The Effect of Fertility on Women's Labor Supply: Heterogeneity by Gender Norms

Elisa Taveras*
Binghamton University

December 2, 2021

Abstract

This paper asks whether the effect of fertility on women's labor supply depends on gender norms. To separate the role of gender norms from institutional features, I compare the labor supply response to children among women living in the United States but born in different countries. Hence, I compare native to immigrant women and, within immigrants, those born in less and more gender-egalitarian countries. I instrument for the variable of interest, having more than two children, with the sex composition of the first two children. The findings show that women from all countries reduce employment due to having more than two children. Yet, this effect is substantially larger for women born in less gender-egalitarian countries. In particular, women from countries with the least egalitarian gender norms have an employment response three times larger than the employment response of natives. Thus, the negative effect of fertility on the labor supply decreases with gender egalitarianism.

Keywords: Gender norms, Employment, Fertility, Immigrants

JEL Classification Numbers: J13, J15, J16, J22

*Email: etavera2@binghamton.edu. I am grateful for comments from Solomon W. Polachek and David Slichter. All remaining errors are my own.

1 Introduction

The literature studying the effect of fertility finds a negative effect of children on women's labor market outcomes (e.g. Polachek (1975), Rosenzweig & Wolpin (1980), Angrist & Evans (1998), Cristia (2008), Lundborg *et al.* (2017)). Time allocated to childrearing expected from mothers relates to employment interruptions, affecting human capital investment, on-the-job training, and cumulative work experience (Mincer & Polachek (1974)). At the same time, attitudes toward the role women in the family and labor market can affect women's fertility and labor supply choices (e.g. Antecol (2000),Fernández (2007), Blau *et al.* (2011), Fernandez & Fogli (2009), Blau *et al.* (2013),Neuman (2018)). Furthermore, using Akerlof & Kranton (2000) definition of individuals' identity being part of a social category, we can think of within-gender identity as one way to categorize women. In this context, women in different gender norms categories might respond to fertility differently to meet their category's prescribed behavior.[1] Childcare responsibility would skew more toward women in the presence of less gender-egalitarian norms. At the same time, the expectation that women would follow a more traditional role in society would limit their labor market participation and more so after having children. Therefore, women's "gender norms" type would affect how responsive their employment is to fertility changes.

In this paper, I ask to what extent the effect of fertility on female labor supply depends on gender norms. To do so, I use the American Community Survey (ACS) 2006-2018 and estimate the effects of fertility using the Angrist & Evans (1998) sex ratio instrument variable (IV) approach. I use variation in the sex composition of the first two children to instrument for the probability of having *more than two* children. To look at the differences by gender views, I group women between those born in the United States (US) (natives) and those born outside the US (immigrants). For immigrants, I match their country of origin with their gender gap measured by the World Economic Forum's Global Gender Gap Index (GGI). Thus, the GGI measures immigrants' country of origin gender norms. Using the sample GGI's mean, I group immigrant women between those with *low GGI* and *high GGI*. Consequently, the groups represent different levels of gender attitudes. This paper will use the terms "gender norms" and "gender attitudes" interchangeably.

My main findings are as follows. First, all women reduce employment in response to having more than two children by 6.86 percentage points (hereafter, *pp*). Second, the less egalitarian gender attitudes, the more negative the effect is. Native women's employment decreases the least by having more than two children (-5.61 *pp*). At the same time, the effect is about twice as large for immigrants (-10.34 *pp*). Finally, within immigrants, *High GGI* women have somewhat comparable point estimation than natives (-7.12 *pp*)

[1] Bertrand *et al.* (2015) uses this framework to study the "man" and "woman" categories prescribed behavior that a man's income should be higher than his wife's.

although not significant, while *Low GGI* women have a negative impact of more than three times that natives (-17.38 *pp*). Therefore, the more egalitarian gender attitudes are, the least responsive women's labor supply is to fertility changes.

I then look at two additional analyses. First, I explore if children also reduce other labor market outcomes. I find a similar pattern in the effect of children on hours of work. In contrast, there is only a marginally significant negative effect for *Low GGI* immigrants in terms of earnings. Thus, fertility mainly determines the extent of women's employment. The second analysis concentrates on immigrants, replicating the analysis by age at migration. Assuming that the age at migration would affect women's cultural assimilation, the effect of children on employment should be smaller for women that moved to the US when they were younger. My findings support this assumption, with a smaller effect for women who migrated when less than 20 years old.

The remainder of the paper concentrates on looking at the robustness of the main results. First, I explore the heterogeneity of the main results by educational attainment. I find that the effect of more than two children on employment varies more for women with lower educational attainment. Furthermore, since the IV estimation shows the results for the compliers, I explore the characteristics of the compliers and find similar characteristics across groups with a small imbalance in educational attainment. Additionally, I look at the sensitivity to the sample selection by first replicating the analysis for the subgroup of 21-35 years old's women and, second, estimating the natives and immigrants results using the 1% 1980 and 1990 Census. In addition, I analyze the sensitivity of the results to controlling for family characteristics.

After that, I concentrate on immigrants, first exploring a different rule to partition immigrants based on their GGI. Second, I control for the educational attainment distribution of immigrants relative to their country of origin's education. Next, I explore whether the different effects across GGI categories hold even when controlling for other countries' characteristics. Finally, I analyze if children's sex preference plays a role in the main finding by looking at the country of origin's sex ratio. Across all these robustness checks, the results are qualitatively similar to the main results and consistent with the effect being due to gender attitudes.

This paper builds on the literature looking at the negative effect of fertility on labor market outcomes. Within this literature, this paper especially relates to those that examine the effect of fertility across countries. In this literature, Aaronson *et al.* (2021) finds that the negative effect of fertility in women's labor supply appears as countries' economic development increases. Papers concentrating across developed countries find that the negative effect of fertility on earnings (Harkness & Waldfogel (1999), Gangl & Ziefle (2009)) and employment (Gutiérrez-Domènech (2005)) varies by countries, as well as the mechanism (Gangl & Ziefle (2009)) and it's long-term effect (Kleven *et al.* (2019a)).

The paper also speaks to the stream of research that explores the effect of gender

norms on labor supply. Francesco Giavazzi *et al.* (2013) compare across countries and find that culture affects women's employment rate, even after controlling for countries' policy structure. Muller *et al.* (2020) finds that in European countries with high female labor force participation, women's labor force participation across different family trajectories converges to each other. Since policy structure varies by country, many papers limit one country and look at cultural variation through immigrants' source country. These papers find that the country of origin's labor force participation has a positive relationship with immigrants' labor force participation (Antecol (2000),Fernández (2007), Blau *et al.* (2011), and Neuman (2018) in the context of Sweden). The effect of culture is smaller but persists for second generations of US-born women (Fernández (2007), Fernandez & Fogli (2009)). Blau & Kahn (2015) builts on these findings, showing that the effect of country of origin's labor force participation on immigrant women's US labor force participation persists even after controlling for pre-migration labor market information. In terms of assimilation, Blau *et al.* (2011) finds that as immigrant women spend more time in the US, the natives-immigrants labor supply gap decreases, even closing for women from high labor supply countries. Furthermore, for fertility, papers find that the country of origin's fertility rate positively relates with immigrants' fertility rate (Cygan-Rehm (2014) in the context of Germany; second generation US-born: Fernandez & Fogli (2009), Blau *et al.* (2013)). Finally, using the GGI to measure country of origin gender norms, papers find it affects the equality of households' distribution of nonmarket work (Blau *et al.* (2020a), Marcén, Miriam; Morales (2019)).

This paper contributes to the literature in a few ways. First, the paper explores the gender attitudes' fertility gradient on labor supply. Although the literature looks at culture as a determinant of fertility and labor market outcomes, it does not consider that cultural variation creates heterogeneity in the sensitivity of employment to fertility. Second, I analyze the cultural gradient for women within a given country. In contrast to across countries analyses in which family policies and gender norms vary, limiting women living in a country keeps family policies constant. Thus, it isolates gender norms' heterogeneity in employment changes due to fertility changes. Finally, the paper finds a natives-immigrants gap and a within-immigrant gap in the effect of children on employment. Thus, it is clear that looking at the complete population's fertility effect mask meaningful differences. Understanding how the underlying population respond to fertility changes is relevant to how this would affect the impact of family policies on improving women's labor supply.

The paper follows the following order. First, Section 2 presents the description of the main sample. Section 3 presents the empirical strategy and the results. The robustness checks follow in Section 4. Lastly, Section 5 concludes the paper.

2 Data

To implement my analysis, I use the 2006–2018 ACS.[2] The sample consist of those women who are the head or spouse of the household. I match women to their children, defined as those labeled "child" in the household's relationship to the household head. For households with multiple families, I use the subfamily relationship in the same way. From my sample, I exclude children who are part of a "multiple birth," measured as children with the same age and quarter of birth within the family. I limit the sample to the women that are 21–40 years old and that, after excluding multiple birth children, their number of children match their reported number of own children in the household. Finally, I keep only those women who had their first child when they were 15 years or older, that the oldest child is 18 years or younger, and, finally, have two or more children.

To explore the gender norms gradient, I group women between natives and immigrants, based on being born in the US/outlying area or outside of the US. Furthermore, I exploit immigrants' gender attitudes variation due to differences in the country of origin's gender attitudes, measured by the GGI.

GGI.– The World Economic Forum creates the GGI, a comparable gender gap measure across countries.[3] The GGI index measures the gender gap in a broad sense since it considers various areas where women might be falling short. Precisely, the GGI base the gender gap on four subindexes: (1) Economic Participation and Opportunity; (2) Educational Attainment; (3) Health and Survival; (4) Political Empowerment. Each subindex is a weighted average of various indicators that measure gender equality in each subindex category. For example, the *Economic Participation and Opportunity* subindex includes, among other measures, the labor force participation rate, wage equality in similar jobs, and participation in specific occupations like legislators and managers. The final GGI index is a simple average of the four subindexes. The GGI index goes from 0 (gender imparity) to 1 (gender parity).[4] Moreover, one key advantage of the GGI is that since it measures gender gaps, it disassociates from countries' development.

When available, I assign the GGI to immigrants based on their country of birth.[5] Using the GGI, I group the immigrants based on their relative gender egalitarianism, using the average GGI as the reference level. Thus, I define four groups: (1) natives, (2) immigrants, (3) *LowGGI*: immigrants with GGI just or below the mean (relative gender imparity), and (4) *HighGGI*: immigrants with GGI above the mean (relative gender

[2]Ruggles *et al.* (2021)

[3]Hausmann, R., Tyson, L. D., & Zahidi (2006)

[4]For more detail in the GGI, see https://reports.weforum.org/global-gender-gap-report-2018/the-global-gender-gap-index-2018/.

[5]In a few cases, I assigned the GGI as follows. For Taiwan, Hong Kong and Macao, I assign China's GGI. For Czechoslovakia, I took the average from the Czech Republic and the Slovak republic. For Yugoslavia, I take the average of Croatia, Montenegro, Serbia, Bosnia Herzegovina, and Slovenia.

parity).

Table 1: Summary Statistics

	(1)	(2)	(3)	(4)	(5)
				Immigrants	
	All	Natives	Immigrants	*LowGGI*	*HighGGI*
Number of Children	2.52	2.52	2.54	2.55	2.52
	(0.81)	(0.81)	(0.82)	(0.81)	(0.80)
More than Two	0.37	0.37	0.39	0.40	0.38
	(0.48)	(0.48)	(0.49)	(0.49)	(0.48)
1st Born Boy	0.51	0.51	0.51	0.51	0.52
	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)
2nd Born Boy	0.51	0.51	0.51	0.51	0.51
	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)
Two Boys	0.26	0.26	0.26	0.26	0.27
	(0.44)	(0.44)	(0.44)	(0.44)	(0.44)
Two Girls	0.24	0.24	0.24	0.24	0.24
	(0.43)	(0.43)	(0.43)	(0.43)	(0.43)
Same Sex	0.50	0.50	0.50	0.50	0.50
	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)
Age	33.59	33.46	34.13	33.89	34.39
	(4.55)	(4.59)	(4.33)	(4.40)	(4.23)
Age at 1st Child	23.89	23.83	24.15	23.92	24.35
	(4.52)	(4.52)	(4.51)	(4.41)	(4.61)
Employed	0.63	0.66	0.52	0.47	0.56
	(0.48)	(0.47)	(0.50)	(0.50)	(0.50)
GGI			0.68	0.65	0.71
			(0.04)	(0.03)	(0.03)
N	1,400,784	1,125,822	274,962	127,313	131,970

Note.– Sample means with standard deviations in parentheses of the main variables. Columns: *All:* include all the sample; *Natives:* include those born in the US (including US outlying territory); *Immigrants:* those born outside from the US; *LowGGI:* those born outside of the US, that have GGI information, and GGI is below the average value; *HighGGI:* those born outside of the US, that have GGI information, and GGI is above the average value. Rows: *Number of Children:* number of children that the women has; *More than 2 Children:* equal to 1 if women has more than 2 children and 0 otherwise; *1st Born Boy:* equal to 1 if first born child is a boy; *2nd Born Boy:* equal to 1 if second born child is a boy; *Two Boys:* equal to 1 if first two children are boy; *Two Girls:* equal to 1 if first two children are girl; *Same Sex:* equal to 1 if first two children have same sex; *Age at 1st Child:* women's age at the birth of their first child; *Employment:* equal to 1 if their employment status is employed; *GGI:* average immigrants country of origin's gender gap index

GGI and Gender Attitude Measures.– One possible concern with the GGI is whether it correlates with individual's gender attitudes. To explore this, I compare the GGI with gender attitudes constructed from the General Social Survey (GSS) and

Joint European Value Survey and World Values Survey (EVS/WVS). To do so, I use individuals' answers to a series of questions regarding gender attitudes and create the average by country of origin. I find that the GGI correlates with these alternative gender attitudes measures, easing concerns regarding the GGI as a measure of women's gender attitudes. Section A.1 discuss this analysis in more detail.

