Gender differences in economic preferences and gender equality are yet unrelated: a replication of Falk and Hermle (*Science*, 2018)

Sara Cerioli[1] and Andrey Formozov[2]

[1]Independent researcher, Hamburg, Germany
[2]Research Group Synaptic Wiring and Information Processing, Center for Molecular Neurobiology Hamburg, University Medical Center Hamburg-Eppendorf, Hamburg, Germany
Correspondence: sara.cerioli@outlook.com, formozoff@gmail.com

Abstract

This study replicates and extends the key components of the analysis of Falk & Hermle (2018a), where the gender differences in economic preferences (defined as time preference, altruism, willingness to take risk, negative and positive reciprocity, and trust) were related to economic development and gender equality of the countries. During the replication, we identified and investigated some potential issues in the original analysis methodology, particularly with the construction of the joint Gender Equality Index as a measure of gender equality in a country. Moreover, we revealed a strong and statistically significant correlation between economic development and several gender equality indexes, only briefly mentioned on a qualitative level in the original paper. This correlation implies that only conditional regression with control on economic development may uncover the impact of gender equality on gender differences in economic preferences. It goes against the results presented in the original article, where simple, unconditional correlations were retained as the main quantitative finding, presented in the graphical abstract. We conducted a further conditional regression analysis using as gender equality index the Global Gender Gap Index from the WEF, the Gender Inequality Index and the Gender Development Index from the UNDP. When analyzing the data using these indicators, we confirmed the strong and statistically significant correlation of gender differences in economic preferences with economic development conditioned on gender equality, as found in the original analysis. The correlation between gender differences in economic preferences and gender equality, conditioning on economic development, has been done using three indexes for gender equality. For one index, the correlation is weak and statistically significant at 5% level, while for the other two indexes the correlation is not statistically significant. A more detailed investigation of single preferences further confirmed this conclusion, demonstrating mild and statistically significant correlation for only two preferences out of 18 preference-gender equality index combinations. Our findings suggest the absence of strong evidence of a correlation between gender differences in economic preferences and gender equality in the country, given the data available, current methodology, and established indexes for measuring gender equality.

JEL: C19 - Econometrics and Statistical Methods: Other, D010 - Microeconomic Behavior: Underlying Principles, C91- Laboratory, Individual Behavior, D630 - Equity, Justice, Inequality, and Other Normative Criteria and Measurement, D64 - Altruism, D810 - Criteria for Decision-Making under Risk and Uncertainty, D910 - Micro-Based Behavioral Economics: Role and Effects of Psychological, Emotional, Social, and Cognitive Factors on Decision Making, F000 International Economics: General

Keywords: replication study, gender differences, economic preferences, cross-country variation

Data availability: The code for this analysis can be found on GitHub at https://github.com/scerioli/Global-Preferences-Survey

1. Introduction

Gender differences in human traits, such, for instance, happiness (Schneider *et al.* 2012), competition (Croson & Gneezy 2009; Gneezy *et al.* 2009; Klonner *et al.* 2021), or work preferences (Beblo & Görges 2018), have been studied in sociology, psychology, and economics for many decades. Published findings on this topic and their relation to gender inequality are frequently used to influence decisions and policy-making, both in the public and private sectors. In turn, gender inequality topics are becoming a more integral part of the agenda for many institutions and organizations, and it is essential for the stakeholders to reveal, estimate, monitor, and prevent gender inequalities on the individual, group, and country levels.

The study of behavioral gender differences on a world scale is challenging. One challenge that hampers the progress is the lack of large and homogeneous data sets across different social groups and countries. The Gallup World Poll 2012 included a Global Preference Survey conducted on almost 80000 people in 76 countries around the world that aimed to fill this gap, covering nearly 90% of the world population representation, with each country having around 1000 participants answering questions related to their time preference (patience), altruism, willingness to take risk, negative and positive reciprocity, and trust. The data set provides a unique insight into the economic preferences of a heterogeneous group of people.

The article that presented a first analysis of this data set was published in the Quarterly Journal of Economics (Falk *et al.* 2018). It focused on general questions about the economic preferences distributions in different countries, exploring several covariates from the Gallup World Poll. The subsequent article (Falk & Hermle 2018a), replicated in this work, focused explicitly on the gender differences highlighted in the previous study and reported evidence for the relationships among gender differences in economic preferences with respect to economic development and gender equality across many countries. The authors proposed two competing hypotheses to be tested. The first one was that the gender differences in economic preferences will decrease for more economically developed and gender-equal countries because social roles related to gender are attenuated. The second hypothesis was, on the contrary, that for more economically developed and gender-equal countries, the gender differences in economic preferences will increase because the gender-neutral goal of subsistence is removed and thus people can pursue their more unconstrained set of preferences. Their analysis showed a positive correlation between gender differences in preferences and economic development (expressed as Log GDP p/c), as well as a positive correlation between gender differences in preferences and gender equality of the countries. Therefore, the authors' conclusion favored the second hypothesis, predicting an increase in the differences as women and men obtain sufficient access to the resources to develop and express their intrinsic preferences independently.

However, an essential relation between the economic development and gender equality indexes was only mentioned in the original text without detailed investigation. At the same time, this relation plays a pivotal role in interpreting the resulting correlations. Another important aspect is the problem of estimation of the gender equality on county level per se. The authors approached this problem composing a custom joint measure of gender equality using four indicators: two officially recognized indexes (the World Economic Forum Global Gender Gap Index, and the United Nations Development Programme Gender Inequality Index) together with two other rather controversial indicators. Doing this, they provide little discussion about the validity and robustness of these supplementary indicators, without establishing a link to the existing literature with this regard. The present work aims to provide a companion set of conclusions based on the indicators validated by recognized institutions of the World Economic Forum (WEF) and the United Nations Development Programme (UNDP). We wish to raise awareness about the potential ramifications of drawing the wrong conclusion about the relationship between gender differences in economic preferences and gender equality when crafting public, corporate, or NGO policy.

