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. On the contrary, for the correlation between gender differences in economic preferences and gender equality indexes, conditioning on economic development, it was not the case. 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.

1. Introduction

Gender differences in human traits, such, for instance, happiness (Schneider *et al.* 2012), competition (Niederle & Vesterlund 2005; 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, FH 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 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, with a total of almost 80000 people surveyed. 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.

In their article, FH 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.

The study focused on the relationship between the gender differences in the economic preferences (summarized into one single variable through the technique of Principle Component Analysis) 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, and in doing that, four different indexes of gender equality have been used: 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, FH performed a PCA on the four above-listed indexes and took 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 analysis in FH and we compare our results to theirs. 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 did not 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 decided to continue the replication study without having access to education level and the income quintile.

Methods and Results

Following FH analysis (see also the Supplementary Material of FH), we built a multilinear regression model to assess the relationship between each of the six economic preferences and the independent variables associated to the individuals across countries:

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

where the subscript i is the index of a survey participant and c is the index for the country. This results in six models – one for each economic preference – with four coefficients, each coefficient being related to an independent variable. The coefficient for the dummy variable *female*, β_1^c, is then used as a measure of the gender difference. The multilinear regression has been conducted independently on each country. Therefore, in total, there is one coefficient representing gender differences for each of the six economic preferences 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 six 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. FH 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 competing hypotheses in FH are reported here (although in short) for more clarity: *Following social role theory, one may hypothesize that gender differences in preferences attenuate in more developed, gender-egalitarian countries (social role hypothesis). [. . .] As a consequence, according to the social role hypothesis, higher economic development and gender equality (and the associated dissolution of traditional gender roles) should lead to a narrowing of gender differences in preferences. In contrast, there is reason to expect that gender differences in preferences expand with economic development and gender equality (resource hypothesis). [. . .] In sum, greater availability of material and social resources to both women and men may facilitate the independent development and expression of gender-specific preferences, and hence may lead to an expansion of gender differences in more developed and gender-egalitarian countries.*

We can formally write the hypotheses using the following linear model:

$$\text{Gender Differences} \propto \beta_{EconomicDevelopment} \text{ Economic Development} + \beta_{GenderEquality} \text{ Gender Equality}$$

The model above takes into account that FH original hypotheses are set on seeing an increase or a decrease of the gender differences in economic preferences related to both the economic development and the gender equality.

Therefore, we would expect that:

1. If the social role hypothesis is correct, the model above will result into negative coefficients for the economic development and the gender equality, and an overall negative correlation between dependent and independent variables, while

2. If the resource hypothesis is correct, then we will have positive coefficients for the economic development and the gender equality, and an overall positive correlation between gender differences in economic preferences with economic development and gender equality in the country.

Following these statements, the null hypothesis is that there is no correlation between gender differences in economic preferences and the economic development and gender equality of the country. Any other scenario is left out of the original hypotheses and therefore would require additional hypotheses and studies.

This model takes also into account the correlation between the economic development and the gender equality of a country, by allowing the conditioning of one over the other. 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 FH. The correlation found is moderately strong (r = 0.544) and statistically significant (p-value < 0.0001).

In practice, the model above can be rewritten using the proxies for economic development and for gender equality, that is Log GDP p/c and Gender Equality Index (the joint index created by FH), to assess the summarized gender differences in economic preferences:

$$\text{SummarizedGenderDiff}_c = \beta_{LogGDPpc} \text{ Log GDP p/c}_c + \beta_{GEI} \text{ GEI}_c + \epsilon_c$$

where c indicates the country-level. Additionally, FH tested their model on the single gender equality indicators, obtaining in this way four more models corresponding to the four single indicators for gender equality - WEF GGGI, UNDP GII, F/M LFP, and TSWS.

Their results can be found in Fig. 2 A-F of Falk & Hermle (2018a), summarized here in Table 1 in comparison with our analysis. The results found are all in agreement with the original ones, although with some differences in terms of p-values.

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 did not differ significantly from the original and the replication analysis (see Table 1 below).

