Homemaker Norms for Married Women: A Cross-Country Comparison

Jay Euijung Lee*

August 21, 2020

Preliminary and incomplete - please do not cite

Abstract

The strength of the norm that prescribes married women to be homemakers varies across countries. This paper quantifies the homemaker norm for married women for seventeen countries. The homemaker norms wedge, calibrated based on the model proposed by Lee (2020), measures how much a 10-dollar market wage is valued at for married women, when they are deciding between working in the labor market versus working at home. The wedges are computed from the gap in the labor force participation of married women and that of similar single women that is not explained by the wage differentials. The homemaker norms wedge is found to be large, amounting to half of the parallel wedge that exists between men and women. If the homemaker norms wedges were to disappear, female labor force participation in the seventeen countries would increase by 21% on average. I show that cultural differences account for 45% of the cross-country variation in the homemaker norms wedge.

*Department of Economics, London School of Economics; email: E.Lee13@lse.ac.uk.

1 Introduction

A low level of female labor force participation implies that a large portion of human talent rests unrealized. Countries around the world are exposed to this issue at widely varying degrees, with female labor force participation ranging from 6% in Yemen to 84% in Rwanda in 2019 (World Bank Statistics). A large literature points to varying gender norms on female employment as a driver of this variation (Alesina, Giuliano, and Nunn, 2013; Jayachandran, 2015, Almond and Edlund, 2008; Abrevaya, 2009; Fernández and Fogli, 2009). Meanwhile, the low rate of female labor force participation is further exacerbated when we consider married women.[1] On top of the differential norm governing the appropriate behavior for men versus women, Lee (2020) suggests that *married* women face a further constraint against participating in the labor market relative to single women.

This paper seeks to highlight the role of marriage as a factor driving international differences in female labor force participation. Adopting Lee (2020)'s computation of the "homemaker norms wedge" applying to married women, I verify that the homemaker norms wedge varies across countries.[2] The norms wedges measure how much a 10-dollar market wage is valued at for married women, when they are deciding between working in the labor market versus working at home. The wedges are computed from the gap in the labor force participation of married women and that of similar single women that is not explained by the wage differentials.

Many theories aiming to explain why the labor force participation of females varies across countries look to differences in economic conditions. First, where men have comparative advantage in the more physically-intensive agriculture and manufacturing sectors and women in the more mentally-intensive service sector, the sectoral composition in the economy affects relative female labor productivity (Goldin, 1995; Galor and Weil, 1996; Pitt, Rosenzweig, and Hassan, 2012; Ngai and Petrongolo, 2017). Second, there are differences in women's earning capacity arising from gender gaps in educational attainment, due to lower expected employment opportunities or son preferences (Heath and Jayachandran, 2017). Third, the burden of home production and child care or child bearing, which keeps women at home, varies by country. Economic development abates these burdens, through electrification (Dinkelman, 2011), invention and spread of household appliances (Greenwood et al., 2005; Coen-Pirani et al., 2010), lower fertility (Miller, 2010), improved child care facilities, and medical advancement in maternal health care (Albaseni and Olivetti, 2009, Jayachandran and Lleras-Muney, 2009).

On the other hand, the homemaker norms wedge of Lee (2020) focuses on the differences not between men and women, but between married and single women. Therefore, sectoral

[1] Among the country censuses of the 2010s compiled in the IPUMS International database, the ratio of married to single women's labor force participation was as low as 0.38 in Iran in 2011. The average ratio is 0.85.

[2] From here on, the homemaker norms wedge is sometimes referred to as plainly, the "norms wedge."

transitions that favor women in general do not affect the norms wedge. Moreover, the norms wedges are computed by matching a group of married women to a group of single women with the same level of education and family composition – number of children under the age 18 and under the age 5 – so that earning capacity or the burden of home production is more or less kept similar in the two groups of women. Nonetheless, even with education and family composition matched, married women's labor force participation typically falls behind single women's. It is the norms wedge that picks up this married-single gap. Thus, the wedges are summative measures that encapsulate all the reasons besides wages for which the gap arise.

Using the country censuses compiled in the IPUMS International database, I compute the homemaker norms wedges for seventeen different countries and find that they exhibit a wide cross-country variation. The norms wedge is almost negligible in Canada, implying that a 10-dollar market wage is valued at very close to 10 dollars for married women in Canada when they make their labor force participation decisions. In contrast, a 10-dollar market wage is only valued at 5.7 dollars by the married women in Indonesia. Because the IPUMS International database includes censuses taken at different points in time, the country-average of the norms wedges over all available census years reflect not solely the cross-country differences but also the time series differences; the norms wedges do change over time in a given country.[3] However, as the wedges are slow to change over time, the cross-country variation eclipses the within-country variation, and so I maintain the main focus of this paper on the cross-country variation. Furthermore, the homemaker norms wedges are large in magnitude, amounting to around a half of the "gender norms wedge", the men-women counterpart to the homemaker norms wedge; the gender norms wedge rationalizes by how much the labor force participation of women falls behind the level predicted if they behaved like similar men.

Although the homemaker norms wedges compound all the reasons besides wages that create a gap in the labor force participation between married and single women, I find that culture plays an important part in explaining the cross-country variation in the wedges. The norms wedges are strongly correlated with answers to questions on the appropriate behavioral prescriptions for men and women, asked in the World Values Survey. This relationship persists even within-country, when I exploit the change in the norms wedges and attitudinal survey answers over time for the countries surveyed at multiple time periods. However, the norms wedges are not correlated with a comparable placebo question about attitudes on immigrants.

To isolate the effect of culture through an epidemiological approach, I repeat the exercise for groups of married and single women of different ancestry among the U.S. population. The idea behind this approach is that the effect of culture can be identified through the variation in economic outcomes of individuals who share the same economic and institutional environment, but whose social beliefs are potentially different (Fernández, 2011). Indeed, the norms wedges

[3]Lee (2020) documents how the homemaker norms wedge changes over time in the United States.

by ancestry are also found to be correlated with gender attitudes, and they account for around 45% of the cross-country variation in the norms wedges.

Finally, I quantify how much of the cross-country variation in female labor force participation is driven by the variation in the homemaker norms wedges. I compute the counterfactual female labor force participation for the scenario where the norms wedges disappear altogether. The labor force participation of married women increases by 38% on average in the counterfactual. Taking married and single women together, female labor force participation increases by 21% on average. The cross-country variance in female labor force participation decreases by 8%.

The rest of this paper is organized as follows. Section 2 explains the data sources used. Section 3 then discusses what the homemaker norms wedges mean and how they are computed. The resulting wedge values are discussed in Section 4. Next, I check that culture is pertinent for the norms wedges in Section 5. Section 6 isolates the effect of culture through the epidemiological approach and discusses how much of the cross-country variation in the norms wedges is due to cultural differences. Lastly, I return to the motivation of the large variation in female labor force participation across countries, and I examine how much of it is driven by differences in the norms wedges. Section 8 concludes.

2 Data

2.1 Country censuses

The data used to compute the norms wedges for each country comes from the IPUMS International database, which compiles the censuses of numerous countries. A few countries' household survey data are also included. The main advantage of this dataset is that the variables present in each country's census are harmonized. This feature allows the model to be applied to each country in a consistent manner. Out of the countries in the database, I select the ones which contain all the variables necessary to apply the model of Lee (2020): market earnings, labor force participation status, marital status, education, children, and the code linking a person to their spouse. Although some 98 countries are listed as part of the IPUMS International database, the earnings variable is not available broadly and therefore is the most limiting factor in the country selection.

The countries used are Brazil, Canada, Colombia, Dominican Republic, India, Indonesia, Israel, Italy, Jamaica, Mexico, Panama, South Africa, Trinidad and Tobago, United States, Uruguay, and Venezuela.[4] Table 1 lists the countries and census years for which all the variables needed to compute the norms wedges are available. The main focus of this paper is the cross-

[4]Germany is another country for which earnings data is available in its 1970 census, but the spousal identifier is missing there.

country variation in the norms wedges, since it markedly dominates within-country time series variation.[5] Therefore, when I conduct the cross-country analysis, I will work with the average over the years by country. I do make explicit use of the country-year panel structure when I demonstrate the strength of the correlation between norms wedges and gender attitudes by showing that the correlation is robust to the addition of country fixed effects in Section 5.1.

Appendix Section A discusses in greater detail how I harmonize the variables are heterogeneous across countries.

