

Changing Gender Norms across Generations: Evidence from a Paternity Leave Reform*

Lidia Farré, Christina Felfe, Libertad González, and Patrick Schneider

January 2023

Abstract: Social norms are an important barrier to gender convergence. We show that public policy designed to promote gender equality at home can pave the way towards gender convergence by shaping gender norms in the next generation. We combine the introduction of paternity leave in Spain with a large-scale lab-in-the field experiment in secondary schools. Following a local difference-in-differences approach, we show that children born after the policy change exhibit more gender egalitarian attitudes and perceive less stereotypical social norms. They are also more likely to engage in counter-stereotypical day-to-day behaviors and to deviate from the male-breadwinner model in the future.

JEL Codes: J08, J13, J16, J18.

Keywords: Gender equality, gender norms, paternity leave permits.

*Farré acknowledges the financial support of the Spanish Ministry of Science and Innovation (grant PID2019-104319RB-I00) and the Government of Catalonia (grant SGR2017-644). González acknowledges financial support from the Spanish Ministry of Economy and Competitiveness, through the Severo Ochoa Programme for Centres of Excellence in R&D (CEX2019-000915-S). Felfe and Schneider acknowledge funding from the Bavarian Research Alliance through the BayIntAn Programme. We are very grateful to Sofía Sierra, Ana Costa, Ana Rodríguez, Tanya Surovtseva, and Claudia Meza, for excellent research and logistical assistance. We also appreciate the comments and suggestions of seminar participants at Banco de España, Goethe University Frankfurt, Institut d'Anàlisi Econòmica, Max Planck Institute for Research on Collective Goods Bonn, University College London, University of Düsseldorf, University of Essen, University of Florida, University of Heidelberg, University of Konstanz, University St. Gallen, University of Lugano, and European University Institute, as well as participants at EALE 2019, ESPE 2022, SKILS 2022 and SAEe 2022. Author affiliations and contacts: Farré (Universitat de Barcelona, IAE-CSIC and IZA, lidia.farre@ub.edu), Felfe (Universität Würzburg, CEPR and CESifo, christina.felfe@uni-wuerzburg.de), González (Universitat Pompeu Fabra and Barcelona School of Economics, libertad.gonzalez@upf.edu), and Schneider (Universität Würzburg, patrick.schneider@uni-wuerzburg.de).

1. Introduction

Despite significant progress in women's labor force participation and educational attainment, large gender gaps in labor market outcomes remain across many countries. In rich nations, most of the remaining gaps can be traced to parenthood, which has been shown to lead to significant and lasting setbacks in women's careers relative to men (Fernández-Kranz et al. 2014, Angelov et al. 2016, Lundborg et al. 2017, Kleven et al. 2019a, Bertrand 2020, Cortes and Pan 2020). Public policies such as subsidized childcare and paid parental leave have not succeeded at mitigating these inequalities (Kleven et al. 2021, Albanesi et al. 2022).¹

Recent work has highlighted the importance of gender norms in hindering progress towards gender convergence (Kleven et al. 2019b, Bertrand 2020). Traditional norms attribute to women the role of main caregivers within the family and limit their labor market choices after having children (Bertrand et al. 2015, Fortin 2015, Goussé et al. 2017, Cortes and Pan 2020, Grewenig et al. 2020, Olivetti et al. 2020, Boelmann et al. 2021, Andresen and Nix 2022). Previous research has shown that gender norms are passed on from parents to children (Fernández et al. 2004, Alesina et al. 2013, Farré and Vella 2013). Such strong intergenerational transmission may perpetuate gender inequality at home and render public policies largely ineffective in eradicating the existing gender differentials in the labor market (Kleven et al. 2021, Albanesi et al. 2022).

We provide novel evidence that public policy has the power to promote gender convergence in the long term by fostering counter-stereotypical behaviors among parents, and through that shaping the gender norms of their children. To that end, we focus on paternity leave, an increasingly

¹ Kleven et al. (2021) show that several large expansions of parental leave and childcare subsidies that have taken place in Austria since the 1950s led the gender gaps in the labor market largely unaffected. In a recent survey for different countries, Albanesi et al. (2022) document limited effects of family policies on women's careers.

popular family policy designed to incentivize less traditional specialization patterns within households. Paternity leave takes place at the onset of a child's life and may trigger persistent changes in the division of paid and unpaid work within households. Indeed, paternity leave has been shown to increase fathers' contribution to childcare and household work (and mothers' employment and earnings) persistently, with evidence from the US (Petts et al. 2020), Canada (Patnaik 2019, Dunatchik and Özcan 2020), Germany (Tamm 2019), Norway (Kotsadam and Finseraas 2011, Rege and Soli 2013) and Spain (Farré and González 2019, González and Zoabi 2021). This shift towards a more gender-neutral home environment (away from the traditional male-breadwinner model) may affect the formation of gender norms at an early age and promote gender equality in the next generation (Bertrand 2020).²

To provide causal evidence, we take advantage of the introduction of paternity leave in Spain in 2007. The reform entitled new fathers to 13 days of fully compensated paternity leave (in addition to the two days that were available before). While this policy change seems small, previous research shows that the reform was well received (take-up was close to 60%, see Figure 1) and triggered important changes within couples, promoting gender equality in the home and in the labor market. Specifically, the introduction of paternity leave led mothers to return to work earlier after childbirth, and induced fathers to become more involved in childcare, an effect which was still detectable several years after childbirth (Farré and González 2019, González and Zoabi 2021).³

² Fernández et al. (2004) show that the wives of men who were brought up in families where the mother worked are themselves significantly more likely to work. Bertrand (2019) also documents that children growing up in non-traditional families display more gender-egalitarian attitudes.

³ Similar evidence on paternity leave promoting a more gender-egalitarian family model exists at the international level, for instance for the U.S. (Petts et al. 2020), Canada (Patnaik 2019, Dunatchik and Özcan 2020), Germany (Tamm 2019), and Norway (Kotsadam and Finseraas 2011, Rege and Soli 2013).

We base our identification on the sharp cut-off date determining fathers' eligibility for paternity leave. The new law was passed shortly before implementation and fathers were entitled to the benefit only if their child was born on or after March 24, 2007. Thus, the reform represents an ideal natural experiment allowing us to isolate the effects of paternity leave on the gender norms and behavior of children whose parents were eligible for the permit.

To take advantage of this quasi-experimental setting, we conducted a large lab-in-the-field experiment with children born in 2006 and 2007, i.e., the reform year and the previous one. We collaborated with secondary schools in the region of Catalonia and surveyed more than 2,000 children between May 2019 and January 2020, when the target children were 11 to 13 years old. We rely on state-of-the-art survey and experimental methods to elicit comprehensive data on children's own attitudes and opinions regarding gender roles in the family and the workplace, as well as their perceptions of the gender norms prevailing among their peers. Both types of measures are relevant, since behaviors may be driven by children's (possibly inaccurate) beliefs about other people's opinions, on top of their own attitudes and views (Bursztyn et al. 2020).

To ensure comparability with the existing evidence on gender attitudes, we first employed a widely used battery of survey questions on the roles of men and women in the labor market and in the home (taken from the International Social Survey Program, ISSP). Based on these questions, we construct an index of what we call "*gender role attitudes*". In addition, we introduced a set of questions focusing explicitly on the preferences regarding parental labor supply when children are below school age. Here we go beyond the literature and asked children not only about their opinion regarding mothers' labor supply, but also about that of fathers'. Thus, we can construct measures for children's individual opinion regarding the preferred labor supply of both parents.

Finally, we elicited children's perception of the social norm regarding mothers' and fathers' labor market behavior. We built upon Krupka and Weber (2013) and implemented an incentivized coordination game asking children about their beliefs regarding the prevailing gender norm in their classroom. To our knowledge, only adults have played incentivized coordination games in previous work. We thus put great care into adapting the game to make it suitable for children, allowing us to provide novel measures of social norms in children.

To enhance our understanding of whether paternity leave has the power to effectively foster gender equality in the next generation, we also asked children about their day-to-day contributions to household chores, and about their expectations regarding their own future work and family life.

Combining our unique lab-in-the-field experiment with the natural experiment, we follow a local difference-in-differences strategy to estimate the causal effect of paternity leave on children's gender norms, behaviors, and expectations. A compelling feature of the paternity leave reform is that it became effective on March 24, which lies in the middle of the school year (the school year cut-off date is December 31). As a result, children born shortly before and after the cut-off date attend the same school grade, while only children born after the cut-off are directly affected by the paternity leave expansion. Our baseline specification relies on comparing children born in a 12-week window before and after the paternity leave introduction, i.e., children born between January 1 and June 12 of 2007. To isolate any relative age or season of birth effects, we use children born in the same time window in the pre-reform year (2006) as a control group. We also control for a set of individual background characteristics, as well as school fixed effects.

Our results show that the introduction of paternity leave led to children displaying significantly more gender-egalitarian views. We estimate an effect of 0.26 standard deviations on the children's attitudes index about gender roles. Moreover, children's perception of the prevailing

social norm is also affected. Paternal eligibility leads to an 18 percentage-point (a 28%) increase in the share of children stating that it is “socially appropriate” for a woman with young children to work, and a 16 percentage-point (a 36%) increase in the share stating that it is “socially appropriate” for a father to work less than full time to take care of a child. We thus find that the introduction of paternity leave affected children’s norms regarding both mothers and fathers to a similar extent. The results on children’s gender role attitudes and norms are remarkably stable in size and across a battery of robustness checks, including alternative sample specifications and estimation approaches (e.g., a specification with class fixed effects and a regression discontinuity design).

Turning to revealed behaviors, we find that the introduction of paternity leave promoted less specialized gender patterns in children’s day-to-day contributions to home production. Specifically, children born after the reform are 14 percentage points (24%) more likely to engage in counter-stereotypical household tasks: girls get more involved in male chores, such as small repairs and grocery shopping, while boys increase their contribution to female household chores, such as cleaning. Finally, we show that children born after the reform are 16 percentage points more likely to report counter-stereotypical expectations regarding their own future work and family life. Boys are less likely to expect working full-time when they have children, while girls are more likely to expect to be full-time working mothers.

All in all, our findings provide compelling evidence that the introduction of paternity leave in Spain induced the next generation to step away from traditional gender norms and stereotypical behaviors. It is still too early to learn about the effects of this policy on children’s labor supply choices. Nevertheless, taking children’s expectations regarding their own future work and family life at face value, our results suggest that a policy that fosters less gendered behaviors among parents

may promote more egalitarian views among their children and thus narrow the gender gap in labor market outcomes in the next generation.

Our study provides novel evidence on the extent to which public policy can shape gender norms across generations. There is only limited previous evidence showing that public policy can causally influence gender norms, and much less intergenerationally. Focusing on the Earned Income Tax Credit implemented in the U.S. in 1975, Bastian (2020) documents how the subsequent rise of working mothers changed the U.S. economy and the role of women in society. He provides suggestive evidence that the influx of working mothers led to a higher approval of working women in the same generation.⁴ Our study focuses instead on a policy shifting parents' behavior and its potential to change gender norms and labor market decisions in the next generation. As such, we also contribute to a scarce, but growing literature on spillover effects of public policies (Dahl et al. 2014, I Brollo et al. 2020, Dahl and Gielen 2021).

More broadly, our paper contributes to the growing literature on the determinants of gender norms. We provide causal evidence on the role of parental behavior in shaping children's gender norms, or in other words, on vertical socialization mechanisms. As such, our research relates to recent work on horizontal socialization mechanisms, such as school and peers. Dhar et al. (2021) evaluate the impact of a school-based randomized intervention in India that engaged adolescents in classroom discussions about gender equality. The two-and-a-half-year-long program not only fostered more progressive gender attitudes, but also induced more gender-egalitarian behaviors. The effect sizes of their intervention are comparable to ours.

⁴ There is also cross-country evidence on the impact of childcare provision (Neimanns 2021) and on the impact of same-sex relationship recognition policies on attitudes in Europe (Aksoy et al. 2020).

Garcia-Brazales (2021) studies the impact of female peers on gender role attitudes, perceived gender ability, and gender-related behaviors in Vietnam. He finds that exposure to more female peers erodes traditionalism both for men and women, which translates into actual behavior (e.g., an increase in female participation in college and in male participation in home production).⁵ For Japan, Hara and Rodríguez-Planas (2021) show that an educational reform that eliminated gender-segregated topics classes (e.g., housekeeping and shop classes) in secondary schools led to less gendered behaviors at home and in the labor market. Dahl et al. (2021) study the effect of young men's exposure to women in a traditionally male-dominated environment. Their context is the military in Norway, where they randomly assigned female recruits to some squads but not to others during boot camp. While living and working with female colleagues for eight weeks induced more egalitarian gender attitudes, the effects did not persist in the long run. In sum, while the number of studies on the role of peers and teachers in shaping (adolescents') gender norms is growing, we provide a novel causal study on the role of parents.⁶

The remainder of the paper is organized as follows. The following section provides background information on our data collection, the questionnaire, and the sample. Section 3 describes the paternity leave reform, section 4 explains the empirical approach, section 5 shows the results, and section 6 concludes.

⁵ In related work, Olivetti et al. (2020) provide causal evidence that labor force participation of high school peers' mothers affects adult women's labor force participation, above and beyond the effect of their own mothers.

⁶ A recent study by Brouwer et al. (2022) shows that parents enforce and comply more with norms when their children are present compared to when they are not.

2. Data

2.1. Setting and study implementation

Our central idea is to use comprehensive data on gender norms among children in a quasi-experimental framework. Our identification strategy relies on the introduction of paternity leave in Spain. A necessary condition for the implementation of our empirical approach is a large sample of children born in a narrow window around the reform's enactment date (March 24, 2007). To reach this group of children, and to ensure a large enough number of observations for children from all family backgrounds, we opted to collect our own data and to run a survey in schools.

We collaborated with 16 selected secondary schools in the region of Catalonia (Spain) in 2019-20.⁷ We targeted all children attending 5th to 7th grade (predominantly born in 2006 and 2007 and 11-13 years old) and thus children at the onset of adolescence, a critical time in the development of identity and social norms formation (Carter and Patterson, 1982, Kohlberg 1976, Markus and Nurius 1986, Waylen and Wolke, 2004). Our sampling design allowed us to reach all children within designated social networks (school cohorts and school classes).

Data collection occurred in two phases. The pilot phase took place between May 20 and May 24, 2019. During this phase, we collected data in 3 schools, all together 15 classes with 401 students. The second (main) phase happened between January 13 and February 7, 2020. We visited 13 additional schools, 80 classes with 1,926 participants altogether.⁸

⁷ Our sample includes a diverse set of public and private schools in urban and rural areas, as well as higher- and lower-income neighborhoods. Note that our identification strategy relies on within-school variation only.

⁸ A third phase was scheduled for March 16-April 3, 2020, but was canceled due to the outbreak of the Covid-19 pandemic (schools were closed in Spain on March 14, and in-person access of researchers to classrooms was not allowed afterwards).

The study was pre-registered (at OSF)⁹ and approved by the ethics and data protection office at Universitat Pompeu Fabra. A week before data collection, schools informed parents about the study by email, giving them the option to withdraw their consent regarding their children's participation.¹⁰ The survey took place during regular instruction time, which ensured that we could reach all students (except in case of illness or other excused absences from school). At the beginning of the survey, we informed students about the general purpose of the study without providing details about our specific research question.¹¹ We also gave them the option to not participate or to drop out at any point during the survey. Students were paid a show-up fee and could earn points in several incentivized parts of the survey, which were later exchanged for vouchers valid at a nearby stationary store. Participants received on average 5.73€, with a minimum of 3€ (show-up fee) and a maximum of 12€.

Data collection took place in a designated room inside the school (e.g., the cafeteria, the library, or the workshop), where we installed 35 laptops. Students came to the room together with their classmates, in groups of 14 to 32 students. To ensure privacy, we set up cardboard screens between the students (see Figure A.1 in the Appendix A). To avoid priming effects, the order of the questions was randomized. Students answered all questions individually and at their own pace. They needed on average 27 minutes to complete the questionnaire. Students who completed the questionnaire early were asked to stay in the room reading in silence (we provided reading materials).

⁹ Created on 2020-02-10 under “Can public policy change gender norms? How paternity leave affects children”.

¹⁰ The ethics and data protection office agreed to this opt-out/informed dissent approach as the data collection took place in the premises of schools.

¹¹ The precise information given to the students is shown in Appendix B.

