

Extension of the analysis of the article “Relationship of gender differences in preferences to economic development and gender equality”

Output of the Analysis

The result of the analysis is written into two csv files, and these files can be used to reproduce the plots and for comparison to any other analysis based on this method.

The files are:

- **extended_data_aggregatedByCountry_preferencePCA_genderIndexPCA.csv** Data aggregated by country containing the Average Gender Differences (the first component of the PCA made on the six preferences for the gender coefficient), all the indicators of the economic development and gender equality, plus their Standardization, and the residuals that can be used to build Fig. 2. The data consists of 42 variables:
 - *country* is a character
 - *avgGenderDiff* is numeric and it is the result of the PCA on the gender coefficients at country level
 - *isocode* is a character
 - *logAvgGDPpc* is numeric and it is the logarithm of the average GDP of the country in the period from 2003 to 2012, in 2010 US dollars
 - *Date* is integer and corresponds to the time of women’s suffrage
 - *ScoreWEF* is numeric and corresponds to a score extracted from the WEF Global Gender Gap Index
 - *avgRatioLabor* is numeric and is the average of the labor of females divided by the labor of males in the country
 - *ValueUN* is numeric and corresponds to a score extracted from the UNDP Gender Inequality Index
 - *region* is a character and corresponds to the region of the world in which the country belongs
 - *telephone* is logic, and it is TRUE if the survey has been performed by telephone in that country
 - *personal* is logic, and it is TRUE if the survey has been performed face-to-face in that country
 - *avgGenderDiffRescaled* is the same measurement as *avgGenderDiff*, but rescaled using the min-max method.
 - *GenderIndex* is numeric and it is the first component extracted from the PCA performed on the four measurements of Gender Equality Indexes used in FH article (*Date*, *ScoreWEF*, *ValueUN*, and *avgRatioLabor*)
 - *logAvgGDPpcStd* same, but standardized
 - *ScoreWEFStd* same, but standardized

- *ValueUNStd* same, but standardized
- *DateStd* same, but standardized
- *avgRatioLaborStd* same, but standardized
- *GenderIndexStd* same, but standardized
- *GenderIndexRescaled* same, but rescaled using a min-max method
- *residualsavgGenderDiffStd_GDP* is the variable created from the *avgGenderDiffStd* residualised using *logAvgGDPpcStd*
- *residualslogAvgGDPpcStd_GEI* is the variable created using the *logAvgGDPpcStd* residualised using *GenderIndexStd*
- *residualsavgGenderDiffStd_GEI* is the variable created from the *avgGenderDiffStd* residualised using *GenderIndexStd*
- *residualsGenderIndexStd* is the variable created from *GenderIndexStd* residualised using *logAvgGDPpcStd*
- *residualslogAvgGDPpcStd_WEF* is the variable created using the *logAvgGDPpcStd* residualised using *ScoreWEFStd*
- *residualsavgGenderDiffStd_WEF* is the variable created using the *avgGenderDiffStd* residualised using *ScoreWEFStd*
 - *residualsScoreWEFStd* is the variable created from the *ScoreWEFStd* residualised using *logAvgGDPpcStd*
 - *residualslogAvgGDPpcStd_UN* is the variable created using the *logAvgGDPpcStd* residualised using *ValueUNStd*
 - *residualsavgGenderDiffStd_UN* is the variable created using the *avgGenderDiffStd* residualised using *ValueUNStd*
 - *residualsValueUNStd* is the variable created from the *ValueUNStd* residualised using *logAvgGDPpcStd*
 - *residualsavgRatioLaborStd* is the variable created from the *avgRatioLaborStd* residualised using *logAvgGDPpcStd*
 - *residualsDateStd* is the variable created from the *DateStd* residualised using *logAvgGDPpcStd*
 - *residualslogAvgGDPpcStd_LFP* is the variable created using the *logAvgGDPpcStd* residualised using *avgRatioLaborStd*
 - *residualsavgGenderDiffStd_LFP* is the variable created using the *avgGenderDiffStd* residualised using *avgRatioLaborStd*
 - *residualslogAvgGDPpcStd_TSWs* is the variable created using the *logAvgGDPpcStd* residualised using *DateStd*
 - *residualsavgGenderDiffStd_TSWs* is the variable created using the *avgGenderDiffStd* residualised using *DateStd*
 - *residualslogAvgGDPpcStd_GDI* is the variable created using the *logAvgGDPpcStd* residualised using *GDIStd*
 - *residualsavgGenderDiffStd_GDI* is the variable created using the *avgGenderDiffStd* residualised using *GDIStd*
 - *residualsGDIStd* is the variable created from the *GDIStd* residualised using *logAvgGDPpcStd*

