Problem Set 1

MACS 30010, Dr. Evans
Mengchen Shi
Jan 7th, 2018

Problem 1

Part (a) The model I found interesting is from the article Culture, Culture, Ethnicity, and Diversity, published in the American Economic Review. This article examines the relationship between ethnic identity and cultural attitudes. The exercise requires individual-level data on answers to questions on norms, values, and preferences, and corresponding data on the respondent's ethnic or linguistic identity.

Part (b)

Desmet, K., OrtuÒo-OrtÌn, I., & Wacziarg, R. (2017). Culture, ethnicity, and diversity. *AmericanEconomicReview*, 107(9), 2479-2513.

Part (c)

$$Q_m = \alpha + \sum_{s=1}^{S} \beta_s D_m^S + \gamma' \mathbf{X}_m + \varepsilon_m \tag{1}$$

where m denotes a respondent, s = 1,...,S indexes ethnolinguistic groups.
Q_m is individual m's answer to the question about culture under consideration.
D_m^S is equal to one if respondent m is part of group s, zero otherwise.
$\mathbf{X}_m$ is a vector of controls. Control variables consist of the respondent's age, sex, education, and household income.
ε_m is an error term.

Part (d). Here in the model, Q_m is the endogenous variable. D_m^S, $\mathbf{X}_m$, and ε_m are exogenous variables.

Part (e). This model is static, linear, and stochastic. It is statics because there is no time dimension in the model. It is stochastic because it has an error term.

Part (f). Place of residence could be a valuable feature that the model is missing. For example, whether the respondent live in a multi-ethnic area or single-ethnic area, which obviously influences the respondent's opinion about norms, values, and preferences.

Problem 2

Part (a, b & c). Since the predicted variable is binary, I use logistic regression model to explain the decision to get married.

$$Pr(M_i = 1|Age_i, Gen_i, ..., Pare_i) =$$
$$F(\alpha_0 + \alpha_1 Age_i + \alpha_2 Gen_i + \alpha_3 Edu_i$$
$$+\alpha_4 Inc_i + \alpha_5 Rela_i + \alpha_6 Pare_i + \sum_{p=1}^{P} \beta_p Eth_i^p + \sum_{q=1}^{Q} \gamma_q Reli_i^q + \varepsilon_i)$$

where

$$F(x) = \frac{1}{1 + e^{-x}} \tag{2}$$

Here, M is the outcome variable whose value is 1 (get married) and 0 (not married);
Age: age of a person, a continuous numerical variable;
Gen: gender of the person;
Edu: years of higher education of the person;
Inc: the natural logarithm of yearly income of the person;
Rela: relationship status, categorized into two groups, single or in a relationship;
Pare: whether $person_i$'s parents are divorced or not;
Eth: ethnicity. Eth_i^P is equal to one if person i is part of group p, zero otherwise, where p = 1, ..., P indexes kinds of ethnicity.
Reli: religion. $Reli_i^Q$ is equal to one if person i is part of group k, zero otherwise, where q = 1, ..., Q indexes kinds of religion;
ε: the error term.

Part (d & e). I think income, education, and marriage of parents (divorced or not) are key factors that influences this outcome. The reasons why I choose those variables are as follows.

First, as known to all, marriage is relevant to wealth share, and the risk of wealth loss due to divorce exists. Hence, people with more income might be more careful about marriage decision.

Second, people with different education levels might hold various opinions about marriage, which affects their decisions to get married.

Last, whether people are positive or negative about marriage is strongly affected by their parents' marriage life (Amato, et al., 1991), which influences people's decision to get married as well.

Part (f). We could do a preliminary test by conducting a small-scaled survey. We can pick people randomly, ask their marriage decisions, and record their personal characters (age, gender, ethnics, and religion, education level, income, relationship status, parents' marriage). With data collected and cleaned, we can run regression to test our model and add/delete variables accordingly to improve the model.

Reference

Amato, P. R., & Keith, B. (1991). Parental divorce and the well-being of children: A meta-analysis. $PsychologicalBulletin$, 110(1), 26-46.