Problem Set #1

Ari Boyarsky
aboyarsky@uchicago.edu

January 18, 2018

Problem 1

a. Philippe Aghion & Ufuk Akcigit & Angus Deaton & Alexandra Roulet, 2016. "Creative Destruction and Subjective Well-Being," American Economic Review, American Economic Association, vol. 106(12), pages 3869-3897, December.

b. We choose the estimation model from pg. 13:

$$SW\,B_{i,m,t} = \alpha CD_{m,t} + \beta U_{m,t} + \delta X_{i,t} + T_t + \epsilon_{i,t}$$

c. $SWB_{m,t}$ which is subjective wellbeing for individual, i, in metropolitan statistical area (MSA, m, in year, t, is the endogenous variable.

The exogenous variables are creative destruction, $CD_{m,t}$. This variable (from Business Dynamics Statistics) denotes job creation vs. job destruction rates in an area. Unemployment rate $U_{m,t}$. $X_{i,t}$ is a vector of individual controls (gender, age, age squared, race, education level, and family status, and income bracket.). T_t are year and month fixed effects. α, β, δ are OLS parameters.

d. This model is dynamic (time component), linear (OLS estimation), and stochastic (random error term).

e. One possible factor that is left out here is an interaction term between overall level of education, say $E_{m,t}$, in an MSA and creative destruction. This maybe useful as education levels may influence the speed at which someone is abel to retrain for a new job. This could be added to the model at $E_{m,t}CD_{m,t}$.

Problem 2

a. Since the outcome of to get married (1) or not to get married is a binary variable. We use a probit regression model to examine this phenomena. Thus are model takes the following

form:

$$P(Y_i = 1|A_i, S_i, r_i, Y_i, W_i, \omega_i) = \Phi(\beta_0+\beta_1 A_i+\beta_2 S_i+\beta_3 A_i S_i+\beta_4 Y_i r_i+\beta_5 W_i+\beta_6 W_i \omega_i r_i+\beta_7 \omega_i r_i+\epsilon_i)$$

Where, $\Phi(\cdot)$ is the CDF and $(\beta_0 + \cdots + \beta_7\omega_i r_i + \epsilon_i)$ is treated as a Z-score.

b. The endogenous variable, Y_i, is an individual, i, descion to get married or not to get married.

$$Y_i = \begin{cases} 1 & y_i^* > 0 \text{ (Get Married)} \\ 0 & y_i^* \leq 0 \text{ (Do Not Get Married)} \end{cases}$$

where $y^* = \beta\mathbf{X} + \epsilon$. And, $\mathbf{X}$ denotes our exogenous variables in matrix form.

The exogenous variables are A_i, which is individual age. S_i the sex of an individual, this is categorical (0 - Male, 1 - Female, 2 - Other). r_i is an individuals relationship status (0 - identifies not in a relationship, 1 - identifies in a relationship). Y_i is the number of years an indiviudal has been in a relationship. W_i is the wealth of an indiviudal, that is all assets (property, savings, etc.) minus all liabilities (loans, credit card debt, etc.). ω_i is partner wealth, calculated the same as above. $\epsilon_i \sim N(0, 1)$ is the error term.

c. My guess is that the key factors here are age, relationship status, and length of relationship. This is because generally speaking a person needs to be in a relationship to get married. Second, a person generally considers their age when they get married, this also acts as a proxy for things such as out of college, more established profesionally, and other life factors that are often correlated with age. Finally, years in a relationship since, the longer someone is in a relationship the more we would expect that they get married as they have invested more and more time with their partner.

d. I choose the factors above because they generally have some relationship to an individuals decision to get married. As we discussed above, age, relationship status, and length of relationship have very obvious correlations to an individuals marriage decsion. Notice in our model we use an interaction term for years in a relationship and partners wealth since if they are not in a relationship these factors have no meaning. We also consider sex, and the interaction of age and sex since in many cultures there is a social desirability bias that proscribes when individuals of a certain sex should get married. We also look at individual wealth, partner wealth and the interaction of the two since people are generally more likely to embark on marriage if they feel financially comfortable. Additionally, there is rationality component here where if they are not financially secure and their partner is secure, or vice-versa, they may be more or less willing to get married. For instance, if one's partner is wealth relative to themselves they may be more willing to get married to that person. Finally, our normally distributed error terms completes the data-generating process and allows us to account for variations. These factors also act for proxy's such as income level, employment status, cultural background, sexual orientation, and others. Hence, we choose these variables as we expect them to be correlated with other useful

factors. We also believe the causal process of how income may affect the marriage descion may actually act through age and wealth rather than be direct.

e. We could test this model by conducting an online survey on mech. turk where we ask for age, sex, est. income level, est. debt level, relationship status, est. partner income, est. partner debt, length of relationship, and then "Would you say yes if your partner proposed today?" with "Yes or No" as option answrs. We would also ask other questions such as classification questions to distract them for what the survey may be about. We will also ask for other demographic data - race, region, education level,etc., so that we can adjsut our data to the U.S. census. Then we will run our probit regression on our model to estimate the β_i parameters via MLE. This will tell us how each term is correlated with our outcome. Finally we can examine psuedo R^2, fraction correctly predicted, and p-values to identify which factors are strong predictors and if we have an overall good preliminary model.