Out of the 2,327 students present in the classroom on the day of the study, 9 did not consent to participate, 8 dropped out during the survey, 9 were not able to answer the questions without help, and information for 1 participant was not stored.

2.2. The questionnaire and the outcome variables

To yield a comprehensive set of measures for children's gender norms, we employed both classical survey and incentivized experimental methods. We further included questions asking students about their day-to-day behaviors, their expectations regarding their own future employment and fertility, and their parents' engagement in the home and in the labor market.¹² Table A.1 in Appendix A provides an overview of our main outcome variables, including summary statistics at baseline for the pre-reform sample (children born in 2007 prior to the introduction of paternity leave, i.e., those born between January 1 and March 24, 2007).

Our first measure of interest are children's attitudes about gender roles. For this purpose, we employed a standard battery of questions, developed by the International Social Survey Programme (ISSP) and widely used in the economics literature on gender norms (see for instance Giuliano 2018 and Bertrand et al. 2021). Specifically, we asked students to rate on a 5-point Likert scale whether they agreed or disagreed with a series of statements regarding the role of men and women in the labor market and in the home (see Question I.1 in the Questionnaire in Appendix B). Table A.1

¹² The full questionnaire can be found in Appendix B. In addition to the questions described in the paper, we asked the students to state their desired future occupation. We lack precision when estimating the reform effect on occupational aspirations. We therefore deviate from our original plan (see also our preregistration) and do not include these results in the paper.

Panel A, column (1) shows the share of pre-reform children with a non-traditional view regarding each statement.¹³

Children in our pre-reform sample largely agree with statements on women's labor market participation in general (e.g., 91% agree or strongly agree with the statement (b) "*Both men and women should contribute to household income*", 98% disagree or strongly disagree with the statement (g) "*A man's job is to earn money and a woman's job is to look after the home and the family*"). However, children are less in agreement when it comes to the consequences that women's labor market participation may have for children and family life (e.g., only 59% and 63% disagree or strongly disagree with the statements (c) "*A preschool child suffers when his or her mother works*" and (d) "*All in all, family life suffers when the woman has a full-time job.*").¹⁴

Second, we elicited children's individual opinions on whether a woman with a child below school age should work full-time, part-time, or not at all (see Question I.2 in the Questionnaire in Appendix B).¹⁵ We again drew on a widely used question from the ISSP (see for instance, Doepke and Kindermann 2019, Kleven et al. 2019a). We also asked children about their opinion on whether a man with a child below school age should work full-time, part-time, or not at all. Doing so allows us to assess and compare children's answers for both parents and to acknowledge that children may

¹³ For non-traditional statements such as (a) "*A working mother can establish just as warm and secure relationship with her children as a mother who does not work*" or (b) "*Both men and women should contribute to the household income*", the statistics refer to the share who "agrees" or "strongly agrees". For the remaining (traditional) statements such as (c) "*A pre-school child is likely to suffer if his or her mother works*" or (g) "*A man's job is to earn money, a woman's job is to look after the home and the family*", the statistics refer to the share who "strongly disagrees" or "disagrees".

¹⁴ Among the adult Spanish population participating in the ISSP 2012, we observe slightly more traditional views: while 93.2% agree or strongly agree with (b), 81.3% disagree or strongly disagree with (g), and in the case of (c) and (d) the respective shares are 47.3% and 38.7%.

¹⁵ In Spain children are granted public education from the year during which they turn 3 years old. As a result, 98% of all 3-year-olds are enrolled in school (Farré and Ortega 2018).

be egocentric and ideally want their parents to work less and spend more time with them all together. As shown in Table A.1, Panel B, most pre-reform children want both men and women with young children to work part-time (64% in the case of mothers, 73% in the case of fathers). Yet, there is still a sizable gap between the share of pre-reform children wanting mothers of young children to not work at all (25%) and the share wanting fathers of young children to not work at all (14%). Only a negligible fraction states full-time to be the preferred workload for men or women with young children (3% and 2%, respectively).

Finally, we introduced an incentivized elicitation method to get at children's perception of the prevailing social norm on mothers' and fathers' labor supply (see Part III of the Questionnaire in Appendix B). We adapted the incentivized coordination game developed by Krupka and Weber (2013) to be appropriate for children. We explained the game to the children as follows: *"You will now play a little game with one of your classmates without knowing who s/he is. In this game we will ask you both the same question. If you both give the same answer, you will both earn 4 points. If you do not give the same answer, no one will get any points."* After ensuring that the children understood the coordination aspect of the game using several control questions, we asked the children to rate on a 4-point Likert scale how *"socially appropriate"* it is that *"a woman with a child below school age works full-time"*, that *"a woman with a child below school age works part-time"*, that *"a father with a child below school age works part-time"*, and that *"a father with a child below school age does not work at all"*. We were very careful in explaining the meaning of *"socially (in)appropriate"* as *"a behavior that most people believe to be correct or good (incorrect or bad)"*. Children could choose on a scale ranging from *"appropriate"*, *"fairly appropriate"*, *"fairly inappropriate"* to *"inappropriate"*. Importantly, presenting these questions as an incentivized coordination game encourages children to consider not only their own opinions, but also the

opinions of others and how they align with each other. As such, these questions reveal their perception of the social norms prevailing among their classmates.

As shown in Table A.1 (Panel C), while most pre-reform children perceive it as socially appropriate (22%) or fairly appropriate (42 %) that a mother with a young child works part-time, the share is much lower when it comes to working full-time (4% and 6%, respectively). When it comes to fathers, 27% (46%) of pre-reform children settle with the answers (fairly) appropriate that a father with a young child works part-time, and still 34% (34%) with the answers (fairly) appropriate that a father with a young child does not work at all.

Turning to gender-stereotypical behaviors, we collected information on children's participation in household chores. We asked them how often they help at home with the following tasks: laundry, grocery shopping, small repairs, cleaning the house, and cooking. Answer categories range from "*at least once a week*", "*occasionally*" (less than once a week but more than once a month), "*almost never*", to "*never*" (see also Question VI.2 in the Questionnaire in Appendix B). Table A.1 (Panel D) shows the percentage of pre-reform children who participate "*at least occasionally*" in each task. Both boys and girls are equally likely to contribute to tasks such as doing the laundry or cooking (84% and 53%, respectively). In contrast, there are substantial gender differences regarding the other three tasks. Boys are much more likely than girls to help with small repairs (59% versus 34%) and grocery shopping (89% versus 79%), while the participation of girls in cleaning is much higher than that of boys (84% versus 66%).

Finally, to get as close as possible to children's future labor supply choices, we collected information on children's expectations about their own future family and work life. We asked them "*How do you see yourself in 20 years from now?*" (see Question IV.1 in the Questionnaire in Appendix B). Answer categories were "*working full-time and having children*", "*working part-time*

and having children”, “*not working and having children*”, “*working and not having children*” and “*not working and not having children*”. As shown in Table A.1, Panel E, the most popular choice among pre-reform boys and girls is to have children and work part-time (35% and 47%, respectively). Yet, there is a substantial gender gap when it comes to having children and working full-time (25% among the boys and only 13% among the girls). A substantial share of pre-reform children, both boys and girls, see themselves as not having children (34% while working and 2% while not working).

2.3 Sample

We exclude from the sample all children not born in 2006 or 2007 (167 children) and those not born in Spain (146 children). The reason for the second restriction is that the parents of children born abroad were not subject to the paternity leave reform. Our final sample contains 1,987 children born in Spain in 2006 or 2007.

Table A.2 shows the summary statistics for the full sample (see Column 1). Children in our sample are on average 13 years old and are almost equally split by gender (47% are boys). By construction, all children are born in Spain, but roughly one tenth has at least one parent born abroad (in 13% of the cases the mother and in 12% the father). Most children live with both parents, but a non-negligible share lives with their father only on some days (15%) or not at all (6%). Among the fathers, 67% work full-time and 24% work part-time, 5% do not work. Among the mothers, 50% work full-time, 37% work part-time, and 10% do not work. Half of the mothers went to college, and slightly fewer fathers did so (39%). Yet, a non-negligible share of the children does not know whether their mother or their father went to college (16% and 21%, respectively).

3. The Paternity Leave Reform

3.1. *Institutional background*

Paternity leave policies, designed to promote fathers' participation in childcare and foster gender equality in the home and in the labor market, are now prevalent in many countries. We focus on the introduction of paid paternity leave in Spain in 2007.

Prior to the reform, the Spanish labor market displayed a significant gender gap in employment patterns.¹⁶ In 2006, the employment rate among 30–45-year-old men reached almost 90%, while only two thirds of all women in this age range were working. At that time, Spain granted 6 weeks of compulsory maternity leave (at full pay), plus 2 days of paid job absence for fathers.¹⁷ In addition, families were granted 10 weeks of parental leave, also at full pay, which could be taken by mothers or fathers, or shared between them.¹⁸ Yet, as shown in Figure 1, very few fathers took parental leave (see the dashed line for Spain and the solid one for Catalonia, where our schools are located). In contrast, roughly two thirds of all mothers in Spain, and three quarters of all mothers in Catalonia were taking maternity leave and subsequently parental leave (dashed and solid blue line, respectively).¹⁹ Employment rates for mothers were low: only 55% of all women with children aged 0-2 were working (60% of all women with children aged 3 years and older).

¹⁶ All numbers in this section stem from <https://www.oecd.org/els/family/database.html> accessed on January 3, 2022.

¹⁷ The exception were public sector workers (roughly 15% of the total workforce), who enjoyed 10 days of paternity leave since 2005 (“Plan Concilia” implemented on December 27, 2005), and public employees of the Catalan government, who were entitled to 4 weeks of paternity leave since 2006 (Article 13 of the law 8/2006 implemented on July 5, 2006).

¹⁸ Parental leave was initially regulated by the Law 8/1980, March 10, (Estatuto de los Trabajadores 1980). Eligibility was tied to a formal work contract or to being officially registered as unemployed.

¹⁹ This seemingly low share can be explained by the low share of women under a formal contract in Spain in 2006 (66% of all women aged 30-45 years old).

At the end of 2006, the national parliament approved a bill (“*proyecto de ley*”) introducing a paid paternity leave permit. The law (“*ley orgánica 3/2007 de 22 de marzo, para la igualdad efectiva de mujeres y hombres*”) was published on March 23, 2007 and enacted immediately thereafter (the next day). The new 13-day paternity leave period (on top of the two days of paid job absence) was fully compensated, and it could be taken by fathers either at the same time or immediately after the mother’s leave period. New fathers were eligible starting from births taking place on March 24, 2007, if they were affiliated to Social Security and had worked for at least 180 days over the previous 7 years. As shown in Figure 1, take-up was high, with 47% (54%) of new fathers in Spain and 53% (61%) of new fathers in Catalonia using it in 2007 (2008) (dashed and solid red line, respectively).²⁰ The paternity leave permit was extended several times, thereafter, reaching 16 weeks in January 2021.

3.2. Paternity leave inducing a persistent shift in parental behavior

In this paper, we study the reduced form effect of the introduction of the 2-weeks paid paternity leave in Spain in 2007 on the gender norms, behaviors as well as employment and fertility expectations of the next generation. We interpret the paternity leave reform as a natural experiment shifting children’s direct exposure to counter-stereotypical behavior of parents. This interpretation is in line with the common aim of paternity leave policies to trigger persistent changes within families that go beyond the mere take-up of paternity leave. In what follows, we provide supportive evidence for the case of Spain.

Starting with mothers’ labor market outcomes, Farré and González (2019) show that women whose partners were eligible for paternity leave were more likely to be employed 6, 12, and even

²⁰The numbers reported for 2007 refer to new fathers whose child is born on or after March 24.

24 months after childbirth. The employment rate of mothers 6 months after having a child rose by up to 7% from a pre-reform employment rate of about 55%. In a follow-up analysis, we provide evidence that the effect on mothers' employment fades out in the medium- to long-run (see Appendix C, Table C.1 for details).

Fathers' labor market outcomes were unaffected by the two weeks of paternity leave. However, fathers' involvement in childcare activities increased not only after the birth of the child, but also several years later. Farré and González (2019) estimate an increase of almost an hour more childcare per day in 2009-10, a remarkable increase given an average involvement of 1.3 hours per day prior to the reform (according to the Spanish Time Use Survey in 2002-03). The increase did not come from reductions in housework, but rather from reductions in recreational time, such as leisure or sleep. Using our data, we show that even 12 years after childbirth, fathers were more likely to help at home (see Appendix C, Table C.2).

Strikingly, the reform also led to a delay in subsequent births. Farré and González (2019) show that couples eligible for paternity leave took up to 38 days longer to have their next child than ineligible couples (a delay of 3% considering the mean birth interval of about 1,250 days, or 3 and a half years). This delay resulted in fewer births in the following years, especially among older couples (with mothers being 30 years and older in 2007). In line with the decrease in fertility, there was a reduction in men's desired fertility leading to a reversal of the pre-reform pattern of men desiring more children than women. We interpret this result as an increase in men's awareness of the costs of child rearing.²¹

²¹ This interpretation is in line with recent evidence that in settings where men desire more children than women and may also enjoy higher bargaining power, providing information to fathers regarding the cost of having children can lower their desired fertility, as well as families' actual completed fertility (Ashraf et al. 2017).

There is also mounting international evidence on the persistent effects of paternity leave permits on the time that fathers devote to childcare and household chores. Early evidence from Norway shows that the introduction of a 4-weeks paternity leave in 1993 caused a decrease in fathers' earnings and an increase in fathers' time and effort invested in home production (Rege and Solli 2013).²² Kotsadam and Finseraas (2011) showed that this translated into a reduction in parental conflicts over division of labor and an increase in parents' likelihood to equally divide household chores. Similar results are available for Canada, where the introduction of a 5-weeks paternity leave in 2006 resulted in a less sex-specialized division of household chores (Patnaik 2019), and Germany, where the 2 daddy months introduced in 2007 boosted fathers' involvement in childcare beyond the leave period (Tamm 2019).

In sum, there is ample empirical evidence that paternity leave indeed causes a positive and persistent shift in fathers' engagement in childcare and in the home. Thus, children born after the introduction of paternity leave are likely to grow up in a more gender-neutral family model and to be exposed to less gendered stereotypical behaviors.

4. Empirical Strategy

Our identification strategy is based on the introduction of paternity leave in Spain applying to fathers of children born on March 24, 2007, or later. We employ a difference-in-differences model comparing children born shortly before and after the cut-off date in the reform (or treated) year,

²² Analyzing the same reform in Norway, Cools et al. (2015) find an increase in paternity leave take-up, but no effects on fathers' earnings. Focusing on the introduction of a daddy month in Sweden, Ekberg et al. (2013) do not find any change in fathers' likelihoods to take leave for sick children.

drawing upon children born in the same window of birthdates in a control year.²³ The treated cohort consists of children born in 2007, while the previous cohort (children born in 2006) serves as the control cohort. All children born on or after March 24, 2007, are possibly affected by the paternity leave reform. Using the same window of birthdates in 2006 allows us to net out potential age trends. Hence, we estimate the following equation:

$$Y_i = \beta_0 + \beta_1 Cohort2007_i + \beta_2 PostMarch24_i + \beta_3 Cohort2007_i * PostMarch24_i + X_i + \varepsilon_i \quad (1)$$

where Y_i stands for the gender norms, behaviors or expectations exhibited by child i . $Cohort2007_i$ is a dummy variable indicating whether child i is born in 2007 (versus 2006), and $PostMarch24_i$ represents a dummy variable taking the value 1 if child i is born on or after March 24 (independently of the calendar year). The interaction term $Cohort2007_i * PostMarch24_i$ takes value 1 for any child born after the introduction of paternity leave. We can thus interpret the coefficient β_3 as the intent-to-treat effect of paternity leave on children's gender norms, expectations, and behaviors.²⁴

We restrict the baseline sample to the 873 children born +/- 82 days around the cut-off date of March 24 (i.e., between January 1 and June 12).²⁵ Hence, we compare children who belong to the same school cohort and are thus subject to the same school cohort specific factors (e.g., the school curriculum). In our preferred specification, we control for individual background

²³ The natural experiment under study is obviously suited for a Regression Discontinuity Design (RDD). Given the limited sample size, only some of our outcomes are robust to an RDD specification, while we lack power when it comes to others (please refer to Section 5.2 and Table 2 for details). Similar identification strategies have been used by Lalive and Zweimüller (2009), Dustmann and Schönberg (2012), Danzer and Lavy (2018), and Schönberg and Ludsteck (2014) in the context of (other) parental leave reforms.

