

## Problem Set #1

MACS 30000, Dr. Evans

Haowen Shang

### Problem 1: Classify a model from a journal

**Part (a).** I find a statistic model from an article named “*The Long-Term Effects of Management and Technology Transfers*” in the *American Economic Review*.

The statistic model examines the long-run causal effects of management and technology transfers on firm performance. Management and technology transferred from US firms to Italian firms through the *United States Technical Assistance and Productivity Program (1952–1958)*. For each firm participated in the Productivity Program, managers and engineers were sent to US firms to study management and technology and thus management and technology transferred from US firms to Italian firms. The author collected and digitized balance sheets from five years before to fifteen years after the Productivity Program and linked them to firms’ application records. Using these data, the author compared the performance of firms that applied for and eventually received the management or the technology transfer (treated firms) with that of firms applying for the same transfer, but not receiving it due to the budget cut (comparison firms). Then she found that “performance of Italian firms that sent their managers to the United States increased for at least fifteen years after the program; performance of companies that received new machines increased, but flattened out over time” (Giorcelli, 2019, p.121).

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

Giorcelli, Michela. 2019. “The Long-Term Effects of Management and Technology Transfers.” *American Economic Review*, 109 (1): 121-152.

**Part (c).** The statistic model is

$$outcome_{it} = \alpha_i + v_t + \sum_{\tau=-5}^{15} \delta_{\tau} [Treat_i \cdot (\text{Years After Treat} = \tau)] + \varepsilon_{it}$$

$outcome_{it}$  is the performance of firm which is “metrics of logged (deflated) sales, number of employees, and TFPR of firm  $i$  in year  $t$ ” (Giorcelli, 2019, p.130);

$\alpha_i$  is the firm fixed effects which “controls for variation in outcomes across firms constant over time” (Giorcelli, 2019, p.130);

$v_t$  is the year fixed effects and “controls for variation in outcomes over time that is common across all firms” (Giorcelli, 2019, p.130);

$Treat_i$  is “an indicator that equals one if firm  $i$  is located in a treatment province, eventually selected to participate in the Productivity Program” (Giorcelli, 2019, p.130-131);

Years After Treat =  $\tau$  equals to “the difference between the calendar year  $t$  and the year in which firm  $i$  participated in the Productivity Program” (Giorcelli, 2019, p.131);

$\varepsilon_{it}$  is the error term.

**Part (d).** The exogenous variables in this model are  $Treat_i$  and Years After Treat =  $\tau$  ; The endogenous variable in this model is  $outcome_{it}$ , which is the the output of the model.

**Part (e).** This is a dynamic, linear and stochastic model, because the model is related to time, uses a linear approach and has a random error term.

**Part (f).** I think the variable of *the total numbers of managers and engineers to be sent in US firms* in the treatment group is valuable, because it may have influence on outcomes of firm performance. Firms perform better when they learned more management and technology. And the number of people they sent is a factor influence how much they learned. The modified model is:

$$outcome_{it} = \alpha_i + v_t + \sum_{\tau=-5}^{15} \delta_{\tau} [Treat_i \cdot (\text{Years After Treat} = \tau + TotalNumbers_i)] + \varepsilon_{it}$$

### Reference:

Giorcelli, Michela. 2019. "The Long-Term Effects of Management and Technology Transfers." *American Economic Review*, 109 (1): 121-152.

### Problem 2: Make your own model

**Part (a).** I'll use logistic regression to create a logit model of whether someone decides to get married.

$$Outcome^* = \beta_0 + \beta_1 \cdot Age + \beta_2 \cdot Gender + \beta_3 \cdot Ethnicity + \beta_4 \cdot Religion + \beta_5 \cdot Income + \epsilon$$

then,

$$Y = \begin{cases} 1 = \text{get married,} & \text{if } Outcome^* > 0 \\ 0 = \text{not get married,} & \text{otherwise} \end{cases}$$

Or

$$p(\text{get married}) = \frac{e^{\beta_0 + \beta_1 \cdot Age + \beta_2 \cdot Gender + \beta_3 \cdot Ethnicity + \beta_4 \cdot Religion + \beta_5 \cdot Income + \epsilon}}{1 + e^{\beta_0 + \beta_1 \cdot Age + \beta_2 \cdot Gender + \beta_3 \cdot Ethnicity + \beta_4 \cdot Religion + \beta_5 \cdot Income + \epsilon}}$$

then,

$$Y = \begin{cases} 1 = \text{get married,} & \text{if } p(\text{get married}) < 0.5 \\ 0 = \text{not get married,} & \text{otherwise} \end{cases}$$

*Age* is a continuous variable.

*Gender* is a dummy variable. 1 = *female* and 0 = *male*.

*Ethnicity* is a dummy variable. 1 = *same ethnicity with dating partner* and 0 = *different ethnicity with dating partner* .

*Religion* is a dummy variable. 1 = *same religion and beliefs with dating partner* and 0 = *different religion and beliefs with dating partner*.

*Income* is a continuous variable.

**Part (b).** The dependent endogenous variables in this model is Y, which represents whether people decide to get married(1 = get married and 0 = not get married).

**Part (c).** The model is a complete data generating process. I can simulate data from the model given all the parameters and relationships. Firstly, I can use some

sample data to estimate parameters  $\beta$ . Then after knowing all the parameters and relationships, I can use this model to estimate whether people decide to get married when we have data of their age, gender, ethnicity, religion and income.

**Part (d).** The key factors that influence this model are *Age* and *Income*, because people decide to get married when they are at certain age and have a relative high income to afford the expense of marriage. For example, it's a relatively low possibility that people get married under 20 years old. And it's also a relatively low possibility that people get married when they can't afford the rent of the house. *Gender* may influence marriage decision, but it's not a key factor. *Ethnicity* and *Religion* capture the similarity between the dating couples, but some couples get married because they are similar with each other and some couples get married because they are different and thus they can complement each other.

**Part (e).** Marriage decision is related to person's characteristics and similarities with dating partner. *Age*, *Gender*, *Income* are most representative for person's characteristics. Other characteristics like *Educational Level* and *Experience* may be linear with *Age*, *Gender* and *Income*. People with higher income always have high educational level. And people always earn richer experience when they get older. *Ethnicity* and *Religion* are most representative for couple's similarity. The similarity in *cultural background* may be linear with these two variables because people with same *Ethnicity* and *Religion* always have same *cultural background*.

**Part (f).** The measure of preliminary test are as following:

1. Do a survey with questionnaires. The questions include: "Whether you decide to get married?", "How old are you?", "What's your gender, income, ethnicity and religion?", and "What's your dating partner's ethnicity and religion".
2. Use the data got from the survey and the method of Maximum Likelihood Estimation to estimate the parameters of the model.
3. Find whether the factors in this model are significant in real life.