I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 14898

Gender Norms and the Motherhood Employment Gap

Simone Moriconi
Núria Rodríguez-Planas

NOVEMBER 2021

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 14898

Gender Norms and the Motherhood Employment Gap

Simone Moriconi
Univ. Lille, CNRS, LEM and CESifo Munich

Núria Rodríguez-Planas
City University of New York, Barnard College at Columbia University, Queens College - CUNY and IZA

NOVEMBER 2021

Any opinions expressed in this paper are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but IZA takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The IZA Institute of Labor Economics is an independent economic research institute that conducts research in labor economics and offers evidence-based policy advice on labor market issues. Supported by the Deutsche Post Foundation, IZA runs the world's largest network of economists, whose research aims to provide answers to the global labor market challenges of our time. Our key objective is to build bridges between academic research, policymakers and society.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Gender Norms and the Motherhood Employment Gap*

Using individual-level data from the European Social Survey, we study the relevance of gender norms in accounting for the motherhood employment gap across 186 European NUTS2 regions (over 29 countries) for the 2002-2016 period. The gender norm variable is taken from a question on whether "men should have more right to a job than women when jobs are scarce" and represents the average extent of disagreement (on a scale 1 to 5) of women belonging to the "grandmothers" cohort. We address the potential endogeneity of our gender norms measure with an index of the degree of reproductive health liberalization when grandmothers were 20 years old. We also account for the endogeneity of motherhood with the level of reproductive health liberalization when mothers were 20 years old. We find a robust positive association between progressive beliefs among the grandmothers' cohort and mothers' likelihood to work while having a small child (0 to 5 years old) relative to similar women without children. No similar association is found among men. Our analysis underscores the role of gender norms and maternal employment, suggesting that non-traditional gender norms mediate on the employment gender gap mainly via motherhood.

JEL Classification: J16, J22

Keywords: gender norms, motherhood employment gap, instrumenting for motherhood

Corresponding author:
Núria Rodríguez-Planas
Queens College - CUNY
Economics Department
Powdermaker Hall
65-30 Kissena Blvd.
Queens, New York 11367
USA

E-mail: nuria.rodriguezplanas@qc.cuny.edu

* The authors would like to thank Purushottam Mohanty for precious research assistance; and Joanna Tyrowitz for bringing the dataset on world reproductive laws by Finlay, Canning, and Po (2012) to our attention. The authors also would like to recognize the comments that helped us improve the paper from participants at IESEG EQM department seminar, the 2019 Gender Economics Workshop in Banco de España, the 1st Workshop in Gender and Economics at the University of Luxembourg, the Women's Applied Microeconomics Workshop at Columbia University, the first Manheim-Bonn Labor and Public Economics workshop, Passau University, the 2021 meetings Society of Economics of the Household, and the LEM ADP/MOF Webinar series. Simone Moriconi received financial support by the French National Research Agency for the project "MALYNES" (ANR-18-CE26-000, AAPG 2018), as well as LEM-CNRS 9221 ("appel à projet recurrent"). Both institutions are gratefully acknowledged.

I. Introduction

Since the 1990s, adult women's and men's convergence in labor force participation rate and earnings has stalled and even reversed in many OECD countries (Goldin 2014; Blau and Kahn 2006; and England, Levine and Mishel 2020). Recent studies underscore the relevance of motherhood for the persistence of gender inequality in labor market outcomes (Kleven *et al.* 2019). For instance, the earnings loss mothers experience compared to non-mothers (Waldfogel 1997, 1998; Anderson, Binder and Krause 2002, 2003; Gangl and Ziefle 2009; Fernández-Kranz, Lacuesta and Rodríguez-Planas 2013) or fathers (Angelov, Johansson and Lindhal 2016; Kuziemko *et al.* 2018; Kleven, Landais, and Søgaard 2019a; Kleven *et al.* 2019b) is one of the leading explanations for the gender earnings gap.[1] At the same time, the reduction in maternal labor employment (at both the intensive or extensive margin) is largely responsible for most of the motherhood earnings penalty as shown by Fernández-Kranz, Lacuesta and Rodríguez-Planas (2013) in Spain; Angelov, Johansson and Lindhal (2016) in Sweden; and Kleven et al. (2019b) in Scandinavian and Germanic countries, the UK and the US. In the current paper, we analyze the role gender norms play in explaining mothers' decision to work outside the household.[2]

According to Goldin (1995), with industrialization, women's role in society got relegated to being a wife and a mother and working mostly inside the household. Regardless of whether women internalized their assigned gender role, working outside the home was socially stigmatized as gender norms set up "acceptable behavioral boundaries for men and women, congruent with the gender division of labor and male power" (Seguino 2007). As societies evolved and gender norms were relaxed, women were gradually allowed to engage in traditionally male activities including formal education and paid employment in the labor market. While many have studied the role of gender norms on female labor force participation or fertility (Antecol 2000; Fortin 2005; Fernández and Fogli 2009; Blau et al. 2013; Bertrand, Kamenica and Pan 2015; Olivetti, Patacchini

[1] In a recent study, Kleven, Landais, & Søgaard (2019a) estimate that having a child is associated with 80 percent of the gender earnings gap in 2013 in Denmark. While the magnitude of this effect varies across countries, others have also underscored the relevance of motherhood in accounting for the persistence of gender inequality in earnings—Angelov, Johansson and Lindhal (2016) for Sweeden; Kuziemko *et al.* (2018) for US and UK; Kleven *et al.* (2019), for Sweeden, Denmark, the US, the UK, Austria and Germany; and De Quinto, Hospido and Sanz (2020) for Spain.

[2] There is an extent complementary literature on how institutions such as taxes, transfers, parental leave, and childcare policies are related to female labor supply and gender gaps (Olivetti and Petrongolo 2017; Fernández-Kranz and Rodríguez-Planas 2021).

and Zenou 2020; Rodríguez-Planas, Sanz-de-Galdeano and Terskaya 2018; Rodríguez-Planas and Tanaka 2021), few have analyzed whether gender norms mostly affect women's decision to work through motherhood. This is the objective of this paper.

To study the role of gender norms on maternal employment, we exploit geographic (186 European NUTS2 regions over 29 countries), cohort (2 cohorts) and time (2002-2016) variation in gender norms on the employment status of mothers with 0-5 years old children. We measure gender norms with gender role beliefs from individual survey data, in our case from the European Social Survey (2004, 2008, 2010 and 2016). Our measure of non-traditional gender norms is the average extent of disagreement (on a scale 1 to 5) with the statement: "*men should have more right to a job than women when jobs are scarce".* To address the reverse causality problem between contemporaneous peers' gender norms and maternal employment decisions, we focus on women belonging to the grandmothers' cohort to identify gender norms of the mothers. This is based on evidence that gender norms are acquired from women during childhood from their mother or their mothers' peers (Farré and Vella 2013; Olivetti, Patacchini and Zenou 2020; Rodríguez-Planas, Sanz-de-Galdeano, and Terskaya 2018; Kleven et al. 2019a).

Our baseline model blends the peer effects literature[3] with the motherhood gap literature to identify the effect of non-traditional norms in the grandmothers' cohort on mothers' decision to work relative to childless women. From the peer effects literature, we exploit idiosyncratic variation in average gender norms of the grandmothers' cohort across cohort, NUTS2 region, and survey year. In addition to individual, family, and regional controls, we include NUTS2-region fixed effects to control for endogenous sorting of individuals across regions, and country-and-year fixed effects to control for time-varying unobserved factors that might confound the relationship between grandmothers' non-traditional beliefs and the women's decision to work within a region. From the motherhood gap literature, we estimate the differential effect of grandmothers' gender norms on the employment decision of mothers of young children relative to similar-age childless women. By taking the difference, we "wash out" any confounding effects related to female employment (as opposed to maternal employment). Crucially, to address concerns that our social

[3] See Angrist and Lang (2004); Friesen and Krauth (2007); Hanushek et al. (2002); Hoxby (2000); Lavy and Schlosser (2011); Lavy, Paserman and Schlosser (2012); Bifulco, Fletcher and Ross (2011); and Olivetti, Patacchini, and Zenou (2020) for papers using peer effects in the field of education. Recently, Rodríguez-Planas and Tanaka (2021) have applied peer effects to explain Japanese women's employment decisions.

norms estimates may still be endogenous, we use an index on the degree of liberalization of reproductive health laws in the country when the grandmothers were in their 20s. Furthermore, we explicitly model the endogeneity of the motherhood decision of women in fertile age with the level of exposure the respondent had to reproductive-health liberalization laws when she was in her 20s.

We find a robust and consistent positive association between non-traditional norms in the grandmothers' cohort and mothers' likelihood to work while having a small child (0 to 5 years old) relative to childless women. As no such association is found among fathers of small children relative to childless men, our analysis stresses the relevance of gender norms on maternal employment (but not on paternal employment). Our results are robust to a battery of sensitivity tests.

Our preferred specification finds that a one-standard deviation increase in the predicted average non-traditional gender norm (that is 0.626) is associated with a reduction of 47 percent in the average motherhood employment gap in our sample of 12 percentage points. To put it differently, if mothers from Andalusia (ES23) in Spain, with an average disagreement to traditional gender norms among grandmothers of 2.36 and a motherhood employment gap of 9 percentage points in 2004, were characterized by the mean gender norms in Sweden (e.g. SE21, Småland med öarna), which is equal to 3.77, the statistical model would suggest that the motherhood employment gap in Andalusia would decrease by 12 percentage points, which would fully reverse the observed motherhood employment gap between Småland med öarna and Andalusia (equal to 9 percentage points). Our analysis underscores that non-traditional gender norms mediate on the employment gender gap mainly via motherhood, and hence, to the extent that the motherhood employment gap is largely responsible of the motherhood wage gap, our results highlight the relevance of gender norms on the gender wage gap in Europe.

Our paper contributes to two strands of the literature: the motherhood penalty; and female labor force participation and culture. Within the first strand, we focus on whether gender norms mediate in explaining the gender gap between mothers and non-mothers (but not between fathers and non-fathers). The few studies that mention the role of gender norms on the motherhood gap have been mostly descriptive and gender norms have not been the focus of their analysis. For

example, using data from six countries[4], Kleven et al. (2019b) find that countries that feature larger child penalties are also characterized by much more gender-conservative views.[5] Similarly, Berniell et al. (2021) document that the more conservative views the country holds, the larger the negative effects of motherhood on employment using 15 non-Eastern European countries from SHARE.[6] At the same time, Andresen and Nix (2019) examine child penalties among heterosexual, adopting, and same-sex couples in Norway, with the objective to disentangle the potential sources of the child penalty. The authors find that the child penalties vanish after two years of the first child among same-sex couples but not among heterosexual and adopting couples. After ruling out that child penalties are not driven by costs associated with giving birth, time spent with the child, or relative skill differences across spouses, the authors conclude that child penalties are likely driven by gender norms and gender differences in preferences for childcare.[7] Our analysis combines a battery of approaches that, taken together, suggest a causal association between gender norms and the motherhood employment penalty, even after we account for the endogeneity of the motherhood decision.

We contribute to the second strand of the literature by focusing on the extent to which motherhood is a mediating factor between female employment and culture. In addition to measuring new causal estimates of the effect of gender norms on women's decision to work, we find that, in European countries, the effect of culture passes through mostly the motherhood channel. A recent paper explores a similar question by exploiting migration across the East and West Germany border. Boelmann, Raute and Schönnberg (2020) document a strong asymmetry in the persistence of the culture in which women were raised. They document that the motherhood penalty in terms of female labor supply is smaller among East German mothers relative to West German mothers. Our work complements this other study by using an alternative identification strategy and adding external validity to European regions across 29 countries.

[4] The six countries are Austria, Denmark, Germany, the UK, the US, and Sweden.
[5] This paper main objective is to measure child penalties in female and male earnings in different countries using event studies around the birth of the first child. The study finds striking similarities in the qualitative effects of children across countries, but also sharp differences in the magnitude of the effects.
[6] The authors find that this strong correlation vanishes when considering only ex-Communist countries. Their paper focuses on the effects of motherhood on the allocation of talent in the labor market.
[7] Cortés and Pan (2020) arrive to a similar conclusion for the United States after observing that the motherhood earnings penalties persist for women who are considerably more skilled than their husbands.

The remainder of this paper is organized as follows. Section II discusses the data and our definition of non-traditional gender norms. Section III presents the identification strategy. Section IV present our baseline findings. Section V presents estimates controlling for the endogeneity of motherhood before concluding in Section VI.

II. Data and Measuring Non-Traditional Gender Norms

We use individual-level data from the repeated cross-sectional *European Social Survey* (hereafter ESS), a multi-country survey conducted every other year since 2002, whose main objective is to "assemble data on Europe's social condition, including the shifting beliefs, values, perceptions and behavior patterns among citizens in different countries."[8] The ESS ensures that high quality data are collected in each participating country and that they are comparable across participating countries. In addition, ESS data are representative of both the participating countries and the concepts under study. For this purpose, in each participating country, a representative sample of all persons aged 15 and over living in private households are selected using random probability methods.

The survey collects detailed information on respondents' socio-demographic characteristics, including their age, education level, number and age of children, household income, and NUTS2 region of residence. It also collects information on respondents' labor market characteristics such as their employment status or previous unemployment experience. The ESS does not have any information on individuals' wages or earnings. While the ESS gathers detailed information on respondents' parental background, such as parental education, employment status, and occupation when the respondent was 14 years old, it does not link parents to children, so even if two generations were to be surveyed by chance in a particular year, we would not be able to match them. Finally, the data include detailed information on respondents' preferences and beliefs, including beliefs on work and family values.