²⁴ To infer the effect on the treated from the intent-to-treat effect, we must consider the actual take-up rate of paternity leave (53% among new fathers in Catalonia in 2007).

²⁵ Ideally, we would include June 13. Yet, for data protection issues we could ask children only whether their birthday fell into windows of 8-12 days (day 1-12 of the month, day 13-23 and 24-31) and not for their exact birthdate.

characteristics (e.g., gender, date-of-birth fixed effects,²⁶ parental education and migration background) and for school fixed effects allowing us to abstract from selection into schools.

The identifying assumption is a common date-of-birth trend in our outcome variables across the two cohorts in the absence of the paternity leave reform. To allow for a concise illustration of the credibility of this assumption and to improve precision, we rely on aggregate measures of our main outcome variables. First, we reduce the dimensionality of the seven questions related to gender role attitudes using a principal component analysis and extract a “*gender role attitudes index*” using the standardized first component (see Table A.1 Panel A).²⁷ Second, we summarize children’s opinions about the preferred parental labor supply. Accordingly, we create an indicator that takes value 1 if reported that a mother with a child below school age should work part-time or full-time and 0 if she should not work. We refer to this indicator as “*Individual opinions about mothers*”. For fathers, the indicator takes value 1 if reported that a father with a child below school age should work part-time or not at all, and 0 if he should work full-time. We refer to this indicator as “*Individual opinions about fathers*” (see Table A.1, Panel B). Third, we create measures for children’s perception of the social norms prevailing in their classrooms. To that end, we aggregate children’s responses regarding the social appropriateness of mothers working full-time or part-time

²⁶ We make use of the information about the range of days when a child was born (day 1-12, 13-23 and 24-31 of the respective month). The date-of-birth fixed effects thus indicate whether a child is born in a window of 8-12 days within a calendar year and apply to all children born in the respective window independently of the year of birth (2006 or 2007).

²⁷ We first code the answers to each statement such that a value of 0 corresponds to the most traditional view and a value of 4 to the most non-traditional view. Accordingly, for questions (a) and (b) in Panel A of Table 1 we assign a value of 4 if the respondent “*strongly agrees*”, 3 if “*agrees*”, 2 if “*neither agrees nor disagrees*”, 1 if “*disagrees*” and 0 if “*strongly disagrees*”. For questions (c) to (g), a value of 0 if “*strongly agrees*”, 1 if “*agrees*”, 2 if “*neither agrees nor disagrees*”, 3 if “*disagrees*” and 4 if “*strongly disagrees*”. We then conduct a principal components analysis. The first principal component is positively loaded on all items and explains 26.1% of the overall variance. We standardized this first principal component to have mean 0 and standard deviation 1 and refer to it as gender role attitudes index.

and fathers working part-time or not at all. We refer to these indicators as “*Social norms about the mother*” and “*Social norms about the father*” (see Table A.1., Panel C). Employing a median split, we consider both answers “*appropriate*” and “*fairly appropriate*” for the social norms about working mothers, but only the answer category “*appropriate*” for the social norms applying to fathers. This different aggregation scheme results from children judging more harshly on mothers than on fathers when deviating from stereotypical gender roles (which is indeed the case when looking at children’s individual opinion about the optimal labor supply of women and men with children below school age, see Panel B).

We also create an indicator for children’s engagement in counter-stereotypical household tasks (see Table A.1., Panel D). Specifically, the indicator “*Counter-stereotypical behavior*” equals 1 if a boy engages “*at least occasionally*” in cleaning (a primarily female chore) and 0 otherwise, or if a girl engages “*at least occasionally*” in grocery shopping and small repairs (primarily male chores) and 0 otherwise (if she engages in either of the two chores, the variable equals 0.5). Finally, we summarize children’s expectations regarding future work and family life in an indicator measure that we refer as “*Counter-stereotypical expectations*” (see Table A.1., Panel E). The variable equals 1 if a girl expects to work full-time and have children (and 0 otherwise), or if a boy expects to not work full-time and have children (and 0 otherwise).

Figure 2 displays the date-of-birth trend (for all children born between January 1, 2006, and December 31, 2007) in our five main outcome variables. We observe the following patterns: first, younger children exhibit clearly more traditional gender role attitudes and norms than older children. Second, we observe a common trend for children born before and after the reform in gender role attitudes (Panel a) and social norms about mothers and fathers (Panel c and e). The same pattern also applies to children’s counter-stereotypical behavior and expectations (Panel f and g).

Yet, the pattern is less clear for children’s individual opinions regarding mothers’ and fathers’ preferred labor supply, which are much noisier (Panel b and d).

We also assess whether the pre-and post-reform children are balanced in terms of observable covariates. Table A.2 shows the pre-and post-cut-off date means for a series of individual background characteristics in our baseline sample, columns 2-3 for the children of the treated cohort (i.e. between January 1 and June 12, 2007) and columns 4-5 for the control cohort (i.e. between January 1 and June 12, 2006). Column 6 reports the difference-in-differences estimates for all individual background characteristics. Reassuringly, none of the estimates is significant at the conventional levels.

5. Results

5.1 Effects on children’s gender role attitudes, opinions, and norms

This section describes the estimated effects of the reform on our measures of gender role attitudes, opinions and norms using the empirical model in equation (1).²⁸ Table 1 displays the point estimates for β_3 , the coefficient on the interaction between the reform indicator (*PostMarch24*) and the dummy indicating the treated cohort (*Cohort2007*). Column 1 shows the coefficient when estimating equation (1) without controls, column 2 when adding individual background characteristics (date-of-birth fixed effects, gender, parental education, and migrant status), and column (3) when further including school fixed effects.

²⁸ We cluster standard errors at the class level to account for the sampling procedure as in Abadie et al. 2022.

Panel A reports the results for the reform effect on the gender role attitudes index, the individual opinion and the perceived social norm regarding the labor supply of mothers. Starting with the first row in Panel A of Table 1 and the most parsimonious specification (column 1), we find a positive (and marginally significant) reform effect on children's gender role attitudes index of 0.21 standard deviations. The magnitude of the coefficient increases to 0.26 standard deviations when controlling for individual background characteristics (column 2) and after including school fixed effects (column 3), with an accompanying improvement in precision. In what follows, we will only refer to the most conservative specification in column (3).

Turning to children's individual opinions, we find that the introduction of paternity leave exerts a strong effect on children's individual views about mothers: the share of children stating that a mother with a child below school age should work (either part- or full-time) increases by 14.6 percentage points. Given the share of pre-reform children supporting deviations from the stereotypical maternal behavior (66.4%), the estimated reform effect corresponds to an increase by 22% or 0.31 standard deviations in the support towards working mothers.

Children's perceptions regarding the social norm about working mothers are also affected by introduction of paternity leave: the share of children perceiving it as socially "appropriate" or at least "fairly appropriate" that a mother works (full-time or part-time) increases by 17.9 percentage points, corresponding to an increase by 28% (from a baseline share of 63.7%) or 0.37 standard deviations.²⁹

Panel B displays the reform effects on children's opinions and norms about the labor market behavior of men when having a child below school age. To begin with, we find no significant effect

²⁹ The incentivized coordination game was only introduced after the pilot phase, which is why the baseline specification only draws upon the data from the main data collection phase.

of the introduction of paternity leave on children's individual opinions regarding fathers' labor supply. To recall, in the eyes of most pre-reform children, fathers should work either part-time or not at all (87.2%). If we consider this extremely high baseline, it is not surprising that the introduction of paternity leave did not exert any effect at this margin. Turning to children's perception of the social norm, we find a significant increase in the share perceiving it as "socially appropriate" that a father works part-time or not at all. This share rises by 16.3 percentage points, which corresponds to an increase by 36% (from a baseline of 45.1%) or 0.33 standard deviations.

Overall, the results in Panel A and B of Table 1 provide compelling and robust evidence of substantial spillover effects of paternity leave on to the next generation. No matter which measures we look at, we find that the policy promoted gender-egalitarian role attitudes and norms among children of eligible parents.³⁰ In fact, the policy shifted norms applying to mothers and fathers to the same extent. Going one step further, parents seem to exert a strong influence on the formation of children's gender norms, that may go beyond that of peers at this early stage of life. If children were perfectly informed about their peers' gender norms, we would not be able to detect any effect on children's perception of the norms prevailing among their classmates. Yet, instead, children seem to internalize what they observe at home and draw conclusions from their own experiences about the experiences of others (a phenomenon known as "false consensus effect", Ross et al. 1977).

To put our findings into perspective, we compare our results to the ones of two recent studies on horizontal socialization mechanisms. The school-based randomized intervention in India studied by Dhar et al. (2021), that engaged adolescents for two and a half years in classroom discussions about gender equality, increased counter-stereotypical gender attitudes of 17-year-olds by 0.18

³⁰ We also studied the effects of the policy separately for boys and girls. The resulting estimates are not significantly different from each other and are available upon request.

standard deviations. Studying the impact of classroom composition on gender role attitudes of 15-year-olds in Vietnamese schools, Garcia-Brazales (2021) finds that a 10-percentage point increase in the proportion of female classmates decreases traditional views by about 0.15 standard deviations. Thus, our study lines up with a recent literature showing that gender role attitudes among adolescents are amenable to change and contributes to this literature by shedding novel light on the power of public policies to shape social norms by affecting parental behavior.

5.2 Robustness

Table 2, Panel A and B report the results from a battery of robustness tests regarding children's gender role attitudes, opinions, and norms. For comparison, Table 2 Column 1 repeats the estimates from our most preferred specification, which includes individual control variables and school fixed effects. It replicates the estimates in Table 1, column 3.

When analyzing the effects of a policy change, one major concern is anticipation or selective sorting. Our treated sample only comprises children born between January 1 and June 12, 2007, and thus conceived in September 2006 or earlier. The policy change was discussed in parliament and the media earliest by mid-December 2006. Thus, strategic planning of the pregnancy to enjoy paternity leave within our sample is unlikely to impossible. Postponement of the actual birth is difficult and given medical concerns not possible for more than a couple of weeks. Yet, to take this possibility into account, we estimate our baseline specification using a sample where we exclude the births right around the cut-off date (between March 13 and 31). Results are robust and shown in Table 2, Panel A and B, column 2.

Fathers were entitled to paternity leave for any child born on or after March 24, 2007. As such, children may have directly benefitted from their father taking paternity leave after their own

birth as well as indirectly after the birth of a younger sibling. The latter also applies to children born prior to the reform. To ensure a “clean” control group, we estimate our baseline specification using a sample excluding all children born prior to the reform with younger siblings. Results on gender role attitudes and social norms are robust and indicate, if anything, stronger effects (Table 2, Panel A and B, column 3). The results on individual opinions, however, loose in size and precision.³¹

We probe robustness to two further alternative sample specifications, first excluding the data collected during the pilot phase (data from 3 schools, resulting in a sample of 754 observations) and second including all children born in 2006 and 2007 (resulting in a sample of 1987 observations for attitudes and individual norms, and 1708 for social norms). The resulting estimates are robust (except for the individual opinions regarding mothers’ labor force participation), but slightly less precise (see Table 2, Panel A and B, column 4 and 5, respectively).

We also investigate the robustness of our estimates to alternative estimation strategies. First, we include a set of class fixed effects considering potential sorting into classes and spillover effects within classes (which if anything should bias our results downward). Results are extremely robust and shown in Table 2, Panel A and B, column 6. Second, we exploit the cut-off date and employ a regression discontinuity design (RDD). Specifically, we add the running variable (relative age within each cohort) to our difference-in-differences specification which results in a so-called RDD-DiD specification. Again, results are remarkably robust to this specification, both in terms of magnitude and precision (see Table 2, Panel A and B, Column 7). We then turn to a classical RDD design using the full sample (children born in 2006 and 2007) and controlling for the running

³¹ It would be interesting to analyze whether the effects vary depending on direct or indirect or even multiple experiences to paternity leave. Unfortunately, our sample is not big enough to conduct such heterogeneity analysis.

variable as a first order polynomial (Table 2, Panel A and B, column 8).³² We find robust estimates for the reform effect on individual opinions and social norms. The estimate for gender role attitudes, however, loses in magnitude (by up to 50%) and precision.

We have further probed the robustness of our results when using alternative aggregation schemes to generate the gender role attitudes index. Table 2, Panel A and B, column 9 shows the difference-in-differences results when using a principal component analysis on all 7 original items from the ISSP questionnaire and 2 further items that we added to the questionnaire.³³ Table 2, Panel A and B, column 10 shows the results when using the unweighted average of the answers given to the 7 original items taken from the ISSP questionnaire and column 11 when using the weighted average with weights constructed by normalizing the variables to have the same standard deviation and then recovering the weights from the inverse covariance matrix (Anderson 2008).

We also assess the sensitivity of our estimates to acknowledge possible alternative correlation structures between potential outcomes not only within classes, but also within schools (see Table A.3, Column 2). Moreover, we consider the close topicality of our outcome variables and adjust for two multiple hypothesis testing (see Table A.3, Column 3 and 4). Results are robust across all specifications.

Finally, we conduct a series of placebo tests. First, we estimate our baseline specification using a placebo outcome, children's perception of the social norm on cheating in an exam. We introduced the question on the social appropriateness of cheating in an exam merely to check

³² We use the `rdrobust` command in Stata and rely on all children born in 2006 and 2007 to determine the optimal bandwidth.

³³ We asked children to rate on a 5-point Likert scale whether they agreed to the following two statements: "When a woman earns more than her husband, there are certainly problems" and "Both mother and father should take a parental leave permit from work after the birth of their son or daughter".

whether the children had understood the coordination game (see questions III.3 in the Questionnaire in Appendix B). The introduction of paternity leave and the subsequent counter-stereotypical behavior of parents should have left the perception of the social norm on cheating unaffected. Indeed, as shown in Figure 3a, the reform effect is indistinguishable from zero. Second, we replicate our baseline specification using alternative placebo cut-off dates when no reform took place. Specifically, we use samples of +/- 82-days-windows around a series of hypothetical placebo cut-off dates in 2007, starting from June 12 (the first possible date not including any of the children belonging to the original pre-reform children) until October 12 (the last placebo cut-off date to include children exclusively belonging to the same school cohort). Out of the 64 placebo estimations only one is significant at the 5% significance level and two at the 10% significance level (see Figure 3b, for gender role attitudes, individual opinions and social norms on mothers' labor supply and Figure A.2 for all other outcome variables). This battery of robustness check reinsures the validity of our findings.

5.3 Effects on children's behaviors and expectations

An open question is to which extent the paternity leave reform has the power to induce gender-egalitarian behaviors and choices contributing to a closure of the gender gap in the future. Children in our sample are obviously too young to explore the effects of the reform on their labor market outcomes. Yet, the collected information on children's day-to-day contributions to several household chores can enhance our understanding of whether the paternity leave reform may shift gender specialization at home already at a young age. Information on children's expectations regarding their own future employment and family plans allows us to get a first glimpse on the potential effects of paternity leave on the future gender gap in labor market outcomes.

Table 1, Panel C reports the reform effect on children's engagement in counter-stereotypical household chores. To recall, counter-stereotypical behavior refers to boys engaging at least occasionally in primarily female chores (i.e., cleaning) and girls engaging in primarily male chores (i.e., small repairs and grocery shopping). According to our most preferred specification shown in column 3, if fathers are eligible for paternity leave, children are 14.2 percentage points more likely to engage at least occasionally in counter-stereotypical household chores. This reform effect corresponds to an increase by 24% or 0.35 standard deviations (from a pre-reform mean of 60.3%, with a standard deviation of 0.40).

Breaking the results down by gender (Table 3, Panel A and B) reveals that girls are more likely to contribute to male chores, such as going grocery shopping and doing small repairs. In our most preferred specification in column 3, the reform effect corresponds to an increase by 12.5 percentage points or 22%. Boys' contribution to male chores, in contrast, remains unchanged. Note, however, the gender-specific estimates regarding the contribution to male chores are not significantly different from each other (p-value when testing the null hypothesis that both estimates are equal equals 0.168). Boys, in turn, are more likely to contribute to female chores, such as cleaning. The reform effect corresponds to 15 percentage points or 23% and significantly differs from the reform effect on girls' contribution to female chores (which decreases by 6.4 percentage points, yet not significantly though).