This article is organized as follows. In the Section 2, we give a brief summary of the original article. In Section 3, we conduct a replication of the reviewed analysis to extract gender differences in economic preferences from the available data sets and to correlate them to the economic development and the gender equality of the countries. Section 4 shows how strong and statistically significant the correlation between economic development and gender equality of the countries is, and addresses the issues of using custom indexes for the gender equality measurement. After this, we demonstrate that findings on the relationship between gender differences in economic preferences and gender equality lack evidence and do not hold when the conditional

analysis is carried out.

The code used to perform the analysis, the input, and the output data are publicly available (or referenced to be downloaded) at https://github.com/scerioli/Global-Preferences-Survey.

2. Summary of the original article

In this section, we summarize the analysis and main findings of the original article. The authors used the Gallup World Poll 2012 Global Preference Survey to measure the gender differences in economic preferences across 76 countries. The economic preferences are defined as time preference (also referred as patience in the study), altruism, willingness to take risk, negative and positive reciprocity, and trust.

The people participating at the survey were asked to answer qualitative and quantitative questions and their score on each preference was assigned based on a weighted mean of the answers given (for more details, we refer to the original study Falk & Hermle (2018a), section "Extended Materials and Methods"). Therefore, for each person in the data set, a score in each of the six economic preferences was given. For every person, moreover, additional variables indicating their age, sex, education level, subjective math skills (as a proxy for cognitive skills), and household income quintile were given. The authors summarized the differences in all six economic preferences through the technique of the Principle Component Analysis (PCA), using the first component as a joint index.

The study focused on the relationship between the gender differences in the economic preferences and the economic development of the countries, using Log GDP p/c as a proxy. The study also explores the relationship between gender differences in economic preferences and gender equality of the countries. For this, the authors decided to use four different indexes of gender equality: the World Economic Forum Gender Global Gap Index (WEF GGGI), the United Nations Development Programme Gender Inequality Index (UNDP GII), the ratio of female and male labor force participation (F/M LFP), taken from the World Bank database, and the time since women suffrage (TSWS), from the Inter-Parliamentary Union Website. To summarize the effect of the gender equality of the countries using a single indicator, the choice of the authors has been to use the PCA on the four above-listed indexes and to take the first component as a custom Gender Equality Index.

The study reported a large and statistically significant correlation between gender differences in economic preferences and Log GDP p/c ($r = 0.67$, p-value < 0.0001), and between gender differences in economic preferences and Gender Equality Index ($r = 0.56$, p-value < 0.0001), reflected in the *Research article summary* and in the graphical abstract of the original study. The authors also conducted a conditional analysis to isolate the impact of economic development and gender equality of the countries. The correlation was found to be of a slightly smaller magnitude but still rather strong and statistically significant ($r = 0.53$, p-value < 0.0001) when gender differences were related to Log GDP p/c conditioned on the Gender Equality Index, and moderate and statistically significant ($r = 0.32$, p-value $= 0.003$) when relating with the Gender Equality Index and conditioning on Log GDP p/c.

The authors concluded that the evidence indicates that higher levels of economic development and gender equality favor the manifestation of gender differences in preferences across countries, *highlighting the critical role of availability of material and social resources, as well as gender-equal access to these resources, in facilitating the independent formation and expression of gender-specific preferences* (Falk & Hermle 2018a).

3. Replication of the original analysis

In this section, we describe the methodology used to replicate the original article analysis and compare our results to the original authors' ones. Additionally, we make use of the robust linear regression model on the same data to take into account the non-normality of the data set. We didn't find any substantial difference from the original authors' results.

Data

To conduct the replication, we downloaded the Gallup World Poll Global Preferences Survey data set from the briq - Institute on Behavior & Inequality. The full data set is under restricted access, and education level and household income quintile on the individual level of the participants are not available in the open-access version (for more information, see Appendix, Section 2, "Data Collection, Cleaning, and Standardization"). Falk & Hermle (2018b) provide a complementary analysis where all the independent variables (except the gender) were dropped, and the results were coherent with what found in their main analysis. Therefore, we continued the reproduction without having access to education level and the income quintile.

Methods and Results

To assess the gender differences in each economic preference, the following model was used:

$$\text{preference}_i^c = \beta_1^c \text{female}_i + \beta_2^c \text{age}_i + \beta_3^c \text{age}_i^2 + \beta_4^c \text{subjectiveMathSkills}_i + \epsilon_i$$

where the subscript i is the index of a survey participant and c is the index for the country. This results in 6 models – one for each economic preference – with 4 coefficients, each coefficient being related to an independent variable. The coefficient for the dummy variable *female*, β_1^c, is used as a measure of the gender difference. The linear regression has been conducted independently on each country. Therefore, in total, there are 6 coefficients representing the economic preference differences related to gender for 76 countries.

To summarize the gender differences among the six economic preferences, a principal component analysis (PCA) is performed on the gender coefficients. The PCA is a dimensionality-reduction technique that allows a reshaping of the 6 coefficients into orthogonal components that maximize the sample variance. The first component of the PCA has then been used as a summary index of gender differences in preferences. In their article, the original authors refer to this summarized index as "average gender differences". We find this nomenclature potentially confusing, and therefore we keep referring to it either as joint index, or as summarized index, rather than average.

The PCA technique has also been applied on the four gender equality indexes to get a joint index that the authors called "Gender Equality Index" (GEI), as already described in the section 2 of this paper.

The summary index of gender differences in economic preferences has been then regressed on Log GDP p/c to obtain an estimation of the correlation with the economic development, finding a correlation coefficient r = 0.6830 with p-value < 0.001 (see Table 1). The summary index of gender differences has been also regressed on the Gender Equality Index, obtaining a correlation coefficient of r = 0.6079 and p-value < 0.001. To have an estimation of the similarity of the results, we compute the z-scores between our correlation coefficients and the original authors' ones. The z-scores couldn't exclude any presence of statistically significant difference between the results.

