Reproduce the results of the article "Relationship of gender differences in preferences to economic development and gender equality"

Sara Cerioli   Andrey Formozov

Highlights

- To be added.
- Code and related infrastructure

Abstract

1. Introduction

This study reproduces the results of the article Falk and Hermle (2018) and partially its supplementary material.

Gender differences are nowadays extensively used as arguments and counter-arguments for decision and policy making, and these differences concerning the economic behaviors (such risk taking, patience, or altruism, for instance) are being studied both in economics and in psychology. *we need a citation here*

One of the problems common for many experiments in social sciences is the lack of large and heterogeneous data-sets that can be used to check for such differences reducing some of the bias induced, for example, by having students or specific sets of people interviewed.

In the Gallup World Poll 2012 there was included a Global Preference Survey conducted on almost 80000 people in 76 countries all around the world, that aimed to fill this gap: Covering almost 90% of the world population representation, with each Country having around 1000 participants answering questions related to their time preference (patience), altruism, will of risk taking, negative and positive reciprocity, and trust.

The data-set provides a unique insight in the economic preferences of a heterogeneous amount of people. The original study published in the Quarterly Journal of Economics (Falk et al. 2018, 133 (4) pp. 1645-1692) focused on more general questions about the economic preferences distributions in different countries, trying to explore different covariates from the Gallup World Poll. While, the subsequent article, replicated in this work, focused specifically on the gender differences arising from the previous study.

The main question that the article wants to answer is the old one nature versus nurture *(underline the resource based factor)*: Are gender differences arising from some kind of biological differences, or from social stereotypes? The first hypothesis means that the differences could potentially be masked by the necessity of fulfilling basic needs for survival reasons, and therefore in less developed Country we would see less gender differences (because people aim for survival first), while in most developed Countries we would see more gender differences (because of the liberation of the women - and men - from basic, granted needs). On the other hand, if it is society that creates those differences, we should see less differences in the more developed Countries, where people are freed from stereotypes and can freely express themselves. The conclusion of the article is that the trends in the data shows a positive correlation of gender differences with GDP p/c of the Countries, and thus "confirming" the first hypothesis (nature is the reason).

The motivation for the replication study came from the wish to apply different analysis beyond the OLS, as for instance multilevel cumulative link models *[cit. link]*, and methods from the Machine Learning toolbox.

In order to achieve that, one needs first to replicate the original analysis since the original code is written by Falk and Hermle (FH) in Stata. *[Check further what's available here https://www.briq-institute.org/global-preferences/downloads#]*

Motivation for reproduction. Other opinions, Open science perspective, cross and meta studies, alternative methodology (beyond OLS, ML).

We would need to add a section for the critics to the Gender Equality Index

2. Methods

Overview

We replicate the results using the R programming language version 4.0.3 (2020-10-10), and its open-source IDE RStudio for an easy access of the code. The following packages with respective version are used:

Package	Version
data.table	1.13.2
bit64	4.0.5
bit	4.0.4
plyr	1.8.6
dplyr	1.0.5
haven	2.4.2
ggplot2	3.3.2
missMDA	1.18

Data Collection, Cleaning, and Standardization

The data used by the authors is not fully available because of two reasons:

1. **Data paywall:** Some sociodemographic variables (for instance, education level or income quintile) are not part of the Global Preference Survey, but of the Gallup World Poll data-set. Check the website of the briq - Institute on Behavior & Inequality for more information on it.

2. **Data used in study is not available online:** This is what happened for the log GDP p/c calculated in 2005 US dollars (which is not directly available online). We decided to calculate the log GDP p/c in 2010 US dollars because it was easily available, which should not change the main findings of the article.

An additional issue that we faced while trying to reproduce the results of the article has been the missing data. We will treat this specific issue later on because it requires a bit of background.

The procedure for cleaning is described for each data-set in the corresponding section below. After manually cleaning the data-set, we standardized the names of the countries and merged the data-sets into one within the function PrepareData.r.

Global Preferences Survey

This data is protected by copyright and can't be given to third parties.

To download the GPS dataset, go to the website of the Global Preferences Survey in the section "downloads". There, choose the "Dataset" form and after filling it, we can download the dataset.

Hint: The organisation can be also "private".

GDP per capita

From the website of the World Bank, one can access the data about the GDP per capita on a certain set of years. We took the GDP per capita (constant 2010 US$), made an average of the data from 2003 until 2012 for all the available countries, and matched the names of the countries with the ones from the GPS data-set.