Table 1: List of census or survey

Country	Census/survey year(s)
Brazil	1980, 1991, 2000, 2010
Canada	2011
Colombia	1973
Dominican Republic	1981, 2002
India†	1983, 1987, 1993, 1999, 2004
Indonesia	1976, 1995
Israel	1972, 1995
Italy†	2011‡
Jamaica	1982, 1991, 2001
Mexico	1970, 1990, 1995, 2000, 2010, 2015
Panama	1970, 1980, 1990, 2000, 2010
Puerto Rico	1990, 2000, 2005, 2010
South Africa	1996, 2001, 2007, 2011
Trinidad and Tobago	1970, 2000
United States	1980, 1990, 2000, 2010
Uruguay	2006
Venezuela	1971, 1981, 1990, 2001

Notes: † Surveys, not censuses

‡ I combined the survey data from years 2011, 2012, ..., 2018, due to the small sample size in each year.

2.2 Attitudinal measures

The second dataset I employ is the World Values Survey (WVS), which asks various attitudinal survey questions to respondents in countries around the world. Of different surveys on attitudes, the WVS is well known to be the one with the greatest coverage globally. The WVS allows

[5] I compare the cross-country and time series variation later in Section 5.2.

me to check whether the norms wedges are in line with conventional measures of attitudes. An example of the ideal attitudinal survey question is, "Do you approve or disapprove of a married woman earning money in business or industry if she has a husband capable of supporting her?", asked in the General Social Survey. The question directly asks about the appropriate role of the married woman. Unfortunately, such a question is not asked in the WVS.[6]

Instead, to measure gender attitudes, I make use of the survey question that asks whether the respondent agrees or disagrees with the statement: "When jobs are scarce, men have more right to a job than women." This particular question is especially appealing for two reasons. Firstly, it is consistently asked for the greatest number of survey waves and therefore allows more country- and year-specific norms wedges to be matched to the attitudinal measure.[7] Secondly, it is asked just alongside a similar yet very different statement: "When jobs are scarce, employers should give priority to [own nationality] people over immigrants." Both statements revolve around job scarcity and are phrased similarly, and the answer categories are the same. Yet, the former asks about gender and the latter about immigrants. Therefore, the latter would be useful as a check for whether any correlation between norms wedges and attitudes on gender, if present, is a fluke. If a similar correlation exists between the wedges and the attitudes on immigrants, then it might merely be that the wedges reflect *general* conservative values.

2.3 Ancestry information among U.S. immigrants

To what extent can the variation in norms wedges by country be explained by differences in culture? In order to break apart culture from other economic and institutional factors affecting the norms wedge values, I consider computing the norms wedges by ancestry among the U.S. population. The idea is that groups of workers in the U.S. who operate within the same economic and institutional conditions might display different labor force participation patterns due to different cultural beliefs that stem from different ancestry. This procedure is performed on the United States censuses in the IPUMS International database, using the variable *ancest*. *Ancest* records the respondent's self-reported ancestry or ethnic origin, where typical examples include German, English, Irish, and Mexican. If the respondent gave multiple ancestries, then *ancest* records the first response. Since *ancestr* is self-reported, it captures the ethnic origin that the respondent identifies with. It therefore would sensibly represent the culture that the respondent's actions and choices are influenced by.

[6]The gender-related questions that are asked in the WVS are typically about what is expected from a woman in general, or a (working) mother in general, not distinguishing married and single women.

[7]The question was asked in all WVS waves but the first. Each survey wave covers 3-5 years, and the second wave covered 1990-1994.

3 Computation of norms wedges

This section discusses how the homemaker norms wedges are computed for each country and year. Before I proceed on to the computation results, I provide a brief background for what the norms wedge means, how it is derived from the model, and which assumptions need to be made. In particular, I highlight which factors are assumed to be the same across countries, and which factors are allowed to be country-specific, when I take the one model to different country datasets.

3.1 Derivation of the homemaker norms wedge

To explain what the norms wedge means and how it is derived, I reproduce the key pertinent parts of the model in Lee (2020). The norms wedge measures how much a 10-dollar market wage is valued at for married women, when they are deciding between working in the labor market versus working at home. A married woman f's labor force participation decision is represented by:

$$L_f^* = \mathbb{1}\left[(1-\tau)w_f \geq h_f\right] \tag{1}$$

where w_f denotes her market wage, h_f her home productivity, and τ the norms wedge. For example, with $\tau = 0.4$, a married women who would make \$20 by working in the market behaves as if she would earn \$12. Furthermore, she would only work if her home productivity is lower than \$12.

The labor force participation decision of married women can be contrasted against that of a single woman i:

$$L_i^* = \mathbb{1}\left[w_i \geq h_i\right] \tag{2}$$

In order to identify the τ in equation (1), I need to know the labor force participation decision and potential market earnings and home productivity of married women. Labor force participation is observed in the data. Also observed in the data are market earnings of those who chose to work, from which potential market earnings can be deduced accounting for selection into the labor market. However, home productivity is *not* observed. Therefore, it is necessary to find a counterfactual group from which home productivity can be extracted. As shown in equation (2), single women's labor force participation decision in the model is based on market earnings and home productivity alone, with no wedge involved, so I can back out the potential home productivity of single women. Assuming that single and married women who are similar in terms of education and family composition – number of children under the age 18 and under the age 5 – share the same level of home productivity, I can finally back out the value of τ.

All in all, the computation of the norms wedges relies on figuring out what the norm wedge

must have been, to rationalize by how much the labor force participation of married women falls behind the level predicted if they behaved like similar single women. The norms wedges are defined as:

$$\tau = 1 - \frac{avrwage_{\text{ single women}}}{avrwage_{\text{ married women}}} \left(\frac{1 - LFP_{\text{single women}}}{1 - LFP_{\text{married women}}} \right)^{\frac{1}{\theta}} \tag{3}$$

where *avrwage* denotes the average wage of the workers among a group of married or single women, and LFP the labor force participation rate of that group. θ is an inverse measure of the dispersion of idiosyncratic abilities in the population, which parameterizes the selection into market work.

3.2 Taking the model to different countries

When I compute the norms wedges across countries, I take equation (3) and compute it using the data for each country.

Is comparing the norms wedge value of country A and country B tantamount to comparing apples and oranges? From equation (3), it is clear that the norms wedge value has a closed form solution. Thus, when I compare the norms wedges of two countries, I are comparing exactly that in each country. The norms wedge can therefore be thought of as a summative measure the encapsulates all the reasons for which the labor force participation rate of a group of married women differs from that of a group of similar single women in a country, apart from wage differentials. The fact that the norms wedge is a summative measure is an advantage in itself, allowing for a simple way to compare different countries.

Moreover, cross-country differences in numerous factors would not confound the norms wedge values, as long as those factors affect married and single women of a country in the same way. For instance, total factor productivity (TFP) and thus the wage levels of the labor force are different by country, but that does not affect the value of the norms wedge, as long as similar married and single women operate in comparable labor markets. This feature also means that differences in data collection methodology by country do not matter for the norms wedge value. Moreover, in matching a group of married women to a group of single women by education and family composition, the model already takes into account that there are differences in education levels and child-bearing propensities by country.

However, which exact ingredients go in to diverging norms wedges by country remains a black box, without tailoring the model in substantial ways to each country. Therefore, the extent to which the norms wedge value can stay true to its name – that it can be interpreted as a measure of the societal norm against married women working – would vary by country. For example, if there are positive returns to experience and the risk of divorce is high in a country, a

married woman would be more likely to work in the market *not* because of modern gender role values but as a precautionary action. Depending on the likelihood of divorce by country, how well the norms wedge value truly reflects the societal norm would vary. As another example, if the national welfare system is weak and market earnings are volatile (such as in seasonal or agricultural jobs), a married woman would be more likely to work in order to secure an additional source of income for the rainy days when the husband cannot provide. Then, the norms wedges would also reflect differences in the national welfare system.[8]

4 Norms wedge values

4.1 Aggregating homemaker norms wedges at the country level

The calibration exercise involves backing out the value of τ for each (education-family composition) match between married and single women. Hence, the value of τ is sensitive to the size of each group, or cell, of married and single women. As it can be seen in equation (3), τ is based on group means: the labor force participation rate and the average of the wages earned by the market workers. The larger the cell size of both married and single women in a given (education-family composition) match, the more precise and stable is τ for that match. On the contrary, with small cell sizes, the more likely it is that the τ value is unreliable. In fact, as evident in equation (3), τ is not constrained to be between 0 and 1; rather, the range of τ is $(-\infty, 1]$. It is thus possible that an imprecise τ takes a very low negative value that is clearly implausible. In the IPUMS International database, different countries have different numbers of observations, and so the countries with small populations or small census sampling would be more prone to this issue.

