Culture, Institutions, and the Child Penalty: Empirical Evidence from Developing Countries*

Ayush Shahi
(Advisor: Dr. Anisha Sharma)

Ashoka University

May 14, 2025

Abstract

This study examines the impact of gender norms and work-family policies on the motherhood penalty across a large sample of developing countries. Using data from the Demographic and Health Surveys, I find that regressive cultural norms about the appropriate role of women— shaped by ancestral adoption of the plough— significantly reduce women's labour force participation (consistent with Alesina et al. 2013). I show that these effects, robust to an instrumental variable analysis, are pronounced for women of childbearing age, mothers, and estimates of the motherhood penalty. I also run a staggered Difference-in-Difference to assess whether implementation of gender-equal laws can attenuate this penalty. My results provide mixed evidence to this end: while broader institutional shifts towards equality have limited impact, targeted work-family policies like provision of paid maternity leave can have significant positive effects on women's employment in the long run. Overall, my work has important implications for policymakers solving for the gender gap in employment in developing countries, and the persisting influence of cultural factors therein.

Keywords: gender norms, work-family policies, motherhood penalty, plough, DHS

*Thesis submitted in partial fulfilment of requirements for the ASP degree in the Department of Economics at Ashoka University. A note of thanks to Prof. Anisha Sharma for her patient advice through the year and to professors from the department for their comments during the proposal/final presentations. I would also like to acknowledge Prof. Kundan Sen and my peers from the Thesis Seminar for their useful inputs on communicating my findings. Lastly, I am ever grateful to my friends and family for their constant support and words of encouragement.
Any errors are mine alone. A replication package (with code and datasets) can be shared upon request.

1 Introduction

Childbirth shapes women's lives in more ways than one, with changes to their labour market trajectories being an economically important outcome. Consequently, the motherhood penalty has been defined as the empirically observed adverse impact of childbirth on labour market outcomes of mothers. Solving for such inequalities in the labour market requires evidence on how the penalty varies across cultural and institutional contexts. However, existing research is highly skewed towards a small sample of developed countries which show little variation along these margins. This study builds on a recent literature that constructs an atlas of child penalties and captures considerable heterogeneity in the same across continents, countries, and the rural-urban divide (Kleven et al. 2024). My focus is on developing countries in particular, and explaining differences seen therein through the lens of culture and institutions.

Starting from the seminal work of Goldin (1994), a growing strand of literature has emphasized the salience of norms in determining female labour force participation. At an individual level, Akerlof and Kranton (2000) use a gendered identity model to describe how socially prescribed behaviour dictates what women and mothers should and should not do (for example, engage in household work but not in the labour market). To the extent that these norms are produced historically (Alesina et al. 2013), they become part of a cultural belief that is sticky over time but varies widely across national and sub-national levels. Women may freely participate in the labour force in some places; not at all in others; and in others yet, may work until they become mothers— post which they are expected to prioritize their role as caregivers.

While policymaking itself is endogenous to the given cultural beliefs in a country, global legal convergence spurred by international conventions has led to laws becoming gender-equal before an equivalent shift towards egalitarian attitudes. For example, Egypt has had laws prohibiting gender discrimination in employment since 2003, despite only 17.6% of its labour force in 2024 being women (*The World Bank* n.d.) and 89.4% of its respondents to the World Values Survey in 2018 agreeing with the statement "Jobs scarce: men should have more right to a job than women". When it comes to improving maternal labour supply, policies can play an important role through legal provisions that support mothers with the combination of work and caregiving.

In this study, I link the two aforementioned themes for a large sample of developing countries. Given differences in policy implementation across years and countries, I use a staggered Difference-in-Difference approach to estimate whether institutional shifts towards gender equality and implementation of work-family policies can attenuate the penalty.

Moreover, to show evidence on the persisting impact of cultural norms, I provide OLS and instrumental variable estimations that link traditional plough use with lower labour force participation of women, specifically mothers. The empirical strategy for this is borrowed from Alesina et al. (2013), who test Ester Boserup's (1970) hypothesis that pre-industrial societies which adopted the plough are more likely to have regressive gender norms today. Previous literature has extensively looked at these two strands of norms and policies separately. Bringing them together can help answer important questions about the determinants of the motherhood penalty in developing countries, and how policymakers can solve for it within a given cultural environment.

My OLS and IV results validate the findings of Alesina et al. (2013). OLS estimates show that countries with greater traditional use of the plough, equivalent to 1 standard deviation, have worse gender norms and reduced probability of women's labour force participation by 7.5 percentage points. I show that these disparities are starkest for mothers and women of childbearing age, with employment probability of mothers falling by 8.82 percentage points. My evidence on the attenuating impact of work-family policies shows mixed results: broader institutional shifts towards gender equality have no short-run effects but increase long-run employment probability of women (aged 18-40) by 7.5-8.75 percentage points. On the other hand, targeted work-family policies like provision of paid maternity leave show much more promising long-run effects: 8-10 years post policy implementation, women of childbearing age are around 20 percentage points more likely to be in the labour force.

The rest of this paper is organized as follows: Section 2 provides background and existing literature on the role of norms and work-families in shaping the penalty. Section 3 describes the data sources used and Section 4 describes in detail my empirical strategy to capture the impact of norms and policies respectively. Section 5 presents the results and Section 6 concludes with a discussion of these results and their policy implications.

2 Background

2.1 Gender norms and maternal labour supply

A vast economic literature has conceptualized culture (for a review, see Fernández 2011; Alesina and Giuliano 2015; Guiso et al. 2006) and its transmission over time (Bisin and Verdier 2001, 2011, 2023; Giavazzi et al. 2019). Borrowing from Boelmann et al. (2025), I define culture as "systematic differences in both values (or preferences) and beliefs that vary across social or geographic groups". Gender norms, within this framework, can then be understood as cultural beliefs— dictating the appropriate role of men and women in society.

Data from the World Values Survey highlights how such gender norms manifest for working mothers. In its latest survey round, 42.4% of respondents agreed with the statement "When a mother works for pay, the children suffer". However, there is considerable heterogeneity across countries, with the proportion in agreement ranging from 87.8% in Bangladesh to 8.8% in Denmark. These cultural differences in beliefs about how central child-rearing is to a woman's identity translate directly into differences in employment outcomes of women across countries (Fortin 2005), and their impact is likely to intensify post-motherhood.

Existing cross-country evidence on the relationship between gender norms and the motherhood penalty is limited to only Kleven et al. (2019), who document a strong positive correlation between restrictive norms and the long-run penalty in earnings for a sample of 6 developed countries. The lack of cross-country causal evidence reflects the challenges in controlling for policy and institutions across countries. Consequently, most of the literature focuses instead on variation in gender norms using within-country natural experiments (Boelmann et al. 2025; Steinhauer 2018), individual-level data (Zhang et al. 2024; Cavapozzi et al. 2021; Kleven 2023) or experimental evidence (Bedi et al. 2018). This study expands the literature to a wide sample of developing countries (N=82) and addresses causality through an instrumental variable approach exploiting differences in traditional plough use as an exogenous determinant of variation in present-day gender norms (Alesina et al. 2013).

2.2 The role of work-family policies

Policies supporting women's employment post-motherhood usually work through two channels: by providing job security and continuity (for example, through paid maternity leave) or by making it easier for mothers to return to work by supporting the combination of caregiving and work (for example, through public provision of childcare). As with norms, the literature here is divided into studies providing cross-country evidence, exploiting within-country variation, or engaging in micro-level analysis (for a review, see Olivetti and Petrongolo 2017).

Evidence on the impact of these policies is mixed. Cross-country studies point towards a persistent positive effect of childcare provision and maternity leaves on reducing the motherhood penalty (Keck and Saraceno 2013; Boeckmann et al. 2015; Grönlund et al. 2016; Pettit and Hook 2005, 2009; Cukrowska-Torzewska 2017). Within-country studies from Austria (Lalive and Zweimüller 2009; Kleven et al. 2024; Lalive et al. 2014), Germany (Schönberg and Ludsteck 2014), Norway (Dahl et al. 2016), and USA (Baum 2003) find no such impact, with Baker et al. (2008) providing a notable exception with their study from Canada. There are two important caveats to the existing literature. First, they are restricted to a small sample of high-income developed countries. Second, cross-country results

are likely to be biased by observable and unobservable differences across countries that cannot be controlled. My work pushes the literature to examining the impact of these policies in developing countries, using large-scale standardized survey data to employ a staggered Difference-in-Difference approach that controls for differences across individuals, countries, and time.

2.3 Contribution to the literature

Missing in the existing economic literature is how work-family policies interact with different norm environments. To the best of my knowledge, only Budig et al. (2012, 2015) offer evidence, using microdata from 22 industrialized countries, that parental leave and childcare provision are most helpful in reducing the penalty where cultural norms also support maternal employment. It is important to note that given the nature of their sample, there is not significant variation in the explanatory variables of culture and policy— high and middle-income countries are likely to have generous policies as well as more egalitarian norms. Expanding this analysis to a larger sample of countries, specifically from the Global South where the burden of gender norms is expected to be higher, can provide us rich insights into how policymakers solve for gender inequality, and the cultural limitations they may face in this endeavour.