Measuring Non-Traditional Gender Norms

To measure beliefs and values that are likely to be relevant for mothers' decision to work, we follow Fortin (2005) and Rodríguez-Planas and Tanaka (2021) and focus on women's belief

[8] See ESS Survey Specification webpage:
https://www.europeansocialsurvey.org/methodology/ess_methodology/survey_specifications.html

towards the priority of male employment when jobs are scarce. Because beliefs on women's role in the workplace shape women's economic decisions, but such decisions also shape beliefs, using contemporaneous beliefs of the economic agents for whom we study their employment choices raises endogeneity concerns that would lead to overestimating the effect of beliefs on mothers' decision to work. To address such endogeneity issues, we use instead contemporaneous beliefs of the cohort of women in the same age range as the mothers of our economic agents (hereafter the "grandmothers"). In our baseline model, we assume that our gender norms variable (constructed with the beliefs of "grandmothers") is uncorrelated with contemporaneous labor market decisions of young mothers and affects young mothers' decision to work only through society's beliefs towards women working.[9] Later on, we will relax this assumption exploiting arguably exogenous variation in gender norms.

To measure non-traditional gender norms we construct, for each ESS wave, an EU-NUTS2 regional indicator taking the average extent of disagreement (on a 1 to 5 scale with higher values indicating higher disagreement) to the statement *"when jobs are scarce, men should have more right to a job than women*" among grandmothers residing in given EU-NUTS2 region. This indicator is constructed for the following two cohorts of "grandmothers": women 48 to 57 years old and women 58 to 68 years old as women in the same age range as the mothers of our economic agents aged 20 to 29, and 30 to 40, respectively. More precisely, we compute the non-traditional norm indicator as follows:

$$NTN_{crt}^{a} = \frac{\sum_{j=1}^{M_{crt}^{a}} Attitudes_{crt}^{a}}{M_{crt}^{a}}$$

where $Attitudes_{crt}^{a}$ reflects the extent of disagreement on the priority of male employment when jobs are scarce of women in cohort *a*, region *r* in country *c*, and survey wave *t*; M_{crt}^{a} is the total number of women in cohort *a*, region *r* in country *c*, and survey wave *t*. The superscript *a* indicates the age range of women *j* in the two cohorts, namely 48 to 57 and 58 to 68 years old. Our non-traditional norm indicator is derived as the cohort/region/wave average disagreement on the priority of male employment when jobs are scarce among "grandmothers" and is coded on a 1 to 5 scale such that increasing values denote a more progressive belief towards women.

[9] As explained earlier, as in the ESS data we do not observe the actual mother of our 20 to 40 year-old respondents, our grandmothers are not necessarily the mothers of the young mothers, hence, concerns on whether they are available to provide free childcare assistance is not an issue here. In the empirical analysis, we control for the fertility rate in the residence region, which proxies the local demand for childcare assistance.

While the ESS covers years 2002-2018 for 38 countries, we are constrained to use those waves that collect information on beliefs towards the priority of male employment when jobs are scarce. As this question is only available in waves 2004, 2008, 2010, and 2016, we lose five countries (Albania, Kosovo, Luxembourg, Montenegro, and Serbia) because they were not surveyed in any of those four waves. In addition, we lose Israel, Russia, Ukraine, and Turkey because we lacked regional controls. Our final sample is an unbalanced regional panel covering 29 European countries over the 2004 to 2016 period. In our sample, we have six Continental European countries (Germany, Austria, France, Netherlands, Belgium, and Switzerland), three Anglo-Saxon countries (United Kingdom, Ireland, and Iceland), four Nordic countries (Sweden, Denmark, Finland, and Norway); eleven Central-Eastern countries (Bulgaria, Croatia, Hungary, Czech Republic, Latvia, Lithuania, Romania, Slovakia, Poland, Slovenia, Estonia); and five Mediterranean countries (Italy, Spain, Portugal, Cyprus, and Greece).

Our non-traditional norm indicator is composed of 1,488 cells, covering two cohorts of grandmothers across 186 NUTS2 regions (in 29 countries), over 4 survey waves. Our identification strategy relies on estimating the effects on mothers' decision to work (relative to childless women) of the time varying components of grandmothers' non-traditional norms using the within-NUTS2 region and cohort variation of the country/wave effects, after controlling for time-invariant regional unobserved heterogeneity (including institutions and policies) with NUTS2 region fixed effects and for time-varying country unobserved heterogeneity (including institutions and policies) with country-and-year fixed effects. Hence, we must show that our identifying variable has enough time variation (and thus grandmothers' gender norms are not only a regional fixed effect) and regional variation (and thus it is not captured by a common year effect). Figures 1 and 2 and Appendix Table A.1 show that our measure of non-traditional gender norms varies both over time, across NUTS2 regions, and across cohorts.

Figure 1 displays the evolution of the non-traditional norms variable over the 2004-2016 period for the two cohorts. We clearly observe variation over time with norms becoming more progressive over the 12-year period. In addition, the prevailing norms for the older cohort are, on average, more conservative than those of the younger one. Figure 2 displays the regional variation of the non-traditional norms variable by plotting the average NUTS2 regional variable over the 2004-2016 period for the younger cohort (on the horizontal axis) against the corresponding value for the older cohort (on the vertical axis). The data points are classified into the following five

groups based on geographical location: Continental Europe, Anglo-Saxon countries, Nordic countries, Central-Eastern Europe, and Mediterranean countries. The figure shows considerable dispersion especially among Mediterranean regions, but also among the more moderate regions of Continental and Central-Eastern Europe. In contrast, both Nordic and Anglo-Saxon regions tend to be more progressive and display less variation. As in Figure 1, Figure 2 also shows that the majority of the data points lie below the 45-degree line, visually confirming that, at the regional level, the younger cohorts of grandmothers are more progressive than the older ones.

Appendix Table A.1 shows the amount of variation that is left in our non-traditional gender norms variable of interest after controlling for region, country-by-year, and cohort fixed effects. The first row in Appendix Table A.1 shows descriptive statistics of our main explanatory variable. The average disagreement on the priority of male employment when jobs are scarce among the grandmothers' sample is 3.713. The standard deviation of this gender norms variable is 63 percent. The second and third rows shows residuals after removing birth cohort, year and NUTS2 fixed effects (for row 2) and birth cohort, country-by-year and NUTS2 fixed effects (for row 3). Doing so reduces the standard deviation of the gender norms variable to 27 percent and 18 percent, respectively.

Sample Restrictions

As our analysis focuses on employment decisions of mothers (fathers) of young children relative to their childless women (men) counterparts, we restrict our analysis to native men and women between 20 and 40 years old as those are the most common childrearing years. We define as native someone who is born in the country of residence, and whose mother is born in the same country. We adopt this definition because our non-traditional gender norm variable is estimated at the cohort/region/wave level using the contemporaneous beliefs of the cohort of grandmothers living in the same NUTS2 region. The presumption here is that gender norms are acquired from women during childhood from their mother or their mothers' peers (Fortin 2005; Farré and Vella 2013; Rodríguez-Planas, Sanz-de-Galdeano, and Terskaya 2018), and hence including immigrants would increase measurement error as it is less likely that they would have been influenced by the older women living in their region of residence.

As we aim at understanding the relevance of non-traditional norms in mothers' decision to work, we further restrict our analysis to mothers (and fathers) whose first child is 0 to 5 years old.

As we assume the transmission of the norm takes place through the maternal channel, we also exclude individuals whose mother was dead or absent at the age of 14 years old. In addition to merging ESS individual-level data with our non-traditional gender norms variable at the cohort/NUTS2 region/wave level, we also merge them with other aggregate variables from *Eurostat* (such as the fertility rate, total population, the unemployment rate, the share of the population with tertiary education) at the NUTS2 by year level (see on-line Data Appendix for details on the precise definition of the main control variables). We exclude regions-by-year cells with ten observations or less. Having more than 10 observations within a cell is not uncommon in papers using a similar methodology in a different context (see, for example, Bifulco, Fletcher and Ross 2011; Olivetti, Patacchini and Zenou 2018; Rodríguez-Planas and Tanaka 2021). We also exclude all observations for which there are missing individual or regional controls to have a stable sample across various specifications.

Thus, our final sample includes 30,474 individuals aged 20 to 40 years old —of which 14,476 are females, 4,163 are mothers of young children, and 3,103 are fathers of young children— living in 29 countries during the period 2004-2016. In each regional-by-year cell there are on average 54 individuals (25 women), ranging between 11 and 456 individuals (the median being 36 individuals).

Descriptive Statistics

Appendix Table A.2 presents descriptive statistics for our sample of mothers and childless women (Panel A), and fathers and childless men (Panel B). On average, 54 percent of mothers in our sample are employed, a considerably lower share than childless women (two thirds work), childless men (70 percent work) and fathers (91 percent work). On average, mothers (fathers) in our sample are three (or four) years older and more likely to be married than childless women (childless men). While both parents are less likely to have secondary education than childless adults, they are more likely to have tertiary education than them. In contrast, mothers have longer past unemployment experiences than childless women, but the opposite is true for fathers and childless men. On average, non-traditional gender norms of mothers are 14.7 percentage points (or 3.6 percent) more conservative than childless women; whereas fathers' non-traditional norms are 10 percentage points (or 2.6 percent) more progressive than that of childless men. In the robustness

section, we show that the association between grandmothers' cohort non-traditional norms and the first-child employment gap holds even after controlling for the respondent's own gender norms.

Figure 3 reflects the relationship between first-child employment penalties (the motherhood penalty) and our non-traditional gender norms based on the beliefs of the grandmothers' cohort. The figure plots on the y-axis the average differences in employment between mothers and childless women (panel A) and between mothers and fathers (panel B) against the average extent of disagreement with the statement above among the "grandmothers" in 184 NUTS2 regions (or non-traditional gender norm, NTN), on the x-axis. It shows that NUTS2 regions with smaller negative employment differences (that is, smaller motherhood penalties) are also those with less traditional views (that is, higher NTN). Correlations are sizeable and always statistically significant at conventional levels. While not necessarily causal, this evidence underscores the association between gender norms and maternal employment.

III. Estimation Strategy

In this section, we first describe the identification strategy of our baseline model, which exploits variation across birth-cohorts and survey year within NUTS2 regions to obtain an estimate of non-traditional gender norms on the decision to work of young mothers relative to childless women; and we investigate the validity of our identification strategy. To address concerns that our social norms estimates may be endogenous, we exploit variation in reproductive health laws at the country level when the grandmothers were in their 20s. Lastly, we will exploit variation in reproductive health laws at the country level when the mothers were in their 20s to account for the fact that the decision to have a child is endogenous.

3.1 Baseline Model: The Motherhood Gap Within Estimate

Identification Strategy

Pooling mothers and young women in our sample, we estimate the following motherhood penalty gap model within cohort, region and country-by-year fixed effects model:

$$Y_{icrt}^{a-28} = \alpha_0 + \alpha_1 NTN_{crt}^{a} + \alpha_2 Small\ Child_{icrt}^{a-28} + \alpha_3 (NTN_{crt}^{a} * Small\ Child_{icrt}^{a-28}) + X_{icrt}^{a-28\prime}\gamma + Z'_{crt}\delta + \lambda_a + \varphi_r + \phi_{ct} + \varepsilon_{icrt}^{a-28} \quad (1)$$

where $Y_{icrt}^{a-28}$is equal 1 if individual *i* from birth cohort *(a-28)*, living in NUTS2 region *r* of country *c*, is working in survey year *y*, zero if the individual is not working. NTN_{crt}^{a} is the average disagreement on the priority of male employment when jobs are scarce of women in cohort *a*, region *r* of country *c*, and survey year *t*. $Small\ Child_{icrt}^{a-28}$is equal 1 if individual *i* from birth cohort *(a-28)*, living in NUTS2 region *r* of country *c*, has a child 0 to 5 years old in survey year *y*, zero if she is childless. $X_{icrt}^{a-28\prime}$ is a vector of individual-specific covariates measured at survey year *t*, namely age, age squared, two education dummies[10], two dummies for marital status[11], a dummy for being in a low-income household[12], one dummy for having experienced unemployment in the last five years, and two dummies for parental employment (one for the father and one for the mother) when the respondent was 14 years old. $Z'_{crt}$is a set of regional/year covariates measured contemporaneously such as the unemployment rate, the fertility rate, the total population, the share of tertiary educated individuals in the population.[13] We also include a dummy for the younger cohort of respondents, aged 20-30 y.o. (λ_a), NUTS2 region fixed effects (φ_r), as well as country-by-year fixed effects (ϕ_{ct}). Standard errors are clustered at the NUTS2 regional level.

Our estimate of interest, $\hat{\alpha}_3$, captures the association between non-traditional norms of the grandmothers' cohorts and mothers' decision to work relative to childless women, also known as the motherhood employment gap, conditional on the included covariates. A positive and statistically significant $\hat{\alpha}_3$ would indicate that mothers living in NUTS2 regions with more progressive grandmothers are more likely to work than those living in regions with less progressive grandmothers relative to childless women living in regions with more versus less progressive grandmothers. Hence, a positive and statistically significant $\hat{\alpha}_3$ would indicate that more

[10] Dummies for secondary and tertiary education, with primary education as the reference group.
[11] Dummies for never married and no longer married, with currently married as the reference group.
[12] Dummy for being in a household with an income below 15,000 euros per year (households with income higher than 15,000 euros per year are the reference group).
[13] In addition to these variables, we also considered many other controls such as the female-to-male differences in the unemployment rate and in the share of tertiary educated individuals, the median age in the region, GDP per capita, and an own computation of female-to-male Bartik wage ratios based on Eurostat SES, EU-SILC, or LFS Data (see on-line Data Appendix for details). We do not include these indicators in the baseline set of controls as they do not turn out as statistically significant, and/or entail a significant loss of NUTS2 regions. We also considered the number of births, as an alternative to the fertility rate (as the two are strongly correlated). All these results are available to the authors upon request.

progressive gender norms among grandmothers' cohort is associated with a smaller motherhood employment gap. In addition, to control for a rich set of individual characteristics known to be associated with employment (including individuals' human capital investment, prior unemployment history, household income and parental background), as well as a rich set of time-varying regional characteristics that may influence the outcome of interest, model (1) is able to control for time-varying unobserved heterogeneity (including institutions and policies) that affects employment decisions at the NUTS2 regional level of both mothers and childless women in the same way. At the same time, the country/year fixed effects control for cross-country differences in employment levels and business cycle that could potentially affect female employment generally.