Therefore, in putting the τ values by group together at the country-year level, many summary statistics can be considered. The summary statistics I consider include the mean and median that are unweighted or weighted by the cell size – either the raw observation number per cell or the combined person weights in the census per cell. In addition, I consider first truncating the values of the group-specific τ at the top and bottom 1%, 5%, and 10%, and then looking at the unweighted or weighted mean and median. The last summary statistic I consider is the unweighted or weighted mean only among τ values computed with cells in which the number of observations exceeds an arbitrary threshold, 100.

Table 2 demonstrates how correlated the various summary statistics are, where each summary statistic is at the country-year level. The mean values tend to be quite sensitive to the weighting, truncation, or restriction criteria, due to the presence of outlier values. The median

[8] In Appendix Section A, I discuss the caveat that the earnings variable in the IPUMS International censuses are before-tax numbers.

values on the other hand, are quite robust to the different criteria. Moreover, it can be seen that truncating the top and bottom values of τ within a country barely makes a difference, whereas weighting does. Overall, the median values are highly correlated with one another. The selection of which median value to use consistently throughout the rest of the analysis therefore becomes tricky. I select the one that is the most stable over time within countries, i.e. the one with the lowest variance, within country and across years. This statistic is the unweighted median τ conditional on cell sizes exceeding 100 observations, which corresponds to ⑩ in Table 2.

To arrive at the norms wedge value at the country level, I take the average of the country-year summary statistic ⑩ across years for each country.

Table 2: Correlation coefficients of various summary statistics for norms wedges

centrality	mean	median	mean	median	mean	median	mean	median	mean	median
weighting	-	-	-	-	obs no.†	obs no.†	census wt.‡	census wt.‡	-	-
truncation	-	-	5%	5%	5%	5%	5%	5%	-	-
restriction	-	-	-	-	-	-	-	-	>100§	>100§
	①	②	③	④	⑤	⑥	⑦	⑧	⑨	⑩
①	1.000									
②	0.781	1.000								
③	0.921	0.948	1.000							
④	0.761	0.997	0.938	1.000						
⑤	0.783	0.900	0.914	0.902	1.000					
⑥	0.655	0.859	0.828	0.863	0.965	1.000				
⑦	0.779	0.901	0.912	0.905	0.998	0.967	1.000			
⑧	0.660	0.853	0.825	0.858	0.959	0.995	0.966	1.000		
⑨	0.828	0.951	0.964	0.949	0.955	0.888	0.956	0.884	1.000	
⑩	0.706	0.976	0.893	0.982	0.889	0.860	0.892	0.856	0.932	1.000

Notes: ①, ②, ..., ⑩ indicate summary statistics that differ by measure of centrality, weighting, sample truncation, and sample restriction.

† obs no.: the number of observations in each cell

‡ census wt.: the total sum of person weights in each cell

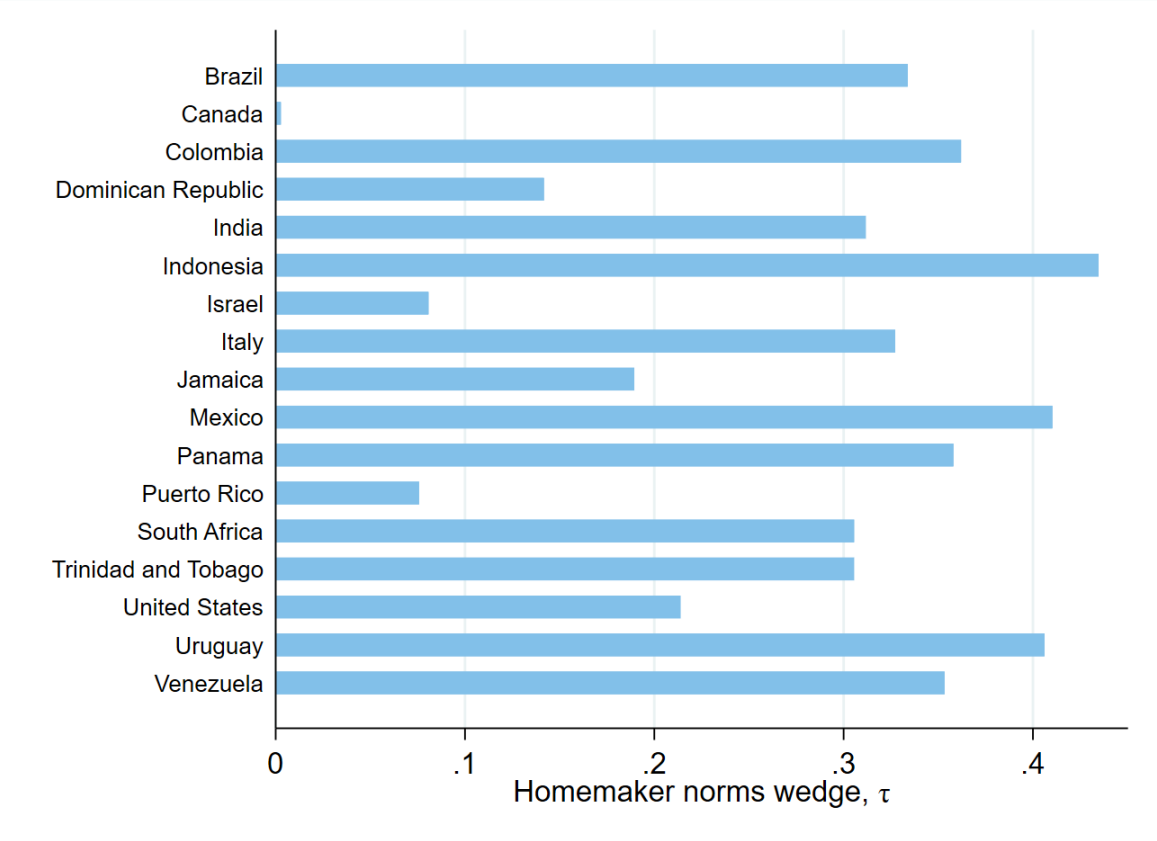
§ >100: only cells with strictly more than 100 observations

4.2 Norms wedges across countries

Figure 1 depicts the values of the norms wedges by country. There is substantial cross-country variation. As societal norms have changed over time, the fact that the IPUMS International

database covers censuses of countries from different points in time contributes to this variation. For instance, the only year in which the data for Canada is available is 2011, whereas the only year in which the data for Colombia is available is 1973.[9] However, differences in years used by country are far from the main drivers of the variation. I show later in Section 5.2 that, at least for the sample period of interest, the time series variation within country is eclipsed by the cross-country variation.

Figure 1: Homemaker norms wedge by country



Notes: This figure illustrates the values of the homemaker norms wedges by country. For each country, I take the average of the country-year summary statistic (10).

4.3 Magnitudes of norms wedges

Are the values of the norms wedges, as seen in Figure 1, large or small? It would be useful to have a benchmark to compare the norms wedges to.

To this end, I compute the "men vs. women" counterpart of the norms wedge: what is the "gender norms wedge" that rationalizes by how much the labor force participation of *women*

[9]Table 1 provides information on which years are covered for each country.

falls behind the level predicted if they behaved like similar *men*? The gender norms wedge can be juxtaposed to the "married women vs. single women" comparison that is the norms wedge.