This study, while not directly investigating the interaction effect of policies and norms, provides important evidence on the cost of restrictive gender norms and the impact of work-family policies on the motherhood penalty. I use a combination of micro-data and country-level indicators to causally estimate these impacts and interpret them from a comparative perspective. In doing so, I also corroborate the results of Alesina et al. (2013), using alternative data, on the impact of ancestral plough use on norms and female labour force participation today. The broader motivation of my work is to open up the literature to questions along the lines of Budig et al. (2012). For example, in countries with regressive gender norms, do policies have limited impact on the penalty? Conversely, where cultural attitudes are gender egalitarian, do they have a differentially positive effect? Importantly, can we *outpolicy* norms?

3 Data and Descriptive Statistics

My primary dataset for measuring women's labour force participation is the Demographic and Health Surveys (DHS). The DHS programme, conducted by the *US Agency for International Development* (USAID), involves nationally representative household-level surveys

undertaken at intervals of ∼5 years and providing valuable data about health, nutrition, and other characteristics of household members. I use information from the woman's questionnaire on her employment (usually at the weekly or biweekly level), motherhood status, and other individual characteristics. Appending all standard DHS survey rounds conducted in 82 countries from the period of 1985-2023, gives me a total of 5,485,900 observations. Figure 3 from Appendix B depicts the countries included in my final dataset: mostly low and middle-income countries, from Africa and Asia. Panel A of Table 1 provides summary statistics of the relevant variables used from the DHS.

For my policy analysis, I supplement the DHS with the Women, Business, and Law (WBL) dataset provided by *The World Bank*. The WBL is an atlas of laws and policies enabling women's economic opportunity at the country-year level from the period of 1970 to 2023. It also provides scores for gender equality on a legal basis, based on the existence of policies, along 8 different dimensions: mobility, workplace, pay, marriage, parenthood, entrepreneurship, assets, and pension. Since my focus is on measuring the impact of a broader policy environment shift, I consider their flagship WBL index which aggregates the 8 subcomponent scores. I also estimate the impact of some individual policies that are relevant to the literature such as the provision of at least 14 weeks of maternity leave and the prohibition of dismissing pregnant employees. I am unable to look at the impact of childcare provision policies, which the literature has focused on previously, since WBL does not have time-series data on the same. Panel B of Table 1 provides summary statistics of the relevant indicators used.

To measure the norm environment in a country, I use the World Values Survey (WVS), which provides nationally representative survey data measuring human beliefs and values from 1981-2022. These surveys are conducted usually at a gap of five years, but only waves 6 and 7 (conducted post 2010) are included here given data availability on my measures of interest. I use survey questions measuring people's attitudes towards traditional gender roles and construct an index from the same to proxy gender norms. For details on questions used and index construction, see Appendix A. Panel C of Table 1 provides summary statistics of the individual questions and constructed index. An important caveat is that while the coverage of the WVS is of 78 countries, only 35 match with my DHS country sample.

Assessing the causal impact of norms on the penalty requires exogenous variation in cultural beliefs. To this end, I borrow data on ancestral plough use constructed by Alesina et al. (2013), which is used for descriptive evidence and the instrumental variable estimation. The dataset provides national-level estimates of the fraction of the population currently living in a country with ancestors that traditionally engaged in plough agriculture. Also included are historical characteristics, like presence of large domesticated animals and proxies of

Table 1: Summary Statistics of Indicators Used, Disaggregated by Data Source

	Mean	SD	Min	Max	N
Panel A: DHS					
Female labor force participation	0.46	0.50	0.00	1.00	4156553
Female age at first marriage	29.78	9.65	15.00	49.00	5465900
Female education	6.20	5.01	0.00	20.00	5457328
Mother indicator	0.69	0.46	0.00	1.00	5485900
Panel B: WBL					
WBL index	59.96	18.75	17.50	100.00	10260
14 weeks paid leave indicator	0.41	0.49	0.00	1.00	10260
Pregnancy dismissal prohibit indicator	0.54	0.50	0.00	1.00	10260
Panel C: WVS					
Job scarce: men right over women	0.47	0.50	0.00	1.00	150597
Housewife fulfilling as working for pay	0.64	0.48	0.00	1.00	175260
Men better political leaders	0.45	0.50	0.00	1.00	176295
University more important or boys	0.24	0.43	0.00	1.00	178929
Children suffer when working mother	0.46	0.50	0.00	1.00	178076
Men better business executives	0.40	0.49	0.00	1.00	177048
Gender norms index	-0.01	0.64	-1.16	0.91	138917
Panel D: Alesina et al. (2013)					
Plough use	0.48	0.48	0.00	1.00	228
Agricultural suitability	0.54	0.33	0.00	0.98	215
Tropical climate	0.74	0.42	0.00	1.00	212
Large domesticated animals	0.93	0.21	0.00	1.00	228
Political hierarchies	3.30	1.04	1.00	5.00	228
Economic complexity	6.38	1.38	1.00	8.00	228
Panel E: Kleven et al. (2024)					
Child penalty estimate	22.70	16.67	-16.00	64.00	134
Gender gap in employment	0.43	0.22	0.05	0.91	134

economic development and political complexity, that are important to control for since they may correlate with both plough use and present-day gender norms. Summary statistics of these indicators is provided in Panel D of Table 1.

Lastly, I use estimates from the Child Penalty Atlas as an added robustness indicator for measuring the impact of norms on the motherhood penalty. Kleven et al. (2024) provide an atlas of child penalties in employment based on micro data from 134 countries. To solve for data limitations, especially in developing countries, they use a novel estimation strategy to calculate the penalties involving pseudo-event studies of first childbirth using cross-sectional data. The pseudo-event studies are also validated against true event studies using panel data for a subset of countries. Their estimates allow me to present additional evidence regarding the motherhood penalty in particular, apart from my broader evidence on the impacts on maternal employment. Moreover, there is a strong match with the country sample from DHS (68 out of 82), corroborating my evidence on a largely similar sample of developing countries. Panel E of Table 1 shows summary statistics on the two country-level indicators that I have used: the penalty estimate and a simple measure of the gender gap in employment.

4 Empirical Strategy

To estimate the causal impact of policies and norms separately, I run a staggered difference-in-difference (section 4.1) and an instrumental variable (section 4.3) specification respectively. Before the IV estimation, I also provide descriptive evidence using OLS (section 4.2) on how plough use affects norms and FLFP today, at both the individual (section 4.2.1) and country-level (section 4.2.2). The empirical strategy for each is described in the following subsections.

4.1 Work-Family Policies: Staggered Difference-in-Difference

Since policy adoption varies across countries and time, I run a Two-Way Fixed Effects model which accounts for the staggered nature of treatment. As discussed earlier, the motivation of this study is to measure the impact of a broader policy environment shift— proxied here by the WBL index crossing the threshold of 50 on a scale of 100. This threshold is chosen based on the simple intuition that the index crossing 50 implies that of all the available policies enabling women's economic opportunity (at least those documented by the WBL), a country has brought into legislation a majority of them than not. From the DHS sample of 82 countries, 25 crossed the threshold at some point between 1991 and 2023. Apart from changes to the aggregate score, I also consider introduction of some policies highlighted by the literature. These include provision of at least 14 weeks of maternity leave (implemented

by 14 countries during the reference period) and prohibition of dismissing pregnant workers (implemented by 19 countries during the reference period).

What sample of women is appropriate for this analysis? Instead of looking at the impact of policy shifts on employment of all women in the DHS or only that of mothers, I look at women aged 18-40— considered as the prime age for motherhood, especially in developing country contexts. This choice is informed by the fact that women's employment decisions are not isolated from their expectations of post-motherhood employment. Women of childbearing age may anticipate the penalty and choose to drop out of the labour force pre-emptively (before motherhood), especially if the policy environment for the same is not conducive. However, improvement of work-family policies might make them change this decision leading to higher labour force participation also among non-mothers who appear in the cross-sectional DHS data. These effects would not be captured by restricting my sample to mothers only.

To create an event study of women's employment relative to time of 'treatment' (for each of the policy changes mentioned above), I run the following regression:

$$flfp_{ict} = \beta_0 + \sum_{k \neq 0} \delta_k . 1\{time\_to\_policy(= k)\}_{ict} + X'_{ict} + \alpha_c + \lambda_t + \varepsilon_{ict} \quad (1)$$

where $flfp_{ict}$ is the labour force participation dummy indicator (taking value 0 or 1) outcome for woman i in country c in year t. This is regressed on dummies $1\{time\_to\_policy(= k)\}_{ict}$, which take value 1 if the policy is in force for individual i in country c and year t. I include dummies for k ranging from -5 to 10, giving me an event study from 5 years before treatment to 10 years after[1]. The dummy for k=0 is omitted as observations from that period serve as the base group. X'_i are individual level controls: respondent's current age, place of residence (urban or rural), years of education, wealth index, number of living children, and marital status. I also include country fixed effects α_c, to control for differences across countries that are constant across time, and year fixed effects λ_t to wash away differences across years that are constant for all countries. Standard errors are clustered at the country level.

