Final Report

Todd Albertson

Jenny Fingles

Max McDonald

Austin Sell

March 12, 2019

Introduction and Motivation

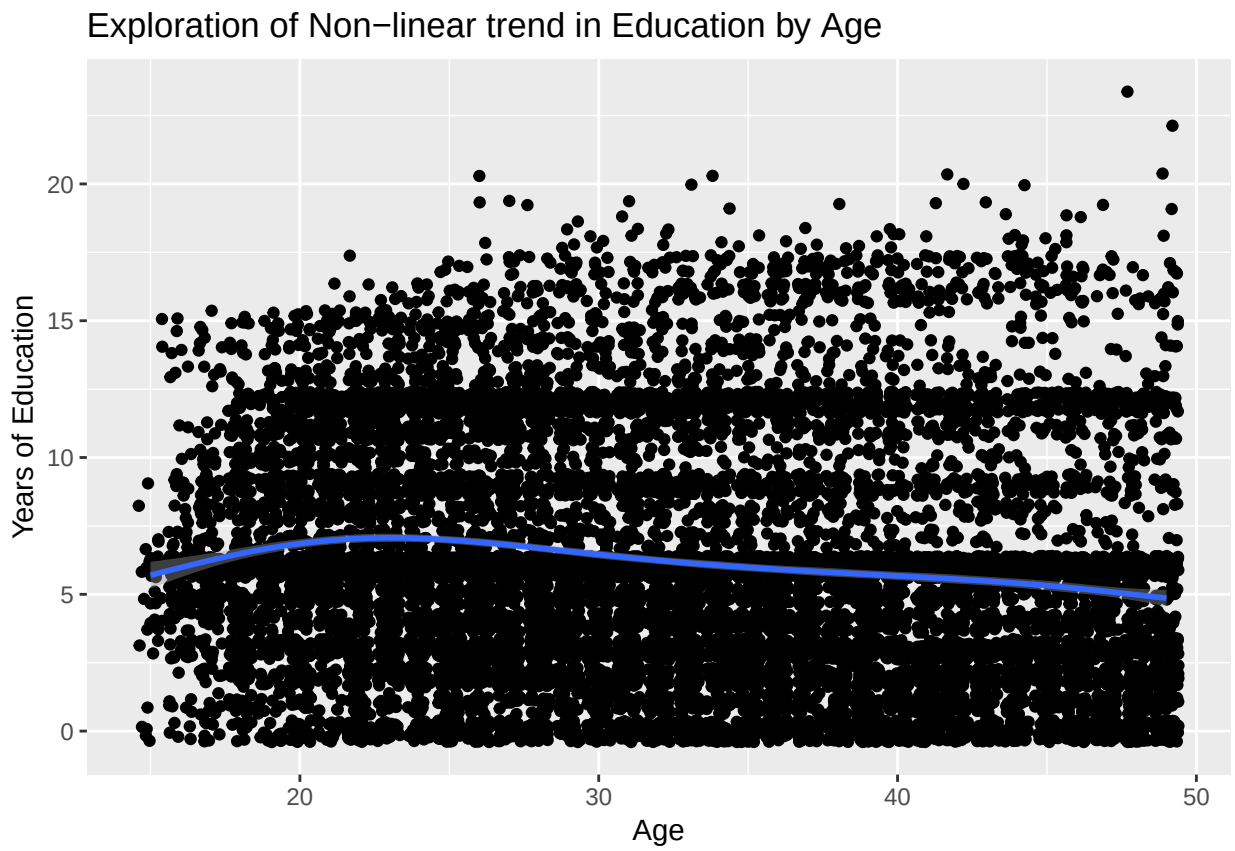
Our research examines the educational attainment of women in Honduras. While there are many things that can play a role in education, we strive to examine the impact of religion and wealth.

Data

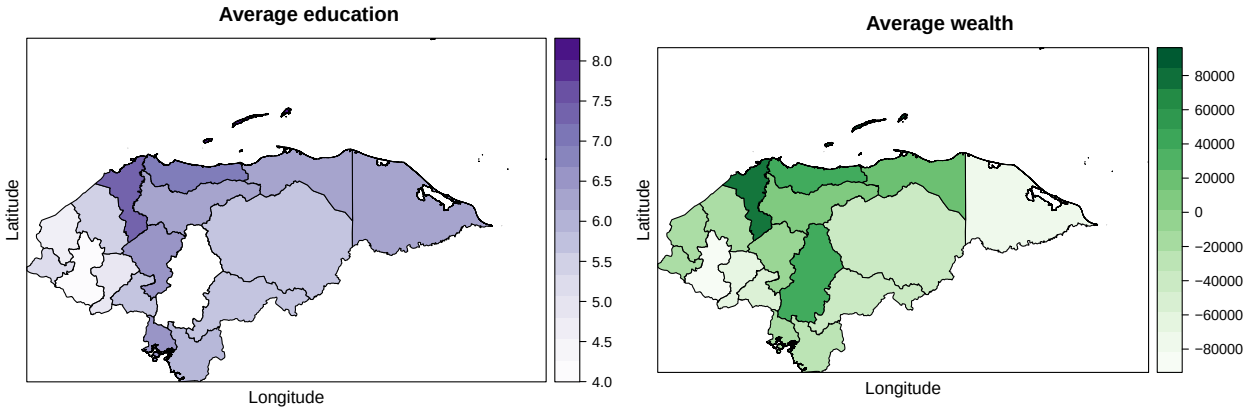
Our data comes from Demographic and Health Surveys (DHS). Specifically, we pull from the responses of over 15,000 Honduran women in 20XX. The data that we use in our analysis includes years of education, age, religion, wealth, urbanicity, and region.