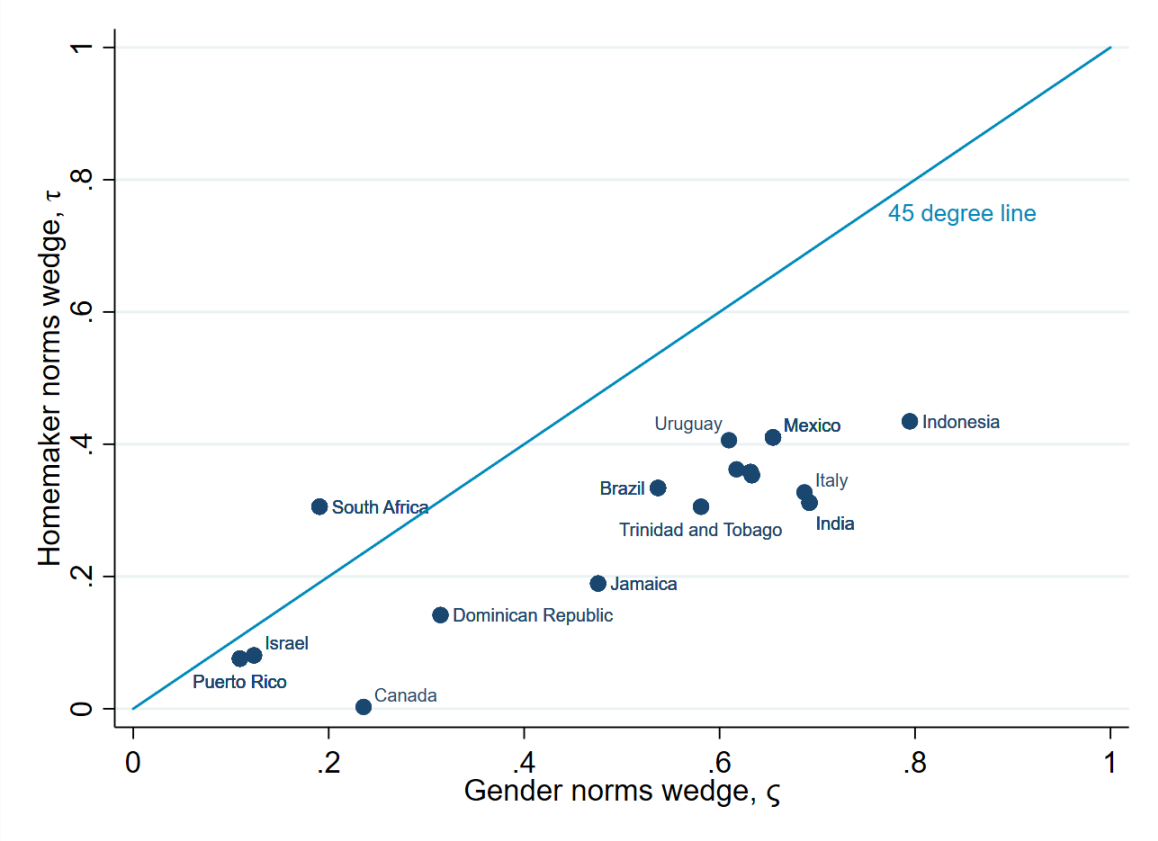
The gender norms wedge is defined as:

$$\varsigma = 1 - \frac{avrwage_{\text{ men}}}{avrwage_{\text{women}}} \left(\frac{1 - LFP_{\text{men}}}{1 - LFP_{\text{women}}} \right)^{\frac{1}{\theta}} \tag{4}$$

For the gender norms wedge, men and women are matched on their education and family composition categories, similarly to the norms wedge computation.

Figure 2 plots the values of the norms wedges against the gender norms wedges by country. For all countries but South Africa, ς is greater than τ. In fact, the magnitude of τ approximates around half of the magnitude of ς. Moreover, there is greater degree of cross-country variation in ς than in τ.

Figure 2: Homemaker norms wedge and gender norms wedge by country



Notes: This figure compares the plots the homemaker norms wedge and the gender norms wedge for each country. The gender norms wedge is the "men vs. women" counterpart of the "married women vs. single women" comparison that is the homemaker norms wedge.

A caveat to keep in mind is that the assumption that men and women share a similar level of

home productivity is less plausible than it is for married and single women. When children are small, in particular, biological gender differences contribute to the greater need for women to stay home than men. There is also a vast literature on gender differences in mental or physical tendencies, such as risk-taking (Charness and Gneezy, 2012), competitiveness (Niederle and Vesterlund, 2007), opportunities to rise to the top (Athey, Avery, and Zemsky, 2000), and how their performance is assessed (Mengel, Sauermann, and Zölitz, 2018). Therefore, where these gender differences contribute to why females participate less in the labor market relative to males than what the wage differentials suggest, they would contribute to the large magnitude of the gender norms wedge. It is less likely that the gender norms wedge serves as a measure of a social norm than the homemaker norms wedge.

Nonetheless, the gender norms wedges is still useful as a benchmark value. That there are vast gender differences in numerous dimensions, in fact, functions to highlight the size of the homemaker norms wedge, amounting to a half of the gender norms wedge.

5 Is culture pertinent for the norms wedge?

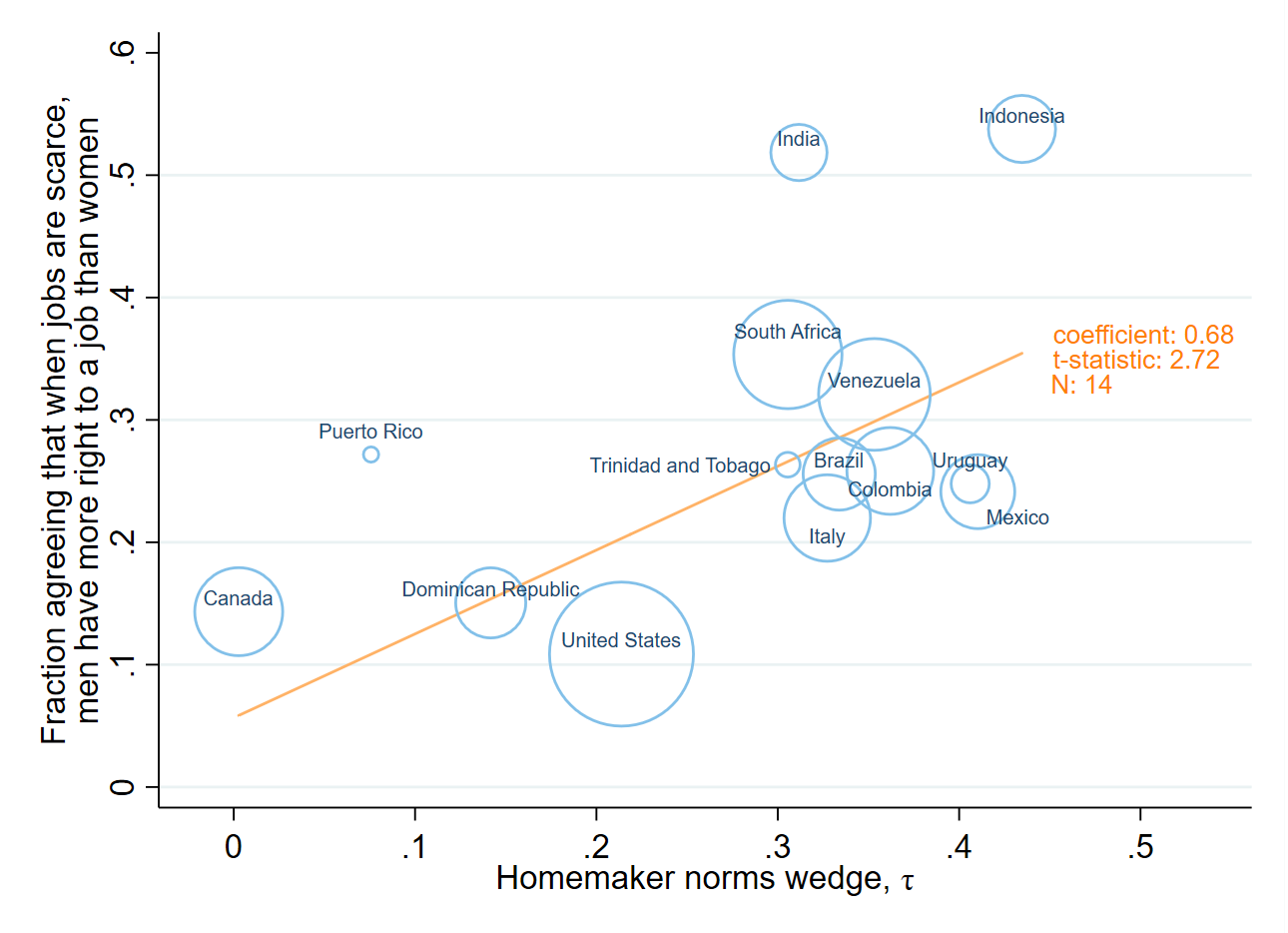
5.1 Correlation between norms wedges and attitudes

The norms wedges are summative measures that encapsulate all the reasons besides wages for which the gap in the labor force participation of married women and that of similar single women arises. Therefore, the cross-country variation in the norms wedges naturally reflect differences in economic structure and institutions. However, could differences in culture also drive the cross-cross variation? If the norms wedges are driven by economic or institutional differences alone, then the nomenclature "norms" wedge would be misleading.

Taking attitudinal survey data as proxies for cultural norms, I gauge whether the norms wedges are in line with attitudes regarding gendered behavioral prescriptions. To this end, I correlate the τ's and the answer to the WVS question, "Do you agree with the following statement: men have more right to a job than women." The τ and the attitudinal survey answers are aggregated at the country level.[10] The result is plotted in Figure 3. The correlation shows that in spite of all the cross-country differences the norms wedges fail to take into account, the countries with large τ's are indeed the ones that tend to have more conservative gender attitudes.

[10]The values of τ aggrefated at the country level are in Figure 1.

Figure 3: Cross-country correlation between norms wedges and attitudes



Notes: The attitudinal survey data originate from the World Values Survey, and the data used to compute the norms wedges are from the IPUMS International database. The reported t-statistic is based on robust standard errors. The size of the bubble for each country represents the size of the sample used in the computation of τ.