An important caveat regarding this analysis: since the DHS is not conducted at a yearly frequency and in the same years for all countries, I do not have a balanced panel across my event study years. For example, to date, the DHS has been conducted in India 5 times: round 1 in 1992-93, round 2 in 1998-99, round 3 in 2005-06, round 4 in 2015-16, and round 5 in 2019-21. For the event study of the policy guaranteeing at least 14 weeks of paid leave, which was passed in India in 2017, observations from round 1, 2, and 3 of the

[1]Observations preceding event study year -5 and following event study year 10 are bunched together at the respective boundaries.

Indian DHS[2] would be bunched together at k= -5. Round 4 observations would occupy k= -2 or k= -1 on the event study, while round 5 would sit at k=2, k=3, or k=4 depending on when an individual's interview was conducted during the round. Other country-year survey observations will similarly sit at different k periods relative to the policy shock. This also means that in a few rare cases, there are some event study years for which I have no observations and they are thus omitted from the plot. Moreover, if for any particular event study year I only have one country's observations, then its dummy becomes collinear with the fixed effects and is omitted. These are some considerations to be kept in mind while interpreting the results.

4.2 Gender Norms: OLS estimates

To explain the widespread differences in gender norms, and its impact on maternal employment, I rely on the work of Alesina et al. (2013). They test Ester Boserup's (1970) hypothesis that variation in cultural attitudes on the appropriate role of women in the economy has its origin in the form of agriculture practiced in the pre-industrial period. Boserup (1970) argues that since plough cultivation required greater strength, men had a comparative advantage in farming in societies where the plough was used. This led to specialization in activities along gender lines— men worked outside in the fields while women worked inside the home. Over time, this division of labour solidified into gendered norms— persisting even after a society transitioned away from agriculture— which dictated that the appropriate place for women is within the home and limited their economic empowerment.

I validate the results found by Alesina et al. (2013) using data from the DHS and an expanded set of norm indicators from the WVS. My OLS estimates provide descriptive evidence on the links between plough use, present-day norms, and maternal employment. I run two specifications for this purpose— at the individual-level and at the country-level— which are discussed below.

4.2.1 Individual OLS estimates

To estimate the impact of ancestral plough use on current female employment, I run the following regression separately for the sample of all women, women aged 18-40, and mothers in the DHS:

$$flfp_{ict} = \beta_0 + \beta_1 Plough_c + X_c^H + X_i^P + X_c^P + \lambda_t + \varepsilon_{ict} \tag{2}$$

where $flfp_{ict}$ is the labour force participation dummy indicator for woman i in country c in year t, regressed on $Plough_c$ which is the proportion of a country's population whose

[2]Referred to as the *National Family Health Survey* (NFHS) in India.

ancestors used the plough. Also included are historical controls X_c^H as discussed in Section 3, contemporary control X_c^P of a country's GDP per capita, and individual-level characteristics that may determine a women's labour force participation. Year fixed effects λ_t is included to control for time trends since the sample contains multiple cross-sections of the same country. Standard errors are clustered at the country level.

The impact on present-day norms is assessed using the WVS dataset. I run the following regression only for individuals in countries from the WVS that are also in the DHS in order to be consistent with making my arguments relevant to developing countries:

$$norm_{idc} = \beta_0 + \beta_1 Plough_d + X_d^H + X_i^P + X_c^P + \alpha_c + \varepsilon_{idc} \tag{3}$$

where $norm_{idc}$ is the constructed gender norms index and individual norm indicators for individual i in district d and country c. The major difference between equation (3) and (2) apart from the dependent variable is that plough use proportions and historical controls are available here at the subnational level, as given by $plough_d$ and X_d^H respectively. Also included are individual level controls, as well as country fixed effects since the plough indicator does not get washed away with its inclusion now. Standard errors are clustered at the district level.

4.2.2 Country OLS estimates

While the individual-level OLS estimates provide rich evidence on norm and employment differences, it can be argued that norms do not amount to much at an individual level. They are likely to vary across many observable characteristics like education, income, and age. However, by definition, norms should be considered as the cultural beliefs prevalent at a regional level— here, taken as the country. The norm and labour force participation indicators are thus aggregated at the country-year level, and the following regressions are run:

$$flfp_{ct} = \beta_0 + \beta_1 Plough_c + X_c^H + X_c^P + \lambda_t + \varepsilon_{ct} \tag{4}$$

$$norm_{ct} = \beta_0 + \beta_1 Plough_c + X_c^H + X_c^P + \lambda_t + \varepsilon_{ct} \tag{5}$$

where $flfp_{ct}$ and $norm_{ct}$ are the estimates from the DHS and WVS respectively in country c in year t. These are regressed on the same country-level indicators from equation (2), which includes the plough use measure along with historical and contemporary controls. Time fixed effects are included to wash away differences in the dependent variable across survey years that are constant for all countries. Standard errors are clustered at the country level.

To provide more pointed evidence regarding the impact of norms on specifically the

motherhood penalty, I run the following regression using the penalty estimates from the Child Penalty Atlas (Kleven et al. 2023):

$$penalty_c = \beta_0 + \beta_1 Plough_c + X_c^H + X_c^P + \varepsilon_c \tag{6}$$

where $penalty_c$ is the motherhood penalty, reported as the percentage point decrease in mothers' employment in the long-run, in country c. Interpretation of independent variables is same as previous country-level estimations and standard errors are clustered at the country level. I also run the same regression for the gender gap in employment measure, reported as the percentage point difference in labour force participation rates of men and women. Together, they provide a robustness check to my results on women's employment from the DHS.

4.3 Gender Norms: Instrumental Variable

OLS estimates are likely to be biased if adoption of the plough is correlated with other country-specific factors that are also influencing present-day norms and female employment. For example, if societies with regressive attitudes were more likely to adopt the plough, the OLS estimates would be biased upwards. Conversely, if economically developed societies were more inclined towards plough adoption then our estimates would be biased downwards since the equivalent present-day countries are likely to be richer and consequently, have more gender equal norms.

To solve for this, I adopt the instrumental variable approach used by Alesina et al. (2013). The identification strategy relies on the hypothesis of Pryor (1985) that the plough adoption was driven by the characteristics of crops grown in a region. Crops whose cultivation benefited greatly from adoption of the plough, referred to as 'plough-positive' crops, included wheat, barley, and rye while those that did not face such benefits ('plough-negative' crops) included foxtail millet, pearl millet, and sorghum. Since geo-climatic conditions drove the cultivation of these crops in different regions, we can use suitability of a location for growing plough-positive and plough-negative crops as an exogenous instrument for plough adoption. The first stage regression for testing this relationship is given below:

$$plough_c = \beta_0 + \beta_1 ploughpositive\_c + \beta_1 ploughnegative\_c + X_c^H + \varepsilon_c \tag{7}$$

where $plough_c$ has the same interpretation as in equation (6). This is regressed on $ploughpositive\_c$ and $ploughnegative\_c$ which are country-level averages of the area suitable for cultivation of each type of crop, normalized by the overall arable land. Historical controls are included as

before, and standard errors are clustered at the country level.

For the second stage, predicted values of plough use are plugged into our original country-level equations:

$$flfp_{ct} = \beta_0 + \beta_1 \widehat{plough}\_c + X_c^H + X_c^P + \lambda_t + \varepsilon_{ct} \tag{8}$$

$$norm_{ct} = \beta_0 + \beta_1 \widehat{plough}\_c + X_c^H + X_c^P + \lambda_t + \varepsilon_{ct} \tag{9}$$

$$penalty_c = \beta_0 + \beta_1 \widehat{plough}\_c + X_c^H + X_c^P + \varepsilon_c \tag{10}$$

where equation (8) captures the causal impact of plough use on labour force participation of all women, women aged 18-40, and mothers from the DHS; equation (9) looks at the impact on the aggregate gender norms index and some individual indicators from the WVS; and equation (10) looks at the impact on the motherhood penalty estimates and gender gap in employment figures provided by Kleven et al. (2023). Since the DHS and WVS include country surveys from multiple years, time fixed effects are included in equations (8) and (9) but not for equation (10). The interpretation of the historical and contemporary controls is the same as before, with standard errors clustered at the country level for all equations.

5 Results

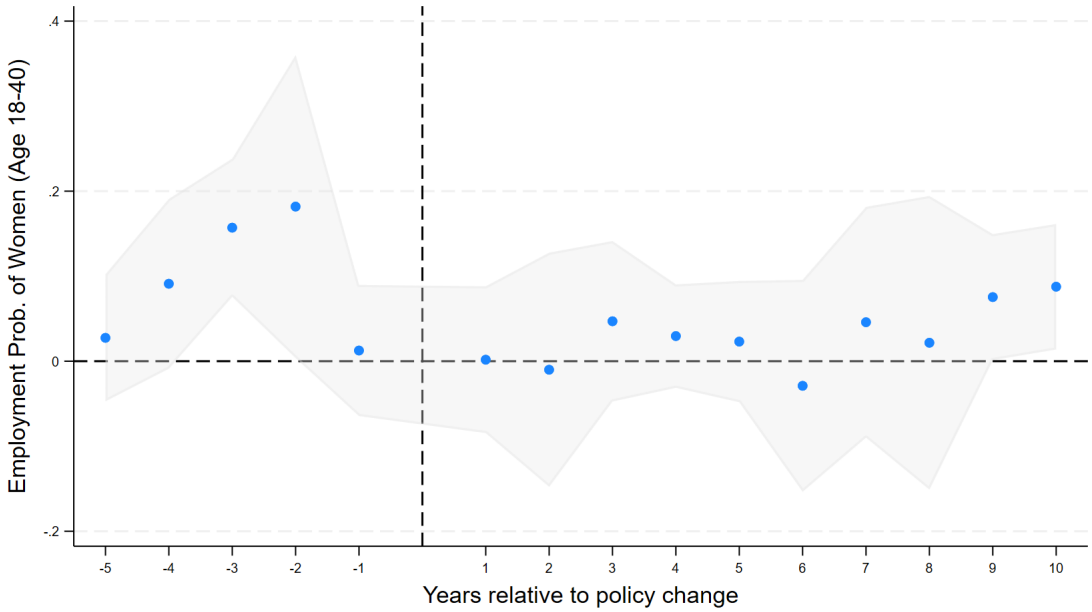
5.1 Work-Family Policies

Results from estimating equation (1) are provided in Figures 1 and 2. Figure 1 depicts the impact of countries' policy environment becoming more gender egalitarian— proxied by the WBL index (reported on a scale of 0-100) crossing the midpoint threshold of 50— on the labour force participation of women of childbearing age (18-40). Coefficients δ_k are plotted for all event study years from -5 to 10, taking year 0 as the base group, along with 95% confidence intervals.

