

Problem Set #1

MACS 30100, Dr. Evans

Xiang Zhang

E-mail: snzhang@uchicago.edu

Problem 1

Part (b): Paper Citation

Chetty, Raj, Nathaniel Hendren, and Lawrence F. Katz. "The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment." *American Economic Review* 106, no. 4 (2016): 855-902.

Part (c): Model Settings

The main model specifications of the paper are written below:

$$y_i = \alpha + \beta_E^{ITT} \text{Exp}_i + \beta_s^{ITT} \text{S8}_i + \gamma \mathbf{X}_i + \delta s_i + \varepsilon_i \quad (1)$$

In this model, y_i represents one of the five outcomes of the paper¹. Exp and S8 are indicator variables for being randomly assigned to the experimental and "Section 8" groups respectively, \mathbf{X} is a vector of baseline control variables (including age at which moved and a set of household head characteristics), s is a set of indicators for randomization sites, α is a constant term and ε is the error term.

$$y_i = \alpha + \beta_E^{TOT} \text{TakeExp}_i + \beta_s^{TOT} \text{TakeS8}_i + \gamma_T \mathbf{X}_i + \delta_T s_i + \varepsilon_i^T \quad (2)$$

In this model, y_i represents one of the five outcomes of the paper. TakeExp and TakeS8 are indicators for taking up the experimental and "Section 8 vouchers", respectively. \mathbf{X} , s , α and ε have the same meaning as equation (1).

Part (d): Endogenous/exogenous Variables

In both models, the endogenous variable is y , which is the outcome variable. In model (1), the exogenous variables are Exp, S8, control variables \mathbf{X} , and s . In model (2), the exogenous variables are TakeExp, TakeS8, control variables \mathbf{X} , and s .

Part (e): Model Classification This is a Static, Linear, and Stochastic model.

Part (f): Adding Variables

One possible missing variable is the **marriage status of the child's parents**. According to both psychology and economic literature (see for example, Abidin (1992)), marriage status of the parents could have significant influence on child's lifetime income. However, one thing to note is, this paper uses data from an experiment. In this way, we should not expect the estimation results to change significantly even with additional controls.

¹There are five outcome variables in the paper: individual income, individual college attendance, quality of the college attend, marriage status and fertility.

Problem 2

Part (a) to Part (c): Model Specification

I build a linear probability model to explain whether someone decides to get married². The model takes the following form.

$$y_i = \beta + \beta_1 \text{Age}_i + \beta_2 \text{Gender}_i + \beta_3 \text{Age}_i \times \text{Gender}_i + \beta_4 \text{Educ}_i + \beta_5 \text{Log}(\text{Income})_i + \beta_6 \text{Religion}_i + \beta_7 \text{Ethnicity}_i + \varepsilon_i \quad (3)$$

In this model, the outcome variable is a 0/1 variable indicating whether one decides to get married, with 0 stands for not getting married and 1 for getting married. On the right-hand-side of the model, I control for five variables: age of the individual (Age), gender of the individual (Gender), years of education (Educ), log annual individual income (Log(Income))³, religious belief (Religion) and ethnicity (Ethnicity) of the individual (both variables are categorical variables). Additionally, I add an interaction term between Age and Gender (Age \times Gender) to allow for heterogeneous effects among different age and gender. Finally, β is a constant term, and ε is an error term.

The model I write is a linear probability model. In this model, I am actually estimating the probability of deciding to get married (exactly the same as standard Probit or Logit regression models). The probability of observing a 0 or 1 (whether decide to get married) depends on the right-hand-side (RHS) explanatory variables. The interpretation of the coefficients on the RHS should be: holding everything else constant, one unit increase in variable X will increase the probability of decide to get married by β_i .

Part (d): Key Factors

There are three factors in my model that I think are important: age of the individual (Age), years of education (Educ), and one's individual income (Log(Income)). Conceptually, age is an important determinant of one's marriage decision, and it is generally taken into consideration when one is making his or her marriage decision. Education plays an important role in understanding individual marriage decision. People with different education attainment levels may have different attitudes towards marriage, through which their marriage decision could be affected. Finally, individual income is also a key factor that might influencing one's marriage decision. For example, Becker(1973) suggests that men's and women's decision on getting married instead of staying single depends on their incomes.

Part (e): Explanation for the Choice of Factors

In the previous section, I explain in detail why I choose to control for individual age,

²Conceptually, I could use standard Probit or Logit model to explain one's marriage decision. However, linear probability model yields much better interpretability. Also, it is faster to estimate a linear probability model econometrically.

³Since individual income is skewed distributed, I take the log form of individual income here.

years of education and income level. In this section, I briefly describe why I choose to control for other factors. First of all, gender should be considered when modeling individual marriage decisions (also, it is important to add the interaction term between age and gender). In many places of the world, men and women exhibit large difference in their marriage decisions. For example, in places with “missing girls”, marriage market is largely disturbed by the imbalance sex ratio (see for example, Ebenstein 2010), which heavily influences people’s marriage decisions. Also, one’s religious belief should be incorporated into the model, since people of different religion may have different attitude towards marriage. Finally, one’s ethnical group should be considered, since individuals of different ethnical group may have different customs and thus their marriage decisions can be influenced.

Part (f): Preliminary Test

We could do small-scale surveys to test for the significance of our model. For example, we could post a survey on Amazon MTurk. In the survey, we ask for their marriage decision (whether decide to get married) and document their age, gender, ethnicity, religious belief and individual income. After collecting the data, we could run regression using data collected and see whether we can get significant coefficients and great explanatory power.

Reference

Abidin, Richard R. "The Determinants of Parenting Behavior." *Journal of Clinical Child Psychology* 21, no. 4 (1992): 407-412.

Chetty, Raj, Nathaniel Hendren, and Lawrence F. Katz. "The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment." *American Economic Review* 106, no. 4 (2016): 855-902.

Becker, Gary S. "A Theory of Marriage: Part I." *Journal of Political Economy* 81, no. 4 (1973): 813-846.

Ebenstein, Avraham. "The Missing Girls of China and the Unintended Consequences of the One Child Policy." *Journal of Human Resources* 45, no. 1 (2010): 87-115.