However, it is difficult to see the correlation in Figure 3 as a causal relationship. One of the most representative confounding variable would be GDP.[11] Richer countries might have more progressive attitudes on all fronts in general, which would result in a lower fraction agreeing that men have more right to a job then women when jobs are scarce. Richer countries might also have better transport infrastructure which allows a married woman – who might live near their husband's workplace, unlike a similar single woman who can live near their own workplace – to travel with ease to work, resulting in a lower norms wedge.

Therefore confirm that the relationship between attitudes and norms wedges is robust, I construct a country-year panel dataset. A panel allows me to address country-specific factors that influence the correlation, such as the level of economic development, sectoral composition

[11]Goldin (1995) documents a U-shaped cross-country relationship between female labor force participation and economic development.

of the economy, infrastructure, and legal or social institutions. Instead of matching country-level means of attitudes and τ, I match the two variables by the closest year.[12] The advantage of using this particular WVS question becomes apparent here. As explained in Section 2.2, it is the question that got asked the most consistently over time in the WVS that relates to gender. It was asked in all five waves after the first wave, covering the years 1990 to 2014. Because the number of countries for which the norms wedge can be computed is quite small, it is critical to include as many time periods as possible.

Moreover, the other advantage of this WVS question is that there is a directly comparable question on the topic of immigrants. Both are similarly phrased and answered, and are on the same issue of job scarcity. The two questions were asked side by side in the survey questionnaire, so a country-year panel can be constructed for the exact same country-year pairs.

Table 3 reports the results of regressing the attitudinal measures on the norms wedge, augmented by country fixed effects. Clearly, the sample is still very small in this panel dataset.[13] Nonetheless, even with the country fixed effects, τ and the WVS question on gender are correlated. What is reassuring, too, is that the coefficient value remains similar with the addition of country fixed effects. Without the country fixed effects, the coefficient value is 0.68, as depicted in Figure 3. The coefficient value with country fixed effects, reported in column (1) of Table 3, is 0.65. When the fixed effects for the World Value Survey waves are added on top, in column (2), the coefficient size decreases but remains statistically significant. Overall, the relationship between the homemaker norms wedges and gender attitudes appear to be robust.

On the other hand, τ and the WVS question on immigrants are found not to be, in columns (3) and (4). Hence, it is not the case that the correlation between norms wedges and attitudes on gender, depicted in Figure 3, represents a correlation between the norms wedges and conservative values in general.

[12] Appendix Table A2 lists how the years of the IPUMS International censuses and the years of the World Values Survey are matched.

[13] The number of countries with multiple years is also very small. Thus, the regression results in Table 3 include both the cases when the standard errors are clustered by country and when they are not. The t-statistics are similar.

Table 3: Norms wedges against attitudes on female or immigrant workers

	Fraction agreeing that when jobs are scarce, ...			
	men have more right to a job then **women**		employers should prioritize **own nationals** over **immigrants**	
	(1)	(2)	(3)	(4)
Homemaker norms wedge (τ)	0.648**	0.453*	0.639	0.268
	(2.39)	(1.99)	(1.30)	(0.44)
	[2.75]	[1.99]	[1.55]	[0.64]
Country fixed effects	✓	✓	✓	✓
Survey wave fixed effects		✓		✓
No. of countries	14	14	13	13
No. of countries with >1 waves	7	7	7	7
N	29	29	28	28

Notes: t statistics from standard errors clustered by country in parentheses, and t statistics from robust standard errors in square brackets; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
When the sample for columns (1) and (2) are fixed at the sample for columns (3) and (4), the regression results barely change. The reason is that the country fixed effect absorbs the one country that is present in the former sample and missing in the latter sample: Colombia.

5.2 Cross-sectional vs. time series variation in norms wedges

Although Table 3 demonstrates that there is certainly variation in the norms wedges and attitudes within a country over time, the more interesting variation remains the cross-country one – at least for the relatively short time range covered by the data. Table 4 decomposes the variation in the norms wedges and the gender attitudes into cross-country and time series components. In Panel A, the Shapely decomposition provides an additive decomposition of the R squared in the OLS regression of the norms wedge on country and year indicators. It quantifies the relative contribution to the R squared of the group of country dummies and the group of year dummies. Although there are far greater number of year dummies than country dummies, the share of contribution to the R squared lies much more heavily on the country dummies. The country dummies contribute to over three times the contribution of the year dummies. Similarly, in Panel B, the WVS question on gender also varies much more greatly across countries than across years. The Shapley decomposition shows that the country dummies contribute 87% of the variation in the attitudinal survey answer.

The between-within decomposition decomposes a variable (x_{it}, of country i and year t) into between-country ($\bar{x}_i \equiv \frac{1}{T}\sum_{t=1}^{T} x_{it}$) and within-country ($x_{it} - \bar{x}_i + \bar{\bar{x}}$, with the global mean

$\bar{\bar{x}} \equiv \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T} x_{it}$ added back in to make the results comparable) components. For both the homemaker norms wedge and the gender attitude measure, the between-country variance is over five times the within-country variance.

Therefore, although the panel structure was useful to demonstrate the robustness of the relationship between norms wedges and attitudes, I focus on the cross-country variation in the sections to follow.[14]

Table 4: Cross-sectional and time series decomposition of norms wedges and attitudes

***Panel A*: Homemaker norms wedge (τ)** [N=51]

	Shapley decomposition		**Between-within decomposition**			
	Number of categories	Share of contribution to R^2		Standard deviaion	Min.	Max.
Countries	17	76.05%	Between-country	0.130	0.003	0.434
Years	26	23.95%	Within-country	0.057	0.158	0.399

***Panel B*: Fraction agreeing that when jobs are scarce, men have more right to a job than women** [N=152]

	Shapley decomposition		**Between-within decomposition**			
	Number of categories	Share of contribution to R^2		Standard deviaion	Min.	Max.
Countries	59	86.78%	Between-country	0.203	0.040	0.877
Years	22	13.22%	Within-country	0.053	0.222	0.574

Notes: The Shapely decomposition provides an additive decomposition of the R squared of the OLS regression. It quantifies the relative contribution of each group of regressors to the R squared. The between-within decomposition decomposes the standard deviation of a variable into between and within components. The value of N refers to the number of observations for each variable.

[14] In addition, adding country fixed effects necessitates that 7 out of 14 countries "drop out" from the variation, since they are observed only in one time period.

6 How much of the cross-country variation in norms wedges can culture account for?

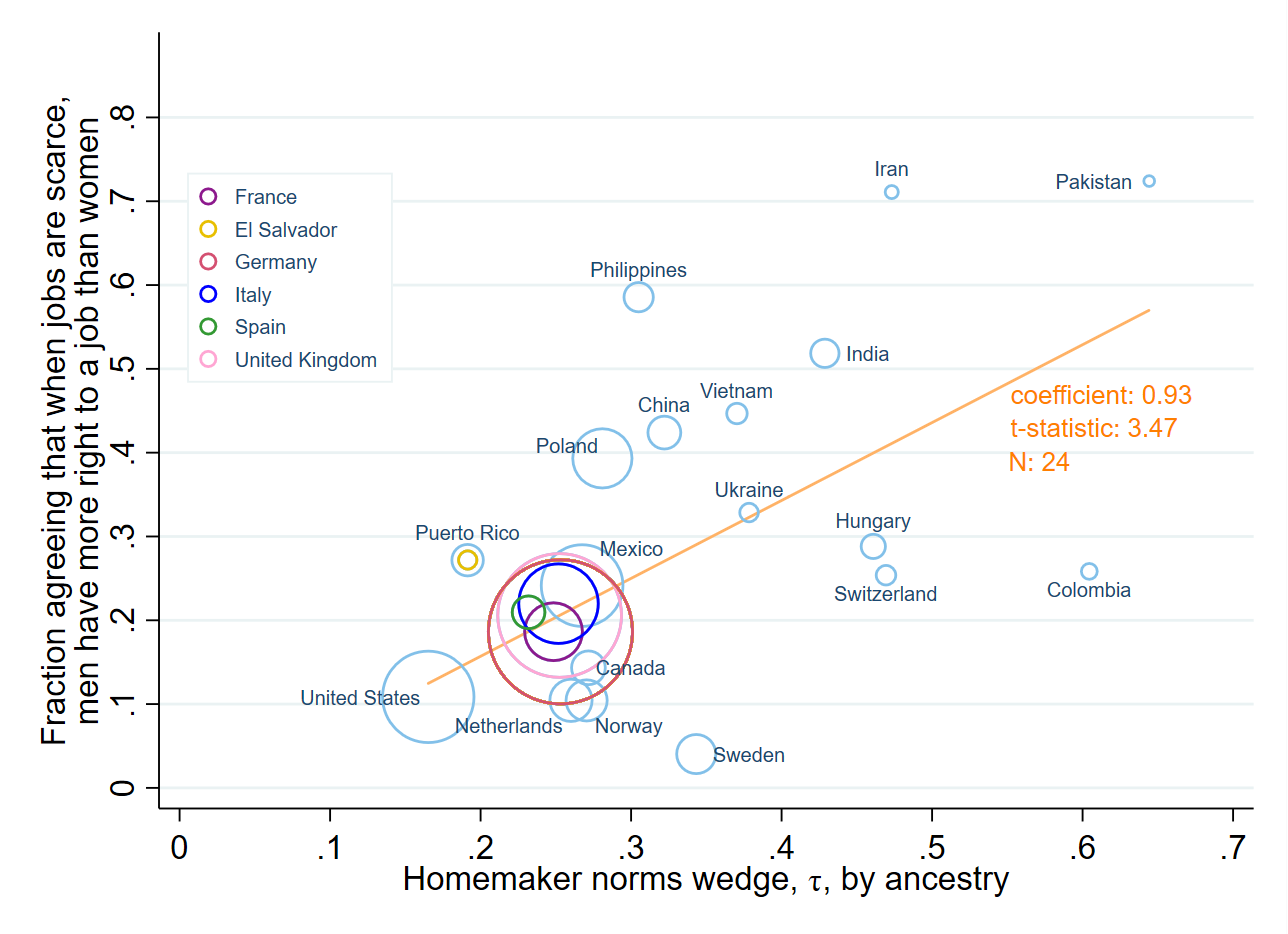
It is not sufficient to show that the homemaker norms wedges are correlated with attitudes, to establish that culture plays a role in explaining the cross-country variation in the wedges.