Table 1: Correlation between PCA-summarized gender differences in economic preferences vs Log GDP p/c and joint Gender Equality Index. Significance ≤ 0.001 (***), ≤ 0.01 (**), ≤ 0.05 (*)

	Original	Replication	*z-score*
$\text{corr}_{EconomDevelop}$	0.6685***	0.6830***	-0.161
$\text{corr}_{GenderEquality}$	0.5580***	0.6079***	-0.449

In the original article, the authors performed an ordinary least squares (OLS) regression on residual plots to study the correlation of gender differences in economic preferences with economic development, separating the contributions of gender equality in the country, and vice versa. The theorem from Frisch–Waugh–Lovell (Frisch & Waugh 1933; Lovell 1963) guarantees that the coefficients found from this residual analysis are the same as those found for multiple regression of the gender differences on both economic development and gender equality index of the countries (as done in the formula above). Therefore, alternatively, to study the effect of both Log GDP p/c and gender equality, one could incorporate all three variables into a multiple

linear regression model with both factors as explanatory variables to study the effect of both Log GDP p/c and gender equality:

$$\text{SummarizedGenderDiff}_c = \beta_{EconomDevelop}\ \text{EconomDevelop}_c + \beta_{GenderEquality}\ \text{GenderEquality}_c + \epsilon_c$$

where c indicates the country-level. This results in 5 models, one for each gender equality indicator (one that is the summarized Gender Equality Index - GEI, and the four single indicators for gender equality - WEF GGGI, UNDP GII, F/M LFP, and TSWS).

The original authors have regressed the summarized gender differences in economic preferences on the Log GDP p/c conditioning on the Gender Equality Index, in order to estimate the effect of solely the economic development. Then, they did the same but conditioning on the Log GDP p/c to isolate the effect of gender equality. They also regressed the gender differences in economic preferences on the single gender equality indexes, conditioning on Log GDP p/c, to estimate the effect of the single gender equality indexes without confounders coming from the economic development.

Their results can be found in Fig. 2 A-F of the original article, summarized here in Table 2 in comparison with our analysis. For clarity, we used the name of the coefficients referred in the formula above, making clear which variable where the gender differences have been regressed on, and the ones where they have been conditioned on.

The results found are all in agreement with the original ones, both in terms of order of magnitude of the slope coefficients, and in the p-value. The only model giving an almost two-fold difference in the slope for the gender coefficient is the one related to Time Since Women Suffrage (TSWS). It is not a surprising result to see because this index had many ambiguous data-points, which is hard to cross-check with the original work, which might have generated this difference. We refer to the Appendix, Section "Data Collection, Cleaning, and Standardization", for a more in-detail explanation.

Table 2: Comparison of the conditional analysis results for the original and replicated study. Reported are the **slopes** of the linear regressions and the corresponding p-value. Significance levels: $\leq$ 0.001 (***), $\leq$ 0.01 (**), $\leq$ 0.05 (*).

Coefficient	Regression on	Conditional on	Original	Replication
$\beta_{EconomDevelop}$	Log GDP p/c	GEI	0.5258***	0.5003***
$\beta_{GenderEquality}$	GEI	Log GDP p/c	0.3192**	0.3358***
$\beta_{GenderEquality}$	WEF GGGI	Log GDP p/c	0.2327**	0.2234*
$\beta_{GenderEquality}$	UNDP GII	Log GDP p/c	0.2911	0.3180
$\beta_{GenderEquality}$	F/M LFP	Log GDP p/c	0.2453*	0.2206*
$\beta_{GenderEquality}$	TSWS	Log GDP p/c	0.2988**	0.1879*

Robust Linear Regression

A simple linear regression is conducted when the variable that one wants to predict is a continuous one. For categorical, and especially ordered categorical variables, linear regression is not the best model choice. One reason among the others is that the distance between two adjacent categories is unknown (Agresti (2010), Greene (2003), Liao (1994), Williams (2016)).

Within the Global Preference Survey, the way economic preferences has been measured is mixed between qualitative and quantitative responses: For all the preferences, a qualitative question about each own's level of economic preference has been measured with a Likert scale between 0 and 10, while a quantitative measurement has been performed on all the preferences but trust (please refer to Falk & Hermle (2018b) for further details). This mixed approach of semi-continuous and ordered categorical variables has lead us questioning the appropriateness of the OLS method on the data.

A diagnostic test on the data for each preference and each country, carried out using a Shapiro-Wilk test, indicated the presence of non-normality for all the measured economic preferences. In all cases, the distribution of the data has been detected to be non-normally distributed.

Based on this outcome, we ran the previous analysis using the robust linear regression instead of ordinary linear regression, to mitigate potential downstream biases. The results obtained with the robust linear regression didn't differ significantly from the original and the replication analysis (see Table 3 and 4 below), but we kept using this method for the extended analysis reported below as well.

Table 3: Correlation between PCA-summarized gender differences in economic preferences vs Log GDP p/c and aggregated Gender Equality Index using the robust linear regression (RLR) for the replication study. Significance ≤ 0.001 (***), ≤ 0.01 (**), ≤ 0.05 (*)

		Log GDP p/c	Gender Equality Index
Original (OLS)		0.6685***	0.5580***
Replication (RLR)		0.6733***	0.5905***
	z-score	-0.053	-0.288

Table 4: Comparison of the conditional analysis results for the original study (where OLS was used) and our replication using the robust linear regression (RLR). Reported are the **slopes** of the linear regressions and the corresponding p-value. Significance levels: ≤ 0.001 (***), ≤ 0.01 (**), ≤ 0.05 (*).

Coefficient	Regression on	Conditional on	Original (OLS)	Replication (RLR)
$\beta_{EconomDevelop}$	Log GDP p/c	GEI	0.5258***	0.4862***
$\beta_{GenderEquality}$	GEI	Log GDP p/c	0.3192**	0.3432**
$\beta_{GenderEquality}$	WEF GGGI	Log GDP p/c	0.2327**	0.2106*
$\beta_{GenderEquality}$	UNDP GII	Log GDP p/c	0.2911	0.3017
$\beta_{GenderEquality}$	F/M LFP	Log GDP p/c	0.2453*	0.2034*
$\beta_{GenderEquality}$	TSWS	Log GDP p/c	0.2988**	0.1929*

4. Extended analysis

Gender equality indexes and potential issues

During the replication analysis, we have encountered potential issues related to the gender equality indicators that we considered worthy to analyze further.

