More on the influence of gender equality on gender differences in economic preferences

Sara Cerioli[1] and Andrey Formozov[2]

[1]Independent researcher, Mannheim, Germany
[2]Department of Neurophysiology, MCTN, Medical Faculty Mannheim, Heidelberg University, 68167 Mannheim, Germany
Correspondence: sara.cerioli@outlook.com, formozoff@gmail.com

Abstract

This study replicates and extends the work of Falk and Hermle (2018), who hypothesized that gender differences in economic preferences (patience, altruism, willingness to take risks, negative and positive reciprocity, and trust) were related to economic development and gender equality. While replicating their analysis, we identified various aspects that question the original approach, encouraging us to expand the scope of the research. Specifically, the use of an ad hoc gender equality index built by the authors lacked systematic justification, which led us to employ solely well-established indexes from gender studies in the subsequent analysis. This analysis confirmed a positive and statistically significant association between aggregated gender differences in economic preferences and economic development conditional on gender equality. However, in contrast to the original article, the evidence of the relationship between gender differences and gender equality conditional on economic development was weak. We also investigated the relationships for the separate economic preferences and found that economic development predicts gender differences in all six preferences, whereas gender equality seems to have a negligible or null influence on most of them. Our findings provide a more nuanced view of the gender differences in economic preferences, with possible implications in policy-making.

1. Introduction

Published findings on gender differences in human perceptions and behaviors, such as happiness (Schneider et al. 2012), competition (Croson and Gneezy 2009; Gneezy, Leonard, and List 2009; Niederle and Vesterlund 2005; Klonner, Pal, and Schwieren 2021), and work preferences (Beblo and Görges 2018), and their relation to gender inequality, are frequently used to influence decisions and policy-making, both in the public and private sectors. Moreover, gender inequality topics are becoming a more integral part of the agenda for many institutions and organizations, and stakeholders need to reveal, estimate, monitor, and prevent gender inequalities on individual, group, and nationwide levels.

The study of behavioral gender differences on a world scale is challenging. One challenge hampering progress in the field is the lack of large and homogeneous data sets across different social groups and countries. In an influential article published in the Quarterly Journal of Economics (Falk et al. 2018), a world scale data set on economic preferences, the Global Preference Survey within the Gallup World Poll 2012, was analyzed. The study focused on general questions about the distributions of economic preferences - defined as patience, altruism, willingness to take risks, negative and positive reciprocity, and trust - in different countries, relating them to several variables from the Gallup World Poll, such as age, gender, education level, and others. The subsequent article (Falk and Hermle 2018a) which we abbreviated in the following text as FH, used the same data set but focused explicitly on the gender differences highlighted in the previous study and reported evidence for the relationships of gender differences in economic preferences with economic development and gender equality.

Following the approach described in FH and analyzing the same data set, we were able to replicate their work and found very similar results in terms of the magnitude and significance of the regression coefficients, as described below. Although statistically significant and socially relevant, FH's findings pose additional questions that the authors give only a cursory treatment and that we focused on in our extended analysis.

The first relevant issue we addressed is the robustness of empirical findings to various indexes for gender equality, as several indexes have been used in FH study. In this regard, one point of concern is FH's introduction of a customized index with insufficient justification and limited final interpretability. In our analysis, we provided a list of possible shortcomings in this custom aggregated index and some of its components. We then restricted our analysis exclusively to the indexes widely used by academic, governmental, and other institutions (World Economic Forum 2006, 2015).

The study's main focus was general hypothesis testing, applying dimensionality-reduction techniques to aggregate gender differences in separate preferences into one joint index, and incorporating a variety of standardizing and normalizing procedures to the measurables. However, investigating in details these gender differences by focusing on separate preference measures – in contrast to the aggregated ones – allows to reveal the possible major contributors to such differences with respect to economic development and gender equality. Moreover, from the perspective of policy-making and follow-up studies, it is more informative to measure the magnitude of the differences (McCartney and Rosenthal 2000) rather than just conducting hypothesis discrimination. The present study aims to fill this gap.

The article is structured as follows: Section 2 presents the summary of FH's original article, while Section 3 contains our replication of FH's main findings. The first part of Section 4 is dedicated to the issues related to measuring gender inequality – particularly the custom index FH built for this purpose – while the second part reports the magnitude of the effects of gender differences on aggregated and separate preferences. Finally, in the Discussion and Conclusions (Section 5), we evaluate the relevance of our results for further studies and practical use in various areas.

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 (Falk and Hermle 2018a). The authors used the Gallup World Poll 2012 Global Preference Survey to measure gender differences in economic preferences across 76 countries, representing nearly 90% of the world population, with a total of almost 80,000 people surveyed and each country having around 1,000 participants.

Survey participants 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 Falk and Hermle (2018b), section "Extended Materials and Methods"). Therefore, for each person in the data set, each of the six economic preferences was scored. Additional individual-level variables indicating age, sex, education level, subjective math skills (as a proxy for cognitive skills), and household income quintile were also collected.

The authors proposed two competing hypotheses to be tested:

1. Social role hypothesis: "Following social role theory, one may hypothesize that gender differences in preferences attenuate in more developed, gender-egalitarian countries [...]. 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."
2. Resource hypothesis: "In contrast [to hypothesis 1], there is reason to expect that gender differences in preferences expand with economic development and gender equality [...]. 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."

