Problem Set #1

MACS 30100, Dr. Evans
Sushmita V Gopalan

Part 1(b). Bettinger, Eric P., and Bridget Terry Long. "Do faculty serve as role models? The impact of instructor gender on female students." The American Economic Review 95, no. 2 (2015): 152-157.

Part 1(c).

$$y_{ik} = \alpha + \beta(FemaleInstructor)_{ik} + \gamma X_i + \delta Z_{ik} + \epsilon_{ik}$$

Here, y_{ik} is the proportion of courses that student i took in subject k after the semester in which she was first exposed to the subject.

$FemaleInstructor_{ik}$ is the proportion of courses that student i took in subject k in the first semester she was exposed to the subject that were taught by a female faculty member.

X_i includes controls for student demographics and ability such as age, gender, race, state of residency and ACT score.

Z_{ik} includes controls for student i's interaction with subject k, including whether the subject is in the students intended major, the semester the student took the course, and the number of credit hours students attempted that semester.

Part 1(d). The exogenous variable is y_{ik}, as described in (c). The endogenous variables are the explanatory variables described in (c), i.e. $FemaleInstructor_{ik}$, X_i and Z_{ik}.

Part 1(e). The model is linear - it uses the Ordinary Least Squares method. The model is stochastic - it contains an error term. The model is static.

Part 1(f). The model is fairly comprehensive in that covers all of the key variables that I think are relevant. One variable that could add value would be a dummy variable that reflects whether or not the student has a mother who is a member of the workforce - there is weak evidence to suggest that female students with working mothers are more likely to see women as role models. One control that I believe should be included is the student i's average grade in subject k as this is likely to strongly correlate with the subsequent number of courses taken in that subject.

Part 2(a).

$$y_i = \alpha + \beta_1(G_i) + \beta_2(R_i) + \beta_3(A_i) + \beta_4(T_i) + \beta_5(X_i) + \epsilon_i$$

y_i is the predicted lifespan of musician i. G_i is a categorical variable reflecting the genre of music musician i is associated with. R_i is the number of times musician i has checked into a rehabilitation facility. A_i is the age at which musician i had their first hit single. T_i is a dummy variable that reflects whether or not the musician i is

past the age of 27. X_i includes factors such as gender, ethnicity and age, country of residence.

Part 2(d). In my model, the key factors affecting the lifespans of famous musicians are the genre they belong to, history of drug and alcohol abuse, how early they began to achieve fame and success and whether or not they are 27 years of age. The remaining variables such as age, gender, ethnicity serve as controls - these factors affect the lifespans of all human beings, not only famous musicians. **Part 2(e).**

The key factors in my model were chosen for the following reasons:

Genre - This model includes a categorical variable to reflect the genre of music practiced by the particular musician such as pop, rock, country, jazz, classical, rap, hip-hop, etc. Different genres of music tend to have different lifestyles and cultures associated with them, which have a bearing on lifespan.

Number of rehab visits - A preliminary examination of media reports suggests that drug alcohol use is often associated with the deaths of famous musicians. In my model, I include a variable reflecting the number of rehab visits undertaken by a musician as a proxy for the degree of seriousness of their drug use problem. This variable could however, be problematic in two ways - rehab visits are not always public knowledge and the incidence of rehab visits could either indicate intense substance abuse or a proactive intention to deal with a serious issue - each of which would influence lifespan in opposite directions.

Age at first hit - I include this variable because early success and the associated adoption of a particular lifestyle at a young, impressionable age could potentially have an impact on a musicians lifespan.

Past 27 - There exists a theory in popular culture, backed up by some anecdotal evidence, that several famous musicians tend to die at the age of 27. If there is indeed a spike in the probability of death at 27, getting past that hurdle could increase predicted life expectancy.

Other factors that could influence lifespan but are either hard to measure or require information that is not in the public domain are family stability, fluctuations in success levels, preexisting medical conditions.

Part 2(f). A preliminary test to examine whether these factors are significant would be to draw a sample of 20 famous musicians who have died in the last five years and 20 that remain alive, generate predictions for lifespan using this model as a data generating process, using their age in 2010. We would then compare these figures to their actual lifespans (for dead musicians) and check whether musicians whose death was predicted by the model remain alive. This examination would give us some idea as to how our model holds up.