A first point of concern is related to the reason why the purely custom Gender Equality Index has been taken as a measure for gender equality of the countries, when many internationally recognized, studied, and adopted indicators are already available. The justification for using GEI built from the PCA rather than the widely used indexes was omitted in the original article. To characterize the structure of the GEI, we visualized its composition with the diagram shown in Figure 2. As mentioned above, the authors built this joint measure by using the PCA technique on four gender-equality indicators and taking the first component as a summary index of gender equality. Two of these indicators are indexes officially approved by international organizations, the WEF GGGI, and the UNDP GII; one is the ratio of female and male labor force participation, taken from the World Bank database, a quantity widely used but yet outdated as a representative measure of gender equality. The other is an indicator newly constructed by the authors the time since women suffrage, with

the data taken from the Inter-Parliamentary Union website, presumably to track long-term influences of the guaranteed right to vote as a proxy of gender equality. The WEF GGGI has a total of fourteen sub-indexes, grouped and weighted into four categories: economic participation and opportunity, political empowerment, educational attainment, and health and survival. The UNDP GII follows a similar logic to cover the same categories describing several aspects of human life, but using only five sub-indexes in total: two for health and reproduction-related issues, and three others for the remaining categories.

As one can see in Figure 2, the components of the GEI used in the original study contain repetitions. The two indexes WEF GGGI and UNDP GII share three sub-indexes, here indicated with different colors: ratio of female and male labor force participation (purple), the share of seats in parliament (green), and enrollment into secondary education (blue). As a third variable to construct the GEI, the authors used the ratio of female and male labor force participation, already included in the previous two indexes as a weighted sub-index. While the PCA technique in some cases permits the aggregation of the indexes even in presence of large correlations among the inputs, in the present case, such a procedure may lead to an imbalance in favor of these specific repetitive indexes (especially female and male labor force participation) over other factors, which were already balanced in the design of WEF GGGI and UNDP GII indexes. The other critical point for the use of PCA is then the interpretability of the index, which is a central question when it comes to build an index that can measure differences in the society without loosing its descriptive power and the ability to identify effective policies for closing the gender gap (World Economic Forum 2015).

Looking at the available literature, it is possible to get an idea of how controversial the indicators of gender equality can be. Here, we summarize the main pain points for each of the indexes chosen by the authors. Despite the identified limitations, these indexes remain the most homogeneous and standardized estimates of gender inequality on the world scale. The UNDP GII has been criticized by several authors (Sen & Anand 1996; Permanyer 2011; Klasen 2017). This index is said to be highly related to economic development, as it includes reproductive health indicators that can penalize less-developed countries, and it has a measure of the welfare loss associated to inequality based on a calculated gender equality measure not documented publicly. Instead, the power and limitation of WEF GGGI index remains mostly undiscussed in the academic literature, with a few exceptions (Barns & Preston 2010; Piper 2019). As our investigation reveals (Figure 2), one point of concern could be the inclusion of a subjective measure, based on a best experts' guess, called "wage equality between men and women for similar work". This measure of economic participation and opportunity represents a substantial ~30% of the sub-index. The index is thought to be the least dependent on the economic development of a country since it measures the gap between male and female access to resources and opportunities (World Economic Forum 2015). However, as we will show below in this paragraph, this dependence exists and is not negligible.

Furthermore, the *Time Since Women Suffrage* index was introduced by the authors to track the long-lasting effects of the right to vote. This is based on the assumption that, during the time, development has always had a monotonic effect and its magnitude is proportional to the time since women suffrage was established. The data on the year suffrage is available on a global scale but provides a very limited overview of gender disparities in politics, as discussed in Carmichael *et al.* (2014). Indeed, it can be argued that even after the right to vote has been granted, many discriminating laws may be still present, and the alignment of law together with the executive branch of the government and elimination of discrimination takes more time – for example, despite gaining the right to vote, the right to work can be suppressed for several decades. The assumption that suffrage played a long-lasting effect on the balance in gender equality sounds reasonable but requires further investigation to be used as a robust estimator. It is also worth mentioning the topic of suffrage and race, which for many countries is strongly intertwined (Carruthers & Wanamaker 2015; Yang 2020).

Along with WEF GGGI and UNDP GII, other indexes exist, such as the GDI earlier introduced, that was discussed in Klasen (2017) and recently added to the UNDP report. It captures three dimensions in terms of health, knowledge, and living standards, separately for males and females. This index is defined as the ratio of the Human Development Index for females divided by that for males. The life expectancy, the expected year of schooling and mean years of schooling, and GNI per capita are calculated within these dimensions. We include this index in our extended analysis.

