# The Daughter Penalty*

Sonia Bhalotra[†] Damian Clarke[‡] Angelina Nazarova[§]

March 31, 2025

**Abstract**

Looking at the earnings profiles of men and women after their first child is born, a number of studies establish that women suffer a larger penalty in earnings than men—a child penalty. Leveraging randomness in the sex of the first birth, we show that the child penalty in the UK is larger when the first born child is a girl. We label this the daughter penalty. Exploiting rich longitudinal survey data, we examine behavioural responses to the birth of a daughter vs. a son to illuminate the underpinnings of the daughter penalty. We find that the birth of a daughter triggers more household specialisation than the birth of a son, with mothers taking on a larger share of household chores and childcare. Mothers suffer a daughter penalty in mental health, while fathers report more satisfaction with their relationship. Our findings imply that girls and boys in the UK are, on average, growing up in different home environments, with girls growing up in households that, by multiple markers, are more gender-regressive. This is potentially a mechanism for the inter-generational transmission of gendered norms.

**JEL classification**: J2, J7, I3
**Keywords**: gender, child penalty, gender wage gap, mental health, parental involvement

*We thank seminar audiences at the Institute for Social and Economic Research, University of Essex, MiSoC, Barcelona Graduate School of Economics, EEA-ESEM and the RES for helpful comments. We are grateful to Dan Brown, Edith Aguirre, Michaela Benzeval, Emilia Del Bono, Thomas Cornelissen, Paul Fisher, Laura Fumagalli, Libertad González, Angus Holford, Laura Hospido, Renee Luthra, Alita Nandi, Arnaud Philippe and Birgitta Rabe for comments and suggestions. Clarke acknowledges the ANID Millenium Institute for Research in Market Imperfections and Public Policy, ICM IS130002 for financial and institutional support. Angelina Nazarova's contribution to this research was supported by the ESRC Research Centre for Microsocial Change at Essex Award ES/S012486/1 and Sonia Bhalotra's time was partially supported by ESRC ES/S012486/1 and partially by European Research Council Advanced Grant Evidence-VAW 885698.

[†]University of Warwick, IFS, CAGE, CEPR, IEA, IZA, CESifo. E-mail: Sonia.Bhalotra@warwick.ac.uk.
[‡]University of Exeter, Universidad de Chile, IZA & MIPP. E-mail: dclarke@fen.uchile.cl.
[§]Institute for Social and Economic Research, University of Essex, MiSoC, EEA, SIdE. E-mail: an21010@essex.ac.uk.

# 1 Introduction

A significant body of research has established that entry to motherhood has a substantial and enduring negative impact on women's labour market outcomes, in contrast to small impacts on fathers. The impact of childbirth on the outcomes of the mother relative to the father is called the child penalty, discussed in, among other places, Waldfogel (1997); Bertrand et al. (2013); Lundborg et al. (2017); Adda et al. (2017); Hotz et al. (2018); Kleven et al. (2019b,a, 2023); Berniell et al. (2024).

Using longitudinal survey data from the UK (the UK Household Longitudinal Study), we investigate, for the first time, whether the gender of the first child influences the size of the child penalty. Our identification strategy is based on the assumption that the gender of the firstborn child is random conditional on the decision to have a child, to defend which we show balance on a rich set of predetermined characteristics between families with a firstborn son vs. a firstborn daughter.[1] We then primarily follow the event-study approach of Kleven et al. (2019b), separately considering the effects of the birth of a daughter and the birth of a son on the labour market outcomes of their parents. In order to test for the significance of any differences we also estimate pooled models in which the birth event is interacted with the gender of the birth.[2]

In this particular sample of UK births occurring from 2010 onwards, we find a striking difference in the child penalty by the sex of the first child, which emerges immediately after birth and is persistent through the next five years. Averaged over the five years after birth, we estimate that the earnings penalty is approximately 3% among mothers of sons and 26% among mothers of daughters. This is mirrored in differential employment rates. The employment penalty is approximately 6% for mothers of sons and 20% for mothers of daughters. We label the difference in the child penalty (which is already differenced between mothers and fathers) between daughters and sons as the daughter penalty.

These intriguing findings imply that girls grow up in a systematically different household environment than boys. In particular, if during the early and influential years of life, mothers of girls are less likely to work than mothers of boys, this may contribute to women (compared with men) coming onto the labour market with different exposures to the notion of a working mother and hence with different preferences over work and family.[3]

[1] In the main analysis, we remove ethnic minorities from the data as the UKHLS oversamples them and we wish to avoid that manipulation of child sex through sex-selective abortion among Asians affects our findings for the broader population.

[2] Some of the identification concerns with the Kleven et al. (2019b) strategy do not apply in our setting as we are interested in the daughter-son difference conditional on birth, but we will discuss identification concerns.

[3] Our results, similar to the existing child penalty literature, flow from the first birth. Data from the UK Office of National Statistics (May 2023) indicate that of all families with dependent children, 44% have one child. If the sex of each child in a family is independent, then there is a 50% chance that in families with two children, a firstborn girl is followed by a girl. Thus, a large share of the population of girls grow up in households with the environment that we describe as following from a firstborn girl.

In their pioneering study set in Denmark, Kleven et al. (2019a) show that child penalties are transmitted through generations, from parents to daughters, but not sons. In other words, they show that girls growing up in families with a more traditional division of labour between the parents incur larger child penalties when they themselves become mothers.[4] All existing estimates constrain the child penalty in the first generation to be the same when daughters and sons are born. Our results identify a mechanism by which girls in fact are *more likely* than boys to grow up in families with a traditional division of labour. Given the heavy weight of early life exposures in shaping preferences, this could contribute to the stylized fact that women continue to make family-friendly adjustments to their careers (Cortés and Pan, 2023). Our results also extend the observation of Fernández et al. (2004) that men who grow up with working mothers are more likely to have wives that work. Our results identify a mechanism by which boys are *more likely* to grow up in households where mothers work.

So as to shine some light on what else is different between families in which the first child is a girl vs. a boy, we investigate how the sex of the first birth influences the division of household work and childcare between the parents, their social and political conservatism and attitudes towards gender-equality, their subsequent fertility, relationship quality and their mental health. Although event study estimates for some of these additional outcomes are imprecise, they reveal a fairly consistent pattern whereby families make a number of correlated adjustments that point in the direction of mothers taking on a more traditional home-maker role when a daughter (rather than a son) is born. As a result, daughters are, on average, growing up in more traditional gendered environments than sons. This provides a new perspective and rationale for the persistence of gender norms over time.

Descriptive statistics for our sample show that mothers dedicate more time to chores than fathers, irrespective of child gender, consistent with a vast literature (see e.g. Hancock et al., 2024), and our estimates show, as expected, that the burden of domestic chores jumps after the first child is born. What we find that is *new* is that the disproportionate burden of housework (and childcare) that falls upon mothers is larger among mothers of daughters than it is among mothers of sons. Following the birth of a first son, the gap between mothers and fathers in the share of household chores they report doing alone increases by 27%. However, following the birth of daughters, this gap is estimated to increase by 43%. These differences are just statistically significant, there is a clear pattern, the magnitudes are substantial, and a similar pattern emerges from different questions in the survey that probe time and responsibility in home production. For example, mothers are observed to be 57% more likely to self identify as the main child carer rather than fathers following the birth of a son. However, following the birth of a daughter, this quantity increases to 67%. This 10 percentage point difference in care

[4] Of relevance to this paper, Kleven et al. (2019b) make reference to this finding of Kleven et al. (2019a) and write "These findings are consistent with an influence of the family environment in the formation of women's preferences over family and career."

roles is both large and statistically significant.

We find that mothers compared with fathers become less progressive after a (first born) daughter is born (this is relative to their pre-birth outcome and relative to mothers whose first born is a son). Progressiveness is captured by an index of attitudes to gender equality (one of which asks the respondent the extent to which they agree that the husband should earn and the wife should stay at home), the social and political leanings of the news-sources they turn to, and whether they report having voted liberal or conservative. These findings cohere with our finding that parents of daughters have a more unequal division of market work and home production than parents of sons. They suggest that the sex of the child moves preferences.

