

Spring 2014

Instructions

You have 4 hours to complete this exam.

This is a closed book examination. No written materials are allowed. You can use a calculator.

YOU MUST OBTAIN AT LEAST 75% IN TWO OUT OF THE THREE PART TO PASS THE FIELD EXAM.

Please answer Parts I, II and III in separate booklets.

Part I: Population (Lleras-Muney)

Answer both questions. Partial credit will be given whenever possible.

1-Education and Life Expectancy (40 points)

Jayachandran and Lleras-Muney (2010) estimate the effect of of increases in life expectancy on education investments using data from Sri Lankan districts in 1946 and 1953. This study looks at how the duration of life affects the incentives to invest in schooling. It uses the large declines in maternal mortality (MMR) that took place during this period (due to medical innovation) to investigate the question.

a. (5 points) does theory predict that longer life expectancies will always increase education investments?
b. (5 points) Education and health are positively associated in many data sets. What alternative explanations are there for the association between education and life expectancy?

The data comprise 76 observations corresponding to a gender (2), district (19), and year (2). Using these data they estimate the following equation

$$e_{dgt} = \beta_0 + \beta_1 MMR_{dt} * female_g + \mu_{dg} + \gamma_{dt} + \nu_{gt} + \varepsilon_{dgt}$$

where *d* denotes district, *t* year, and *g* gender. The dependent variable is years of schooling and independent variable of interest is the maternal mortality rate (MMR). The specification includes a full set of double interactions, namely district-gender (μ_{dg}), district-year (γ_{dt}), and gender-year fixed effects (ν_{gt}). ε_{dgt} is a random disturbance term and *female* is a dummy variable equal to one.

c. (5 points) What are the identifying assumptions needed to estimate the causal effect of MMR on education?
d. (5 points) Why is this estimation strategy preferable to only using data on women? (what model would you estimate if you only had data on women?)
e. (5 points) if you had data by age groups at each time period t, how would you use it?
f. (5 points) the investigators had access to similar data for other periods. Why not include it?
g. (5 points) if you could collect additional data, what other outcomes would you want to investigate and why?
h. (5 points) the study finds that the elasticity of years of schooling with respect to life expectancy is about 1. Life expectancy grew in the US by about 30 years in the last century, start from 45 in 1900 to about 75 in 2000. What is the implied effect on years of schooling, assuming in 1900 average years of schooling were around 8?

2-The statistical value of life (40 points)

Using wage and job risk data a researcher estimates from the following equation:

$$ln(W)_{ij} = \alpha + \beta * (Annual\ MR)_j + e_{ij}$$

where $ln(W)_{ij}$ is the log of annual wages of individual i in industry j and "Annual MR" is the annual mortality rate in industry j.

a. (5 points) Under what theoretical assumptions can we interpret β as the statistical value of life?
b. (5 points) What variables would it be important to include in the regression analysis and why?
c. (5 points) Why is the specification in logs and does it matter?
d. (5 points) Should the standard errors be clustered and why?
e. (5 points) If you had data over time, what model would you estimate and why?

Other estimates of the statistical value of like can be obtained using calibration methods such as the one used in Becker et al (2005).

f. (5 points) Explain the intuition of the calibration method used in Becker et al. (2005)
g. (5 points) What are the advantages and disadvantages of calibration methods compared to the method used here?
h. (5 points) What are these estimates of the value of life used for?

Part III: The question sheets for Part III are found directly inside the exam answer booklet labeled "Part III."

Answer the questions in the spaces provided on the question sheets.

- This test has a total of 125 points and has a total of 8 questions.
- This is a closed book exam.
- **Show your work!** No credit will be given for correct answers if you do not justify your argument.
- Please make sure that your handwriting is **legible**!
- In free response questions, **be precise but brief**. If a correct reply is hidden among wrong, or irrelevant, arguments, you will not get full credit (I have given plenty of space following each question but there is no expectation that all of the assigned space needs to be filled, it is there to account for different handwriting sizes).

Shorter Questions

1. (10 points) Does the *separability result* for agricultural household models (e.g. as outlined in Benjamin (1992)) imply that consumption decisions should be independent of production decisions? Be as explicit in your answer as you can.

2. (10 points) Discuss briefly the evidence on credit being an important constraint for firms in developing countries.

3. (10 points) Does the failure of separability necessarily imply anything about the failure or completeness of specific markets (e.g. can one conclude from the failure of separability to hold that labor markets are incomplete).

4. (10 points) The results in Deaton and Paxson (1998) and Benjamin (1992) assume that household size is exogenous for their outcomes of interest. Provide an example where endogenous household size may alter their conclusions (you need only do this for one paper but you have to be specific).

5. (10 points) Discuss briefly the findings of Karlan and Zinman (2009) with respect to the existence of moral hazard in their sample. In particular distinguish between the different interventions designed to test for moral hazard.

6. (10 points) Assume that self-contained villages in a rural economy achieve pareto efficient risk sharing among all households within the village. Assume further that all agents have CARA utility functions

$$u_i(c) = \frac{1}{\sigma_i} \exp(\sigma_i c)$$

note that the CARA parameter σ_i varies across agents. In such an economy, will regressing individual consumption on village level consumption yield a consistent test of risk-sharing? Be as explicit (formal) in your answer as you can.

Longer Questions

Answer questions concisely but completely. If a correct response is hidden among wrong, or irrelevant arguments, you will not get full credit.

Question 7 .... *15 points*

A traditional view among development economists has been that the poor are not particularly different from the wealthy in their decision making ability and do the best they can given their straitened circumstances (though these outcomes may not be efficient for various reasons). This view, sometimes called the "poor but neo-classical" view contrasts with a newer view that argues that poverty directly affects the decision making ability of the poor. Which of these two view-points do you find more compelling and why? Please include in your answer a discussion of the relevant papers.

