14.03/14.003 Problem Set 3
Fall 2014

Professor David Autor

Due Thursday, 10/23/2014 at 5pm EST

1 Partial Equilibrium and Consumer Theory (30 points)

1. Consider the labor market for high skill workers characterized by the following supply and demand:
   $S(w) = -50 + 5w$
   $D(w) = 1000 - 10w$

   (a) Calculate the equilibrium wage in this market. Draw the supply and demand curves on a graph, and calculate worker and employer surplus. **[4 points]**

   (b) There is another group of workers who may be high skilled, but have not been tested. The government institutes a free testing service which verifies workers as being skilled. The result is that a second, identical group of high skilled workers join the workforce, doubling the supply of high skilled labor. What is the new supply equation that the employer faces? Draw the new supply curve on your previous graph (and verify that at each level of wage the supply of labor doubled). What is the new equilibrium wage? **[4 points]**

   (c) At the new equilibrium, calculate the employer's surplus, the surplus for the original group of high skilled workers and the surplus for the new group of high skilled workers. How would each group be willing to individually pay to implement this government testing service? **[4 points]**

   (d) The original group of workers rightfully views this new development as a wage decrease for themselves. Associated with this wage decrease is a transfer of wealth to the employer (since each working employee is now getting paid less) as well as a loss of welfare from workers who were productive but no longer have a job. On your graph label the areas associated with the transfer and job loss. What is the sum of the two areas? **[3 points]**

   (e) If the program costs K to implement, for what values of K does this service increase total welfare? If $K = 0$, what are the deadweight losses associated with this policy change? **[3 points]**

2. Return to our original market for labor
   $S(w) = -50 + 5w$
   $D(w) = 1000 - 10w$
   Suppose that the government implements an earned income tax credit. Any worker in the market that is employed gets a government subsidy of 4, so our new labor supply is $S(w) = -50 + 5(w + 4)$.

   (a) For subsidies, we are often interested in the incidence, which describes how much each party gains from the subsidy. To calculate this, find the the new equilibrium wage. What is the difference in the amount the worker gets paid? What is the difference in the amount the employer pays? The ratio of the difference for the worker and the employer is the incidence. This also the ratio of the elasticities of supply and demand. What is this value? **[4 points]**

(b) Is there deadweight loss associated with this program? On a new graph draw the shift in supply and demand and label the deadweight loss associated with the subsidy. Calculate the change in total welfare and the cost of the subsidy to determine the deadweight loss. **[4 points]**

(c) Imagine instead that the labor supply curve was $S(w) = -400 + 10w$. Recalculate the incidence as in part (a) and compare to your answer above. Explain intuitively why the incidence of the subsidy on equilibrium wages is higher or lower in this case. **[4 points]**

2 General Equilibrium in an Exchange Economy (30 points)

Romeo is an MIT undergraduate who lives in East Campus. Juliet is a Wellesley student. Suppose Romeo has utility

$$u_r(R, C) = R \ C^{1-}$$

where R are roses and C are chocolates. Juliet has utility

$$u_j(R, C) = R \ C^{1-}$$

with $\in (0,1)$ and $\in (0,1)$. Initially Romeo and Juliet do not know each other, and so are unable to trade the roses and chocolates they own, only consuming their own endowment. Romeo begins with 1 rose and 9 chocolates ($\bar{R}_r = 1$, $\bar{C}_r = 9$), Juliet begins with 1 chocolate and 9 roses ($\bar{C}_j = 1$, $\bar{R}_j = 9$).

1. What are the initial utility levels of Romeo and Juliet? **[2 points]**

2. Suppose that Romeo bumps into Juliet at Toscanini's and after talking about life over ice cream, they decide to enter into a trading agreement. Draw an Edgeworth box and show clearly all feasible allocations of chocolates and roses for Romeo and Juliet. Draw Romeo and Juliet's indifference curves within the box [you can assume $=$ $= \frac{1}{2}$ for this part if you would like, but the indifference curves need not be completely accurate] and mark the endowment allocation. Shade in the allocations which would be Pareto improvements over the endowment allocation. **[5 points]**

3. Suppose that Romeo and Juliet sell/buy roses at price p_R and sell/buy chocolates at price p_C. We can normalise and set $p_R = 1$. What relative price $\frac{p_C}{p_R} = p_C$ clears these mini-markets for roses and chocolates? [Note that chocolates and roses are divisible for the purposes of this question, that is, you can buy a fraction of a rose or take a small bite of a chocolate.] **[6 points]**

4. Assume now that $=$. Suppose that Romeo and Juliet trade at the relative price p_C that you just found. Can you show that the utility of each person is higher than it was in part (1) (i.e. that we have a Pareto improvement, a.k.a. gains from trade)? **[7 points]**