Broadly, Figure 1 shows that there are no short-run effects emerging from crossing the threshold. In the long run, that is year 9 and 10 post treatment, the labour force participation is higher by 7.54 and 8.76 percentage points respectively— with statistical significance at the 5% level. However, an important caveat to these results is that we can observe upward pre-trends in women's employment from event study year -5 to -2. These could be biasing upwards our interpretation of the impact of a policy environment shift. However, the results are still meaningful for two reasons. First, the trends do not hold from year -1 onwards, with the estimates holding stable near zero for a long period before a minimal upward trend from years 8-10 as discussed above. Second, the pre-trends could be driven by differential

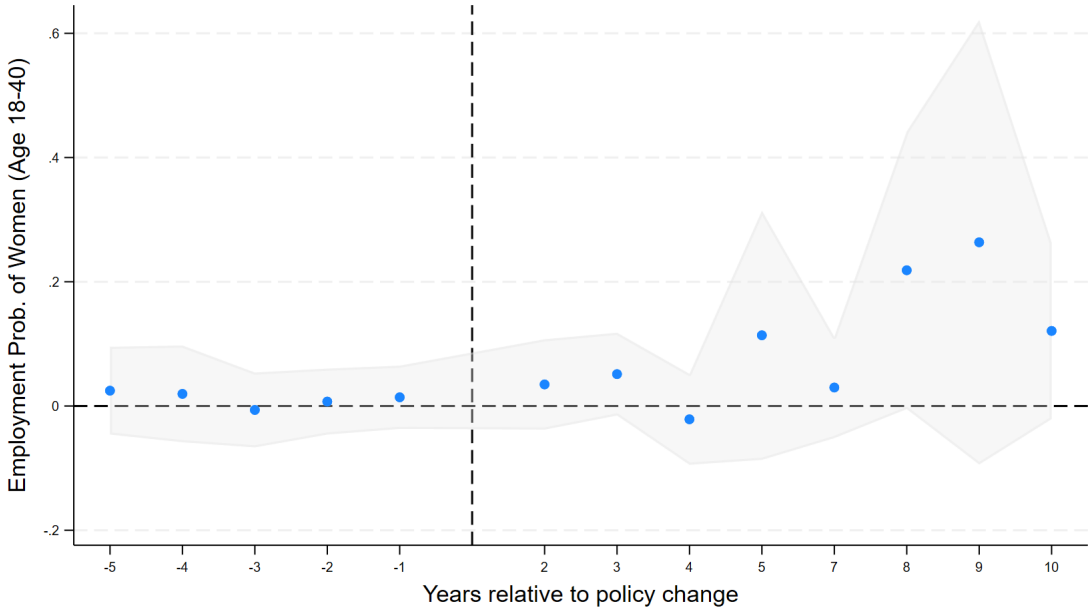
labour force participation rates only in the countries that were surveyed during event period -4 to -2. As an added check, I also provide TWFE event study estimates from the WBL index crossing the median threshold of 66. The results for the same are provided in Figure 4 in Appendix B, and they validate my results here that a policy environment shift towards gender equality has negligible to no impact on employment of women of childbearing age.

Figure 1: **Impact of WBL index crossing score of 50 (on scale of 100)**



Note: Missing ES years had nil observations or no country-year variation (so collinear with FE).

Figure 2: **Impact of policy guaranteeing atleast 14 weeks of paid leave to mothers**



Note: Missing ES years had nil observations or no country-year variation (so collinear with FE).

I also separately estimate the impact of paid maternity leave provision, depicted in Figure 2, since it has been emphasized by the literature on work-family policies (Lalive and Zweimüller 2009; Schönberg and Ludsteck 2014; Dahl et al. 2016; Baum 2003; among others). Coefficients δ_k are plotted for all event study years from -5 to 10, except k=1 and k=6 which are missing due to reasons discussed in Section 4.

No pre-trends appear for the maternity leave provision policy, providing confidence that factors other from the policy shock are not driving my results here. As before, Figure 2 shows that there is no impact of the policy in the short run but strong upward trends appear post year 5. In year 8 post treatment, female labour force participation is 21.86 percentage points higher than the year of the policy implementation (statistically significant at the 5% level). The long-run estimates are not only statistically significant but also hold significance since average labour force participation of women in the sample stands at 45%. I also provide estimates of another policy, prohibition of dismissal of pregnant workers, in Figure 5 provided in Appendix B.

5.2 Norms: OLS estimates

5.2.1 Individual-level OLS estimates

Results from estimating equation (2) are provided in Table 1. Each specification is run with and without region fixed effects to control for other systemic differences across regions that

may be driving our results. Estimates controlling for region fixed effects are reported in the even numbered columns.

Results from Table (1) validate the findings of Alesina et al. (2013) that traditional plough use continues to negatively impact present day female labour force participation. The results hold for all three samples of women that I run my estimation equation on, with the greatest impact on employment of mothers. The coefficient on plough use falls upon including region fixed effects but remains statistically significant at the 1% level as well as economically significant. For example, based on the estimates from column (6), a 1 standard deviation increase in traditional plough use (=0.441) is associated with a fall in maternal employment of 8.82 (20 x 0.441) percentage points, which is equivalent to 17.19% of the sample mean of mother's labour force participation and 17.65% of its standard deviation.

Table 2: Individual-Level OLS Estimates, Current FLFP (DHS) on Plough Use

	DHS FLFP sample					
	All women		Women age 18-40		Mothers	
	(1)	(2)	(3)	(4)	(5)	(6)
Plough use	-0.33***	-0.17***	-0.34***	-0.18***	-0.38***	-0.20***
	(0.04)	(0.05)	(0.04)	(0.05)	(0.05)	(0.05)
Historical Controls:						
Agricultural suitability	-0.06	-0.11**	-0.06	-0.11**	-0.04	-0.10*
	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
Tropical climate	-0.02	0.10	-0.02	0.11	-0.02	0.09
	(0.08)	(0.10)	(0.08)	(0.11)	(0.09)	(0.11)
Presence of large animals	-0.13	-0.30***	-0.14	-0.34***	-0.13	-0.33***
	(0.11)	(0.11)	(0.12)	(0.12)	(0.12)	(0.12)
Political hierarchies	0.05***	0.06***	0.05***	0.07***	0.05**	0.07***
	(0.02)	(0.01)	(0.02)	(0.01)	(0.02)	(0.02)
Economic complexity	0.04***	0.05***	0.04***	0.05***	0.04***	0.05***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Contemporary Controls:						
GDP per capita	-0.48	-0.09	-0.52	-0.10	-0.40	0.05
	(0.38)	(0.40)	(0.40)	(0.43)	(0.42)	(0.44)
GDP per capita sq.	0.02	0.00	0.03	0.00	0.02	-0.01
	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)	(0.03)
Individual-level controls	yes	yes	yes	yes	yes	yes
Region fixed effects	no	yes	no	yes	no	yes
Year fixed effects	yes	yes	yes	yes	yes	yes
Observations	2878128	2878128	2062763	2062763	1996647	1996647

Results from estimating equation (3) presented in Table 2 emphasize that this negative impact on women's present day employment works through the channel of restrictive gender

norms. I alternate between continent and country fixed effects for each specification, with the latter reported in even numbered columns. Since I am only focusing on the sample of WVS respondents from countries that are also in the DHS, and limited by data availability on plough use from Alesina et al. (2013), these results come from respondents in a smaller subset of 20 countries.