Turning to children's family and labor market expectations, we find that children born after the reform are 15.5 percentage points more likely to expect deviating from the traditional male-breadwinner model (see Table 1, Panel D, column 3). To recall, this implies for girls to expect working full-time and having children and for boys not working full-time and having children. The

reform effect corresponds to an increase by 42% from the pre-reform mean or 0.32 standard deviation.

Table 3, Panel C reports the results for family and labor market expectations separately for boys and girls. The results exhibit a gender-specific pattern. Boys whose father was eligible for paternity leave are 19.8 percentage points or 80% less likely to see themselves working full-time and having children. This effect is quite sizeable, mostly because of the relatively low share of pre-reform boys planning to work full-time and having children in the future (24.7%).³⁴ Girls, in contrast, are more likely to see themselves working full-time and having children. The estimated reform effect corresponds to an increase in 8.1 percentage points or 60%, again a sizeable effect driven by the relatively low share of pre-reform girls planning to work full-time and having children. Yet, the reform effect for girls is not statistically significant at the conventional levels.³⁵

Overall, the results provide strong and robust evidence that paternity leave is a powerful tool to shift not only the gender role attitudes and norms of the next generation, but to induce gender convergence both at home and in the labor market. As such, this policy promises to sustainably combat the gender gap in the labor market and to close the child penalty in the long-run.

6. Conclusions

Despite remarkable improvements, gender gaps in employment and earnings are still sizable and persistent in all countries. Much of the current debate on the mechanisms underlying the remaining differentials deals with the existence of social norms on the role of women in childcare and home

³⁴ We do not find any effect on children's fertility intentions. The results are available upon request.

³⁵ The entire battery of robustness tests – alternative sample specifications (i.e., excluding the observations around the cut-off, excluding control children with younger siblings, excluding the pilot sample, using the full sample of children born in 2006 and 2007) and alternative estimation strategies (i.e., additional class fixed effects, the RDD-DD, a classical RDD) is shown in Table 2, Panel C and D.

production. Many scholars go as far as proclaiming gender norms to constitute an important barrier to gender equality that render family policies largely ineffective (Bertrand 2020, Cortes and Pan 2020, Kleven et al. 2021 and Albanesi et al. 2022). Yet, gender norms are very persistent and transmitted from one generation to the next.

We study the extent to which family policy can shape gender norms and induce gender equality in the next generation, by promoting counter-stereotypical behaviors among parents. We exploit the introduction of paternity leave in Spain as an exogenous shock that increased fathers' participation in childcare and housework persistently. We conduct a large, targeted lab-in-the-field experiment with 11-13-years old children to elicit their attitudes towards gender roles, their individual opinions and perceived social norms on working mothers and fathers. We also ask them about their engagement in day-to-day behaviors and expectations regarding their future labor market participation. We follow a local difference-in-differences approach that compares children (in the same school class) born around the date of the introduction of paternity leave on March 24, 2007, using the previous cohort (children born in 2006) as controls. We find that children whose fathers were eligible for paternity leave display significantly more gender egalitarian norms and behaviors at age 11-13. Children born after the reform also report less stereotypical expectations regarding their own labor market involvement after parenthood.

Our results speak to the presence of long-run effects of public policies operating through slow-moving changes in preferences and norms that existing quasi-experimental studies have missed (Kleven et al. 2021). Paternity leave permit may have the power to foster gender convergence in the future by shaping the gender norms of the next generation and thus removing a key obstacle on the path to gender equality.

References

- Abadie, Alberto, Susan Athey, Guido Imbens, and Jeffrey Wooldridge. "When Should You Adjust Standard Errors for Clustering" mimeo (2022)
- Aksoy, Cevat G., Christopher S. Carpenter, Ralph De Haas, and Kevin D. Tran. "Do laws shape attitudes? Evidence from same-sex relationship recognition policies in Europe." *European Economic Review* 124 (2020): 103399.
- Albanesi, S., Olivetti, C. and B. Petrongolo (2022). "Families, Labor Markets and Policy" manuscript prepared for the Handbook of Family Economic
- Alesina, Alberto, Paola Giuliano, and Nathan Nunn. "On the origins of gender roles: Women and the plough." *The Quarterly Journal of Economics* 128.2 (2013): 469-530.
- Andresen, Martin E., and Emily Nix. "What causes the child penalty? Evidence from same sex couples and policy reforms." *Journal of Labor Economics* 40.4 (2022).
- Anderson, Michael L. "Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry preschool, and early training projects." *Journal of the American Statistical Association* 103.484 (2008): 1481-1495.
- Angelov, Nikolay, Per Johansson, and Erica Lindahl. "Parenthood and the Gender Gap in Pay." *Journal of Labor Economics* 34:3 (2016): 545-579
- Ashraf, N., E. Field, G. Rusconi, A. Voena, and R. Ziparo. "Traditional Beliefs and Learning about Maternal Risk in Zambia." *American Economic Review* 107.5 (2017): 511-515
- Bastian, Jacob. "The rise of working mothers and the 1975 earned income tax credit." *American Economic Journal: Economic Policy* 12.3 (2020): 44-75.
- Bertrand, Marianne, Patricia Cortes, Claudia Olivetti, and Jessica Pan. "Social norms, labour market opportunities, and the marriage gap between skilled and unskilled women." *The Review of Economic Studies* 88.4 (2021): 1936-1978.
- Bertrand, Marianne. "The gender socialization of children growing up in nontraditional families." *AEA P&P* 109 (2019): 115-121.
- Bertrand, Marianne. "Gender in the twenty-first century." *AEA P&P* 110 (2020): 1-24.
- Bertrand, Marianne, Emir Kamenica, and Jessica Pan. "Gender identity and relative income within households." *The Quarterly Journal of Economics* 130.2 (2015): 571-614.
- Boelmann, Barbara, Anna Raute, and Uta Schönberg. "Wind of Change? Cultural Determinants of Maternal Labor Supply." (2021).
- Brollo, Fernanda, Katja M. Kaufmann, and Eliana La Ferrara. "Learning Spillovers in Conditional Welfare Programs: Evidence from Brazil." *The Economic Journal* 130.628 (2020): 853-879.
- Brouwer, Thijs, Fabio Galeotti, and Marie Claire Villeval (2022) "Teaching Norms: Direct Evidence of Parental Transmission." Working Papers hal-03725575
- Buckles, Kasey S., and Daniel M. Hungerman. "Season of birth and later outcomes: Old questions, new answers. " *Review of Economics and Statistics* 95.3 (2013): 711-724.

- Bursztyn, Leonardo, Alessandra González, and David Yanagizawa-Drott. "Misperceived social norms: Women working outside the home in Saudi Arabia." *American Economic Review* 110.10 (2020): 2997-3029
- Carter, D. and Patterson, C. "Sex roles as social conventions: The development of children's conceptions of sex-role stereotypes." *Developmental Psychology*, 18.6 (1982)
- Clarke, Damian, Joseph P. Romano, and Michael Wolf. "The Romano-Wolf Multiple Hypothesis Correction in Stata." *The Stata Journal* 20(4) (2020): 812-843.
- Cools, S., J. H. Fiva, and L. J. Kirkebøen. "Causal Effects of Paternity Leave on Children and Parents." *Scandinavian Journal of Economics* 117.3 (2015): 801–828
- Cortes, P. and Pan, J. "Children and the Remaining Gender Gaps in the Labor Market." NBER Working Paper 27980 (2020)
- Dahl, Gordon B., and Anne C. Gielen. "Intergenerational spillovers in disability insurance." *American Economic Journal: Applied Economics* 13.2 (2021): 116-150.
- Dahl, Gordon B., Andreas Kotsadam, and Dan-Olof Rooth. "Does integration change gender attitudes? The effect of randomly assigning women to traditionally male teams." *The Quarterly Journal of Economics* 136.2 (2021): 987-1030.
- Dahl, Gordon B., Katherine V. Løken, and Magne Mogstad. "Peer Effects in Program Participation." *American Economic Review* 104.7 (2014): 2049-2074.
- Danzer, Natalia, and Victor Lavy. "Paid parental leave and children's schooling outcomes." *The Economic Journal* 128.608 (2018): 81-117.
- Dhar, Diva, Tarun Jain, and Seema Jayachandran. "Reshaping adolescents' gender attitudes: Evidence from a school-based experiment in India." NBER Working Paper Series No. 25331 (2021).
- Doepke, Matthias, and Fabian Kindermann. "Bargaining over babies: Theory, evidence, and policy implications." *American Economic Review* 109.9 (2019): 3264-3306.
- Dunatchik, Allison, and Berkay Özcan "Reducing mommy penalties with daddy quotas." *Journal of European Social Policy* 31.2 (2020): 175-191.
- Dustmann, Christian, and Uta Schönberg. "Expansions in maternity leave coverage and children's long-term outcomes." *American Economic Journal: Applied Economics* 4.3 (2012): 190-224.
- Ekberg, John, Rickard Eriksson and Guido Friebel. "Parental leave - A policy evaluation of the Swedish "Daddy-Month" reform." *Journal of Public Economics* 97 (2013): 131-143.
- Farré, Lúdia, and Francis Vella. "The intergenerational transmission of gender role attitudes and its implications for female labour force participation." *Economica* 80.318 (2013): 219-247.
- Farré, Lúdia, and Libertad González. "Does paternity leave reduce fertility?" *Journal of Public Economics* 172 (2019): 52-66.
- Farré, Lidia, Francesc Ortega, and Ryuichi Tanaka. "Immigration and the public–private school choice." *Labour Economics* 51 (2018): 184-201.

- Fernández, Raquel, Alessandra Fogli, and Claudia Olivetti. "Mothers and sons: Preference formation and female labor force dynamics." *The Quarterly Journal of Economics* 119.4 (2004): 1249-1299.
- Daniel, Fernández-Kranz, Aitor Lacuesta, and Núria Rodríguez-Planas. "The motherhood earnings dip: Evidence from administrative records." *Journal of Human Resources* 48.1 (2013): 169-197.
- Fortin, Nicole M. "Gender role attitudes and women's labor market participation: Opting-out, aids, and the persistent appeal of housewifery." *Annals of Economics and Statistics/Annales d'Économie et de Statistique* 117/118 (2015): 379-401.
- Garcia-Brazales, Javier. "Unlearning Traditionalism: The Long-run effects of Schools on Gender Attitudes." CEMFI Manuscript (2021).
- Giuliano, Paola. "Gender: A Historical Perspective". In Susan L. Averett, Laura M. Argys, and Saul D. Hoffmann editors: *The Oxford Handbook of Women and the Economy*. Oxford University Press (2018).
- González, Libertad, and Hosny Zoabi. "Does Paternity Leave Promote Gender Equality within Households?" CESifo Working Paper No. 9430 (2021).
- Goussé, Marion, Jacquemet, Nicolas and Jean-Marc Robin. "Marriage, Labor Supply, and Home Production." *Econometrica* 85.6 (2017): 1873-1919
- Grewenig, Elisabeth, Philipp Lergetporer, and Katharina Werner. "Gender norms and labor-supply expectations: Experimental evidence from adolescents." CESifo Working Paper No. 8611 (2020).
- Hara, Hiromi and Núria Rodríguez-Plana. "Long-Term Consequences of Teaching Gender Roles: Evidence from Desegregating Industrial Arts and Home Economics in Japan". IZA Working Paper 14611 (2021).
- Kleven, Henrik, Camille Landais, and Jakob E. Søgaaard. "Children and gender inequality: Evidence from Denmark." *American Economic Journal: Applied Economics* 11.4 (2019): 181-209.
- Kleven, Henrik, Camille Landais, Johanna Posch, Andreas Steinhauer, and Josef Zweimüller. "Child penalties across countries: Evidence and explanations." *American Economic Review P&P* 109 (2019): 122-126.
- Kleven, Henrik, Camille Landais, Johnna Posch, Andreas Steinhauer and J. Zweimuller. "Do Family Policies Shape the Evolution of Gender Inequality? Evidence from 60 Years of Policy Experimentation" NBER Working Paper No. 28082 (2021).
- Kohlberg, Lawrence. "Moral stages and moralization: The cognitive-development approach." *Moral development and behavior: Theory research and social issues* (1976): 31-53.
- Kotsadam, Andreas, and Henning Finseraas. "The state intervenes in the battle of the sexes: Causal effects of paternity leave." *Social Science Research* 40 (2011): 1611-1622
- Krupka, Erin, and Roberto A. Weber. "Identifying social norms using coordination games: Why does dictator game sharing vary?" *Journal of the European Economic Association* 11.3 (2013): 495-524.

Lalive, Rafael, and Josef Zweimüller. "How does parental leave affect fertility and return to work? Evidence from two natural experiments." *Quarterly Journal of Economics* 124.3 (2009): 1363-1402.

Markus, Hazel, and Paula Nurius. "Possible selves." *American psychologist* 41.9 (1986): 954.

Neimanns, Erik. "Unequal benefits – diverging attitudes? Analysing the effects of an unequal expansion of childcare provision on attitudes towards maternal employment across 18 European countries." *Journal of Public Policy* 41.2 (2021): 251-276.

Olivetti, Claudia, Eleonora Patacchini, and Yves Zenou. "Mothers, peers, and gender-role identity" *Journal of the European Economic Association* 18.1 (2020): 266-301.

Patnaik, Ankita. "Reserving time for daddy: The consequences of fathers' quotas." *Journal of Labor Economics* 37.4 (2019): 1009-1059.

Petts, Richard J., Chris Knoester, and Jane Waldfogel. "Fathers' paternity leave-taking and children's perceptions of father-child relationships in the United States." *Sex Roles* 82.3 (2020): 173–188.

Rege, Mari, and Ingeborg F. Solli. "The impact of paternity leave on fathers' future earnings." *Demography* 50.6 (2013): 2255-2277.

Ross, Lee, David Greene, and Pamela House. "The "false consensus effect": An egocentric bias in social perception and attribution processes." *Journal of experimental social psychology* 13.3 (1977): 279-301.

Schönberg, Uta, and Johannes Ludsteck. "Expansions in maternity leave coverage and mothers' labor market outcomes after childbirth." *Journal of Labor Economics* 32.3 (2014): 469-505.

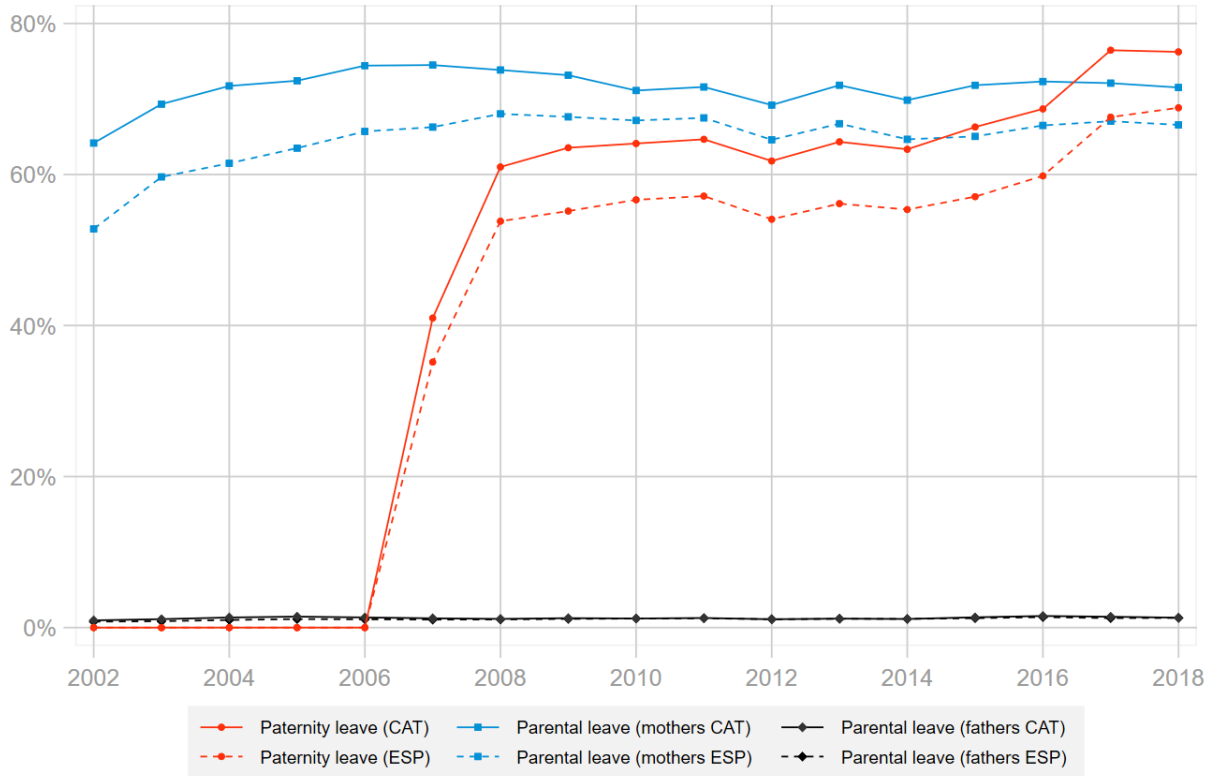
Tamm, Marcus. "Fathers' parental leave-taking, childcare involvement and labor market participation." *Labour Economics* 59 (2019): 184-197.