We then considered whether the chances that parents have a second child depends upon whether the first is a girl or a boy. There is a vast literature documenting son preference in fertility in India and China, where families with a first-born girl are more likely to continue fertility, presumably with a view to having a son. However, son preference appears not to be unique to South and South East Asia. Dahl et al. (2008), for instance, show that shotgun marriages in the United States are more likely when the mother is carrying a son, indicating son preference among men. There is evidence from Denmark that fathers (not mothers) are a lot happier after a son is born (Kohler and Mencarini, 2016). In our UK data sample, we find that, for both parents, the chances of a subsequent child within five years of the first birth are higher when the first birth is a girl.[5] This result coheres with our other results—if parents anticipate building a larger family, they may be more likely to specialize, with one more focused on market work and the other on home production, and women in these same families may report less gender-progressive values.

We finally turned to measures of wellbeing, including relationship quality (ever considered divorce, or regret getting married) and mental health (from the SF12 scale). While relationship quality experienced by mothers does not depend on the child's gender, fathers of daughters are significantly more satisfied with their relationships than fathers of sons. This is consistent with the mothers of daughters taking on more household work and child care and reporting more traditional attitudes to family life. While the traditional division of work appears to be good for father marital satisfaction, it appears to exert a toll on women. Looking at a clinically verified mental health indicator, we see that mothers of girls suffer a larger drop in mental health than mothers of boys, while the mental health of the father is largely invariant to child gender.

To summarise, our primary contribution is that we provide new evidence showing that the child penalty in earnings and employment in the UK varies with the gender of the first born child. This has not been discussed before, as far as we know.[6] However it is not implausible

[5] This is similar to findings from India, for example, see Bhalotra and van Soest (2008), but note that our result is from a UK sample from which we have removed ethnic minorities.

[6] Lundberg and Rose (2002)show that father's labour supply in the United States responds to the birth of a son

a priori as child gender has been documented to impact a number of outcomes.[7] Concerns that arise in the child penalty literature over the identifying assumption that the timing of birth is idiosyncratic (Melentyeva and Riedel (2023), Bensnes et al. (2023)) are mitigated in our setting by our differencing between boy and girl births. Nevertheless, we investigate robustness of our estimates to allowing for time-varying entry to parenthood.

Our second contribution is to show that the daughter penalty on earnings is mirrored in a daughter penalty in the burden of housework and childcare, with an increase in the chances of continued fertility, improved marital satisfaction among fathers, and deterioration of mental health among mothers. While we would not want to overstate this, there is some evidence that preferences shift as mothers of daughters reveal less progressive gender attitudes and select into less progressive news-consumption relative to their pre-birth levels, relative to mothers of sons, and relative to fathers.

Our third contribution is to suggest that the daughter penalty (or, put differently, the larger child penalty associated with daughters relative to sons) is a potential mechanism for the inter-generational transmission of norms. Girls are more likely than boys to have mothers (relative to fathers) who have weaker career attachment, who engage in a more gender-stereotypical division of household work, who suffer postnatal depression, and who are less likely to have gender-equal views. They are also more likely to have a younger sibling, which possibly reinforces the nurturing or caring responsibilities that are often associated with women.

Many of the estimates of child penalties in the literature are on administrative data—which have the advantage of size—but which typically do not contain information on housework and childcare, wellbeing or preferences. Our data constrain us with relatively small sample sizes, but they contain this additional information. Using the rich question bank, we are able to reveal the inner workings of the household that underlie (or are associated with) the child penalty in earnings.

Although we are underpowered to examine differences across subgroups in detail, we establish that the daughter penalty holds for parents with a college-degree and for parents without. Indeed, while the child penalty is smaller among degree-educated parents, the daughter penalty is larger. The daughter penalty is, similarly, larger among parents who report more progressive gender norms (indicated by disagreement with the statement that the family suffers if the mother

rather than a daughter. However we know of no previous work documenting responses of mothers and fathers to child gender, over time, and with analysis of correlated decisions on household production, together with indicators of stress and marital satisfaction.

[7] These include relationship formation and dissolution (Bedard and Deschênes, 2005; Dahl et al., 2008; Kabátek and Ribar, 2020), future fertility (Das Gupta et al., 2003; Jayachandran, 2017; Bhalotra and Cochrane, 2010; Pollard and Morgan, 2002), views on women's rights (Powdthavee and Oswald, 2010; Washington, 2008; Ronchi and Smith, 2024), the risk that the father commits crime (Dustmann and Landersø, 2018), and extraversion and earnings of the father (van Lent, 2022).

works full-time).

While the notion that women bear the brunt of child birth is by no means new, recent evidence establishing the persistence of the child penalty through as long as ten years after birth (Kleven et al., 2019b) has caught the imagination of economists. Replications and extensions of the analysis show that the penalty is pervasive, being evident across a range of cohorts and countries (Kleven et al., 2023; Berniell et al., 2021, 2024), albeit with varying magnitude. Evidence that the child penalty exists among adoptee parents as much as among biological parents (Andresen and Nix, 2022b), and that it varies with gender norms (Kleven et al., 2019b; Kleven, 2022) and policy interventions (Andresen and Nix, 2022a) implies that it is driven by behavioural choices. There is a large and growing literature which indicates that the child penalty (in OECD countries) explains most of the (average) gender gap in earnings. There is no simple, well-identified explanation of the variation in the child penalty across countries but, in the OECD, it is correlated with gender identity considerations and social norms (Berniell et al., 2021; Kleven et al., 2019b; Cortés and Pan, 2023).[8] Our findings suggest that, at least in the UK, daughters trigger greater gender differentiation among the parents and, notably, the daughter penalty (albeit not the child penalty) is larger among parents with a degree education and parents who report more progressive gender norms. The identifying assumptions of the child penalty literature have been discussed and, in cases, challenged (see Section 3).

The remainder of this paper is structured as follows: In Sections 2 we discuss the UK Household Longitudinal Survey data used here, along with the definition of key variables and a presentation of key summary statistics. In Section 3 we lay out our empirical approach. Section 4 presents our findings, including potential correlated mechanisms. We briefly conclude in Section 5.

# 2 Data and Descriptive Statistics

## 2.1 Data Sources and Estimation Sample

Our analysis is conducted with data from the UK Household Longitudinal Study (UKHLS), a long-standing survey produced by the Institute of Social and Economic Research at the University of Essex. This survey has been conducted in a representative sample of UK households across 13 waves between 2009 and 2023. Each wave comprises interviews over a 24 month time frame, for example the first wave has interviews conducted in the period of 2009–2011 and the last wave in 2021–2023. The survey measures fertility histories, employment, hours,

[8] The child penalty is relatively high in the UK, just short of that in Germany and Austria, higher than that in the US and a lot higher than that in Sweden and Denmark (Kleven et al., 2019b). Looking at the World Values Survey (Wave 7, 2017-2022), we observe that the UK stands out as having a relatively high fraction of respondents agree with the statement that the family is very important. This fraction is 92.3% in the UK, while the corresponding figures are 90.7% in Sweden, 88% in Germany, 87.6% in Denmark, 85.4% in France and 76.7% in the Netherlands.

earnings, education and demographics, and it has a rich battery of questions pertaining to time use and preferences. Respondents are asked whether they have ever had any biological children before the start of the sample period. The labour market outcomes refer to the twelve months preceding the survey date.