As a falsification test, we estimate model (1) using fathers of small children and childless men. Importantly, finding no evidence of association between the gender norms in the grandmothers' cohort and the employment gap between fathers and childless men would suggest that beliefs towards the appropriateness of female employment among the grandmothers' cohort mediates only via motherhood, not fatherhood, and hence are not due to systematic differences between parents and non-parents within a region.

Identification comes from variation in grandmothers' gender norms across birth-cohorts over time within a NUTS2 region. This variation can be considered as quasi-random if the following two conditions are met. First, being in one cohort or another is beyond one's control as one does not choose the year (cohort) one is born into. Second, the difference in unobserved heterogeneity across cohorts within a NUTS2 region for a given survey year is not driven by unobserved factors that may also influence a mother's decision to work in a wage and salary job.

Validity of the Identification Strategy

In Table 1, we present balancing tests for our non-traditional gender norms variable separately for mothers, childless women, fathers, and childless men. They show the coefficient on the non-traditional norms of the grandmothers' cohort variable in a regression of individuals' socio-demographic characteristics or regional characteristics on the grandmothers' cohort non-traditional gender norms variable, a cohort dummy, NUTS2 regions fixed effects, and country/wave fixed effects. Focusing on our subgroup of interest, mothers, column 1 in Table 1 reveals that one of fourteen coefficients are statistically significantly different from zero at the

10% level, and none are statistically significantly different at the 5% level. In both cases, this is less than what we would expect by chance. These tests reveal that controlling for birth-cohort, NUTS2 region and country/wave fixed effects is likely to be sufficient to isolate variation in our grandmothers' gender norms variable that is not systematically related to mothers' socio-demographic composition within NUTS2 regions. Nonetheless, we will add many individual and time-varying regional controls. The balancing tests show that there are only two of fourteen systematic relationships (statistically significant at the 10% level) between our grandmothers' gender norms variable and the socio-demographic composition of childless young women within NUTS2 regions. The balancing tests for fathers or young childless men reveal that only one of fourteen coefficients significantly different from zero for fathers and none for childless men.

Despite the individual and regional controls included in model (1) and the fact that our gender norms variable is measured using the grandmothers' cohort (as opposed to the cohort of the individual for which we are studying employment decisions), concern that $E(NTN_{crt}^{a}\ \varepsilon_{icrt}^{a-28}) \neq 0$ may still remain. One concern is that there is reverse causality as preferences of the grandmothers' cohort change when observing the employment and motherhood decisions of their daughters' or their daughters' friends/peers. Alternatively, endogeneity of our main explanatory variable may be driven by omitted region-level or individual factors that may be associated with both the non-traditional gender norms of the grandmothers' cohort and the maternal employment of their daughters' cohort, for reasons unrelated to gender norms. To address such endogeneity concern, we take the following approach.

3.2 Exploiting Variation in Reproductive Health Laws to Address Endogenous Non-Traditional Gender Norms

To address the potential endogeneity of the grandmothers' cohort non-traditional norms variable, NTN_{crt}^{a}, we exploit institutional variation in countries' reproductive health laws when the grandmothers' cohort was 20 years old. The rationale for exploiting such variation lies in Goldin and Katz (2002)'s argument that reproductive-health liberalization laws (such as the contraceptive pill) raised women's awareness, thus favoring the development of less traditional gender norms. This variation is as good as exogenous as it is unlikely that each individual grandmother, *per se*, influenced the legalization of the pill or the abortion in her country. Our cohorts of grandmothers

were aged 20 years old between 1956 and 1988. Crucially, regularization laws that took place during this "exposure window" are random with respect to the counterfactual employment decision to work of younger women (born after these liberalizations took place) in the period 2004 to 2016 and living in regions with more or less progressive gender norms.

Finlay, Canning, and Po (2012)[14] developed an original dataset on Reproductive Health Laws Around the World for the period 1960-2009, which we have updated to describe the environment our grandmothers were exposed to during the period 1956 to 2016.[15] We call $Liberalization_{ct}$ the degree of liberalization of contraceptive pill and abortion laws[16] in country *c* at time *t* available for our sample of countries. We assign to each grandmother *j,* (i.e. to each woman between 48-68 years old) the degree of liberalization she was exposed to in her country *c* at the age of 20 (that is, at $t - age + 20$). This is the degree of liberalization observed in each country between 1956 and 1988. After that, we construct an index of average grandmothers' exposure to reproductive health laws at the region *r* by cohort level *a* in country *c* at survey wave *t*:

$$GrandExposure20^{a}_{crt} = \frac{\sum_{j=1}^{M^{a}_{crt}} Liberalization_{cj(t-age+20)}}{M^{a}_{crt}}$$

where M^{a}_{crt} is the total number of grandmothers from cohort *a=48/58, 58/68* in region *r* of country *c* in survey wave *t*. $GrandExposure20^{a}_{crt}$, constructed in this way measures the degree of liberalization of reproductive health laws (contraceptive pill and abortion) grandmothers from cohort *a* resident in region *r* of country *c* at time *t* were exposed to when they were 20 years old (i.e. at time *t-age+20)*. This indicator ranges between zero (indicating grandmothers were exposed to a full prohibition of contraceptive pill and abortion) and one (indicating grandmothers were exposed to full liberalization of abortion, as well as a full commercialization of contraceptive pill).

[14] The working paper is available here: https://cdn1.sph.harvard.edu/wp-content/uploads/sites/1288/2013/10/PGDA_WP_96.pdf.
[15] We made three major updates of data from Finlay, Canning, and Po (2012). First, we extended the degree of liberalization observed in 1960 back to 1956, after verifying that no country had introduced relevant reforms according to the taxonomy proposed by Finlay, Canning, and Po (2012). Second, we accounted for abortion liberalizations during the 1970s in former GDR/DDR (to which German grandmothers were exposed to), and former Yugoslavia (to which Croatian and Slovenian grandmothers were exposed to). Third, we accounted for further contraceptive pill liberalizations during 2009-2016 in Portugal, Slovenia, Slovakia, and Czech Republic.
[16] A third dimension of reproductive health legislation is available in the data, which concerns the legalization and commercialization of Intrauterine devices (IUD). We decide not to consider it as it proves largely irrelevant in the present analysis. Further details are available to the authors upon request.

The time and regional variation of exposure to reproductive-health laws of grandmothers' cohort when they were young is shown in Figure 4. For each country, we show the regional variation in exposure overtime for each of the two cohorts. Each dot represents the average degree of exposure to health liberalization laws for a given cohort of grandmothers in region *r* (red squares refer to *a=58-68 years old*, while blue triangles refer to *a=48-58 years old*). The Figure 4 highlights the two main sources of variation of the $GrandExposure20^{a}_{crt}$ indicator. The first one is variation between the young and the old cohort of grandmothers observed in the same year. In general, as reproductive-health liberalization reforms in most countries occurred during the 1970s, the younger cohort of grandmothers (that is, those aged 48 to 57 years old) at the age of 20 were exposed to a higher degree of liberalization in the use of pill and abortion, compared to the older cohort (aged 58- to 68-year). Second, besides variation across cohorts, the $GrandExposure20^{a}_{crt}$ indicator also displays variation overtime as we move from wave 2004 to wave 2016. The reason for this is that, as an example, a 50-year-old grandmother (belonging to the cohort 48-58 years old) will be exposed to different degrees of liberalization in survey wave 2004 and 2016 because they will be assigned the degrees of liberalization of 1974 (the former) and 1986 (the latter). The $GrandExposure20^{a}_{crt}$ indicator reflects this source of overtime variation within each cohort during the 2004-2016 period too. In most countries, younger grandmothers were exposed to more liberal environment in most countries, so that $GrandExposure20^{a}_{crt}$ is generally monotonically increasing over the 2004-2016 period. We find some important exceptions in countries from Central and Eastern Europe (CEE) such as Bulgaria, Hungary, Romania, and the Baltic countries, which experienced temporary reversals in the degree of reproductive-health liberalizations during the 1970s.[17] This induces a non-monotonic trend in the exposure to liberalizations, which is particularly pronounced in the case of Estonia. Notice that in most countries, $GrandExposure20^{a}_{crt}$ also displays some variation across regions within a country-cohort-year cell due to the change in grandmothers' age composition across regions for a given year.

Our assumption is that the degree of exposure to reproductive health liberalization laws in the country when the grandmother was 20 years old is correlated with her beliefs on women's role

[17] In Bulgaria abortion was prohibited in the mid-70s and was gradually re-introduced in the following years. In Hungary, abortion requests were initially legal, but prohibited in 1973 to be legalized again in 1993. In Romania, both the use of contraceptive pill and abortion was restrained in the 1960s before being reintroduced again in the late 1980s. Finally, in the Baltic countries the contraceptive pill was also legal back to 1956, and then prohibited between 1974 and 1982. See Finlay, Canning, and Po (2012) for details. These patterns can be seen in Figure A.2 in the Appendix, which displays the sample variation of the original variation in Finlay, Canning, and Po (2012) by country.

in the workplace (namely, the gender norm transmitted to the next generation), but it only impacts the employment decisions of the next generation of mothers through such beliefs, once we control for time-varying controls, and fixed effects.

This conjecture leads us to estimate the following auxiliary regression stage at the region-cohort-year level:

$$NTN^{a}_{crt} = \gamma_0 + \gamma_1 GrandExposure20^{a}_{crt} + Z'_{crt}\delta + \varphi_r + \phi_{ct} + \varepsilon^{a}_{crt} \quad (2)$$

where all the variables and vectors have been previously defined. Panel A in Appendix Table A.3 shows results from three different specifications of these auxiliary regression. In column 1, $GrandExposure20^{a}_{crt}$ is computed on the basis of both abortion and pill liberalization laws; in column 2, $GrandExposure20^{a}_{crt}$ only uses abortion liberalization laws; and in column 3, $GrandExposure20^{a}_{crt}$ is entered separately in the equation based on either pill or abortion liberalization laws. Estimates in Panel A, Appendix Table A.3. show that greater exposure to reproductive health laws when grandmothers were in their 20s is associated with more progressive gender norms among the grandmothers' cohorts about three to four decades later. In columns 1 and 2, the coefficient $\hat{\gamma}_1$ is statistically significant at the 1% level. In column 3, the coefficient on exposure to abortion liberalization is statistically significant at the 1% level, and that on exposure to pill liberalization is not statistically significant. The adjusted R-squared in columns [1]-[3] are always above 0.63, which indicates the auxiliary model (2) has a very good explanatory power. Moreover, the inclusion of the $Exposure^{a}_{crt}$ indicator(s) increases the adjusted R-squared of the auxiliary model by up to about 2% (from 0.625 to 0.635), implying that pill and abortion liberalizations have a non-negligible contribution to both regional and time variation of the non-traditional gender norms variable.

Using the auxiliary regression, we estimate a predicted norm $\widehat{NTN}^{a}_{crt}$, which we use as the main regressor in the following model:

$$Y^{a-28}_{icrt} = \alpha_0 + \alpha_1 \widehat{NTN}^{a}_{crt} + \alpha_2 Small\ Child^{a-28}_{icrt} + \alpha_3 (\widehat{NTN}^{a}_{crt} * Small\ Child^{a-28}_{icrt}) + X^{a-28\prime}_{icrt}\gamma + Z'_{crt}\delta + \lambda_a + \varphi_r + \phi_{ct} + \varepsilon^{a-28}_{icrt} \quad (3)$$

Appendix Table A.4 shows descriptive statistics of $\widehat{NTN}^{a}_{crt}$ and compares it with the observed NTN^{a}_{crt}. The appendix table shows that both observed and predicted norms have very similar

distributions, means and standard deviations. In equation (3) above, time variation within the same region is guaranteed by the change in the age structure of the population of grandmothers between 2004 and 2016 (that is, the exact age composition of grandmothers in the group M_{crt}^{a}). This allows the inclusion of region fixed effects. In this case, $\hat{\alpha}_1$ and $\hat{\alpha}_3$ are identified by the variation in the age composition of grandmothers within each region over time.

