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

Sara Cerioli   Andrey Formozov

Abstract

This study attempts to replicate and extend the results of the article of Falk & Hermle (2018a) measuring the gender differences economic preferences relating them to economic development and to gender equality of the countries. In the original paper, the authors use data from the Gallup World Poll 2012, which included a Global Preference Survey conducted on almost 80000 people in 76 countries all around the world. The dataset covers almost 90% of the world population representation, with each country having around 1000 participants answering questions related to their time preference (patience), altruism, will of risk taking, negative and positive reciprocity, and trust. The dataset is available in its integrity only with a license to be paid. The free version has only partial data that can also be used for this purpose because, according to the study, the gender differences can be studied also only taking in consideration a smaller number of predictors (according to the supplementary material, see (Falk & Hermle 2018b)). In this replication study, therefore, we use only a subset of predictors that are made publicly available to check whether the results can still be reproduced and are consistent. The outcome of the replication is that we see similar results as the ones obtained by the original authors for the relationship of gender differences and the economic development, but with differences (some times minor, some times significantly large) regarding the gender equality, especially when comparing the results of the single indexes building the general Gender Equality Index. Beyond the replication, we have extended the analysis using a robust linear regression instead of the OLS used by the authors, finding that the results are not robust under this change of model. Moreover, using a different measurement for the gender equality of the countries, the Gender Development Index from the UN Human Development reports, the correlation between gender differences and Gender Development of the country disappears, when a conditional analysis on the Economic Development is performed.

1. Introduction

Gender differences concerning the economic behaviors, such as 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. Any scientific knowledge about gender differences and their relation to gender inequality is used nowadays as arguments and counter-arguments for decision and policy-making. Gender-related issues such as gender inequality are becoming an integral part of the agenda for many public and private institutions and organizations. Therefore, it is essential for the stakeholders to reveal, estimate, monitor, and prevent gender inequalities on individual, group and country levels.

The study of gender differences on a world scale per se is challenging. One of the challenges that hampers the progress is the lack of large and heterogeneous datasets across different social groups and countries. The Gallup World Poll 2012 included a Global Preference Survey conducted on almost 80000 people in 76 countries all 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, will of risk-taking, negative and positive reciprocity, and trust. The dataset provides a unique insight into the economic preferences of a heterogeneous number of people.

The original study published in the Quarterly Journal of Economics (Falk *et al.* 2018) focused on more general questions about the economic preferences distributions in different countries, trying to explore several covariates from the Gallup World Poll. While, the subsequent article (Falk & Hermle 2018a), replicated in this work, focused explicitly on the gender differences arising from the previous study and reported the evidence for the relationship between gender differences in economic preferences, economic development, and gender equality across many countries. The authors propose two hypotheses to be tested: The first one is that the gender differences will decrease for more economically developed and gender-equal countries, because social roles related to gender are attenuated. The second hypothesis is that, on the contrary, for more economically developed and gender-equal countries, the gender differences will increase because the gender-neutral goal of subsistence is removed and thus people can pursue their real preferences. The data reveals a positive correlation of gender differences in preferences with GDP p/c and with the gender equality of the countries, thus favoring the hypothesis that predicts the increase of the differences as women and men obtain sufficient access to the resources to develop and express their intrinsic preferences independently.

Throughout the study, the authors construct and use a joint measure of gender equality of different countries, coming from different sources (see also the Appendix), providing little to no discussion about the reason in doing so. This is seen as a critical point, since some of the indicators used, such as WEF Global Gender Gap and UNDP Gender Inequality Index are far from being a complete representation of gender inequality, and were heavily criticized by several authors (Anand & Sen 1995; Permanyer 2013; Klasen 2017).