Gender Equality Index

Check the many updates since then! https://eige.europa.eu/gender-equality-index/2020/SE

The Gender Equality Index is composed of four main data-sets. Here below we describe where to get them (as originally sourced by the authors) and how we treated the data within them, if needed.

- **Time since women's suffrage:** Taken from the Inter-Parliamentary Union Website. We prepared the data in the following way. For several countries more than one date where provided (for example, the right to be elected and the right to vote). We use the last date when both vote and stand for election right were granted, with no other restrictions commented. Some countries were colonies or within union of the countries (for instance, Kazakhstan in Soviet Union). For these countries, the rights to vote and be elected might be technically granted two times within union and as independent state. In this case we kept the first date. It was difficult to decide on South Africa because its history shows the racism part very entangled with women's rights [citation]. We kept the latest date when also Black women could vote. For Nigeria, considered the distinctions between North and South, we decided to keep only the North data because, again, it was showing the completeness of the country and it was the last date. Note: USA data doesn't take into account that also up to 1964 black women couldn't vote (in general, Blacks couldn't vote up to that year). We didn't keep this date, because it was not explicitly mentioned in the original dataset. This can be seen as in contrast with other choices made though.

- **UN Gender Inequality Index:** Taken from the Human Development Report 2015. We kept only the table called "Gender Inequality Index".

- **WEF Global Gender Gap:** WEF Global Gender Gap Index Taken from the World Economic Forum Global Gender Gap Report 2015. For countries where data was missing, data was added from the World Economic Forum Global Gender Gap Report 2006. NOTE: We modified some of the country names directly on the csv file, that is why we provide this as an input file.

- **Ratio of female and male labour force participation:** Average International Labour Organization estimates from 2003 to 2012 taken from the World Bank database (http://data.worldbank.org/indicat or/SL.TLF.CACT.FM.ZS). Values were inverted to create an index of equality. We took the average for the period between 2004 and 2013.

Main issue About Missing Data

During the reproduction of the article, we found that the authors didn't write in details how they handled missing data in the indicators.

They mention on page 14 of the Supplementary Material, that (quoting): "For countries where data were missing data were added from the World Economic Forum Global Gender Gap Report 2006 (http://www3.weforum.org/docs/WEF_GenderGap_Report_2006.pdf)."

However, there are two problems here:

- Regarding the year when women received the right to vote in a specific country. The missing values are the ones coming from the United Arab Emirates and Saudi Arabia, that neither in 2006 (when the WEF Global Gender Gap Report that the authors quote as a reference for the missing values) nor now (in 2021) have guaranteed the right to vote for women yet.

- There are missing data also in the other sources that the authors quote. So a quick search for the missing countries of the WEF report of 2015, shows us that these countries can't be found in the report of 2006 either.

These two unclear points, even though in our understanding not crucial for the replication of the analysis, are not desirable, but couldn't be further clarified with the authors. *[Try to ping again the authors]*

The problem of missing data for a given countries often does not influence much the overall trends of found correlations, however, very relevant if one would to see the implecations of the study with respect to one cirtain coutry of interest.

Data Analysis

The article uses several methods commonly accepted in the field:

- Linear regression for each Country for each preference to extract the gender coefficient as a measure of the gender differences.

- Principal Component Analysis on 6 gender coefficients to summarize an overall measure of the gender differences, and 4 gender equality indexes of the countries to summarize an overall Gender Equality Index.

- Variable Conditioning to separate further between economic development and gender equality in the country.

Linear Model on Each Country for Each Preference

Starting from the complete dataset (meaning with removed NA rows), we wanted to reproduce the data plotted in Fig. S2. regarding the gender differences and economic development by preference and by country.

As already mentioned in the previous paragraph, part of the data to reproduce the article is under restricted access: education level and household income quintile on the individual level are not available in open access version.

As the FH article addresses the gender differences, the main focus is on that individual variable and all the others provided in the dataset (education level, income quitile, age, and subjective math skills) are taken as control variables, meaning that the presence of these variables may not affect the result of the correlation.

In FH *(supplementary material link)* they check that the role of these variables to any extent negligible in the overall correlation and we therefore decided to continue the analysis without using the two variables education level and household income quintile for building the linear models.

Note that the code can be easily modified to include this variable once the full dataset is available.