To get closer to causality, I adopt the idea behind the epidemiological approach in the literature assessing whether culture matters for economic decisions, reviewed by Fernández (2011). The epidemiological approach attempts to identify the effect of culture through the variation in economic outcomes of individuals who share the same economic and institutional environment, but whose social beliefs are potentially different. For immigrants in the United States operating in the same labor market, differences in their labor force participation is a reflection of different cultural heritage. Thus I am able to separate culture from differences in economic and institutional environment, by computing the τ separately for U.S. immigrants from different origin countries.

I compute the τ separately by self-reported ancestry in the U.S. censuses of the IPUMS International database. To illustrate, the τ corresponding to the Italian heritage represents the gap in the labor force participation rates of married Italian American women to similar single Italian American women, that is not accounted for by the wage differentials. Groups of married and single Italian American women are matched by education and family composition, as before. After identifying the group-specific τ for each census year, I take all those values of τ, and aggregate them at the country level by taking the unweighted median of the τ's computed with cells of sizes greater than 100. This statistic is the same as the one used to aggregate the norms wedges by country in Section 4.

Do these norms wedges by ancestry make sense? I check whether they are correlated with attitudes on gender. Figure 4 reproduces Figure 3 for the norms wedges by ancestry. I could compute and aggregate the norms wedge by ancestry for 24 origin countries, more than the 14 countries for which I could compute the country-level norms wedge. Similarly to norms wedges at the country level, the norms wedges by ancestry among the U.S. population are strongly correlated with gender attitudes.

Figure 4: Correlation between norms wedges by ancestry and gender attitudes



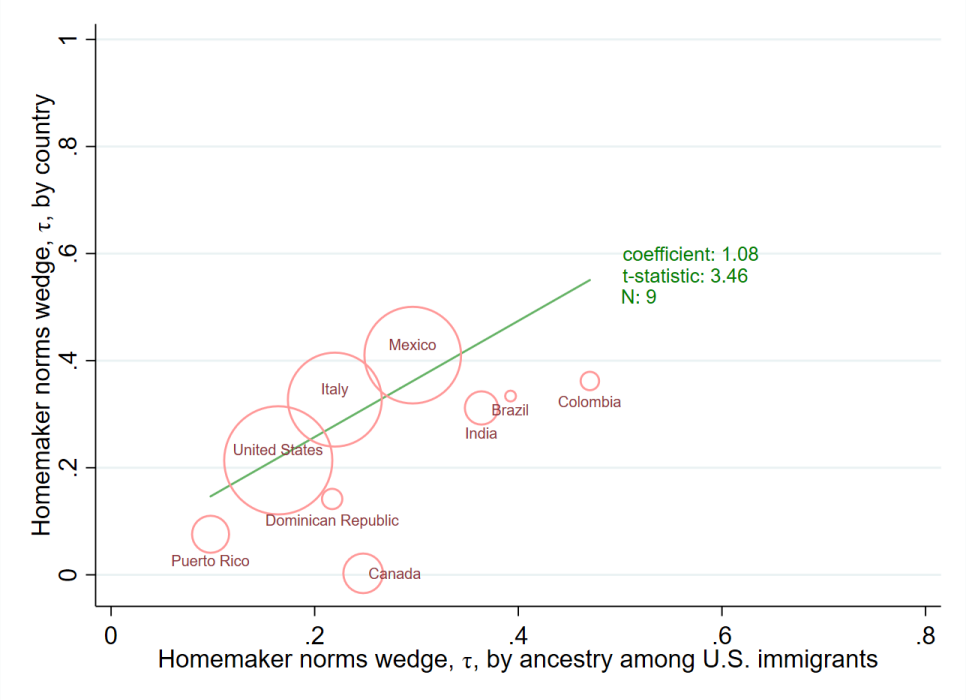
Notes: The attitudinal survey data originate from the World Values Survey, and the data used to compute the norms wedges are from the IPUMS International database. The reported t-statistic is based on robust standard errors. The size of the bubble for each country represents the size of the sample used in the computation of τ by ancestry. Some of the bubbles are colored differently for visual clarity.

Then, how much of the cross-country variation in norms wedges be explained by culture, as measured through the epidemiological approach? To this end, I plot the norms wedges by country against the norms wedges by ancestry among U.S. immigrants in Figure 5.[15] Plot (A) of Figure 5 depicts the scatter plot of the two types of norms wedges when the same statistic, unweighted median conditional on cell sizes exceeding 100, is used for both. The number of countries for which the two types of norms wedges are matched is rather small at 9. The reason is that the U.S. immigrant population in the 5% census sample cannot be large for all ancestral backgrounds. For example, where the size of the Indonesian American population is not large,

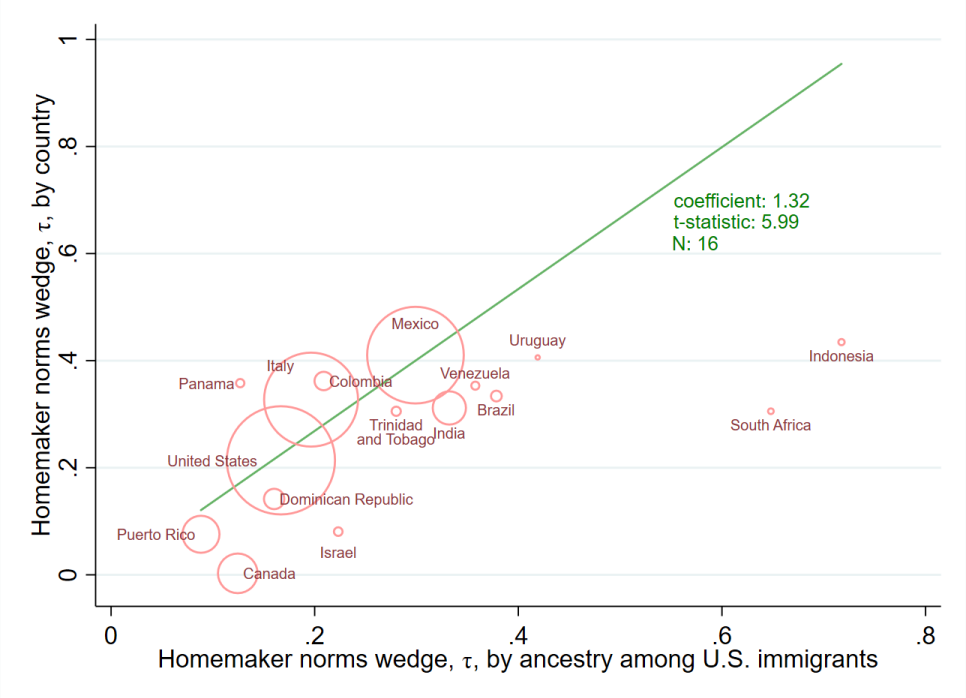
[15]Appendix Table A3 reports the regression results. For greater comparability, Appendix Table A3 also reports the results of the regression when the cell size restriction is lifted for the norms wedges by country, too. The results remain similar.

it becomes difficult to meet the restriction that τ is to be calculated only if the cell sizes are large enough. Hence, in Plot (B) I also repeat the scatter plot for when the norms wedges by ancestry are computed with the cell size restriction lifted.