We restrict our sample to individuals aged 16-64 who have not had a biological child before the sample period begins i.e. we drop individuals who are already parents when the survey starts in 2009. We further exclude those who have a firstborn child in the first wave because we require pre-child birth measures of labour market outcomes. We then drop individuals who do not give birth at all during the analysis period, as we will study outcomes of parents, comparing those with first-born girls vs. first-born boys. In our analysis, we focus on the period before the pandemic, thus our sample consists of ten waves over the period from 2009 to 2019. We construct a sample that is balanced across the 10 observation years, such that we observe each parent in each wave. For a basic schema of the data structure see Table A1 though note that our balanced panel does not imply balance across all periods relative to child bearing, given that we will not observe many leads for individuals giving birth in early waves, or many lags for individuals born in later waves. We report estimates for 3 waves pre-birth and 5 waves post-birth. We also estimated the main equations on the larger unbalanced sample. Our main findings, reported below, hold in this sample too (See Appendix B, Figures B1–B3).

The balanced sample includes 7,480 parent×year observations of individuals who have at least one child during the sample period. Similar to Kleven et al. (2019b), we include parents irrespective of whether they are married, cohabiting or single. In this sample, 87% of individuals are living with a partner in the household (this partner may or may not be the biological parent of their child) and 13% do not reside with the parent of the child or any other partner.[9] We exclude the small proportion of adopted children (1.5% of the sample) because the gender of the adoptee cannot necessarily be assumed to be random. We drop ethnic minorities, (10% of all parents as the UKHLS oversamples minorities) given evidence of son preference in such groups, even in societies where son-preference is uncommon (Kaushal and Muchomba, 2018). However, including the ethnic minority sample in the main analysis does not change the broad pattern of the results we report below.[10]

**Labour market outcomes** Regarding labour supply outcomes we focus on three variables: (i) employment, (ii) total hours worked, and (iii) monthly earnings. Descriptive statistics of these and all variables at baseline are provided in Appendix A, Table A2. Employment is a dummy variable that takes the value 1 if an individual is employed, and 0 if they are not. To construct

[9] At birth 5% of mothers did not have a partner in the household (12% one wave before birth) in comparison with 0% of fathers at birth (5.5% of fathers one wave before birth)

[10] The ethnic minority sample is not large enough for us to analyse it independently. We do however report results from all main analyses for the combined sample which *does not* remove these individuals from the sample; see Appendix B Figures B4–B6.

this variable, we follow the Office for National Statistics (ONS) definition of employment. An individual is employed if they report either doing paid work in the previous week or if they did not report having done paid work but report having a paid job. An individual is also employed if they are on a government training scheme, or are doing unpaid work for a family business.[11] Total hours are measured by the weekly hours that an individual has been working. In order to generate this variable we add together regular weekly hours with the number of overtime hours. This value is positive if an individual is in paid employment and 0 otherwise. We exclude self-employed individuals from our principal measure of hours worked given that hours worked are collected in a slightly different way for self-employed individuals.[12] Monthly earnings is the earnings of an individual adjusted by yearly inflation (base year 2015), where the year refers to the year of the interview.

In Appended analyses, we additionally consider the usual monthly net pay reported by individuals. The main reason is that net pay offers a reflection of disposable income available for everyday living expenses, savings and investments. By considering both real and net earnings, we have a better overview of parental financial situation. We are also interested in the probability that an individual is engaged in part-time work to understand whether parents tend to switch to part-time hours due to increased childcare responsibilities, and whether this can partially explain the decrease in intensive margin measures of labour supply after birth.

**Home production** We measure the parent's involvement in domestic work and childcare. To measure domestic work, we rely on the reported contribution that each parent makes to household chores which is the fraction of six household chores that a respondent reports completing mostly themselves. The six chores are grocery shopping, cooking, cleaning, washing/ironing, gardening, and DIY jobs. As a robustness check, we also examine the number of hours spent on housework, which excludes childcare. In order to measure parental involvement in childcare, we examine how care responsibilities are shared among partners. Specifically, we generate a binary indicator measuring whether an individual reports being the main person who cares for their child (rather than care being provided by their partner, or being shared). These categories are reported by respondents following these three options defined in the UKHLS. Less than 5% of respondents report being co-main carers, and as such results are not sensitive to considering whether an individual reports being the main *or* co-main carer.

**Family formation** We measure fertility continuation and cohabitation. We create an indicator for whether an individual has an additional child following the birth of the first child.

[11] In our main sample only 7 people are in unpaid family work, and so it is relatively innocuous to include these observations in estimation.

[12] We also create an alternative measure of weekly hours worked, which includes the self-employed (using a separate variable for the self-employed from the UKHLS, variable description in Data Appendix A). We discuss the results using this measure at the end of Section 4.

**Progressive parental beliefs** We use data from the UKHLS on two dimensions of progressiveness: voting behaviour and progressiveness in terms of gender equality. Respondents are asked which party they voted for in the last election, and we code an indicator for having voted for a right-wing party. We then define being progressive as 1 (not voting for the right-wing party) and 0 otherwise.[13]

We combine answers to five attitude questions to create an index measuring progressiveness in attitudes to gender. The questions ask respondents whether they believe that (a) pre-school children suffer if mothers work; (b) the family suffers if mothers work full-time; (c) the husband should earn and the wife should stay at home; (d) if the husband and wife should contribute to household income; and (e) employers should help mothers combine jobs and childcare. For each question, the respondent is provided with a Likert scale for the degree to which they agree. We create an additive index such that the higher the score, the more gender-progressive the individual is.

We create another indicator of progressiveness derived from information on which newspapers the respondent reads. We use the Leveson Report to categorize news sources (Leveson, 2012). The Leveson Inquiry was a public inquiry into the cultures, practices, and ethics of the British press, and its report names five news sources that tend to publish damaging portrayals of women. The UKHLS questions provides a list of news sources for the respondent to check and four of these five regressive news sources is included in the list. We create an indicator for reading progressive news if the respondent does not read news from any of these four sources, which are the newspaper or website version of The Sun, The Daily Mail, The Daily Express, or The Daily Star.

**Relationship satisfaction** We also model the effect of having a boy or girl on satisfaction with and, implicitly, stability of the marital relationship. We use several variables which are available for those respondents who have a partner in the household. First, we create a dummy variable taking the value of 1 if the parent has never considered divorce/separation after having the first-born child, or 0 otherwise. Second, we create a dummy variable taking the value of 1 if the parent reports never regretting getting married after having the first-born child, or 0 otherwise. Finally, we create an additive relationship happiness index, which comprises each individual's response to the following questions: (i) whether they never regret getting married/living together, (ii) whether they never get on each others' nerves, (iii) whether they never consider divorce/separation, (iv) whether they never argue, (v) whether they report being very happy with the relationship. By construction, the higher the value, the happier with the relationship the parent is.

[13] We include parties represented in the House of Commons as well as parties without representation in the House of Commons, but with representation in other UK legislatures. This information comes from the party registries of Electoral Commission.

**Parental mental health and well-being** To evaluate mental health, we draw on a widely-used series of questions from the UKHLS, which measures mental health with the SF-12 scale (Ware et al., 2001). This scale ranges from 0 to 100, where higher values indicate better mental health. This continuous measure allows for a more detailed understanding of the psychological state rather than a simple binary assessment. A high score suggests a better capacity to cope with daily challenges, whereas lower scores may indicate mental health struggles that could affect other aspects of life. We consider a range of additional measures of well-being as Appended results. These are a measure of an individual's life satisfaction, whether they have recently suffered depression, have feelings of worthlessness, and their perceived ability to face problems. These are all gleaned from questions which ask individuals precisely about these factors. Exact definitions of each question are provided in Appendix A.2.

## 2.2 Descriptive Statistics

In Table 1, we present summary statistics reporting key characteristics of all individuals in our estimation sample at baseline (wave 1, collected in 2009-2011), which is when they are childless by definition. In Appendix A, Table A2 we present full summary statistics. The sample is predominantly urban and UK-born. Men in our sample are slightly older than women (27 years compared to 25 years), and earn higher monthly wages (2,385 GBP compared to 1,682 GBP), work more hours on average in a given week (35 versus 32), though have lower rates of employment (86 and 90 percent among men and women respectively). These values are in agreement with values reported by the Office for National Statistics (2021) when considering young women and men without dependent children.