The linear model for each country is created using the equation:

$p_i = \beta_1^c female_i + \beta_2^c age_i + \beta_3^c age_i^2 + \beta_4^c subjective\ math\ skills_i + \epsilon_i$

This resulted in 6 models – one for each preference measure, p_i – having intercept and 4 weights, each of the weight being related to the variable in the formula above. The weight for the dummy variable "female", β_1^c, is used as a measure of the country-level gender difference. Therefore, in total, we have 6 weights that represent the preference difference related to the gender for 76 countries.

Principal Component Analysis

To summarise the average gender difference among the six economic preferences, we performed a principal component analysis (PCA) on the gender coefficients from the linear models. The PCA is a dimensionality reduction technique which allows to "reshape" the 6 coefficients into other mixed components that maximise the variance. The first component of the PCA has then been used as a summary index of average gender differences in preferences.

We performed a PCA also on the four datasets used for Gender Equality, to extract a summarised Gender Equality Index (more on the structure of the constacted Equality Index in section).

Variable Conditioning

[add the theorem] Frisch–Waugh–Lovell theorem? https://bookdown.org/ts_robinson1994/10_fundamental_theorems_for_econometrics/frisch.html

To separate the effects of the economic development and the gender equality, a conditional analysis was performed. To generalise, if one wants to estimate the correlation of x and y conditioning on z, one needs to perform a double linear regression:

- First, regressing x on z and extracting the residuals

- Second, regressing y on z, and extracting the residuals.

In the end, one needs to take the so calculated residuals of x on z and of y on z, and make a last regression to calculate the correlation.

In practice, if we are interested in checking the influence of the economic development on the summarised gender differences, conditioning on the gender equality, we would need to regress the economic development on the gender equality index, then the average gender differences regressed on the gender equality index, and finally regress the residuals of the average gender differences on the residuals of the economic development.

3. Comparison to the Original Article

In this section, we describe how to reproduce the plots and compare the results in terms of z-scores. Note that there are additional plots produced together with the Supplementary Material: We reproduce them but we will not enter here into the details of the work. *[Do we need to have a Supplementary Material section?]*

Reproducing the Plots of the Main Article

To reproduce the plot of Fig. 1A, we grouped the countries in quartiles based on the logarithm of their average GDP p/c, extracted the mean of each preference from the gender coefficients (the β_1^c) of the countries for each quartile, after standardizing them. The same method was applied to the Gender Equality Index in correlation to the gender differences for each economic preference, to reproduce the plot in Fig. 1C.

Then, we related the magnitude of the summarised gender difference coefficients (the first component of the PCA) with the logarithm of the average GDP per capita to see the effect of the economic development. This reproduced Fig. 1B of the original article. We used a linear model to fit the correlation and extract the p-value, and for the plot the variables on the y-axis were additionally transformed as (y-y_min)/(y_max-y_min). We applied the same method to extract the correlation between the Gender Equality Index and the summarised gender preference, to see the effect of the gender equality in the countries (Fig. 1D). Note that here also the Gender Equality Index is transformed to be on a scale between 0 and 1.

We finally reproduced the plots in Fig. 2A-F using the variable conditioning analysis. This has been done for the economic development, for the Gender Equality Index, and for each of the four indicators building the Gender Equality Index. The variable used on the y-axis is the first Principal Component of the PCA made on the gender differences on the six preferences. Using the residuals, built as described in the Data Analysis section of the Method paragraph, we performed a linear regression on the data points, and we extracted correlation coefficients and p-values.

Tables and z-scores

The comparison to the original article has been done using the z-scores on the correlation coefficients, and checking if the statistical significance was at the same level. *[Do we need to explain why and what the z-scores? I don't think so]*

Here below we report the tables with the corresponding values of the correlation for the original article, our replication study, and the z-scores calculated from them. The sample size of the data, needed for the calculation of the z-scores, is always 76 (the number of the countries involved in this studies), except when

using the Gender Equality Index, where due to the missing data, the number of countries in the sample was reduced to 68. *[Check here the sample size of the FH and the number of countries used in the other gender equality indeces!]* We also indicate the significance level for each correlation using the following scheme:

Signifincance $\leq$ 0.001 (***), $\leq$ 0.01 (**), $\leq$ 0.05 (*)

Table 1: Correlations between Log GDP p/c and country-level gender differences

Variable	Corr. original article	Corr. this analysis	z-score
Altruism	0.58***	0.6***	-0.19
Trust	0.59***	0.53***	0.53
Positive Reciprocity	0.31***	0.29*	0.13
Positive Reciprocity	0.35***	0.42***	-0.50
Risk Taking	0.37***	0.33***	0.28
Patience	0.38***	0.49***	-0.82