Analysis

Two primary concerns motivated the model construction. First, we examined the possibility of a curvilinear impact of age on education. The figure below reveals that there is a slight increase in education before a gradual decline in the older generation. Notably, the variance of education also increases with age. We think that there are likely two effects present here. The first is that young adults now are far more likely to pursue education now than they were a few decades ago. Thus, the older generations have lower levels of education. The second effect is that as you age, the probability of attaining additional years of education increases, which is evidence from the initial bump that we observe. Given the cross-sectional nature of our data, we cannot disentangle these effects. As such, we model age as a quadratic function which we believe (and our model diagnostics show) is the best fit for the mix of these effects.



The second concern we have is the heterogenous distribution of key variables across the administrative regions of Honduras. As can be seen in the two figures below, we observe very noticeable differences in both education and wealth across the different regions of Honduras. Additionally, these patterns seem to overlap. To account for the variation across the administrative regions, we fit random intercepts for these regions. We also considered the possibility of included region-fixed effects, but leave-one-out cross-validation revealed that random intercepts performed the best.



The final model that we fit is defined as the following:

$$EDUC_{ij} = \beta_0 + \beta_1 AGE_{ij} + \beta_2 AGE_{ij}^2 + \beta_3 WEALTH_{ij} + \beta_4 RELIGION_{ij} + \beta_5 URBAN_{ij} + \mu_j + \epsilon_{ij}$$