Table 1: Summary Statistics at Baseline (Wave 1)

| | Men | Women | Total | | | |
|---|---|---|---|---|---|---|
| **Individual Characteristic** | Mean | Mean | Mean | SD | Min | Max |
| Age | 27.02 | 25.05 | 25.93 | 5.49 | 16 | 48 |
| Married | 0.60 | 0.58 | 0.59 | 0.49 | 0 | 1 |
| Urban | 0.82 | 0.81 | 0.82 | 0.39 | 0 | 1 |
| UK born | 0.92 | 0.92 | 0.92 | 0.27 | 0 | 1 |
| Has a degree | 0.35 | 0.38 | 0.37 | 0.48 | 0 | 1 |
| GSCE only | 0.25 | 0.21 | 0.23 | 0.42 | 0 | 1 |
| Hours worked | 35.27 | 32.10 | 33.45 | 17.11 | 0 | 91 |
| Employed | 0.86 | 0.90 | 0.88 | 0.32 | 0 | 1 |
| Monthly earnings | 2385.04 | 1682.52 | 1972.40 | 1307.17 | 25.98 | 9237.88 |
| Observations | 307 | 441 | 748 | | | |

*Notes:* Summary statistics in wave-1 (2009-2011) for the analysis sample—a balanced panel of individuals observed in all waves from 2009-2019 who are childless in the first wave and who become parents in the sample window. The sample excludes ethnic minorities and parents of adopted children. Longitudinal weights from wave 10 are used.

**Balance in baseline parental characteristics by gender of firstborn child** Of all firstborn children in our sample 49% are boys and 51% girls. Our analysis rests on the premise that the gender of the first child is random conditional on the decision to have a child.[14] Table 2 shows weighted means for parental characteristics in wave 1 by the eventual gender of parents' first-born child, and the $p$-values for tests of differences in means. The weight we use is a longitudinal weight covering all individuals in the adult main interview given that our analysis uses data from 10 consecutive waves. The weights in the UKHLS dataset adjust for unequal selection probabilities, differential non-response, and potential sampling error.[15] We cannot reject equality across child gender at the 5% level in all variables considered, and for a single variable (whether an individual has a degree), observe a marginally statistical difference across samples. For mothers we also observe a slight mis-balance in a derived relationship index. However, this is one or two of 18 tests considered for each of mothers and fathers, and hence in line with rejection rates one would expect even if all null hypotheses were true. Namely, across a large sample of tests, if the null hypothesis of equality of means was correct, we would expect 10% of tests to result in a $p$-value of $\leq 0.10$. Across 36 tests, we would expect 3.6 rejections at 10% if the null is true, which is relatively close (and indeed slightly higher) than the number of rejections we observe here (3).

While in line with expected rates of type I error, in empirical models we consistently control for education, which has the additional theoretical benefit of increasing the precision of estimation and inference. In Appendix Results we will additionally consistently document estimates separately stratifying by parents' education. Namely, we will present results for parents with a degree and without a degree for all models. If results are observed to hold *within* educational groups, this would suggest that any documented results pooled across groups do not simply owe to mis-balance, as estimating in separate groups resolves this mis-balance. As well as providing a test for mis-balance, this will provide a richer description of heterogeneity, and allow us to assess the weight of our findings across quite different individuals.

[14] In fact, this assumption is sufficient for us to estimate causal differences in labour market penalties by the gender of the child for either mothers *or* fathers. If we instead refer to differences in penalties between mothers and fathers, we rely on the slightly weaker assumption than any non-random selection of child gender is identical across fathers and mothers. Because we conduct formal tests of equality of penalties by child gender for mothers and fathers separately, we generally maintain the prior assumption.

[15] More information about the weights can be found in the UKHLS weighting guidance and main survey user guide.

Table 2: Balance Tests by Gender of the First Child at Baseline

| | Fathers | | | Mothers | | |
|---|---|---|---|---|---|---|
| | Daughter | Son | p-value ($H_0$: Diff=0) | Daughter | Son | p-value ($H_0$: Diff=0) |
| Age | 27.13 | 26.92 | 0.774 | 25.02 | 25.08 | 0.913 |
| Married | 0.93 | 0.93 | 0.837 | 0.85 | 0.81 | 0.310 |
| Urban | 0.81 | 0.83 | 0.766 | 0.82 | 0.80 | 0.578 |
| Has a degree | 0.30 | 0.40 | 0.078 | 0.34 | 0.43 | 0.063 |
| **Labour market outcomes** | | | | | | |
| Monthly earnings | 2422.97 | 2348.08 | 0.707 | 1697.23 | 1665.68 | 0.791 |
| Weekly hours | 35.41 | 35.13 | 0.905 | 31.41 | 32.91 | 0.372 |
| Weekly hours c. on employment | 41.85 | 41.48 | 0.798 | 35.53 | 35.99 | 0.737 |
| Employed | 0.86 | 0.87 | 0.821 | 0.89 | 0.92 | 0.346 |
| **Mental health** | | | | | | |
| Mental health (SF12) | 52.29 | 53.01 | 0.464 | 50.38 | 48.85 | 0.113 |
| **Home production** | | | | | | |
| Fraction of chores alone | 0.35 | 0.33 | 0.595 | 0.38 | 0.34 | 0.295 |
| Childcare mostly on self | 0.05 | 0.03 | 0.208 | 0.69 | 0.65 | 0.189 |
| **Family formation** | | | | | | |
| Having an additional child | 0.11 | 0.10 | 0.595 | 0.12 | 0.11 | 0.380 |
| Never consider divorce/separation | 0.96 | 0.98 | 0.493 | 0.99 | 0.99 | 0.994 |
| Never regret marriage/cohabitation | 0.96 | 0.98 | 0.410 | 0.99 | 0.99 | 0.422 |
| Relationship happiness index | 4.28 | 4.39 | 0.416 | 4.25 | 4.41 | 0.081 |
| **Parental values and attitudes** | | | | | | |
| Progressive news source | 0.71 | 0.78 | 0.211 | 0.76 | 0.67 | 0.140 |
| Progressive voting | 0.54 | 0.44 | 0.209 | 0.65 | 0.68 | 0.631 |
| Gender equality | 19.04 | 18.47 | 0.402 | 19.46 | 18.89 | 0.334 |
| Parent × Year | 1490 | 1580 | 3070 | 2350 | 2060 | 4410 |

*Notes.* Sample as in Table 1. $p$-values report a test of difference of means, one variable at a time. Characteristics are measured in the first wave of the survey which, by construction, is before the first child is born. Certain characteristics are not available in the first wave due to data availability (fraction of chores alone, gender index), and as such are measured in the wave prior to birth. Being married is also measured in the wave prior to birth. Characteristics available only in the post-birth period (childcare, having an additional child) are measured across five post-birth periods. Sample excludes ethnic minorities. The longitudinal weight from wave 10 *j_indinus_lw* was used for all weighted means.