To aggregate gender differences among the six economic preferences, the dimensionality-reduction technique of principal component analysis (Jolliffe 2002), also known as PCA, was performed on the gender coefficients. The first component of the PCA was then used as a summary index for gender differences in economic preferences. The logarithm of GDP per capita (Log GDP p/c) was used as a proxy for the economic development of the countries under study, while for gender equality FH used a customized gender equality index, which they called the Gender Equality Index. This index is built using the first component of a PCA applied to four different indexes for 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 to male labor force participation (F/M LFP), taken from the World Bank database, and the Time Since Women's Suffrage (TSWS), from the Inter-Parliamentary Union Website.

The study reported a positive, 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 the custom Gender Equality Index proposed by the authors ($r = 0.56$, p-value < 0.0001), as reflected in the Research article summary and the graphic abstract of FH study. The authors also conducted a conditional analysis to isolate the impact of economic development and gender equality. This time, they reported the regression coefficient being large and statistically significant (slope coefficient $= 0.53$, p-value < 0.0001) when gender differences were related to Log GDP p/c conditioned by Gender Equality Index, while moderately weak and statistically significant (slope coefficient $= 0.32$, p-value $= 0.003$) when relating to Gender Equality Index and controlling for Log GDP p/c.

Based on this evidence, the authors concluded 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 and 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 on the same data to take into account the non-normality of the data set.

3.1. Data

To conduct the replication, we downloaded the Gallup World Poll 2012 Global Preferences Survey data set from the briq - Institute on Behavior & Inequality website. The full data set is under restricted access, and education level and household income quintile on the individual level are not available in the open-access version (for more information, see Supplementary Material "Data Collection, Cleaning, and Standardization"). In Falk and Hermle (2018b), FH provide a complementary analysis where all the independent variables (except for gender) are dropped, and the results are coherent with what was found in their main analysis. Therefore, we decided to continue the replication study without having access to education level and income quintile.

3.2. Methods and Results

Following the analysis conducted by FH, we built a multilinear regression model to assess the relationship between each of the six economic preferences, standardized at the global level to exhibit a mean of 0 and a standard deviation of 1, and the independent variables associated to the individuals across countries:

$$\text{preference}_i^c = \beta_1^c \text{female}_i + \beta_2^c \text{age}_i + \beta_3^c \text{age}_i^2 + \beta_4^c \text{subjectiveMathSkills}_i + \epsilon_i \quad (1)$$

The subscript i is the index of a survey participant and c is the index for a country. This results in six models – one for each economic preference – with four coefficients. The coefficient for the dummy variable *female*, β_1^c,

is used as a measure for gender difference, showing how many standard deviations far apart are men and women for a given economic preference in a certain country.

We performed PCA on the six coefficients for gender differences of the separate preference measures and used the first component to obtain a single measure for gender differences. FH referred to this summarized index as "average gender differences". We find this nomenclature potentially confusing, therefore we refer to it as either "aggregated index" or as "summarized index", rather than "average". The PCA technique has also been applied to the four gender equality indexes to get the joint index (Gender Equality Index) already described in Section 2 of this paper.

The competing hypotheses proposed by FH and described in Section 2 can be formally written using the following multilinear model:

$$\text{Summarized Gender Diff} = \beta_{EconDevelop}\,\text{Econ Develop} + \beta_{GenderEquality}\,\text{Gender Equality} \tag{2}$$

Where the variable *Econ Develop* is always Log GDP p/c, while *Gender Equality* can be either the Gender Equality Index or one of its sub-indexes (WEF GGGI, UNDP GII, F/M LFP, and TSWS). All the variables were standardized at the global level to show a mean of 0 and a standard deviation of 1. After standardization, $Summarized\ Gender\ Diff$ shows how many standard deviations away is a certain country from the global average gender difference.

From the equation above, we would expect that:

1. If the social role hypothesis is correct, the model above will result in negative coefficients for economic development and gender equality.
2. If the resource hypothesis is correct, then we will have positive coefficients for economic development and gender equality.

Any other scenario (for example, when one coefficient in the model is positive and the other is negative) is left out of FH's original research design and would require the formulation of additional hypotheses and further studies.

Since the correlation between economic development and gender equality is a known effect (Duflo 2012; World Economic Forum 2015), we checked the correlation between their proxies, regressing Log GDP p/c on the Gender Equality Index built by FH. The correlation found is moderately strong ($r = 0.54$) and statistically significant (p-value < 0.0001), as one can see in our Supplementary Material, Figure 3. The multilinear regression takes into account this correlation, and the theorem from Frisch–Waugh–Lovell (Frisch and Waugh 1933; Lovell 1963) guarantees that the coefficients found are the same as those found in the residual analysis, as performed in FH.

In Table 1, we summarize the comparison of our analysis to the one performed in FH (Figures 2 A-F of Falk and Hermle (2018a)). The results found are all in agreement with those of the original study (although with some differences in p-values), except for the coefficient found for TSWS. The difference is not surprising, as TSWS was one of the most difficult indicators to replicate because of a lack of clear instructions in FH (see also our Supplementary Material, "Data Collection, Cleaning, and Standardization"). Note also that the coefficients for economic development conditional on the four single indexes for gender equality are not provided in FH's original analysis.

3.3. Robust Linear Regression

Within the Global Preference Survey, economic preferences were measured with both qualitative and quantitative responses. For all the economic preferences, a qualitative question based on a Likert scale between 0 and 10 was used, while a quantitative measurement was performed for every preference, excluding trust (please refer to Falk and Hermle (2018b) for further details). This mixed approach of semi-continuous and ordered categorical variables has led us to the question of the appropriateness of the OLS method for the data analysis.