Table 3: Individual-Level OLS Estimates, Current Norms from WVS on Plough Use

	WVS respondents from DHS countries					
	Norms Index		Job Scarce		Working Mom	
	(1)	(2)	(3)	(4)	(5)	(6)
Plough use	-0.18**	-0.15**	0.07	0.09	0.32***	0.06
	(0.07)	(0.07)	(0.06)	(0.09)	(0.07)	(0.07)
District-level historical controls	yes	yes	yes	yes	yes	yes
Contemporary country controls	yes	n/a	yes	n/a	yes	n/a
Individual-level controls	yes	yes	yes	yes	yes	yes
Fixed effects	continent	country	continent	country	continent	country
Observations	18170	18170	19209	19209	22430	22430

The results are again in line with the hypothesis that traditional plough use led to evolution of more traditional norms regarding the appropriate role of women and mothers in society. The coefficient on my constructed norm index is in the expected direction (lower scores imply worse norms), and is robust to inclusion of continent and country fixed effects. Based on estimates from column (2), a 1 standard deviation increase in plough use (=0.459) leads to a reduction in the gender norms index amounting to more than 10% of its standard deviation (0.069 = 0.15 x 0.459). The coefficients on the individual indicators are also in the expected direction (for each, 1 is coded as agreement with a regressive statement) but lose statistical significance upon inclusion of fixed effects.

5.2.2 Country-level OLS estimates

As mentioned previously, cross-country regressions are run on the intuition that unobservable individual characteristics may be biasing my results from the previous section. Country regressions capture variation in norms that are prevalent at a broader geographical level, and are thus likely to be more accurate in capturing the impact on norms and women's employment. Results from equation (4) are given in Table 3, with the coefficient on plough use closely matching that of the individual-level regressions and robust to the inclusion of region fixed effects.

Results from estimating equation (5), providing country-level regressions of norms on traditional plough use, are provided in Table 4. Consistent with the aforementioned

Table 4: Country-Year Level OLS Estimates, Current FLFP (DHS) on Plough Use

	DHS FLFP sample					
	All women		Women age 18-40		Mothers	
	(1)	(2)	(3)	(4)	(5)	(6)
Traditional plough use	-0.26***	-0.17***	-0.28***	-0.18***	-0.31***	-0.20***
	(0.03)	(0.04)	(0.03)	(0.04)	(0.03)	(0.05)
Historical Controls:						
Agricultural suitability	-0.05	-0.14***	-0.02	-0.12***	-0.02	-0.11***
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
Tropical climate	-0.04	0.06	-0.03	0.06	-0.06	0.05
	(0.06)	(0.09)	(0.06)	(0.09)	(0.07)	(0.10)
Large animals	-0.15	-0.18	-0.21*	-0.24*	-0.19	-0.24*
	(0.11)	(0.13)	(0.12)	(0.13)	(0.12)	(0.13)
Political hierarchies	0.05***	0.05***	0.05***	0.05***	0.05***	0.06***
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)
Economic complexity	0.04***	0.04***	0.03***	0.04***	0.03***	0.04***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Contemporary Controls:						
GDP per capita	-0.53*	-0.28	-0.47	-0.18	-0.41	-0.10
	(0.29)	(0.31)	(0.30)	(0.32)	(0.31)	(0.33)
GDP per capita sq.	0.03	0.01	0.02	0.01	0.02	0.00
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Region fixed effects	no	yes	no	yes	no	yes
Year fixed effects	yes	yes	yes	yes	yes	yes
Observations	275	275	275	275	275	275

argument that norms are more relevant at the country level than individual, coefficients here are much stronger (in the expected direction) compared to results from equation (3). Based on estimates from column (2), a 1 standard deviation increase in plough use (=0.454) is associated with a worsening of the gender norms index equivalent to 34% of its standard deviation. This results is also statistically significant at the 5% level.

As an added robustness check, and to provide pointed evidence on the motherhood penalty, I also run country-level OLS regression of child penalty estimates and the gender gap in employment figures (from Kleven et al. 2023) on traditional plough use. Results for the same are consistent with my estimates from the DHS and are provided in Table 9 of Appendix B.

Table 5: Country-Year Level OLS Estimates, Current Norms from WVS on Plough Use

	WVS countries also in the DHS					
	Norms Index		Job Scarce		Working Mom	
	(1)	(2)	(3)	(4)	(5)	(6)
Traditional plough use	-1.12***	-0.54**	0.31***	0.07	0.23***	0.42***
	(0.16)	(0.22)	(0.06)	(0.07)	(0.06)	(0.11)
Historical Controls:						
Agricultural suitability	-0.15	0.45	0.07	-0.19*	0.02	0.21
	(0.23)	(0.30)	(0.09)	(0.10)	(0.08)	(0.16)
Tropical climate	-0.13	-0.17	0.08	0.05	0.20***	0.20***
	(0.20)	(0.21)	(0.09)	(0.08)	(0.06)	(0.07)
Large animals	-0.96**	0.32	0.51***	-0.04	-0.19*	0.16
	(0.43)	(0.55)	(0.16)	(0.16)	(0.10)	(0.23)
Political hierarchies	0.07	-0.12	-0.03	0.04	0.03	-0.00
	(0.10)	(0.11)	(0.04)	(0.04)	(0.02)	(0.04)
Economic complexity	0.22***	0.17***	-0.08***	-0.04***	-0.02	-0.04*
	(0.05)	(0.05)	(0.02)	(0.01)	(0.01)	(0.02)
Contemporary Controls:						
GDP per capita	-7.17***	-8.13***	1.49**	1.91***	1.10*	0.78
	(1.98)	(1.42)	(0.65)	(0.43)	(0.58)	(0.58)
GDP per capita sq.	0.41***	0.46***	-0.09**	-0.11***	-0.06*	-0.05
	(0.11)	(0.08)	(0.04)	(0.02)	(0.03)	(0.03)
Region fixed effects	no	yes	no	yes	no	yes
Year fixed effects	yes	yes	yes	yes	yes	yes
Observations	54	54	54	54	54	54

5.3 Norms: Instrumental Variable

As discussed in Section 4, the OLS estimates may be biased if plough adoption is correlated with other factors that also impact gender norms today. To solve for this, I instrument plough adoption with 'plough-positive' and 'plough-negative' crop suitability measures at the country-level. Results from equation (7) provided in Table 5 confirms the relevance of the instrument for my DHS country sample:

Table 6: First Stage Regression: Confirming Relevance of Instrument

	(1)
	Traditional plough use
Plough positive crops	0.85***
	(0.16)
Plough negative crops	0.16
	(0.34)
Historical controls:	
Agricultural suitability	-0.45***
	(0.14)
Tropical climate	0.11
	(0.17)
Large animals	0.32
	(0.38)
Political hierarchies	0.15***
	(0.04)
Economic complexity	0.03
	(0.03)
Observations	68
R-squared	0.70

Having established that the instrument holds relevance, in accordance with the hypothesis of Pryor (1985), I move onto showing my second stage results. Estimates from equation (8) are provided in Table 6 below.

My IV estimates measuring the the impact of plough use on women's employment are similar in magnitude and direction to my OLS results. Indeed, some of the coefficients are larger on an absolute scale— implying that my OLS estimates were biased downward. One way in which this could have manifested is that traditional societies which were more economically advanced were also more likely to adopt the plough, with these societies also being more likely to be better off today and have more progressive gender norms (as a result of their economic development). Estimates from column (5), for example, show that maternal employment is affected most severely: a 1 standard deviation increase in plough

Table 7: Country-Year Level IV Estimates, Current FLFP (DHS) on Predicted Plough Use

	DHS FLFP sample					
	All women		Women age 18-40		Mothers	
	(1)	(2)	(3)	(4)	(5)	(6)
Traditional plough use	-0.31***	-0.19	-0.34***	-0.18	-0.36***	-0.21
	(0.11)	(0.13)	(0.11)	(0.14)	(0.12)	(0.14)
Historical Controls:						
Agricultural suitability	-0.06	-0.14**	-0.04	-0.12**	-0.03	-0.12*
	(0.05)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)
Tropical climate	-0.06	0.06	-0.05	0.06	-0.08	0.04
	(0.08)	(0.10)	(0.09)	(0.11)	(0.09)	(0.11)
Large animals	-0.13	-0.18	-0.19	-0.24	-0.17	-0.24
	(0.18)	(0.15)	(0.19)	(0.16)	(0.20)	(0.16)
Political hierarchies	0.06**	0.05**	0.06**	0.05**	0.06**	0.06**
	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Economic complexity	0.03**	0.04***	0.03**	0.04***	0.03**	0.04***
	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.02)
Contemporary Controls:						
GDP per capita	-0.47	-0.28	-0.41	-0.18	-0.35	-0.10
	(0.37)	(0.39)	(0.41)	(0.44)	(0.41)	(0.45)
GDP per capita sq.	0.02	0.01	0.02	0.01	0.02	0.00
	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)	(0.03)
Region fixed effects	no	yes	no	yes	no	yes
Year fixed effects	yes	yes	yes	yes	yes	yes
Observations	275	275	275	275	275	275

use (=0.439) leads to a 15.8 percentage points decrease in the probability of labour force participation of mothers. However, as opposed to the OLS estimates, inclusion of region fixed effects removes the statistical significance of these coefficients. This means that there is wide heterogeneity in impact on female employment across regions, with norms mattering more in some regions and not so much in others.

Similarly, my IV results for impact on the gender norms index shows enlarged coefficients—hinting at downward biased OLS estimates consistent with IV results on employment— that lose significance upon including region fixed effects. A notable exception is the coefficient on the dummy indicator checking agreement with the statement "When a mother works for pay, the children suffer", which becomes larger upon inclusion of the regional control. Interpreting the estimate from column (6): a 1 standard deviation increase in traditional plough use (=0.454) causes an increase in agreement with the aforementioned statement by 24.97 percentage points, which is 1.34 times the standard deviation of the indicator.