# 3 Empirical Strategy

**Baseline model** We begin by estimating the canonical child penalty model of Kleven et al. (2019b). This is:

$$Y_{ist}^g = \mu_0 + \sum_{j \neq 1} \alpha_j^g \cdot \mathbb{1}[j = t] + \sum_k \beta_k^g \cdot \mathbb{1}[k = age_{is}] + \sum_j \gamma_y^g \cdot \mathbb{1}[y = s] + \delta X_{is}^g + \varepsilon_{ist}^g, \quad (1)$$

estimated for parent $i$ in survey wave $s$ and event time (time to birth) $t$, stratified by parent gender $g$. Coefficients $\alpha$ capture time to birth effects, $\beta$ capture age fixed effects, and $\gamma$ capture survey wave (mother's birth cohort) fixed effects. Baseline covariates $X_{is}$ consist of indicators of whether mothers or fathers have a tertiary degree, as discussed in Section 2. All coefficients are separately identified up to an omitted baseline reference group (arbitrarily set to 1 year prior

to birth in the case of $\alpha$) given that there is variation in the age at which women of any birth cohort give birth in each survey wave. Time to birth effects capture dynamics around child-birth, and the statistic of interest is the proportional change in earnings surrounding child birth:

$$P_t^g \equiv \frac{\widehat{\alpha}_t^g}{E[\widetilde{Y}_{ist}^g|t]}.$$

Following Kleven et al. (2019b), each coefficient $\widehat{\alpha}_t^g$ is standardised by the predicted earnings of a parent of age $k$ at calendar year $y$, omitting the contribution of time to birth dummies, namely:

$$\widetilde{Y}_{ist}^g = \widehat{\mu}_0 + \sum_k \widehat{\beta}_k^g \cdot \mathbb{1}[k = age_{is}] + \sum_y \widehat{\gamma}_y^g \cdot \mathbb{1}[y = s].$$

Thus, $P_t^g$ measures proportional salary changes at event year $t$ compared with the regression-generated counterfactual outcome for the same individual absent children. The counterfactual is estimated off all other parents from the data who at one point are aged $k$, and whose salary is also observed in calendar year $y$. In baseline models we report the series of coefficients $P_t^g$ over years $-3, \ldots, 5$, omitting $t = -1$ as the baseline reference category. Note that because $P_t^g$ refers to a sequence of *proportional* changes in parental salary around time to childbirth $t$, this implies that even if there are pre-birth differences in salaries between fathers and mothers, the estimated coefficients will be comparable across models as proportional declines surrounding childbirth. This would not be guaranteed if instead we reported the un-adjusted coefficients $\widehat{\alpha}_t^g$.

We present coefficient estimates graphically, consistently plotting $P_t^g \; \forall \; t$ separately for $g \in \{mother, father\}$ along with 95% confidence intervals based on heteroscedasticity-robust standard errors. Along with point estimates and confidence intervals for each group $g$, we calculate the child penalty for women relative to men at event time $t$ as:

$$P_t = \frac{\widehat{\alpha}_t^{father} - \widehat{\alpha}_t^{mother}}{E[\widetilde{Y}_{ist}^{mother}|t]}. \tag{2}$$

This quantity measures the percentage by which any labour market decline for women exceeds the decline observed for men at event time $t$. As well as reporting event-specific quantities $P_t$, we report this quantity averaged over the 5 years following birth in resulting plots. This is accompanied by standard errors based on a block bootstrap resampling individuals in the panel.

**Heterogeneity by child gender** Our point of departure from the literature lies in considering whether parental labour market and household production choices respond differently to the birth of a first child depending upon the child's gender. To investigate this, we fully interact the specification from Kleven et al. (2019b) in Equation (1) with an indicator for the child's gender,

specifically $Girl_{is} = 1$ indicating that the first born is a girl rather than a boy. Formally, this is:

$$\begin{aligned} Y_{ist}^{g} &= \sum_{j\neq 1} \alpha_j^g \cdot \mathbb{1}[j=t] + \sum_{j\neq 1} \left(\alpha_j^{Girl,g} \cdot \mathbb{1}[j=t] \cdot Girl_{is}^{g}\right) \\ &+ \sum_{k} \beta_k^g \cdot \mathbb{1}[k=age_{is}] + \sum_{k} \left(\beta_k^{Girl,g} \cdot \mathbb{1}[k=age_{is}] \cdot Girl_{is}^{g}\right) \\ &+ \sum_{j} \gamma_y^g \cdot \mathbb{1}[y=s] + \sum_{j} \left(\gamma_y^g \cdot \mathbb{1}[y=s] \cdot Girl_{is}^{g}\right) \\ &+ X_{is}^{g\prime}\delta + \left(X_{is}^{g\prime} \cdot Girl_{is}\right)\delta^{Girl} + \mu_0 + \mu_0^{Girl} \cdot Girl_{is} + \varepsilon_{ist}^{g} \end{aligned} \tag{3}$$

Here, in the spirit of Chow (1960) we are interested in testing the equality of coefficients across models.[16] In particular, we wish to determine whether dynamic responses to childbirth differ by firstborn child gender. Any differential responses will be captured by interactions of time to birth indicators with $Girl_{is}$, with negative values indicative of *larger* trade-offs following girl, rather than boy, children. This suggest a formal test of equality of child penalties by gender, which is the hypothesis:

$$H_0 : \alpha_0^{Girl,g} = \alpha_1^{Girl,g} = \cdots = \alpha_5^{Girl,g} = 0$$

against the alternative that $H_0$ is not true. In the main results, we will present estimates for wage dynamics around the birth of boys from Equation (3) (coefficients $\alpha_j^g$), wage dynamics around the birth of girls ($\alpha_j^g + \alpha_j^{Girl,g}$), and then present tests for the equality of child penalties across boys and girls ($\alpha_j^{Girl,g}$). These will be presented for both $g \in \{mother, father\}$. We conduct identical standardisations as laid out in (2) so that penalties are expressed in proportional terms for mothers of boys and girls (and fathers of boys and girls). Thus, as discussed above, our patterns of interest capture *proportional* declines in salary, and are comparable even if salaries across particular groups within the data are *not* identical in the lead up to childbirth.

Finally, and analogously to (2), we calculate the daughter penalty for women relative to men as:

$$P_t^{daughter} = \frac{\widehat{\alpha}_t^{Girl,father} - \widehat{\alpha}_t^{Girl,mother}}{E[\widetilde{Y}_{ist}^{mother}|t]}. \tag{4}$$

This is the *differential* effect of a daughter compared to a son on mothers compared to fathers. In all graphical output we report both event-specific quantities $P_t^{daughter}$, as well as this quantity averaged over the 5 years following birth, with block bootstrap standard errors.

**Identifying assumptions** The identifying assumptions of the Kleven et al. (2019b) model are laid out in the original paper. The additional assumption that we make is that the gender of

[16] Alternatively, this specification is equivalent to stratifying on child gender and estimating Equation (1) separately. While we could proceed with separate models, we implement this joint model as it provides a simple direct test of coefficient equality.

the firstborn child is random. To defend this, we demonstrate balance on observables between families whose first child is a son vs. a daughter. We may nevertheless be concerned that child gender is correlated with a relevant unobservable, i.e. an unobservable that is also correlated with the labour market outcomes of mothers vs. fathers. A natural candidate for this is maternal health—Trivers and Willard (1973) argued that there is an evolutionary basis for sons to be more likely to be born to healthier women—and this does hold in US census data (Almond and Edlund, 2004). The UKHLS contains a measure of self-reported physical health. We model this as an outcome and find no difference between parents of daughters vs. sons.

**Other outcomes** We investigate a number of secondary outcomes (which we loosely refer to as mechanisms) to understand any differential labour market responses to child gender. These include family formation, home production, parental mental health and parental attitudes and values. These are, strictly, not mechanisms but other behaviours that are, in principle, jointly determined with labour market choices after birth — and after the sex of the birth is revealed. We consider these outcomes alongside labour market outcomes with a view to illuminating the decision-making that might lie behind any variation in the child penalty by child gender.

Where these measures are systematically measured before and after child birth—for example time spent on household chores—we simply estimate Equations (1) and (3), with the mechanism variable as the outcome. However, some potential mechanisms are, by their nature, not observed prior to birth, for example, time spent on childcare. In these cases, we anchor at 0 all outcomes in pre-birth periods when estimating Equation (1) and Equation (3). In these cases, we do not examine pre-birth periods as trends will be parallel and fixed at 0 by construction, nor do we conduct the standardisation described in (2) given that the denominator will be close to 0 in these cases. As a result, estimates for these outcomes will be presented in absolute changes rather than in standardised changes.