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 with robust linear regression (Fox 2015) instead of ordinary linear regression to mitigate potential biases introduced by outliers. 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 from FH study (where OLS was used) and our replication using the OLS and the RLR. Reported are the coefficients of the linear regressions and the corresponding p-values. In parenthesis, we indicate the standard error of the coefficient. Note that the errors related to FH study are missing because they were not reported in the article. 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)***	0.49 (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)*	0.42 (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)***
$\beta_{GenderEquality}$	GEI	Log GDP p/c	0.3192**	0.36 (0.09)***	0.34 (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)**	0.20 (0.09)*
$\beta_{GenderEquality}$	TSWS	Log GDP p/c	0.2988**	0.19 (0.09)*	0.19 (0.10)*

4. Analysis of Gender Equality Indexes and Separate Preference Measures

This section brings attention to various unresolved concerns regarding the Gender Equality Index built by FH, especially the lack of interpretability associated with this index. We then extended the analysis to state-of-the-art indexes such as WEF GGGI, UNDP GII, and UNDP Gender Development Index (GDI). We also delved deeper into the relationship between the separate preference measures and gender equality to better understand the source of the largest differences. Doing so sheds light on the found associations and provokes new research questions.

4.1. Gender Equality Indexes and Potential Issues

During the replication analysis, we encountered potential issues related to the gender equality indexes that we considered worthy of analyzing further.

One concern is the way FH built their Gender Equality Index and why they propose it as a measure of gender equality. The justification for using this custom index rather than internationally recognized, studied, adopted, and already available indexes was omitted in FH's study. To characterize the structure of the Gender Equality Index, we visualized its composition with the diagram shown in Figure 1. We will briefly summarize the main issues found below.

- The most critical point for the use of PCA is the interpretability of the index. This is a central question for measuring differences in society without losing descriptive power and is crucial for identifying effective policies for closing the gender gap (World Economic Forum 2015).