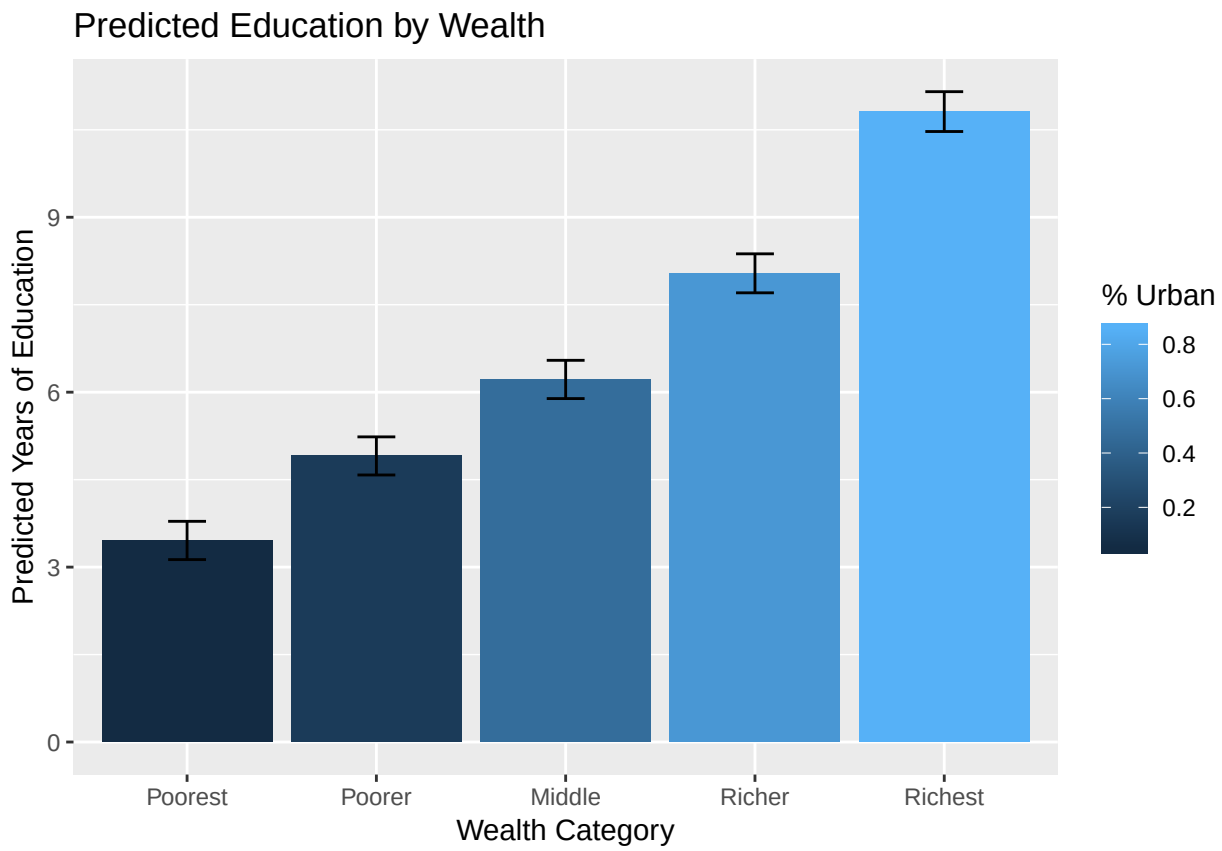
$$\epsilon_{ij} \sim N(0, \sigma^2)$$

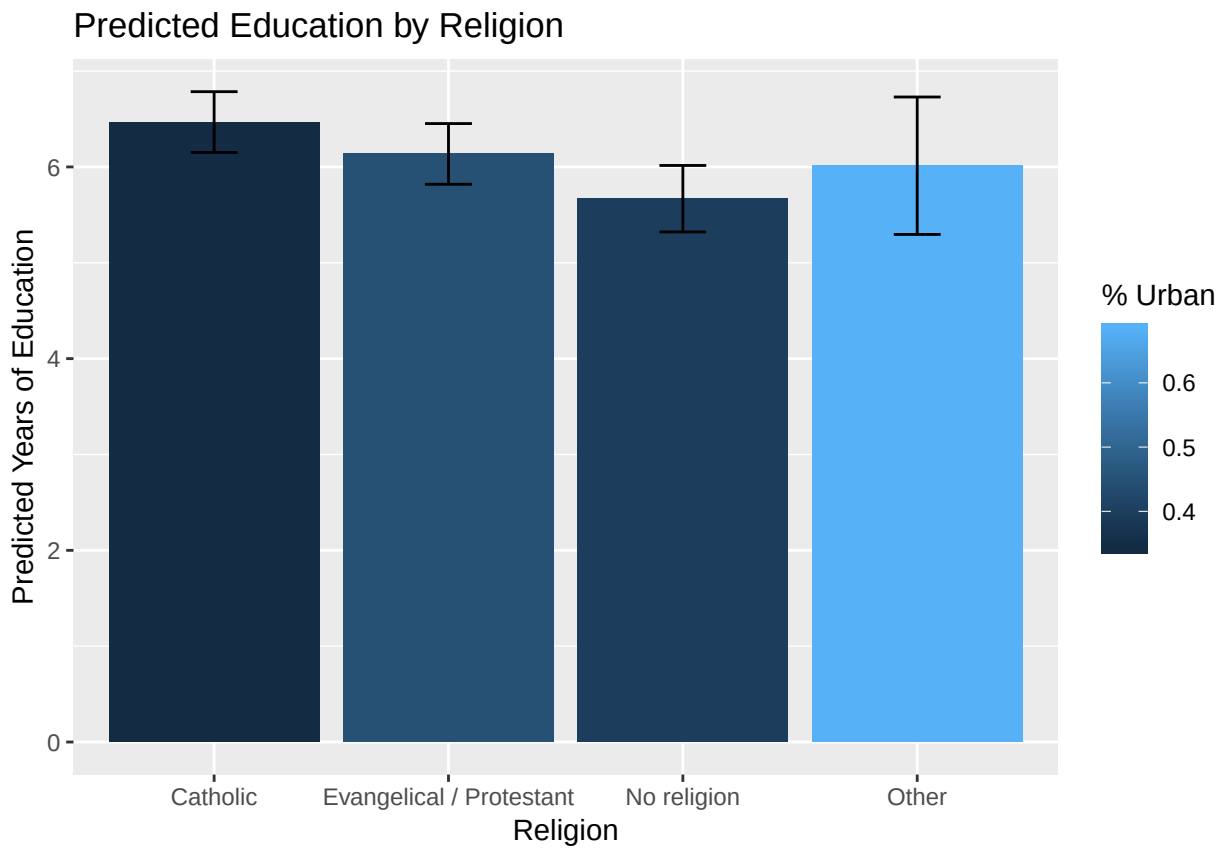
$$\mu_j \sim N(0, \tau^2)$$

The years of education for individual i in region j are predicted by the above model where $WEALTH$ and $RELIGION$ are factor variables, and μ_j is the random intercept for each region.

Results

The two primary takeaways are below.


Predicted Education by Wealth
Predicted Years of Education
0
3
6
9
Poorest
Poorer
Middle
Richer
Richest
Wealth Category
% Urban
0.8
0.6
0.4
0.2



Conclusion

Discussion of the substantive interpretation of the model, including any meaningful impact of findings as well as limitations of model.

References