IV. Main Findings

Column 1 in Panel A in Table 2 shows the baseline model: the motherhood gap within cohort, region, and country-by-year fixed effects. Columns 2 and 3 present results from estimating model (3) in which the potentially endogenous non-traditional gender norms variable has been replaced with its predicted measure, $\widehat{NTN}_{crt}^{a}$, estimated from the first-stage equation (2). The difference in the specification used in columns 2 and 3 is that the former exploits variation in exposure to both the pill and the abortion liberalization laws, whereas the latter only exploits variation in exposure to abortion liberalization laws. In all three specifications, $\hat{\alpha}_3$is positive and statistically significant at the 5% level suggesting that living in regions with more progressive gender norms among the grandmothers' cohort is associated with a higher likelihood of maternal employment relative to childless women relative to the motherhood gap in regions with more traditional gender norms among the grandmothers' cohort. Moving from the baseline model (column 1) to the ones that address endogeneity (columns 2 and 3) increases the size of $\hat{\alpha}_3$from 5.3 percentage points to 5.5 and 5.7 percentage points. Columns 4 and 5 in Panel A in Table 2 address concerns that the exclusion restriction may be violated if our respondents were born before the legalization of the reproductive health laws. To do so, we re-estimate the model used in column 2 but only including in the sample respondents who were born after such legalizations took place (abortion only in Column [4], and both abortion and pill in Column [5]). Moving from column 2 to column 4 or column 3 to column 5 shows a stronger and more precise relationship between predicted gender norms among the grandmothers' cohort and the motherhood employment gap. Column 6 presents our preferred specification, as we re-estimate the specification in column 4, excluding respondents from countries in which grandmothers were not affected by any changes in the reproductive health laws when they were in their 20s. Doing so has a minimal impact on the size, the sign, and the precision of $\hat{\alpha}_3$. This reassures us that our results are not driven by countries that do not display any variation in the instrument.

Estimates in column 6 in Panel A in Table 2 show that non-traditional norms of the grandmothers' cohort are directly related to maternal employment. More specifically, we find that an increase in the predicted average disagreement on the priority of male employment when jobs are scarce of women in the grandmothers' cohort by one standard deviation is associated with an increase in the likelihood that mothers with small children work of 5.6 percentage points[18], the equivalent of a reduction of 46.7 percent in the average motherhood employment gap in our sample of 12 percentage points.[19] To put it differently, if mothers from Andalusia (ES23) in Spain, with an average disagreement to traditional gender norms among grandmothers of 2.36 and a motherhood employment gap of 9 percentage points in 2004, were characterized by the mean gender norms in Sweden (e.g. SE21, Småland med öarna), which is equal to 3.77, the statistical model would suggest that the motherhood employment gap in Andalusia would decrease by 12 percentage points, which would fully reverse the observed motherhood employment gap between Småland med öarna and Andalusia (equal to 9 percentage points).[20]

Our preferred specification (in column 6, Panel A in Table 2) also reveals a motherhood employment gap of 71 percent (or 46.2 percentage points) of the average employment among childless women in our sample (equal to 65.5 percent as shown in Table A-2). In addition, we also find that gender norms of the grandmothers' cohort are also associated with an increase in the employment rate of childless women, albeit the coefficient is only statistically significant at the 10 percent level and is not robust to accounting for the endogeneity of motherhood as discussed in Section V.

Falsification Test

Panel B in Table 2 replicates the above exercise but using fathers of young children and childless men. Focusing first on the differential impact of gender norms of the grandmothers' cohort on the fatherhood employment gap, we observe that there is no effect as $\hat{\alpha}_3$ is mostly zero. Evidence of no association between the gender norms in the grandmothers' cohort and the employment gap

[18] This is calculated as: $\hat{\alpha}_3 * \widehat{NTN}_{StDev} = 0.087 * 0.644 = 0.056$

[19] This is calculated as $\frac{\hat{\alpha}_3 * \widehat{NTN}_{StDev}}{Motherhood\ Gap_{Mean}} = \frac{0.087 * 0.644}{-0.116} = \frac{0.056}{-0.116} = -0.467$

[20] (Motherhood_Gap*(ES23)* − Motherhood_Gap*(SE21)*)* $\hat{\alpha}_3$ = (3.77 -2.36)*0.087 = 1.41*0.087=0.123.

between fathers and childless men suggests that beliefs towards the appropriateness of female employment among the grandmothers' cohort mediates only via motherhood, not fatherhood, and hence our findings on the motherhood employment gap are not due to systematic differences between parents and non-parents within a region.

As others have found, there is no penalty for having a first child among men as $\hat{\alpha}_2$ is small and not statistically significant (with the exception of a positive and marginally significant effect in column 5). Interestingly, there is a positive association between childless men and non-traditional gender norms of the grandmothers' cohort ($\hat{\alpha}_1$) of similar size to that found among childless women although measured with less precision in columns (1)-(5). In our preferred specification, we observe that an increase in the predicted average disagreement on the priority of male employment when jobs are scarce of women in the grandmothers' cohort by one standard deviation is associated with an increase in the likelihood of employment *among childless men* of 12.5 percentage points[21], the equivalent of an increase of about 18 percent in the average employment rate of childless men in our sample of 71 percent.[22]. This is not the first paper to find that progressive gender norms impact both men and women. For instance, Rodríguez-Planas and Nollenberger (2018) find that "*youth whose parents come from more gender-equal societies perform better in exams regardless of gender or subject type*". Rodríguez-Planas, Sanz-de-Galdeano, and Terskaya (2018) also find that "*a greater proportion of peers with non-traditional mothers during high school curbs males' risky behaviors*" during high school and in their early 20s. In addition, we find that progressive gender norms seem to mitigate the child employment penalty, inexistent among fathers.

Robustness Checks

Using our preferred specification, Table 3 presents a battery of robustness checks for women 20- to 40-year old (column 1) and for men in the same age range (column 2). Panel A adds as additional covariates the own respondent's disagreement on the priority of male employment when jobs are scarce, and the gender norms of the respondent's cohort as additional covariates in the preferred

[21] This is calculated as: $\hat{\alpha}_1 * \widehat{NTN}_{StDev} = 0.195 * \ 0.644 = 0.1256$

[22] This is calculated as $\frac{\hat{\alpha}_1 * \widehat{NTN}_{StDev}}{Employment_{Mean}} = \frac{0.195 * 0.655}{0.708} = \frac{0.1256}{0.708} = 0.1773$

specification. We do this to explore the extent to which the gender norms of grandmothers' cohort may be picking up the respondents' own beliefs. While own cohort's gender norms are relevant in women's decision to work, the relevance of the gender norms of the grandmothers' cohort remains large, positive and statistically significant at 7.7 percentage points. Concerns that our main findings are driven by respondents' religiosity are ruled out in Panel B, which presents results from a specification that controls for respondent's religiosity.

Panels C and D replace our left-hand-side variable (being employed or not) with two alternative outcomes: being in the labor force or not[23], and hours worked, respectively. In both cases, $\hat{\alpha}_3$ remains positive and statistically significant at the 1% level. Interestingly, the finding that the gender norms of the grandmothers' cohort is negatively associated with the motherhood employment gap holds at the intensive margin. More specifically, we find that an increase in the predicted average disagreement on the priority of male employment when jobs are scarce of women in the grandmothers' cohort by one standard deviation is associated with a reduction in the motherhood *hours worked* gap of 20 percentage points[24], the equivalent of a reduction of 44 percent in the average motherhood hours worked gap in our sample.[25]

Panel E and F re-estimates our preferred specification without regional (Panel E) and parental controls (Panel F). Doing so has little effect on the size and precision of our coefficient of interest $\hat{\alpha}_3$. In Panel G, we account for country-specific time varying factors that may affect differently parents and non-parents by replacing the country-by-year fixed effects with country specific controls for the total number of weeks of maternity leave, and the paid weeks of maternity leave legislated in each country at time *t*. Doing so leaves results basically unchanged, with a negative (non-statistically significant) association between the number of paid or unpaid maternity leave weeks and female employment.

[23] This is considered as being either employed or unemployed as opposed to being in education or housework as main economic status.

[24] This is calculated as: $\hat{\alpha}_3 * \widehat{NTN}_{StDev} = 0.321 * 0.644 = 0.206$

[25] This is calculated as $\frac{\hat{\alpha}_3 * \widehat{NTN}_{StDev}}{Motherhood\ Hours\ Worked\ Gap_{Mean}} = \frac{0.321 * 0.644}{-0.469} = \frac{0.206}{-0.469} = -0.441$

In Panel H, we re-estimate the model bootstrapping standard errors (200 replications), which accounts for the fact that our main regressor comes from a predicted variable ($\widehat{NTN}^{a}_{crt}$). Again, our results are robust to this sensitivity analysis.

In Panel I, we present estimates from a two-stage least squares (2SLS) model using ($\widehat{NTN}^{a}_{crt} * Small\ Child^{a-28}_{icrt}$) as an instrument for ($NTN^{a}_{crt} * Small\ Child^{a-28}_{icrt}$). As discussed by Wooldridge (2002), this 2SLS estimator shares better asymptotic properties compared with the OLS estimator of equation (3), as $\widehat{NTN}^{a}_{crt}$ enters the specification non linearly through the inclusion of its interaction term with $Small\ Child^{a-28}_{icrt}$ (Wooldridge, 2002, pp. 235).[26] These estimates confirm that non-traditional norms of the grandmothers' cohort are directly related to maternal employment. The coefficient of the interaction term is about ten times larger than the corresponding coefficient reported in Table 2 (column 6). We interpret this as *Local Average Treatment Effect* (LATE), in the presence of a compliance rate equal to 0.092, as measured by the first stage coefficient.[27] We obtain the corresponding *Intention-to-Treat* (ITT) effect as 0.084=0.092*0.921. This estimated ITT effect is fully consistent with estimates of equation (2) presented in Table 2 (Column 6). These are indeed reduced-form results of a regression of the outcome on the instrument, regardless of the compliance rate in the underlying population.

Panel L re-estimates our preferred specification once more with a probit model . Our results remain robust to this alternative functional form. The marginal effects reported in Table 3 suggest very similar magnitudes to those of the linear probability model. It is noteworthy that, for the sample of men, $\hat{\alpha}_3$ is zero and not statistically significant for all the robustness checks performed in Table 3 (shown in column 2).

A final concern in our estimates is measurement error due to internal migration of grandmothers between regions, possibly in a way that is not random relative to gender norms.[28]

---

[26] In practice, we use equation (1) as a 'stage 0', that predicts $\widehat{NTN}^{a}_{crt}$, and use ($\widehat{NTN}^{a}_{crt} * Small\ Child^{a-28}_{icrt}$) as an instrument in the first stage of the two-stage least squares IV model. This is implemented using *ivreghdfe* in stata, which is a version of ivreg2 which allows to absorb fixed effects (Correia 2018).

[27] In this setting, the compliers are those mothers, whose norms actually respond/correlate with the treatment they are assigned to i.e. the predicted norm $\widehat{NTN}^{a}_{crt}$. Accordingly, the compliance rate is indeed the coefficient of $\widehat{NTN}^{a}_{crt} * Small\ Child^{a-28}_{icrt}$ in the first stage (see Angrist and Pischke, 2009 p.164).

[28] We perform additional robustness exercises on the sample of mobile 20-40 y.o. women. In Figure A-1, we plot the average norm by current NUTS2 region of residence (on the y-axis) against the average norm in the region of origin (on the x-axis) for women with a young child (Panel A) and without children (Panel B). The two averages for each region are equally scattered around the 45 degrees line. In consecutive Panels C to F, we provide similar evidence for

To avoid possible bias, we should measure the non-traditional gender norms in the region of birth, which however it is not available in the ESS. To address such concern, we re-estimate the fixed-effects model using gender norms measured in the region of residence at age 14. For this exercise we use a different dataset, the 2008 wave of the European Value Study (EVS), which includes the region of birth. Appendix table A.5 replicates the analysis using both the region of residence at time of the survey (Columns 1 and 2) and when the respondent was 14 years old, (Columns 3 and 4). As our main finding remains unchanged with this alternative measure of non-traditional norms it is unlikely that internal migration is driving our results. Importantly, estimates in Appendix Table A.5 suggest that the effect of non-traditional norms is larger and estimated more precisely when measured in the region of origin, compared to the region of residence.[29]

Subgroup Analysis

Finally, Table 4 presents some heterogeneity analysis using our preferred specification. Columns (1) and (2) show that the salience of gender norms on the motherhood employment gap holds in regions with both high or low unemployment rate. Perhaps not surprisingly, the effect is stronger when the labor market is tight as there are fewer labor-demand constraints. Our main findings also hold regardless of women's educational level, as shown in columns (3) and (4). In contrast, women's parental education seems to matter as gender norms association with the motherhood employment gap is driven by those women with non-college educated parents, which represent two-thirds of the sample.

V. Addressing Endogeneity of Motherhood

A well-known fact in the motherhood gap literature is that the decision to have a child is endogenous and that "*individual unobserved heterogeneity may still prevail, as women deciding*

average female employment rates, and female-to-male Bartik wage ratios, respectively. This evidence is globally consistent with no systematic mobility pattern from more traditional regions, characterized by worse female labor outcomes, to less traditional ones, characterized by better work prospects for women.

[29] Notice that coefficients estimated on EVS data in Column (2) of Table A-5 are not immediately comparable to their counterparts estimated on the ESS data in column (1) of Table 2. In fact, estimates in Table A-5 are identified only by cross-regional variation in the EVS sample for 2008. Conversely, estimates in Table 2 are based on within-region variation available from the multiple waves of the European Social Survey. Accordingly, in estimates in Table A-5 we do not include region FEs, but only FEs for the 24 countries in the EVS 2008 sample (see columns 3 and 6). Standard errors are always clustered at the regional level.

to have children may have different tastes and preferences about work than childless female workers. If there are unobserved quality differences between mothers and childless female workers, results from (repeated) cross-sectional studies of the family gap will reflect an omitted variable bias" (Fernández-Kranz, Lacuesta, and Rodríguez-Planas 2013). To address this concern, we measure the exposure to reproductive health liberalization laws for young women (that is, 20- to 40-year-old women) when they were 20 years old using the same dataset on reproductive health laws from Finlay, Canning, and Po (2012) and updating when need be. We then use this individual measure of exposure to reproductive health laws when respondent was 20 years old to estimate the following auxiliary regression using the sample of all 20- to 40-year-old women in the seven ESS waves covering the 2002 to 2016 period:[30]

$$Small\ Child_{icrt} = \rho_0 + \rho_1 Liberalization_{icr(t-age+20)} + X_{icrt}{}'\gamma + Z'_{crt}\delta + \varphi_r + \phi_{ct} + \varepsilon_{icrt} \quad (4)$$

In equation (4), we exploit variation in exposure to liberalization laws across individuals within a country and for a given survey wave.