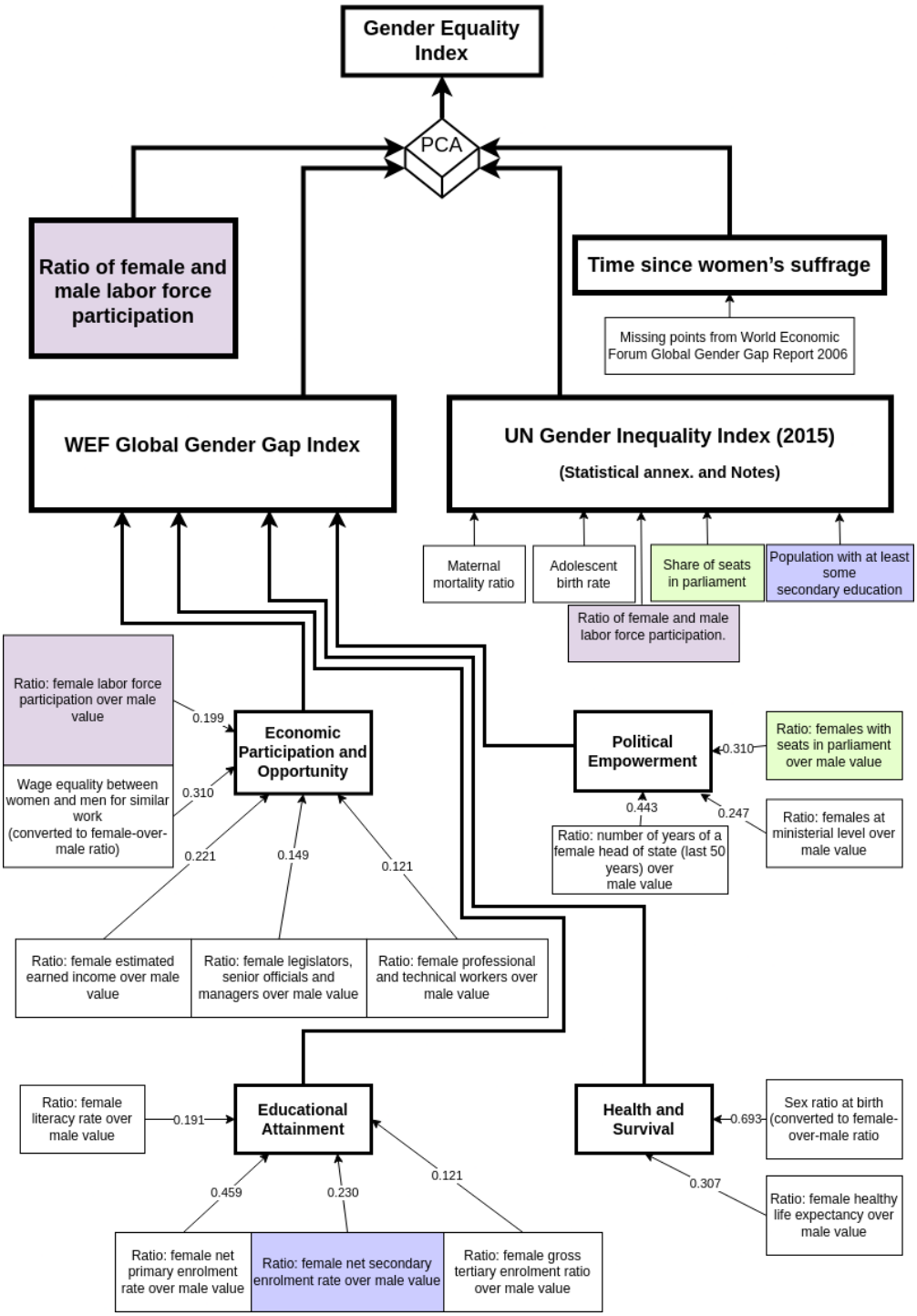

Figure 1: The custom Gender Equality Index decomposed into its sub-indexes, as built by FH. The repeated indexes and sub-indexes are highlighted with different colors. In the index "Time since Women's Suffrage", there are missing points that have been taken from another source, WEF Global Gender Gap Report 2006, following the approach used in FH study. Note that for each sub-index WEF GGGI calculates a weight to balance its impact on the overall index, while UNDP GII treats the sub-indexes without extra weighting. See also the technical notes of the World Economic Forum (2015) and of United Nations Development Programme (2021).

- As seen in Figure 1, the components of the Gender Equality Index used in FH contain repetitions. The two indexes, WEF GGGI and UNDP GII, share three sub-indexes, indicated here with different colors: *ratio of female to 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 Gender Equality Index, FH used the *ratio of female to 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 variables even in the 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 TSWS indicator, introduced by FH, is used as a proxy to track the long-lasting effects of the right to vote. The use of this proxy is based on the assumption that, over time, development has a monotonic effect and its magnitude is proportional to the time since women's suffrage was established. The data on the year of suffrage is available on a global scale but provides a minimal overview of gender disparities in politics, as discussed in Carmichael, Dili, and Rijpma (2014). Indeed, even after the right to vote has been granted, many discriminatory laws may still be present and enforced by the executive branch of the government. Elimination of discrimination takes more time – despite gaining the right to vote, there may not be an improvement in gender equality in other areas (Yang 2020), such as suppression of the right to work. As an example, married women in Western Germany needed permission of their husbands to work until 1977 (Grebe 2010), although their right to vote was granted in 1918. The assumption that suffrage played a long-lasting effect on the balance of gender equality sounds reasonable but requires further investigation to be used as a robust estimator.

For all these reasons, we decided to discontinue the use of the FH custom index, exclude the F/M LFP and the TSWS indicators from further analysis, and instead carry on only with WEF GGGI and UNDP GII.

Although these indexes are widely used by many researchers and policymakers, they also have their weaknesses. The UNDP GII has been criticized by several authors (Klasen 2017; Sen and Anand 1996; Permanyer 2011), as being highly correlated with economic development and including reproductive health indicators that can penalize less-developed countries. Moreover, this index measures welfare loss associated with inequality based on a calculated gender equality measure, which is not documented publicly.

The power and limitations of WEF GGGI are discussed in Piper (2019), Barns and Preston (2010), Worsdale and Wright (2021). One of the main critics is that this index is truncated in such way that it does not allow any country to be more favorable for women than for men. Another weak point highlighted by our investigation (Figure 1) is the inclusion of a subjective measure, based on an expert panel opinion, called "wage equality between men and women for similar work", representing a substantial part (~30%) of one of the indicators composing the GGGI. Additionally, although the WEF GGGI index is thought to be the least dependent on economic development since it measures the gap between male and female access to resources and opportunities (World Economic Forum 2015), this dependence exists and is not negligible (see Supplementary Material, Figure 3).

We conducted extensive research in the literature to identify and understand the current options available for evaluating gender equality. In addition to the two above-mentioned indexes, 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 to males, and it captures three dimensions in terms of health, knowledge, and living standards, separately for males and females. Life expectancy, the expected year and mean years of schooling, and GNI per capita are calculated within these dimensions. This index is praised by Klasen (2017) for its clarity in interpretability and focus on gender equality (rather than female relative achievements). Therefore, we included this index in our extended analysis, along with WEF GGGI and UNDP GII.

4.2. Conditional Analysis of Gender Differences in Economic Preferences and Their Relationship to Economic Development and Gender Equality

In this section, we explore the relationship between the aggregated gender differences in economic preferences with economic development and gender equality, using the gender equality indexes referred to above (WEF GGGI, UNDP GII, and UNDP GDI). We run the same robust linear regression using the model in Eq. 2 on these indexes and present the results in Figure 2. One can see that gender differences in economic preferences have a strong, positive, and statistically significant correlation with economic development when the conditional analysis is performed on the individual gender equality indexes. Conversely, the correlation between gender differences in economic preferences and gender equality, conditioned on 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 5% confidence level.

To investigate the role of economic development and gender equality on separate preferences, we used a multilinear regression model for each preference, similar to the multilinear regression using the aggregated gender differences described earlier in Eq. 2:

$$\text{Gender Diff}^p = \beta^p_{EconDevelop}\,\text{Econ Develop} + \beta^p_{GenderEquality}\,\text{Gender Equality} + \epsilon \tag{3}$$

where the index p indicates economic preferences. $Gender\ Diff^p$ indicates gender differences in the original six economic preferences, and $Gender\ Equality$ indicates the three individual indexes (WEF GGGI, UNDP GII, and UNDP GDI) for gender equality. As done in Eq. 2, all the variables were standardized at the global level to show a mean of 0 and a standard deviation of 1. After standardization, $Gender\ Diff^p$ shows how many standard deviations away is a certain country from the global average gender difference for a certain preference.

