

ECONOMICS 172: Issues in African Economic Development
Problem Set 1

Due Tuesday February 14, 2017 **by 2 pm** on bCourses
[Please turn in your own solutions – do not solve in teams]

1. African Economic Development and Economic Growth in Asia [4 points]

Discuss at least one reason why rapid economic growth in Asia over the past three decades might have helped promote economic growth in Sub-Saharan Africa, and at least one reason why Asian growth might have led to slower African growth. In your view, which of these effects has been more important overall for African development, and why? Your answer should be no more than 2 pages, with 12 point font and double-spacing. (Please draw on the course readings, lectures, and any additional materials you find relevant.)

2. Child Health and School Attendance [6 points]

a) Why was the health intervention in the Primary School Deworming Project (Miguel and Kremer 2004) randomized across schools? How does the project's randomized design affect the estimation of treatment effects and help address omitted variable bias? (Please illustrate your points with the discussion presented in class, including the equations that illustrate potential bias.) [2 points]

b) Download "Econ-172_PSet1-data.csv" from the bCourses page, and use the "read.csv" command to open it in R (or RStudio). Using the "lm" command, determine the average difference between 1998 treatment schools (group1=1) and comparison schools (group1=0) in the following three characteristics: Proportion of female children ("female"), Average child year of birth ("yob"), and Involvement in another school assistance program ("sap").

Then separately determine the average difference between Group 2 (group2=1) versus Group 3 (group3=1) schools in the same three characteristics. (In this analysis, restrict attention to just these two groups of schools using a logic statement.) There are several ways to use subsets of data in R with logical commands. For example, you can use the following options:

```
data=worms[worms$group1==0,]  
or  
data=subset(worms,group1==0)
```

where here we have named our rds file "worms" and we are restricting to observations in the data which have the group1 variable equal to 0.

Report the regression output for the six regressions, and interpret the coefficients. Please also discuss the standard errors and t-statistics. You should electronically turn in the actual R "history" file print-out (including the file time stamp and computer directory path, etc.) along with formatted tables using the output from "stargazer" or a similar command. Taken together, did the randomization appear to succeed in creating comparable groups at baseline? [1 point]

c) Determine the average difference between treatment and comparison schools in: Average school participation in 1998 after the program had started (“part98”).

Report the regression results – you should again electronically turn in the history file and formatted tables – and interpret the coefficients. What is the impact of attending a treatment school on average school participation? Is this significantly different than zero at 95% confidence? Then determine the average difference between Group 2 and Group 3 schools in “part98” (again restricting analysis to just these two groups). Would you expect to find differences between Groups 2 and 3, and do you find any? [1 point]

d) Re-run the two regressions in part (c), but include “female”, “yob” and “sap” as additional explanatory variables, again using the “lm” command in R. Report and interpret the regression results. Does the inclusion of these control variables (also called covariates) change the conclusions in (c)? Is this surprising? [1 point]

e) In R, graphically plot the relationship between “female” (horizontal axis) vs. “part98” (vertical axis), using the “plot” command. Include the best linear regression fit for this relationship (using the “abline” command). Include this graph in your solution, and speculate on what factors might help explain the observed pattern. Then create an analogous plot but restricting attention to Group 1 schools alone. Are there any noticeable differences in the patterns? What might account for any differences? [1 point]