The summary statistics are available in Table 1. In the table, the complete sample information is in column (1). Out of the entire sample, 80% are born within the US territory, natives, and their summary statistics are in columns (2). The rest of the sample are immigrants, in column (3). From those that are immigrants, 94% have GGI information and are split 49% with *LowGGI*, column (4), and 51% with *HighGGI* in column (5). Table 1 rows present both the characteristics of the women and their children composition.

The information in terms of the number of children is similar across the groups, with a slightly higher number of children for immigrants, especially for *LowGGI* women. This higher number of children translates to a higher percentage of women with more than two children for those immigrants. Nevertheless, these characteristics are not statistically different across groups of women. Thus the percentage of women with more than two children is comparable across different levels of gender attitudes.

In terms of their children's gender composition, just above half of the sample has a first and second boy, with 26% of the population with two boys and 24% two girls, resulting in half of the sample having the first two children of the same sex. Importantly, since the identification strategy uses the sex composition of the first two children, the percentage of women with the first two children of the same sex is constant across groups.

In the sample, women are around 34 years of age and had their first children when they were around 24 years old. Moreover, more than half of these women are employed. However, when comparing across groups, there is significant variation in terms of employment. Columns (2) and (3) show that natives women are 14 *pp* more likely to be employed than immigrants. Within immigrants, women with *HighGGI* are nine *pp* more likely to be employed than those with *LowGGI* countries. Finally, comparing natives with *LowGGI* and *HighGGI*, the differences are 19 and 10 *pp*, respectively. Therefore, immigrant women are less likely to be employed when compared to natives. Nevertheless, this difference varies by gender attitudes, with immigrant women with relatively more egalitarian gender views being more comparable with natives.

To compare immigrants with natives is important to understand gender attitudes in the US. In 2018,[6] the US's GGI was 0.720, high enough to occupy the 51st rank globally. To put this ranking in perspective, the 1st place went to Iceland with a GGI of 0.858, while the lowest rank was 149th place, corresponds to Yemen with a GGI of 0.499. The

[6]The 2018 World Economic Forum's Global Gender Gap Report is available here: `https://reports.weforum.org/global-gender-gap-report-2018/results-and-analysis/`

median rank, the 74th-75th place, was shared by the Dominican Republic and the Russian Federation with a GGI of 0.701. The US's GGI is not only relatively high in the GGI's ranking, but it is also significantly above average for my sample of immigrants, 0.6794. This immigrants' average fell between the 97th (Azerbaijan) and 98th (Armenia) place. Finally, note that the average GGI for those from *HighGGI* countries is closer to the US's at 0.712, which aligns with *HighGGI* women being closer in characteristics to natives.

From the data, we can see that although all women have a similar number of children, the likelihood of being employed varies significantly with gender attitudes. Women with more egalitarian gender attitudes work for pay at a higher rate than those with less so. This pattern suggests that women with less traditional gender attitudes see motherhood as less incompatible with a career, instead decide on labor force participation based on their preference. Thus, their labor supply should be less affected by changes in fertility. On the contrary, women with more traditional gender attitudes would feel more pressure to exit the labor force due to changes in fertility. As a consequence, the effect of children should be larger for women with less egalitarian gender attitudes. In the next section, I empirically explore this by studying the effect of children on employment for women in each group defined above.

3 Analysis

The goal of this analysis is to see the effect of children on women's employment. Due to women's making simultaneous decisions in terms of fertility and the labor market, there are endogeneity concerns with the correlation using OLS estimation. This section discusses the identification strategy to overcome this, followed by applying this strategy to the data. Then, I implement additional analyses that are relevant to the question.

3.1 Identification Strategy

The regression of interest is

$$Employed_i = \beta_0 + \beta_1' MorethanTwo_i + \beta_2' X_i + \epsilon_i, \tag{1}$$

where the covariates in X_i are the women's age, age at first child, Black, Hispanic and other races dummies (white as a reference group), dummies for 1st Born Boy and 2nd Born Boy.[7] I also include the survey year and state dummies. $Employed_i$ is a dummy variable equal to 1 if women work for pay and 0 otherwise. The variable of interest is the dummy $MorethanTwo_i$, which is equal to 1 if more than two children and 0 otherwise.

[7]These variables match the included in Angrist & Evans (1998).

Since women decide the number of children to have while choosing their labor supply, $MorethanTwo_i$ is endogeneous. To overcome $MorethanTwo_i$ endogeneity, I implement a Two-Stage Least Squares (2SLS) estimation using the Angrist & Evans (1998)'s sex composition instrument variable (IV). This IV is motivated by parents preferring mixed-sex children. This preference leads parents with two children of the same sex to be more likely to have a third child in the hope of an opposite sex's child. Thus, I exploit the variation in the sex composition of the first two children as an instrument for having more than two children.

In practice,the first stage is estimating $MorethanTwo_i$ by

$$MorethanTwo_i = \gamma_0 + \gamma_1^{'} SameSex_i + \gamma_2^{'} X_i + \nu_i, \tag{2}$$

where the IV $SameSex_i$ is a dummy equal to 1 if the first two children have the same sex.

First Stage.– To see how relevant this instrument is, the results of estimating equation (2) are available in Table A.1, for the complete sample in column (1) and the four groups in columns (2)–(5). Overall, women whose first two children are of the same sex are 5.21 *pp* more likely to have more than two children. The effect is similar across all the groups, quantitatively being the lowest for *LowGGI* women. Moreover, the effect is very significant, and there is a strong *F-statistics* across all samples, suggesting that the instrument is relevant across all the groups. With the first stage results being strong and comparable across groups, I show the second stage results next.

3.2 Results

The correlation (OLS estimation) and the 2SLS results of equation (1) are in Table 2. Panel A presents the OLS estimation, while Panel B presents the 2SLS second stage estimation. In the following order, the columns present the results for the entire sample, natives, immigrants, *LowGGI* immigrants, and *HighGGI* immigrants.

The hypothesis suggests that, for women with relatively more egalitarian gender attitudes, having more children would affect their likelihood of being employed less than for women with more gender unequal attitudes. Thus, the impact of children on women's employment should be less negative for natives than for immigrants and, within immigrants, should be more negative for *LowGGI* women. *LowGGI* should have the largest negative effect across all women. Moreover, given similar gender views for *HighGGI* immigrants and natives, *HighGGI* and natives should have an approximately similar effect.

Table 2: Effect of Having More Than Two Children on Women's Employment Status

	(1)	(2)	(3)	(4)	(5)
				Immigrants	
	All	Natives	Immigrants	*LowGGI*	*HighGGI*
A: OLS					
MorethanTwo	-0.1410***	-0.1395***	-0.1407***	-0.1392***	-0.1397***
	(0.0044)	(0.0012)	(0.0089)	(0.0099)	(0.0091)
B: 2SLS					
MorethanTwo	-0.0686***	-0.0561***	-0.1034**	-0.1738**	-0.0712
	(0.0209)	(0.0209)	(0.0498)	(0.0837)	(0.0464)
N	1,400,784	1,125,822	274,962	127,313	131,970

Note.– OLS and 2SLS estimated coefficient of *MorethanTwo* in equation (1). Covariates in the model but excluded from the table for brevity: age, age at first child, Black, Hispanic and Other races dummies (white as a reference group), 1st Born Boy, 2nd Born Boy, year and states dummies. Columns: *All:* include all the sample; *Natives:* include those born in the US (including US outlying territory); *Immigrants:* those born outside from the US; *LowGGI:* those born outside of the US, that have GGI information, and GGI is below the average value; *HighGGI:* those born outside of the US, that have GGI information, and GGI is above the average value. Standard errors clustered at birthplace in parentheses, except for natives in column (2); * p<0.10, ** p<0.05, *** p<0.01.

OLS estimation.– As we can see in the OLS estimation in Panel A, the correlation between having more than two children and being employed is 14.10 *pp* for the complete sample. The results across gender attitude groups are very similar and not significantly different from each other. The results for natives in column (2) and immigrants in column (3) are about the same, being -0.12 *pp* more negative for immigrants. When looking at the *LowGGI* and *HighGGI*, the difference is -0.05 *pp* more negative for *HighGGI*.

From these OLS results, there seem to be no meaningful differences across groups. Thus we might conclude that the effect of having *MorethanTwo* children on employment does not vary across gender attitudes. However, due to the OLS endogeneity, we can't conclude that the causal relationship between *MorethanTwo* and employment is similar across different levels of gender attitudes. Perhaps these similarities reflect differences in endogeneity that obscure the true relationship.

IV Results.– The 2SLS results in Panel B show a different story. We can see that after controlling for endogeneity, the effect of *MorethanTwo* children varies across the different groups. Looking at the complete sample, having more than two children decreases the likelihood of employment by 6.86 *pp*. Compared with the OLS results, the effect of *MorethanTwo* decreases by 7.24 *pp*, with endogeneity overestimating the effect

of childbearing.

The effect for the complete sample mask important variations. Looking at natives in column (2), we can see that the resulting number is even smaller at 5.61 *pp*, decreasing 8.34 *pp* compared with its OLS result. Similarly, for immigrants in column (3), the effect decreases but not as much as for natives (-3.73 *pp*) to an effect of having *MorethanTwo* children on decreasing employment by 10.34 *pp*. Moreover, within immigrants, *LowGGI* (*HighGGI*) women in column (4) (column(5)) shows that having *MorethanTwo* children decreases employment by 17.38 *pp* (7.12 *pp*, not statistically significant). These effect are 3.46 *pp* more negative than their correlation result for *LowGGI*, while the *HighGGI* decreases 6.85 *pp*. Therefore, after controlling for endogeneity, the effect decreases for all but *LowGGI* women.

Now that we have compared the IV results with the OLS results, we can further compare the IV results across groups. This comparison is important to understand how the results reflect the groups' gender attitudes, as hypothesized previously.

Natives versus Immigrants.– Since the US has more egalitarian gender attitudes than the average immigrant's country of origin, children should less negatively affect employment for natives (column (2)) than immigrants (column (3)). As expected, natives women fair better when having *MorethanTwo* children. In the 2SLS results, *MorethanTwo* children decrease employment by almost twice as much for immigrant women than for natives born women. Precisely, the effect is more negative for immigrants than natives by -4.73 *pp* . Based on the means in Table 1, this difference represents 9.10% of the percentage of employed immigrants (0.52). In other words, the -4.73 differences represent about a third of the employment gap (14 *pp*) between natives and immigrants.

Variation Within Immigrants.– Since immigrants are not homogeneous, it is relevant to compare with immigrants by gender attitudes. For *LowGGI* and *HighGGI* women, the expected relationship is that the effect is more negative for *LowGGI* due to having more traditional gender views than *HighGGI*. The results show this relationship when comparing columns (4) and (5). Precisely, having *MorethanTwo* children decreases the likelihood of being employed by 10.26 *pp* more for *LowGGI* than for *HighGGI*.

Compared with natives, the effect of *HighGGI* in column (5) is not significant, but quantitatively the effect is very closely similar to the natives results, with a differences of just -1.51 *pp*. On the contrary, the gap between *LowGGI* and natives results is the largest at -11.77 *pp*. As we saw Table 1, the natives to *LowGGI* employment gap is -19 *pp*. Therefore, *MorethanTwo* gap represents around 0.6 of the total employment gap.

When having more children, women's employment is more at the margin for those women with more traditional gender attitudes. These traditional gender views tend to center women as the primary children's caregivers. Having more children will pressure

them to function in that expected role and decrease their labor market participation. Thus, the results reflect the expected pattern of women's employment decreasing with children depending on their gender attitudes.

The effect is the smallest for natives women and reasonably close to *HighGGI* women, whose average GGI is very close to the US's. At the same time, there is a large gap between native and *LowGGI*'s, with *LowGGI* women being from countries with significantly more traditional gender views than the US. Hence, since GGI reflects gender attitudes, the results align with the expectation that fertility changes pressure on employment varies with gender attitude, being lower for women from countries with less traditional gender attitudes.

There are two important aspects to highlight when comparing IV and correlation results. First, because the OLS results are very similar across groups while the IV results are not, endogeneity plays different roles for both natives and immigrants. Precisely, we saw that the effect decreases much more for natives than for any other group, followed by *HighGGI* immigrants when accounting for endogeneity. For women with more egalitarian gender attitudes, like natives/*HighGGI* women, their fertility and labor choices reflect more so their preference. Women who decide to have more children might also simultaneously choose to be less career-oriented, which means that the correlation overestimates the effect of children on employment. Thus, the effect of children on employment largely decreases after considering endogeneity.

The second important point to highlight is the differences in endogeneity signs. Precisely, for all except for *LowGGI*, the endogeneity sign is negative, with the IV coefficient being less negative than the correlation ones. For *LowGGI*, however, the coefficient becomes more negative, which means a positive endogeneity and correlation underestimate the real effect of fertility on employment.