The coefficient $\hat{\rho}_1$ in equation (4) is identified if women in country *c* have been exposed to different degrees of liberalization based on their age. The variation we exploit to estimate equation (4) can be better understood by looking at the variation of the baseline $Liberalization_{ct}$ indicator by Finlay Canning and Po (shown in Figure A-2 in the Appendix).

Panel B in Appendix Table A.3 shows results from three different probit estimates of auxiliary regression (4). In column 1, $Liberalization_{icr(t-age+20)}$ uses both abortion and pill liberalization laws; in column 2, $Liberalization_{icr(t-age+20)}$ is entered separately in the equation based on either pill or abortion liberalization laws; and in column 3, $Liberalization_{icr(t-age+20)}$ only uses abortion liberalization laws. Estimates in Panel B, Appendix Table A.3. show that greater exposure to reproductive health laws when respondents were in their 20s is associated with lower likelihood of being a mother up to two decades later with coefficients always statistically significant at the 1% level. The results suggest that the $Liberalization_{icr(t-age+20)}$ indicator is

[30] As this auxiliary regression does not include the non-traditional norms variable, we are able to use all seven waves, increasing sample size and maximizing exogenous variation to identify the effect of the degree of reproductive health liberalization laws on the decision to have a child.

important to explain maternity decisions. Alone, it increases the pseudo R-squared of the auxiliary model by 13% to 20%, implying that individual exposure to pill and abortion liberalizations give a non-negligible contribution to actual motherhood decisions.

From the first stage in equation (4) above, we estimate the predicted probability of having a child, $\widehat{Small\ Child}_{icrt}^{a-28}$ for all 20- to 40-year-old females in our sample, and use such predicted estimator in the equation (5) below:

$$Y_{icrt}^{a-28} = a_0 + a_1\widehat{NTN}_{crt}^{a} + a_2\widehat{Small\ Child}_{icrt}^{a-28} + a_3(\widehat{NTN}_{crt}^{a} * \widehat{Small\ Child}_{iicrt}^{a-28}) + X_{icrt}^{a-28\prime}\gamma + Z'_{crt}\delta + \lambda_a + \varphi_r + \phi_{ct} + \varepsilon_{icrt}^{a-28} \quad (5)$$

Our estimate of interest, $\widehat{a_3}$, captures the association between predicted non-traditional norms of the grandmothers' cohorts and predicted motherhood probability relative to childless women (conditional on the included covariates). We argue that the actual realization of a motherhood outcome depends on the direct exposure of our female respondents to reproductive-health liberalization laws implemented in their country at they were 20 years old. Our assumption is that the country's degree of reproductive-health liberalization when the respondent was 20 years old is correlated with women's decision to become mothers, however, contrary to the actual decision to have a child, the country's degree of reproductive-health liberalization when respondent was 20 (i) is not (contemporaneously) affected by labor market outcomes and (ii) is likely to affect women's decision to work only through motherhood.

While $\widehat{\alpha_3}$ estimated from equation (3) measured the effect of the predicted non-traditional gender norm by motherhood status, $\widehat{a_3}$ estimated from equation (5) measures the effect of the predicted non-traditional gender norm by predicted exposure to motherhood based on liberalization laws when respondent was 20 years old. Notice also that identification of both $\widehat{NTN}_{crt}^{a}$ and $\widehat{Small\ Child}_{icrt}^{a-28}$ in equation (5) comes from exposure to reproductive health liberalization laws from equations (2) and (4). $\widehat{NTN}_{crt}^{a}$ is identified by changes in the degree of liberalization taking place between 1956 and 1988, while $\widehat{Small\ Child}_{icrt}^{a-28}$ is identified by changes in liberalization occurring between 1982 and 2016. This implies that any change in liberalization taking place in country *c* between 1982 and 1988 actually contributes to the identification of both indicators. Such an overlap is not an issue in terms of identification as there

are still reforms during 1956-1982 and 1988-2016, which allow to separately identify the effect of $\widehat{NTN}^{a}_{crt}$ and $\widehat{Small\ Child}^{a-28}_{icrt}$ in equation (5).

Table 5 re-estimates Table 2 but replacing observed maternity status with the predicted exposure to maternity, $\widehat{Small\ Child}^{a-28}_{icrt}$, estimated from the auxiliary regression (4). In Column 1 we report results using the observed NTN^{a}_{crt}. Columns 2 and 3 present results from estimating model (5) in which, in addition to addressing endogenous motherhood, the potentially endogenous non-traditional gender norms variable has been replaced with its predicted measure, $\widehat{NTN}^{a}_{crt}$, estimated from the auxiliary equation (2). The difference in the specification used in columns 2 and 3 is that the former exploits variation in exposure to both the pill and the abortion liberalization laws, whereas the latter only exploits variation in exposure to abortion liberalization laws. Column 4 in Panel A in Table 5 re-estimates the model used in column 2 but only including in the sample respondents who were born after such legalizations took place, and column 5 re-estimates the specification in column 4 excluding respondents from countries in which grandmothers were not affected by any changes in the reproductive health laws when they were in their 20s.

In five out of six specifications, $\widehat{a_3}$ is positive and statistically significant at the 10 percent level or lower consistent with earlier findings that living in regions with more progressive gender norms among the grandmothers' cohort is associated with a smaller motherhood employment gap. Crucially, when all six specifications are estimated using fathers and childless men (shown in Panel B in Table 5), $\widehat{a_3}$is zero and not statistically significant.

Based on our preferred specification, shown in column 6, Panel A in Table 5, a one-standard deviation increase in the predicted average disagreement on the priority of male employment when jobs are scarce of women in the grandmothers' cohort is associated with an increase in the maternal employment of 9.17 percentage points—note that the predicted probability of becoming mothers averages0.28 as shown in Table A-4 in the Appendix.[31] This is equivalent to a reduction of about 80 percent in the average motherhood employment gap in our sample of 12 percentage points.[32]

After addressing the endogeneity of motherhood, the effect of grandmothers' cohort non-traditional norms increases by 70 percent—as a one-standard-deviation increase in non-traditional

[31] This is calculated as: $a_3 * NonTraditionalBeliefs_{StDev} * \widehat{Small\ Child}_{StDev} = 0.515 * 0.646 * 0.28 = 0.0917$

[32] This is calculated as $\frac{a_3 * \widehat{NTN}_{StDev} * \widehat{Small\ Child}_{StDev}}{Motherhood\ Gap_{Mean}} = \frac{0.515 * 0.644 * 0.281}{-0.116} = \frac{0.093}{-0.116} = -0.803$

norms reduced the motherhood employment by 47 percent prior to correcting for the endogeneity of motherhood (column 6, Panel A in Table 2), and by 80 percent after correcting for motherhood endogeneity (column 6, Panel A in Table 5).

Once we correct for the endogeneity of motherhood, the association between grandmother's cohort gender norms and childless women employment (captured by $\widehat{a_1}$) is considerably smaller and no longer statistically significant. In our preferred specification, $\widehat{a_1}$ goes from 0.156 (significant at the 10 percent level) in column 6, Panel A in Table 2 to 0.026 (not statistically significant) in column 6, Panel A in Table 5. This suggest that non-traditional gender norms mediate on female employment mainly via motherhood, which brings a novel insight in this literature on culture and female employment.

VI. Conclusion

Using individual-level data from the European Social Survey, we study the relevance of gender norms in accounting for the motherhood employment gap across 186 European NUTS2 regions (over 29 countries) for the 2002-2016 period. The gender norm variable is taken from a question on whether "*men should have more right to a job than women when jobs are scarce"* and represents the average extent of disagreement (on a scale 1 to 5) of women belonging to the "grandmothers" cohort. We address the potential endogeneity of our gender norms measure with an index of the degree of reproductive health liberalization when grandmothers were 20 years old. We also account for the endogeneity of motherhood with the level of reproductive health liberalization when mothers were 20 years old. We find a robust positive association between progressive beliefs among the grandmothers' cohort and mothers' likelihood to work while having a small child (0 to 5 years old) relative to similar women without children. No similar association is found among men. Our analysis underscores the role of gender norms and maternal employment, suggesting that non-traditional gender norms mediate on the employment gender gap mainly via motherhood.

References

Anderson. Deborah J,, Melissa Binder, and Kate Krause. 2002, "The Motherhood Wage Penalty: Which Mothers Pay it and Why?" *American Economic Review Papers and Proceedings 92{3):354-5.*

Anderson. Deborah J,, Melissa Binder, and Kate Krause.. 2003, "The Motherhood Wage Penalty Revisited: Experience, Heterogeneity, Work Effort and Work-Schedule Flexibility," *Industrial and Labor Relations Review* 56 (2):273-94.

Andresen, Martin E. and Emily Nix, "What Causes the Child Penalty? Evidence from Same Sex Couples and Policy Reforms," Technical Report, Discussion Papers 902, Statistics Norway, Research Department 2019.

Angrist, J. D. and K. Lang. 2004. Does School Integration Generate Peer Effects? Evidence from Boston's Metco Program. *American Economic Review*, 945: 1613–34.

Angrist & Jörn-Steffen Pischke, 2009. "Mostly Harmless Econometrics: An Empiricist's Companion," Economics Books, Princeton University Press, edition 1, number 8769, April.

Nikolay Angelov, Per Johansson, and Erica Lindahl, "Parenthood and the Gender Gap in Pay," Journal of Labor Economics 34, no. 3 (July 2016): 545-579.

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

Berniell, I., Berniell, L., De la Mata, D., Edo, M. and Marchionni, M., 2021. Gender gaps in labor informality: The motherhood effect. *Journal of Development Economics*, *150*, p.102599.

Bertrand, M., Kamenica, E., & Pan, J. (2015). Gender identity and relative income within households. *The Quarterly Journal of Economics*, *130*(2), 571-614.

Bifulco, R., Fletcher, J. M., and S. L. Ross. 2011. The effect of classmate characteristics on post-secondary outcomes: Evidence from the Add Health. *American Economic Journal: Economic Policy*, 31: 25-53.

Blau, F. D., & Kahn, L. M. (2006). The US gender pay gap in the 1990s: Slowing convergence. *ILR Review, 60*(1), 45-66.

Blau, Francine D., Lawrence M. Kahn, Albert Yung-Hsu Liu, and Kerry L. Papps. 2013. "The Transmission of Women's Fertility, Human Capital, and Work Orientation across Immigrant Generations." *Journal of Population Economics,* 262: 405–35.

Boelmann, Barbara, Anna Raute, and Uta Schönberg, "Wind of Change? Cultural Determinants of Maternal Labor Supply," Working Paper 2020.

Correia, Sergio, 2018. "IVREGHDFE: Stata module for extended instrumental variable regressions with multiple levels of fixed effects," Statistical Software Components S458530, Boston College Department of Economics.

Cortés, Patricia and Jessica Pan. 2020. "Children and the Remaining Gender Gaps in the Labor Market" in preparation for the *Journal of Economic Literature.* NBER Working Paper #27980

Fernández, Raquel. 2007. "Women, Work and Culture." *Journal of the European Economic Association* 24(4): 329–30.

Fernández, Raquel, and Alessandra Fogli, and Claudia Olivetti. 2004. "Mothers and Sons: Preference Formation and Female Labor Force Dynamics." Quarterly Journal of Economics, 119 (4), 1249-99.

Fernández, Raquel, and Alessandra Fogli. 2009. "Culture: An Empirical Investigation of Beliefs, Work, and Fertility." *American Economic Journal: Macroeconomics* 1(1): 146–77.

Fernández-Kranz Daniel & Aitor Lacuesta & Núria Rodríguez-Planas, 2013. "The Motherhood Earnings Dip: Evidence from Administrative Records," *Journal of Human Resources*, vol. 48(1), pages 169-197.

Fernández-Kranz Daniel & Núria Rodríguez-Planas, 2021. "Too Family Friendly? The Consequences of Parents' Right to Request Part-Time Work **."** *Journal of Public Economics*, 2021, Vol. 197.

Finlay, Jocelyn, David Canning, & June Po. 2012. "Reproductive Health Laws Around the World 1960-present", https://doi.org/10.791/DVN/1WKMXK, Harvard Dataverse, V3.

Fortin, N. M. 2005. Gender role attitudes and the labour-market outcomes of women across OECD countries. *Oxford Review of Economic Policy*, 21(3): 416-438.

Goldin C. The U-Shaped Female Labor Force Function in Economic Development and Economic History. In: Schultz TP Investment in Women's Human Capital and Economic Development. University of Chicago Press ; 1995. pp. 61-90.

Goldin, C. 2014. A grand gender convergence: Its last chapter. *The American Economic Review*, 1044: 1091-1119.

Goldin, Claudia and Lawrence F. Katz. 2002. The power of the pill: Oral contraceptives and women's career and marriage decisions. Journal of Political Economy 110(4): 730-770.

Kleven, H., C. Landais and J. E. Søgaard (2019a). "Children and Gender Inequality: Evidence from Denmark", American Economic Journal: Applied Economics, 11(4), 181-209.