Table 1: Comparison of the conditional analysis results for the original study (where OLS was used) and our replication using the OLS and the robust linear regression (RLR). Reported are the **slopes** of the linear regressions and the corresponding p-value. In parenthesis, we indicate the standard error of the coefficient. Note that we do not know the errors related to FH study. Significance levels: ≤ 0.001 (***), ≤ 0.01 (**), ≤ 0.05 (*).

Coefficient	Regression on	Conditional on	Original (OLS)	Replication (OLS)	Replication (RLR)
$\beta_{EconDevelop}$	Log GDP p/c	GEI	0.5258***	0.50 (0.09)***	0.40 (0.10)
$\beta_{EconDevelop}$	Log GDP p/c	WEF GGGI	-	0.62 (0.09)***	0.63 (0.09)
$\beta_{EconDevelop}$	Log GDP p/c	UNDP GII	-	0.40 (0.20)*	0.40 (0.20)
$\beta_{EconDevelop}$	Log GDP p/c	F/M LFP	-	0.66 (0.08)***	0.65 (0.09)
$\beta_{EconDevelop}$	Log GDP p/c	TSWS	-	0.64 (0.09)***	0.63 (0.09)

Coefficient	Regression on	Conditional on	Original (OLS)	Replication (OLS)	Replication (RLR)
$\beta_{GenderEquality}$	GEI	Log GDP p/c	0.3192**	0.36 (0.09)***	0.30 (0.10)
$\beta_{GenderEquality}$	WEF GGGI	Log GDP p/c	0.2327**	0.22 (0.09)*	0.21 (0.09)
$\beta_{GenderEquality}$	UNDP GII	Log GDP p/c	0.2911	0.30 (0.20)	0.30 (0.20)
$\beta_{GenderEquality}$	F/M LFP	Log GDP p/c	0.2453*	0.22 (0.08)**	0.21 (0.09)
$\beta_{GenderEquality}$	TSWS	Log GDP p/c	0.2988**	0.19 (0.90)*	0.20 (0.10)

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 way the purely custom Gender Equality Index has been built and the reason for it to be 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 FH. To characterize the structure of the GEI, we visualized its composition with the diagram shown in Figure 2 and here below we briefly summarize the main issues found.

- As one can see in Figure 1, 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, FH 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.

- Another critical point for the use of PCA is 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).

- The *Time Since Women Suffrage* index introduced by FH. This last index has been introduced to track the long-lasting effects of the right to vote, and it 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. Note that within this indicator, some countries have not been recorded and therefore FH take the missing points from the WEF Global Gender Gap Report of 2006.