Table 8: Country-Year Level IV Estimates, WVS Gender Norms on Predicted Plough Use

	WVS countries also in the DHS					
	Norms Index		Job Scarce		Working Mom	
	(1)	(2)	(3)	(4)	(5)	(6)
Plough use	-1.19***	-0.97	0.35***	0.29	0.39***	0.55***
	(0.33)	(0.65)	(0.12)	(0.23)	(0.12)	(0.19)
Historical Controls:						
Agricultural suitability	-0.16	0.14	0.08	-0.03	0.06	0.31
	(0.25)	(0.53)	(0.09)	(0.19)	(0.10)	(0.20)
Tropical climate	-0.13	-0.16	0.09	0.05	0.21***	0.20**
	(0.21)	(0.23)	(0.09)	(0.10)	(0.06)	(0.07)
Large animals	-0.93*	-0.14	0.49***	0.19	-0.26**	0.31
	(0.46)	(0.95)	(0.17)	(0.33)	(0.12)	(0.31)
Political hierarchies	0.08	-0.03	-0.04	-0.00	0.01	-0.03
	(0.11)	(0.19)	(0.04)	(0.07)	(0.04)	(0.06)
Economic complexity	0.23***	0.21**	-0.08***	-0.06*	-0.03	-0.06*
	(0.06)	(0.09)	(0.02)	(0.03)	(0.02)	(0.03)
Contemporary Controls:						
GDP per capita	-7.28***	-7.97***	1.57**	1.83***	1.35**	0.73
	(2.17)	(1.79)	(0.74)	(0.62)	(0.60)	(0.63)
GDP per capita sq.	0.42***	0.46***	-0.09**	-0.10***	-0.08**	-0.04
	(0.12)	(0.10)	(0.04)	(0.03)	(0.03)	(0.03)
Region fixed effects	no	yes	no	yes	no	yes
Year fixed effects	yes	yes	yes	yes	yes	yes
Observations	54	54	54	54	54	54

The impact of norms on maternal employment is corroborated by IV estimates using data from the Child Penalty Atlas (Kleven et al. 2024). These results, provided in Table 10 of Appendix B, show economically large impacts that stay in place even after inclusion of continent fixed effects. Among the sample of 69 countries that are both in the Atlas and the DHS, going from a country with no ancestral plough use to one where everyone used the plough increases the motherhood penalty by 24.25 percentage points and the raw gender gap in employment by 35 percentage points.

6 Discussion and Conclusion

I find that cultural norms— determined historically based on the form of agriculture practiced by societies— have a significant and lasting impact on attitudes towards the appropriate role of women and mothers in society, and consequently on their labour force participation. Consistent with the findings of Alesina et al. (2013), in countries where ancestral plough adoption was greater, women are less likely to be in the labour force and attitudes towards women are more regressive. These effects are magnified for mothers, as seen in the results from my DHS subsample of mothers and the child penalty estimates provided by Kleven et al. (2024), suggesting that gender norms become stricter for women after childbirth. Indeed, countries with greater plough adoption also had more people agreeing with the sentiment that children suffer when mothers go out for work.

The link between regressive cultural attitudes and actual reduced maternal employment could manifest in many ways. Take, for instance, an employer who holds the aforementioned sentiment regarding working mothers. Given information on the motherhood status of a candidate, the employer is much more likely to discriminate and not hire mothers (Correll et al. 2007). Another possibility is that mothers who internalize these views may self-select out of the labour market because of "guilt" arising from not engaging in their socially prescribed role as caregiver (Fortin 2005). Finally, a more long-drawn cause could be gender differences in human capital that stem from childhood: regressive gender norms that give lower value to a girl child may lead to families reducing investment in their daughters' education and health, as has been documented extensively in the literature.

Can work-family policies attenuate this penalty? My results provide mixed evidence to this end, with some policies working better than others. While broader shifts in the policy environment towards gender equality have negligible short-run and long-run impacts, more directed policy efforts like the legal guarantee of paid maternity leave can increase the employment probability of women of childbearing age by almost 20 percentage points in the long run. Other policies like free provision of childcare services are also likely to

have important effects for maternal employment but my analysis is limited by the lack of cross-country and time-series data on them.

This study thus provides robust evidence on the long shadow of gender norms and the moderating role played by work-family policies in a large sample of developing countries. For policymakers, my work has two important implications. First, while recent changes to the foreign aid space have led to dismantling of the DHS programme, I show the value behind funding these large-scale data collection exercises in developing countries. They provide rich insights into demographic transitions, fertility dynamics, health indicators, as well as the labour force participation of women and mothers (as evidenced in this study)— information that has great economic significance (Hsieh et al. 2019) but which would not be known if not for the DHS. Second, and more importantly, the comparative analysis between norms and policies provided here can open up the literature to more nuanced questions about the interaction effects between them. It allows policymakers to design work-family interventions that are not isolated from their cultural environment, but rather informed by them.

References

1. Akerlof, George A., and Rachel E. Kranton. 2000. "Economics and Identity." *Quarterly Journal of Economics* 115 (3): 715–53. doi:10.1162/003355300554881.

2. Alesina, Alberto, and Paola Giuliano. 2015. "Culture and Institutions." *Journal of Economic Literature* 53 (4): 898–944. doi:10.1257/jel.53.4.898.

3. Alesina, Alberto, Paola Giuliano, and Nathan Nunn. 2013. "On the Origins of Gender Roles: Women and the Plough." *The Quarterly Journal of Economics* 128 (2): 469–530. doi:10.1093/qje/qjt005.

4. Baker, Michael, Jonathan Gruber, and Kevin Milligan. 2005. Universal Childcare, Maternal Labor Supply, and Family Well-Being, December. doi:10.3386/w11832.

5. Baum, Charles L. 2003. "The Effects of Maternity Leave Legislation on Mothers' Labor Supply after Childbirth." *Southern Economic Journal* 69 (4): 772–99. doi:10.1002/j.2325-8012.2003.tb00533.x.

6. Bedi, Arjun S., Tanmoy Majilla, and Matthias Rieger. 2018. "Gender Norms and the Motherhood Penalty: Experimental Evidence from India." SSRN Electronic Journal. doi:10.2139/ssrn.3137490.

7. Bisin, Alberto, and Thierry Verdier. 2001. "The Economics of Cultural Transmission and the Dynamics of Preferences." *Journal of Economic Theory* 97 (2): 298–319. doi:10.1006/jeth.2000.2678.

8. Bisin, Alberto, and Thierry Verdier. 2011. "Chapter 9 - The Economics of Cultural Transmission and Socialization" *Handbook of Social Economics* 1: 339–416. doi:10.1016/B978-0-444-53187-2.00009-7.

9. Bisin, Alberto, and Thierry Verdier. 2023. "Advances in the Economic Theory of Cultural Transmission." *Annual Review of Economics* 15 (1): 63–89. doi:10.1146/annurev-economics-090622-100258.

10. Boeckmann, Irene, Joya Misra, and Michelle J. Budig. 2014. "Cultural and Institutional Factors Shaping Mothers' Employment and Working Hours in Postindustrial Countries." *Social Forces* 93 (4): 1301–33. doi:10.1093/sf/sou119.

11. Boeckmann, Irene, Joya Misra, and Michelle J. Budig. 2014a. "Cultural and Institutional Factors Shaping Mothers' Employment and Working Hours in Postindustrial Countries." *Social Forces* 93 (4): 1301–33. doi:10.1093/sf/sou119.

12. Boelmann, Barbara, Anna Raute, and Uta Schönberg. 2025. "Wind of Change? Cultural Determinants of Maternal Labor Supply." *American Economic Journal: Applied Economics* 17 (2): 41–74. doi:10.1257/app.20220567.

13. Boserup, Ester, Su Fei Tan, and Camilla Toulmin. 2013. Woman's Role in Economic Development, July. doi:10.4324/9781315065892.

14. Budig, Michelle J., Joya Misra, and Irene Boeckmann. 2012. "The Motherhood Penalty in Cross-National Perspective: The Importance of Work-Family Policies and Cultural Attitudes." *Social Politics: International Studies in Gender, State & Society* 19 (2): 163–93. doi:10.1093/sp/jxs006.

15. Burda, Michael, Daniel S. Hamermesh, and Philippe Weil. 2013. "Total Work and Gender: Facts and Possible Explanations." *Journal of Population Economics* 26: 239–261. https://doi.org/10.1007/s00148-012-0408-x.

16. Campaña, Juan Carlos, Jose Ignacio Giménez-Nadal, and José Alberto Molina. 2017. "Gender Norms and the Gendered Distribution of Total Work in Latin American Households." *Feminist Economics* 24 (1): 35–62. doi:10.1080/13545701.2017.1390320.

17. Cavapozzi, Danilo, Marco Francesconi, and Cheti Nicoletti. 2021. "The Impact of Gender Role Norms on Mothers' Labor Supply." *Journal of Economic Behavior & Organization* 186 (June): 113–34. doi:10.1016/j.jebo.2021.03.033.