Table 2: Correlations between Gender Equality Index and country-level gender differences

Variable	Corr. original article	Corr. this analysis	z-score
Altruism	0.51***	0.49***	0.16
Trust	0.41***	0.47***	-0.44
Positive Reciprocity	0.13	0.20	-0.42
Positive Reciprocity	0.40***	0.30**	0.67
Risk Taking	0.34***	0.26*	0.56
Patience	0.43***	0.48***	-0.37

Table 3: Correlation between Log GDP p/c and Gender Equality Index, and summarised gender differences

Variable	Corr. original article	Corr. this analysis	z-score
Log GDP p/c	0.67***	0.71***	-0.46
Gender Equality Index	0.56***	0.59***	-0.26

Table 4: Conditional analysis to separate the impacts of economic development and gender equality on gender differences

Variable	Residualized on	Corr. original article	Corr. this analysis	z-score
Log GDP p/c	Gender Equality Index	0.53***	0.56***	-0.24
Gender Equality Index	Log GDP p/c	0.32***	0.3***	0.13
WEF Global Gender Gap	Log GDP p/c	0.23***	0.26*	-0.18
UN Gender Equality Index	Log GDP p/c	0.29	0.17	0.72
F/M in Labor Force Participation	Log GDP p/c	0.25*	0.21	0.24
Years since Women Suffrage	Log GDP p/c	0.30**	0.19	0.67

4. Discussion of the Results

Comparing the results of our analysis to the one from the original paper, starting with the single preferences correlations to the economic development, we see that our analysis brings us to very similar results in terms

of correlation coefficients (see Table 1). The p-values, although different, are all indicating a statistically significant correlation, as in the original paper, and when calculating the z-scores thanks to Fisher's r to z transformation, we see that each one is below 2 (which is usually taken as threshold to be statistically significant). This means that our correlations were not statistically significantly different from the ones in the original article.

We do the same for the Gender Equality Index (Table 2), and again we don't find any large difference in the correlation. The p-values tend to be different with respect to the original article, but all of them are pointing towards the same direction. Also the z-scores are in absolute values below 2.

In Table 3, we compare the two core concepts of the article, where the summarised gender differences are regressed on the log GDP p/c and on the Gender Equality Index. The correlations found are similarly positive, strong, and statistically significant. Again, the correlations found in our analysis are not statistically significantly different from the correlations found in the original article.

Lastly, we have the conditional analysis (Table 4). For the two main country-level variable, we see that the values tend to agree and be on the same direction (similar r coefficients, significant p-value, and low z-score). But when we start to check for the single indeces, we see that there are some differences which are worthy to discuss.

The first thing to say is that we had to make choices on how to impute data and also how to handle the missing data (see discussion above in paragraph "Methods"). The main imputation on missing data has been done on the "time since women's suffrage" data-set", that is where we see a substantial difference in the results. Other data-sets, on the other hand, has not been treated for missing data but still they present some difference. For instance, the data-set"F/M in Labor Force Participation" in our analysis has a non-statistically significant correlation, while in the original paper they found a correlation with p-value less than 0.05.

A first thought was that this might be the result of using a different data-set for the GDP (the 2010 USD instead of 2005), but in our opinion this can't be an explanation but rather a check about how robust the results are. So this question about the differences that were found is kept open.

Critics

- Gender Equality Index robustness and validity
- What about cross-calibration for personal interview vs telephone interview? Was it verified that the results for economic preferences are the same in both cases? Where? I would expect it different. There are should be some info in literature about that (XiaoChi Zhang, 2017).
- Cohen's effect size?
- Not raw data available (and in general not all data available)

5. Conclusions

The study indicates that higher economic development and higher gender equality are associated with an increase in the gender differences in preferences, and therefore rules out the social-role theory over the post-materialistic one: When more resources are available to both men and women, the expression of the gender specific preferences can be seen. Our replication leads to the same conclusions, but we have some open questions regarding unexplained differences that might lead to further checks on the results' robustness.

References

Falk, A., A. Becker, T. Dohmen, B. Enke, D. Huffman, and U. Sunde. 2018. "Global Evidence on Economic Preferences."

Falk, A., and J. Hermle. 2018. "Relationship of Gender Differences in Preferences to Economic Development and Gender Equality." https://doi.org/10.1126/science.aas9899.