In this work, we first analyse the gender equality index that was used by the original authors and the sub-indexes involved into the study, highlighting its problematic. Second, we conduct a pure replication of the article in R language, challenged by the fact that most of the data were either not available anymore, or could not be provided for free (see Appendix, Section 2, "Data Collection, Cleaning, and Standardization"). Third, we extended the original article using a robust linear regression to address the non-normality of the data. Fourth, we demonstrate that the fundamental finding is not stable when trying to replace the gender equality index introduced by the authors with the Gender Development index proposed by Klasen (2017) and recently introduced by the UNDP.

2. Results

A composed Gender equality Index and related problematics

The authors in the original article compose a joint measure of gender equality that they denoted as Gender Equality Index (GEI). We visualized its composition using a diagram shown in Figure 1. Four measurables were used to compose this joint measure: Two of them are indexes officially approved by world organizations, the Global Gender Gap from the World Economic Forum Global Gender Gap Report 2015, and the Gender Inequality Index from the UNDP Human Development Report 2015; one widely used measurable, the ratio of female and male labor force participation, taken from the World Bank database; and lastly, a measurable newly constructed by the authors, the time since woman suffrage, taken from the Inter-Parliamentary Union Website, presumably to track low-term influences of the guaranteed right to vote as a proxy of gender equality. The WEF Global Gender Gap has a total of fourteen sub-indexes, grouped and weighted into four categories: economic participation and opportunity, political empowerment, educational attainment, health and survival. UNDP Gender Inequality Index follows a similar logic to cover the same categories describing different aspects of the human life, but using only five sub-indexes in total, that are two for health and reproduction-related issues, and three others for the remaining categories.

Both indexes were criticized by several authors (Anand & Sen 1995; Permanyer 2013; Klasen 2017). To
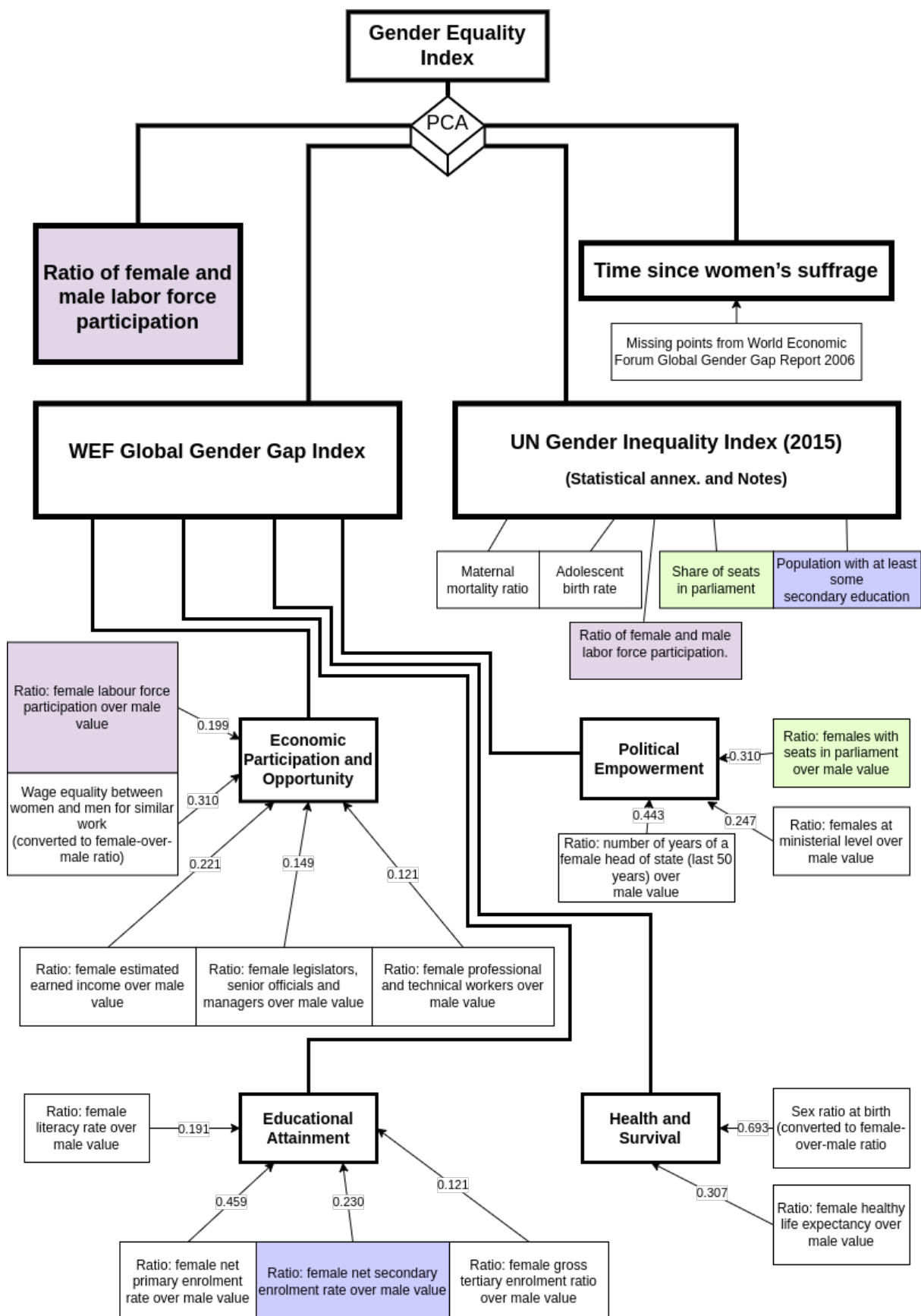
Figure 1: Gender Equality Index decomposed in its sub-indexes.