As one can see in Table 2, the regression coefficient related to economic development (written as $\beta^p_{LogGDPpc}$) is in most cases positive and statistically significant. On the other hand, when we look at the coefficients related to the gender equality index ($\beta^p_{GenderEquality}$ of Table 3), we see that in 16 of the 18 regressor-regressee pairs, no statistically significant coefficients were observed. Only in two cases statistically significant coefficients were found: Between the pairs WEF GGG-Altruism and UNDP GII-Risk Taking.

This analysis suggests that the association between gender differences in separate preference measures and economic development still holds (Table 2). Regression coefficients for separate economic measures show an increase in gender differences with respect to the global average in the range between 0.22 and 0.55 standard deviations for one standard deviation change in Log GDP p/c, when conditioning on gender equality.

At the same time, the results do not support an association between gender differences in separate preference measures and gender equality indexes for the absolute majority of the preferences (Table 3). Together with the results of the aggregated gender differences, as seen in Figure 2, this absence of consistency leads us to the conclusion that either there is an association but it is weak, or that such association does not exist at all. In the latter case, the set of hypotheses should not be limited to the two alternatives that were proposed as the main hypotheses in FH.

It is important to highlight that FH also explored the separate preference measures by conducting a conditional analysis on Log GDP p/c and their custom Gender Equality Index. They found a statistically significant correlation between the Gender Equality Index and only three of the six preferences, while all the preferences were correlated with Log GDP p/c conditional on the Gender Equality Index (refer to Figures S5 and S6 in Falk and Hermle (2018b)). To compare these results with ours, we presented the correlation coefficients in the Supplementary Material (Tables 3 and 4). Note that the corresponding slope coefficients were not reported in FH article.

Thus far, our analysis has been performed by standardizing gender differences in separate preferences ($Gender\ Diff^p$ in Eq.3). This imparts slope coefficients with the meaning of indicating a specific country's deviation from the global average gender difference, measured in standard deviations. To assess the actual magnitude of modulation of gender differences for each separate preference, one needs to remove the standardization, yielding to slope coefficients that quantify the extent to which men and women differ in a