5. Still assuming that $=$, what happens to the relative price of chocolates p_C and equilibrium consumption levels as increases (i.e. as preferences of both shift toward roses)? Can you give any intuition for these effects? **[5 points]**

6. Now suppose we start with a scenario where $\neq$ and Romeo has 9 roses and 9 chocolates ($\bar{R}_r = 9$, $\bar{C}_r = 9$), Juliet has 1 rose and 1 chocolate ($\bar{R}_j = 1$, $\bar{C}_j = 1$). The endowments are very unequal. Are there still gains from trade? What does this say about the possibility of true love between the rich and poor? **[5 points]**

3 Subsidizing or Cost sharing (40 points)

In this exercise, you will replicate some of the results from the paper "Free distribution or cost sharing: Evidence from a randomized Malaria prevention experiment" by Pascaline Dupas and Jessica Cohen. You might find it helpful to read the paper to familiarize yourself with the context.

For this part of the problem set, you will need to submit the stata code (the do-file) you used to generate the results. Please also include the relevant portion of the stata results in the content of your problem set solutions. (We don't want to read the do files line by line, but we do want to see them.)

In Kenya, Insecticide Treated Bednets (ITNs) are sold through clinics where pregnant women come for their routine prenatal check ups and vaccinations. Bednets are sold in the market at 50 Ksh per piece. The authors designed an experiment to find out how the demand of bednets change when price changes. The authors have 20 clinics which they randomly assigned into one of five groups:

- Group 1; 4 clinics: control group
- Group 2; 5 clinics: bednets to be provided free for anyone visiting the clinic
- Group 3; 5 clinics: bednets to be sold at 10 Ksh
- Group 4; 3 clinics: bednets to be sold at 20 Ksh
- Group 5; 3 clinics: bednets to be sold at 40 Ksh

For this problem, you will use four stata datasets: **baseline.dta**, **buypatient.dta**, **sales.dta**, and **salesMarket.dta**. baseline.dta contains clinic level data collected during baseline. buypatient.dta contains patient (all women visiting the clinic) level information collected after the program is implemented. sales.dta contains weekly sales data for the 6 weeks after the program is implemented. **salesMarket.dta contains data on bed net prices and bed net sales at 100 clinics across Kenya collected *prior to* the experiment (this is non-experimental, AKA observational, data).**

The relevant variables in each dataset are described below:

Dataset	Variables	Description
baseline.dta		
	`clinicid`	clinic ID
	`district`	District in which clinic is located
	`treatment`	=1 if clinic received any kind of subsidy, 0 otherwise
	`netprice`	Price at which bednets are sold (randomized), for the control group this variable is Ksh 50
	`group`	Group assignment (described as above)
	`anc_fee`	Fee charged by clinic for ANC services
buypatient.dta		
	`respid`	ID of patients in the clinic
	`clinicid`	clinic ID
	`district`	District in which clinic is located
	`treatment`	=1 if clinic received bednet distribution program
	`netprice`	Price at which bednets are sold (randomized)
	`group`	Group assignment
	`buy`	=1 if Purchased net during prenatal care visit
sales.dta		
	`clinicid`	Clinic ID
	`week`	Week since net distribution program start
	`weeklynetsales`	Total number of nets sold in the week
	`district`	District in which clinic is located
	`treatment`	=1 if clinic received bednet distribution program
	`netprice`	Price at which bednets are sold (randomized)
	`group`	Group assignment
salesMarket.dta		
	`clinicid`	clinicid
	`p`	Bed net price on day of observation
	`salesMarket`	Number of bed nets sold on day of observation

1. **[10 points]** In Table 1, Cohen and Dupas show that baseline characteristics are balanced even in their small sample of clinics. You will replicate results from row 3. For this part, load **baseline.dta** into stata

   (a) **[2 points]** Find the average fees charged by clinics for ANC (Ante-natal care) services for each of the five groups of clinics (the group information is given by variable `group`). Find the standard deviations as well.

   (b) **[1 point]** Test whether fees charged by clinics the control group (Group 1) is different from fees charged by clinics that receive full subsidy (Group 2). What is the **t-statistic**? Is the difference statistically significant at the 95% significance level?
   (Hint: you can use the `ttest` command in stata, but be sure to *restrict your data only to the two groups that you are comparing*)

   (c) **[3 points]** Repeat the previous exercise for the following groups: Group 1 vs Group 3; Group 1 vs Group 4; Group 1 vs Group 5. Are any of the differences statistically significant at the 95% significance level

   (d) **[3 points]** A more succinct way to do this test is using a regression. Generate indicator variables for each of the groups (for example, `GR1` should be 1 if the clinic is in gorup 1 and 0 otherwise). Regress fees variable against the four indicator variables (omit `GR1`, the control group). Are any of the coefficients significantly different from zero at 95% significance?