summarise some of the main critics to them: The Gender Inequality Index from the UNDP is said to be very highly related to the economic development, as it includes reproductive health indicators that can penalize less-developed countries, and it has a "inexistent" measure of welfare loss of inequality, because it is based on a calculated gender equality measure that is reported nowhere. One of the coefficients that has a highest priority (~30 % of the sub-index) in calculation of the economic participation and opportunity, "wage equality between men and women for similar work" sub-index (that indicates how much women receive more with respect to man for the same work) is a based on a best experts' guess, not corrected on the difference in occupation (Halkos *et al.* 2021).

The WEF Global Gender Gap has a total of fourteen sub-indexes (some of them being the same as for the UNDP GII) where a cap to 1 is applied for those countries where the ratio is higher than 1. This means that countries treating equally men and women are ranked the same as countries disfavoring men.

The ratio of female and male labor force participation counts as the overall labor participation rate, and thus, once approaching 100% as a characteristic of developed countries (Tasseven 2017), the gender gap is automatically reduced. This measure does not take into account on the part-time jobs occupation as in case of the reduction , including involuntary part time job occupation in developed countries (Rosenfeld & Kalleberg 1990; Pech *et al.* 2021).

Lastly, the "time since women suffrage" index was introduced by authors to track long-lasting effects of the right to vote. This is based on the assumption that during the course of the time development has always a positive effect and its magnitude is proportional to the time since women suffrage was established. Moreover, it can be argued that, even after the right to vote, many discriminating laws were still in presence and the alignment of law together with executive branch of the government and elimination of discrimination – for example, despite to the right to vote, the right to work can be suppressed for several decades. Worthy to mention also the question about suffrage and race, that are in many cases strongly connected (Carruthers & Wanamaker 2015; Yang 2020). Thus, that 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.

Another important aspect is that the comparison of the four indexes used to build the Gender Equality Index shows many repetitions in the datasets used, as one can see in Figure. 1. For instance, the two indexes from WEF and UNDP share three sub-indexes here indicated with different colors: ratio of female and male labor force participation (yellow), share of seats in parliament (green) and enrollment into secondary education (blue). As a third variable to construct their Gender Equality Index, the authors also used the ratio of female and male labor force participation, already included into the previous two indexes. Such a procedure leads to the imbalance and prioritization of these indexes over other factors that needs to be justified.

