Replication Instructions

This provides the instructions for the results presented in the manuscript titled "The Historical Sources of Language Policy". The analysis has been carried out using Stata 15.

1. Replication for the Figure 1 titled "Official language choice, human capital, political engagement and assets" and Table A1 titled "Association between official language choice and human capital, political engagement and assets".
   a.i) Please register at the website https://dhsprogram.com/ to access the datasets.
   a.ii) The data included is men and women survey from the latest round of the standard DHS from all countries.
   a.iii) The list of the data rounds included is in folder labelled "**DHS_data_list**" and the files are titled "mendata.txt", "mendatalit.txt", "mendata.txt" and "mendata.txt". The directory of the folder is as follows: "**ReplicationMaterial/Data/DHS_data_list**"
   a.iv) Please run the do-files available in the folder titled "Preparation" and titled **prep_data.do** and **prep_data_v133.do** to create the final data files.
   a.v) The Figure 1 and Table A1 are then reproducible by running the do-file titled **Figure1_MainManuscript_TableA1_OnlineAppendix.do** and available in the analysis folder. For data replication for final acceptance of manuscript, the constructed dataset titled "**master_data_women_men**" is also included but this is from the Demographic and Health Survey (DHS) and thus also codes to construct the data from the DHS website is provided above.

2. Replication of Tables 2 and 3 and Figure 3 and 4 in the main manuscript.
   a.i) Please use the dataset titled **rep_data_histlpolicy.dta** available in the Data Folder.
   a.ii) The figures and tables are reproducible by running the do-file titled **Replication_MainManuscript_Table2&3_Figures_3&4.do** available in the analysis folder.

3. Replication of Online Appendix Tables A2-A12 and Figure A1 in the Online Appendix.
   a.iii) Please use the dataset titled **rep_data_histlpolicy .dta** available in the Data Folder.
   a.iv) The figures and tables are reproducible by running the do-file titled **Replication_Appendix_TablesA2toA12_FiguresA1_OnlineAppendix.do** available in the analysis folder.

4. Replication of Table 1 titled "Language repertoires and national cultures"
   a.i) Use the data file titled **Table1.dta** available in the data folder.
   a.ii) The tables is reproducible by running the do-file titled **Table1_MainManuscript.do** available in the analysis folder.
   a.iii) The raw data from Afrobarometer employed in Table 1 is tiled "**merged4.dta**" and is stored in the folder titled

"AfrobarometerDatatoPrepare" with the directory "ReplicationMaterial/Data/AfrobarometerDatatoPrepare"

a.iv) The two key explanatory variables used in Table 1 are created using the do-files titled "**Repoitores_Only_colonial_Lang.do**" and "**Repoitores_W_colonial_Lang.do**" and available under the directory "ReplicationMaterial/DoFiles/Codes/Preparation/Afrobarometercommunicationmeasures"

5. The replication folder also contains a folder called "**codebook**" with the names of all variables and a definition of each variable.