Kleven, H., C. Landais, J. Posch, A. Steinhauer and J. Zweimüller (2019b). "Child Penalties across Countries: Evidence and Explanations", American Economic Review Papers and Proceedings, vol. 109, 122-126.

Kuziemko, Ilyana, Jessica Pan, Jenny Shen, and Ebonya Washington. 2018. "The Mommy Effect: Do Women Anticipate the Employment Effects of Motherhood?" NBER Working Paper 24740.

Lavy, V., and Schlosser, A. 2011. Mechanisms and impacts of gender peer effects at school. *American Economic Journal: Applied Economics*, 32: 1-33.

Lavy, V., M. D. Paserman and A. Schlosser 2012. Inside the Black Box of Ability Peer Effects: Evidence from Variation in the Proportion of Low Achievers in the Classroom. *Economic Journal*, 122 559: 208–237.

England, P., Levine, A., & Mishel, E. (2020). Is the Gender Revolution Stalled? An Update. *Proceedings of the National Academy of Sciences, forthcoming*, 1-37.

Farré, L. and Vella, F., 2013. The intergenerational transmission of gender role attitudes and its implications for female labour force participation. *Economica*, *80*(318), pp.219-247.

Friesen, J. and B. Krauth. 2007. Sorting and inequality in Canadian schools. *Journal of Public Economics*, 9111: 2185-2212.

Gangl Markus, and Andrea Ziefle. 2009. "Motherhood, Labor Force Behavior and Women's Careers: An Empirical Assessment of the Wage Penalty for Motherhood in Britain, Germany and the U.S." *Demography* 46(2):341-69.

Hanushek, E. A., Kain, J. F., Markman, J. M., and S. G. Rivkin. 2003. Does peer ability affect student achievement? *Journal of Applied Econometrics*, 185: 527-544.

Hoxby, Caroline. 2000. Peer Effects in the Classroom: Learning from Gender and Race Variation. National Bureau of Economic Research, Working Paper 7867.

Olivetti, C., Patacchini, E., and Y. Zenou. 2020. Mothers, peers, and gender-role identity. *Journal of the European Economic Association*, 18(1), 266-301.

Rodríguez-Planas N. and N. Nollenberger. 2018. "Let the Girls Learn! It's not Only about Math… It is About Gender Social Norms." *Economics of Education Review*. Vol. 62: 230-253.

Rodríguez-Planas, Núria, and Anna Sanz-de-Galdeano and Anastasia Terskaya. 2018. "Independent Thinking and Hard Working, or Caring and Well Behaved? Short- and Long-Term Impacts of Gender Identity Norms." IZA Discussion Paper 11694.

Rodríguez-Planas, Núria, and Ryuichi Tanaka. 2021. "Gender norms and women's decision to work: evidence from Japan." *Rev Econ Household* (2021). https://doi.org/10.1007/s11150-021-09543-0.

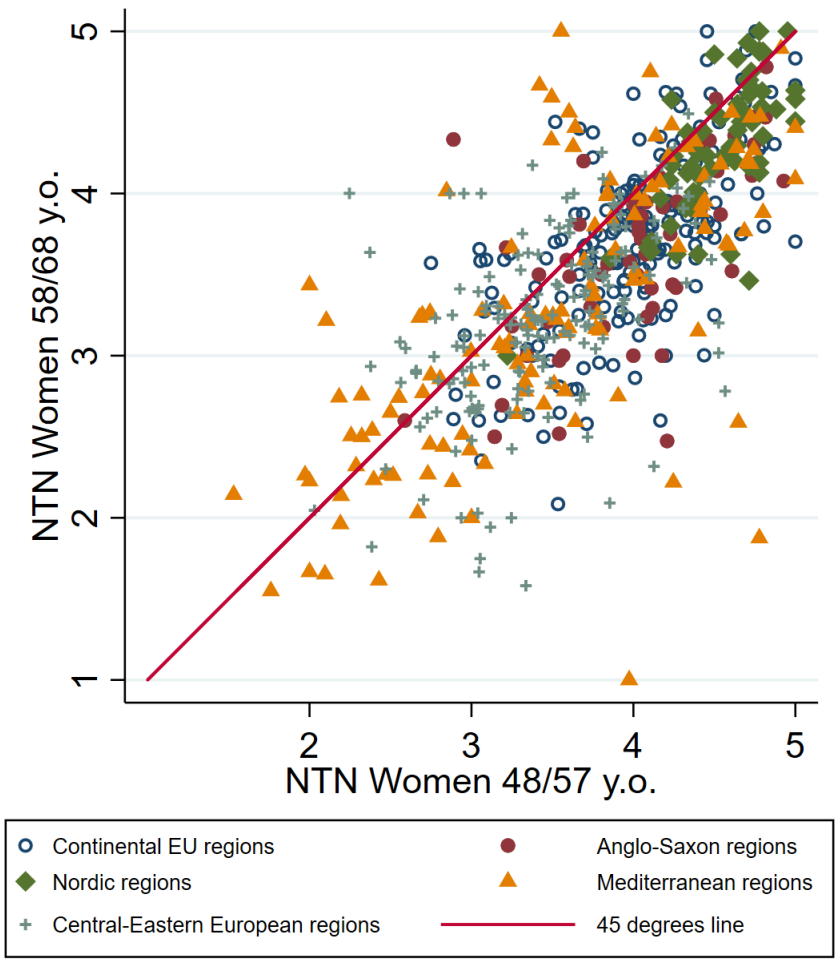
Seguino Stephanie. 2007. "Plus Ça Change? evidence on global trends in gender norms and stereotypes", Feminist Economics, 13:2, 1-28, DOI: 10.1080/13545700601184880

Waldfogel, Jane, 1997. "The Effect of Children on Women's Wages," *American Sociological Review 62(2): 209-211.*

Figure 1: NTN Sample Average 2004-2016 by Cohort

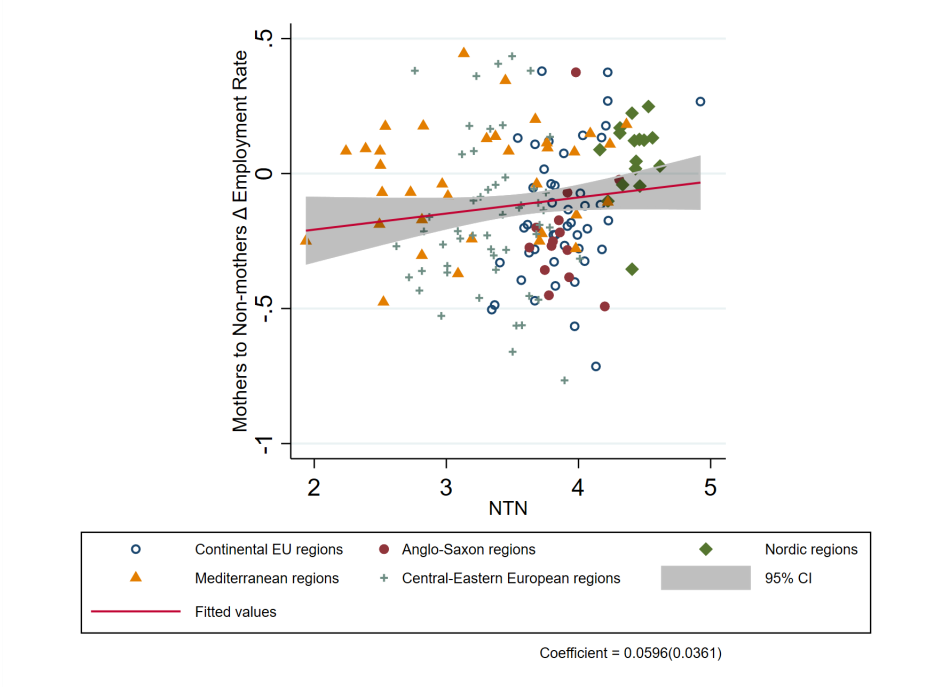
Notes: Non-traditional norm (NTN) represented in the Y axis refers to the average extent of disagreement (1-5 scale) to the statement *'when jobs are scarce, men should have more right to a job than women'* within the ESS region at the NUTS-2 level. Non traditional norm is measured as the average extent of disagreement (1-5 scale) to the statement 'when jobs are scarce, men should have more right to a job than women' within the ESS region. The lines plot average norms over ESS round 2004, 2008, 2010 and 2016 for the two cohorts of grandmothers aged 48-57 y.o., and grandmothers aged 58-68 y.o.

Figure 2: NTN, Cross-Regional Variation by Cohort

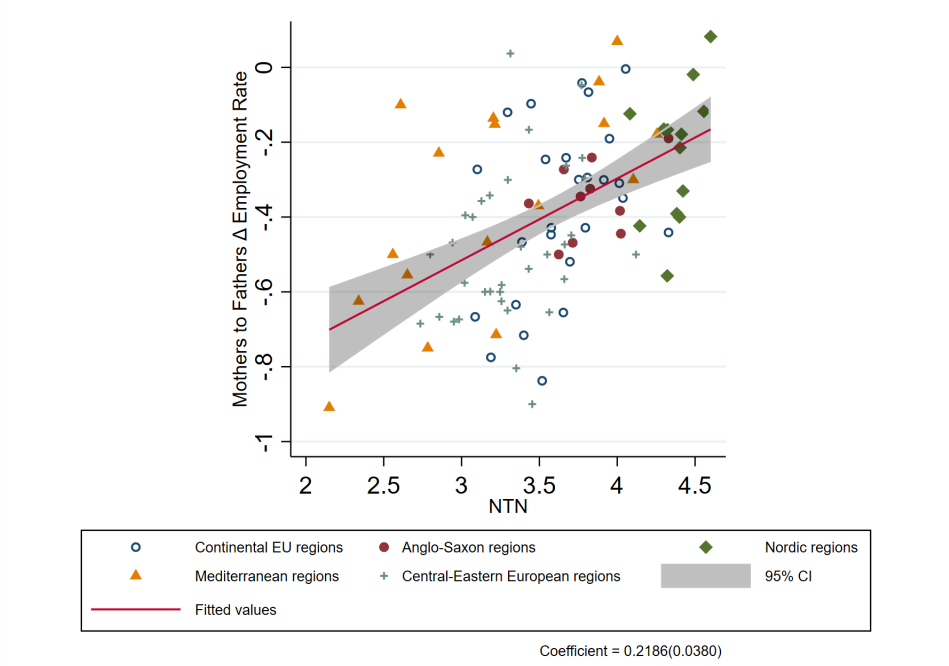


Notes: Non-traditional norm (NTN) represented in the Y axis refers to the average extent of disagreement (1-5 scale) to the statement *'when jobs are scarce, men should have more right to a job than women'* within the NUTS2 region, in a given year (i.e. 2004,2008,2010,2016). On the Y axis the average norm measured across grandmothers aged 58-68 Y.O. On the X-axis the average norm across grandmothers aged 48-57 Y.O. The graph plots the total variation (regional and overtime) within the sample.

Figure 3: First Child Employment Penalties and Elicited Gender Norms



(a) Non-mothers to mothers



(b) Fathers to mothers

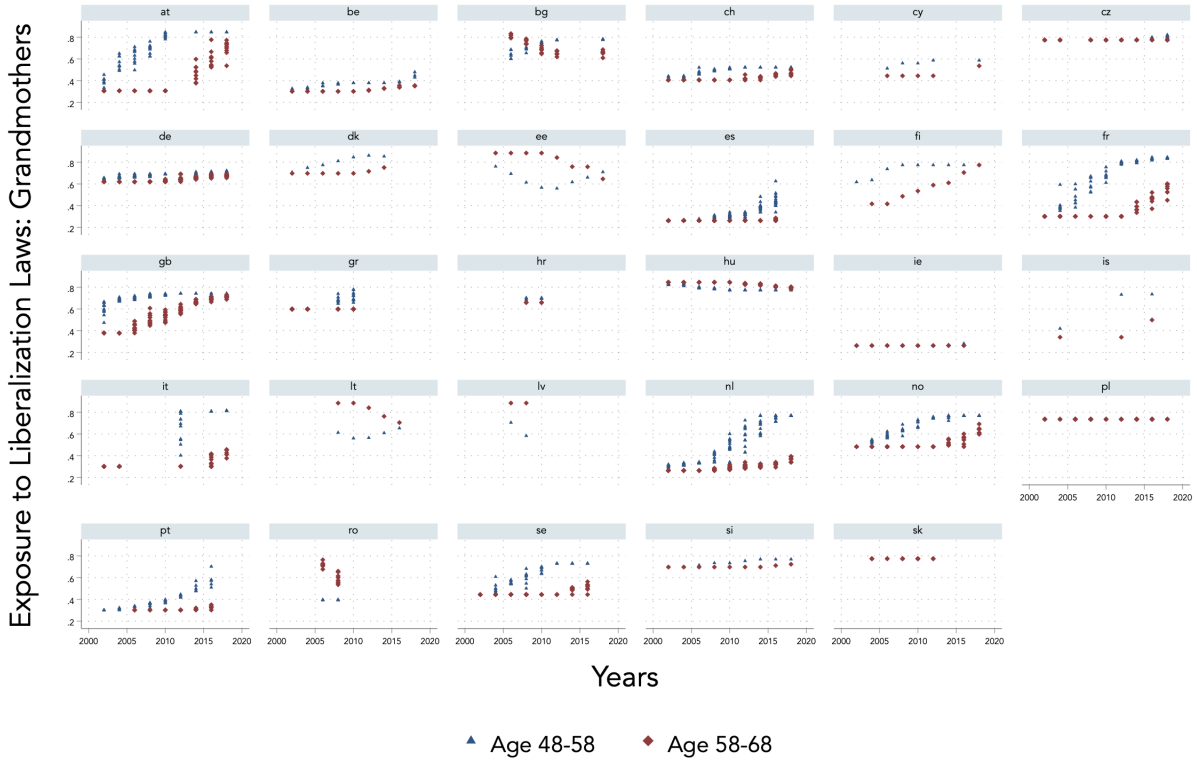
Notes: Non-traditional norm (NTN) represented in the X axis is measured as the average extent of disagreement (1-5 scale) to the statement *'when jobs are scarce, men should have more right to a job than women'* within the ESS region at the NUTS-2 level. The Y axis in figure 3a shows differences in average employment rates between mothers (first child 0-5 Y.O.) and non-mothers within each region. The Y axis in figure 3b shows the differences in average employment rates between mothers (first child 0-5 Y.O.) and fathers.