Pure replication and extended analysis

In Table 1, we summarise the results of the conditional analysis. We indicate the significance level for each correlation using the following scheme:

Significance ≤ 0.001 (***), ≤ 0.01 (**), ≤ 0.05 (*)

Table 1: Conditional analysis to separate the impacts of economic development and gender equality on gender differences in economic preferences. Reported are the slopes of the linear regressions.

Variable	Residualized on	Original	Replication	Extended
Log GDP p/c	Gender Equality Index	0.5258***	0.4956***	0.4862***
Gender Equality Index	Log GDP p/c	0.3192***	0.3483***	0.3432**
WEF Global Gender Gap	Log GDP p/c	0.2327**	0.2172*	0.2106*
UN Gender Equality Index	Log GDP p/c	0.2911	0.3057	0.3017
F/M in Labor Force Participation	Log GDP p/c	0.2453*	0.2249*	0.2034*
Years since Women Suffrage	Log GDP p/c	0.2988**	0.1851*	0.1929*

Variable	Residualized on	Original	Replication	Extended
Log GDP p/c	Gender Development Index	–	0.6271***	0.6612***
Gender Development Index	Log GDP p/c	–	0.1052	0.0353

Since the correlation between Log GDP p/c and Gender Index is not negligible (see Figure. 2), we can't ignore the effect of the one on the other. Thus, to understand the effect of the gender differences regressed on economic development, we perform every regression conditioning on the gender equality, and vice-versa. The theorem from Frisch–Waugh–Lovell (Frisch & Waugh 1933)(Lovell 1963) guarantees us that the coefficients found from this conditional analysis are the same as the ones found for a regression of the gender differences on both economic development and gender equality index of the countries.
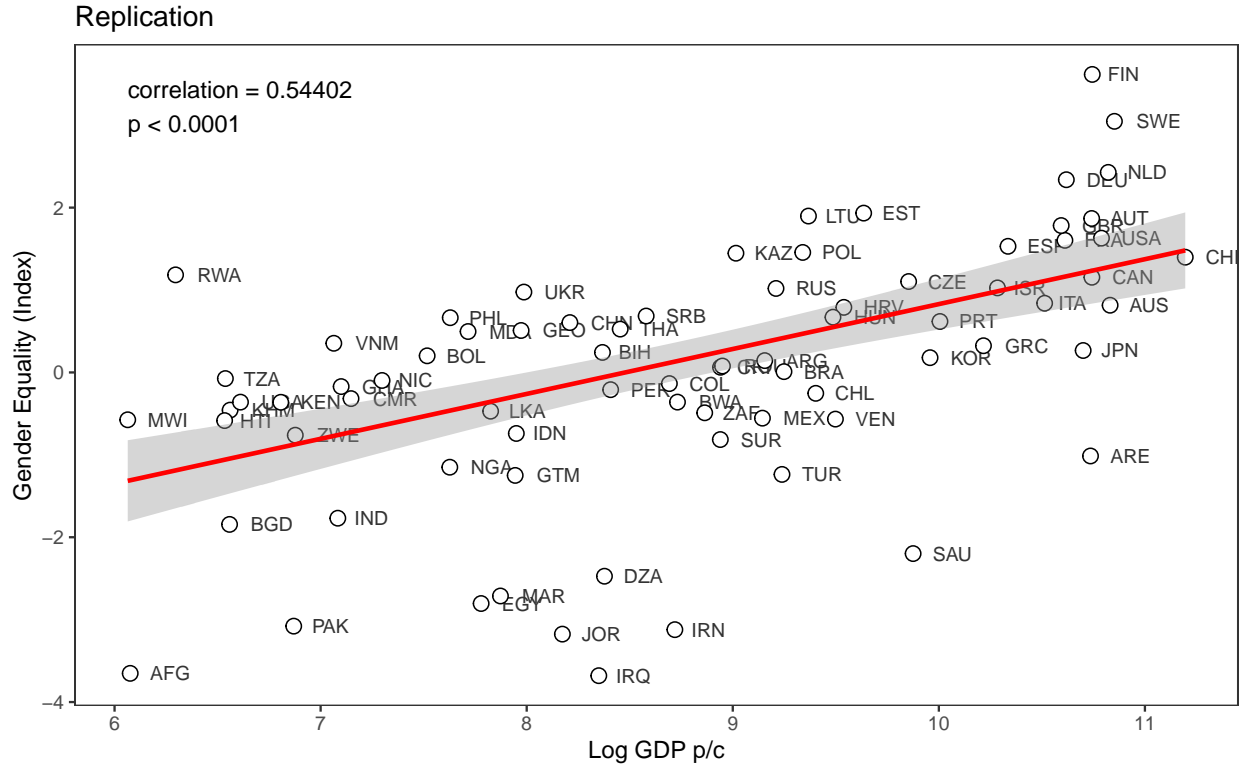