For all the reasons above, we decided to exclude the *ratio of female and male labor force participation* and the *Time Since Women Suffrage* indicators from further analysis, and instead to continue only with the WEF GGGI and UNDP GII. These indicators are internationally recognized, studied, and adopted from many researchers and policy makers (!!! QUOTE !!!), although are not exempt from critics.
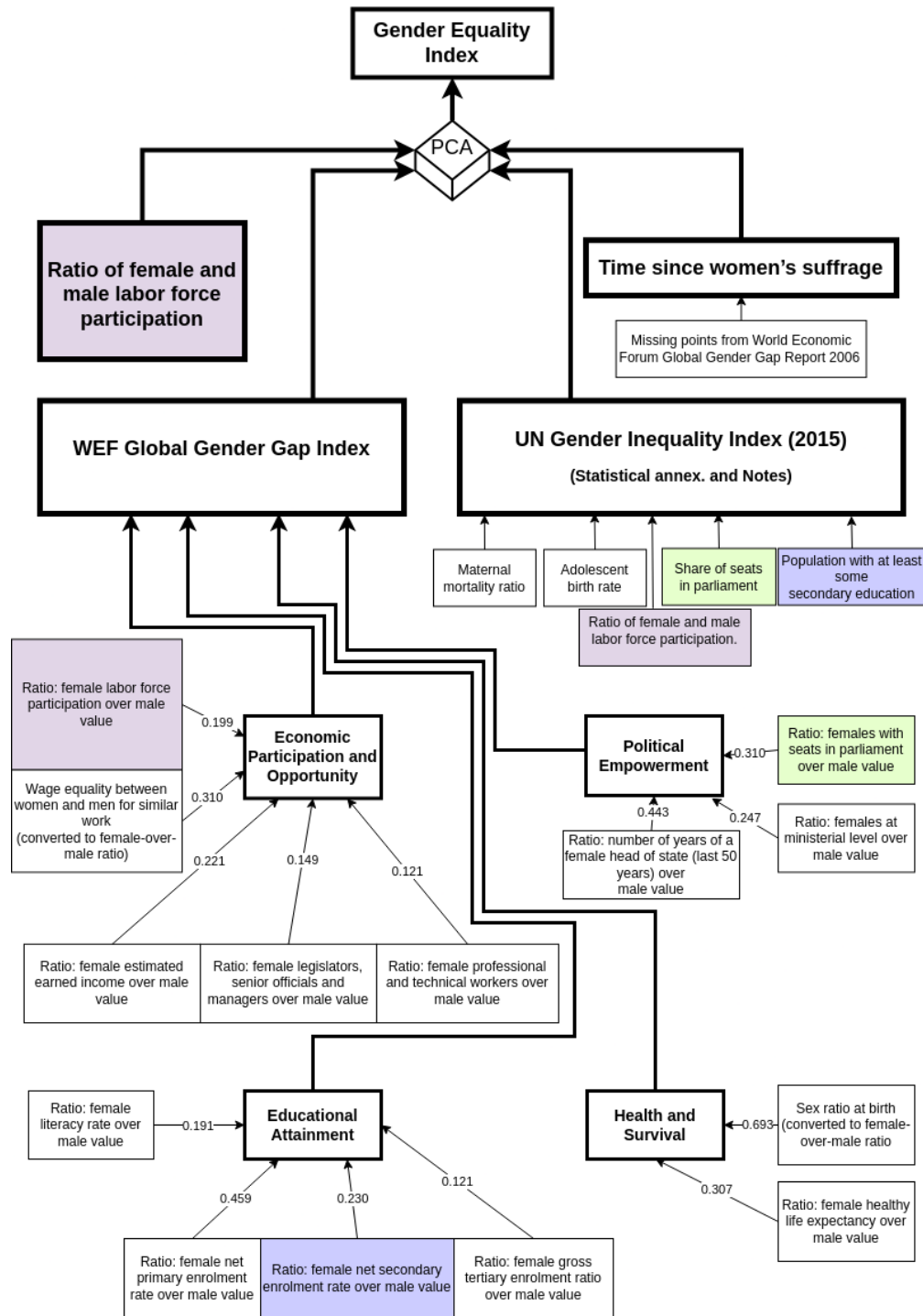
Figure 1: The custom Gender Equality Index decomposed in its sub-indexes, as used by the original authors. The repeated indexes and sub-indexes are highlighted with different colors. Note that for each sub-index the WEF GGGI calculates a weight to balance its impact on the overall index, while in the 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)

The UNDP GII has been criticized by several authors (Sen & Anand 1996; Permanyer 2011; Klasen 2017), 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 1), 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, this dependence exists and is not negligible (see Appendix).

Our approach to counteract the effects of the above-mentioned problematic has been to conduct an extensive research in the literature to understand what are the best options available at this time to evaluate the gender equality in the countries. We found that the Gender Development Index added from 2014 to the UNDP report is a good candidate for this kind of evaluation. This index is defined as the ratio of the Human Development Index for females divided by that for males, and it captures three dimensions in terms of health, knowledge, and living standards, separately for males and females. The life expectancy, the expected year of schooling and mean years of schooling, and GNI per capita are calculated within these dimensions. This index has been discussed in Klasen (2017), being praised for the good interpretability and for the focus on gender equality (rather than the sole female relative achievements). Therefore, we included this index in our extended analysis, together with the WEF GGGI and the UNDP GII.

Conditional analysis of gender differences in economic preferences 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 2, the gender differences in economic preferences 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 (r = 0.28, 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 gender difference in the six single economic preferences, and the GenderEquality$_c$ indicates the three indicators for gender equality in the countries, with the index c indicating the country-level. 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. The standard errors of the correlation coefficient have been calculated using the formula $se_r = \sqrt{\frac{1-r^2}{n-2}}$, with r^2 being the correlation coefficient and n being the sample size for that set of data. We refer to the Appendix for the corresponding value of each index.