**Time-varying entry to parenthood** The models of Kleven et al. (2019b) – and earlier implementations such as Korenman and Neumark (1992) – have been influential and widely-adopted in a broader literature on child penalties (see, e.g. Berniell et al., 2021; Kleven et al., 2019a; Li, 2021; Andresen and Nix, 2022b; Kleven et al., 2023; Fiva and King, 2024). However, recent papers have noted a number of issues with such pooled-event study style models, and Kleven et al. (2019b) themselves noted concerns with extrapolating coefficients over multiple post-birth periods. Given the popularity of Kleven et al. (2019b) and the relevance of using this as a baseline model to benchmark estimates and compare them to a wider literature, in our main analysis, we follow Kleven et al. (2019b), which implies using data only on parents. Earnings (or other outcome) counterfactuals $\widetilde{Y}_{ist}^{g}$ are thus based on individuals who have a first child either earlier or later than the index parent, implying that, potentially, some counterfactuals are contaminated by dynamic effects of childbirth. Given this, we conduct a specification test, presenting addi-

tional estimates obtained by comparing individuals who have given birth with individuals who have not yet given birth, thus avoiding so called "forbidden comparison" problems which is well known in cases of staggered event adoption with heterogeneous effects over time (Sun and Abraham, 2021; Borusyak et al., 2024). Specifically, we follow Callaway and Sant'Anna (2021), in grouping time-to-birth specific quantities:

$$ATT(b,\ell) = E[Y_{b+\ell} - Y_{b-1}|\text{Parent}_b = 1] - E[Y_{b+.\ell} - Y_{b-1}|\text{Not Yet Parent}_{b+\ell} = 1],$$

where $b$ refers to birth cohorts (parents giving birth at a specific time), and $\ell$ lags and leads relative to these birth dates. This quantity thus calculates a double-difference comparing all parents who give birth at time $b$ comparing their outcomes at time $b + \ell$ to those the year before birth, to non-parents over the same time period. We aggregate these quantities across all birth cohorts $b$, and report average pre- and post-birth effects in a panel event study style framework, with inference following Callaway and Sant'Anna (2021). This procedure can be viewed as analogous to the difference-in-differences approach presented by Kleven et al. (2019b) as an appended result, accounting for the forbidden comparisons problem which Melentyeva and Riedel (2023) show is potentially relevant in this setting. While these estimates will be based on the identical sample of parents as that used previously, these are capturing absolute effects of child-bearing rather than proportional changes as in (1), and so are presented as Appendix A robustness checks focused only on child gender-specific responses among mothers and fathers (Figures B7-B10).

A second recently-documented concern is that child bearing itself is endogenously timed. Both Melentyeva and Riedel (2023) and Bensnes et al. (2023) note that the timing of child-birth can interact with labour market wages profiles in a complex way depending on both parental age (Melentyeva and Riedel, 2023), as well as parent's expectations regarding future wage paths (Bensnes et al., 2023). While flat pre-trends are often used to argue that such concerns can be ignored, Bensnes et al. (2023) have documented that if parents time births to occur at precisely the point when their labour market profiles flatten out, we would observe both flat pre-birth trends, and divergent counterfactual post-birth trends, in violation of the assumptions of these models. While models such as (1) *are not* robust to this critique, if such endogenous timing decisions are identical between parents who go on to have girls and those who go on to have boys—as is likely the case given that gender is as good as random—results which present *differences* by child gender should purge such violations of parallel post-birth trends. Given that our interest in this paper is precisely in documenting differential responses to girls compared to boys, we argue that the concerns documented by Melentyeva and Riedel (2023) and Bensnes et al. (2023) are likely to be minimised, as we are effectively differencing them out in equations such as (3) and (4).

# 4 Results

## 4.1 The child penalty in parental labour market outcomes

We estimate the effects of parenthood on earnings, employment, and work hours, with results presented in Figure 1. The event-time-specific effects represent the estimated coefficients $\widehat{\alpha}_t^{mother}$ (green line) and $\widehat{\alpha}_t^{father}$ (grey line) and denote the average proportional changes at event time $t$. We present 95% confidence intervals along with point estimates.

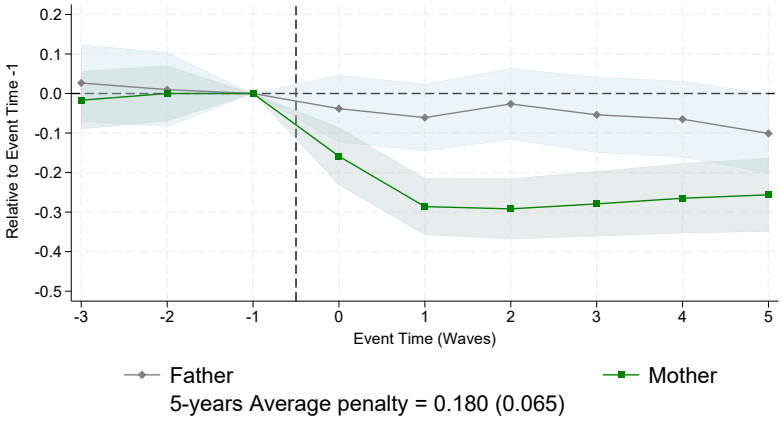
After partialling out life-cycle and calendar year effects, earnings, the probability of being employed, and total working hours follow similar trajectories for mothers and fathers before their first birth. Figure 1(a) shows that in the year of birth, women's earnings drop by almost 20%, while men's earnings decrease slightly (around 2%). Over the 5 years following birth, the motherhood penalty is estimated at 18%, and this penalty remains large even five waves after birth. While large, this average 18% earnings penalty across five waves post-birth is smaller than that reported by Kleven et al. (2019a). In comparison, Kleven et al. (2019a) report a 10-year earnings penalty of 44% and a participation penalty of 44% for women in the UK. Kleven et al. (2019a) use the BHPS data and focus on individuals aged 20-45, who are observed at least eight times over the event-study window, as well as at least once before and once after birth between 1991 and 2009. We follow Kleven et al. (2019a) in using gross labour earnings, excluding taxes or transfers, specified in levels, however, we use monthly gross earnings instead of annual earnings and we adjust for inflation. Also, instead of defining the participation as having positive earnings in the previous year, we follow the ONS definition of employment—whether the person was engaged in paid work in the previous week or reported having a paid job.[17] Additionally, our study follows people each year between 2009 and 2019 and excludes adopted children and ethnic minorities, whereas Kleven's estimates encompass a population without these specific exclusions.

Figure 1(b) show an employment penalty that averages to about 13% across the five years following the first child's birth. This effect is observed immediately in the wave when individuals gave birth, and is sustained over the horizon examined. Similar patterns are observed for total hours conditional on employment in panel (c). Here we observe an even sharper drop, with an estimated motherhood penalty of 24.2% when averaged over the entire post-birth period considered. Once again this is large and sustained, driven by declines in employment among women.

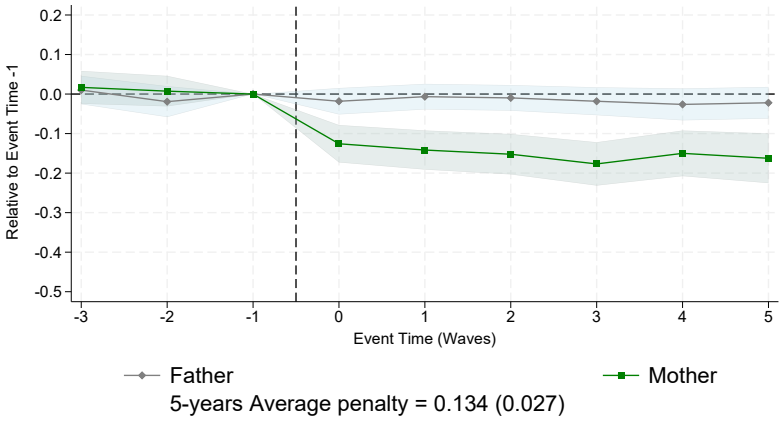
[17] More information on all points related to variable definitions can be found in Appendix A.2.