18. Correll, Shelley J., Stephen Benard, and In Paik. 2007. "Getting a Job: Is There a Motherhood Penalty?" *American Journal of Sociology* 112, no. 5 (March): 1297-1338. https://doi.org/10.1086/511799.

19. Cukrowska-Torzewska, Ewa, and Anna Matysiak. 2020. "The motherhood wage penalty: A meta-analysis." *Social Science Research* 88-89, no. 102416 (May). https://doi.org/10.1016/j.ssresearch.2020.102416.

20. Cukrowska-Torzewska, Ewa. 2017. "Cross-Country Evidence on Motherhood Employment and Wage Gaps: The Role of Work–Family Policies and Their Interaction." *Social Politics: International Studies in Gender, State & Society* 24 (2): 178–220. doi:10.1093/sp/jxx004.

21. Dahl, Gordon B., Katrine V. Løken, Magne Mogstad, and Kari Vea Salvanes. 2016. "What Is the Case for Paid Maternity Leave?" *Review of Economics and Statistics* 98 (4): 655–70. doi:10.1162/rest_a_00602.

22. Fernández, Raquel. 2011. "Chapter 11 - Does Culture Matter?" *Handbook of Social Economics* 1: 481–510. doi:10.1016/b978-0-444-53187-2.00011-5.

23. Fortin, Nicole M. 2005. "Gender Role Attitudes and the Labour-Market Outcomes of Women across OECD Countries." *Oxford Review of Economic Policy* 21 (3): 416–38. doi:10.1093/oxrep/gri024.

24. Giavazzi, Francesco, Ivan Petkov, and Fabio Schiantarelli. 2019. "Culture: Persistence and Evolution." *Journal of Economic Growth* 24 (2): 117–54. doi:10.1007/s10887-019-09166-2.

25. Goldin, Claudia. 1994. The U-Shaped Female Labor Force Function in Economic Development and Economic History, April. NBER Working Paper 4707, doi:10.3386/w4707.

26. Grönlund, Anne, Karin Halldén, and Charlotta Magnusson. 2016. "A Scandinavian Success Story? Women's Labour Market Outcomes in Denmark, Finland, Norway and Sweden." *Acta Sociologica* 60 (2): 97–119. doi:10.1177/0001699316660595.

27. Guiso, Luigi, Paola Sapienza, and Luigi Zingales. 2006. "Does Culture Affect Economic Outcomes?" *Journal of Economic Perspectives* 20 (2): 23–48. doi:10.1257/jep.20.2.23

28. Hsieh, Chang-Tai, Erik Hurst, Charles I. Jones, and Peter J. Klenow. 2019. "The Allocation of Talent and U.S. Economic Growth." *Econometrica* 87, no. 5 (September): 1439-1474. https://doi.org/10.3982/ECTA11427.

29. ICF. 2004–2017. Demographic and Health Surveys (Various) [Datasets]. Rockville, MD: ICF. Funded by USAID.

30. Inglehart, Ronald, Christian Haerpfer, Alejandro Moreno, Christian Welzel, Kseniya Kizilova, José Diez-Medrano, Marta Lagos, Pippa Norris, Eduard Ponarin, and Bi Puranen, et al., eds. 2020. World Values Survey: All Rounds – Country-Pooled Datafile. Madrid and Vienna: JD Systems Institute and WVSA Secretariat. http://www.worldvaluessurvey.org/WVSDocumentationWVL.jsp.

31. Keck, W., and C. Saraceno. 2013. "The Impact of Different Social-Policy Frameworks on Social Inequalities among Women in the European Union: The Labour-Market Participation of Mothers." *Social Politics: International Studies in Gender, State & Society* 20 (3): 297–328. doi:10.1093/sp/jxt005.

32. Kleven, Henrik, Camille Landais, and Gabriel Leite-Mariante. 2024. "The Child Penalty Atlas." *Review of Economic Studies*, October. doi:10.1093/restud/rdae104.

33. Kleven, Henrik, Camille Landais, Johanna Posch, Andreas Steinhauer, and Josef Zweimuller. 2019. "Child Penalties across Countries: Evidence and Explanations." *AEA Papers and Proceedings* 109: 122–26. doi: 10.1257/pandp.20191078

34. Kleven, Henrik, Camille Landais, Johanna Posch, Andreas Steinhauer, and Josef Zweimuller. 2024. "Do Family Policies Reduce Gender Inequality? Evidence from 60 Years of Policy Experimentation." *American Economic Journal: Economic Policy* 16 (2): 110–49. doi: 10.1257/pol.20210346

35. Kleven, Henrik. 2022. The Geography of Child Penalties and Gender Norms: A Pseudo-Event Study Approach. NBER Working Paper 30176, doi:10.3386/w30176.

36. Kling, Jeffrey, Jeffrey Liebman, and Lawrence Katz. 2007. "Experimental Analysis of Neighborhood Effects." *Econometrica* 75 (1): 83–119.

37. Lalive, Rafael, Analia Schlosser, Andreas Steinhauer, and Josef Zweimuller. 2013. "Parental Leave and Mothers' Careers: The Relative Importance of Job Protection and Cash Benefits." *The Review of Economic Studies* 81 (1): 219–65. doi:10.1093/restud/rdt028.

38. Lalive, Rafael, and Josef Zweimüller. 2009. "How Does Parental Leave Affect Fertility and Return to Work? Evidence from Two Natural Experiments." *Quarterly Journal of Economics* 124 (3): 1363–1402. doi:10.1162/qjec.2009.124.3.1363.

39. Olivetti, Claudia, and Barbara Petrongolo. 2017. "The Economic Consequences of Family Policies: Lessons from a Century of Legislation in High-Income Countries." *Journal of Economic Perspectives* 31 (1): 205–30. doi:10.1257/jep.31.1.205.

40. Pettit, Becky, and Jennifer Hook. 2005. "The Structure of Women's Employment in Comparative Perspective." *Social Forces* 84 (2): 779–801. doi:10.1353/sof.2006.0029.

41. Pettit, Becky, and Jennifer L. Hook. 2009. "Gendered Tradeoffs: Women, Family, and Workplace Inequality in Twenty-One Countries". Russell Sage Foundation. JSTOR, http://www.jstor.org/stable/10.7758/9781610446785. Accessed 13 May 2025.

42. Pryor, Frederic L. 1985. "The Invention of the Plow." *Comparative Studies in Society and History* 27 (4): 727–43. doi:10.1017/s0010417500011749.

43. Schönberg, Uta, and Johannes Ludsteck. 2014. "Expansions in Maternity Leave Coverage and Mothers' Labor Market Outcomes after Childbirth." *Journal of Labor Economics* 32 (3): 469–505. doi:10.1086/675078.

44. Sevilla-Sanz, Almudena. 2010. "Household Division of Labor and Cross-Country Differences in Household Formation Rates." *Journal of Population Economics* 23: 225–249. https://doi.org/10.1007/s00148-009-0254-7.

45. Steinhauer, Andreas. 2018. "Working Moms, Childlessness, and Female Identity". CEPR Discussion Paper No. DP12929, Available at SSRN: https://ssrn.com/abstract=3178100

46. Zhang, Mingxue, Yue Wang, and Lingling Hou. 2024. "Gender Norms and the Child Penalty in China." *Journal of Economic Behavior & Organization* 221 (May): 277–91. doi:10.1016/j.jebo.2024.03.011.

Appendix A: Norms Index - WVS

For constructing the gender norms index, I use the following questions from the WVS:

1. When a mother works for pay, the children suffer
2. On the whole, men make better political leaders than women do
3. A university education is more important for a boy than a girl
4. On the whole, men make better business executives than women
5. Being a housewife is just as fulfilling as working for pay
6. When jobs are scarce, men should get preference over women

These questions are chosen based on how the previous literature has operationalized gender norms using the World Values Survey (Burda et al. 2007; Campana et al. 2017; Sevilla-Sanz 2009; Alesina et al. 2013), and data availability.

I use the method described by Katz et al. (2007) in constructing my method. First, the 3/5 point agreement scale for each question is converted to a binary indicator (agree-disagree), where 1 is assigned to agree and 0 for disagree. Two of these binary individual indicators are also used in my analysis alongside the aggregated index. Next, I standardize each indicator, i.e. create z-scores. The norm index for each individual is then the sign-oriented average of the 6 standardized norm indicators. The final index is made such that higher values represent better norms.