A possible explanation for the endogeneity sign could be differences in human capital as women's education is partly a consequence of their cultural reality due to the following reasons. First, women from countries with higher gender disparity would have less access to education and thus lower education on average. Moreover, the lower education could also reflect their compliance to their birthplace's gender norms. *LowGGI* women are from countries where women would have a more traditional role in society. Choosing to invest less in their human capital could go hand in hand with their expectation to not participate or exit the labor market after having children. Lower educational attainment and work experience would limit their access to the type of work that would be more family-friendly. Thus, employment and motherhood would be more in conflict. Therefore, having more children would have a larger effect on employment. This possible effect is in line with the literature findings, in which Angrist & Evans (1998) finds that the effect of children on labor market outcomes decreases with women's education. Moreover, using the type of occupation to classify women by skills level, Hupkau & Leturcq (2016) finds

that endogeneity upward bias the results to "low-skill" women.

Differences in access to educational attainment for women of different gender attitudes would affect the distribution of women's education. These differences might be driving the change in the endogeneity sign. I discuss the relationship with education further when looking at the robustness checks in Section 4.

3.3 Other Labor Market Outcomes

Additional to employment, there could be other labor characteristics that could be relevantly related to childbearing. Precisely, the effect could drive the number of hours work (including those working 0 hours if not working for paid) as well as earnings. Table A.2 shows the summary statistics of these two variables, as well as the log of earnings for only those employed.

Similarly to employment, the variables' averages show a negative relationship between a more egalitarian gender attitude and labor market outcomes. Natives women work more hours and have higher earnings than immigrants and *LowGGI* women work the least amount of hours and have lower earnings than the rest. Within immigrants, *HighGGI* have hours work and earnings closer to natives.

The analysis of the effect of children on these labor outcomes and the OLS and 2SLS results are in Table A.3. Panel I presents the results for usual hours of work per week, while the results for labor income and its log are in Panel II and Panel III. The OLS results in each of the Panels show similar effects across the different groups, similarly to the employment results. Parallel to the main results, the 2SLS results show different effects for each group, with the most negative effect for women with more gender unequal attitudes. In terms of statistical significance, labor income is affected by having *MorethanTwo* children only in some cases.

For hours of work per week in Panel I, the 2SLS results are significant for all but the *HighGGI* group. The overall effect of having *MorethanTwo* children is reducing the hours worked by three and a half hours. This effect is -3.0 for natives while the overall effect for immigrants is 1.9 hours more negative than for natives. Within immigrants, *LowGGI* decreases hours work per week by -7.5 when having more than two children. In contrast, *HighGGI* effect is not significant and quantitatively similar to that of natives. Looking at Panels II and III for earnings, the effect is not significant for most groups except for those from *LowGGI* countries, especially when looking only at the log of earnings.

Qualitatively, *MorethanTwo* coefficient negatively relates with gender egalitarianism attitudes. Thus, results are in the same line as the main results. It is important to note that the results suggest that children play a more relevant negative impact by determining selection into employment.

3.4 Immigrants: Age of migration

Previous literature finds that as immigrants spend more time in the US and assimilate, their labor supply gets closer to that of natives (Blau *et al.* (2011)). Women who migrated at an older age spent less time in the US at a given age would be "closer" to their country of origin due to higher exposure to said culture. Thus, a variable of interest is the age at migration.

I divide the immigrants' sample based on their age of migration, setting the cutoff at 20 years old. Importantly, this would divide women that have been in the US most of their adult life from those who live in their birthplace most of it. The results for the split samples are in Table A.4. Columns (1)-(3) present the information for immigrants that came before age 20, with the first column showing all the immigrants and, when GGI is available, the results for those from *LowGGI* and *HighGGI* are in column (2) and column (3). Similarly, columns (4)-(6) present the results for immigrants that came when they were 20 years old or older. The first row presents the correlation results, while the second row presents the 2SLS results.

The OLS results are similar as before, with the correlation being comparable across groups. Looking at the 2SLS coefficients, they are quantitatively more negative for those women that migrate when they are older. Although not significant, the coefficient is the largest for *LowGGI* women that came when 20 years or older. In line with the main findings, within each sample, the effect is more negative for *LowGGI*.

Comparing younger (columns (1)-(3)) to older migrants (columns (4)-(6)), the coefficients for those that came to the US being older is 8.34 *pp* more negative than those that migrate before being 20 years old, for all immigrants. The same differences for *LowGGI* and *HighGGI* is 9.60 and 6.68 *pp* more negative, respectively.

From the findings, Table A.4, the pattern across gender views exists for women, no matter how old they came to the US. Nevertheless, the relationship is in line with women who migrate when they are younger, assimilating more to the more egalitarian gender attitudes of the US. Therefore the effect of children on employment is weaker and quantitatively closer to natives result for women who migrate at a younger age. These results suggest that younger immigrants who assimilate to the US are more comparable to natives.

4 Robustness Check

In this section, I explore the robustness of the results, with a large part of the analysis concentrating on immigrants. First, I explore the possible reasons behind endogeneity differences across groups in Section 4.1. I follow this analysis by exploring the effect of sample choices in Section 4.2. I then look at the family's characteristics in Section 4.3.

After this, I concentrate on immigrants by first looking at the results' robustness by using a different GGI group in Section 4.4.1 and controlling for country of origin's educational attainment in Section 4.4.2. Finally, I explore the effect of country of origin's characteristics in Section 4.5.

4.1 Endogeneity Differences

The main results showed that the endogeneity sign is different for immigrants with different gender attitudes. I suggested that, in part, this might be due to differences in human capital. To explore this, first, I look at heterogeneity by education. Second, I explore the possibility that because 2SLS depends on the population's compliers, the observed differences could reflect that the compliers across different types of women are not comparable.

Heterogeneity by Educational Attainment.– I classify women between those with the educational attainment of *"High School or less"* and *"More than High School"*. I then estimate the effect of *MorethanTwo* on women's employment in (1) for these two separate groups. The results are in Table A.5, where Panel I present the results for those with high school or less educational attainment. Panel II presents the results for those with More than High School.

Similar to the main results, the OLS results in Panel I.A show that correlations are fairly similar across all groups. This pattern is different in Panel II.A, where the correlation for natives is about four *pp* more negative than immigrants, especially so for *LowGGI* immigrants. Thus, suggesting that the effect of children on women's employment might vary by their educational attainment.

While the 2SLS results for more than high school education (Panel II.B) show comparable results across gender egalitarianism, the effect for women with high school or less education (Panel I.B) aligns with the pattern in the main results. Comparing the 2SLS to the OLS results, the endogeneity sign is the same across all groups for more than high school (Panel II.B). Thus, the differences in the sign for *LowGGI* in the main findings is due to high school or less educated women (Panel I.B).

The results discussed above suggest that the endogeneity difference is due to educational attainment differences for women of different gender attitudes. Given this, I discuss further the distributional differences across each group's complier next.

Compliers' Characteristics.– When comparing IV results across different groups, one possible concern is that perhaps the compliers might not be completely comparable. Although not directly observable, we can "find" compliers using the method in Abadie (2003). More detail in the method is available in Section A.2. I calculate the compliers'

characteristics following this method, comparing them with the population across the different groups.

The results are in Table A.16. Panel A presents the average characteristics of the population, while Panel B presents the average characteristics of the compliers. The complier-to-population ratio is in Panel C. Overall, the compliers' characteristics are similarly comparable across all the groups. It is important to note, however, that there is a salient difference regarding educational attainment. Specifically, the immigrants' compliers have a lower education relative to the population they represent. More immigrant's compliers have a "high school or less" educational attainment than native's compliers, compared with their population. This difference is especially so for those immigrants coming from *LowGGI*. These results are not surprising, given what we saw in Table A.5, where the *LowGGI* endogeneity sign reflects the one founded for those with "high school or less" education.

It is important to note that, aside from educational differences, comparing the characteristics of the compliers to the population shows a similar pattern across groups of women. This finding eases concerns that the results represent very different samples across different levels of gender egalitarianism.

4.2 Sample Choice

4.2.1 Women from Age 21-35

In the main analysis, I estimate the results for women that are 21-40 years old. To explore if the results are sensitive to the sample use, I limit the sample to the age range of 21-35 and estimate (1) using this subsample. This age range is the same as in Angrist & Evans (1998), which I expanded in the main data to accommodate that having children at an older age is more common nowadays than in the 1980s/1990s.

The results are in Table A.6, where Panel A and Panel B present the OLS results and the 2SLS results, respectively. The results are qualitatively similar to the main findings. Quantitatively, the 2SLS results are larger for everyone than the main results, but more so for the *LowGGI* women.

The results 2SLS being more negative for this group of women suggest that younger women's employment status is more affected by children than older women. Note that since I exclude older women, as a consequence, this excludes the women that start having children when they are older. These women tend to favor education and employment, meaning they are more career-driven than women having children at a younger age. Simultaneously, by having children when they are older, they would be more established in their careers. Therefore, their employment status would be less affected by childbearing.

It is important to note that qualitatively the conclusion does not change. This result suggests that sample selection regarding age is not a concern in terms of the main

outcomes.

4.2.2 Previous Data: Census 1980 and 1990 Data

To further explore the results, I use the 5% state sample 1980 and 1990 Census. Doing this analysis serves two purposes. First, It helps to estimate the results for a period of, on average, higher fertility and lower women's labor market participation. Thus, childbearing should decrease employment more for these samples than for the main results in Table 2. Second, it helps to check possible measurement errors concern regarding women's total fertility. Precisely, I measure total fertility in the main sample using the number of own children in the household. This variable is prone to measurement errors if women have other children who do not live with them. In the case of the 1980 and 1990 databases, the information of *"Children ever born"* is available, which is a more precise measurement of the total number of children. Thus, if results using the 1980 and 1990 samples are comparable to the main results, it will alleviate measurement error concerns.

I process the databases the same way as the main sample, except that I change the variable used to measure total fertility. Precisely, I limit the sample to women whose number of children matches their reported number of children ever born, instead of the number of own children in the household. For the 1990 sample only, "multiple birth" is defined with age only, instead of both age and quarter of birth since the quarter of birth is not available. Finally, I only concentrate on natives and immigrants due to a comparable gender attitude measure not available for the years covered by these samples.

The summary statistics for both samples are in Table A.7. The first three columns are of the 1980 sample, while columns (4)-(6) present the information for the 1990 sample. The columns are in the following order: all, natives, and immigrants. Overall, the observed variables' distribution is similar to the main sample in Table 1 for both census samples. Immigrant women have slightly more children, with around 30% of the women having more than two children. The sex composition of the first two children is such that half of the women have two children of the same sex. The sample skew slightly younger than the main sample, which aligns with these samples being in years in which younger women have two children at a higher rate.

In terms of employment, we can see the change in women's overall employment over time. While only 48% of the women worked for pay in 1980, this number grew to 60% in 1990. Comparing natives to immigrants, natives work 4 *pp* more than immigrants in 1980. This difference widens to 10 *pp* 1990. Thus, native women's employment gain widens the native to immigrant employment gap.

Using these samples, I estimate the model in (1) using the OLS and 2SLS approaches. The results are in Table A.8. The columns show the information for Census 1980 in (1)-(3), while the Census 1990 information is in (4)-(6). The first column presents the

results for all women, followed by natives and immigrants in the following two columns for each sample. The rows show the OLS followed by the 2SLS approach first and second stages.

Looking at the OLS results, correlations are similar across all the groups. This pattern is in line with the main findings. Moreover, when looking at the 2SLS results, the first stage shows that the IV is relevant and that the effect is similar for natives and immigrants. We then look at the second stage results.

In the second stage, we can see that in 1980, the effect of having *MorethanTwo* children reduces women's employment by 11.8 *pp*. The effect for natives is fairly similar (-10.92 *pp*), while immigrants have a more negative effect by 11.50 *pp* compared with natives. Similarly, for 1990 data, the children's effect for all women is -12.48%, while the effect is -11.91 *pp* for natives and immigrants effect is 6.69 *pp* more negative.

It is important to note that results are qualitatively similar to the main results in Table 2. Immigrant women are more negatively affected by having more children than native women. The pattern of the gap and the direction of the results are the same, with improvements over time. Comparing the Table 2 coefficients, the fertility effect are significant smaller nowadays than in Table A.8. This decrease of negative impact is consistent with women's gain in the labor force, weakening gender roles over time, and family policies implementation. Simultaneously, we see that although the native-immigrant gap persists, it is smaller than in the past. This reduction could be in part explained by worldwide gender equality gains, which weakens traditional gender attitudes across immigrants as a whole.

Concerning measurement errors in total fertility, the results ease this concern. The estimated pattern for natives and immigrants persists in the 1980 and 1990 data when using the *"Children ever born"* as total fertility. Since the results are qualitatively similar to the main results, it reassures that measurement errors do not drive the results.

4.3 Family Characteristics

There are reasons to think that the effect of children on employment could also be affected by family characteristics. One example of this effect is the number of other adults in the household, which can help ease the opportunity cost of working since other adults can help with childcare. Moreover, women's marital status might also affect how likely women are to be in the labor force primarily for two main reasons. First, having a partner would mean that they might need to bargain their roles in the household. If their partner wants a stay-at-home mother, this will pressure women to leave the labor force, aside from their views on gender roles. Second, if their partner is employed, the family has access to a certain income level, which eases the need for an additional income and might reduce women's need to be working for earnings. In a similar vein to the latter,

family income might also affect how likely women are to work for pay, aside from marital status.