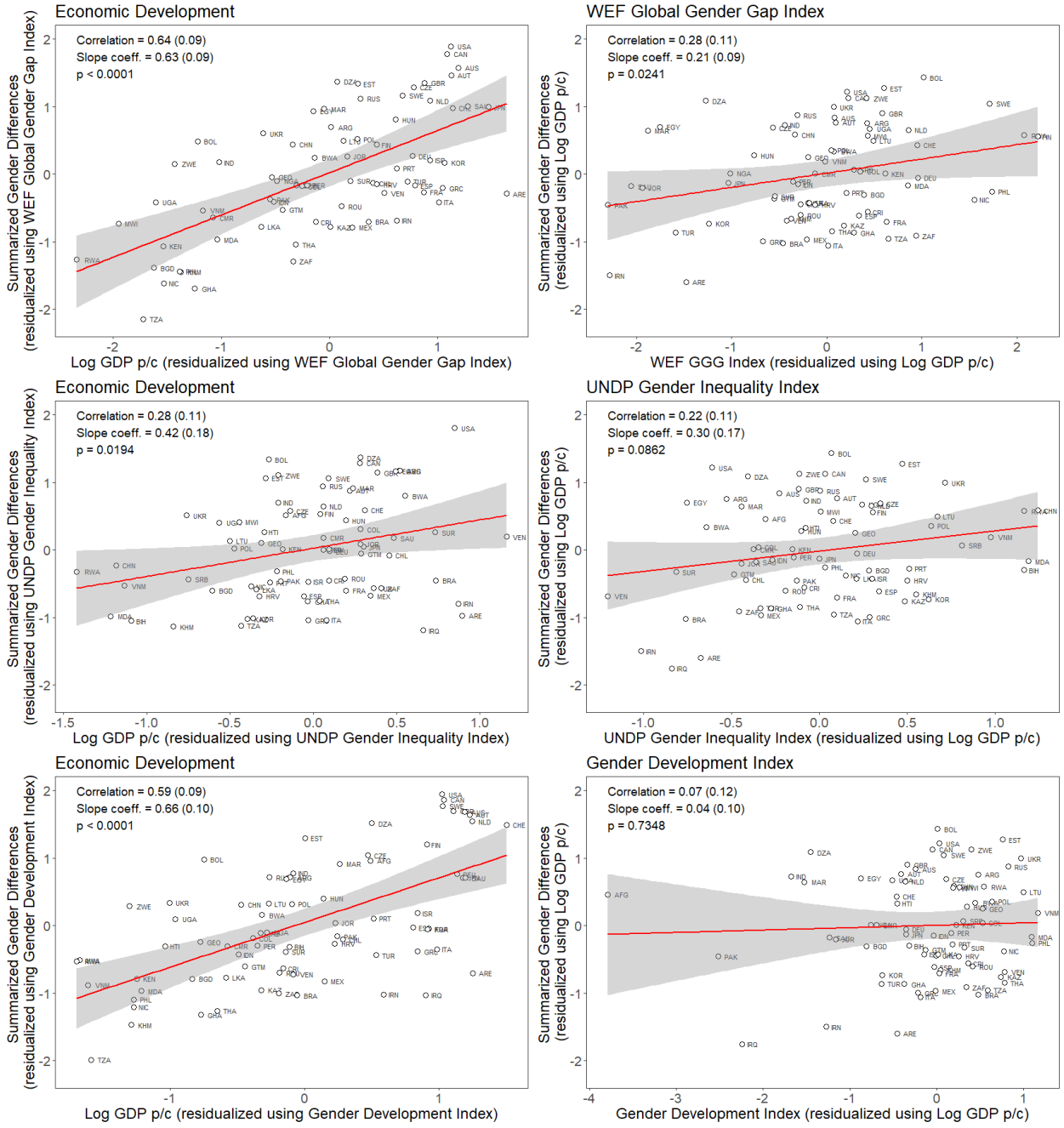

Figure 2: Correlation and slope coefficients between summarized gender differences in economic preferences and economic development, and between summarized gender differences in economic preferences and gender equality indexes, using the residuals plots. On the left, gender differences are regressed on economic development conditioned on gender equality for the different indexes (WEF GGGI, UNDP GGI, and GDI). On the right, the corresponding values of gender differences are regressed on gender equality indexes conditioned on economic development. We also report the corresponding p-values.

given preference in terms of standard deviations. We reported the corresponding slope coefficients in Tables 4 and 5. As one can see, the increase in gender differences in separate preferences is in the range between 0.03 and 0.08 standard deviations for one standard deviation change in Log GDP p/c, when conditioning on gender equality. As already noted above, for gender equality indexes only altruism for WEF GGGI and risk-taking for UNDP GII give a statistically significant result, which is not higher than 0.05 standard deviations increase.

Table 2: The slope coefficients for economic development from the eighteen multilinear regression models for separate economic preferences and three distinct choices of gender equality indexes (WEF GGGI, UNDP GII, UNDP GDI). As in FH, the symbols (+)/(-) indicate the general direction of the difference. (+) indicates that women exhibited higher levels of the respective preference compared to the global average (-) indicates that men on average exhibited higher levels of the respective preference. The regression coefficients, their standard errors in brackets, and significance levels ≤ 0.001 (***), ≤ 0.01 (**), ≤ 0.05 (*) are reported. RLR method is used.

Preference	$\beta^p_{LogGDPpc}$ (WEF GGGI)	$\beta^p_{LogGDPpc}$ (UNDP GII)	$\beta^p_{LogGDPpc}$ (UNDP GDI)
Trust (+)	0.51 (0.11)***	0.28 (0.19)	0.57 (0.12)***
Altruism (+)	0.55 (0.10)***	0.66 (0.19)***	0.53 (0.12)***
Pos. Recip. (+)	0.22 (0.10)*	0.25 (0.17)	0.15 (0.10)
Neg. Recip. (−)	0.36 (0.11)**	0.10 (0.19)	0.43 (0.12)***
Risk Taking (−)	0.33 (0.11)**	0.04 (0.19)	0.39 (0.12)**
Patience (−)	0.29 (0.09)**	0.15 (0.16)	0.33 (0.10)**

Table 3: The slope coefficients for gender equality from the same eighteen multilinear regression models for separate economic preferences and three distinct choices of gender equality indexes (WEF GGGI, UNDP GII, UNDP GDI). As in FH, the symbols (+)/(-) indicate the general direction of the difference. (+) indicates that women exhibited higher levels of the respective preference compared to the global average (-) indicates that men on average exhibited higher levels of the respective preference. The regression coefficients, their standard errors in brackets, and significance levels ≤ 0.001 (***), ≤ 0.01 (**), ≤ 0.05 (*) are reported. RLR method is used.

Preference	$\beta^p_{GenderEq}$ (WEF GGGI)	$\beta^p_{GenderEq}$ (UNDP GII)	$\beta^p_{GenderEq}$ (UNDP GDI)
Trust (+)	0.11 (0.11)	0.36 (0.19)	0.04 (0.12)
Altruism (+)	0.27 (0.10)**	−0.11 (0.19)	0.11 (0.11)
Pos. Recip. (+)	0.04 (0.09)	−0.02 (0.17)	0.18 (0.10)
Neg. Recip. (−)	0.14 (0.11)	0.32 (0.19)	−0.10 (0.12)
Risk Taking (−)	0.08 (0.11)	0.40 (0.19)*	−0.10 (0.12)
Patience (−)	0.16 (0.09)	0.25 (0.16)	0.06 (0.10)

Table 4: The slope coefficients that quantify the magnitude of increase in gender differences for separate preferences per one standard deviation change in Log GDP p/c. As in FH, the symbols (+)/(-) indicate the general direction of the difference. (+) indicates that women exhibited higher levels of the respective preference compared to the global average (-) indicates that men on average exhibited higher levels of the respective preference. The regression coefficients, their standard errors in brackets, and significance levels ≤ 0.001 (***), ≤ 0.01 (**), ≤ 0.05 (*) are reported. RLR method is used.