As one can see in the following Table 2, 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
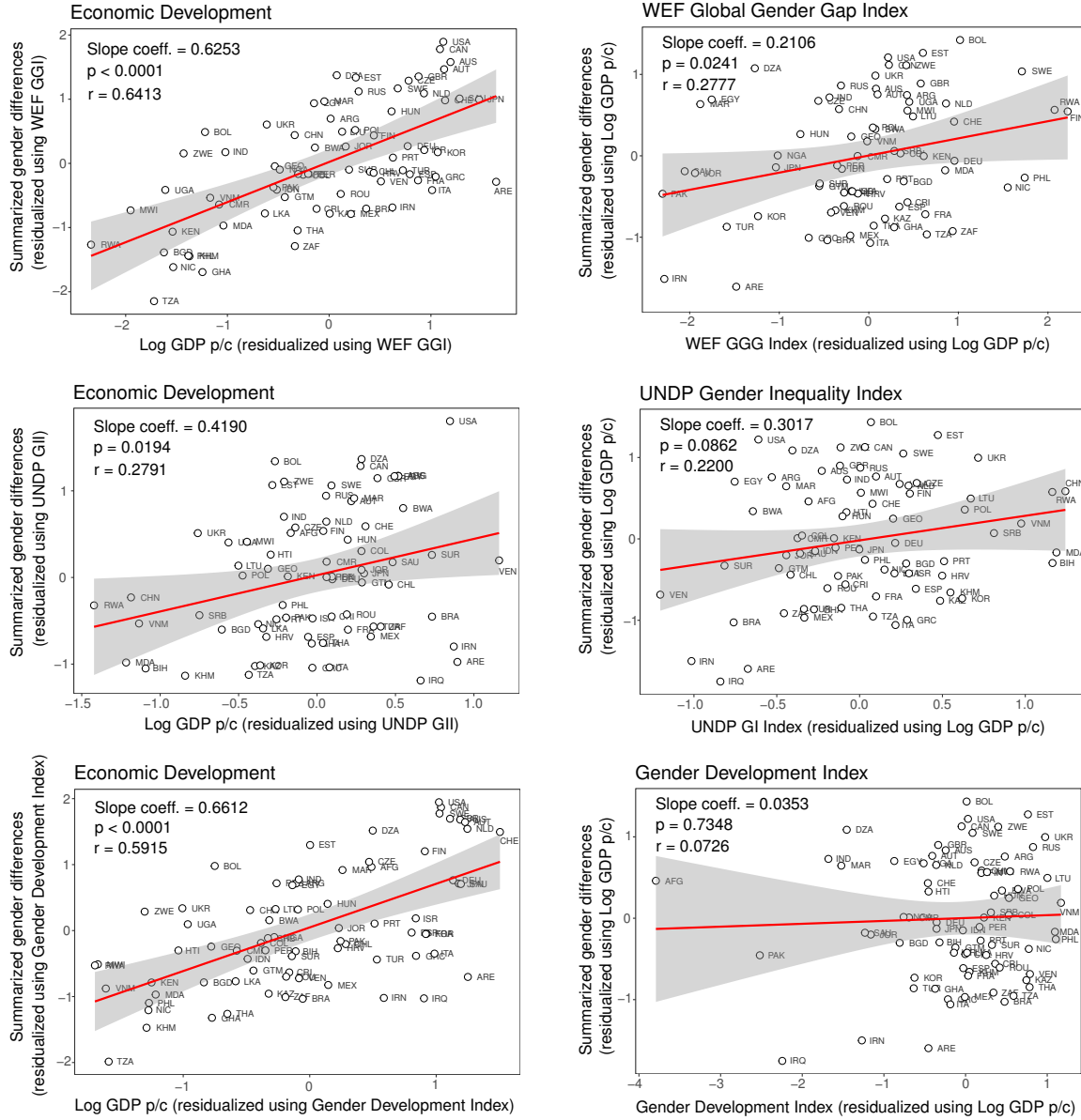
Figure 2: Correlation between gender differences in economic preferences and economic development, and between gender differences in economic preferences and gender equality indexes, using the residuals plots. On the left, gender differences are regressed on economic development conditioned on gender 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.

each gender equality indicator conditioning on economic development ($r_{GenderEquality}$ of Table 3), we see that in 17 out of 18 regressor-regressee pairs, no statistically significant correlations were observed. In one case only a moderate but statistically significant correlation was found, between the pairs WEF GGG-Altruism.