(Continued)

Question 8 .... *50 points*

Consider a population that is divided in to S strata and the relative size of each stratum is π_s with $\sum_{s=1}^{S} \pi_s = 1$. Assume that the strata weights π_s are known. Suppose, that for each stratum there is a linear model

$$Y_{is} = X_{is}'\beta_s + \epsilon_{is}$$

where β_s and X_{is} are $k \times 1$ vectors. We observe an i.i.d. sample $\{Y_{is}, X_{is}\}_{i=1}^{n_s}$ for each stratum $s = 1, \ldots, S$. In addition *across strata* observations are independent (though not necessarily identically distributed).

We assume that $E(\epsilon_{is}|X_{is}) = 0$ so that one can run stratum-by-stratum OLS regressions to estimate $\{\beta_s\}_{s=1}^{S}$. Further, assume that each stratum is large enough so that we can let $n_s \to \infty$ for each $s = 1, \ldots, S$. Further, suppose that $n = \sum_{t=1}^{S} n_s$ and that

$$\lim_{n_j \to \infty\ \forall j=1,\ldots,S} \frac{n_s}{n} = \lambda_s$$

Suppose that the object of interest is

$$\theta = \sum_{s=1}^{S} \pi_s \beta_s$$

Consider the following estimators for θ:

- $\hat{\theta}_1 = \sum_{s=1}^{S} \pi_s \hat{\beta}_s$ where $\hat{\beta}_s$ is the OLS coefficient in a regression of Y on X in stratum s (i.e. stratum-by-stratum OLS).
- $\hat{\theta}_2$ is the coefficient on X in an unweighted OLS regression of Y on X that pools all the data across strata (i.e. runs one regression after pooling all the data).
- $\hat{\theta}_3$ is the coefficient on X in an weighted OLS regression of Y on X that pools all the data across strata where each observation in stratum s gets a weight of $\sqrt{\pi_s/n_s}$.[1] In other words one estimates the regression

$$\sqrt{\pi_s/n_s}Y_{is} = \sqrt{\pi_s/n_s}X_{is}'\beta + \sqrt{\pi_s/n_s}\upsilon_{is}$$

[1]To implement this in STATA for instance you would specify the `aweight` option and set it equal to $\sqrt{\pi_s/n_s}$.

(a) (5 points) Will $\hat{\theta}_1$ be consistent for θ? Prove or disprove.

(b) (5 points) Will $\hat{\theta}_2$ be consistent for θ? Prove or disprove.

(c) (5 points) Will $\hat{\theta}_3$ be consistent for θ? Prove or disprove.

(d) (5 points) Suppose now that the data is also identically distributed across strata. How do your answers to the previous parts change?

(e) (5 points) Do you have a preference for using $\hat{\theta}_2$ or $\hat{\theta}_3$. Why?

(f) (5 points) Suppose that the stratum sample sizes n_s are chosen so that

$$\frac{n_s}{n} = \pi_s$$

so that $\lambda_s = \pi_s$ (a sampling design that satisfies this requirement is said to be "self-weighting"). Do your answers to part (b) or (c) above change?

(g) (10 points) Researchers often attempt a compromise between computing stratum-by-stratum estimators and a single OLS estimator by adding stratum indicators as right hand side variables. Consider therefore the model

$$Y_{is} = \delta X_{is} + \alpha_s + \epsilon_{is}$$

where the $\{\alpha_s\}_{s=1}^S$ are a full set of stratum indicators. Suppose for simplicity that X_{is} is a scalar binary variable ($X_{is} \in \{0, 1\}$). To make the calculations simpler, let

$$\frac{1}{n_s} \sum_{i=1}^{n_s} X_{is} \equiv p_s$$

where the equality holds for *all sample sizes* n_s.

In this part of the problem we will derive a relationship between the stratum effects $\{\beta_s\}_{s=1}^S$ and δ. Begin by showing that the OLS estimator for δ can be written as

$$\hat{\delta} = \left(\sum_{s=1}^{S} \sum_{i=1}^{n_s} \tilde{X}_{is}^2 \right)^{-1} \left(\sum_{s=1}^{S} \sum_{i=1}^{n_S} \tilde{X}_{is} Y_{is} \right)$$

where

$$\tilde{X}_{is} = X_{is} - p_s$$

(h) (5 points) Next, show that you can write

$$\sum_{s=1}^{S}\sum_{i=1}^{n_S}\tilde{X}_{is}Y_{is} = \sum_{s=1}^{S}\frac{n_s}{n}p_s(1-p_s)\beta_s - \sum_{s=1}^{S}\frac{n_s}{n}\widehat{Cov}(X_s,\epsilon_s)$$
$$\to \sum_{s=1}^{S}\lambda_s p_s(1-p_s)\beta_s$$

where

$$\widehat{Cov}(X_s,\epsilon_s) = \frac{1}{n_s}\sum_{i=1}^{n_s}X_{is}\epsilon_{is} - p_s\frac{1}{n_s}\sum_{i=1}^{n_s}\epsilon_{is}$$

Hint: Use the fact that X_{is} is binary so that $X_{is}^2 = X_{is}$

(i) (5 points) Show that

$$\sum_{s=1}^{S}\sum_{i=1}^{n_s}\tilde{X}_{is}^2 = \sum_{s=1}^{S}\frac{n_s}{n}p_s(1-p_s) \to \sum_{s=1}^{S}\lambda_s p_s(1-p_s)$$

;

Question	Points	Score
1	10	
2	10	
3	10	
4	10	
5	10	
6	10	
7	15	
8	50	
Total:	125	