Figure 1: The child penalty – Labour market outcomes

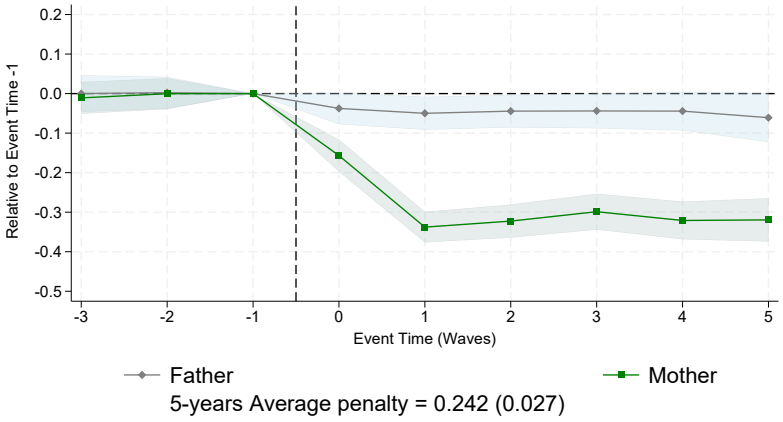
(a) Monthly earnings




(b) Employment




(c) Working hours conditional on employment



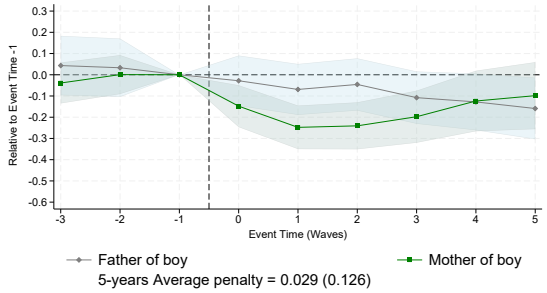

**Notes**: Estimates of $\widehat{\alpha}_j^g$ are presented following equation (1) and (2) with outcomes indicated in caption titles. Estimates are presented separately for mothers (green line) and fathers (grey line). The estimation is on a sample of parents who had their first child between 2009 and 2019. Shaded blue and green areas represent 95% confidence intervals based on robust standard errors.

## 4.2 Heterogeneity by the gender of the firstborn child

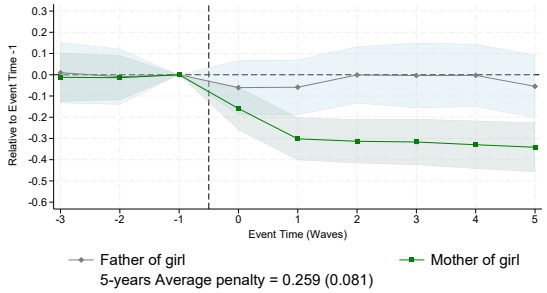
We first estimate Equation (1) separately for firstborn sons and firstborn daughters, and then estimate the interaction model from Equation (3) to test for the significance of child gender.

Figure 2: The child penalty in labour market outcomes by gender of the firstborn child

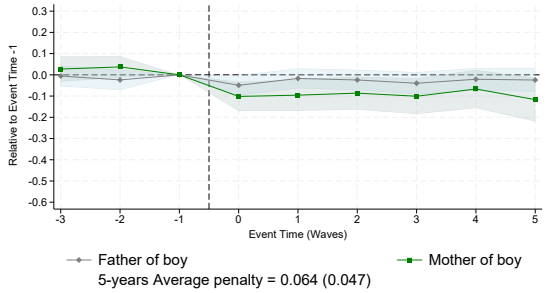
(a) Monthly earnings: Boy



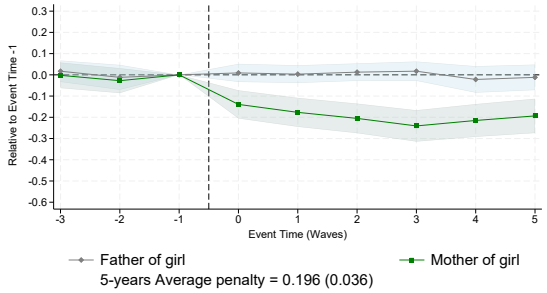

(b) Monthly earnings: Girl



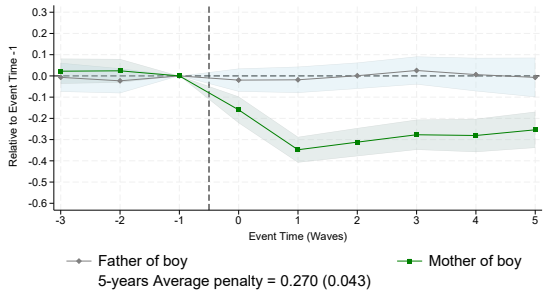

(c) Employment: Boy



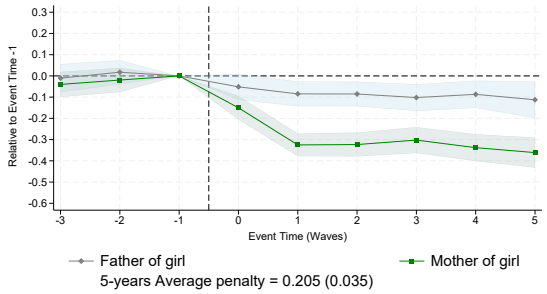

(d) Employment: Girl




(e) Hours conditional on employment: Boy




(f) Hours conditional on employment: Girl




**Notes**: Refer to notes to Figure 1. Identical estimates are presented, however now estimating separately for parents of boys (left-hand column) and parents of girls (right-hand column). All other details follow those described in Figure 1.

### 4.2.1 Labour market outcomes by child gender

Figure 2 documents the estimated child penalty for mothers (relative to fathers), separating the sample into parents whose first child is a son (left-hand columns) vs. a daughter (right-hand column). Mothers of girls experience larger declines in earnings and employment than mothers of boys. However, conditional on employment, the decline in hours worked is similar. As is already known, father's labour market outcomes are much less responsive to birth than for mothers. Focusing now on any differences by child gender, we see that father's earnings and

employment decline a slightly after a boy is born, supporting findings of Bhalotra and Brown (2018), and that father's adjust hours downward conditional on employment after a girl is born. The child earnings penalty—the difference between mothers and fathers as described in equation (4)—averaged over the entire post-birth period studied—is 25.9% after a daughter versus 2.9% after a son. The employment penalty is 19.6% after a daughter versus 6.4% after a son.

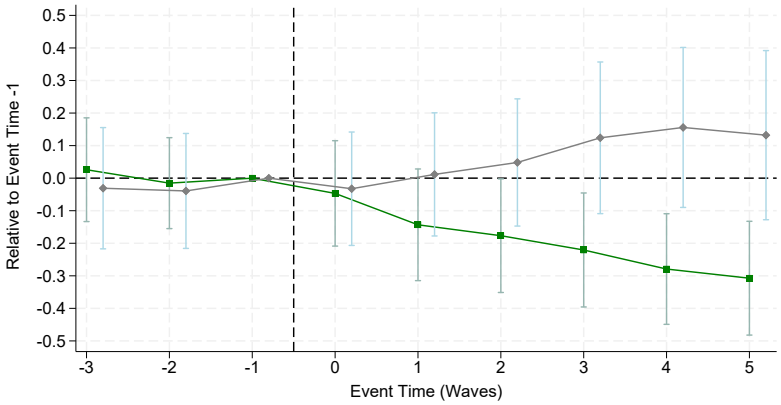
#### 4.2.2 Interaction results by child gender

We combine the sub-sample estimates shown in the preceding section in a single equation by estimating the fully-interacted model in Equation (3). This allows for a formal test of the significance of the difference in labour market responses between parents with sons vs. daughters. These results are presented in Figure 3. Here the green line with square points shows son versus daughter differentials for mothers, while the grey line with diamond-shaped points shows similar son versus daughter differentials for fathers. The size of the 'daughter penalty' is calculated by taking the difference between proportional changes in the respective mothers' and fathers' outcomes. Figure 3 panel (a) shows that a mother's earnings and the likelihood of employment declines significantly more following a daughter than following a son. This gap increases through at least the first four waves after birth for both outcomes. Father's earnings and employment, if at all, tend to increase if they have a daughter rather than a son, but this difference is generally not significant. Estimated effects are large. For example, considering earnings responses, by the fourth and fifth wave after a birth occurs, mothers are in this sample are observed to have around 30% lower wages if their first birth was a daughter than if their first birth was a son, with the average penalty for mothers compared to fathers being equal to 27.8%. Much of this is likely mediated by employment itself (panel (b)), as rates of employment are observed to be around 10-15% lower for mothers of daughters than for mothers of sons. The penalty is estimated at 14.6% across five years.

