Problem Set #[1]

MACS 30000, Dr. Evans
Qianqian Xiong 12161987

Problem 1
(a). "Are There Environmental Benefits from Driving Electric Vehicles? The Importance of Local Factors" from *American Economic Review*.

(b).
Holland, S.P., E.T. Mansur, N.Z. Muller and A.J. Yates (2016), "Are There Environmental Benefits from Driving Electric Vehicles? The Importance of Local Factors," *American Economic Review*, 106 (12): 3700-3729.

(c).
This is a theoretical discrete choice transportation model in which consumers in the market for a new vehicle choose between a gasoline vehicle and an electric vehicle. Consumers obtain utility from a composite consumption good x (with price normalized to one) and from miles driven over the life of the selected vehicle, either gasoline miles g or electric miles e. There are several policy variables. The government may provide a subsidy s for the purchase of an electric vehicle, place a tax t_g on gasoline miles, a tax t_e on electric miles, or some combination of these policies. We hold fuel and vehicle prices fixed.

The indirect utility of purchasing a gasoline vehicle is

$$V_g = \max_{x,g} x + f(g) \text{ such that } x + (p_g + t_g)g = I - p_\psi$$

where p_ψ is the price of the gasoline vehicle, p_g is the price of a gasoline mile, I is income, and f is a concave function.

Likewise, the indirect utility of purchasing an electric vehicle is

$$V_e = \max_{x,e} x + h(e) \text{ such that } x + (p_e + t_e)e = I - (p_\Omega - s)$$

where p_Ω is the price of the electric vehicle, p_e is the price of an electric mile, and h is a concave function. Any difference in attributes between the gasoline and electric vehicle are captured by differences in the functions f and g.

We assume that the choice of vehicle is influenced by i.i.d. random variables ϵ_g and ϵ_e drawn from a common extreme value distribution with zero expected value and

standard deviation that is proportional to a parameter μ.

Accordingly, we define

$U_g=V_g+\epsilon_g$ and $U_e=V_e+\epsilon_e$

A consumer selects the gasoline vehicle if $U_g > U_e$.

(d).

Exogenous: x, p_g, t_g, g, I, p_ψ, p_e, t_e, e, s, p_Ω, ϵ_g, ϵ_e

Endogenous: V_g, V_e, U_g, U_e

(e).
This model is a static model because this model has nothing to do with time and describe the system at a specific time.
This model is a non-linear model because the function g and h are non-linear.
This model is a stochastic model because it includes random variables ϵ_g and ϵ_e.

(f).
I think the model is missing a variable of the performance and comfort level of the vehicle which might be valuable to the utility of the consumer because vehicles with different performance and comfort level can provide the consumers with different level of utility.

Problem 2
(a).
We need to include dummy variables because the outcome is binary dependent variable. The outcome is denoted as Y (if someone decides to get married Y=1; otherwise Y=0). And the exogenous variables are also dummy variables. D_1: whether the person has a marriage partner (If yes, D_1=1; if no, D_1=0); D_2: whether the person wants to get married (If yes, D_2=1; if no, D_2=0); D_3: whether the marriage partner has the willing to get married with the person (If yes, D_3=1; if no, D_3=0); D_4: whether the person and his/her marriage partner reach the legal age of marriage (If yes, D_4=1; if no, D_4=0). Thus, all of the exogenous variables are dummy variables.

Thus, the model should be: $D = \beta_0 + \beta_1 D_1 + \beta_2 D_2 + \beta_3 D_3 + \beta_4 D_4 + \varepsilon$ and $\beta_0 + \beta_1 + \beta_2 + \beta_3 + \beta_4 = 1$ which means that one will get married (Y=1) when all of the criteria are meet.

Y=1, if D=1; Y=0, if D≠1.

We additional include "Logit Model".

$\Pr(Y = 1|D) = G(\beta_0 + \beta_1 D_1 + \beta_2 D_2 + \beta_3 D_3)$ and $G(w) = \Lambda(w) = \frac{\exp(w)}{1+\exp(w)}$

Thus, we can also calculate the probability that someone decides to get married.

(b).
Y is the dependent endogenous variable that satisfies the requirement.

(c).
The model is a complete data generating process because we could simulate data from the model given all the parameters and relationships.

(d).
The key factors that influence the outcome are D_2 and D_4.

(e).
Because these factors are determined by the person himself/herself. The other factors can be meet by other methods. Thus, D_2 and D_4 are important to the outcomes.

(f).
I can regress the “Logit Model” for the preliminary test and to see whether the factors are significant in real life instead of using the original model for the preliminary test because the OLS of the original model can be biased.