OECD. "From premises to action: Addressing discriminatory social institutions to accelerate gender equality in G-20 countries." 2021

Waylen, A. and Wolke, D. "Sex'n'drugs'n'rock'n'roll: the meaning and social consequences of pubertal timing." *European Journal of Endocrinology*, 151.3 (2004): 151-159.

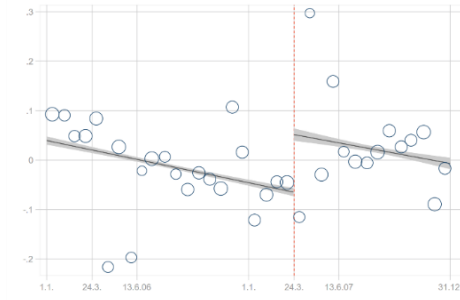
Figures and Tables

Figure 1: Take-up of parental leave and paternity leave in Catalonia and Spain

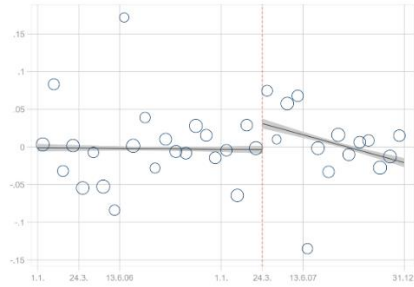


Note: The blue lines represent the percentage of mothers on paid parental leave, the black lines represent the percentage of fathers who used at least some weeks of the shared parental leave permit, and the red lines represents the percentage of fathers on paternity leave. The solid lines refer to the take-up rates in Catalonia while the dashed lines to that in Spain. Data are obtained from the administrative registers of the Spanish National Security System for several years.

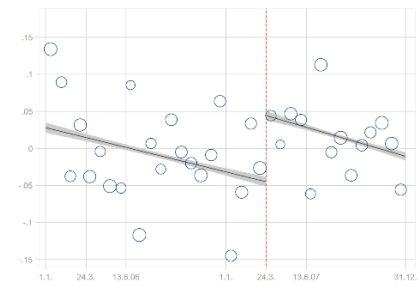
Figure 2: Date-of-birth trends in children's gender attitudes, norms, behavior, and expectations



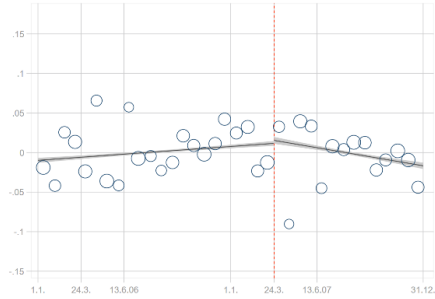
(a) Gender role attitudes index



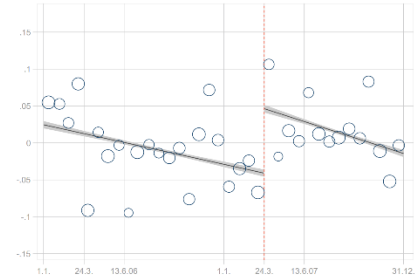
(b) Individual opinions about mothers



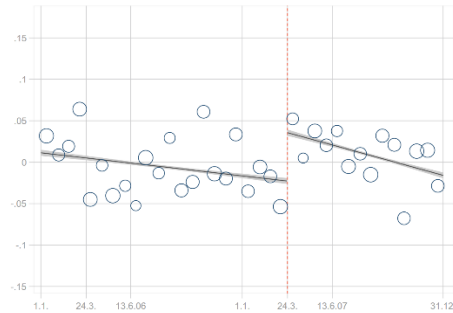
(c) Social norms about mothers



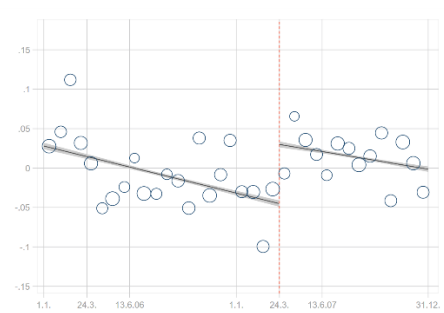
(d) Individual opinions about fathers



(e) Social norms about fathers



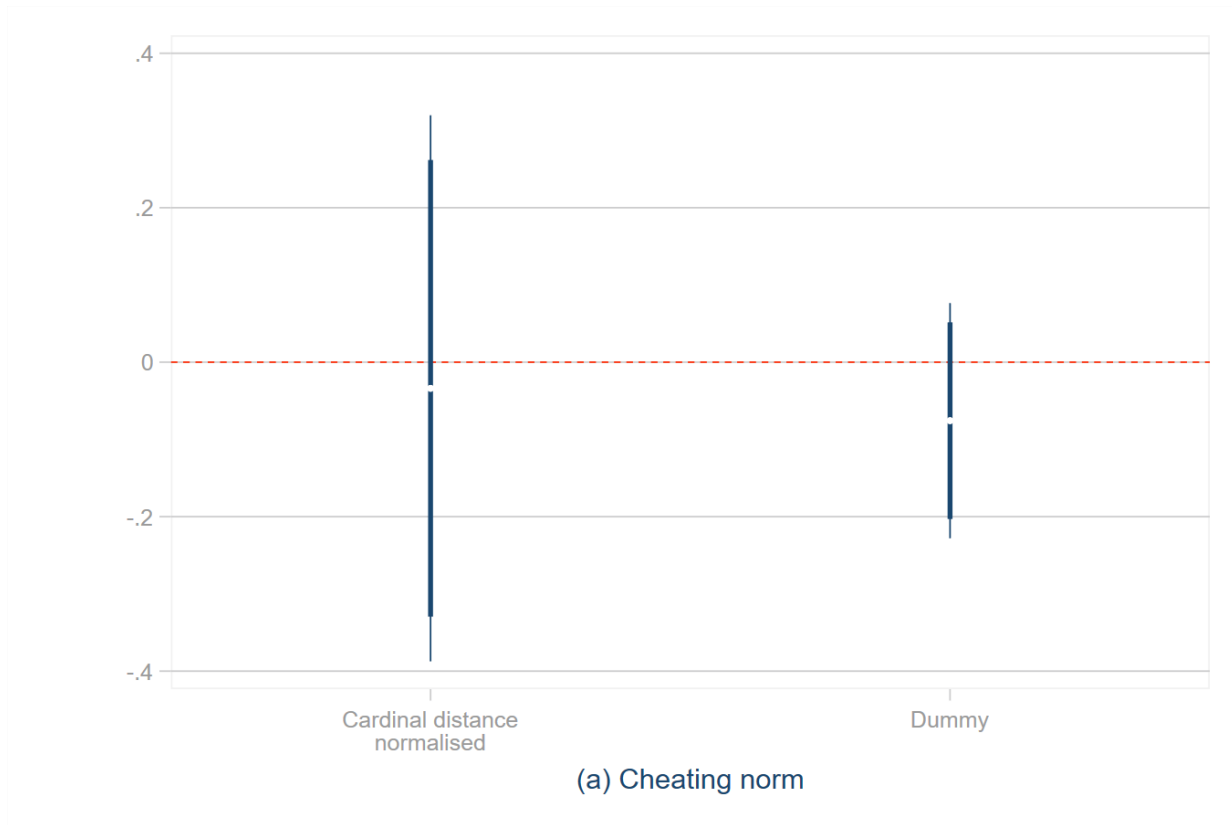
(f) Counter-stereotypical behavior



(g) Counter-stereotypical expectations

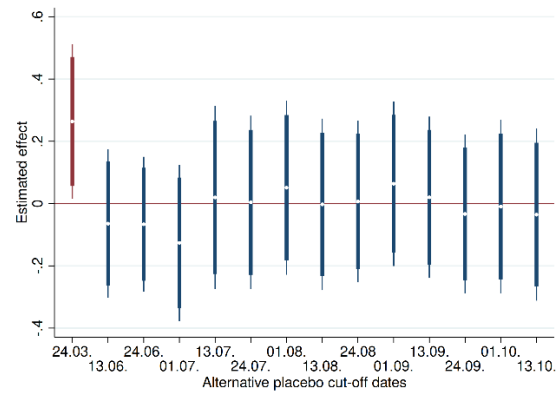
Note: Each figure plots the OLS residuals from our preferred specification in equation (1) excluding the interaction between the *Cohort2007* and the *Post March 24* indicators and using children born between 01.01.2006 to 31.12.2007. Each dot represents the average of the residuals over two consecutive birthdate bins. The size of the dot indicates the number of observations per bin. The red line indicates the date of the reform introduction (i.e. March 24, 2007). The Figures also display the linear prediction of the residuals to the left and the right of the cut-off date of the reform (March 24, 2007). The shaded area indicates the 95% confidence interval.

Figure 3a: Placebo tests with alternative measures

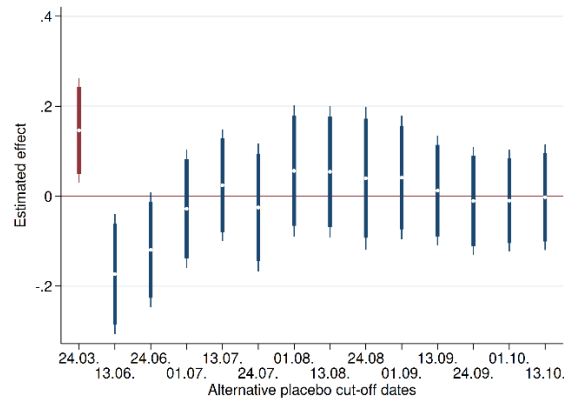


Note: The figure shows the reform effect on the perception of the social norm regarding “cheating during an exam” using our preferred specification (controlling for individual characteristics and school fixed effects) and sample (+/- 82 days around the cut-off date). Specifically, children are asked “*Copying in an exam is: appropriate, fairly appropriate, fairly inappropriate or inappropriate*”. Note that we show the reform effect using two alternative ways of aggregating children’s answers. On the left, we show the result when using the four answer categories (“inappropriate” as 0, “fairly inappropriate” as 1, “fairly appropriate” as 2, and “appropriate” as 3) and standardizing the variable to have a mean 0 and a standard deviation 1 for pre-reform children. On the right, we show the results when creating a dummy being equal to 0 for the category “inappropriate”, and 1 otherwise. The latter aggregation scheme corresponds to a median split and thus considers the fact that most children choose the answer category “inappropriate”.

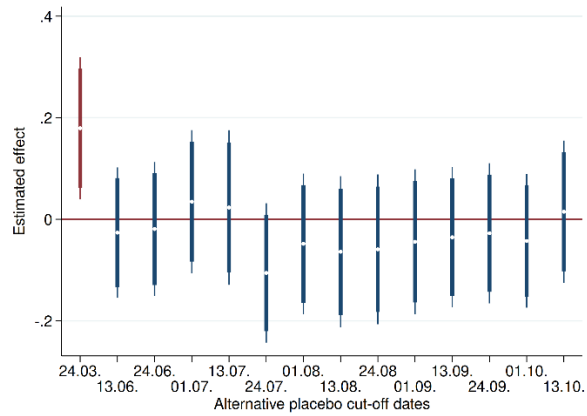
Figure 3b: Placebo tests for alternative cut-off dates



(a) Gender role attitudes index



(b) Individual opinions about mothers



(c) Social norms about mothers

Note: The figure shows a series of placebo reform effects for gender role attitudes (a), individual opinions (b) and social norms on mothers' labor supply (c) using our preferred specification (estimating equation (1) using a ± 82 days window around alternative cut-off dates and controlling for individual characteristics and school fixed effects). The red bar shows the estimated reform effect and the respective 95% confidence interval using the actual cut-off date (March 24, 2007), while the 13 blue bars show the estimated placebo effects when shifting the cut-off dates successively further (starting with the first possible cut-off date on June 12, 2007 until the last possible cut-off date – to only consider children belonging to the same school cohort – on October 12). The results for the other main outcomes variables: opinions and norms on fathers' labor supply and counter-stereotypical behavior and expectations are displayed in Figure A.2.

Table 1: Reform effects on children's gender role attitudes, norms, behavior, and expectations

| | (1) | (2) | (3) |
|---|----------|----------|----------|
| Panel A: Children's norms about mothers | | | |
| Gender role attitudes index | 0.205* | 0.263** | 0.264** |
| Pre-reform mean [sd] = 0 [1] | (0.123) | (0.123) | (0.125) |
| <i>N</i> | 873 | 873 | 873 |
| Individual opinions about mothers | 0.109* | 0.151*** | 0.146** |
| Pre-reform mean [sd] = 0.664 [0.473] | (0.057) | (0.057) | (0.058) |
| <i>N</i> | 873 | 873 | 873 |
| Social norms about mothers | 0.163** | 0.180*** | 0.179** |
| Pre-reform mean [sd] = 0.637 [0.482] | (0.065) | (0.068) | (0.070) |
| <i>N</i> | 754 | 754 | 754 |
| Panel B: Children's norms about fathers | | | |
| Individual opinions about fathers | 0.023 | 0.034 | 0.027 |
| Pre-reform mean [sd] = 0.872 [0.334] | (0.047) | (0.047) | (0.048) |
| <i>N</i> | 873 | 873 | 873 |
| Social norms about fathers | 0.163** | 0.169** | 0.163** |
| Pre-reform mean [sd] = 0.451 [0.499] | (0.077) | (0.078) | (0.080) |
| <i>N</i> | 754 | 754 | 754 |
| Panel C: Children's counter-stereotypical behavior | | | |
| Engagement in counter-stereotypical chores | 0.145*** | 0.141*** | 0.142*** |
| Pre-reform mean [sd] = 0.603 [0.401] | (0.050) | (0.051) | (0.052) |
| <i>N</i> | 750 | 750 | 750 |
| Panel D: Children's counter-stereotypical expectations | | | |
| Full-time work and children | 0.176** | 0.145** | 0.155*** |
| Pre-reform mean [sd] = 0.367 [0.483] | (0.082) | (0.061) | (0.057) |
| <i>N</i> | 754 | 754 | 754 |
| Individual controls | | Yes | Yes |
| School FE | | | Yes |