To the extent that different types of women have different types of families, it might be that the differences in the effect of fertility on employment are due to variations in family characteristics. To ease this concern, I include family-related variables in my regression. Precisely, I include a married dummy, the number of other adults (18–65 years old) in the household and the family income, defined as the sum of the other adults' labor income. The results with these controls are in Table A.9. The table presents both the OLS and the 2SLS results for the complete sample and the subgroups.

The OLS results show similar correlations across subgroups of women, which is in line with the main results in Table 2. Moreover, the 2SLS results are qualitatively similar to the main findings. The effect of having more than two children on employment decreases as gender egalitarianism increases. The main difference is for *HighGGI* immigrants, which is now significant and quantitatively larger than in the main results. Nevertheless, the results are not statistically different from each other.

Albeit small quantitative differences with the main results, the general pattern of the findings while controlling for family characteristics mimics the results for the complete sample. In this way, the relationship of the effect of childbearing on employment and gender attitudes still holds.

4.4 Immigrants' Characteristics

4.4.1 Using a different GGI measurement

In the main results, the grouping between *LowGGI* and *HighGGI* is determined using the average GGI for the entire sample. To explore sensitivity to this definition, I now use the average GGI per year as the cutoff to create new *LowGGI* and *HighGGI* groups.

I estimate (1) for these new groups. The results are in Table A.10. Columns (1) and (2) present the OLS results, while the 2SLS results are in columns (3) and (4). The results are qualitatively very similar to the main findings, both OLS and 2SLS. Because of the redistribution of women between *LowGGI* and *HighGGI*, we see that the effect decreases for *LowGGI* and it increases in *HighGGI*. Nevertheless, compared with the main results, in Table 2, results are not qualitatively different. Thus, the definition of the relatively less and more egalitarian gender attitudes does not affect the main results.

4.4.2 Immigrants Educational Attainment Distribution

From before, educational differences are playing a role in the differences in endogeneity sign between *LowGGI* and *HighGGI*. One possibility is that since immigrants are self-selected, countries' differences in selection into migration based on educational at-

tainment might be behind these differences. I exploit the information on the education distribution by country from the Barro-Lee (BL) estimation to create the immigrants' relative educational attainment distribution.[8]

I use the BL's educational attainment to compare with the educational distribution of this country's immigrants in the US. For each level of education, I create "educational attainment ratios", equal to the percentage of a country's population divided by the percentage of immigrants from that country for a given level of education. For example, if for a given country 20% of the population are in the education category of "no schooling" while 10% of their immigrants in the US are, then the ratio is $20\%/10\% = 2$. We can interpret this ratio as an individual from the country being two times more likely to have no schooling than their immigrants in the US. In this example, immigrants are relatively more educated and hence positively selected on education.

I use the ratios by education category to create the variable Education Ratio (*EducRatio*), which assigns to each individual the educational attainment ratio corresponding to their own. I estimate (1) for immigrant women using the *EducRatio* as an additional control variable.

The idea of this robustness check is that if, for example, *LowGGI* tend to have lower education than *HighGGI* relative to their country of origin, selection might drive the within-immigrants differences in the effect of having more than two children on employment. The results are in Table A.11, with the results for all the immigrants in column (1). The results for *LowGGI* and *HighGGI* immigrants are in column (2) and (3), respectively. The results are qualitatively similar to the main results, thus suggesting that differences in educational attainment selection across different gender attitudes' countries do not affect the results.

4.5 Country of Origin's Characteristics

It could be that the resulting pattern for immigrants is due to differences in their country of origin's characteristics, such as education, development, and fertility. These characteristics might correlate with GGI and thus determine both women's gender egalitarianism and labor force participation.[9]

To understand the effect of birthplace characteristics in my results, I collect country of origin's information from the World Bank. Precisely, I use the following variables: GDP per capita (international dollars converted by purchasing power parity (PPP)), the fertility rate (number of children per woman), and the sex ratio at birth (male-to-female

[8]Barro & Lee (2013), retrieved from :http://www.barrolee.com/ I use the percentage in the population in the following educational attainment categories: (1) no schooling; (2) Primary Schooling; (3) Complete Primary Schooling; (4) Secondary Schooling; (5) Complete Secondary Schooling; (6) Tertiary Schooling; (7) Complete Tertiary Schooling.

[9]Even though the GGI disassociates from development by construction by looking at countries' gender gap, it is still worth it exploring the effect of such characteristics.

births) when available.[10] Moreover, I include BL's educational distribution by country as an additional key characteristic.

Additionally to the variables described above, since families and close societal peers' views would affect individuals' views, peers are one of the key channels of culture transmission. This channel means that the more immigrants from a country in the US, the more accessible this country's culture is in the US. In this sense, the rate of migration from a given country is relevant. To take this into account, I measure the *Immigrants Rate*, as the total of immigrants in the US from a given country divided by its population, when available. I retrieved the information on countries' populations from the World Bank.[11]

It is not only important how many immigrants but also how concentrated they are in the US. Immigrants who live around a large population from the same country would connect more with their birthplace's culture than immigrants who live in states with fewer people who share the same culture. Individuals around a more diverse group would assimilate faster than those closer to a group with a similar background. To account for this, I measured *Immigrants Concentration*, as the percentage of immigrants from a given country that lives in a given state-year out of the total people in that given state-year.

I use the variables described above in the following subsections. I first explore the impact of the country of origin's development. Moreover, given that children's sex preference might be a driver in the number of children a woman has, I then concentrate on the effect of the country of origin's sex ratio.

4.5.1 Development, Fertility and Education

To understand how a country's characteristics might affect the results, I implement the following two-step analysis:

1. I regress GGI on country characteristics→

2. I get residual from step 1, GGI_{res}, and use it to classify women as *LowGGI* if GGI_{res} is negative, and *HighGGI* if GGI_{res} is positive. Using this new grouping, I estimate (1).

This process aims to classify women based on a GGI after considering other countries' characteristics that might be driving the main results. If the country of origin's characteristics drives the main results, estimation results from step 2 should be significantly different from the main findings in Table 2.

[10] In the following cases, I assign a number as follows. First, I use the information of China for Taiwan. Second, for Czechoslovakia (Yugoslavia), I use the average from the Czech Republic and the Slovak Republic (Croatia, Montenegro, Serbia, Bosnia Herzegovina Slovenia, and Kosovo).

[11] In the following cases, I assign the value as follows. I use China's information for Taiwan. For Czechoslovakia (Yugoslavia), I use the total population from the Czech Republic and the Slovak Republic (Croatia, Montenegro, Serbia, Bosnia Herzegovina Slovenia, and Kosovo).

Table A.12 shows the estimation results from step 2. The first two columns present the for *LowGGI* and *HighGGI* immigrants, controlling for the GDP per Capita and Fertility, both as a 4th polynomial. Columns (3) and (4) present the results controlling for BL's distribution of educational attainment. The final two columns also include the migration rates. The rows present the OLS and 2SLS results, respectively.

The results show the same pattern that the main results. First, the OLS results are comparable across groups. Second, in line with the main findings, the effect *MorethanTwo* decreases with gender egalitarianism. When comparing *LowGGI*'s *MorethanTwo* coefficient to the *HighGGI* one, the coefficient gap observed in the main findings still holds. Since the pattern does not change, it eases concerns that the main results are due to differences in countries' characteristics.

4.5.2 Son preference

Another important concern is that children's gender preference is not homogeneous across different cultures since some countries have a defined preference for one gender. In that case, women with a child of their preferred sex might stop having children earlier than women who prefer mixed-sex children. Consequently, these variations in children's gender preferences might create selection bias in my sample. To explore this possibility, I use the World Bank measures of sex ratio at birth (*SexRatio*) to understand how this might affect the main findings.

SexRatio measures male to female birth ratio average over the last five years in a given country. Regarding the value that this ratio would have, note that the natural sex ratio has an average of 1.05 and varies from 1.03 to 1.07.[12] Thus, there are more boys than girls being born. Because some countries have sex preferences, the *SexRatio* range from 1.018 to 1.17.

To explore if women's gender preference is a concern, I exclude women from countries with an imbalanced sex ratio using two different cutoffs. In the stricter cutoff, I exclude women from countries with *SexRatio* of more than 1.07, which is the upper bound of the considered natural sex ratio.[13]. For the second cutoff, I use the 90th percentile in my sample, which corresponds to a *SexRatio* of 1.099.[14]

Table A.13 presents the results excluding women from countries with a very imbalanced sex ratio. Panel I presents the results using the 1.07 ratio as the cutoff, while Panel II uses 1.099 instead. Within each Panel, the rows show the OLS and 2SLS results.

[12]For various biological reasons, such as the probability of miscarriage by the sex of the child at different stages of pregnancy, there is a male-biased at birth. For a further discussion of sex ratio, see https://ourworldindata.org/gender-ratio#sex-ratio-at-birth

[13]This rule eliminates the following countries: Albania, Macedonia, Croatia, Montenegro, Serbia, Armenia, Azerbaijan, Republic of Georgia, China, Hong Kong, Taiwan, Singapore, Vietnam, India, Pakistan, Cyprus, and Micronesia

[14]This rule excludes Albania, Armenia, Azerbaijan, Republic of Georgia, China, Hong Kong, Taiwan, Vietnam, India

In terms of the columns, columns (1) - (3) present the results without controlling for *SexRatio* for all immigrants, *LowGGI* and *HighGGI* in that order. Similarly, columns (4) - (6) do the same analysis by further including the *SexRatio* as a control. Panel I and II OLS results show a similar correlation across groups. These results are in line with the main findings. In terms of the 2SLS, the results are qualitatively similar to the main results. Nevertheless, it is important to note that most of the results are more negative than the main results in Table 2. Therefore, when excluding women with more sex-biased preferences regarding children, the effect of *MorethanTwo* on employment is more negative, especially for less gender-egalitarian women.

From the results Table A.13, we can conclude that, if anything, not considering that some countries might have preferences for boys would underestimate the negative effect of fertility on employment. This result is especially true for *LowGGI* women. This effect is possibly due to the correlation between GGI and *SexRatio* being negative at -0.1181. This relationship means that women with less gender-egalitarian attitudes would also have a higher son preference, lowering compliance with the instrument. If the first two children are boys, they are more willing to stop due to a stronger preference for boys than for having mixed-sex children. Thus, when excluding skews son preference countries, the results are more negative for women with less gender egalitarianism.

5 Conclusion

This paper asks if the effect of fertility on employment varies by women's gender attitudes. To do so, I use the 2006-2018 ACS and divide women between those born in the US (natives) and immigrants, where native women are from a country with more egalitarian gender attitudes than immigrants. Because immigrants are not homogeneous, I measure gender attitudes using their country of origin's gender gap, the GGI. Matching the GGI to immigrant women's birthplace, I classify immigrant women between those from countries with relatively less (more) egalitarian gender attitudes, *LowGGI* (*HighGGI*).

Using the first two children's sex composition as an IV, I estimate the effect of having more than two children on employment for women of different gender attitudes. The findings show that having more than two children decreases women's employment, but it masks heterogeneity. The negative effect of more than two children decreases as gender egalitarianism increases. For immigrants, the effect is almost double that of natives. Among immigrants, the effect of having more than two children on employment for women with less egalitarian gender attitudes is three times more negative than that of natives.

Although the results are robust to a series of analyses, it is important to note that there could still be some selection concerns due to self-selection into migration. That said, it might be that election would bias the coefficient toward zero if migrants are less attached to their culture than those that stay back in their countries. Moreover, for

the differences in the immigrants' results to be due to selection, the effect must have the opposite sign for *LowGGI* and *HighGGI* women. Thus, easing the self-selection concerns.

In 2018, the US labor force participation's gender gap was in 59th place globally,[15] falling behind many developed countries. As the US advance with families policies that aim to improve gender equality in the labor force participation, this paper draws attention to a key characteristic: variations on the importance of fertility by gender attitudes. Understanding the gender attitudes' composition of the policies' target population is key to recognizing its relevance and reach. Thus, implementing culturally sensitive family policies could improve the overall impact of policies aiming to ease the effect of fertility on women's employment.

References

Aaronson, Daniel, Dehejia, Rajeev H., Jordan, Andrew, Pop-Eleches, Cristian, Samii, Cyrus, & Schulze, Karl. 2021. The Effect of Fertility on Mothers' Labor Supply over the Last Two Centuries. *The Economic Journal*, **131**(633), 1–32.

Abadie, Alberto. 2003. *Semiparametric Instrumental Variable Estimation of Treatment Response Models*. Vol. 113.

Akerlof, George A., & Kranton, Rachel E. 2000. Economics and identity. *Quarterly Journal of Economics*, **115**(3), 715–753.

Akerlof, George A., & Kranton, Rachel E. 2010. *Identity Economics: How Our Identities Shape Our Work, Wages, and Well-Being*.

Anderson, Deborah J, Binder, Melissa, & Krause, Kate. 2002. The Motherhood Wage Penalty: Which Mothers Pay It and Why? *American Economic Review*, **92**(2), 354–358.

Anderson, Deborah J., Binder, Melissa, & Krause, Kate. 2003. The Motherhood Wage Penalty Revisited: Experience, Heterogeneity, Work Effort, and Work-Schedule Flexibility. *Industrial and Labor Relations Review*, **56**(2), 273–294.

