Explanatory Memo

|  |  |  |
| --- | --- | --- |
|  | File | Description |
| 1 | parent\_survey\_raw\_deid.dta | The deidentified raw data in wide format, including variables from the two waves of the data collection. |
| 2 | deid\_data\_prep.do | Stata commands performed on the data to clean the data variables and create indices. The file starts with cleaning the deidentified version of the data, reshapes the data to long format to create indices across the two waves and then reshapes back to wide format as this is the format that will be used for the analysis. It creates a csv output file that will be further processed in R |
| 3 | parent\_survey\_deid\_wide.csv | Output data from stata |
| 4 | log\_stata.smcl | Log file of the code in deid\_data\_prep.do |
| 5 | deid\_data\_analysis.R | R script performed on the data for further cleaning and performing analyses of the research. It first filters out observations with missing values in the COVID-19 accina variable (covid\_vaccine) and runs the following analyses:   1. Missing data imputation: It examines the percentage of missing values in each variable. Then applies multiple imputation using predictive mean matching method in the MICE package. 2. Logistic regression: Performed to study the determinants of COVID-19 vaccination. Odds ratios are reported with their associated confidence intervals. 3. Propensity score matching: It first starts by examining the balance across vaccinated and unvaccinated participants in the data. After that, nearest neighbor matching with caliper 0.25 is performed. The matching method is performed by checking the standard mean differences between vaccinated and unvaccinated groups using demographic variables in the matched data. 4. Treatment effects estimation: Multiple regression is performed on the matched data to estimate effects of COVID-19 vaccination, accounting for a number of controls. 5. Supplementary analyses: act as robustness checks for matching method and selection of controls in the treatment effect estimation. |
| 6 | logofRcode.txt | Log file of the R script deid\_data\_analysis.R |
| 7 | Codebook.xlsx | Data codebook |