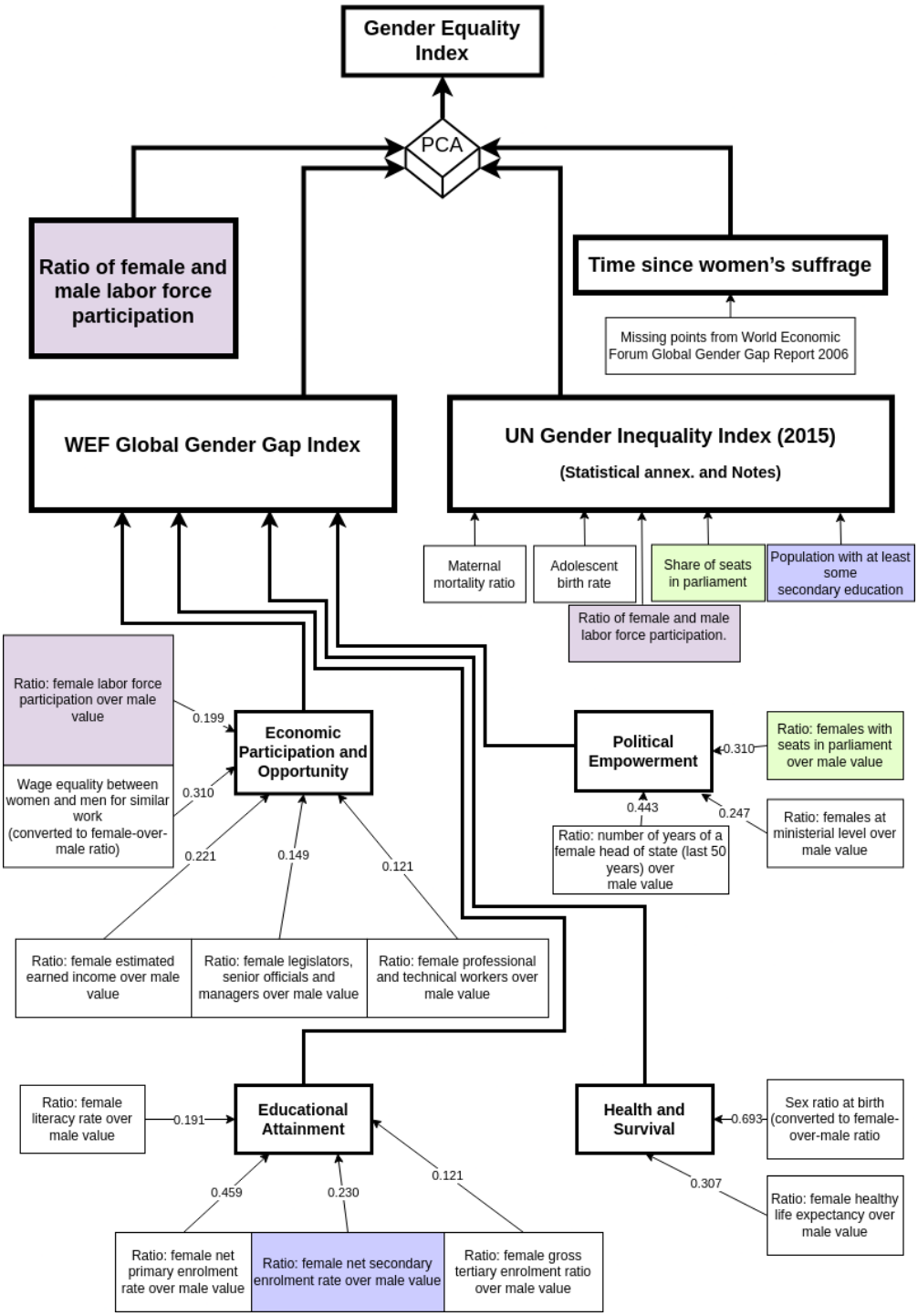

Figure 1: The custom Gender Equality Index decomposed in its sub-indexes, as used by the original authors. The repeated indexes and sub-indexes are highlighted with different colors. Note that for each sub-index the WEF GGGI calculates a weight to balance its impact on the overall index, while in the UNDP GII treats the sub-indexes without extra weighting. See also the technical notes of the World Economic Forum (2015) and of United Nation Development Programme (2021)

As already briefly anticipated, an additional concern rises when studying the correlation between economic development and gender equality index, using linear and not logarithmic scale for economic development expressed in GDP.. The fact that there is a correlation between them is revealed (Duflo 2012) and reported in the World Economic Forum (2015). To check the correlation between economic development and gender equality of the countries analyzed in the study, we regressed the Log GDP p/c on the joint Gender Equality Index used by the authors in the original article. In addition, we explored the correlation of Log GDP p/c with other three indexes for the measure of gender equality: two are indexes already introduced by the original authors, as the WEF GGGI from the World Economic Forum Global Gender Gap Report 2015, and the UNDP GII Human Development Report 2015. The third index to be analyzed in its correlation with Log GDP p/c was the UNDP Gender Development Index (GDI), which is also later used for the extended analysis. The correlation between economic development expressed in Log GDP p/c and gender equality indexes in the countries is very strong (see also Figure 1): for the Gender Equality Index, we found a correlation with r = 0.5440 and a p-value < 0.0001; for WEF GGGI, r = 0.2926 and p-value = 0.013; for UNDP GII, r = 0.8542 and p-value < 0.0001; finally, for the UNDP GDI, a correlation of r = 0.5316 with p-value < 0.0001.

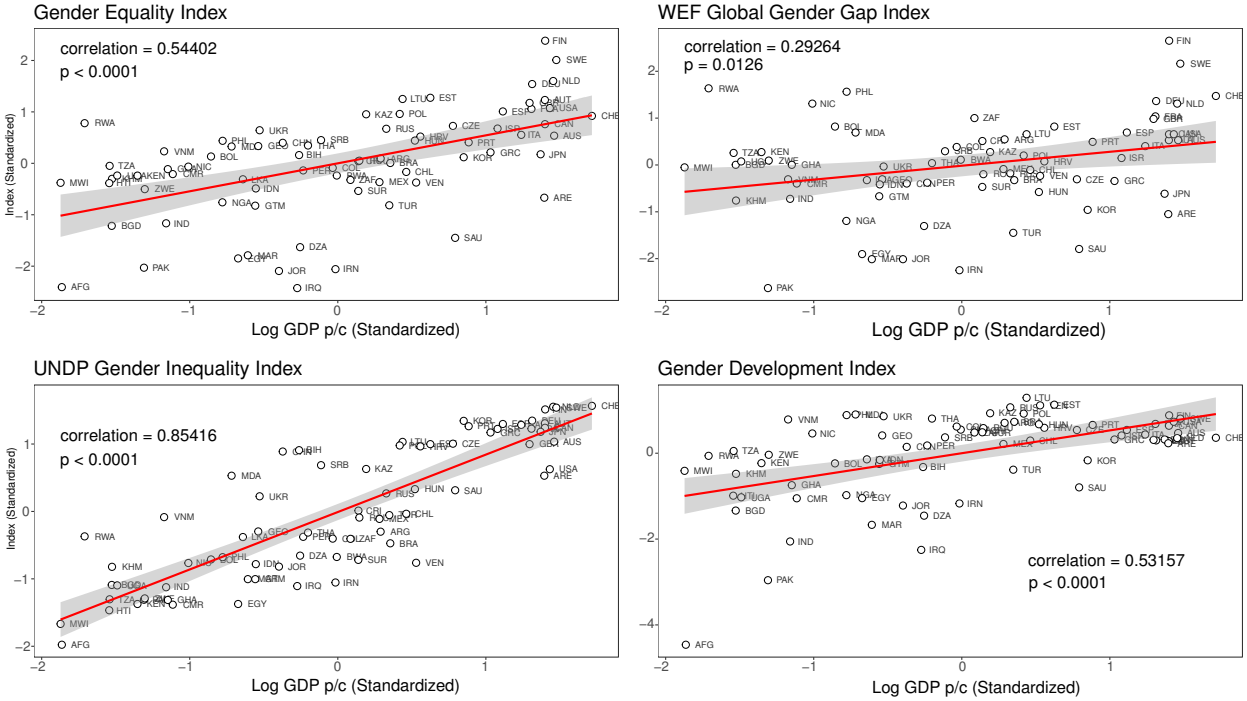

Figure 2: Correlation between gender equality indicators and economic development by country. Note that only the countries participating in the original study are included.