Angrist, Joshua D., & Evans, William N. 1998. Children and Their Parents' Labor Supply: Evidence from Exogenous Variation in Family Size. *American Economic Association*, **88**(3), 450–477.

Antecol, Heather. 2000. An Examination of Cross-Country Differences in the Gender Gap in Labor Force Participation Rates. *Labour Economics*, **7**(4), 409–426.

[15]Detail available here Global Gender Gap Report 2018:USA

Barro, Robert, & Lee, Jong-Wha. 2013. A New Data Set of Educational Attainment in the World, 1950-2010. *Journal of Development Economics*, **104**, 184–198.

Battu, Harminder, Mwale, McDonald, & Zenou, Yves. 2007. Oppositional Identities and the Labor Market. *Journal of Population Economics*, **20**(3), 643–667.

Bertrand, Marianne, Kamenica, Emir, & Pan, Jessica. 2015. Gender Identity and Relative Income within Households. *Quarterly Journal of Economics*, 571–614.

Black, Dan, Joo, Joonhwi, LaLonde, Robert, Smith, Jeffrey, & Taylor, Evan. 2017. *Simple Tests for Selection: Learning More from Instrumental Variables*.

Blau, David M, & Robins, Philip K. 1989. Fertility, Employment, and Child-Care Costs. *Demography*, **26**(2), 287–299.

Blau, Francine D. 1992. The Fertility of Immigrant Women: Evidence from High- Fertility Source Countries. *Pages 93 – 134 of: Immigration and the Workforce: Economic Consequences for the United States and Source Areas*.

Blau, Francine D., & Kahn, Lawrence M. 2015. Substitution Between Individual and Source Country Characteristics: Social Capital, Culture, and US Labor Market Outcomes Among Immigrant Women. *Journal of Human Capital*, **9**(4), 439–482.

Blau, Francine D, Kahn, Lawrence M, & Papps, Kerry L. 2011. Gender, Source Country Characteristics, and Labor Market Assimilation among Immigrants. *The Review of Economics and Statistics*, **93**(1), 43–58.

Blau, Francine D., Kahn, Lawrence M., Liu, Albert Yung Hsu, & Papps, Kerry L. 2013. The Transmission of Women's Fertility, Human Capital, and Work Orientation Across Immigrant Generations. *Journal of Population Economics*, **26**(2), 405–435.

Blau, Francine D., Kahn, Lawrence M., Comey, Matthew, Eng, Amanda, Meyerhofer, Pamela, & Willén, Alexander. 2020a. Culture and gender allocation of tasks: source country characteristics and the division of non-market work among US immigrants. *Review of Economics of the Household*, **18**(3), 907–958.

Blau, Francine D., Kahn, Lawrence M., Brummund, Peter, Cook, Jason, & Larson-Koester, Miriam. 2020b. Is There Still Son Preference in the United States? *Journal of Population Economics*, **33**, 709–750.

Boca, Daniela Del, Pasqua, Silvia, & Pronzato, Chiara. 2005. Fertility and employment in Italy, France, and the UK. *Labour*, **19**(SUPPL. 1), 51–77.

Bongaarts, John, Blanc, Ann K., & McCarthy, Katharine J. 2019. The links between women's employment and children at home: Variations in low- and middle-income countries by world region. *Population Studies*, **73**(2), 149–163.

Buckles, Kasey S., & Hungerman, Daniel M. 2013. Season of Birth and Later Outcomes: Old Questions, New Answers. *Review of Economics and Statistics*, **95**(3), 711–724.

Budig, Michelle J, & England, Paula. 2001. The Wage Penalty for Motherhood. *American Sociological Review*, **66**(2), 204–225.

Budig, Michelle J., & Hodges, Melissa J. 2010. Differences in Disadvantage: Variation in the Motherhood Penalty Across White Women's Earnings Distribution. *American Sociological Review*, **75**(5), 705–728.

Campaña, Juan Carlos, Giménez-Nadal, J. Ignacio, & Molina, José Alberto. 2020. Self-employed and Employed Mothers in Latin American Families: Are There Differences in Paid Work, Unpaid Work, and Child Care? *Journal of Family and Economic Issues*, **41**(1), 52–69.

Chung, YoonKyung, Downs, Barbara, Sandler, Danielle H, & Sienkiewicz, Robert. 2017. *The Parental Gender Earnings Gap in the United States*.

Clark, Roger, Ramsbey, Thomas W, & Adler, Emily Stier. 1991. Culture, Gender, and Labor Force Participation. *Gender and Society*, **5**(1), 47–66.

Conley, Dalton, & Glauber, Rebecca. 2006. Parental Educational Investment and Children's Academic Risk: Estimates of the Impact of Sibship Size and Birth Order from Exogenous Variation in Fertility. *Journal of Human Resources*, **41**(4), 722–737.

Correll, Shelley J., Benard, Stephen, & Paik, In. 2007. Getting a Job: Is There a Motherhood Penalty? *American Journal of Sociology*, **112**(5), 1297–1338.

Cristia, Julian P. 2008. The effect of a first child on female labor supply: Evidence from women seeking fertility services. *Journal of Human Resources*, **43**(3), 487–510.

Cygan-Rehm, Kamila. 2014. *Immigrant Fertility in Germany: The Role of Culture*.

Danielle Sandler, & Nichole Szembrot. 2019. *Maternal Labor Dynamics: Participation, Earnings, and Employer Changes*.

Datar, Ashlesha. 2006. Does Delaying Kindergarten Entrance Give Children a Head Start? *Economics of Education Review*, **25**(1), 43–62.

Del Boca, Daniela, Pasqua, Silvia, & Pronzato, Chiara. 2009. Motherhood and Market Work Decisions in Institutional Context: A European Perspective. *Oxford Economic Papers*, **61**(SPEC. ISS.), 147–171.

Dobkin, Carlos, & Ferreira, Fernando. 2010. Do School Entry Laws Affect Educational Attainment and Labor Market Outcomes? *Economics of Education Review*, **29**(1), 40–54.

Ekert-Jaffe, Olivia, & Stier, Haya. 2009. Normative or economic behavior? Fertility and women's employment in Israel. *Social Science Research*, **38**, 644–655.

Fernández, Raquel. 2007. Alfred Marshall lecture women, work, and culture. *Journal of the European Economic Association*, **5**(2-3), 305–332.

Fernández, Raquel. 2011. Does Culture Matter? *J. Benhabib, M. O. Jackson and A. Bisin, editors, Handbook of Social Economics*, **1A**, 481–510.

Fernandez, Raquel, & Fogli, Alessandra. 2006. Fertility : The Role of Culture and Family Experience. *Journal of the European Economic Association*, **4**(2/3), 552–561.

Fernandez, Raquel, & Fogli, Alessandra. 2009. Culture: An empirical investigation of beliefs, work, and fertility. *American Economic Journal: Macroeconomics*, **1**(1), 146–177.

Florian, Sandra M. 2018. Racial Variation in the Effect of Motherhood on Women's Employment: Temporary or Enduring Effect? Sandra. *Social Science Research*, **73**, 80–91.

Francesco Giavazzi, Schiantarelli, Fabio, & Serafinelli, Michel. 2013. Attitudes, Policies, and Work. *Journal of the European Economic Association*, **11**, 1256–1289.

Furtado, Delia, & Hock, Heinrich. 2010. Low Skilled Immigration and Work-Fertility Tradeoffs Among High Skilled US Natives. *American Economic Review*, **100**(2), 224–228.

Gangl, Markus, & Ziefle, Andrea. 2009. Motherhood, labor force behavior, and women's careers: An empirical assessment of the wage penalty for motherhood in Britain, Germany, and the United States. *Demography*, **46**(2), 341–369.

Gay, Victor, Hicks, Daniel L, & Shoham, Amir. 2018. Decomposing culture : an analysis of gender , language , and labor supply in the household. *Review of Economics of the Household*, 879–909.

Glauber, Rebecca. 2007. Marriage and the Motherhood Wage Penalty Among African Americans, Hispanics, and Whites. *Journal of Marriage and Family*, **69**(4), 951–961.

Goldin, Claudia, & Polachek, Solomon. 1987. Residual Differences by Sex: Perspectives on the Gender Gap in Earnings. *The American Economic Review*, **77**(2), 143– 151.

Gorry, Devon. 2019. *Heterogeneous Consequences of Teenage Childbearing.*

Guner, Duygu, & Uysal, Gokce. 2014. *Culture, Religiosity and Female Labor Supply.*

Gutiérrez-Domènech, Maria. 2005. Employment after Motherhood: A European Comparison. *Labour Economics*, **12**(1), 99–123.

Harkness, Susan, & Waldfogel, Jane. 1999. *the Family Gap in Pay: Evidence From Seven Industrialized Countries.*

Hausmann, R., Tyson, L. D., & Zahidi, S. 2006. The Global Gender Gap Report 2006. *Geneva, Switzerland: World Economic Forum.*

Hupkau, Claudia, & Leturcq, Marion. 2016. Fertility and Labor Supply : New Evidence from the UK,Working Paper,. *Centre for Economic Performance, London School of Economics.*

Joona, Pernilla Andersson. 2017. Are mothers of young children more likely to be self-employed? The case of Sweden. *Review of Economics of the Household*, **15**(1), 307–333.

Kahn, Joan R., García-Manglano, Javier, & Bianchi, Suzanne M. 2014. The Motherhood Penalty at Midlife: Long-Term Effects of Children on Women's Careers Joan. *Journal of Marriage and Family*, **76**(1), 56–72.

Kleven, Henrik, Landais, Camille, Posch, Johanna, Steinhauer, Andreas, & Zweimüller, Josef. 2019a. Child Penalties across Countries: Evidence and Explanations. *AEA Papers and Proceedings*, **109**, 122–126.

Kleven, Henrik, Landais, Camille, & Søgaard, Jakob Egholt. 2019b. Children and gender inequality: Evidence from Denmark. *American Economic Journal: Applied Economics*, **11**(4), 181–209.

Lu, Yao, Wang, Julia Shu Huah, & Han, Wen Jui. 2017. Women's Short-Term Employment Trajectories Following Birth: Patterns, Determinants, and Variations by Race/Ethnicity and Nativity. *Demography*, **54**(1), 93–118.

Lundberg, Shelly, & Rose, Elaina. 2000. Parenthood and the Earnings of Married Men and Women. *Labour Economics*, **7**, 689–710.

Lundborg, Petter, Plug, Erik, & Rasmussen, Astrid Würtz. 2017. Can Women Have Children and a Career? IV Evidence from IVF Treatments. *American Economic Review*, **107**(6), 1611–1637.

Marcén, Miriam, & Morales, Marina. 2019. Live together: does culture matter? *Review of Economics of the Household*, **17**(2), 671–713.

Marcén, Miriam; Morales, Marina Working. 2019. Gender division of household labor: How does culture operate? *GLO Discussion Paper*, **373**(Global Labor Organization (GLO)).

Mincer, Jacob, & Polachek, Solomon. 1974. Family Investments in Human Capital : Earnings of Women, Marriage, Family Human Capital, and Fertility. *Journal of Political Economy*, **82**(2), S76–S108.

Muller, Joanne S., Hiekel, Nicole, & Liefbroer, Aart C. 2020. The Long-Term Costs of Family Trajectories: Women's Later-Life Employment and Earnings Across Europe. *Demography*, **57**(3), 1007–1034.

Nekby, Lena, & Rödin, Magnus. 2007. *Acculturation Identity and Labor Market Outcomes*.

Neuman, Emma. 2018. Source Country Culture and Labor Market Assimilation of Immigrant Women in Sweden: Evidence from Longitudinal Data. *Review of Economics of the Household*, **16**(3), 585–627.

Polachek, Solomon William. 1975. Differences in Expected Post-School Investment as a Determinant of Market Wage Differentials. *International Economic Review*, **16**(2), 451–470.

Polachek, Solomon William. 2004. How the Human Capital Model Explains Why the Gender Wage Gap Narrowed. *IZA Discussion Papers 1102. Institute of Labor Economics (IZA)*.

Reimers, Cordelia W . 1985. Cultural Differences in Labor Force Participation Among Married Women. *American Economic Review*, **75**(2), 251–255.

Reshid, Abdulaziz Abrar. 2019. The Gender Gap in Early Career Wage Growth: The Role of Children, Job Mobility, and Occupational Mobility. *Labour*, **33**(3), 278–305.

Rosenzweig, Mark R, & Wolpin, Kenneth. 1980. Life-Cycle Labor Supply and Fertility: Causal Inferences from Household Models. *Journal of Political Economy*, **88**(2), 328–348.

Ruggles, Steven, Flood, Sarah, Goeken, Ronald, Grover, Josiah, Meyer, Erin, Pacas, Jose, & Sobek, Matthew. 2021. *IPUMS USA: Version 11.0 [ACS]. Minneapolis, MN: IPUMS, 2021. https://doi.org/10.18128/D010.V11.0*.

Sanders Korenman, & Neumark, David. 1992. Marriage, Motherhood, and Wages. *The Journal of Human Resources*, **27**(2), 233–255.

Sandra M. Florian. 2018. Motherhood and Employment Among Whites, Hispanics, and Blacks: A Life Course Approach. *Journal of Marriage and Family*, **80**(1), 134–149.

Schoeni, Robert F. 1998. Labor market outcomes of immigrant women in the United States: 1970 to 1990. *International Migration Review*, **32**(1), 57–77.