- **extended_data_aggregatedByCountry_singlePreference_genderCoefficients.csv** Data aggregated by country but separating each of the six preferences gender difference values. The data set includes 90 variables, here below described:
 - *country*, *isocode*, *logAvgGDPpc*, *GenderIndexRescaled*, *GDIStd*, *ScoreWEFStd*, and *ValueUNStd*, are the same variables as described above
 - *gender* is the coefficient related to the robust linear regression on the preference
 - *genderOrig* is the original value of the *gender* coefficient, before inverting for those preferences requiring it (negative reciprocity, risk taking, and patience)
 - *preference* is the same as described above, and in this dataset we kept the six preferences distinguished and not combined into a PCA
 - *residualsgenderGEI_trust*, *residualsgenderGEI_altruism*, *residualsgenderGEI_negrecip*, *residualsgenderGEI_posrecip*, *residualsgenderGEI_risktaking*, *residualsgenderGEI_patience* are built performing a robust linear regression of the gender coefficient of the specific preference and the Gender Equality Index, and then calculating the residuals from it
 - *residualsgenderWEF_trust*, *residualsgenderWEF_altruism*, *residualsgenderWEF_negrecip*, *residualsgenderWEF_posrecip*, *residualsgenderWEF_risktaking*, *residualsgenderWEF_patience* are built performing a robust linear regression of the gender coefficient of the specific preference and the WEF GGI, and then calculating the residuals from it
 - *residualsgenderUN_trust*, *residualsgenderUN_altruism*, *residualsgenderUN_negrecip*, *residualsgenderUN_posrecip*, *residualsgenderUN_risktaking*, *residualsgenderUN_patience* are built performing a robust linear regression of the gender coefficient of the specific preference and the UNDP GII, and then calculating the residuals from it
 - *residualsgenderGDI_trust*, *residualsgenderGDI_altruism*, *residualsgenderGDI_negrecip*, *residualsgenderGDI_posrecip*, *residualsgenderGDI_risktaking*, *residualsgenderGDI_patience* are built performing a robust linear regression of the gender coefficient of the specific preference and the UNDP GDI, and then calculating the residuals from it
 - *residualsgenderGDP_trust*, *residualsgenderGDP_altruism*, *residualsgenderGDP_negrecip*, *residualsgenderGDP_posrecip*, *residualsgenderGDP_risktaking*, *residualsgenderGDP_patience* are built performing a robust linear regression of the gender coefficient of the specific preference and the log GDP p/c, and then calculating the residuals from it
 - *residualslogAvgGDPpc_GEI_trust* is the variable created from *GenderIndex* residualised using *logAvgGDPpc*, and *residualsGenderIndex_GDP_trust* is the variable created from *logAvgGDPpc* residualised using *GenderIndex*, selecting the data set on the specific preference “trust”. Similarly for all the other 10 variables referring to the preferences “altruism”, “posrecip”, “negrecip”, “risktaking”, and “patience”.
 - *residualslogAvgGDPpc_WEF_trust* is the variable created from *ScoreWEFStd* residualised using *logAvgGDPpc*, and *residualsWEF_GDP_trust* is the variable created from *logAvgGDPpc* residualised using *ScoreWEFStd*, selecting the data set on the specific preference “trust”. Similarly for all the other 10 variables referring to the preferences “altruism”, “posrecip”, “negrecip”, “risktaking”, and “patience”.
 - *residualslogAvgGDPpc_UN_trust* is the variable created from *ValueUNStd* residualised using *logAvgGDPpc*, and *residualsUN_GDP_trust* is the variable created from *logAvgGDPpc* residualised using *ValueUNStd*, selecting the data set on the specific preference “trust”. Similarly for all the other 10 variables referring to the preferences “altruism”, “posrecip”, “negrecip”, “risktaking”, and “patience”.
 - *residualslogAvgGDPpc_GDI_trust* is the variable created from *GDIStd* residualised using *logAvgGDPpc*, and *residualsGDI_GDP_trust* is the variable created from *logAvgGDPpc* residualised

using *GDIStd*, selecting the data set on the specific preference “trust”. Similarly for all the other 10 variables referring to the preferences “altruism”, “posrecip”, “negrecip”, “risktaking”, and “patience”.

- *meanGender* and *stdGender* are the variables indicating the average of the gender differences by preferences, and the 95% confidence interval of the gender differences by preferences