Midterm Report

In the course final project, the project I want to carry out will be a causal inference model integrated with machine learning. My research will generally have 2 questions, Q1: Does individuals' level of religiosity causally affect trust in labor unions? Q2: Does this effect change as economic insecurity increases or as income level decreases?

The main source from which the variables in the model will be drawn will be the World Values Survey, and with the data of four waves between the 4th and the 7th wave of this survey, a longitudinal dataset will be created and used as the input and output of the model. It will not be panel data, it will be examined as repeated cross-sections. By fixing wave dependence, fixed effects will also be taken into account.

- Outcome (Y): Trust in labor unions. An index will be created from the questions corresponding to the content and that will be used as the output.
- Treatment (T): While calculating the religiosity score, an index will also be created and a latent score will be created as a continuous index.
- Moderator (M): The economic insecurity score will also be used as a continuous index.
- Covariates (X): age, gender, education, income, employment, marital status, etc. will be considered.
- Fixed effects (FE): Wave fixed effects (if I can obtain meaningful results, region and wave x region FEs will also be examined).

According to the questions in the waves, variables will be organized through the common question set, and the values in question sets where important questions are missing will be filled with the average values within the indices. Values will be reduced to z-scores and normalized, and the reliability of the scores will be measured with various tests (item total correlations, etc.). The output, trust in labor unions, will be handled as a scale, and for robustness it will be checked whether there is a relationship in the same direction with logit/probit. The confounder dataset (X) will be considered in the main model generally as age, gender, education and other values, while as robustness it will also be examined specifically for potentially risky controls (left–right positioning and trust in institutions may generate each other, etc.).

Modeling will be carried out in 3 stages. In the first stage, as a benchmark, a regression model including fixed effects will be established and in this model T, M and TxM will be together. In this way, we will have an idea about the direction and magnitude of moderation. Effects here are in a form that can be reported as marginal effects, and the effect of standardized M in terms of standard deviation can be shown separately.

In the second stage, the main causal result will be examined with the Double/Debiased Machine Learning tool. The purpose of the supervised method here is to estimate the part of the relationship between religiosity and trust in labor unions that is explained by both covariates and contextual differences (wave and region if it is possible) with flexible machine learning models (Nuisance Learning), and to obtain causal parameters from the remaining residual relationship. To bring moderation into the model, a "treatment-alike" term is defined here as D1=T and D2=TxM. The control set (W) is constructed as X, M and wave FE (different FEs can also be examined in robustness). Then, by applying cross-fitting to our data, the data will be split into K folds and the folds will be stratified by wave. The nuisance functions will be trained only on the other folds to estimate:

(i) $g(W) \approx E[Y|W]$ (ii) $m(W) \approx E[T|W]$, (iii) $(W) \approx E[T \cdot M|W]$.

Residuals will be produced in the test fold here, and in the final step by combining all fold residuals, the regression $Y=\theta D1+\kappa D2+u$ will be established and the estimation of the baseline effect of religiosity (θ) and the effect change according to economic insecurity (κ) will be carried out. Since wave fixed effects are "one-hot encoded" into W, the effects including level differences and wave shocks are based on within-wave comparison.

In the third stage, in order to present moderation convincingly, the results will not be left with a single interaction coefficient; the marginal effect function will be reported as $\text{Effect}(M)=\theta+\kappa M$, and the effect of religiosity will be presented separately at low/medium/high levels of M (e.g. -1SD, 0, +1SD).

As the next step, in order to examine moderation more flexibly without squeezing it into a single linear interaction coefficient, a CATE (Conditional Average Treatment Effect) analysis based on Orthogonal / Generalized Random Forests (ORF/GRF) will be conducted. The model will be built with the "honest splitting" principle (tree splits and within-leaf effect estimation will be carried out with different sample parts), the risk of overfitting will be controlled with hyperparameters such as approximately ~2,000 trees and moderate depth, and as outputs (i) CATE maps/heatmaps (e.g., M slices x another important covariate level), (ii) variable importance (which variables explain heterogeneity the most) and (iii) effect vs M partial dependence (the average effect curve across M) will be reported. In addition, by comparing to what extent the "effect vs M" curve obtained with ORF/GRF overlaps with the $\text{Effect}(M)=\theta+\kappa M$ structure from DML in the second stage (is it approximately linear, or is there nonlinear divergence in certain M ranges), the interpretation of the findings will be strengthened.