Smith, James P. 2006. Immigrants and the labor market. *Journal of Labor Economics*, **24**(2), 203–233.

Topel, Robert J . LaLonde, & ., Robert H. 1991. Immigrants in the American Labor Market: Quality, Assimilation, and Distributional Effects. *American Economic Review*, **81**(2), 297–302.

Topel, Robert J. LaLonde, & H., Robert. 1992. The Assimilation of Immigrants in the U.S. Labor Market. *Pages 67–92 of: Immigration and the Workforce: Economic Consequences for the United States and Source Areas.*

Appendix

Table A.1: First Stage Estimates of The Effect of *Same Sex* on *MorethanTwo*

	(1)	(2)	(3)	(4)	(5)
				Immigrants	
	All	Natives	Immigrants	*LowGGI*	*HighGGI*
SameSex	0.0521***	0.0533***	0.0478***	0.0425***	0.0526***
	(0.0010)	(0.0011)	(0.0020)	(0.0033)	(0.0026)
F-Statistics	2,728.3	2,349.5	594.3	168.0	417.7
N	1,400,784	1,125,822	274,962	127,313	131,970

Note.– 2SLS first stage estimated coefficient of *SameSex* in equation (2). Covariates in the model but excluded from the table for brevity: age, age at first child, Black, Hispanic and Other races dummies (white as a reference group), 1st Born Boy, 2nd Born Boy, year and states dummies. Columns: *All:* include all the sample; *Natives:* include those born in the US (including US outlying territory); *Immigrants:* those born outside from the US; *LowGGI:* those born outside of the US, that have GGI information, and GGI is below the average value; *HighGGI:* those born outside of the US, that have GGI information, and GGI is above the average value. Standard errors cluster at the birthplace in parentheses; * p<0.10, ** p<0.05, *** p<0.01.

Table A.2: Summary Statistics: Hours Worked and Labor Income

	(1)	(2)	(3)	(4)	(5)
				Immigrants	
	All	Natives	Immigrants	*LowGGI*	*HighGGI*
			All		
Hours Work Per Week	24.90	25.91	20.76	19.11	22.06
	(18.92)	(18.60)	(19.62)	(19.56)	(19.57)
Earnings	25,165.34	26,333.51	20,382.31	17,927.50	22,678.32
	(37,782.31)	(37,642.77)	(37,975.50)	(36,335.58)	(39,942.01)
N	1,400,784	1,125,822	274,962	127,313	131,970
			Only Employed		
Log of Earnings	10.03	10.04	9.99	9.91	10.05
	(1.19)	(1.19)	(1.17)	(1.19)	(1.16)
N	936,879	786,940	149,939	63,939	76,256

Note.– Sample means with standard deviations in parentheses of the main variables. Columns: *All:* Include all the variables; *Natives:* Include those born in the US (including US outlying territory); *Immigrants:* those born outside from the US; *LowGGI:* those born outside of the US, that have GGI, and GGI is below the average value; *HighGGI:* those born outside of the US, that have GGI, and GGI is above the average value. Rows: *Hours Work per Week:* Usual hours of worked per week; *Earnings:* Labor income, CPI corrected to December 2019; *Log of Earning:* Log of labor income for only employed women

Table A.3: Effect of Having More Than Two Children on Women's Labor Market Outcomes: Hours Worked and Labor Income

	(1) All	(2) Natives	(3) Immigrants	(4) *LowGGI*	(5) *HighGGI*
Panel I: Usual Hours Work per Week					
A: OLS					
MorethanTwo	-6.03***	-6.02***	-5.72***	-5.50***	-5.76***
	(0.17)	(0.05)	(0.30)	(0.34)	(0.30)
B: 2SLS					
MorethanTwo	-3.48***	-2.98***	-4.89**	-7.53**	-2.87
	(0.89)	(0.82)	(2.10)	(3.08)	(2.06)
Panel II: Labor Earnings					
A: OLS					
MorethanTwo	-6,339.69***	-6,411.50***	-5,471.63***	-5,018.69***	-5,188.87***
	(216.15)	(82.86)	(660.13)	(870.49)	(445.19)
B: 2SLS					
MorethanTwo	-2,246.16*	-1,952.89	-2,613.07	-4,788.57*	-964.69
	(1,201.87)	(1,471.67)	(2,158.27)	(2,899.45)	(3,205.84)
F-Statistics	2,728.3	2,349.5	594.3	168.0	417.7
N	1,400,784	1,125,822	274,962	127,313	131,970
Panel III: Log of Labor Earnings					
A: OLS					
MorethanTwo	-0.1806***	-0.1827***	-0.1507***	-0.1534***	-0.1203***
	(0.0065)	(0.0036)	(0.0129)	(0.0118)	(0.0180)
B: 2SLS					
MorethanTwo	-0.0854	-0.0562	-0.2350**	-0.4004**	-0.1131
	(0.0575)	(0.0591)	(0.1052)	(0.1915)	(0.1448)
F-Statistics	2,419.8	1,644.4	391.0	92.59	215.0
N	936,879	786,940	149,939	63,939	76,256

Note.– OLS and 2SLS estimated coefficient of *MorethanTwo* in equation (1) for other labor market outcomes. Covariates in the model but excluded from the table for brevity: age, age at first child, Black, Hispanic and Other races dummies (white as a reference group), 1st Born Boy, 2nd Born Boy, year and states dummies. Columns: *All:* include all the sample; *Natives:* include those born in the US (including US outlying territory); *Immigrants:* those born outside from the US; *LowGGI:* those born outside of the US, that have GGI information, and GGI is below the average value; *HighGGI:* those born outside of the US, that have GGI information, and GGI is above the average value. Standard errors clustered at birthplace in parentheses; * p<0.10, ** p<0.05, *** p<0.01.

Table A.4: Effect of Having More Than Two Children on Immigrant Women's Employment Status: By Age at Migration

	(1)	(2)	(3)	(4)	(5)	(6)
	Age at Migration					
	< 20 years old			**20 years old or older**		
	Immigrants	*LowGGI*	*HighGGI*	Immigrants	*LowGGI*	*HighGGI*
A: OLS						
MorethanTwo	-0.1470***	-0.1480***	-0.1448***	-0.1402***	-0.1368***	-0.1398***
	(0.0042)	(0.0037)	(0.0071)	(0.0100)	(0.0120)	(0.0094)
B: 2SLS						
MorethanTwo	-0.0745*	-0.1164*	-0.0596	-0.1579*	-0.2124	-0.1264
	(0.0427)	(0.0617)	(0.0560)	(0.0879)	(0.1693)	(0.0862)
F-Statistics	219.8	56.40	170.9	438.7	204.3	190.9
N	128,572	54,840	67,179	146,390	72,473	64,791

Note.– OLS and 2SLS estimated coefficient of *MorethanTwo* in equation (1) for other labor market outcomes. Covariates in the model but excluded from the table for brevity: age, age at first child, Black, Hispanic and Other races dummies (white as a reference group), 1st Born Boy, 2nd Born Boy, year and states dummies. Columns: those that migrated before 20 years old (columns (1)-(3)); migrated 20 years old or after (column (4)-(6)). *Immigrants:* those born outside from the US; *LowGGI:* those born outside of the US, that have GGI information, and GGI is below the average valued; *HighGGI:* those born outside of the US, that have GGI information, and GGI is above the average value. Standard errors clustered at birthplace in parentheses; * p<0.10, ** p<0.05, *** p<0.01.

Table A.5: Effect of Having More Than Two Children on Women's Employment Status: By Educational Attainment

	(1)	(2)	(3)	(4)	(5)
				Immigrants	
	All	Natives	Immigrants	*LowGGI*	*HighGGI*
Panel I: High School or Less					
A: OLS					
MorethanTwo	-0.1393***	-0.1310***	-0.1489***	-0.1460***	-0.1450***
	(0.0017)	(0.0020)	(0.0032)	(0.0045)	(0.0049)
B: 2SLS					
MorethanTwo	-0.0890***	-0.0620*	-0.1474**	-0.2268**	-0.0955
	(0.0307)	(0.0354)	(0.0605)	(0.0894)	(0.0861)
F-Statistics	1,156.3	865.7	300.5	137.9	149.7
N	573,939	421,399	152,540	76,183	67,970
Panel II: More than High School					
A: OLS					
MorethanTwo	-0.1379***	-0.1455***	-0.1057***	-0.1075***	-0.1144***
	(0.0014)	(0.0015)	(0.0040)	(0.0065)	(0.0054)
B: 2SLS					
MorethanTwo	-0.0537**	-0.0542**	-0.0323	-0.0525	-0.0503
	(0.0242)	(0.0252)	(0.0754)	(0.1515)	(0.0865)
F-Statistics	1732.9	1511.7	225.8	62.1	158.7
N	826,845	704,423	122,422	51,130	64,000

Note.– OLS and 2SLS estimated coefficient of *MorethanTwo* in equation (1) for other labor market outcomes. Covariates in the model but excluded from the table for brevity: age, age at first child, Black, Hispanic and Other races dummies (white as a reference group), 1st Born Boy, 2nd Born Boy, year and states dummies. Columns: *All:* include all the sample; *Natives:* include those born in the US (including US outlying territory); *Immigrants:* those born outside from the US; *LowGGI:* those born outside of the US, that have GGI information, and GGI is below the average value; *HighGGI:* those born outside of the US, that have GGI information, and GGI is above the average value. Rows: *Panel I:* estimation for women with "High School or Less" educational attainment; *Panel II:* estimation for women with "More than High School" educational attainment. Standard errors in parentheses (not cluster due smaller samples size); * p<0.10, ** p<0.05, *** p<0.01.

Table A.6: Effect of Having More Than Two Children on Women's Employment Status: Age 21-35

	(1)	(2)	(3)	(4)	(5)
				Immigrants	
	All	Natives	Immigrants	*LowGGI*	*HighGGI*
A: OLS					
MorethanTwo	-0.1479***	-0.1474***	-0.1447***	-0.1421***	-0.1425***
	(0.0043)	(0.0016)	(0.0084)	(0.0096)	(0.0079)
B: 2SLS					
MorethanTwo	-0.0927***	-0.0758**	-0.1276***	-0.2392***	-0.0707
	(0.0265)	(0.0305)	(0.0480)	(0.0619)	(0.0830)
F-Statistics	1,388.0	1,178.4	166.8	56.49	158.7
N	842,360	688,553	153,807	74,026	70,692

Note.– OLS and 2SLS estimated coefficient of *MorethanTwo* in equation (1) for other labor market outcomes. Covariates in the model but excluded from the table for brevity: age, age at first child, Black, Hispanic and Other races dummies (white as a reference group), 1st Born Boy, 2nd Born Boy, year and states dummies. Columns: *All:* include all the sample; *Natives:* include those born in the US (including US outlying territory); *Immigrants:* those born outside from the US; *LowGGI:* those born outside of the US, that have GGI information, and GGI is below the average value; *HighGGI:* those born outside of the US, that have GGI information, and GGI is above the average value. Standard errors clustered at birthplace in parentheses; * p<0.10, ** p<0.05, *** p<0.01.

Table A.7: Summary Statistics: 1980 and 1990 Censuses

	(1)	(2)	(3)	(4)	(5)	(6)
	1980 Census			**1990 Census**		
	All	Natives	Immigrants	All	Natives	Immigrants
Number of Children	2.54	2.53	2.62	2.46	2.44	2.57
	(0.83)	(0.82)	(0.95)	(0.75)	(0.73)	(0.88)
More than 2 Children	0.38	0.38	0.41	0.34	0.34	0.39
	(0.49)	(0.49)	(0.49)	(0.47)	(0.47)	(0.49)
1st Born Boy	0.51	0.51	0.51	0.51	0.51	0.51
	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)
2nd Born Boy	0.51	0.51	0.51	0.51	0.51	0.51
	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)
Two Boys	0.26	0.26	0.26	0.26	0.26	0.26
	(0.44)	(0.44)	(0.44)	(0.44)	(0.44)	(0.44)
Two Girls	0.24	0.24	0.24	0.24	0.24	0.24
	(0.43)	(0.43)	(0.43)	(0.43)	(0.43)	(0.43)
Same Sex	0.50	0.50	0.50	0.50	0.50	0.50
	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)
Age	31.57	31.51	32.29	32.50	32.46	32.87
	(4.72)	(4.71)	(4.74)	(4.60)	(4.60)	(4.60)
Age at 1st Child	21.72	21.61	22.95	22.80	22.75	23.24
	(3.33)	(3.27)	(3.71)	(3.95)	(3.94)	(4.01)
Employed	0.48	0.48	0.44	0.60	0.61	0.51
	(0.50)	(0.50)	(0.50)	(0.49)	(0.49)	(0.50)
N	601,794	554,260	47,534	616,481	555,098	61,383

Note.– Sample means with standard deviations in parentheses of the main variables. Columns: *All:* Include all the variables; *Natives:* Include those born in the US (including US outlying territory); *Immigrants:* those born outside from the US. Rows: *Number of Children:* number of children than the women has; *More than 2 Children:* equal to 1 if women has more than 2 children and 0 otherwise; *1st Born Boy:* equal to 1 if first born child is a boy; *2nd Born Boy:* equal to 1 if second born child is a boy; *Two Boys:* equal to 1 if first two children are boy; *Two Girls:* equal to 1 if first two children are girl; *Same Sex:* equal to 1 if first two children have same sex; *Age at 1st Child:* women's age at the birth of their first child; *Employment:* equal to 1 if their employment status is employed.