Figure 2: Gender Equality Index and Economic Development by country.

For the most of the variables in Table 1, we see that the replication values tend to agree and be on the same direction (similar slope coefficients and significant p-value) as the original ones. Regarding the indicator "years since women suffrage", we had to make choices on how to impute data and also how to handle the missing data (see discussion above in paragraph "Methods"), and here it is where we see a substantial difference in the results.

The most interesting part of the analysis arises from the use of the Gender Development Index in place of the Gender Equality index built by the authors. When the variable conditioning analysis regressing on the Log GDP p/c is done, the correlation between gender differences and GDI vanishes, with slope = 0.1052, p-value = 0.294 for the analysis performed using OLS, and slope = 0.0353 with p-value = 0.735 for the analysis performed using robust linear regression (Fig. 3).

As one can see in Table 2, the economic development still has a strong correlation to the single preferences when regressed on the gender equality of the country. Here we report only the conditional analysis made
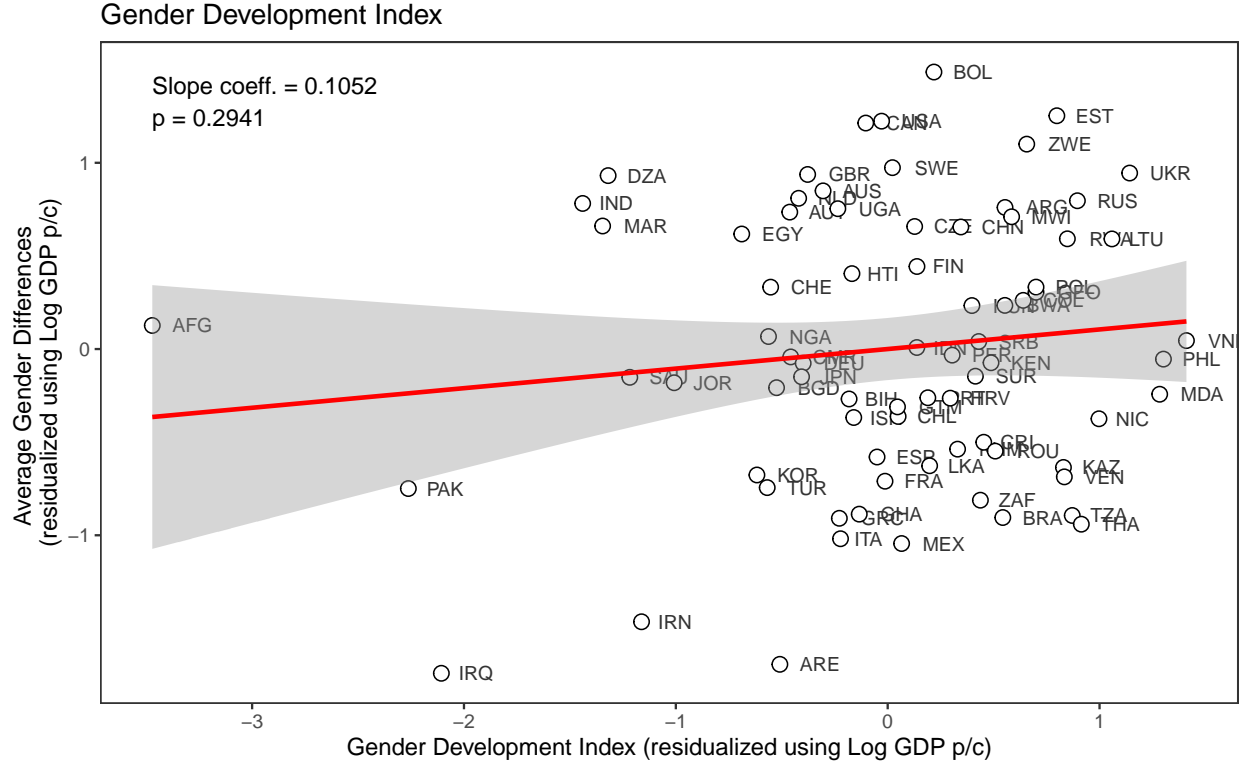
Gender Development Index

