Problem Set #1

MACS 30150, Dr. Evans
Adriana Rakshana

Problem 1 Classify a model from a journal.

Part (a). The paper I chose is *"How Does Household Income Affect Child Personality Traits and Behaviors?"* in the *American Economic Review* from 2018.

Part (b). E The detailed citation of the article:

Akee, Randall, et al. "How does household income affect child personality traits and behaviors?" *American Economic Review*, 108.3 (2018): 775-827.

Part (c). The model outlined in the paper is:

$$\begin{aligned} Y_{it} = \alpha_0 + \beta_1 YoungestCohorts_i + \beta_2 After_t + \beta_3 AmericanIndian_i + \\ \delta_1 YoungestCohorts_i \times After_t + \delta_2 YoungestCohorts_i \times AmericanIndian_i + \\ \lambda YoungestCohorts_i \times After_t \times AmericanIndian_i + \\ X'\mu + \epsilon_{it} \end{aligned}$$

(ϵ_{it} is the error term.)

Part (d). The endogenous variable is Y_{it} which represents the child i's outcomes, including "psychopathologies (emotional/behavior disorders) and personality traits (conscientiousness, agreeableness and neuroticism) at year t" (Akee et al., 2018, p.777).

The exogenous variables in this model are as follows:

$YoungestCohorts_i$: a dummy indicator for whether the child belongs to the youngest or the second youngest;

$After_t$: a dummy indicator for whether the year is after the start of the casino transfer payment;

$AmericanIndian_i$: a dummy indicator for whether the child is American Indian race;

X': a vector which includes "a control for child age, the interaction between age and American Indian race, calendar-month-specific dummies and a count variable for the number of children younger than six in the household" (Akee et al., 2018, p.789)

Part (e). This model is a static, nonlinear and stochastic model. First, there is time component that influences the performance of this model, so it is dynamic; Second, the model is nonlinear, it includes two double-interaction terms and a triple-intersection term; Third, it is stochastic because an error term ϵ_{it} is included in this model.

missing variable that might be valuable here is the initial household income. According the data in the pre-intervention period, we can find that there is a strong

positive relationship between initial household income and initial personality skill endowment across both racial groups. A lower household income may cause a lower investment in child skills. So, it is possible that children from households with lower income (a lower investment in child skills) would exhibit greater human capital gains from the unconditional household income transfer if the marginal effect is decreasing.

Reference:

Akee, Randall, et al. "How does household income affect child personality traits and behaviors?" *American Economic Review*, 108.3 (2018): 775-827. /

Problem 2 Make your own model.

Part (a) - Part (c) The model is a simple logistic model.

$$Marriage = \begin{cases} 0, & if\ P_i < 0.5 \\ 1, & if\ P_i \geq 0.5 \end{cases}$$

where

$$P_i = \frac{1}{1 + e^{-z_i}}$$

where

$$z_i = \beta_0 + \beta_1\ age_i + \beta_2\ restriction_i + \beta_3\ education + \beta_4\ income$$

Marriage: a dummy variable of whether the individual decides to get married, it equals to 1 when the individual decides to get married, otherwise, it equals to 0;

age: the age of the individual;

education: a categorical variable to indicate the education level of the individual;

income: the Logarithmic form of the individual's income;

restriction: the restriction variable is an overall indicator of one's attitude towards marriage. It could be calculated based on personal answers to the questions in a survey.

Part (b). The endogenous variable is *Marriage*, which represents whether the individual decides to get married(1 = get married and 0 = not get married).

Part (c). Based on this model, we could use the collected data to estimate the value of β. Then we could simulate data from this model given all the parameters and relationships. So the model is a complete data generating process.

Part (d) The key factors in the model are restriction, education, income and age. The **restriction** variable quantifies personal attitude towards marriage, which is indispensable in the model. The **income** factor is also important because financial stability is quite an important part of a stable marriage. **Age** is very important for this specific modeling question of when to get married as after a certain age people in a couple are more likely to want to get married. Furthermore, many countries have legal provisions on the age of marriage and people cannot get married before this legal age. Besides, for a particular age stage, people are more likely to get married as they get older. As for **education** level, there is no doubt that education level has important influence on marriage. First, people with different education levels have different requirements and expectations for marriage. Second, it takes time to get a higher education, which may delay the marriage plan of these people.

Part (e) The other potential variables to be included in this model could be parents' wealth. The reason that I do not include wealth variables is because the data of asset value could easily be inauthentic. Wealthy people might hide their wealth and the average person might boast.

Part (f) Firstly, we need to collect a data-set containing over 100 individuals' data. Each individual's data includes whether he or she is married, his or her age, overall restriction level to marriage, income and education level. Then, run the logistic regression of the marriage variable on the other three independent variables. We can see from the p-value whether the three variables are significant or not. We can also use scatterplots matrix to see the relationship.