Problem Set 1 (DUE FRIDAY APRIL 22ND)

Instructions: All answers must be typed (Latex preferred but anything is fine). When appropriate, answer the question by creating concise and well-labeled tables or graphs. Be sure to hand-in your computer code with sufficient annotation. You can use any programming language/software you like.

Question 1: Roy Model

Consider the following Roy model for self-selection of an occupation.

There are two occupations: $k = 1$ and $k = 2$. Each agent i is endowed with two skills S_{i1}, S_{i2}, with $\ln S_{ik} = \mu_k + \epsilon_{ik}$. The joint distribution of $\epsilon = (\epsilon_{i1}, \epsilon_{i2})$ is

$$\epsilon \sim N(0_2, \Sigma),$$

and

$$\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix}$$

It is often more convenient to parameterize the model in terms of the correlation coefficient:

$$\rho = \frac{\sigma_{12}}{\sigma_1 \sigma_2},$$

where $\sigma_k = \sigma_{kk}^{1/2}$

Wages from each occupation k are given by

$$W_{ik} = \pi_k S_{ik}$$

There is a single market, with occupation specific prices π_0, π_1.

Each agent chooses the occupation with the highest wage. Define $D_i = 1$ if occupation $k = 1$ is chosen by agent i ($W_{i1} \geq W_{i2}$). $D_i = 0$ if occupation $k = 2$ is chosen ($W_{i1} < W_{i2}$). The data consist of N realizations of $(D_i, D_i W_{i1} + (1 - D_i) W_{i2})$ (choice and observed wage)

a) Explain why a normalization of $\pi_1 = \pi_2 = 1$ would be without any loss of generality in this model.

b) Write a computer program to compute this model for a given set of parameters and N simulated individuals. The parameters in the model are: $\theta = (\pi_1, \pi_2, \mu_1, \mu_2, \sigma_1, \sigma_2, \rho)$.

c) Find one vector of parameters such that (approximately) 60 percent of your simulated observations choose occupation 1.

d) Write (in math, words) a consistent estimator for the parameters θ. Briefly explain why your estimator is consistent.

e) Using the parameters in c) as the "true" population parameters, write a program to compute your estimator in d) for only the μ_1, ρ parameters, keeping the other parameters fixed at their true values.

f) Create some identification figures which show that your estimator identifies the true parameter values for μ_1, ρ.

g) Report your parameter estimates from e) in a table. Show sample fit at your estimated parameters: for both the data and at your estimated parameters, show fraction choosing occupation 1, average observed wages in both occupations, standard deviation of observed wages in both occupations (5 statistics for data, and 5 statistics for estimated model = 10 total)

h) Using your parameter estimates, find a counterfactual minimum wage policy for occupation 1 $\underline{W}_1$ such that under this policy 70 percent of the population choose occupation 1. The counterfactual policy is that offered wages in occupation 1 are now

$$W_{i1} = \max\{\underline{W}_1, \pi_1 S_{i1}\}$$

For this counterfactual minimum wage policy show how the average and standard deviation of accepted/observed wages differ relative to the baseline with no minimum wage policy. Show your results in a table.

THE REST OF THE PROBLEM SET IS OPTIONAL

i) (OPTIONAL) Now write an estimator in which you **incorrectly** assume the population is composed of two types A, B, with joint skill S_{i1}, S_{i2} support given by $(S_{A1}, S_{A2}), (S_{B1}, S_{B2})$ and probability mass of type A given by $P_A \in (0, 1)$.

j) (OPTIONAL) Assuming the Normal Roy model is the true model and the true parameters are given in c), compute your estimator for the parameters of this mis-specified model $(S_{A1}, S_{A2}, S_{B1}, S_{B2}, P_A)$. In a table, show your parameter estimates. In addition, provide some identification graphs to show that the observed data identifies this mis-specified model.

k) (OPTIONAL) Compare the sample fit of the two estimators based on the correctly specified and incorrectly specified models. For sample fit, show average fraction choosing occupation 1, average observed wages in both occupations, standard deviation of observed wages in both occupations (10 total statistics).

l) (OPTIONAL) Re-compute the minimum wage policy but using your estimated mis-specified model. Use the $\underline{W}_1$ value you found above. In a table, compare the estimated effects under the two models.

m) (OPTIONAL) Now assume you have many markets with different prices π_1, π_2. Continuing to assume the Normal Roy model is the true model and the distribution of S_{i1}, S_{i2} is the same in each market, write a computer program to simulate your model for M markets.

n) (OPTIONAL) Without assuming skills are Normal, write an estimator for the joint distribution of S_{i1}, S_{i2}. Assume you observe the prices in each market. You can parametrize this joint distribution however you like, but try to assume something that is appreciably more flexible than the Normal distribution.

o) (OPTIONAL) Compute your flexible (non-parametric?) estimator in n) for some number of markets $M > 1$ (your choice), report parameter estimates and sample fit.

p) (OPTIONAL) Re-compute the minimum wage counterfactual policy but using your non-parametric estimate of the unobserved skill distribution.