Figure 3: Gender differences and gender development conditional on economic development by country.

using the Gender Equality Index and the Gender Development Index, as a comparison (see also Fig. S5 in the Falk & Hermle (2018b)).

Table 2: Log GDP p/c conditional on Gender Equality Index and on the Gender Development Index, regressed on each single economic preference gender difference coefficient. Here reported are the correlation terms for the original analysis, the replication (using OLS) and the extended analysis (robust linear regression).

Variable	Regressed on	Trust	Altruism	Pos. Rec.	Neg. Rec.	Risk Taking	Patience
Log GDP p/c	GEI (Original)	0.4574***	0.4751***	0.2771*	0.2444*	0.2868*	0.2621*
Log GDP p/c	GEI (Replication)	0.4265***	0.4338***	0.2509	0.2111	0.2256*	0.2288*
Log GDP p/c	GEI (Extended)	0.4450***	0.3959***	0.2524	0.2451*	0.2191*	0.2444*
Log GDP p/c	GDI (Replication)	0.4868***	0.4762***	0.1820	0.3532**	0.3212**	0.3049**
Log GDP p/c	GDI (Extended)	0.5045***	0.4477***	0.1870	0.3858***	0.3200**	0.3337**

To further check that no correlation was found for the Gender Development Index when conditioned on economic development, we regressed the single preferences on it and compared the results of the joint Gender Equality Index (for original, replicated, and extended analysis) and of the Gender Development Index (replicated and extended analysis), when regressed on Log GDP p/c (see Table 3).

Table 3: Gender Equality Index and the Gender Development Index, conditional on Log GDP p/c, regressed on each single economic preference gender difference coefficient. Here reported are the correlation terms for the original analysis, the replication (using OLS) and the extended analysis (robust linear regression).

Variable	Regressed on	Trust	Altruism	Pos. Rec.	Neg. Rec.	Risk Taking	Patience
GEI (Original)	Log GDP p/c	0.2050	0.3304**	-0.0115	0.2788*	0.1973	0.2967*
GEI (Replication)	Log GDP p/c	0.2472*	0.2696*	0.0481	0.2240*	0.1863*	0.2841*
GEI (Extended)	Log GDP p/c	0.2521*	0.2401	0.051	0.2037*	0.1868*	0.2781*
GDI (Replication)	Log GDP p/c	-0.0127	-0.0325	-0.1603	0.1292	0.0659	0.0758
GDI (Extended)	Log GDP p/c	0.0193	-0.0104	-0.1454	0.1049	0.1147	0.0938

Some bla bla to describe the table here. Do we need to add the economic development table too? Or maybe instead having some plots?

