

Problem Set #1

MACS 30100, Dr. Evans

Yinxian Zhang

Question 1 Classify a Model from a Journal

Part (a) and (b). Find a recently published article and give a detailed citation.

Citation: Glass, Jennifer, Robin W. Simon, and Matthew A. Andersson. "Parenthood and Happiness: Effects of Work-Family Reconciliation Policies in 22 OECD Countries 1." *American Journal of Sociology* 122, no. 3 (2016): 886-929.

Part (c). Write down the model.

The fixed-effects model includes three equations as following:

$$Y_{i,j} = \beta_{0,j} + \beta_{1,j}X_{1,i,j} + \beta_{2,j}Z_{i,j} + E_{i,j}, \quad (1)$$

where $Y_{i,j}$ is the self-reported happiness of individual i in country j , $X_{1,i,j}$ is the individual's parental status. $Z_{i,j}$ is a vector of personal attributes, including age, gender, education, income decile, marital and employment status.

And,

$$\beta_{0,j} = \lambda_0 + \lambda_1 X_{1,j}, \quad (2)$$

$$\beta_{1,j} = \alpha_0 + \alpha_1 X_{1,j}, \quad (3)$$

where $X_{1,j}$ is a vector of variables including country-level policies and conditions, e.g. GDP, paid parental leave, work flexibility etc.

Part (d). List which variables are exogenous and which are endogenous.

In this model, individual's happiness $Y_{i,j}$ is endogenous. It is the output of the model. Meanwhile, individual's parental status ($X_{1,i,j}$), other personal attributes ($Z_{i,j}$) such as age, gender, education etc., and country-level variable vector $X_{1,j}$, are all exogenous. They are independent variables outside the model.

Part (e). Classify the model.

First, this model is static rather than dynamic. All individual- and country-level exogenous variables describe situations at one time point without considering impact of previous and/or future situations. Second, this model is a linear model with all three equations being linear regressions. Lastly, this model is stochastic. In equation (1), $E_{i,j}$ is the error term, indicating stochasticity.

Part (f). List a variable of feature that you think the model is missing.

This paper uses fixed-effects model to explain the variation in the parenthood gap in happiness (parents are less happy than non-parents) across countries. The authors mainly consider country-level differences in state policies, such as paid parental leave, child-care cost, average work hours etc. However, societal/cultural factors may be equally important. For instance, Asian countries like Japan may exert more pressures on parents, for example mothers are

expected to cook for children on their own rather than ordering in or eating out. Likewise, parenting in China is more likely to be interfered by grandparents and close relatives. Therefore, we can add a cultural variable to describe to what extent parenting would be judged and interfered by other people in certain culture. Inglehart's country-level index of post-modern values could be a possible data source.

Question 2 Make Your Own Model

Part (a). Write down a model of how long popular musicians live

$$L_{i,j} = t_i + \beta_i(\mu_j - t_i),$$

$$\beta_i = \rho * G_i + \alpha * Inc_i + \lambda * Risk_i$$

where $L_{i,j}$ is the estimated lifespan of individual musician i living in country j . t_i is the current age of the musician. μ_j is the average life expectancy in country j (as reported by WHO). G_i is gender and Inc_i is the income decile of the musician. $Risk_i$ is a vector of dummy variables describing the health risks of the musician, including behaviors of smoking(1 if yes), excessive drinking (1 if yes), drug using(1 if yes), and diagnoses of heart diseases and cancers(1 if yes).

Part (d). Key factors that influence the outcome.

First, I consider the average life expectancy of the country where the musician lives. This variable reflects the quality of healthcare, ongoing wars, HIV infections and so on in a given country. These country-dependent conditions will influence one's life expectancy at a macro level. Then, individual attributes are closely associated with longevity. Females significantly lives longer than male; income level by and large determines lifestyle and the quality of healthcare one could have. Lastly, risk behaviors have a strong negative impact. Popular musicians and entertainment stars are always reported to be involved with smoking, drinking and even drug-using. These behaviors may lead to higher risks of death. On the other hand, being diagnosed with heart disease and/or cancer is highly dangerous. David Bowie died of liver cancer two years after diagnosis.

Part (e). Why these factors but not others?

There are a few other variables which are also important factors of longevity, including race, education, occupation, stress and so on. However, I did not include them in my model. The reason is two-fold. On one hand, my research subjects are popular musicians in the same industry(so occupation is less meaningful). This population are mostly celebrities whose race and educational level may not significantly impact their life quality. Stress can negatively impact musician's life expectancy, but I suspect that the variation in stress of popular musicians may not be significant, as they are working in the same industry and facing similar pressures. On the other hand, I need to keep my model simple and concise. So I mainly focus on key factors such as risk behaviors and certain diseases. According to news reports, these factors are top reasons of death of popular musicians. To complement the lack of other sociodemographic variables, I also include income decile in my model. Taken together, I decided to include those factors in my model and not others.

Part (f). How to do a preliminary test?

First, I can use existing survey data from the Census Bureau to test this model, assuming that these factors will also work (perhaps to a different extent) for average people. This can help me evaluate if certain factor is not significant at all. Then I can collect data of dozens of passed-away musicians to further test the validity of this model applied to the particular population.