Note: Each coefficient stems from a separate regression and corresponds to the OLS estimate of the coefficient on the interaction term (Cohort 2007 * Post-March23) in equation (1). Column (1) shows the results without controls, column (2) when controlling for individual background characteristics (i.e. gender, date-of-birth fixed effects, parental education and migration background), and column (3) when adding school fixed effects. The *gender role attitudes index* is the first element resulting from a principal component analysis of the seven ISSP questions on gender roles. The index has been standardized to have mean 0 and standard deviation 1 for pre-reform children. *Individual opinions about mothers* is a dummy variable that equals 1 if the child believes that a mother with a child below school age should work full-time or part-time and 0 otherwise. *Social norms about mothers* is a dummy variable that equals 1 if the child answers that it is "socially appropriate" or "fairly socially appropriate" that a mother with a child below school age works full-time or part-time, and 0 otherwise. *Individual opinions about fathers* is a dummy variable that equals 1 if the child believes that a father with a child below school age should work part-time or not at all, and 0 otherwise. *Social norms about fathers* is a dummy variable that equals 1 if the child answers that it is "socially appropriate" that a father with a child below school age works part-time or not at all, and 0 otherwise. *Counter-stereotypical behavior* is the weighted average of whether children participate at least occasionally with a counter-stereotypical household chore. *Counter-stereotypical expectations* is a dummy that is 1 for girls (boys) when they (not) expect to work full-time and have children, and 0 otherwise. The lower number of observations in some questions is because they were not collected in the pilot sample. Standard errors are clustered at the session/class level and shown in parentheses: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 2: Robustness checks for the reform effects on children's gender role attitudes, norms, behavior, and expectations

| | Baseline | Alternative sample specifications | | | | Alternative estimation strategies | | | Alternative indices | | |
|--|---------------------|-----------------------------------|--------------------|---------------------|---------------------|-----------------------------------|---------------------|---------------------|---------------------|--------------------|--------------------|
| | (1) Pref spec. | (2) Donut | (3) Excl. sibl. | (4) Excl. pilot | (5) Full smpl. | (6) Class FE | (7) RDD DiD | (8) RDD P1 | (9) PCA 9Q | (10) Equal 7Q | (11) ICW 7Q |
| Panel A: children's gender role attitudes, opinions and norms about mothers | | | | | | | | | | | |
| Gender roles attitudes | 0.264** (0.125) | 0.324** (0.130) | 0.348** (0.157) | 0.287** (0.142) | 0.187* (0.103) | 0.241* (0.129) | 0.266** (0.128) | 0.182 (0.144) | 0.244** (0.122) | 0.362** (0.170) | 0.265** (0.126) |
| <i>N</i> | 873 | 760 | 553 | 754 | 1987 | 869 | 873 | 1987 | 873 | 873 | 873 |
| Individual opinions | 0.146** (0.058) | 0.128** (0.062) | 0.079 (0.081) | 0.123* (0.065) | 0.033 (0.049) | 0.163** (0.062) | 0.153** (0.059) | 0.106* (0.062) | | | |
| <i>N</i> | 873 | 760 | 553 | 754 | 1987 | 869 | 873 | 1987 | | | |
| Social norms | 0.162** (0.069) | 0.180** (0.077) | 0.196* (0.103) | 0.162** (0.069) | 0.139*** (0.050) | 0.170** (0.072) | 0.164** (0.068) | 0.145** (0.062) | | | |
| <i>N</i> | 754 | 657 | 485 | 754 | 1708 | 750 | 754 | 1708 | | | |
| Panel B: children's opinions and norms about fathers | | | | | | | | | | | |
| Individual Opinions | 0.027 (0.048) | 0.006 (0.049) | 0.095 (0.059) | -0.017 (0.048) | -0.014 (0.037) | 0.033 (0.052) | 0.028 (0.049) | 0.036 (0.061) | | | |
| <i>N</i> | 873 | 760 | 553 | 754 | 1987 | 869 | 873 | 1987 | | | |
| Social norms | 0.163** (0.080) | 0.152* (0.081) | 0.383** (0.105) | 0.163** (0.080) | 0.139** (0.060) | 0.183** (0.087) | 0.163** (0.080) | 0.206*** (0.065) | | | |
| <i>N</i> | 754 | 657 | 485 | 754 | 1708 | 750 | 754 | 1708 | | | |
| Panel C: Children's counter-stereotypical behavior | | | | | | | | | | | |
| Household Chores | 0.142*** (0.052) | 0.123** (0.057) | 0.093 (0.068) | 0.142*** (0.052) | 0.084** (0.041) | 0.171*** (0.056) | 0.142*** (0.052) | 0.152*** (0.054) | | | |
| <i>N</i> | 750 | 653 | 483 | 750 | 1698 | 747 | 750 | 1698 | | | |
| Panel D: Children's counter-stereotypical expectations | | | | | | | | | | | |
| Full-time work and children | 0.155*** (0.057) | 0.151** (0.061) | 0.154** (0.072) | 0.155** (0.057) | 0.128*** (0.045) | 0.158** (0.062) | 0.155*** (0.057) | 0.161*** (0.056) | | | |
| <i>N</i> | 754 | 657 | 485 | 754 | 1708 | 750 | 754 | 1708 | | | |

Note: Each coefficient comes from separate regressions. Column (1) shows our preferred specification using equation (1) and including individual controls (i.e. gender, date-of-birth fixed effects, parental education and migration background) and school fixed effects. Column (2) drops all children born closely around the cut-off date, i.e., children born between 13.3- 31.3.2007 and 13.3. - 31.3.2006. Column (3) excludes all children born before the reform who have younger siblings. Column (4) excludes all data collected in the pilot phase. Column (5) draws upon all children born in 2006 and 2007. Column (6) estimates the baseline equation (1) but replaces the school fixed effects by class fixed effects. Column (7) displays the estimates from a RD-DiD design which corresponds to equation (1) but adds a first order polynomial of the running variable (the day of birth which corresponds to the relative age in class). Columns (8) and (9) correspond to the classical RDD specification controlling a first order polynomial or second order polynomial, respectively, of the running variable date of birth and drawing upon all children born in 2006 and 2007. We estimate the RDD specification using the *rdrobust* command in Stata and employing a triangular kernel function for the local-polynomial estimator. Column (10) includes two additional non-ISSP questions on gender attitudes. Column (11) weights each ISSP question equally. Column (12) uses inverse-covariance weighting for the ISSP questions. Standard errors are clustered at the class level and in parentheses: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3: Reform effects by gender on behavior and expectations

| | (1) | (2) | (3) |
|--|---------------------|---------------------|---------------------|
| Panel A: Contribution to male chores | | | |
| Effect for Boys | 0.031 (0.066) | 0.025 (0.067) | -0.021 (0.075) |
| Effect for Girls | 0.120* (0.065) | 0.125* (0.069) | 0.125* (0.071) |
| p-value | 0.331 | 0.314 | 0.168 |
| Panel B: Contribution to female chores | | | |
| Effect for Boys | 0.169** (0.083) | 0.143* (0.079) | 0.150* (0.087) |
| Effect for Girls | -0.042 (0.070) | -0.056 (0.072) | -0.064 (0.076) |
| p-value | 0.066 | 0.080 | 0.075 |
| Panel C: Expectation to work full-time and have children | | | |
| Effect for Boys | -0.208** (0.087) | -0.211** (0.089) | -0.198** (0.086) |
| Effect for Girls | 0.054 (0.087) | 0.075 (0.088) | 0.081 (0.085) |
| p-value | 0.035 | 0.024 | 0.017 |
| Individual controls | | Yes | Yes |
| School FE | | | Yes |

Notes: Each column in each panel stems from a separate regression estimating equation (1) separate by gender. Column (1) shows the results without controls, column (2) when controlling for individual background characteristics and column (3) when adding school fixed effects. *Male household chores* equals 1 if a child helps at least sometimes with making small repairs or grocery shopping, 0.5 if in either of the two, and 0 otherwise. *Female household chores* corresponds to a dummy which equals 1 if a child helps at least sometimes with cleaning. *Full-time work and children* is a dummy that is 1 when a participant expects to work full-time and have children, and 0 otherwise. P-value gives the p-value of a Wald test testing the equality of coefficients estimated separately for boys and girls. Standard errors are clustered at the session/class level and shown in parentheses: * $p < 0.1$, ** $p < 0.05$, ***

FOR ONLINE PUBLICATION ONLY

Appendix A: Additional Figures and Tables

Figure A.1: The data collection process

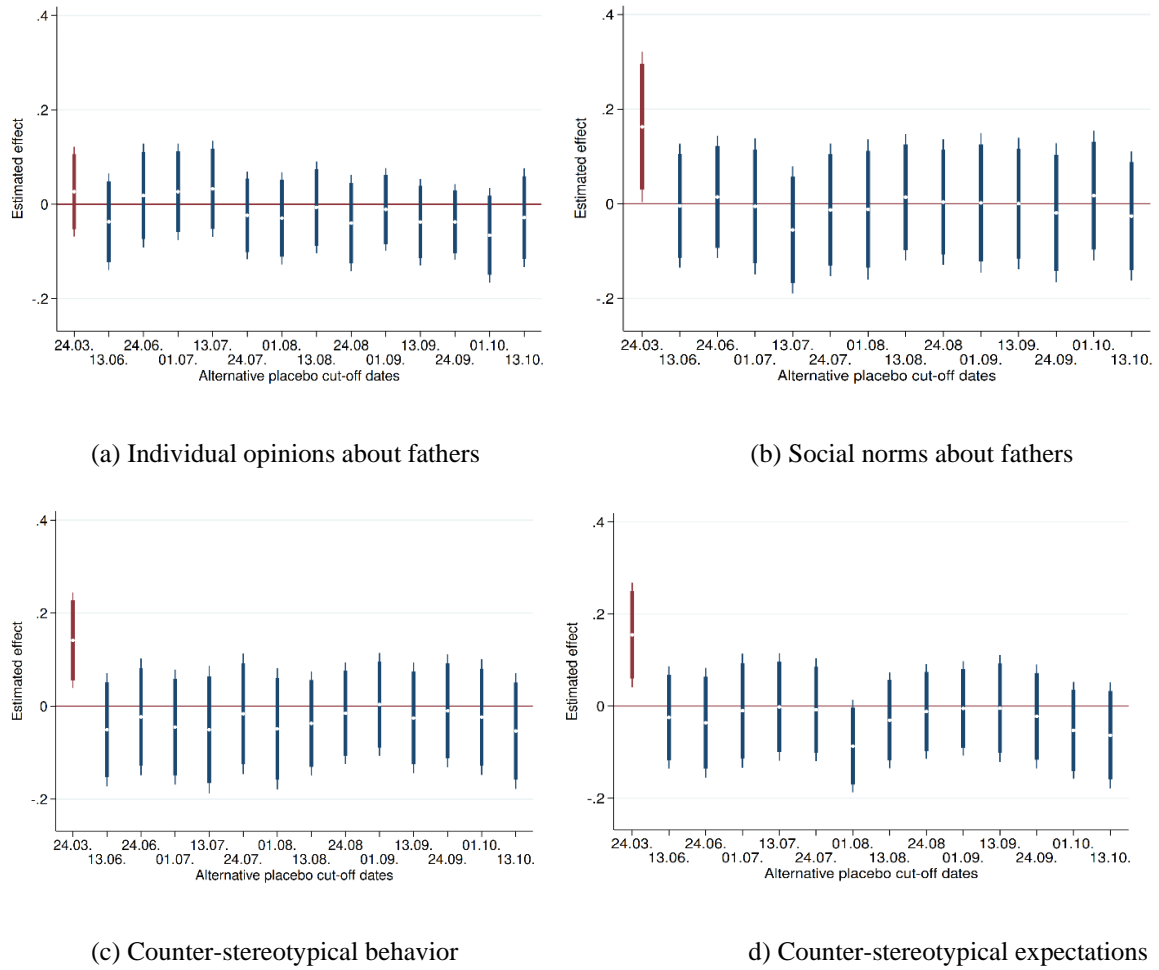


Note: Mobile laboratory installed in one of the schools.



Note: Students at one of the schools answering the questionnaire.

Figure A.2: Placebo tests for alternative cut-off dates (opinions and norms about fathers' labor supply, counter-stereotypical behavior and expectations)



Note: The figure shows a series of placebo reform effects for individual opinions (a) and social norms (b) on fathers' labor supply and counter-stereotypical behaviors (c) and expectations (d) using our preferred specification (estimating equation (1) using a ± 82 days window around alternative cut-off dates and controlling for individual characteristics and school fixed effects). The red bar shows the estimated reform effect and the respective 95% confidence interval using the actual cut-off date (March 24, 2007), while the 13 blue bars show the estimated placebo effects when shifting the cut-off dates successively further (starting with the first possible cut-off date on June 12, 2007 until the last possible cut-off date – to only consider children belonging to the same school cohort – on October 12).

Table A.1: Summary statistics on the outcome variables among pre-reform children

| | (1) Pooled | (2) Girls | (3) Boys | (4) Diff |
|--|---------------------------------|--------------------------------|---------------------------------|----------------------------------|
| Panel A: Gender role attitudes | | | | |
| “Agree” or “strongly agree” with: | | | | |
| (a) A working mother can establish just as warm and secure relationship with her children as a mother who does not work. | 0.668 [0.472] | 0.647 [0.480] | 0.706 [0.458] | -0.059 (0.064) |
| (b) Both men and women should contribute to the household income. | 0.906 [0.292] | 0.893 [0.310] | 0.929 [0.258] | -0.036 (0.040) |
| “Disagree” or “strongly disagree” with: | | | | |
| (c) A pre-school child is likely to suffer if his or her mother works. | 0.587 [0.493] | 0.627 [0.485] | 0.518 [0.503] | 0.109 (0.067) |
| (d) All in all, family life suffers when the woman has a full-time job. | 0.634 [0.483] | 0.653 [0.478] | 0.600 [0.493] | 0.053 (0.066) |
| (e) A job is all right, but what most women really want is a home and children. | 0.698 [0.460] | 0.733 [0.444] | 0.635 [0.484] | 0.098 (0.062) |
| (f) Being a housewife is just as fulfilling as working for pay. | 0.617 [0.487] | 0.640 [0.482] | 0.576 [0.497] | 0.064 (0.066) |
| (g) A man's job is to earn money and a woman's job is to look after the home and family. | 0.983 [0.130] | 0.987 [0.115] | 0.976 [0.152] | 0.010 (0.018) |
| Gender role attitudes index (1 st principal component of the 7 items) | -0.000 [1.000] | 0.116 [0.993] | -0.205 [0.984] | 0.322** (0.134) |
| Panel B: Individual opinions | | | | |
| A mother with a child below school age | | | | |
| should not work. | 0.247 [0.432] | 0.220 [0.416] | 0.294 [0.458] | -0.074 (0.059) |
| should work part-time. | 0.643 [0.480] | 0.660 [0.475] | 0.612 [0.490] | 0.048 (0.065) |
| should work full-time. | 0.021 [0.145] | 0.027 [0.162] | 0.012 [0.108] | 0.015 (0.020) |
| A father with a child below school age | | | | |
| should not work. | 0.140 [0.348] | 0.133 [0.341] | 0.153 [0.362] | -0.020 (0.047) |
| should work part-time. | 0.732 [0.444] | 0.733 [0.444] | 0.729 [0.447] | 0.004 (0.060) |
| should work full-time. | 0.034 [0.182] | 0.033 [0.180] | 0.035 [0.186] | -0.002 (0.025) |
| Individual opinions about mothers (=1 if either part-time or full-time work, 0 otherwise) | 0.664 [0.473] | 0.687 [0.465] | 0.624 [0.487] | 0.063 (0.064) |
| Individual opinions about fathers (=1 if either part-time or not work, 0 otherwise) | 0.872 [0.334] | 0.867 [0.341] | 0.882 [0.324] | -0.016 (0.045) |
| Panel C: Social norms | | | | |
| How socially appropriate is it that | | | | |
| ... a mother with a child below school age works part-time? | | | | |
| fairly appropriate | 0.419 [0.494] | 0.403 [0.492] | 0.444 [0.500] | -0.041 (0.070) |
| appropriate | 0.219 [0.414] | 0.261 [0.441] | 0.148 [0.357] | 0.113* (0.058) |
| ... a mother with a child below school age works full-time? | | | | |
| fairly appropriate | 0.056 [0.230] | 0.075 [0.264] | 0.025 [0.156] | 0.050 (0.032) |
| appropriate | 0.042 [0.201] | 0.052 [0.223] | 0.025 [0.156] | 0.028 [0.028] |
| ... a father with a child below school age works part-time? | | | | |
| fairly appropriate | 0.460 [0.500] | 0.465 [0.501] | 0.453 [0.502] | 0.012 (0.080) |
| appropriate | 0.270 [0.445] | 0.303 [0.462] | 0.219 [0.417] | 0.084 (0.071) |
| ... a father with a child below school age works not at all? | | | | |
| fairly appropriate | 0.335 [0.473] | 0.336 [0.474] | 0.333 [0.474] | 0.002 (0.067) |
| appropriate | 0.344 [0.476] | 0.343 [0.477] | 0.346 [0.479] | -0.002 (0.067) |
| Social norms about mothers (=1 if appropriate and fairly appropriate to work part- or full-time, 0 otherwise) | 0.637 [0.482] | 0.664 [0.474] | 0.593 [0.494] | 0.072 (0.068) |
| Social norms about fathers (=1 if appropriate to work part-time or not at all, 0 otherwise) | 0.451 [0.499] | 0.463 [0.500] | 0.432 [0.498] | 0.031 (0.070) |

Continues on next page ...