The correlation coefficients indicate a moderate to strong correlation of the economic development with each of the gender equality indexes. This means that any result presented without taking this into account - that is, without presenting a model where the gender equality is conditioned on the economic preference of the country - is misleading, since the correlation terms will be influenced by confounders.

Based on the issues here discussed, we decided to carry out a complementary analysis of the data. The difference with the original authors analysis is that we used only the internationally established (although subjects to critic) indicators WEF GGGI, UNDP GII and the newly introduced GDI, studying their relationship to the gender differences in economic preferences and the economic development. The analysis performed took into account the demonstrated correlated nature of the gender equality of the countries and their economic development by studying the gender differences in economic preferences in a multilinear regression model where economic development and gender equality index were simultaneously included.

Conditional analysis of gender differences and their relationship to economic development and gender equality

In this section, we explore the correlations between gender differences in economic preferences (expressed in PCA-based summarized index as well as single preferences), the economic development and gender equality indicators (WEF GGGI, UNDP GII, and UNDP GDI), using a conditional analysis.

As can be seen in Figure 3, the gender differences have a strong and statistically significant correlation with economic development when the conditional analysis is performed on the single-gender equality indicators, suggesting that the economic development of a country plays a key role in the measured gender differences in economic preferences. On the contrary, the correlation between the gender differences in economic preferences and the gender equality of the country, conditioned on the economic development, is only statistically significant for WEF GGGI conditioned on Log GDP p/c (p-value = 0.0241), while for UNDP GII and GDI the correlation is weak to null, with no statistical significance at the 5% confidence level.

To investigate the role of economic development and gender equality on single preferences, we used a multilinear regression model with each preference, similarly to what have done for the summarized gender differences:

$$\text{preference}_c = \beta_{EconomDevelop}\ \text{EconomDevelop}_c + \beta_{GenderEquality}\ \text{GenderEquality}_c + \epsilon_c$$

where preference_c indicates the 6 economic preferences, and the GenderEquality_c indicates the 3 indicators for gender equality in the countries. In order to extract the correlation, one simply uses the relation between slope and correlation term $\beta = r_{xy} \cdot \frac{\sigma_x}{\sigma_y}$, where β is the slope found in the equation above, σ_x and σ_y are the standard deviations of the X and Y variables, and r_{xy} is the correlation term between the two variables.

As one can see in the following Tables 3-5, the correlation between single preferences and economic development conditioning on gender equality (for simplicity here just written as $r_{LogGDPpc}$) is in most cases strong and statistically significant. On the other hand, when we look at the correlation between single preferences and each gender equality indicator conditioning on economic development ($r_{GenderEquality}$), we see that in 16 out of 18 regressor-regressee pairs, no statistically significant correlations were observed. In two cases, weak but statistically significant correlations were found, between the pairs WEF GGG-Altruism and the UNDP GII-Risk Taking.

Taking into account these results, we conclude that there is a lack of a correlation between the gender differences in single economic preferences and gender equality indexes for the absolute majority of the preferences, while the correlation between gender differences in single economic preferences and the economic development of a country holds.

Table 5: **Gender differences in each economic preference regressed on Log GDP p/c and on WEF GGGI**. For all the Tables, the robust linear regression method is used. The correlation terms and their significance levels ≤ 0.001 (***), ≤ 0.01 (**), ≤ 0.05 (*) are reported.

Preference	$r_{LogGDPpc}$	$r_{WEF-GGGI}$
Trust	0.5174***	0.1325
Altruism	0.5255***	0.3561**
Positive Reciprocity	0.2898*	0.0396
Negative Reciprocity	0.3974**	0.1680
Risk Taking	0.3469**	0.0349
Patience	0.3742**	0.2312