Figure 5: Correlation between norms wedges by country and by ancestry



(A)



(B)

Notes: This figure plots the homemaker norms wedge τ by country against τ by ancestry among the U.S. population. Plot (A) depicts the scatter plot of the two types of norms wedges when the same statistic, unweighted median conditional on cell sizes exceeding 100, is used for both. For Plot (B), the norms wedges by ancestry are computed without the cell size restriction. The size of the bubble for each country represents the size of the sample used in the computation of τ by ancestry.

In both plots (A) and (B) of Figure 5, there is clearly a strong positive relationship between the two types of norms wedges. Therefore, taking the norms wedge by ancestry as a good measure of the isolated effect of culture, the norms wedge by country does indeed contain the effects of culture.

To see how much of the cross-country variation in norms wedges can culture account for, I consider two factors: the R squared and the coefficient value. Take a simple linear regression given by

$$y_i = \alpha + \beta x_i + \epsilon_i \tag{5}$$

The R squared value of this regression represents the degree of variation in y that the linear projection of x can account for. Visually, it is displayed by how close the observation points are situated to the line of best fit. The coefficient $\hat{\beta}$, on the other hand, is useful for gauging whether the relationship is positive or negative, and also for comparing the magnitude of the change in y that is associated with a one-unit change in x.

If, for example, the norms wedge by country was always exactly double the norms wedge by ancestry, then the R squared would equal 1, and the coefficient estimate $\hat{\beta}$ (as well as the true parameter value β) would equal 2. The conclusion in that scenario would be that the cross-country variation in the norms wedge is entirely explained by cultural differences, and that the effect of culture is attenuated by half among the U.S. immigrants, perhaps by cultural assimilation within the U.S, post-immigration.

By contrast, if the norms wedge by country had absolutely no correlation with the norms wedge by ancestry, then both the R squared and the coefficient estimate would be very close to 0. The conclusion then would be that factors that are completely orthogonal to culture drive all of the cross-country differences in the norms wedges.

The R squared value of plot (A) is 0.44, while for plot (B), the R squared value is larger at 0.69.[16] Hence, 44 to 70% of the variation in the cross-country differences in the norms wedge can be accounted for by differences in culture. Moreover, the coefficient estimates for both plots are greater 1, which implies that the effect of culture is stronger in the country-level norms wedges. The reason could be that the strength of culture is weakened post-immigration. It could also be that other country variables that are correlated with culture affect the norms wedges in the same direction as culture. To illustrate, countries with a stronger homemaker social norm for married women may tend have inferior transport infrastructure which further impedes married women from traveling to work.

[16]Visually, too, the scatter is tighter around the line of best fit in plot (B) than in plot (A).

7 Female labor force participation across countries

The previous sections showed that a) there is a large degree of variation in the country-level norms wedges, b) norms wedges are sizable, relative to gender norms wedges, and c) a hefty portion of the variation can be accounted for by cultural differences.

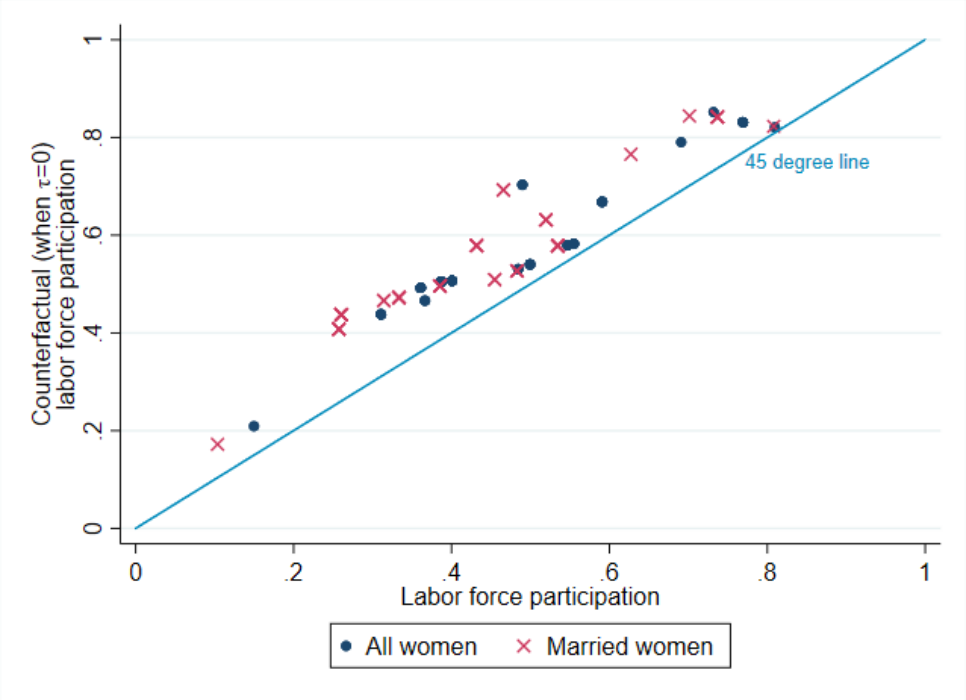
The cross-country variation in the norms wedges demonstrates itself in the large variation in female labor force participation across countries. Figure 6 plots on the horizontal axis the labor force participation rate of all (married and single) women and of married women by country, taking the average across all available IPUMS International census years. The large variation is evident; married women's labor force participation ranges from a low of 10.3% to a high of 80.8%.[17]

To what extent do the norms wedges contribute to the cross-country variation in female labor force participation? I compute how the pattern of female labor force participation change, had the norms wedges disappeared altogether. Those counterfactual values are plotted on the vertical axis of Figure 6. As the scatter points in Figure 6 are situated above the 45 degree line, the figure shows that the counterfactual values are larger than the actual values. On average, the labor force participation rate of married women rises by 38% in the counterfactual. Taking married and single women together, the rise is at 21%. The reason the rise is larger for married women is simple: the norms wedges directly impact the labor force participation decisions of married women only. Proportionally speaking, it is the countries that have low female labor force participation that experience greater increases when the norms wedges disappear.

Moreover, the cross-country variance of married female labor force participation drops by 7.8%. Thus, differences in the norms wedge are far from driving all of the differences in female labor force participation across countries, but they do contribute to a non-negligible degree.

[17] As these numbers are averages across census years, it does mix up the time series variation along with the cross-sectional variation. Specifically, the low of 10.3% corresponds to Colombia in 1973 and the high of 80.8% to Canada in 2011. However, even just taking recent years, there is much divergence across countries. In 2015, the labor force participation rate of Mexico's married women was 34.2%.

Figure 6: Female labor force participation with zero norms wedges



Notes: This figure plots the female labor force participation by country. I differentiate between the labor force participation of married women and that of all (married and single) women. The horizontal axis stands for the empirically observed labor force participation rate, whereas the vertical axis stands for the counterfactual rate computed when the homemaker norms wedges disappear for all countries.

8 Conclusion

A low level of female labor force participation implies that a large part of human talent rests unrealized. This paper demonstrates that the homemaker norms wedge, which glues married women to the home, exists. Not only that, they are substantial in size in many countries. The homemaker norms wedge that married women experience relative to similar single women amounts to a *half* of the gender norms wedge that women in general experience relative to similar men. Against a backdrop of a large and vigorously growing subfield of economics on various gaps by gender, it calls to attention how important marriage is, in creating gaps amongst individuals of the same gender. Married women do face additional constraints against participating in the labor market relative to single women, and their participation would rise by close to 40% if the homemaker norms wedge were to disappear.

The homemaker norms wedge diverging by ancestry among U.S. immigrants also underscores the fact that culture persists even when economic conditions change. Culture is slow to change,

and cultural shifts often lags behind economic progress.[18] Recall that the female labor force participation rate was 6% in Yemen and 84% in Rwanda in 2019. A slow-moving culture implies that it is difficult to expect Yemen to converge to the levels of Rwanda any time soon, even if the two countries were to be placed under the same economic conditions.