Table A.8: Effect of Having More Than Two Children on Women's Employment Status: 1980 and 1990 Censuses

	(1)	(2)	(3)	(4)	(5)	(6)
	1980 Census			**1990 Census**		
	All	Natives	Immigrants	All	Natives	Immigrants
A: OLS						
MorethanTwo	-0.1852***	-0.1856***	-0.1715***	-0.1763***	-0.1745***	-0.1682***
	(0.0029)	(0.0015)	(0.0108)	(0.0037)	(0.0016)	(0.0075)
B: 2SLS						
I: First Stage						
Same Sex	0.0545***	0.0544***	0.0564***	0.0587***	0.0590***	0.0570***
	(0.0014)	(0.0012)	(0.0057)	(0.0013)	(0.0013)	(0.0046)
II: Second Stage						
MorethanTwo	-0.1180***	-0.1092***	-0.2242***	-0.1248***	-0.1191***	-0.1860***
	(0.0203)	(0.0238)	(0.0792)	(0.0248)	(0.0238)	(0.0708)
F-Statistics	1,420.3	2,219.6	96.8	1,915.0	2,034.1	151.2
N	601,794	554,260	47,534	616,463	555,082	61,381

Note.– Covariates in the model but excluded from the table for brevity: age, age at first child, Black, Hispanic and Other races dummies (white as a reference group), 1st Born Boy, 2nd Born Boy, year and states dummies;; Standard errors clustered at birthplace in parentheses; * p<0.10, ** p<0.05, *** p<0.01.

Table A.9: Effect of Having More Than Two Children on Women's Employment Status: Controlling for Family Characteristics

	(1)	(2)	(3)	(4)	(5)
				Immigrants	
	All	Natives	Immigrants	*LowGGI*	*HighGGI*
A: OLS					
MorethanTwo	-0.1369***	-0.1386***	-0.1299***	-0.1300***	-0.1283***
	(0.0037)	(0.0012)	(0.0077)	(0.0088)	(0.0072)
B: 2SLS					
MorethanTwo	-0.0695***	-0.0549***	-0.1198**	-0.1735**	-0.1031**
	(0.0216)	(0.0209)	(0.0514)	(0.0837)	(0.0462)
F-Statistics	2,791.9	2,374.8	601.7	164.5	434.3
N	1,400,784	1,125,822	274,962	127,313	131,970

Note.– OLS and 2SLS estimated coefficient of *MorethanTwo* in equation (1) for other labor market outcomes. Covariates in the model but excluded from the table for brevity: age, age at first child, Black, Hispanic and Other races dummies (white as a reference group), 1st Born Boy, 2nd Born Boy, year and states dummies. Columns: *All:* include all the sample; *Natives:* include those born in the US (including US outlying territory); *Immigrants:* those born outside from the US; *LowGGI:* those born outside of the US, that have GGI information, and GGI is below the average value; *HighGGI:* those born outside of the US, that have GGI information, and GGI is above the average value. Standard errors clustered at birthplace in parentheses; * p<0.10, ** p<0.05, *** p<0.01.

Table A.10: Effect of Having More Than Two Children on Immigrant Women's Employment Status: *LowGGI* and *HighGGI* Using Average GGI per Year

	(1)	(2)	(3)	(4)
	OLS		**2SLS**	
	LowGGI	*HighGGI*	*LowGGI*	*HighGGI*
MorethanTwo	-0.1390***	-0.1364***	-0.1418*	-0.0947*
	(0.0092)	(0.0089)	(0.0782)	(0.0496)
F-Statistics			209.2	382.6
N	146,674	112,609	146,674	112,609

Note.– OLS and 2SLS estimated coefficient of *MorethanTwo* in equation (1) for other labor market outcomes. Covariates in the model but excluded from the table for brevity: age, age at first child, Black, Hispanic and Other races dummies (white as a reference group), 1st Born Boy, 2nd Born Boy, year and states dummies. Columns: *LowGGI:* those born outside of the US, that have GGI information, and GGI is below the average value; *HighGGI:* those born outside of the US, that have GGI information, and GGI is above the average value. Standard errors clustered at birthplace in parentheses; * p<0.10, ** p<0.05, *** p<0.01.

Table A.11: Effect of Having More Than Two Children on Immigrant Women's Employment Status: Controlling for Education Ratio

	(1)	(2)	(3)
		Immigrants	
	Immigrants	*LowGGI*	*HighGGI*
A: OLS			
MorethanTwo	-0.1398***	-0.1416***	-0.1381***
	(0.0026)	(0.0038)	(0.0037)
B: 2SLS			
MorethanTwo	-0.0984**	-0.1637**	-0.0694
	(0.0481)	(0.0786)	(0.0630)
F-Statistics	508.4	194.9	290.8
N	263,066	123,627	130,242

Note.– OLS and 2SLS estimated coefficient of *MorethanTwo* in equation (1) for other labor market outcomes. Covariates in the model but excluded from the table for brevity: age, age at first child, Black, Hispanic and Other races dummies (white as a reference group), 1st Born Boy, 2nd Born Boy, year and states dummies. Columns: *Immigrants:* those born outside from the US; *LowGGI:* those born outside of the US, that have GGI information, and GGI is below the average valued; *HighGGI:* those born outside of the US, that have GGI information, and GGI is above the average value. Standard errors in parentheses (not cluster due smaller samples size); * p<0.10, ** p<0.05, *** p<0.01.

Table A.12: Effect of Having More Than Two Children on Immigrant Women's Employment Status: Controlling for Country Characteristics

	(1) *LowGGI*	(2) *HighGGI*	(3) *LowGGI*	(4) *HighGGI*	(5) *LowGGI*	(6) *HighGGI*
A: OLS						
MorethanTwo	-0.1342***	-0.1405***	-0.1392***	-0.1329***	-0.1391***	-0.1333***
	(0.0111)	(0.0085)	(0.0082)	(0.0143)	(0.0082)	(0.0148)
B: 2SLS						
MorethanTwo	-0.1842**	-0.0797*	-0.1876***	-0.0572	-0.1967***	-0.0538
	(0.0851)	(0.0425)	(0.0707)	(0.0397)	(0.0749)	(0.0405)
GDP per Capita, Fertility	✓	✓	✓	✓	✓	✓
Educational Attainment			✓	✓	✓	✓
Immigration rates					✓	✓
F-Statistics	195.2	428.2	241.5	365.7	232.1	419.8
N	123,028	131,838	127,815	121,637	125,130	124,322

Note.– OLS and 2SLS estimated coefficient of *MorethanTwo* in equation (1) for other labor market outcomes. Covariates in the model but excluded from the table for brevity: age, age at first child, Black, Hispanic and Other races dummies (white as a reference group), 1st Born Boy, 2nd Born Boy, year and states dummies. Columns: *LowGGI:* those born outside of the US, that have GGI information, and GGI is below the average value; *HighGGI:* those born outside of the US, that have GGI information, and GGI is above the average value. Standard errors clustered at birthplace in parentheses; * p<0.10, ** p<0.05, *** p<0.01.

Table A.13: Effect of Having More Than Two Children on Immigrant Women's Employment Status: Excluding Sex Ratio Imbalance Countries

	(1) Immigrants	(2) *LowGGI*	(3) *HighGGI*	(4) Immigrants	(5) *LowGGI*	(6) *HighGGI*
Panel I: Sex Ratio 1.07 or less						
A: OLS						
MorethanTwo	-0.1466***	-0.1428***	-0.1429***	-0.1455***	-0.1427***	-0.1416***
	(0.0028)	(0.0043)	(0.0039)	(0.0028)	(0.0043)	(0.0039)
B: 2SLS						
MorethanTwo	-0.1315**	-0.2390**	-0.0842	-0.1268**	-0.2384**	-0.0807
	(0.0541)	(0.0931)	(0.0692)	(0.0543)	(0.0931)	(0.0694)
Sex Ratio				✓	✓	✓
F-Statistics	387.0	130.9	238.2	384.7	130.8	237.5
N	212546	88420	113847	212546	88420	113847
Panel II: Sex Ratio 1.099 or less						
A: OLS						
MorethanTwo	-0.1444***	-0.1429***	-0.1413***	-0.1449***	-0.1416***	-0.1403***
	(0.0028)	(0.0041)	(0.0039)	(0.0028)	(0.0041)	(0.0039)
B: 2SLS						
MorethanTwo	-0.1091**	-0.1850**	-0.0728	-0.1094**	-0.1861**	-0.0706
	(0.0536)	(0.0909)	(0.0686)	(0.0538)	(0.0906)	(0.0691)
Sex Ratio				✓	✓	✓
F-Statistics	401.8	141.3	244.1	400.7	143.1	242.1
N	225,781	99,421	115,983	225,781	99,421	115,983

Note.– OLS and 2SLS estimated coefficient of *MorethanTwo* in equation (1) for other labor market outcomes. Covariates in the model but excluded from the table for brevity: age, age at first child, Black, Hispanic and Other races dummies (white as a reference group), 1st Born Boy, 2nd Born Boy, year and states dummies. Columns: *Immigrants:* those born outside from the US; *LowGGI:* those born outside of the US, that have GGI information, and GGI is below the average value; *HighGGI:* those born outside of the US, that have GGI information, and GGI is above the average value. Standard errors in parentheses (not cluster due smaller samples size); * p<0.10, ** p<0.05, *** p<0.01.

A.1 Correlation between GGI and Measures of Gender Attitudes

I explore the correlation of the GGI with other relevant gender attitudes measures. I collect gender attitudes measures available on the General Social Survey (GSS) and the Joint European Value Survey and World Values Survey (EVS/WVS).

GSS.– The GSS is a survey in the US for every year until 1994 and every other year after that. The goal of the survey is to understand attitudes in the US, from crime and

civil liberty to stress. I use the GSS for 1972–2018 and advantage questions regarding individuals' gender attitudes. To measure an individual's origin, I use the reported country of birth for the 2004 survey and the individual's ethnicity for the rest of the years.

To measured gender attitudes, I explore the following six questions:

1. *Working Mom can have warm relationship with kids* ($WorkMomWarmRel$)
2. *Preschooler suffer if mom works* ($KidSuffer$)
3. *family suffer if mom works* ($FamSuffer$)
4. *Gender role reversal hurts family* ($GenderRole$)
5. *being housewife as fulfilling as paid work* ($Housework$)
6. *Husb should work; wife should look after home* ($HusbandWork$)

For each question, the interviewer answer from 1 (Strongly agree) to 5 (Strongly disagree).[16] Using the GSS country of origin/ethnicity, I create for each question a variable that is the average answer by country. Matched by country, for those available, I add these variables to my main sample.

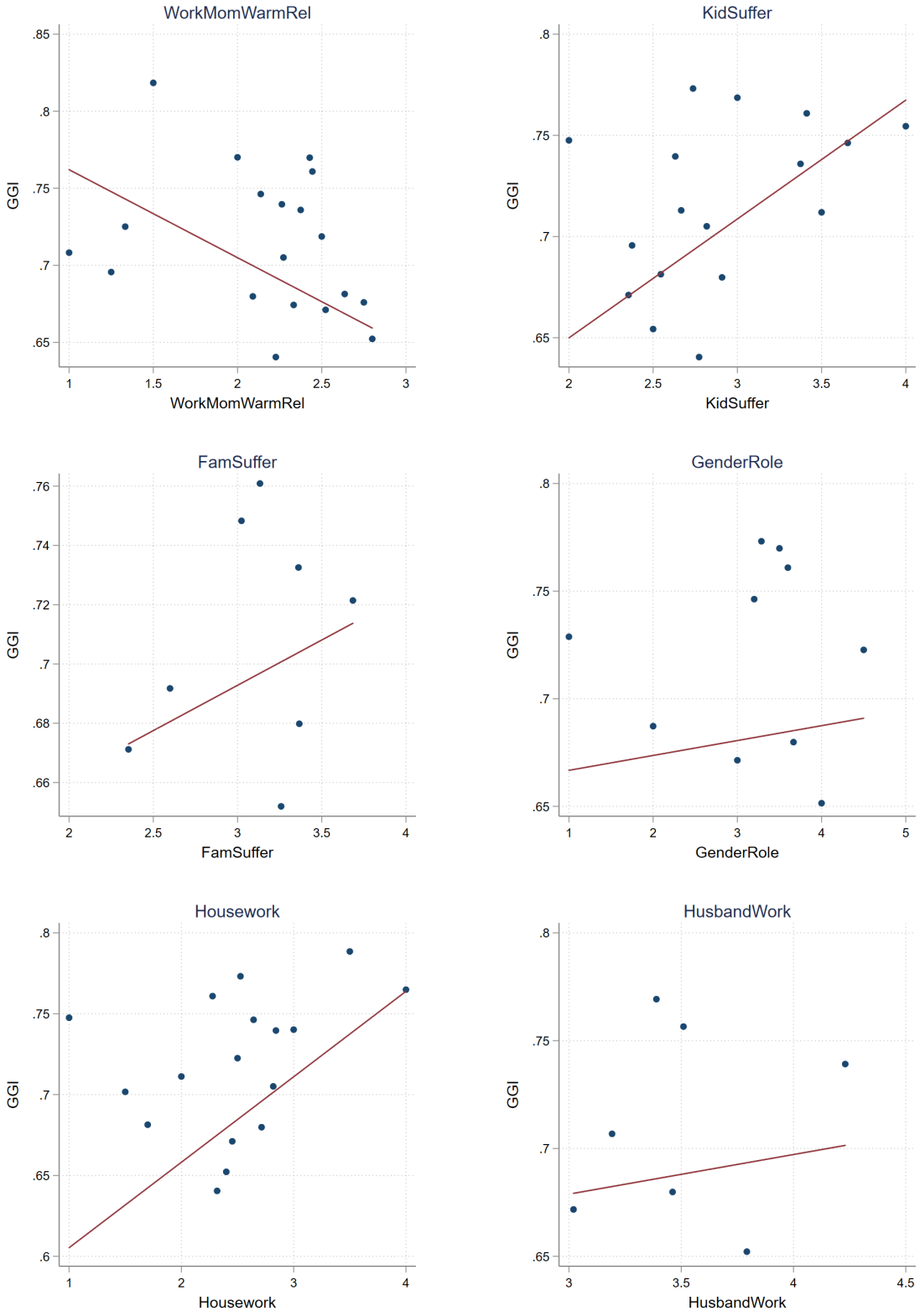
I then explore the GGI relationship with these variables. If the GGI reflects individuals' gender attitudes, then the GGI should be correlated with the gender attitudes measured through the GSS answers. Moreover, based on the phrasing of the questions and because the larger the GGI, the more gender-egalitarian a country is, the correlation between the GGI and $WorkMomWarmRel$ (question 1) should be negative. In contrast, questions 2–6 should positively relate to the GGI.

