Problem Set 1

Fulin Guo

1.

(a) I find a statistical model from the paper "Intergenerational Mobility and Preferences for Redistribution" from the American Economic Review (Alesina, 2018).

(b) Citation: Alberto Alesina, Stefanie Stantcheva, Edoardo Teso. 2018. "Intergenerational Mobility and Preferences for Redistribution." *American Economic Review*, 108(2): 521-554

(c) The model:

$$policy\ preference_i=\beta_0+\beta 1*perception_i+\beta_2\ left_i+\beta_3*(perception_i*left_i)+\varepsilon_i$$ (Alesina, 2018).

The variable *policy preference* is people's preference for redistribution policies (to what extent people support redistribution policies). The variable *perception* is people's perception of intergenerational mobility (how high he or she thinks the intergenerational mobility is). The variable *left* is one's political spectrum; it equals to 1 if the people is left leaning and equals to 0 if he/she is right leaning. *Perception*left* is an interaction term, which is used to test the difference of the causal effect of *perception* on *policy preference* between left-wing and right-wing people (Alesina, 2018).

(d) The variable *perception* and *left* are exogenous. The variable *policy preference* is endogenous.

(e) The model is static since time dimension is not included in the model. It is linear as the equation in (c) is a linear regression equation. The model is stochastic since there is a random variable ε in the equation in (c). That is, an individual's *policy preference* does not only depend on his/her political spectrum and perception of mobility, but also depends on other personal characteristics, like one's education level, ability, etc. Therefore, even though we know one's perception and political spectrum, we cannot accurately predict his/her policy preference (although we can predict the average of the result given the values of the parameters and the independent variables).

(f) I think individual's income is an important variable that the model might miss because wealthy people might not like redistribution policy since they can earn a lot of money by themselves and this policy might actually make them worse. On the other hand, poor people might tend to regard the redistribution policy is necessary because it will improve their life. Therefore, I think people's preference for redistribution policy depends on his/her income.

2. My model is:

$$logit(get\ married)=ln\frac{get\ married}{1-getmarried}=\beta_0+\beta_1*age_i+\beta_2\ income_i+\beta_3*\text{girl\_boyfriend}_i+\varepsilon_i$$

The dependent endogenous variable is *get married*, which equals to 1 if an individual decides to get married and equals to 0 if he/she does not. The variable *age* and *income* are individual's age and income respectively at the time we predict whether he/she decides to get married. The variable *girl_boyfriend* is whether the individual has a boyfriend or girlfriend; it equals to 1 if he/she has and equals to 0 if he/she does not have. ε is a random variable, which represents other unobservable factors that influence the dependent variable *get married*.

The above model is a logit model, the output of the model (*get married*) represents the probability that the people decide to get married. If we know all the parameters, given the values of the independent variables, we could simulate data to see the distribution of the dependent variable *get married*.

I think people's age, income, whether he/she has a boyfriend or girlfriend are key factors that determine the outcome. The older an individual is, the more likely he/she is to decide to get married. People under 15 years old are unlikely to get married while those about 25 years old might think getting married is an important thing that he/she should take time to consider right now. Income is another important factor. If a man is too poor, he might not have ability to support his wife, but for a wealthy person, money will not be a barrier to marriage. The third factor is whether an individual has a boyfriend/girlfriend. If the answer is yes, he/she might have the desire and ability to get married immediately. If not, it might indicate that he/she thinks living alone is better. Also, it would be more difficult for him/her to get married since he/she needs to find a boyfriend/girlfriend first.

Other factors might not be included in the model, like education. Although it seems that few people who are receiving education would decide to get married, adding the variable *education* to the model is not necessary because the reason why students do not get married is just they are too young and they do not have income. These factors have been included in the model as the variable *age* and *income.* Holding the income and age fixed, we might not find a significant relationship between *education* and *get married* (We might include it in the model first and conduct a statistical inference to decide whether we should omit it).

In order to conduct a preliminary test, we could first collect data about people's marital status, age, income, and whether he/she has a boy/girlfriend. Then, we conduct the logit regression to get the estimation of the parameters β_0, β_1, β_2, and β_3. If the estimated value of β_1 is significantly larger than 0, we could reject the hypothesis that age does not impact marital status, then we might think that the factor age is significant in real life. Otherwise, if we could not reject the hypothesis that age does not influence marital status (we could not reject the hypothesis that $\beta_1=0$), we might regard the factor age is not significant. In a similar way, we could test whether variable *income* and *girl_boyfriend* are significant.

Reference:

Alberto Alesina, Stefanie Stantcheva, Edoardo Teso (2018). Intergenerational Mobility and Preferences for Redistribution. *American Economic Review*, 108(2): 521-554