Nonetheless, it is not impossible to accelerate cultural change. Cultural change does occur over time within countries (Lee, 2020), and policies aimed at behavioral and attitudinal change have been found to be effectual (La Ferrara, 2016). Policies can indeed work to hasten the adoption of gender-equal attitudes. At the same time, culture is far from explaining all of the cross-country variation in female labor force participation. Even if the homemaker norms wedges are removed everywhere, 92% of the variation would remain. Hence, other factors are also important. Jayachandran (2015) review that policies aimed at improving the earning potential of women, such as school expansion, microfinance, and business skill training, have effectively contributed to raising female labor force participation in developing countries. Therefore, policies that tackle economics and culture should work in tandem in order to achieve higher female labor force participation across the board worldwide.

[18] Although traditional plough cultivation that requires male brute force have long been largely outmoded, descendants of societies that practiced it still have more conservative gender roles today (Boserup, 1970; Alesina, Giuliano, and Nunn, 2013).

References

Abrevaya, J. (2009). Are there missing girls in the united states? evidence from birth data. *American Economic Journal: Applied Economics 1*(2), 1–34.

Alesina, A., P. Giuliano, and N. Nunn (2013). On the Origins of Gender Roles: Women and the Plough. *The Quarterly Journal of Economics 128*(2), 469–530.

Almond, D. and L. Edlund (2008). Son-biased sex ratios in the 2000 united states census. *Proceedings of the National Academy of Sciences 105*(15), 5681–5682.

Athey, S., C. Avery, and P. Zemsky (2000). Mentoring and diversity. *American Economic Review 90*(4), 765–786.

Boserup, E. (1970). *Women's role in economic development*. London: George Allen and Universe Ltd.

Charness, G. and U. Gneezy (2012). Strong evidence for gender differences in risk taking. *Journal of Economic Behavior Organization 83*(1), 50 – 58. Gender Differences in Risk Aversion and Competition.

Coen-Pirani, D., A. León, and S. Lugauer (2010). The effect of household appliances on female labor force participation: Evidence from microdata. *Labour Economics 17*(3), 503–513.

Dinkelman, T. (2011). The effects of rural electrification on employment: New evidence from south africa. *American Economic Review 101*(7), 3078–3108.

Fernández, R. (2011). Does culture matter? In J. Benhabib, M. O. Jackson, and A. Bisin (Eds.), *Handbook of Social Economics*, Volume 1A, pp. 481–510. The Netherlands: North-Holland.

Fernández, R. and A. Fogli (2009). Culture: An empirical investigation of beliefs, work, and fertility. *American Economic Journal: Macroeconomics 1*(1), 146–77.

Galor, O. and D. Weil (1996). The gender gap, fertility, and growth. *American Economic Review 86*(3), 374–87.

Goldin, C. (1995). *The U-Shaped Female Labor Force Function in Economic Development and Economic History*, pp. 61–90. University of Chicago Press.

Greenwood, J., N. Guner, G. Kocharkov, and C. Santos (2016). Technology and the Changing Family: A Unified Model of Marriage, Divorce, Educational Attainment, and Married Female Labor-Force Participation, journal = American Economic Journal: Macroeconomics. *8*(1), 1–41.

Heath, R. and S. Jayachandran (2017). The causes and consequences of increased female education and labor force participation in developing countries. In S. Averett, L. Argys, and S. Hoffman (Eds.), *The Oxford Handbook of Women and the Economy*.

Jayachandran, S. (2015). The roots of gender inequality in developing countries. *Annual Review of Economics 7*(1), 63–88.

Jayachandran, S. and A. Lleras-Muney (2009). Life Expectancy and Human Capital Investments: Evidence from Maternal Mortality Declines*. *The Quarterly Journal of Economics 124*(1), 349–397.

La Ferrara, E. (2016). Mass Media and Social Change: Can We Use Television to Fight Poverty? *Journal of the European Economic Association 14*(4), 791–827.

Lee, J. E. (2020). Marriage and Misallocation: Evidence from 70 Years of U.S. History. Working paper.

Mengel, F., J. Sauermann, and U. Zölitz (2018). Gender Bias in Teaching Evaluations. *Journal of the European Economic Association 17*(2), 535–566.

Miller, G. (2010). Contraception as development? new evidence from family planning in colombia. *Economic Journal 120*(545), 709–736.

Niederle, M. and L. Vesterlund (2007). Do Women Shy Away From Competition? Do Men Compete Too Much?*. *The Quarterly Journal of Economics 122*(3), 1067–1101.

Pitt, M. M., M. R. Rosenzweig, and M. N. Hassan (2012). Human capital investment and the gender division of labor in a brawn-based economy. *American Economic Review 102*(7), 3531–60.

Appendix A Data

A.1 Harmonizing earnings in the IPUMS International database censuses

The earnings variable is not homogeneous across these countries. Where available, I use the variable *incwage*, which reports the respondent's weekly, monthly, or annual wage and salary income. The next alternative is the variable *incearn*, which reports the respondent's total income from their labor, including wages, business income, and farm income, made in the previous month or year. When neither are available, I use the variable *inctot*, which reports the respondent's total personal income from all sources in the previous month or year. The reason why I utilize the earnings variables in this order is that in the model, the decision to work in the labor market depends on the comparison of potential market earnings and home productivity that one could generate. Therefore, I do not want business profits (particularly negative profits) or other unearned sources of income such as remittances to confound the labor force participation decision. Table A1 lists the census or survey countries and years, and also depicts which earnings variable is used in each sample.

A caveat for all earnings variables for all the countries is that the recorded earnings are before-tax. Of course, countries have different national systems for taxes and benefits, which affects the take-home earnings that really matter for the labor force participation decision. Since to compute the homemaker norms wedge, I compare the wages of married and single women within a country, the fact that some nations have much higher tax rates in general than others would not affect the norms wedge. However, national tax systems would also vary in how differently they treat married individuals from single, which would be included in the norms wedge value.

Appendix B Tables

Table A1: List of census or survey, with earnings variable used

Country	Census/survey year(s)	Earnings variable
Brazil	1980, 1991, 2000, 2010	incearn
Canada	2011	incearn
Colombia	1973	inctot
Dominican Republic	1981	incwage
	2002	inctot
India†	1983, 1987, 1993, 1999, 2004	incwage
Indonesia	1976, 1995	incwage
Israel	1972	incearn
	1995	incwage
Italy†	2011‡	incwage
Jamaica	1982, 1991, 2001	incwage
Mexico	1970	inctot
	1990, 1995, 2000, 2010, 2015	incearn
Panama	1970	incwage
	1980	inctot
	1990, 2000, 2010	incwage
Puerto Rico	1990, 2000, 2005, 2010	incwage
South Africa	1996, 2001, 2007, 2011	inctot
Trinidad and Tobago	1970	inctot
	2000	incwage
United States	1980, 1990, 2000, 2010	incwage
Uruguay	2006	incwage
Venezuela	1971, 1981, 1990, 2001	incearn

Notes: *incwage* reports the respondent's weekly, monthly, or annual wage and salary income; *incearn* reports the respondent's total income from their labor, including wages, business income, and farm income, made in the previous month or year; *inctot* reports the respondent's total personal income from all sources in the previous month or year.

† Surveys

‡ I combined the survey data from years 2011, 2012, ..., 2018, due to the small sample size in each year.

Table A2: List of census or survey, with earnings variable used

	IPUMS	WVS		IPUMS	WVS
Brazil	1991	1991	Puerto Rico	1990	1995
	2000	2006		2000	2001
	2010	2014	South Africa	1996	1996
Canada	2011	2006		2001	2001
Colombia	1973	1997		2007	2006
Dominican Republic	2002	1996		2011	2013
India	1987	1990	Trinidad and Tobago	2000	2006
	1993	1995	United States	1990	1995
	1999	2001		2000	1999
	2004	2006		2010	2011
Indonesia	1995	2001	Uruguay	2006	2006
Italy	2011	2005	Venezuela	1990	1996
Mexico	1990	1990		2001	2000
	2000	2000			
	2010	2005			
	2015	2012			

Notes: *incwage* reports the respondent's weekly, monthly, or annual wage and salary income; *incearn* reports the respondent's total income from their labor, including wages, business income, and farm income, made in the previous month or year; *inctot* reports the respondent's total personal income from all sources in the previous month or year.

† Surveys

‡ I combined the survey data from years 2011, 2012, ..., 2018, due to the small sample size in each year.

Table A3: List of census or survey, with earnings variable used

	Norms wedge by country		
	(1)	(2)	(3)
Norms wedge by ancestry	1.084**	1.325***	1.371***
	(3.46)	(5.99)	(6.02)
Cell size restriction, norms wedge by country	>100	>100	-
Cell size restriction, norms wedge by ancestry	>100	-	-
N	9	16	16
R^2	0.444	0.686	0.704

Notes: t statistics from robust standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Columns (1) and (2) correspond to the regression results plotted in plots (A) and (B), respectively, of Figure 5.