14.74 Problem Set # 1

Due in Box by E52-201 **BEFORE** Lecture on Monday, February 21, 2005

No problem sets will be accepted at lecture, No late problem sets will be accepted

Feel free to work in groups. But please hand in individual problem sets.

Please give clear concise answers (the least amount of words it takes to make your meaning clear). The point of this problem set is to think thoroughly about the models and issues discussed in lectures and recitation and to relate your own insights to the intuition given in class.

1. World Bank Development Report Data

The point of this exercise is to help you understand the faces of poverty on the one hand, and acquaint you with Stata on the other. You will need to download the Stata data set (wbdr.dta; this data set contains 1997 variables, unless otherwise noted) from the course website. There, you will also find instructions for using Stata. *Hand in your do file with the problem set. No need to hand in the log file, although you can hand in the parts that are directly relevant for answering questions below.*

a. What are the mean and standard deviation of GNP per capita, illiteracy rate, infant mortality rate and under 5 mortality rate in 1997 across countries?

b. Restrict your data set to countries for which we have GNP per capita for 1997. What are the mean, minimum, and maximum illiteracy rate, infant mortality rate (for 1997), and under 5 mortality rate among the 50 richest countries? Among the 50 poorest countries? What does this tell you about the relationship between income and illiteracy, and income and mortality

c. Now, find the median GNP per capita. How is different from the mean GNP per capita from part (a) Why might the mean and median be different? What information can we infer about the income distribution (inequality) from this?

Return to using the original (unrestricted) data set.

i. Regress the illiteracy rate on per capita GNP in 1997. Report the coefficient on per capita GNP and its standard error; what do they tell you? Is the sign of the coefficient what you expected? Explain briefly. What is the t-statistic for this coefficient, and what does it tell you? Interpret the 95% confidence interval.

ii. Regress the infant mortality rate in 1997 on GNP per capita in 1997. Is the coefficient on per capita GNP significantly different from zero? How do you know? Interpret the coefficient in terms of a $1000 difference in per capita GNP.

iii. Regress the infant mortality rate in 1997 on the illiteracy rate. Graph a scatter plot of the data as well as the regression line. *(Use the "graphics" option from the Stata toolbar. Stata will output the code as well as the graph. You should then try typing the code in directly instead of the editor. It is much faster.)*

iv. Using the results from part i-iii, can what can we say about the causal relationship between illiteracy, infant mortality, and income (GNP)? In other

words, tell a story of how illiteracy, income, and infant mortality might affect one another. Be concise but clear about what causes what and why.

2. Poverty, Nutrition and Labor Markets

This exercise is based on the lecture and the textbook pp. 272-79, 489-504. Each answer should not be longer than a couple of sentences.

a. In what circumstances will unequal allocation amongst a 2-member household increase total household work capacity? Discuss the importance of the convexity of the relationship between work capacity and income. Please illustrate this in a simple graph.
b. For a workers with no non-labor income (e.g. land) and are getting piece rates, draw the aggregate labor supply curve and a standard downward sloping labor demand curve with piece rates on the y-axis. When will unemployment occur? In equilibrium, why can the piece rate not be bid down so that people who are involuntarily unemployed can obtain work?
c. How does having non-labor income affect the work capacity or willingness to work? (Hint: think of the difference tradeoffs between working and not working for the two different types of workers.)
d. Discuss in words and graphically how voluntary and involuntary unemployment can be generated in an economy where some workers have non-labor income and other workers do not.
e. What policy can the government enact to address involuntary unemployment? Briefly discuss the tradeoffs of your policy suggestion.
f. Recall from class that $\boldsymbol{n_t=g(workcapacity_t)=g(f(n_t-1))}$, algebraically derive the elasticity of income with respect to nutrition. Explain the meaning of this term with words.

Imagine that we have data from supermarket scanners from Star Market. The scanners trace people's food expenditure and we also know some basic income and demographic characteristics (e.g. age, education, etc.) from when each person registers for their card. We want to see if richer people eat more nutritiously.

g. Write the regression that you would run to answer the question above.
h. Can the regression estimate tell us whether more income causes people to eat more nutritiously? Why or why not?
i. Do you predict the coefficient d to be positive or negative if you regress income on nutrition?
   $\mathbf{Income_i=a + d(nutrition\ intake_i) +e}$
j. Do you predict the coefficients b and d to be positive or negative if you regress income or nutrition on education?
   $\mathbf{Income_i=a + b(education_i) +e}$
   $\mathbf{Nutrition_i=c +d(education_i) +m}$
k. How might the results from part (i) and (j) affect the interpretation of the regression you ran in (g)?

l. Give two separate examples of measurement error that would cause us to overestimate and underestimate the effect of income on nutrition intake. Explain.