Table A.1. – continues from previous page

| | (1) Pooled | (2) Girls | (3) Boys | (4) Diff |
|---|--------------------------------|--------------------------------|--------------------------------|------------------------------------|
| Panel D: Household chores | | | | |
| Do you help at least occasionally with ... | | | | |
| Doing the laundry | 0.837 [0.370] | 0.858 [0.350] | 0.802 [0.401] | 0.056 (0.052) |
| Small repairs | 0.437 [0.497] | 0.343 [0.477] | 0.593 [0.494] | -0.249*** (0.068) |
| Grocery shopping | 0.827 [0.379] | 0.791 [0.408] | 0.887 [0.318] | -0.096* (0.053) |
| Cleaning | 0.771 [0.421] | 0.836 [0.372] | 0.662 [0.476] | 0.173*** (0.058) |
| Cooking | 0.533 [0.500] | 0.545 [0.500] | 0.512 [0.503] | 0.032 (0.071) |
| Counter-stereotypical behavior (=1 if small repairs and grocery shopping (cleaning) for girls (boys), =0.5 if small repairs or grocery shopping for girls, =0 otherwise) | 0.603 [0.401] | 0.567 [0.346] | 0.662 [0.476] | -0.095* (0.056) |
| Panel E: Employment and family expectations | | | | |
| In 20-year time, how do you see yourself? | | | | |
| Not working and not having children | 0.023 [0.151] | 0.015 [0.122] | 0.037 [0.190] | -0.022 (0.021) |
| Working and not having children | 0.344 [0.476] | 0.358 [0.481] | 0.321 [0.470] | 0.037 (0.067) |
| Not working and having children | 0.033 [0.178] | 0.022 [0.148] | 0.049 [0.218] | -0.027 (0.025) |
| Working part-time and having children | 0.423 [0.495] | 0.470 [0.501] | 0.346 [0.479] | 0.124* (0.069) |
| Working full-time and having children | 0.177 [0.382] | 0.134 [0.342] | 0.247 [0.434] | -0.113** (0.053) |
| Counter-stereotypical expectations (=1 (0) if full-time work and children for girls (boys), =0 (1) otherwise) | 0.367 [0.483] | 0.134 [0.342] | 0.753 [0.434] | -0.619*** (0.053) |

Note: The information reported is restricted to children born between 01.01.2007 and 23.03.2007 to show the descriptive statistics at baseline (the counterfactual group). Column (1) includes both girls and boys, while column (2) is restricted to girls and column (3) to boys. Column (4) displays the gender gap in the different measures. Panel A displays the share of pre-reform children providing a non-traditional answer to the battery of gender role attitudes questions taken from the ISSP. For question (a) and (b) it shows the percentage who “strongly agree” or “agree”. For questions (c), (d), (f), (g) and (h), it displays the percentage who “strongly disagree” or “disagree”. The **gender role attitudes index** is the first element resulting from a principal component analysis of the seven questions. The index is normalized to have mean 0 and standard deviation 1 for pre-reform children. Panel B displays respondents’ opinion about the optimal labor supply of parents with young children. Missing categories are “I don’t know” and “I don’t want to answer”. **Individual opinion about mothers** is an indicator that takes value 1 if the child answered that a mother with a child below school age should work full-time or part-time. **Individual opinion about fathers** is an indicator that takes value 1 if the child answered that a father with a child below school age should work part-time or not at all. Panel C shows the results of the incentivized coordination game. For every question it displays the percentage of children who answered that it is “fairly appropriate” and the percentage of children who answered that it is “appropriate”. The **social norm about mothers** indicates the share who answered that it is at least “fairly appropriate” that a mother with a child below school age works part-time or full-time. The **social norm about fathers** indicates the share who answered that it is “appropriate” that a father with a child below school age works part-time or does not work at all. Panel D shows the share that contributes “at least occasionally” (i.e. less than once a week but more than once a month) to each domestic task. The variable **counter-stereotypical behavior** is the average of “small repairs” and “grocery shopping” for girls and the value of “cleaning” for boys. Panel E shows the expectations of respondents in 20 years’ time regarding fertility and labor market participation. The variable **counter-stereotypical expectations** is 1 if a girl (boy) does (not) expect to work full-time and have children. The standard deviation is given in brackets. Column (4) displays a Wald test on gender differences and the respective standard errors are in parentheses: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A.2: Balancing test in covariates

| | Pooled | Treated Cohort (2007) | | Control Cohort (2006) | | DiD |
|--------------------------------|-------------------|-----------------------|-------------------|-----------------------|-------------------|-------------------|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| | | Post (24.03-12.06) | Pre (01.01-23.03) | Post (24.03-12.06) | Pre (01.01-23.03) | |
| Age | 13.264 [0.535] | 12.669 [0.191] | 12.888 [0.204] | 13.609 [0.276] | 13.828 [0.271] | 0.000 (0.000) |
| Male | 0.467 [0.499] | 0.544 [0.499] | 0.362 [0.482] | 0.541 [0.499] | 0.441 [0.498] | 0.097 (0.067) |
| Do you live with your mother? | | | | | | |
| Yes | 0.844 [0.363] | 0.845 [0.363] | 0.826 [0.380] | 0.839 [0.368] | 0.868 [0.339] | 0.049 (0.047) |
| Most days | 0.063 [0.243] | 0.057 [0.232] | 0.068 [0.252] | 0.069 [0.254] | 0.057 [0.233] | -0.031 (0.029) |
| Some days | 0.080 [0.272] | 0.083 [0.276] | 0.098 [0.298] | 0.078 [0.269] | 0.062 [0.241] | -0.022 (0.036) |
| No | 0.013 [0.112] | 0.016 [0.124] | 0.009 [0.092] | 0.014 [0.117] | 0.013 [0.114] | 0.004 (0.017) |
| Do you live with your father? | | | | | | |
| Yes | 0.784 [0.412] | 0.777 [0.417] | 0.791 [0.407] | 0.798 [0.402] | 0.767 [0.424] | -0.029 (0.052) |
| Most days | 0.013 [0.112] | 0.010 [0.102] | 0.013 [0.113] | 0.014 [0.117] | 0.013 [0.114] | -0.006 (0.017) |
| Some days | 0.148 [0.355] | 0.155 [0.363] | 0.157 [0.365] | 0.138 [0.345] | 0.141 [0.349] | -0.002 (0.046) |
| No | 0.056 [0.230] | 0.057 [0.232] | 0.038 [0.192] | 0.050 [0.219] | 0.079 [0.271] | 0.037 (0.033) |
| Does your mother work? | | | | | | |
| Full-time | 0.504 [0.500] | 0.487 [0.501] | 0.498 [0.501] | 0.541 [0.499] | 0.489 [0.501] | -0.046 (0.060) |
| Part-time | 0.371 [0.483] | 0.383 [0.487] | 0.383 [0.487] | 0.335 [0.473] | 0.383 [0.487] | 0.047 (0.058) |
| Does not work | 0.105 [0.307] | 0.104 [0.306] | 0.098 [0.298] | 0.110 [0.314] | 0.110 [0.314] | -0.010 (0.038) |
| Does your father work? | | | | | | |
| Full-time | 0.675 [0.469] | 0.648 [0.479] | 0.621 [0.486] | 0.711 [0.454] | 0.718 [0.451] | 0.081 (0.062) |
| Part-time | 0.244 [0.430] | 0.280 [0.450] | 0.298 [0.458] | 0.211 [0.409] | 0.189 [0.393] | -0.071 (0.053) |
| Does not work | 0.049 [0.217] | 0.031 [0.174] | 0.064 [0.245] | 0.046 [0.210] | 0.053 [0.224] | -0.032 (0.033) |
| Did your mother go to college? | | | | | | |
| Yes | 0.498 [0.500] | 0.513 [0.501] | 0.489 [0.501] | 0.509 [0.501] | 0.485 [0.501] | 0.016 (0.057) |
| No | 0.345 [0.476] | 0.347 [0.477] | 0.336 [0.473] | 0.326 [0.470] | 0.370 [0.484] | 0.057 (0.060) |
| Did your father go to college? | | | | | | |
| Yes | 0.394 [0.489] | 0.383 [0.487] | 0.374 [0.485] | 0.427 [0.496] | 0.392 [0.489] | 0.011 (0.062) |
| No | 0.399 [0.490] | 0.352 [0.479] | 0.400 [0.491] | 0.385 [0.488] | 0.449 [0.499] | 0.002 (0.055) |
| Migrant mother | 0.126 [0.332] | 0.171 [0.377] | 0.102 [0.303] | 0.110 [0.314] | 0.128 [0.335] | 0.052 (0.043) |
| ... born in Americas | 0.048 [0.214] | 0.073 [0.260] | 0.043 [0.202] | 0.037 [0.188] | 0.044 [0.206] | 0.021 (0.027) |
| ... born in Africa | 0.034 [0.182] | 0.057 [0.232] | 0.021 [0.145] | 0.037 [0.188] | 0.026 [0.161] | 0.008 (0.023) |
| ... born in Europe | 0.011 [0.106] | 0.010 [0.102] | 0.021 [0.145] | 0.009 [0.096] | 0.004 [0.066] | -0.015 (0.017) |
| ... born somewhere else | 0.015 [0.121] | 0.021 [0.143] | 0.017 [0.130] | 0.000 [0.000] | 0.022 [0.147] | 0.028 (0.017) |
| Migrant father | 0.117 [0.321] | 0.166 [0.373] | 0.098 [0.298] | 0.092 [0.289] | 0.119 [0.324] | 0.074 (0.045) |
| ... born in Americas | 0.036 [0.185] | 0.057 [0.232] | 0.043 [0.202] | 0.009 [0.096] | 0.035 [0.185] | 0.029 (0.023) |
| ... born in Africa | 0.032 [0.176] | 0.057 [0.232] | 0.017 [0.130] | 0.037 [0.188] | 0.022 [0.147] | 0.009 (0.022) |
| ... born in Europe | 0.018 [0.134] | 0.016 [0.124] | 0.021 [0.145] | 0.018 [0.135] | 0.018 [0.132] | 0.002 (0.021) |
| ... born somewhere else | 0.017 [0.130] | 0.031 [0.174] | 0.013 [0.113] | 0.014 [0.117] | 0.013 [0.114] | 0.013 (0.018) |

Notes: Column (1) display the mean of each variable for our baseline sample (i.e. January 1 to June 12, 2006 and January 1 to June 12 in 2007). Column (2) and (3) are restricted to the treated cohort (2007) and Column (4) and (5) to the control cohort (2006). Column (6) shows the DiD estimate for the respective variable using equation (1) in the main text and controlling for birth-date fixed effects and school fixed effect. The respective standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.3: Robustness checks using standard error correction methods

| | (1) Baseline | (2) School cluster | (3) Bonferroni - Holm | (4) Romano - Wolf |
|------------------------------------|-----------------|--------------------------|-----------------------------|-------------------------|
| Gender role attitudes index | 0.0373 | 0.0706 | 0.0208 | 0.0189 |
| Individual opinions about mothers | 0.0141 | 0.0213 | 0.0048 | 0.0072 |
| Social norms about mothers | 0.0128 | 0.0029 | 0.0112 | 0.0072 |
| Individual opinions about fathers | 0.5791 | 0.5766 | 0.4614 | 0.4614 |
| Social norms about fathers | 0.0453 | 0.0832 | 0.0163 | 0.0189 |
| Counter-stereotypical behavior | 0.0076 | 0.0094 | 0.0062 | 0.0048 |
| Counter-stereotypical expectations | 0.0083 | 0.0130 | 0.0035 | 0.0048 |

Notes: Each cell contains a p-value of the reform effect on the main outcomes. Column (1) uses our preferred specification using equation (1) including individual controls (i.e. gender, date-of-birth fixed effects, parental education and migration background), school fixed effects and Liang-Zeger robust standard errors clustered at the class level. Column (2) uses the same specifications as in column (1) but Liang-Zeger robust standard errors are clustered at the school level. Columns (3) and (4) report p-values providing control of the familywise error rate. These are implemented using the *rwolf2* command in Stata using our preferred specification and clustering while relying on 100,000 bootstrap replications (see *Clarke, Romano, and Wolf 2020*).

Appendix B: Questionnaire

START

INFORMATION AND CONSENT

We are conducting a research project to analyze the impact of public policies on the cognitive and non-cognitive development of children and adolescents. The principal investigator on the project is Dr. Lúdia Farré from the University of Barcelona.

I CONFIRM that:

- the information about the research project has been read to me,
- I have been able to ask questions about the project,
- I have received enough information about the project.

I UNDERSTAND that my participation in the project is voluntary and that I can withdraw from it at any time without having to justify my decision. I GIVE MY CONSENT to take part in this research project

Yes – No

I. ATTITUDES

We are now going to ask you some questions. Please answer these questions sincerely. We want to know what you really think. No one will know how you chose to respond. There are no right or wrong answers.

1. Read the following statements and say whether you: Strongly agree / Agree / Neither agree nor disagree / Disagree / Strongly disagree.

- a) A mother who goes out to work can have just as good a relationship with her children as a mother who does not go out to work.
- b) It can be bad for a child under the age of 3 if her mother goes out to work.
- c) When a mother spends the whole day (morning and afternoon) out working, family life can suffer.
- d) Having a job is fine, but what most women want is to form a family and have children.
- e) Taking care of the home and family can make a woman just as happy as having a job.
- f) Both partners, should bring money into the home.
- g) The man should be the breadwinner while the woman should take care of the home and family.
- h) When a woman earns more than her husband, there are certainly problems.
- i) Both the mother and father should take leave from work for a few weeks after the birth of their son or daughter.

2.a) Do you think a **father** should work full-time (mornings and afternoons), part-time (mornings or afternoons only) or not work at all when his child has not yet started school (under the age of 3).

2.b) Do you think a **mother** should work full-time (mornings and afternoons), part-time (mornings or afternoons only) or not work at all when her child has not yet started school (under the age of 3).

- Shouldn't work at all.
- Should work part-time (mornings or afternoons only).
- Should work full-time (mornings and afternoons).
- I don't know
- I prefer not to answer

3. Imagine a family with a child who is still too young to go to school (under the age of 3). What do you consider the best way to organize their family and work life?

- The mother stays at home and the father goes out to work full-time (mornings and afternoons).
- The mother works part-time (mornings or afternoons only) and the father goes out to work full time (mornings and afternoons).
- Both the mother and father go out to work full-time (mornings and afternoons).
- Both the mother and father work part-time (mornings or afternoons only).
- The father works part-time (mornings or afternoons only) and the mother go out to work full-time (mornings and afternoons).
- The father stays at home and the mother go out to work full-time (mornings and afternoons).
- I don't know.

II. CONTEST

In this part of the study you have to complete three different tasks to win points. The number of points you win will depend on the number of correct answers you give when completing the tasks. Before beginning each task, we will give you instructions on how you can win points.

Test run

In the three tasks, you will be asked to add up 3 numbers. To get used to doing the task on the computer, you will now be allowed a 30-second test run. During the test run you cannot win any points.

1. Task 1 - Individual task

You will now have 2 minutes to do the sums. For each sum you get right, you will win 1 point. If you make a mistake, no points are deducted. This task is called the **individual task**.

2. Task 2 - Quiz

You have been put in a group with three boys or girls from your class (but you do not know who they are). As in task 1, you will have 2 minutes to do the sums. The boy or girl from your group that correctly solves most sums will receive 4 points for each correct answer. The rest of the group will not win any points. This task is called the **quiz**.

3. Task 3 - Choose the task type you want to do to win points

As in the previous tasks, you will have 2 minutes to do the sums. But now you have to choose which task you want to do: the *individual task* or the *quiz*. If you choose the *individual task*, you will win 1 point for each sum you get right. If you choose the *quiz*, you will only win points if you solve more sums than those solved by the same boys and girls in your group in task 2 above. If you choose the *quiz*, you will win 4 points for each correct answer.

Which task do you prefer to do to win points?

Individual task – Quiz

a) Task 3 - Individual task

You have chosen the individual task. As before, you have 2 minutes to do the sums. For each sum you get right, you will win 1 point. If you make a mistake, no points are deducted.

b) Task 3 – Quiz

You have chosen the quiz. As before, you have 2 minutes to do the sums. Now you will only win points if you solve more sums than those solved by the same boys and girls in your group in task 2 above. In this case, you will win 4 points for each correct answer.

4. This question is about the **quiz** you took in **task 2**. How well do you think you did in relation to the other members of your group? If you answer correctly, you will earn 4 points.

I finished: First, Second, Third, Last

5. We played this same game with boys and girls in the same grade as you but at a different school. Think again about the **quiz** you took in **task 2**.

How many boys and girls do you think there were among the three who got **the most points** at the **other** school? If you answer correctly, you will win 4 points.

How many boys and girls do you think there were among the three who got **the fewest points** at the **other** school? If you answer correctly, you will win 4 points.