Taking into account these results, we conclude that there is a lack of evidence 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. *The UNDP GII index is the index with the highest correlation with Log GDP p/c and therefore an effect of multicollinearity might take place, deflating the values of the multilinear regression.*

Table 2: **Correlation between the gender differences in each economic preference and Log GDP p/c**. For all the Tables, the multilinear regression method is used. The correlation terms, their standard errors and significance levels ≤ 0.001 (***), ≤ 0.01 (**), ≤ 0.05 (*) are reported.

Preference	WEF GGGI $r_{LogGDPpc}$	UNDP GII $r_{LogGDPpc}$	UNDP GDI $r_{LogGDPpc}$
Trust	0.52 (0.10)***	0.18 (0.12)	0.49 (0.10)***
Altruism	0.55 (0.10)***	0.38 (0.11)***	0.48 (0.10)***
Pos. Recip.	0.29 (0.11)*	0.18 (0.12)	0.19 (0.11)
Neg. Recip.	0.37 (0.11)**	0.04 (0.12)	0.35 (0.11)***
Risk Taking	0.35 (0.11)**	0.04 (0.12)	0.32 (0.11)***
Patience	0.37 (0.11)***	0.10 (0.12)	0.31 (0.11)***

Table 3: **Correlation between the gender differences in each economic preference and on the three gender equality indexes**. For all the Tables, the multilinear regression method is used. The correlation terms, their standard errors and significance levels ≤ 0.001 (***), ≤ 0.01 (**), ≤ 0.05 (*) are reported.

Preference	WEF GGGI $r_{GenderEq}$	UNDP GII $r_{GenderEq}$	UNDP GDI $r_{GenderEq}$
Trust	0.14 (0.12)	0.22 (0.11)	0.08 (0.12)
Altruism	0.38 (0.11)***	-0.05 (0.12)	0.15 (0.11)
Pos. Recip.	0.03 (0.12)	-0.01 (0.12)	0.20 (0.11)
Neg. Recip.	0.16 (0.12)	0.19 (0.12)	-0.08 (0.12)
Risk Taking	0.06 (0.12)	0.21 (0.11)	-0.01 (0.12)
Patience	0.21 (0.12)	0.16 (0.12)	0.12 (0.12)

5. Discussion and conclusions

In the present article, we replicated and extended the results of the work by Falk & Hermle (2018a) that related gender differences in economic preferences to economic development and gender equality.

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 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, while for UNDP GII the correlation was somewhat milder. 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 with the WEF GGGI was statistically significant but moderately weak.

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 GII. 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* when conditioning on UNDP GDI, thus providing little support for the presence of a correlation between single gender differences in economic preferences and gender equality.

These results highlight the fact that the hypotheses under test (the social role theory and the resource hypothesis) are both inadequate to describe the situation shown by the data. In fact, as the hypotheses aim to prove whether gender differences in economic preferences increase or decrease with economic development and with gender equality of a country, we could not reject any of the two hypotheses. The data shows an increase of gender differences in economic preferences strongly correlated with increase of economic development of the countries, but no correlation was found with the gender equality of the countries. In other words, the increase of equality in a country - holding the economic development fixed - does not seem to increase the gender differences measured in economic preferences, while the increase of economic development - holding the gender equality fixed - does.

We can also state that gender differences in economic preferences are larger in gender-egalitarian countries or, equivalently, in more developed countries (and here the terms can be interchangeable because of the correlation that holds between them), but that an increase in gender equality at a given level economic development does not lead to the increase of gender differences in economic preferences. Therefore, the data does not suggest that the implementation of the policies that are targeting closing the gap leads to the increase in gender differences in economic preferences.

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

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

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.

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.

Niederle, M. & Vesterlund, L. (2005). *Do women shy away from competition? Do men compete too much?* National Bureau of Economic Research. Retrieved from http://www.nber.org/papers/w11474

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/