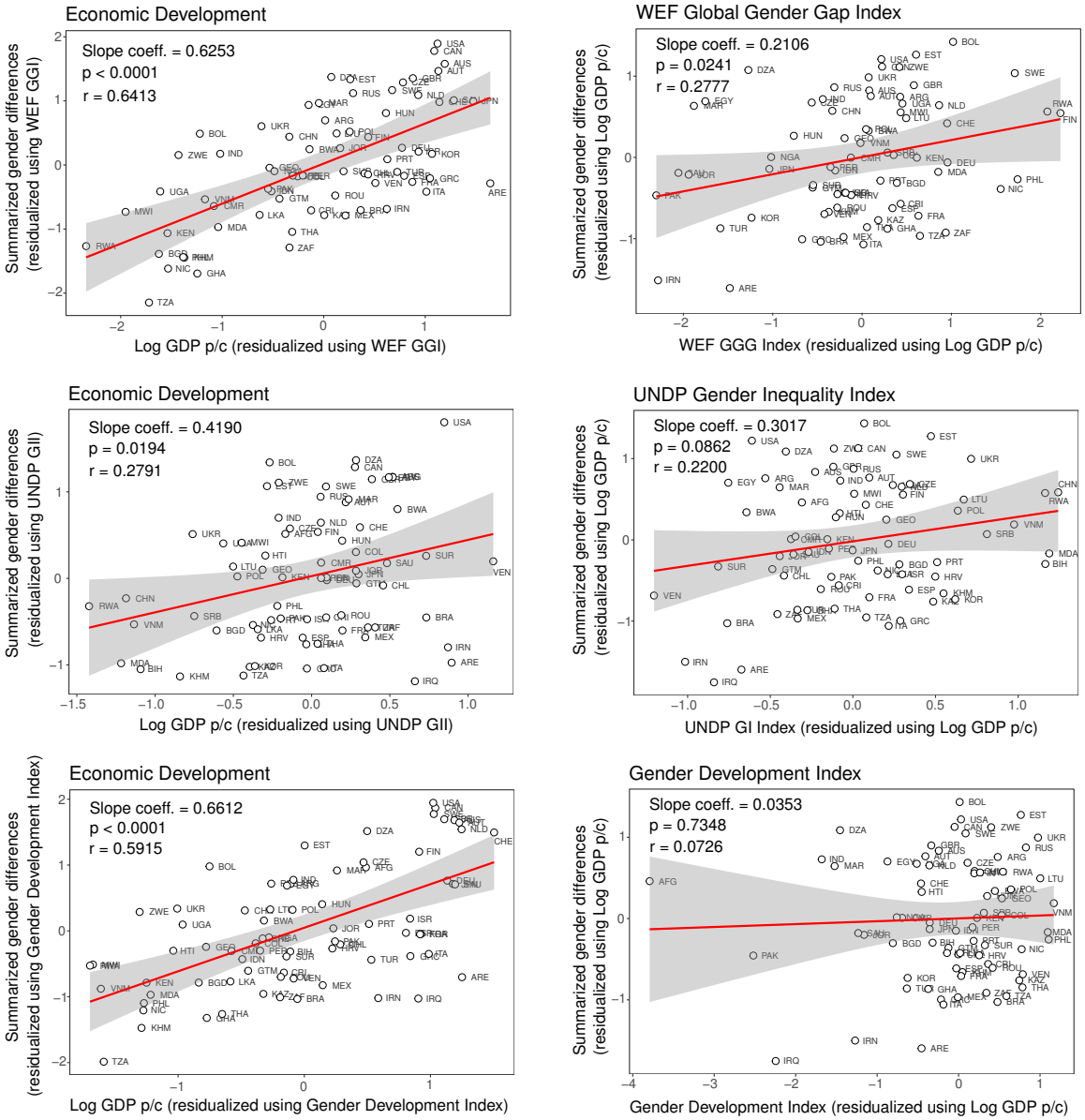

Figure 3: Correlation between gender differences in economic preferences and economic development, and between gender differences in economic preferences and gender equality indexes, using the residuals plots. On the left, gender differences are regressed on economic development conditioned on gender equality for the different indicators (WEF GGGI, UNDP GGI, and GDI). On the right, the corresponding values of gender differences are regressed on gender equality indicator conditioned on economic development.

Table 6: **Gender differences in each economic preference regressed on Log GDP p/c and on UNDP GII**.

Preference	$r_{LogGDPpc}$	$r_{UNDP-GII}$
Trust	0.1930	0.2160
Altruism	0.3527***	-0.0421
Positive Reciprocity	0.2054	-0.0402
Negative Reciprocity	0.0706	0.1742
Risk Taking	0.0262	0.2192*
Patience	0.1046	0.1705

Table 7: **Gender differences in each economic preference (y) regressed on Log GDP p/c and on UNDP GDI**.

Preference	$r_{LogGDPpc}$	$r_{UNDP-GDI}$
Trust	0.5045***	0.0794
Altruism	0.4477***	0.1021
Positive Reciprocity	0.1870	0.1978
Negative Reciprocity	0.3858***	-0.1033
Risk Taking	0.3199**	-0.0261
Patience	0.3337**	0.0812

5. Discussion and conclusions

In the present article, we replicated the results of the work by Falk & Hermle (2018a) that related gender differences in economic preferences to economic development and gender equality. Given the importance of these findings and their impact, and the impressive size and complexity of the analysis, it is of high value to conduct a detailed reconstruction of the study and its methodology, checking the robustness of assumptions and inferences.

As a first milestone, we performed a nearly pure replication, obtaining the gender differences in economic preferences from the Gallup World Poll 2012 Global Preference Survey, using the same methodology as in the original article. Unfortunately, the data set is publicly available only in pre-processed form and partially restricted. Nevertheless, the extracted gender differences were close to the ones from the original article. In addition, we ran the same analysis using robust regression instead of ordinary linear regression as the data revealed signs of non-normality and outliers, but no significant changes in the distribution of gender differences in economic preferences were observed.

We then investigated the indexes used to estimate the gender equality and its relation to economic development. We analyzed the Gender Equality Index built by the authors and its single components. Some methodological issues were identified, and the usage of this custom index over more established, balanced measures lacks justification and remains an open question. Therefore, we conducted our further analysis based on separate, widely-accepted indicators of gender equality (WEF Global Gender Gap Index and UNDP Gender Inequality Index) used in the replicated article, plus an additional indicator, the UNDP Gender Development Index. We also observed a strong correlation between Log GDP p/c and the Gender Equality Index built by the authors ($r = 0.5440$, p-value < 0.0001), and also measured the correlation between Log GDP p/c and the indexes chosen to be investigated in the extended analysis (WEF Global Gender Gap Index having a correlation of $r = 0.2926$ and p-value $= 0.013$; UNDP Gender Inequality Index of $r = 0.8542$, p-value < 0.0001; and UNDP Gender Development Index having $r = 0.5316$ with p-value < 0.0001). In light of these strong correlations, we conclude that only a conditional regression with control on Log GDP p/c may uncover the separate role of gender equality on gender differences in economic preferences. A simple linear regression between the gender

differences in economic preferences and the gender equality of the countries shouldn't be used to test the underlying hypothesis regarding their association.

We then examined gender differences and their relationship to economic development and gender equality using the above-mentioned indexes. Performing a conditional analysis, we found a large and statistically significant correlation between summarized gender differences in economic preferences and economic development, conditioning on WEF GGGI and UNDP GDI (r = 0.6413, p-value < 0.0001, and r = 0.5915, p-value < 0.0001, respectively), while for UNDP GII the correlation was somewhat milder with a larger p-value (r = 0.2791, p-value = 0.0194). On the other hand, when conditioning on economic development, no correlation between UNDP GII or GDI and the summarized gender differences in economic preferences was found. Only the correlation for the WEF GGGI was statistically significant (p-value = 0.0241) but somewhat weaker (r = 0.2777).

We additionally analyzed how single gender differences in economic preferences are related to economic development and gender equality. Along with the results for the summarized gender differences, large and statistically significant correlations were demonstrated when we tested the correlation between gender differences with Log GDP p/c conditioning on WEF, and conditioning on GDI. Interestingly, no preference except *altruism* demonstrated any relation with Log GDP p/c when conditioning on UNDP. Furthermore, among six economic preferences and three gender equality indexes, no statistically significant correlation between economic preferences and Log GDP p/c conditioning on gender equality was found, except for *altruism* (r = 0.3561; p-value = 0.006) conditioning on WEF, and *risk-taking* (r = 0.2192; p-value = 0.045) when conditioning on UNDP, thus providing little support for the presence of a correlation between single gender differences in economic preferences and gender equality.