Table 1: Balancing Test

	Mothers	Non-Mothers	Fathers	Non-Fathers
	(1)	(2)	(3)	(4)
Individual Controls				
Tertiary ed.	0.011	0.017	–0.019	–0.022
	(0.039)	(0.029)	(0.039)	(0.019)
Secondary ed.	–0.029	0.001	0.030	0.027
	(0.039)	(0.024)	(0.038)	(0.026)
Age	0.093	0.135	0.070	–0.179
	(0.211)	(0.151)	(0.253)	(0.108)
Age squared	6.786	10.384	5.383	–9.949
	(13.640)	(8.913)	(16.759)	(7.081)
Never married	0.054	0.034*	0.028	0.009
	(0.033)	(0.020)	(0.031)	(0.017)
No longer married	0.016	0.003	0.019	0.002
	(0.016)	(0.012)	(0.023)	(0.010)
Ever unemployed in 5 yrs	0.027	–0.035*	–0.057**	–0.002
	(0.022)	(0.018)	(0.029)	(0.018)
Household Income	0.043	–0.017	–0.058	–0.006
	(0.055)	(0.054)	(0.082)	(0.034)
Parental Controls				
Father working	0.013	–0.013	0.003	0.010
	(0.022)	(0.012)	(0.030)	(0.015)
Mother working	–0.068*	0.035	–0.041	0.014
	(0.036)	(0.026)	(0.041)	(0.023)
Regional Controls				
Population	0.012	0.017	0.016	0.017
	(0.015)	(0.016)	(0.018)	(0.016)
Fertility rate	0.001	0.001	–0.006	–0.001
	(0.004)	(0.004)	(0.005)	(0.004)
Share with tertiary ed.	0.232	0.047	–0.143	0.056
	(0.148)	(0.126)	(0.155)	(0.121)
Unemployment rate	0.121	0.176	0.099	0.134
	(0.143)	(0.145)	(0.174)	(0.140)

Notes: The table shows the relationship between the non-traditional norm (NTN) which is the dependent variable and the particular individual, parental or regional control over different samples. Column [1] reports the correlation in the sample of mothers (with first child 0-5 Y.O.); column [2] reports the correlation for the sample of non-mothers; columns [3] and [4] consider the samples of fathers (first child 0-5 Y.O.) and non-fathers, respectively. All specifications include a dummy for the age group 20 to 30 year old, as well as country-by-year and regional dummies. Robust standard errors, clustered at the regional level in parentheses. Significance levels: *: 10% **: 5% ***: 1%

Figure 4: Grandmothers' exposure to Contraceptive Pill and Abortion Liberalization



Notes: On the X axis the year of the ESS round 2002-2018. On The Y axis the degree of average degree of reproductive health laws liberalization (contraceptive pill and abortion) to which the cohort of grandmothers in the region was exposed to at the age of 20 y.o. (max=1). Each dot indicates the average exposure of grandmothers in the corresponding cohort in the region. Authors' calculation on data from ESS and Finlay, Canning, and Po (2012).

Table 2: Non-Traditional Norm (NTN) and Employment - Motherhood Gap Estimator

	Panel (A) Dep. variable: employment of women (20-40 Y.O.)					
	(1)	(2)	(3)	(4)	(5)	(6)
NTN	0.031*					
	(0.016)					
NTN x Child	0.053**					
	(0.024)					
$\widehat{NTN}$		0.172**	0.134	0.211**	0.222**	0.156*
		(0.084)	(0.089)	(0.089)	(0.101)	(0.089)
$\widehat{NTN}$ x Child		0.055**	0.057**	0.083***	0.082***	0.087***
		(0.026)	(0.026)	(0.027)	(0.029)	(0.028)
Child	–0.411***	–0.388***	–0.396***	–0.459***	–0.471***	–0.462***
	(0.091)	(0.091)	(0.092)	(0.094)	(0.101)	(0.097)
Observations	14476	14476	14476	8629	7233	8166
	Panel (B) Dep. variable: employment of men (20-40 Y.O.)					
	(1)	(2)	(3)	(4)	(5)	(6)
NTN	–0.008					
	(0.016)					
NTN x Child	0.009					
	(0.009)					
$\widehat{NTN}$		0.147*	0.127	0.143	0.159	0.195**
		(0.085)	(0.090)	(0.095)	(0.099)	(0.094)
$\widehat{NTN}$ x Child		–0.001	–0.001	–0.004	–0.013	0.002
		(0.006)	(0.007)	(0.010)	(0.011)	(0.009)
Child	–0.017	0.019	0.021	0.038	0.066*	0.020
	(0.038)	(0.022)	(0.023)	(0.034)	(0.039)	(0.032)
Observations	15998	15998	15998	9273	7727	8760
$\widehat{NTN}$ predicted from Pill & Abortion	No	Yes	No	Yes	Yes	Yes
$\widehat{NTN}$ predicted from Abortion only	No	No	Yes	No	No	No
Respondent born after abortion legalization	No	No	No	Yes	Yes	Yes
Respondent born after pill legalization	No	No	No	No	Yes	No
no reform countries dropped	No	No	No	No	No	Yes

Notes: The sample includes male and female respondents aged 20-40 years old. In all specifications the dependent variable is a dummy variable equal to 1 if individual is employed and 0 otherwise. Child stands for *Small Child* in equation (3) in the main text. In Column [1], NTN refers to the average NTN for each ESS round (2004, 2008, 2010 and 2016) in the region. In Column [2], we report results when we use $\widehat{NTN}$ predicted from Column [1] of Table A-3. In Column [3], we report results when we use the predicted $\widehat{NTN}$ from Column [3] of Table A-3. For estimates in Column [4], we drop from the sample respondents born before the legalization of abortion, and in Column [5] we also drop respondents born before the legalization of the contraceptive pill. For estimates in Column [6], we drop from the sample countries that do not exhibit variation in the exposure indicator. All specifications include an extensive set of individual characteristics (i.e. tertiary and secondary education dummies, age, age-squared, two dummies for marital status, a dummy for being in a low income household, dummy for own unemployment experience), parental characteristics (two dummies father, mother working when the respondent was 14) and regional characteristics (total population, fertility rate, unemployment rate and share of tertiary educated). All specifications include regional and country by year FE. Standard errors clustered at the regional level in parentheses. Significance levels: *: 10% **: 5% ***: 1%

Table 3: Robustness Checks

	(1) Women 20-40	(2) Men 20-40
Panel A: Controlling for Cohort Non-Traditional Belief		
$\widehat{NTN}$ x Child	0.077***	0.001
	(0.028)	(0.009)
Own Cohort NTN Belief	0.184***	0.018
	(0.040)	(0.038)
Own personal Belief	0.007	–0.008
	(0.008)	(0.009)
Panel B: Controlling for Own Religious Intensity		
$\widehat{NTN}$ x Child	0.087***	0.003
	(0.028)	(0.009)
Own Religious Intensity	–0.005	–0.022
	(0.020)	(0.024)
Panel C: Alternative Outcome: Labour Participation		
$\widehat{NTN}$ x Child	0.086***	–0.006
	(0.026)	(0.009)
Panel D: Alternative Outcome: Log Hours Worked		
$\widehat{NTN}$ x Child	0.321***	0.017
	(0.107)	(0.035)
Panel E: Without Regional Controls		
$\widehat{NTN}$ x Child	0.086***	0.002
	(0.028)	(0.009)
Panel F: Without Parental Controls		
$\widehat{NTN}$ x Child	0.087***	0.002
	(0.028)	(0.009)
Panel G: Maternity Leave Legislation		
$\widehat{NTN}$ x Child	0.090***	0.000
	(0.029)	(0.009)
Maternity: Total Weeks	–0.002	–0.009***
	(0.003)	(0.003)
Maternity: Paid Weeks	–0.001	–0.000
	(0.001)	(0.001)
Panel H: Bootsrapped Standard Errors		
$\widehat{NTN}$ x Child	0.088**	0.001
	(0.033)	(0.016)
Panel I: 2SLS estimator		
NTN x Child	0.921**	–0.005
	(0.404)	(0.089)
K-P F-stat	3.370	4.309
First Stage		
$\widehat{NTN}$ x Child	0.093*	0.105**
	(0.051)	(0.051)
Panel L: Probit Model		
$\widehat{NTN}$ x Child	0.071***	–0.007
	(0.023)	(0.018)

Notes: The sample includes male and female respondents aged 20-40 years old. All specifications include the same set of individual characteristics, parental characteristics, and regional characteristics as in Table 2 above. All specifications include regional and country by year FE, but specifications in Panel G where country-by-year FE are replaced by country-specific controls for maternity leave legislation. Standard errors clustered at the regional level in parentheses in all specifications but in Panel H, where they are boostrapped (200 replications). Significance levels: *: 10% **: 5% ***: 1%

Table 4: Heterogeneity: Economic Conditions and Education Background of the Respondent

	[I] Unemployment Rate		[II] Own Education		[III] Parental Education	
	(1)	(2)	(3)	(4)	(5)	(6)
	Below Median	Above Median	Below Median	Above Median	Below Median	Above Median
$\widehat{NTN}$ x Child	0.174***	0.060***	0.084***	0.080**	0.114***	–0.015
	(0.028)	(0.020)	(0.032)	(0.037)	(0.035)	(0.024)
Observations	3833	4333	4724	3428	5428	2722

Notes: The sample includes female respondents aged 20-40 years old. All specifications include the usual extensive set of individual characteristics, parental characteristics, and regional characteristics. The unemployment rate is omitted from regional controls in Panel [I]. Own education is omitted from specification in Panel [II]. A parental background characterized by tertiary education is defined as at least one of the parents of the respondent having completed tertiary eduction. All specifications also include regional and country by year FE. Standard errors clustered at the regional level in parentheses. Significance levels: *: 10% **: 5% ***: 1%

Table 5: Motherhood gap Estimator: Endogenous Motherhood

	Panel (A) Dep. variable: employment of women (20-40 Y.O.)					
	(1)	(2)	(3)	(4)	(5)	(6)
NTN	–0.002					
	(0.037)					
NTN x $\widehat{Child}$	0.161					
	(0.116)					
$\widehat{Child}$	–0.532	–1.161*	–1.161*	–1.861**	–3.510***	–1.588**
	(0.390)	(0.593)	(0.607)	(0.810)	(1.248)	(0.786)
$\widehat{NTN}$		0.132	0.090	0.084	0.094	0.026
		(0.094)	(0.098)	(0.104)	(0.129)	(0.104)
$\widehat{NTN}$ x $\widehat{Child}$		0.335*	0.334*	0.528**	0.509**	0.515**
		(0.171)	(0.174)	(0.212)	(0.229)	(0.208)
Observations	14476	14476	14476	8629	7233	8166
	Panel (B) Dep. variable: employment of men (20-40 Y.O.)					
	(1)	(2)	(3)	(4)	(5)	(6)
NTN	–0.051					
	(0.036)					
NTN x $\widehat{Child}$	0.159					
	(0.121)					
$\widehat{Child}$	–0.548	–0.254	–0.246	–0.024	–0.226	0.061
	(0.389)	(0.378)	(0.388)	(0.522)	(0.906)	(0.511)
$\widehat{NTN}$		0.130	0.110	0.118	0.117	0.186*
		(0.092)	(0.098)	(0.101)	(0.111)	(0.099)
$\widehat{NTN}$ x $\widehat{Child}$		0.061	0.059	0.083	0.163	0.028
		(0.109)	(0.111)	(0.153)	(0.172)	(0.151)
Observations	15998	15998	15998	9273	7727	8760
$\widehat{NTN}$ predicted from Pill & Abortion	No	Yes	No	Yes	Yes	Yes
$\widehat{NTN}$ predicted from Abortion	No	No	Yes	No	No	No
Respondent born after abortion legalization	No	No	No	Yes	Yes	Yes
Respondent born after pill legalization	No	No	No	No	Yes	No
no reform countries dropped	No	No	No	No	No	Yes

Notes: The sample includes male and female respondents aged 20-40 years old. In all specifications the dependent variable is a dummy variable equal to 1 if individual is employed and 0 otherwise, while $\widehat{Child}$ is the probability of having a child predicted from Column [4] of Table A-3. In Columns [1], the NTN variable refers to the average NTN for each ESS round (2004, 2008, 2010 and 2016). In Column [2] we report results when we use the predicted $\widehat{NTN}$ from Column [1] of Table A-3. In Column [3] we report results when we use the predicted $\widehat{NTN}$ from Column [3] of Table A-3. For estimates in Column [4], we drop from the sample respondents born before the legalization of abortion, and in Column [5] we also drop respondents born before the legalization of the contraceptive pill. For estimates in Column [5], we drop from the sample countries that do not exhibit variation in the exposure indicator. All specifications include the usual set of extensive set of individual characteristics, parental characteristics, and regional characteristics. All specifications include regional and country by year FE. Standard errors clustered at the regional level in parentheses. Significance levels: *: 10% **: 5% ***: 1%

A Appendix

Table A-1: Variation after Removal of Fixed Effects

	(1) Observations	(2) Mean	(3) Standard Deviation	(4) Min	(5) Max
Raw Summary					
NTN	30474	3.713	0.628	1.000	5.000
Residuals after Removing Cohort FE, Region FE, Year FE					
NTN	30474	–0.000	0.275	–1.927	2.192
Residuals after Removing Cohort FE, Region FE, and Country-by-Year FE					
NTN	30474	–0.000	0.183	–1.118	1.739

Notes: Authors calculations on ESS data.