Preference	$\beta^p_{LogGDPpc}$ (WEF GGGI)	$\beta^p_{LogGDPpc}$ (UNDP GII)	$\beta^p_{LogGDPpc}$ (UNDP GDI)
Trust (+)	0.06 (0.01)***	0.03 (0.02)	0.07 (0.02)***
Altruism (+)	0.07 (0.01)***	0.08 (0.02)***	0.07 (0.01)***
Pos. Recip. (+)	0.03 (0.01)*	0.03 (0.02)	0.02 (0.01)
Neg. Recip. (−)	0.04 (0.01)**	0.01 (0.02)	0.05 (0.01)***
Risk Taking (−)	0.04 (0.01)**	0.00 (0.02)	0.05 (0.01)**
Patience (−)	0.04 (0.01)**	0.02 (0.02)	0.04 (0.01)**

Table 5: The slope coefficients that quantify the magnitude of increase in gender differences for separate preferences per one standard deviation change in gender equality (WEF GGGI, UNDP GII, UNDP GDI). As in FH, the symbols (+)/(-) indicate the general direction of the difference. (+) indicates that women exhibited higher levels of the respective preference compared to the global average (-) indicates that men on average exhibited higher levels of the respective preference. The regression coefficients, their standard errors in brackets, and significance levels ≤ 0.001 (***), ≤ 0.01 (**), ≤ 0.05 (*) are reported. RLR method is used.

Preference	$\beta^p_{GenderEq}$ (WEF GGGI)	$\beta^p_{GenderEq}$ (UNDP GII)	$\beta^p_{GenderEq}$ (UNDP GDI)
Trust (+)	0.01 (0.01)	0.04 (0.02)	0.01 (0.01)
Altruism (+)	0.03 (0.01)**	−0.01 (0.02)	0.01 (0.01)
Pos. Recip. (+)	0.00 (0.01)	0.00 (0.02)	0.02 (0.01)
Neg. Recip. (−)	0.01 (0.01)	0.04 (0.02)	−0.01 (0.01)
Risk Taking (−)	0.01 (0.01)	0.05 (0.02)*	−0.01 (0.01)
Patience (−)	0.02 (0.01)	0.03 (0.02)	0.01 (0.01)

5. Discussion and Conclusions

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

First, as a benchmark, we performed a nearly exact replication of FH's analysis, obtaining the data from the Gallup World Poll 2012 Global Preference Survey and using the same methodology as FH. Unfortunately, the data set is publicly available only in a pre-processed form and is partially restricted. Nevertheless, we were able to replicate the analysis and obtained results similar to those in FH's original article. In addition, we ran the same analysis using robust linear regression instead of ordinary linear regression to correct for the non-normality and outliers observed in the data. However, no significant changes in the results were observed.

We then investigated the indexes used to estimate gender equality and their relationship with economic development. We analyzed the Gender Equality Index built by FH and its components. Some methodological

issues were identified, and we concluded that the use of this custom index over more established, balanced indexes lacks justification and remains open to question. Therefore, we conducted a further analysis based on separate, widely accepted indexes of gender equality used in FH's original article (WEF Global Gender Gap Index and UNDP Gender Inequality Index), plus an additional index – the UNDP Gender Development Index.

We examined gender differences in economic preferences and their relationship with economic development and gender equality using the above-mentioned indexes. Performing a conditional analysis, we found a positive, strong, and statistically significant correlation between the aggregated gender differences in economic preferences and economic development when we controlled for WEF GGGI and UNDP GDI; controlling for UNDP GII yielded a somewhat milder correlation. On the other hand, when controlling for economic development, no correlation between UNDP GII or GDI and the aggregated gender differences in economic preferences was found. We did find a statistically significant but weak correlation between aggregated gender differences in economic preferences and WEF GGGI. Therefore, the dependency of gender differences in economic preferences on gender equality can not be consistently supported when only established, commonly recognized indexes are used. This lack of consistency in the results leads us to the conclusion that either there is a weak correlation between gender differences in economic preferences and gender equality or that the correlation does not exist at all. In the latter case, the set of hypotheses should not be limited to the two alternatives that were proposed as the main hypotheses in FH.

We additionally analyzed how gender differences in separate preference measures are related to economic development and gender equality. Interestingly, we observed contrasting patterns in the regression coefficients between gender differences in each separate economic preference and both Log GDP p/c and gender equality indexes. Specifically, we found positive and statistically significant coefficients when examining the relationship between gender differences in separate preference measures and Log GDP p/c, when controlling for WEF GGGI and UNDP GDI. When controlling for UNDP GII, only one economic preference (altruism) shows a statistically significant coefficient, while for the other preferences the coefficients are not statistically significant. Meanwhile, among the six economic preferences studied in relationship with gender equality indexes, the only statistically significant coefficients were for altruism (which exhibited an association with WEF GGGI) and risk-taking (with an association with UNDP GII).

These findings align with our results obtained from the aggregated gender differences. However, the dependencies of gender differences in economic preferences on economic development and gender equality became more nuanced once analyzed for separate economic measures, with a more differentiated picture emerging. This differentiation becomes particularly important in the realms of decision-making and policy formulation.

Meanwhile, the impact of economic development on the manifestation of gender differences in preferences should also be further investigated. It might be worth researching the origins of these gender differences, considering hypotheses encompassing economic development and other potential country-level variables not considered in FH that could explain the disparities observed. For instance, one may speculate that at the higher level of economic development, the market expands and diversifies its offerings to cater to a wider range of consumers, exploiting status quo gender stereotypes as a starting point for the promotion of goods. In this case, a potential accentuation of gender differences by reinforcing gender stereotypical behaviors may occur.

References

Barns, Angela, and Alison Preston. 2010. "Is Australia Really a World Leader in Closing the Gender Gap?" *Feminist Economics* 16 (4): 81–103. https://doi.org/10.1080/13545701.2010.530607.

Beblo, Miriam, and Luise Görges. 2018. "On the Nature of Nurture. The Malleability of Gender Differences in Work Preferences." *Journal of Economic Behavior & Organization* 151: 19–41. https://doi.org/https://doi.org/10.1016/j.jebo.2018.05.002.

