

Exercise 1

I choose to study the model of the paper "Way Down in the Hole: Adaptation to Long-Term Water Loss in Rural India" (Blakeslee, David, Ram Fishman, and Veena Srinivasan. 2020). The paper stated that adverse environmental conditions have the potential to undermine progress in reducing rural poverty. However, the knowledge is limited about the prospects for farmers to adopt adaptation measures to mitigate this threat, especially the opportunities for income diversification through recent non-agricultural growth. The paper used the quasi-randomness of groundwater access determined by geological conditions to study the impacts of water scarcity in India. The drying up of wells has led to a sharp and continuous decline in agricultural income and wealth, with few signs of agricultural adaptation. However, the redistribution of labor to non-farm employment appears to be able to successfully maintain overall income, especially in the more developed manufacturing sectors.

Their estimated regression model is that

$$y_{i,v} = \alpha_1 + \alpha_2 F_i + X_i \Phi + A_v + B_t + u_i$$

where i is a household index and v a village index, $y_{i,v}$ is the outcome of interest, F_i is a binary indicator of "whether the first borewell drilled by the household has failed by the time of the survey", X_i is a vector of household characteristics, A_v are village fixed effects, and B_t are fixed effects for the year t in which household i drilled its first borewell. The household characteristics include the age, caste, and literacy of the household head, as well as the total land inherited by the household (Blakeslee, David, Ram Fishman, and Veena Srinivasan, 2020).

Therefore, we have the exogenous variables are F_i , X_i , and A_v . We have $y_{i,v}$, u_i , and B_t are endogenous variables.

The model is dynamic with the term B_t . And it is an non-linear model. The model is stochastic.

Exercise 2

We have the model as following:

$$y_{i,l,r} = \alpha_1 + \alpha_2 G_i + \alpha_3 D_i + X_i \Phi + M_l + M_r + u_i$$

where i is a index of the person, l is the spacial index of the location where the person lives, r is a index of the religious preference, and $y_{i,l,r}$ is the outcome between 0 and 1, which indicates that the person would get married(1) or not(0). We have G_i is a binary indicator of gender and D_i is binary indicator of whether the person's parents are divorced. We have X_i is a vector of the person's characteristics including age, income, education level, number of previous relationships, and amount of debt. Then we get M_l is the fixed effects of living location, and M_r is the fixed effects of religious belief.

We choose gender and parents' marriage status as two binary indicators since these two variables could directly represent a person's attitude towards marriage. The personal characteristics we choose are based on general social survey questions and an article from the New York Times (<https://www.nytimes.com/interactive/2016/03/23/fashion/weddings/marriage-questions.html>). We also add a spacial factor since in each area there could be cultural differences; besides, we add the effects of religious belief since each religion might have a different definition of marriage.

For the preliminary test, we decide to do a survey of 100 individuals and use the results to test the model. Based on the statistical outcomes we get, we could evaluate which variables are significant in real life and which are not.