Table A-2: Descriptive Statistics

	Panel A: Women			Panel B: Men		
	(1)	(2)	(3)	(4)	(5)	(6)
	0-5 Child	No Child	Difference	0-5 Child	No Child	Difference
Individual Controls						
Employed	0.539	0.655	0.116***	0.910	0.710	–0.199***
	(0.499)	(0.475)	(0.009)	(0.287)	(0.454)	(0.009)
Men Have More Right to Work Than Women: Disagree (Gender Beliefs)	4.038	4.185	0.147***	3.865	3.765	–0.100***
	(1.099)	(1.044)	(0.020)	(1.123)	(1.147)	(0.023)
NTN	3.677	3.721	0.044***	3.695	3.722	0.027**
	(0.624)	(0.631)	(0.012)	(0.636)	(0.625)	(0.013)
Tertiary Education	0.415	0.399	–0.017*	0.372	0.281	–0.091***
	(0.493)	(0.490)	(0.009)	(0.483)	(0.450)	(0.009)
Secondary Education	0.458	0.513	0.055***	0.486	0.578	0.092***
	(0.498)	(0.500)	(0.009)	(0.500)	(0.494)	(0.010)
Age	30.570	27.437	–3.134***	32.529	28.202	–4.328***
	(4.744)	(5.585)	(0.098)	(4.389)	(5.696)	(0.109)
Age Squared	957.041	783.955	–173.086***	1077.427	827.775	–249.651***
	(289.011)	(326.162)	(5.801)	(279.608)	(335.026)	(6.499)
Never Married	0.291	0.788	0.497***	0.239	0.832	0.593***
	(0.454)	(0.409)	(0.008)	(0.427)	(0.374)	(0.008)
No Longer Married	0.050	0.042	–0.008**	0.019	0.048	0.029***
	(0.218)	(0.201)	(0.004)	(0.138)	(0.214)	(0.004)
Ever Unemployed in 5 Years	0.102	0.080	–0.022***	0.057	0.098	0.040***
	(0.303)	(0.271)	(0.005)	(0.233)	(0.297)	(0.006)
Household Income	0.265	0.355	0.090***	0.198	0.335	0.138***
	(0.441)	(0.479)	(0.009)	(0.398)	(0.472)	(0.009)
Parental Controls						
Father Working	0.902	0.908	0.006	0.917	0.910	–0.007
	(0.298)	(0.289)	(0.005)	(0.275)	(0.286)	(0.006)
Mother Working	0.691	0.715	0.024***	0.703	0.705	0.003
	(0.462)	(0.451)	(0.008)	(0.457)	(0.456)	(0.009)
Regional Controls						
Population	3.117	3.082	–0.035	3.120	3.100	–0.020
	(2.635)	(2.746)	(0.050)	(2.680)	(2.833)	(0.056)
Fertility Rate	1.622	1.576	–0.045***	1.622	1.579	–0.043***
	(0.250)	(0.242)	(0.004)	(0.248)	(0.243)	(0.005)
Share with Tertiary Education	27.655	26.938	–0.717***	28.283	26.957	–1.327***
	(9.064)	(9.235)	(0.169)	(9.055)	(9.202)	(0.183)
Unemployment Rate	8.033	8.163	0.130	7.990	8.041	0.052
	(4.485)	(4.703)	(0.085)	(4.614)	(4.638)	(0.093)
N	4163	10313	14476	3103	12895	15998

Notes: Authors calculations on ESS data. Standard errors clustered by region in Parentheses. Significance levels: *: 10% **: 5% ***: 1%

Table A-3: Auxiliary NTN and motherhood regressions

	(A) NTN (Grandmothers 48/68 Y.O. in the Region)			(B) Women 20/40 Y.O.		
	(1)	(2)	(3)	(4)	(5)	(6)
	NTN	NTN	NTN	Child 0-5	Child 0-5	Child 0-5
Exposure to Abortion or Pill Liberalizations	0.691***			–1.434***		
	(0.121)			(0.252)		
Exposure to Abortion Liberalizations		0.378***	0.329***		–0.463***	–0.534***
		(0.068)	(0.076)		(0.091)	(0.086)
Exposure to Pill Liberalizations			0.434		–2.210***	
			(0.303)		(0.435)	
Observations	1250	1255	1250	29879	28981	28981
Adj. R-squared	0.635	0.633	0.634			
Pseudo R-Squared				0.078	0.084	0.074
(Adj./Pseudo) R-squared with no exposure	0.625			0.069		

Notes: Regressions in Columns (1)-(3) are carried out at the regional level and include regional characteristics (total population, fertility rate, unemployment rate and share of tertiary educated). Regressions in columns (4)-(6) include the usual set of individual characteristics (i.e. tertiary and secondary education dummies, age, age-squared, two dummies for marital status, a dummy for being in a low income household, dummy for own unemployment experience), parental characteristics (two dummies father, mother working when the respondent was 14) and regional characteristics (total population, fertility rate, unemployment rate and share of tertiary educated). All specifications also include region and country by year fixed effects. Results from regression in column (1) are used to predict $\widehat{NTN}$ used in Tables 2, and 5, columns [2],[4],[5]. Results from regression in column (1) are used to predict $\widehat{NTN}$ used in Table 2, column [3]. Results from regression in column (3) are used to predict $\widehat{NTN}$ used in Tables 2 and 5, column [3]. Results from regression in column (4) are used to predict $\widehat{Child}$, which is used as a regressor in Table 5 and as an instrument in 2SLS estimates in Table 3, Panel I. Standard errors are clustered at the regional level in parentheses. Significance levels: *: 10% **: 5% ***: 1%

Table A-4: Descriptive Statistics of Observed and Predicted Norm and Motherhood Values

	(1)	(2)	(3)	(4)
	$\widehat{NTN}$	NTN	$\widehat{Child}$	Child
Mean	3.569	3.705	0.281	0.342
Median	3.740	3.758	0.283	0.000
Standard Deviation	0.644	0.626	0.092	0.474
Minimum	0.205	1.000	0.027	0.000
Maximum	4.781	5.000	0.911	1.000

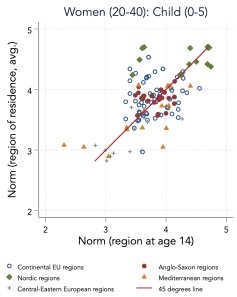
Note: $\widehat{NTN}$ and $\widehat{Child}$ are predicted from the regression reported in Columns 1 and 4 of Table A-3, respectively. Statistics refer to the sample of women only.

Table A-5: Robustness Test: European Values Survey (EVS)

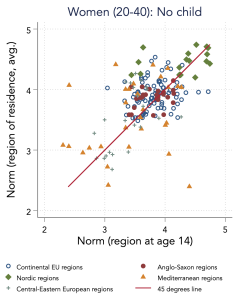
	OLS					
	(1)	(2)	(3)	(4)	(5)	(6)
	Norm Region Residence	Norm Region Residence	Norm Region Residence	Norm Region Age 14	Norm Region Age 14	Norm Region Age 14
NTN	0.082***	0.084***	0.012	0.087***	0.088***	0.019
	(0.027)	(0.028)	(0.035)	(0.028)	(0.028)	(0.037)
Child	–0.281***	–0.279***	–0.275***	–0.317***	–0.315***	–0.308***
	(0.049)	(0.049)	(0.048)	(0.042)	(0.041)	(0.040)
NTN x Child	0.078	0.079	0.078	0.129**	0.129**	0.123**
	(0.064)	(0.063)	(0.063)	(0.054)	(0.054)	(0.053)
Observations	2893	2893	2893	2871	2871	2871
Personal Controls	Yes	Yes	Yes	Yes	Yes	Yes
Unemployment Controls	Yes	Yes	Yes	Yes	Yes	Yes
Parental Work Controls	No	Yes	Yes	No	Yes	Yes
country fixed effects	No	No	Yes	No	No	Yes

Notes: OLS regressions equivalent to those reported in Table 2 (see Column (1)) carried out on data available from the 2008 wave of the European Value Study (EVS). Countries in the EVS 2008 sample are Austria, Belgium, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Netherlands, Norway, Poland, Portugal, Slovak Republic, Slovenia, Spain, Sweden, Switzerland, Great Britain, Northern Ireland. Columns (1) to (3) report results as we compute the norm in the current region of residence of the respondent. Columns (4) to (6) report results as we compute the norm in the region of residence of the respondent at the age of 14 y.o. Standard errors are clustered at the respective region level. Significance levels: *: 10% **: 5% ***: 1%

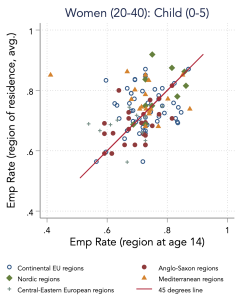
Figure A-1



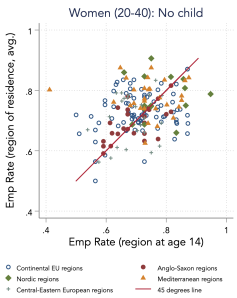
(a) EVS Migration (NTN)



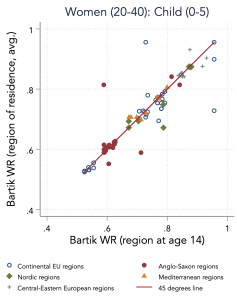
(b) EVS Migration (NTN)



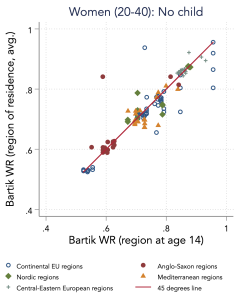
(c) EVS Migration (Employment)



(d) EVS Migration (Employment)



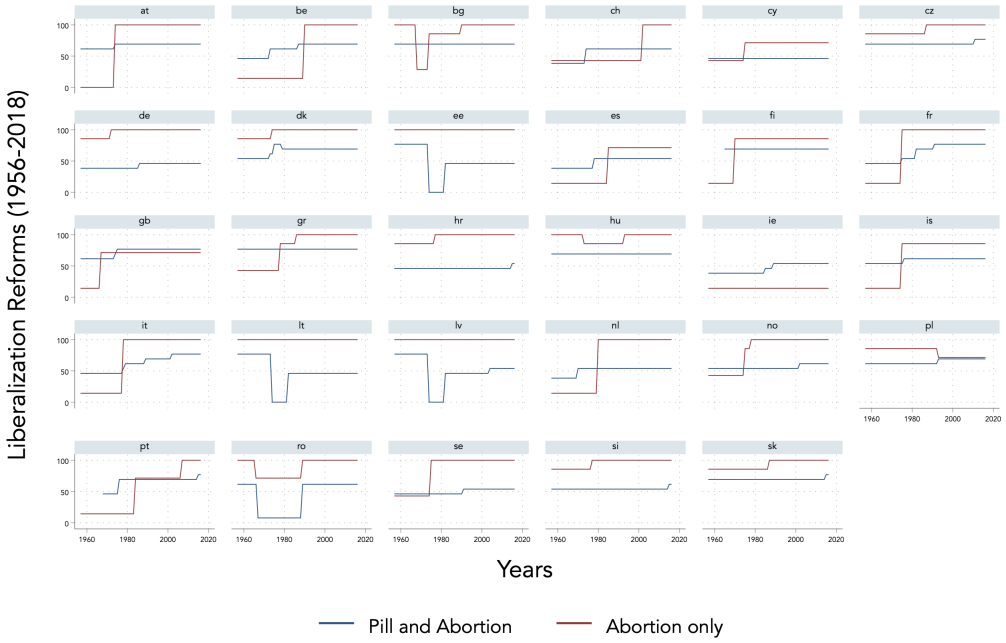
(e) EVS Migration (Bartik Wage Ratio)



(f) EVS Migration (Bartik Wage Ratio)

Notes: All figures have been plotted using the European Value Survey (EVS) for women between the age of 20-40. Figure A-1a and A-1d show the relationship between the average norm at the region of residence and the norm at the region when the respondent was 14 Y.O. for every combination of region of residence and region at age 14. Figure A-1a shows the relationship for mothers (first child 0-5 Y.O.) and A-1d shows the relationship for childless women. The norm in the EVS dataset refers to the response to the following question "Men should have more rightto a job than women (agree or disagree)?". Figure A-1c and A-1d show the relationship between average employment rates at the region of residence and the employment rates at the region when the respondent was 14 Y.O. for each combination of region of residence and region at age 14. Figure A-1e and A-1f show the relationship between average bartik wage ratio at the region of residence and the bartik wage ratio at the region when the respondent was 14 Y.O. for every each combination of region of residence and the region at age 14.

Figure A-2: Exposure to Liberalization Index - Sample Variation



Notes: On The Y axis the degree of average degree of liberalization to in the country between 1960 and 2018. Authors' calculation on data from Finlay, Canning, and Po (2012).