Appendix B: Tables and Figures

Figure 3: **Map of countries included in my DHS sample (N=82)**

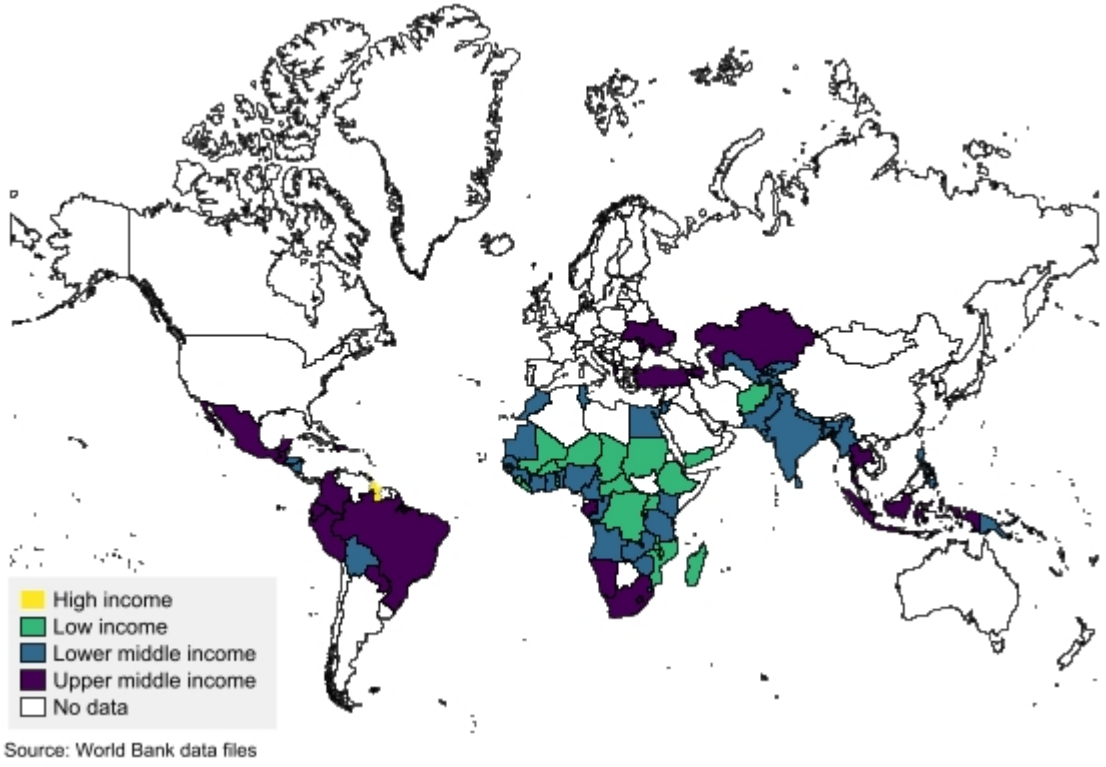
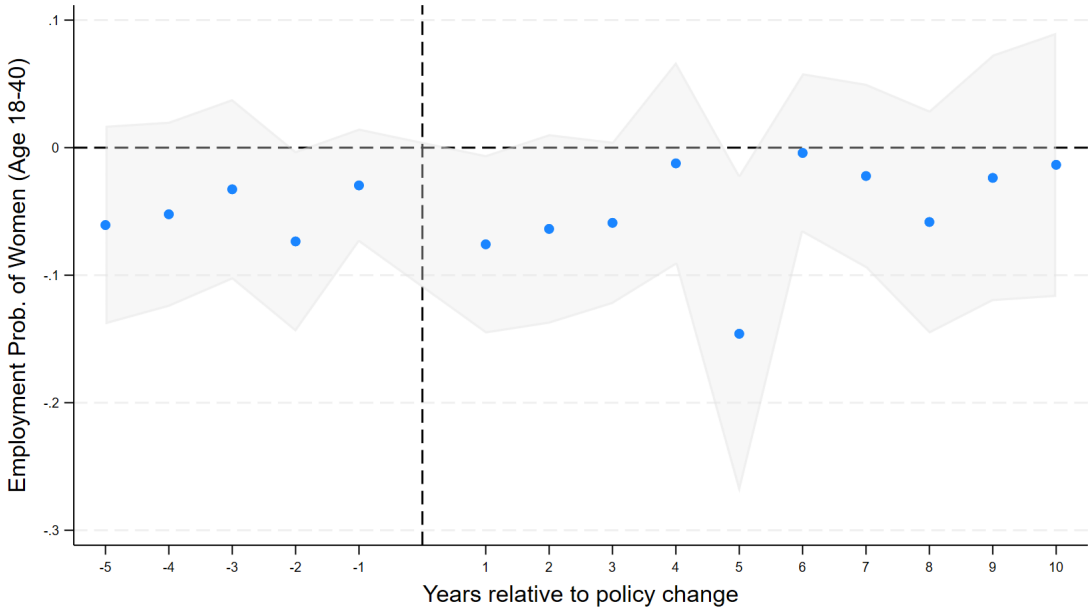
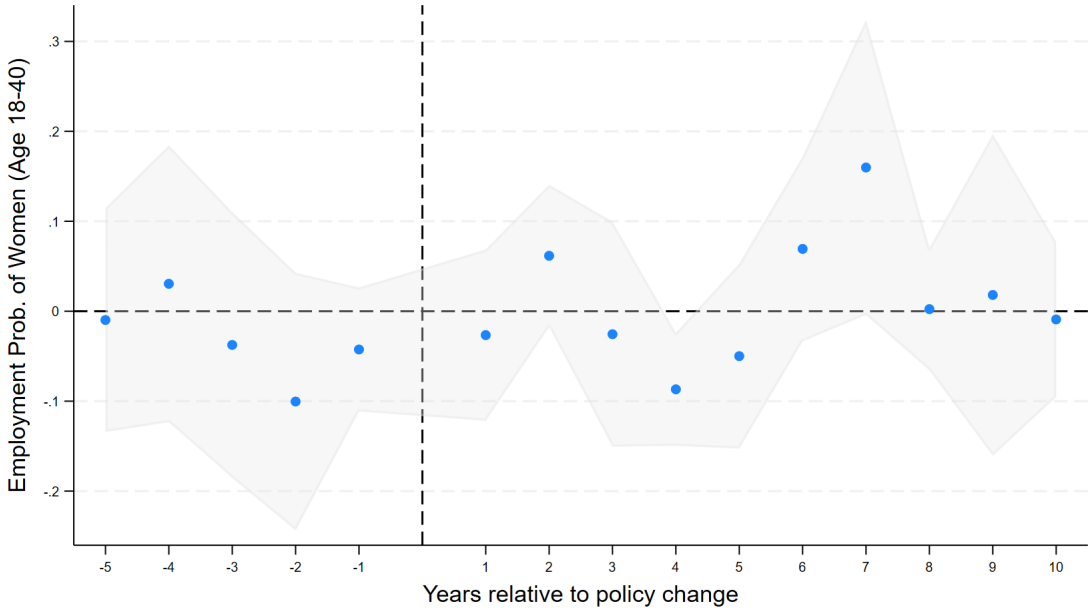


Figure 4: **Impact of WBL index crossing median threshold (=66) on scale of 100**



Note: Missing ES years had nil observations or no country-year variation (so collinear with FE).

Figure 5: **Impact of policy prohibiting dismissal of pregnant workers**



Note: Missing ES years had nil observations or no country-year variation (so collinear with FE).

Table 9: Country-Level OLS Estimates, Estimated Child Penalty on Plough Use

	Child Penalty Atlas overlap with DHS			
	Child Penalty		Gender gap in emp.	
	(1)	(2)	(3)	(4)
Traditional plough use	8.56	13.39*	0.15**	0.13
	(5.33)	(7.61)	(0.06)	(0.08)
Historical controls:				
Agricultural suitability	-4.04	2.36	-0.06	-0.03
	(5.22)	(5.44)	(0.07)	(0.08)
Tropical climate	-4.91	-0.42	0.05	0.06
	(8.95)	(10.06)	(0.10)	(0.10)
Large animals	-27.64***	-7.08	-0.15	0.07
	(8.65)	(11.05)	(0.13)	(0.13)
Political hierarchies	2.20	-1.48	0.01	-0.04
	(2.09)	(2.40)	(0.03)	(0.03)
Economic complexity	0.51	-1.26	-0.03	-0.05**
	(1.42)	(1.23)	(0.02)	(0.02)
Contemporary controls:				
GDP per capita 2010	56.19	40.03	1.36***	1.11**
	(41.01)	(32.04)	(0.51)	(0.48)
GDP per capita 2010 sq.	-2.84	-2.03	-0.08**	-0.06**
	(2.46)	(1.93)	(0.03)	(0.03)
Continent fixed effects	no	yes	no	yes
Observations	69	69	69	69

Table 10: Country Level IV Estimates, Current Child Penalty on Predicted Plough Use

	Child Penalty Atlas overlap with DHS			
	Child Penalty		Gender gap in emp.	
	(1)	(2)	(3)	(4)
Plough use	13.14	24.25*	0.21**	0.35**
	(9.21)	(12.61)	(0.11)	(0.15)
Historical controls:				
Agricultural suitability	-3.06	4.61	-0.05	-0.02
	(5.15)	(5.70)	(0.07)	(0.08)
Tropical climate	-3.88	-2.20	0.06	0.04
	(9.19)	(10.74)	(0.10)	(0.12)
Large animals	-27.95***	-1.65	-0.15	0.10
	(8.27)	(13.11)	(0.12)	(0.15)
Political hierarchies	1.08	-4.06	-0.01	-0.06
	(2.94)	(3.51)	(0.04)	(0.04)
Economic complexity	0.40	-1.35	-0.03	-0.05**
	(1.39)	(1.29)	(0.02)	(0.02)
Contemporary controls:				
GDP per capita 2010	51.60	33.31	1.30***	1.27***
	(39.81)	(31.74)	(0.49)	(0.46)
GDP per capita 2010 sq	-2.60	-1.69	-0.07**	-0.07***
	(2.39)	(1.91)	(0.03)	(0.03)
Continent fixed effects	no	yes	no	yes
Observations	69	69	69	69