3 boys, 2 boys and 1 girl, 1 boy and 2 girls, 3 girls

6. Now, to win points, you have to choose to take part in one of the following lotteries. It's like tossing a coin in the air and seeing if it lands as heads or tails. If you choose Lottery 1, you are guaranteed to win 5 points. In the other lotteries, the number of points you win depends on how lucky you are. If you choose Lottery 2, you can win either 4 point or 8 points, etc.

- Lottery 1 - 5 points or 5 points
- Lottery 2 - 4 points or 8 points
- Lottery 3 - 3 points or 10 points
- Lottery 4 - 2 points or 12 points
- Lottery 5 - 1 point or 14 points
- Lottery 6 - 0 points or 15 points

III. NORMS

You will now play a game with a classmate but you won't know who he or she is. In this game, we will ask you both the same question. For example: **Do you think it will rain tomorrow?** If you both give the same answer, you will each win 2 points. If you give different answers, neither of you will win any points. If you have any doubts, please raise your hand.

Test Question 1: If your partner says "Yes, it'll rain tomorrow", what must you answer to win two points?

Yes, it'll rain tomorrow - No, it won't rain tomorrow

Correct answer. Incorrect answer. To win points, you must give the same answer as your partner.

Test Question 2: If your partner thinks that you will answer: "No, it won't rain tomorrow", what must your answer be to win two points?

Yes, it'll rain tomorrow - No, it won't rain tomorrow

Correct answer. Incorrect answer. If your partner thinks that you will answer: "No, it won't rain tomorrow", he or she will answer: "No, it won't rain tomorrow". When he or she answers, "No, it won't rain tomorrow", you must answer: "No, it won't rain tomorrow" to win points.

Test Question 3: To win points in this game, do you think it's important that it rains tomorrow?

Yes – No

Correct answer. Incorrect answer. In this game, tomorrow's weather (whether it rains or not) is not important to win points. What is important is that you and your partner give the same answer.

1. Let's start the game. We are now going to ask you and your partner if you think it'll rain tomorrow. If you give the same answer, you will each win 2 points. If you give different answers, neither of you will win any points.

Do you think it will rain tomorrow? Yes, it'll rain tomorrow - No, it won't rain tomorrow

2. Let's continue playing the game. We are now going to ask you and your partner if you think it is appropriate to copy in an exam. When we say "appropriate", we refer to behavior that most people think is correct or good. In contrast we say that behavior is "inappropriate" when most people think it is incorrect or bad. If you and your partner give the same answer, you will each receive 4 points. Copying in an exam: Is appropriate, Is fairly appropriate, Is fairly inappropriate, Is inappropriate

3. We are now going to ask you and your partner four more questions. Remember, only if you give the same answer will you each receive 4 points.

- a) That a **mother** goes out to work **full-time** (mornings and afternoons) when her child has not yet started school (under the age of 3).
- b) That a **mother** goes out to work **part-time** (only mornings or afternoons) when her child has not yet started school (under the age of 3).
- c) That a **father** goes out to work **part-time** (only mornings or afternoons) when his child has not yet started school (under the age of 3).
- d) That a **father does not go out to work** at all so as to look after his child when the child has not yet started school (under the age of 3).

Is appropriate, Is fairly appropriate, Is fairly inappropriate, Is inappropriate

4. How do you think most people in your class responded to the following statement: “The man should be the breadwinner while the woman should take care of the home and family”. If you answer correctly, you will win 5 points. Strongly agree / Agree / Neither agree nor disagree / Disagree / Strongly disagree.

IV. OUTCOMES

We are now going to ask you some questions about how you see yourself in the future. Please answer these questions sincerely and, remember, no one will know how you chose to respond.

1. How do you see yourself in 20 years' time?

- I'll have children and not go out to work
- I'll have children and go out to work part-time (only mornings or afternoons)
- I'll have children and go out to work full-time (mornings and afternoons)
- I'll go out to work and I'll have no children
- I'll not go out to work and I'll have no children

2. What do you want to be when you grow up? Choose just ONE of the jobs on the list:

- Football player
- Police officer
- Primary or Secondary school teacher
- Scientist
- Architect
- Engineer
- Doctor
- Firefighter
- Vet
- Computer scientist
- Hairdresser
- Singer or Musician
- Nurse
- Actor
- Gymnast or Dancer
- Designer
- Journalist
- Lawyer or Judge
- Biologist

If the job you'd like to do is not on the list, write it here:

V. DEMOGRAPHICS

- 1.a) What year were you born in?
- 1.b) What month were you born in? January, February, March, April, May, June, July, August, September, October, November, December
- 1.c) What day were you born on? Between the 1st and the 12th, between the 13th and the 23rd, between the 24th and the 31st

2. Are you a boy or a girl? Boy, Girl

- 3.a) Where were you born?
- 3.b) Where was your **mother** born?
- 3.c) Where was your **father** born?

Pilot: In Catalonia, somewhere in the rest of Spain, in another country, I don't know

Main: Morocco, Romania, Ecuador, Another European country, Another African country, Another American country, An Asian country, Other, I don't know

4.a) Do you live with your **mother**?

4.b) Do you live with your **father**?

Yes; No; Yes, some days; Yes, most days

5.a) Does your **mother** have a job?

5.b) Does your **father** have a job?

Yes, a full-time job (mornings and afternoons)

Yes, a part-time job (only mornings or afternoons)

No; I don't know

6.a) Did your **mother** go to college?

6.b) Did your **father** go to college?

Yes, No, I don't know

7.a) How many **older brothers** do you have?

7.b) How many **older sisters** do you have?

7.c) How many **younger brothers** do you have?

7.d) How many **younger sisters** do you have?

0; 1; 2; more than 2

Including stepbrothers and stepsisters

VI. MECHANISM

We now ask you some questions about you and your family. You should tell us who normally does the chores in the house. We also want to know about you and which of the chores you do.

1. Who normally does these chores around your house?

a) Wash the clothes and put them away.

b) Do small repair jobs. For example, hang a picture, repair a door knob, paint a wall.

c) Go to the supermarket or the market.

d) Clean the house.

e) Cook.

Always my mother; My mother more than my father; Always my father; My father more than my mother; Both my father and my mother; Another person; My parents don't live together; I don't know

2. What about you? When do you do these chores or when do you help out with them?

At least once a week; Occasionally (less than once a week); Almost never; Never

3. We want you now to think about your parents,

- a) Who helps you with your homework?
- b) Who stays at home when you are sick?
- c) Who spends more time with you on weekends?
- d) Who's home when you get back from school?
- e) Who do you talk to when you're worried or have a problem?
- f) Who's home when you eat at night?
- g) Who asks you if you've had a good day?

Always my mother; My mother more than my father; Always my father; My father more than my mother; Both my father and my mother; Another person; My parents don't live together; I don't know

3. Now we would like to know if you think your father is a supporter of feminism.

Yes, No, I don't know, I prefer not to answer

VII. OPTIONAL

1. We are now going to ask you some more questions about yourself. Please answer these questions sincerely and, remember, no one will know how you chose to respond.

- a) When someone does me a favor, I am more than willing to return that favor.
- b) I assume people's intentions are always good.
- c) I am in the habit of putting off until later work I know it would be better to finish right away.
- d) This statement does not describe me at all; This statement describes me perfectly
- e) Do you like to take risks or take risky actions?
- f) Do you like to make donations to a good cause without expecting anything in return?
I usually put off until tomorrow the jobs I should get done today

This statement does not describe me at all; This statement describes me perfectly

2. To finish, we would like you to answer the following questions. You should answer either True or False

- a) I always show respect to the elderly:
- b) Sometimes I don't feel like doing what the teacher asks us to do:
- c) I sometimes feel like throwing or breaking things:
- d) I am never disrespectful to my parents or answer them back:
- e) When I make a mistake, I'm the first to admit it:
- f) I sometimes laugh at people:

- g) I always wash my hands before eating:
- h) I sometimes feel like lazing around and not going to school:
- i) I have never been tempted to break the rules or the law:
- j) Sometimes I don't feel like helping my parents out even though I know they need my help around the house:
- k) I sometimes say things just to impress my friends:
- l) I never shout when I'm angry:

3. And last but not least:

- a) What do cows drink?
- b) If I have 3 apples and you take away 2, how many apples do you have?
- c) A bat and a ball cost \$1.10 in total. The bat costs \$1.00 more than the ball. How much does the ball cost?

PAYMENT

Thank you very much for taking part. The following page shows how many euros you have contributed to the voucher we will give to your school.

Appendix C: Mechanisms

In section 3.2, we summarize the existing evidence on paternity leave acting as a trigger for counter-stereotypical behavior among parents. In this Appendix, we want to corroborate these findings using our empirical specification and add some evidence on possible medium- to long-run results. For this purpose, we rely on various datasets, (i) the Spanish Labor Force Survey to document mothers' and father's labor market choices over the first decade of their child's life, (ii) the Spanish Time-Use Survey (2009-10) and (iii) our own data (described in section 2) to study fathers' contribution to childcare and household chores when children are 2-3 years old and 12-13 years old, respectively.

For this exercise, we employ the same empirical specification as in our main analysis, a difference-in-differences model comparing children born shortly before and after the enactment date of the paternity leave reform on March 24, 2007, drawing upon children born in the same window of birthdates in the previous year (2006) as a control group. Hence, we estimate the following equation:

$$P_i = \beta_0 + \beta_1 Cohort2007_i + \beta_2 PostMarch24_i + \beta_3 Cohort2007_i * PostMarch24_i + X_i + \varepsilon_i \quad (2)$$

where P_i stands for the various parental behaviors in the labor market and in the home. $Cohort2007_i$ is a dummy variable indicating whether child i is born in 2007 (versus 2006), and $PostMarch24_i$ represents a dummy variable taking the value 1 if child i is born on or after March 24 (independently of the calendar year). The interaction term $Cohort2007_i * PostMarch24_i$ takes value 1 for any child born after the introduction of paternity leave. We can thus interpret the coefficient β_3 as the intent-to-treat effect of paternity leave on parental behaviors.

Table C.1. reports the difference-in-differences estimates for mothers' and fathers' employment outcomes using equation (2) and controlling for individual and parental background characteristics (e.g., child's gender and birth order as well as parental age, education, and migration background) and region fixed effects. We use two alternative samples, parents whose child was born between January and June in 2006 and 2007 (Table C.1., column (1)) and parents whose child was born at any point in 2006 and June 2007 (Table C.1., column (2)).

In line with Farré and González (2019), we find that women whose partners were eligible for paternity leave worked more in the first three years after childbirth (Table C.1, Panel A). Specifically, women with children aged 0-2 years old are 3 percentage points more likely to be employed and work on average 1 hour more per week. This effect, however, fades out when children are 3 years and older, i.e., as soon as children enter compulsory schooling. The introduction of paternity leave does not alter fathers' labor supply, neither in the short-run nor in the longer-run (Table C.1, Panel B).

Table C.2. displays the difference-in-differences estimates for paternal engagement in childcare and household chores using equation (2). Table C.2., Panel A reports the difference-in-differences estimates using the Spanish Time-Use Survey (of 2009-10) and thus 2-3 years after childbirth. The Spanish Time-Use Survey allows us to control for individual-level control variables (child's gender and birth order as well as parents' age at childbirth, marital status, citizenship, education) and region fixed effects. We again rely on two samples, a narrow sample using only fathers whose child was born January and June in 2006 and 2007, and a broader sample using all fathers whose child was born at any time in 2006 and 2007.

Given the rather small sample size of the Spanish Time Use Survey, we concentrate immediately on the results using the broader sample (see Table C.2, Panel A, column (2)). Fathers eligible for paternity leave spend on average 19 minutes per day more on household chores and childcare. While the combined effect is not significant at the conventional levels, splitting time devoted to childcare into the different activities undertaken with a child reveals that paternity leave induces fathers to spend significantly more time with their children doing recreational activities, such as reading, playing, and talking. On average, eligible fathers increase their time devoted to recreational activities with their child by 45 minutes per day. These effects are sizeable considering that prior to the introduction of paternity leave fathers spend on average 95 minutes per day on childcare and 34 minutes on recreational activities with their children.

Table C.2., Panel B displays the reform effects on children's perception of fathers' contribution to household chores using our own data and thus 12-13 years after childbirth. The reported estimates stem again from a difference-in-difference specification using equation (2) and controlling for individual level controls (gender, parental education, and migrant status) and school fixed effects. We find that fathers eligible for paternity leave are 14 percentage points or 22% more likely to do at least some of the laundry. These findings indicate that paternity leave may have affected fathers' involvement in (some) household tasks even in the long run.

Taken together, these additional results support the claim that the introduction of paternity leave triggered counter-stereotypical behavior among eligible parents not only immediately after childbirth but also in the long-run.

Table C.1: Reform effects on parental labor supply

| | (1) Ages 0-2 | (2) Ages 3-5 |
|------------------------|--------------------|-------------------|
| Employed | 0.032** (0.014) | -0.003 (0.014) |
| Hours worked last week | 1.046* (0.550) | 0.472 (0.541) |
| N | 15,929 | 17,270 |
| Employed | 0.006 (0.009) | 0.005 (0.011) |
| Hours worked last week | 0.639 (0.590) | 0.447 (0.590) |
| N | 15,929 | 17,270 |

Note: Each coefficient comes from a different regression. The sample includes all mothers or fathers from the 2006q1 to 2019q4 EPA who live with a child born between January and June 2006 or between January and June 2007, and who live with the father/mother of their reference child. Coefficients reported correspond to the treatment indicator that takes value 1 for parents who had a child born from April 2007 onwards. All regressions control for region fixed effects, a quadratic trend in the age of the mother and father at the time of the interview, indicators for secondary and college educated mothers and fathers, and foreign-born mothers and fathers, and indicators for mothers whose child was born in 2007 and for mothers whose child was born between April and June. Standard errors are shown in parentheses: *p < 0.1, **p < 0.05, ***p < 0.01

Table C.2: The reform effects on paternal engagement in childcare and household chores

| | (1) Jan-June 2006 & 07 | (2) 2006 & 2007 |
|---|---------------------------|--------------------|
| Panel A: Paternal engagement 2-3 years after childbirth in: | | |
| Childcare & Housework (min/day) | 23.0 | 18.6 |
| Pre-reform mean [sd] = 201 [172] | (67.2) | (48.7) |
| Childcare (min/day) | 13.8 | 22.0 |
| Pre-reform mean [sd] = 95 [99] | (42.7) | (33.4) |
| i) Educational activities (min/day) | 0.6 | 4.4 |
| Pre-reform mean [sd] = 2.4 [16] | (10.2) | (5.6) |
| ii) Recreational activities (min/day) | 26.1 | 44.8*** |
| Pre-reform mean [sd] = 34 [48] | (17.5) | (16.4) |
| iii) Physical care (min/day) | -16.1 | -26.3 |
| Pre-reform mean [sd] = 46 [66] | (27.6) | (21.3) |
| N | 141 | 303 |
| Panel B: Paternal engagement 12-13 years after childbirth in: | | |
| Laundry | 0.142*** | 0.096** |
| Pre-reform mean [sd] = 0.617 [0.487] | (0.053) | (0.043) |
| N | 870 | 1984 |
| Cleaning | -0.037 | -0.064 |
| Pre-reform mean [sd] = 0.598 [0.491] | (0.057) | (0.049) |
| N | 869 | 1977 |
| Cooking | 0.022 | 0.004 |
| Pre-reform mean [sd] = 0.692 [0.463] | (0.059) | (0.046) |
| N | 869 | 1977 |

Note: Each coefficient comes from a different regression. Dependent variables are specified on the first column of each row. Panel A uses data from the 2009-10 Spanish Time-Use Survey, in column (1), a sample limited to parents whose youngest child was born between January and June 2006 and 2007, and in column (2) the sample is extended to include parents whose youngest child was born in 2006 or 2007. The regressions based on the Spanish Time-Use Survey control for second-order polynomials in age at birth of both parents, indicators for whether the parents were married, were of foreign nationality, whether each parent had at least a high school education, whether the child is the first-born and the gender of the child. Panel B uses our own data when children are 12-13 years old. In column (1) the sample is limited to children born between January 1 and June 13 in 2006 and 2007, in column (2) the sample is extended to all children born in 2006 and 2007. All regressions based on our own data control for child's gender, age and migration stats. Robust standard errors in parentheses: *p < 0.1, **p < 0.05, ***p < 0.01