   (e) **[1 points]** Now do a joint-test that all the coefficients are zero. Do we reject the null that all the coefficients are jointly zero? Compare your p-value (`Prob > F`) with the corresponding value reported in Table 1.

(Hint: use `test` command in stata after the `regress` command)

2. **[14 points]** Now, lets compute the effect of the subsidy on whether a patient that visits the clinic buys a bednet or not. The researchers arrived at clinics on random, non-announced days and asked all the visitors present in that day about whether they purchased bednets or not. For this part, load **buypatient.dta** into Stata.

   (a) **[4 points]** Find the average probability that a patient in Group 2 (free distribution) buys ITNs. Find the same for groups 3, 4 and 5. Plot these probabilities in a graph. Compare your graph with Figure 1 in the paper.
   [Hint: In stata you can use the `collapse (mean)` command to collapse your data to an aggregated level that you specify (in the `by(varname)` option) and plot the data using `twoway scatter` command. Or you can simply get the averages into excel and plot it in excel]

   (b) **[3 points]** This data is at the patient level whereas the experiment was randomized at the clinic level. So, it makes sense to aggregate things to the clinic level. Use `collapse` command to collapse individual level purchase data to clinic level averages (your collapsed dataset should have 20 observations, 1 for each clinic). Now test whether purchase rate in Group 2 (free distribution) clinics were different from purchase rates in Group 3 (10 Ksh). Give the t-statistic and indicate whether the difference is significant at 95% significance level.

   (c) **[1 point]** Now, compare purchase rates for Group 2 vs Group 4 and for Group 2 vs Group 5. Can you reject the null that free distribution has no effect on purchase rate compared to clinic that charge higher prices.

   (d) **[3 points]** As in part d in the previous question, use a regression to perform all three comparison in one statistical test. (exclude group 2 in this case). Do your conclusions differ from part b and c? Explain.

   (e) **[3 points]** Instead of averages at the clinic level where you worked with the share of patients that bought the bednets, one could do this exercise with the number of patients that bought the bednets. Go back to using the original **buypatients.dta** but instead of collapsing using means, collapse the sum at the clinic level. That is, use `collapse (sum)` instead of just `collapse`. You will still have 20 observations (one per each clinic), but the variable will now be a count of individuals who bought the bednet rather than share. In this dataset, regress number of patients who bought the bednets on the price which clinics sell ITNs. Report your coefficients as well as standard errors for the slope and the constant terms.

3. **[10 points]** The researchers also collect weekly sales data from the log-books of the clinics. They do so to provide another measure of the demand of bednets. Lets now look at this data and see if we find the same (the drastic reduction in demand when price is increased). For this part, load **sales.dta** into stata.

   (a) **[2 points]** The log-book data records all the sales done by the clinic. In an ideal setting, this measures the sales data perfectly. Why do you think the researchers went on to the trouble of going to the clinics unannounced and collecting the data from those who were present?

   (b) **[2 points]** Regress weekly sales on price at which clinics sell ITNs. Compare your coefficient with Table 3, column 1.

   (c) **[2 points]** Interpret the slope coefficient. What does it tell you? Interpret the intercept.

   (d) **[1 point]** Based on your coefficients, how many bednets would a clinic expect to sell if the price of the bednet is sold at its prevailing price of 50 Ksh?

   (e) **[3 points]** Compare the estimate in 3b with 2e. You would not expect the coefficients to be similar because your dependent variable is different in each case. In part 2e, it measured total number of patients who bought the nets in the few random days the researchers went and collected data. Whereas, in part 3b, it measured the total number of bednets sold by the clinics in a week. To compare the coefficients, you can perform this simple exercise for each of the estimates:

- Assume that $p = 0$, what is the sales?
- Now suppose the price increased to $p = 1$. What is the sales?
- What is the proportional increase in sales/demand?

Perform this exercise for both estimates, and explain whether you think the two measures of bednet demand are giving you the same result.

4. **[6 points]** Now, suppose another researcher had collected data from 100 clinics in the same region before the experiment was conducted. This researcher collected data on weekly sales of ITNs and the price at which the clinic sold the bednets. The data is present in **salesMarket.dta**. (Note that this dataset has nothing to do with the experiment described above)

   (a) **[2 points]** Regress the sale of bednets on the price of bednets. Is the slope coefficient significantly different from zero?

   (b) **[1 point]** Interpret the slope coefficient.

   (c) **[3 points]** Propose an economic hypothesis on why your the estimate of the slope from 4a differs from that of 2e and 3b.

14.03/14.003 Microeconomic Theory and Public Policy
Fall 2025