Carmichael, Sarah, Selin Dili, and Auke Rijpma. 2014. *Gender Inequality Since 1820.* https://doi.org/https://doi.org/https://doi.org/10.1787/9789264214262-16-en.

Croson, Rachel, and Uri Gneezy. 2009. "Gender Differences in Preferences." *Journal of Economic Literature* 47 (2): 448–74. https://doi.org/10.1257/jel.47.2.448.

Duflo, Esther. 2012. "Women Empowerment and Economic Development." *Journal of Economic Literature* 50 (4): 1051–79. http://www.jstor.org/stable/23644911.

Falk, Armin, Anke Becker, Thomas Dohmen, Benjamin Enke, David Huffman, and Uwe Sunde. 2018. "Global Evidence on Economic Preferences*." *The Quarterly Journal of Economics* 133 (4): 1645–92. https://doi.org/10.1093/qje/qjy013.

Falk, Armin, and Johannes Hermle. 2018a. "Relationship of Gender Differences in Preferences to Economic Development and Gender Equality." *Science* 362 (6412): eaas9899. https://doi.org/10.1126/science.aas9899.

———. 2018b. "Supplementary Materials of Relationship of Gender Differences in Preferences to Economic Development and Gender Equality." https://doi.org/http://science.sciencemag.org/content/suppl/2018/10/17/362.6412.eaas9899.DC1.

Fox, J. 2015. *Applied Regression Analysis and Generalized Linear Models*. SAGE Publications. https://books.google.de/books?id=3wrwCQAAQBAJ.

Frisch, Ragnar, and Frederick V. Waugh. 1933. "Partial Time Regressions as Compared with Individual Trends." *Econometrica* 1 (4): 387–401. http://www.jstor.org/stable/1907330.

Gneezy, Uri, Kenneth L. Leonard, and John A. List. 2009. "Gender Differences in Competition: Evidence from a Matrilineal and a Patriarchal Society." *Econometrica* 77 (5): 1637–64. https://doi.org/https://doi.org/10.3982/ECTA6690.

Grebe, Cornelius. 2010. "Reconciliation Policy in Germany 1998-2008." In, 92. https://doi.org/https://doi.org/10.1007/978-3-531-91924-9.

Jolliffe, I. T. 2002. *Principal Component Analysis*. Springer New York, NY. https://doi.org/https://doi.org/10.1007/b98835.

Klasen, Stephan. 2017. "UNDP's Gender-Related Measures: Current Problems and Proposals for Fixing Them." Discussion Papers 220. Göttingen: Georg-August-Universität Göttingen, Courant Research Centre - Poverty, Equity and Growth (CRC-PEG). http://hdl.handle.net/10419/157265.

Klonner, Stefan, Sumantra Pal, and Christiane Schwieren. 2021. "Equality of the Sexes and Gender Differences in Competitiveness: Experimental Evidence from a Traditional Society with Gender-Balanced Norms." *SSRN Electronic Journal*, January. https://doi.org/10.2139/ssrn.3940199.

Lovell, Michael C. 1963. "Seasonal Adjustment of Economic Time Series and Multiple Regression Analysis." *Journal of the American Statistical Association* 58 (304): 993–1010. https://doi.org/10.1080/01621459.1963.10480682.

McCartney, Kathleen, and Robert Rosenthal. 2000. "Effect Size, Practical Importance, and Social Policy for Children." *Child Development* 71 (1): 173–80. http://www.jstor.org/stable/1132230.

Niederle, Muriel, and Lise Vesterlund. 2005. "Do Women Shy Away from Competition? Do Men Compete Too Much?" Working Paper, Working paper series, no. 11474. https://doi.org/10.3386/w11474.

Permanyer, Iñaki. 2011. "Are UNDP Indices Appropriate to Capture Gender Inequalities in Europe?" *Social Indicators Research* 110 (February). https://doi.org/10.1007/s11205-011-9975-6.

Piper, Alan. 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 (3): 1391–1413. https://doi.org/10.1007/s11205-019-02080-.

Schneider, Udo, Christian Pfarr, Brit Schneider, and Volker Ulrich. 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 (February): 251–65. https://doi.org/10.1007/s10198-011-0301-7.

Sen, Amartya, and Sudhir Anand. 1996. "Gender Inequality in Human Development: Theories and Measurement." In *Background Papers: Human Development Report 1995*, 1–20. New York: United Nations Development Programme; United Nations Development Programme.

United Nations Development Programme. 2021. "Human Development Report, Technical Notes." https://hdr.undp.org/sites/default/files/2021-22_HDR/hdr2021-22_technical_notes.pdf.

World Economic Forum. 2006. "Global Gender Gap Report 2006." http://www3.weforum.org/docs/WEF_GenderGap_Report_2006.pdf.

———. 2015. "Global Gender Gap Report 2015." https://reports.weforum.org/global-gender-gap-report-

2015/the-global-gender-gap-index-2015/.
Worsdale, Rosie, and Jack Wright. 2021. "My Objectivity Is Better Than Yours: Contextualising Debates about Gender Inequality." https://doi.org/10.1007/s11229-020-02835-5.
Yang, Mimi. 2020. "An Intimate Dialog Between Race and Gender at Women's Suffrage Centennial." *Humanities and Social Sciences Communications* 7. https://doi.org/https://doi.org/10.1057/s41599-020-00554-3.