To sum up, gender equality indexes and economic development expressed in Log GDP p/c are strongly correlated variables; thus, the correlation between gender differences in economic preferences and gender equality indexes should be determined by conditioning on economic development. The conditional analysis suggests that only a mild and sparse correlation between gender differences in economic preference and gender equality exists (only one preference over six), the magnitude of the correlation coefficient being at least two-fold smaller than the correlation of gender differences and the economic development conditioning on gender equality indexes. Given these results, we conclude that the evidence is not sufficient to support a correlation between gender differences in economic preferences and gender equality in a country. The hypothesis of the role of resources needs to be updated taking into consideration that when a country is more economically developed, the gender differences are increasing, but this fact doesn't seem to hold for more gender equal countries. Gender equality doesn't seem to play a role in the gender differences in economic preferences, while economic development seems to influence the measured differences.

References

Agresti, A. (2010). *Analysis of ordinal categorical data.* Wiley.

Barns, A. & Preston, A. (2010). Is Australia really a world leader in closing the gender gap? *Feminist Economics*, **16**, 81–103. Retrieved from https://doi.org/10.1080/13545701.2010.530607

Beblo, M. & Görges, L. (2018). On the nature of nurture. The malleability of gender differences in work preferences. *Journal of Economic Behavior & Organization*, **151**, 19–41. Retrieved from https://www.sciencedirect.com/science/article/pii/S0167268118301343

Carmichael, S., Dili, S. & Rijpma, A. (2014). *Gender inequality since 1820.* Retrieved from https://www.oecd-ilibrary.org/content/component/9789264214262-16-en

Carruthers, C.K. & Wanamaker, M.H. (2015). *Municipal housekeeping: The impact of women's suffrage on public education.* National Bureau of Economic Research. Retrieved from http://www.nber.org/papers/w20864

Croson, R. & Gneezy, U. (2009). Gender differences in preferences. *Journal of Economic Literature*, **47**, 448–74. Retrieved from https://www.aeaweb.org/articles?id=10.1257/jel.47.2.448

Duflo, E. (2012). Women empowerment and economic development. *Journal of Economic Literature*, **50**, 1051–1079. Retrieved from http://www.jstor.org/stable/23644911

Falk, A., Becker, A., Dohmen, T., Enke, B., Huffman, D. & Sunde, U. (2018). Global Evidence on

Economic Preferences*. *The Quarterly Journal of Economics*, **133**, 1645–1692. Retrieved from https://doi.org/10.1093/qje/qjy013

Falk, A. & Hermle, J. (2018a). Relationship of gender differences in preferences to economic development and gender equality. *Science*, **362**, eaas9899. Retrieved from https://www.science.org/doi/abs/10.1126/science.aas9899

Falk, A. & Hermle, J. (2018b). Supplementary materials of Relationship of gender differences in preferences to economic development and gender equality.

Frisch, R. & Waugh, F.V. (1933). Partial time regressions as compared with individual trends. *Econometrica*, **1**, 387–401. Retrieved from http://www.jstor.org/stable/1907330

Gneezy, U., Leonard, K.L. & List, J.A. (2009). Gender differences in competition: Evidence from a matrilineal and a patriarchal society. *Econometrica*, **77**, 1637–1664. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA6690

Greene, W. (2003). *Econometric analysis*, Forthn. Prentice Hall.

Klasen, S. (2017). *UNDP's gender-related measures: Current problems and proposals for fixing them*. Georg-August-Universität Göttingen, Courant Research Centre - Poverty, Equity and Growth (CRC-PEG), Göttingen. Retrieved from http://hdl.handle.net/10419/157265

Klonner, S., Pal, S. & Schwieren, C. (2021). Equality of the sexes and gender differences in competitiveness: Experimental evidence from a traditional society with gender-balanced norms. *SSRN Electronic Journal*.

Liao, T.F. (1994). *Interpreting probability models: Logit, probit, and other generalized linear models*. SAGE Publishing, United States.

Lovell, M.C. (1963). Seasonal adjustment of economic time series and multiple regression analysis. *Journal of the American Statistical Association*, **58**, 993–1010. Retrieved from https://www.tandfonline.com/doi/abs/10.1080/01621459.1963.10480682

Permanyer, I. (2011). Are UNDP indices appropriate to capture gender inequalities in Europe? *Social Indicators Research*, **110**.

Piper, A. (2019). An Investigation into the Reported Closing of the Nicaraguan Gender Gap. *Social Indicators Research: An International and Interdisciplinary Journal for Quality-of-Life Measurement*, **144**, 1391–1413. Retrieved from https://ideas.repec.org/a/spr/soinre/v144y2019i3d10.1007_s11205-019-02080-5.html

Schneider, U., Pfarr, C., Schneider, B. & Ulrich, V. (2012). I feel good! Gender differences and reporting heterogeneity in self-assessed health. *The European journal of health economics : HEPAC : health economics in prevention and care*, **13**, 251–65.

Sen, A. & Anand, S. (1996). Gender inequality in human development: Theories and measurement. *Background papers: Human development report 1995*, pp. 1–20. United Nations Development Programme; United Nations Development Programme, New York.

United Nation Development Programme. (2021). Human Development Report, technical notes. Retrieved from https://hdr.undp.org/sites/default/files/2021-22_HDR/hdr2021-22_technical_notes.pdf

Williams, R. (2016). Understanding and interpreting generalized ordered logit models. *The Journal of Mathematical Sociology*, **40**, 7–20. Retrieved from https://doi.org/10.1080/0022250X.2015.1112384

World Economic Forum. (2015). Global Gender Gap report 2015. Retrieved from https://reports.weforum.org/global-gender-gap-report-2015/the-global-gender-gap-index-2015/

Yang, M. (2020). An intimate dialog between race and gender at Women's Suffrage Centennial. *Humanities and Social Sciences Communications*, **7**.