Looking at these results, we can deduce that:

1. For countries having the same gender development, the gender differences are depending from the economic development of the country, meaning that richer countries have more gender differences than poorer countries, gender development being equal;

2. For countries having the same economic development, the gender differences are independent from the gender development of the country, meaning that there are no gender differences arising from countries having same economics but different gender development;

From this, one could therefore assume that the economic development is the country-level indicator associated with higher changes in gender differences, rather than the gender development of a country. From such a simple analysis is therefore not possible to extract any information regarding the reason for higher gender differences in more economically developed countries. We can only see that, for countries with similar economics, those differences don't exist regardless of how gender developed that countries are. The reason behind the differences might be related to purely economical conditions, that would also make sense since the gender differences here studied are "economical preferences". Can it be that the differences arise where the economic is more developed because of marketing reasons? After all, all the economics here analysed are based on capitalistic systems, meaning that the bigger the economic, the bigger the market. It is an interesting hypothesis that would require further analysis.

3. Discussion and conclusions

The original study indicates that higher economic development and higher gender equality are associated with an increase in the gender differences in preferences, and therefore rules out the social-role theory over the post-materialistic one: When more resources are available to both men and women, the expression of the gender specific preferences can be seen. Our replication leads to the same conclusions, but we have different results in terms of statistical significance of the single indixes when conditioned on economic development, and there are some open questions regarding unexplained differences that lead to further checks on the results' robustness. Moreover, we prove that, using a different indicator for the gender inequality in different countries, the result found by Falk & Hermle (2018a) does not hold anymore.

We decided to focus on two main checks for the robustness: The first was to change the model to have a

more "relaxed" assumption for the linear regression, and the choice went to robust linear regression. The results are similar to the ones from the replication analysis, meaning that the model is robust enough.

The second check has been the introduction of a different gender equality indicator, the Gender Development Index (from the UNDP). We performed the same analysis done so far, and we could see some differences rising from it:

TO BE REWRITTEN:

1. When conditional analysis is perfomed on gender equality, the results don't hold

2. The single preferences are not correlated with GDI (or with GEI) but only with economic development

3. In fact, the whole assumption is that economic development is causing gender equality, but it is true the contrary. There is a causal correlation, and not taking this into account (and the fact that the two indicators are intertwined) is wrong and should be corrected only by controlling the two variables.

References

Anand, S. & Sen, A. (1995). Gender inequality in human development: Theories and measurement. *Occasional Papers 19, Human Development Report Office, New York.*

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.

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.

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.

Falk, A. & Hermle, J. (2018a). Relationship of gender differences in preferences to economic development and gender equality. *Science*, **362**.

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., Kenneth, L.L. & List, J.A. (2009). Gender differences in competition: Evidence from a matrilineal and a patriarchal society. *Econometrica*, **77**, 1637–1664.

Halkos, G., Alba, J.M. de & Todorov, V. (2021). Economies' inclusive and green industrial performance: An evidence based proposed index. *Journal of Cleaner Production*, **279**, 123516.

Klasen, S. (2017). UNDP's gender-related measures: Current problems and proposals for fixing them.

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.

Lovell, M.C. (1963). Seasonal adjustment of economic time series and multiple regression analysis. *Journal of the American Statistical Association*, **58**, 993–1010.

Pech, C., Klainot-Hess, E. & Norris, D. (2021). Part-time by gender, not choice: The gender gap in involuntary part-time work. *Sociological Perspectives*, **64**, 280–300.

Permanyer, I. (2013). Are undp indices appropriate to capture gender inequalities in europe? *Social Indicators Research*, **110**, 927–950.

Rosenfeld, R.A. & Kalleberg, A.L. (1990). A cross-national comparison of the gender gap in income. *American Journal of Sociology*, **96**, 69–106.

Schneider, U., Pfarr, C., Schneider, B.S. & Ulrich, V. (2012). I feel good! Gender differences and reporting heterogeneity in self-assessed health. *European Journal of Health Economics*, 251–265.

Tasseven, O. (2017). The relationship between economic development and female labor force participation rate: A panel data analysis. *Global financial crisis and its ramifications on capital markets*, pp. 555–568. Springer.

Yang, M. (2020). An intimate dialog between race and gender at women's suffrage centennial. *Humanities and Social Sciences Communications*, **7**.