Figure A.1 shows the binned scatterplots of the average country or origin answer to the GSS's question and the GGI of that country in my main sample. The relationships depicted in the scatterplots show the expected pattern, with individuals' gender attitudes moving in the same direction as the GGI. Furthermore, Table A.14 shows the OLS results of estimating the GGI on the variables from the GSS separately. Even though not all results are significant, the signs are in line with the expected results.

Overall, the results show that the GGI qualitatively relates to other gender attitudes measurements from the GSS. Thus, gender egalitarianism measured as country of origin's gender gap relates to immigrants' gender egalitarianism.

[16] Precisely, 1=Strongly agree; 2=Agree; 3=Neither agree nor disagree; 4=Disagree; 5=Strongly disagree.

Figure A.1: Binned Scatterplots Between the GGI and the GSS variables



Note. Figures show the binscatter graphing the relationship of each of the GSS variables with the GGI. *Observations:* N=(163,692; 163,692; 163,875; 162,521; 163,692; 163,975)

Table A.14: Association Between The GGI and The GSS's Gender Attitudes Measurements

Explanatory Variable	Coef/SE	Observations
WorkMomWarmRel	-0.0571	163,692
	(0.0341)	
KidSuffer	0.0587***	163,692
	(0.0168)	
FamSuffer	0.0306	163,875
	(0.0240)	
GenderRole	0.00692	162,521
	(0.0210)	
Housework	0.0528	163,692
	(0.0393)	
HusbandWork	0.0184	163,975
	(0.0424)	

Note.– OLS estimation on the relationship of the GGI with the GSS variables measuring gender attitudes. Each Row is a separate regression; * p<0.10, ** p<0.05, *** p<0.01; Standard Errors cluster at the country level.

Joint European Value Survey and World Values Survey (EVS/WVS).– Similarly, I use the EVS/WVS for the years 2017–2020. Both surveys do cross-countries surveys that ask questions regarding people's attitudes. Using the country of the survey, I add the information of the average answer per country t to the following four questions:

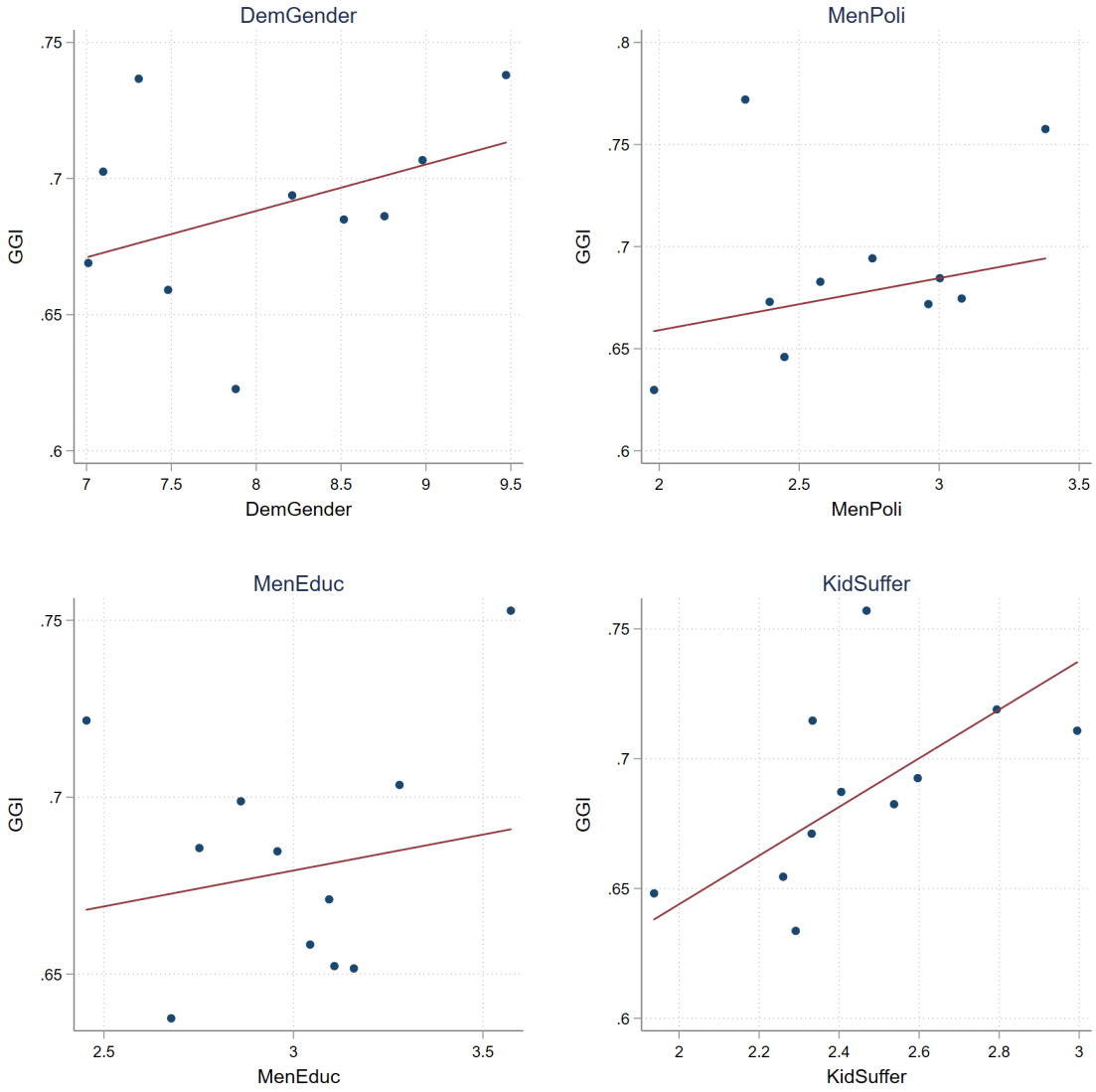
1. *Democracy: Women have the same rights as men.* ($DemGender$)
2. textitMen make better political leaders than women do ($MenPoli$)
3. *University is more important for a boy than for a girl* ($MenEduc$)
4. *Pre-school child suffers with working mother* ($KidSuffer$)

DemGender ranges from 1 ("Not an essential characteristic of democracy") to 10 ("An essential characteristic of democracy"). while the rest of the questions range from 1 (agree strongly) to 4 (strongly disagree).[17] When available, I add the average answer per country. I then look at the relationship between these variables and the GGI.

If the GGI correlated with gender attitudes across countries, the association between the GGI and the EVS/WVS questions should be positive. The resulting binned scatterplots, in Figure A.2, and OLS estimation, in Table A.15, are in line with the finding from the GSS. Thus, it shows that gender egalitarianism measured through the GGI aligns with other gender attitudes measurements.

[17]Precisely, 1=Agree strongly; 2=Agree; 3=Disagree; 4=Strongly disagree.

Figure A.2: Binned Scatterplots Between the GGI and the EVS/WVS variables



Note. Figures show the binscatter graphing the relationship of each of the GSS variables with the GGI. *Observations:* N=(195,677)

Table A.15: Association Between The GGI and The EVS/WVS's Gender Attitudes Measurements

Explanatory Variable	Coef/SE	Observations
DemGender	0.0159***	195,677
	(0.0059)	
MenPoli	0.0264	195,677
	(0.0327)	
MenEduc	0.0205	195,677
	(0.0574)	
KidSuffer	0.0933***	195,677
	(0.0279)	

Note.– Each Row is a separate regression; * p<0.10, ** p<0.05, *** p<0.01; Standard Errors cluster at the country level.

A.2 Compliers Characteristics

Based on the Abadie (2003)'s Theorem 3.1, we can identify the compliers through a weighting function, kappa weight (κ):

$$\kappa = 1 - \frac{D_i(1 - Z_i)}{1 - Pr(z_i = 1|x_i)} - \frac{(1 - D_i)Z_i}{Pr(z_i = 1|x_i)}, \tag{3}$$

where D_i is the endogeneous variable of interest which in this case is $MorethanTwo_i$. Z_i is the IV variable, which here is $SameSex_i$. Moreover, $Pr(z_i = 1|x_i)$ is estimated using a probit model estimating the probability of having the first two children of the same sex, conditional on the characteristic x_i. Then, compliers' average characteristic x_i can be calculated by the weighted average:

$$E[x_i|compliers] = \frac{E[\kappa x_i]}{E[\kappa]}. \tag{4}$$

Using this technique, I calculate the compliers' characteristics and compare them with the population averages. The results are in Table A.16, where each column representing each of the groups as defined before. The characteristics included are women's age, marital status, educational outcomes, and years in the US (for Immigrants). I also include the first two children's age. Panel A presents the population means for these variables, while Panel B presents the compliers' means. To understand how the compliers compare to their population, I take the ratio of compliers to population, which is in Panel C.

Looking at Panel B, we can see that the compliers' characteristics are similarly comparable across all the groups, which eases the concern that the resulting findings represent

very different samples from each other. There is a difference that is important to note in terms of educational attainment. While years of education are similar across groups, when comparing educational attainment distribution, natives complier are more likely to have more than high school educational attainment than immigrants. These differences, however, could represent differences in the population that the compliers represent. Thus, differences do not imply that compliers compare to their population in similar ways. To understand this, we can look at Panel C, which takes the ratio of complier to the population average. In general, we can see that the compliers to population ratios are very similar across the different groups. Therefore, differences in the results do not seem to be a consequence of large differences in samples' compliers.

Although ratios are similar across subsamples in Panel C, It is worth noting that the salient difference in the compliers' education distribution compared with their population.

Table A.16: Compliers Characteristics

	(1)	(2)	(3)	(4)	(5)
	All	Natives	Immigrants	*LowGGI*	*HighGGI*
Panel A: Population					
Age	33.59	33.46	34.13	33.89	34.39
Married	0.79	0.77	0.85	0.87	0.84
Years of Education	13.50	13.80	12.27	11.87	12.65
High School or Less	0.41	0.37	0.55	0.60	0.52
More than High School	0.59	0.63	0.45	0.40	0.48
Years in the US			15.13	14.12	16.28
Age Older Child	9.71	9.64	9.98	9.97	10.04
Age Second Child	6.24	6.23	6.29	6.32	6.28
Panel B: Compliers					
Age	34.45	34.29	35.24	34.81	35.57
Married	0.85	0.84	0.90	0.91	0.90
Years of Education	13.42	13.70	12.08	11.53	12.53
High School or Less	0.42	0.38	0.58	0.64	0.54
More than High School	0.58	0.62	0.42	0.36	0.46
Years in the US			15.85	14.94	16.72
Age Older Child	10.87	10.82	11.08	11.00	11.15
Age Second Child	7.73	7.75	7.66	7.54	7.76
Panel C: Compliers/Population					
Age	1.03	1.02	1.03	1.03	1.03
Married	1.08	1.09	1.06	1.05	1.07
Years of Education	0.99	0.99	0.99	0.97	0.99
High School or Less	1.02	1.03	1.05	1.08	1.05
More than High School	0.99	0.98	0.94	0.88	0.95
Years in the US			1.05	1.06	1.03
Age Older Child	1.12	1.12	1.11	1.10	1.11
Age Second Child	1.24	1.24	1.22	1.19	1.24

Note.– The table report the mean of the population and the compliers. Columns: *All:* include all the sample; *Natives:* include those born in the US (including US outlying territory); *Immigrants:* those born outside from the US; *LowGGI:* those born outside of the US, that have GGI information, and GGI is below the average value; *HighGGI:* those born outside of the US, that have GGI information, and GGI is above the average value. Rows: *Panel A:* include the complete sample average; *Panel B:* include the compliers sample average, calculated by equation (4); *Panel C:* divided the Panel B by Panel A to take the compliers to population ratio. The results are based on the main sample of women 21-40 old women with two or more children. The observations use are $N =$(1,400,784; 1,125,822; 274,962; 127,313; 131,970)