Figure 3 panel (c) shows that the total hours worked by mothers (conditional on being employed) follow the same pattern, declining significantly more after a daughter is born than when a son is born. As seen in the separate sample results in the preceding section, when considering hours worked conditional on employment, the father's hours move in the same direction as mother's hours—falling more after a daughter than after a son. This similar decline for fathers explains why the labour market penalty documented in Figure 1(f)—following girls—is smaller than that in Figure 1(e)—following boys. In the case of daughters, hours worked *conditional on employment* are observed to decline among both mothers and fathers, unlike the case of earnings and employment, where only mothers are observed to have larger responses to daughters than to sons.

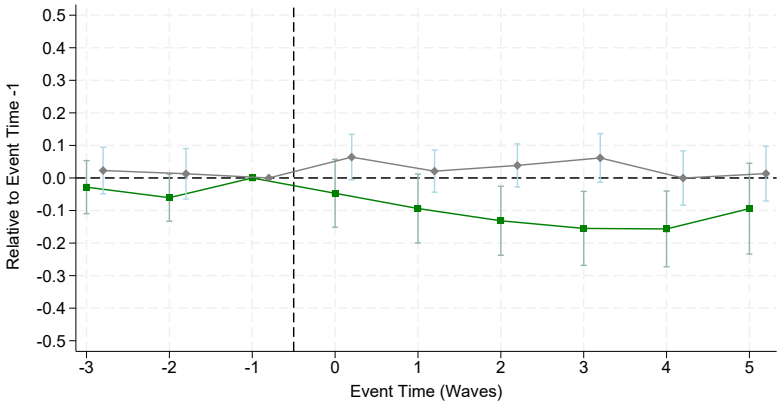
Figure 3: Do parents respond differentially to girls? Mothers versus Fathers Daughter Penalties

(a) Earnings differentials



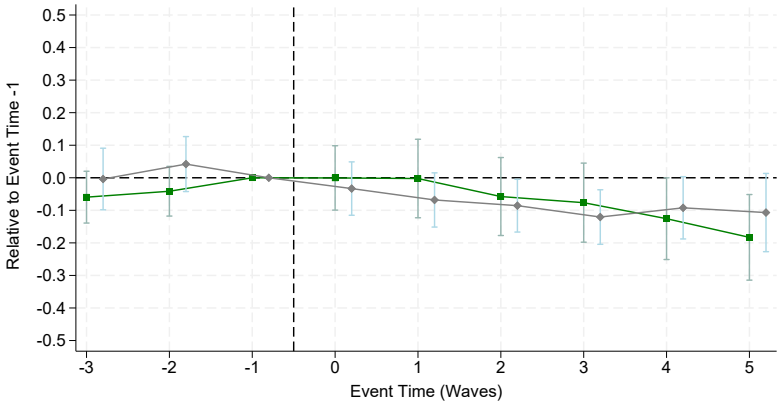

5 year average daughter penalty = 0.278 (0.133)

(b) Employment differentials




5 year average daughter penalty = 0.146 (0.062)

(c) Hours differential




5 year average daughter penalty = -0.053 (0.060)

*Notes:* Estimates of $\widehat{\alpha}_j^{Girl,g}$ are presented following equation (3) with outcomes indicated in caption titles. Estimates are presented separately for mothers (green line) and fathers (grey line). 95% confidence intervals are based on robust standard errors.

### 4.2.3 Alternative Measures of Labour Market Outcomes and Occupational Choice

In Appendix C we examine the sensitivity of observed labour market patterns to alternative measures of income, employment, and hours worked, and test whether child gender drives broader career choice like job transition or occupational change. In Appendix Figure C1 we present results where outcomes consist of a measure of working hours which also includes self-employed individuals (panel (a)), monthly earnings net of taxation and other social service contributions (panel (b)), whether an individual is employed part time (panel (c)), and whether an individual changed job (panel (d)) or changed occupation (panel (e)).

Considering net monthly earnings in place of gross earnings leads to similar conclusions. Namely, a mother's earnings net of tax and national insurance payments—a measure of true purchasing power—declines substantially more following the birth of a daughter than following the birth of a son. Similarly, we observe little evidence suggestive of a corresponding pattern for fathers, if anything observing moderately larger increases following daughters rather than sons. The similarity of these results to models based on gross earnings is unsurprising given the close correspondence between net and gross earnings, however makes clear that any results are similar in magnitude and significance when focusing on take-home pay: likely a more salient measure to households. We also observe that results are similar when focusing on hours worked (conditional on employment) which include self-employed individuals, with a 16.9% larger decline in hours worked for mothers following the birth of a daughter than fathers.

While we did not observe statistically significant gender-related responses in part-time employment immediately after childbirth, there was a notable increase in the probability of part-time employment in the fourth and fifth years after the birth of a daughter for mothers, with no similar movement observed for fathers. The increase in part-time employment is important because it often comes with reduced benefits and career advancement opportunities, reflecting a potential compromise that mothers of daughters might be making between work and childcare responsibilities. This shift toward part-time work coincides with the period when many parents of girls have their second child, suggesting that the cumulative burden of childcare is compelling mothers to adjust their work patterns in a way that could have long-term career implications.

Changes in modes of work, for example from full time to part time, evidence a greater likelihood that mothers seek job flexibility following daughters than sons. This job flexibility may be more pervasive, including potential changes of jobs or occupations to seek areas which offer greater flexibility, even at the cost of worse remunerations (see e.g. Goldin, 2015). Regarding differential shifts in either jobs or entire occupation classes, we observe relatively little evidence to suggest that such shifts are more or less likely following girl children rather than boy children, though note that given the relatively low proportion of job and occupation changes observed in our sample, we do not have power to estimate effects precisely.

## 4.3 Other Outcomes: Home production, attitudes and mental health

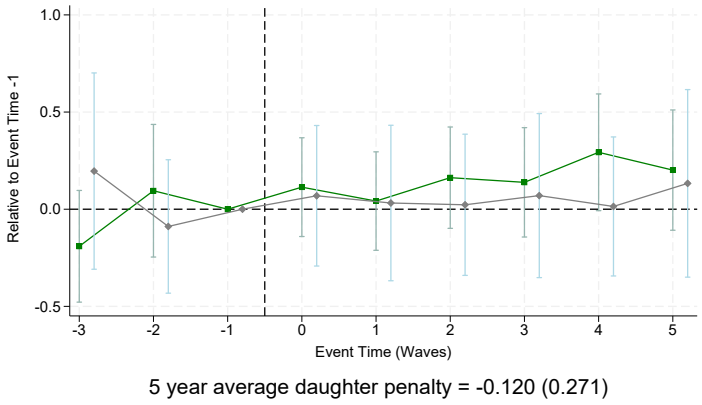
Parenthood brings about profound changes in family life. It affects mental health and well-being, it demands more time spent on domestic chores and in childcare, it can reshape beliefs and attitudes and change preferences about family formation. This section explores the underlying mechanisms that explain why mothers of daughters experience a larger child penalty in the labour market compared to mothers of sons. Given that we are interested in understanding possible drivers of the daughter penalty documented previously, the estimates in this section focus on Equation (3), which captures differences in changes in potential mediators following a daughter compared to following a son. We then document the component effects of this for both sons and daughters separately, as laid out in Equation (1). We first look at whether differences in labour market penalties reflect differences in household production demands. In this domain, we consider time spent on household chores, continuation of fertility, and time spent parenting. To better understand the observed responses on household production, we model cohabitation dynamics as clearly the father's presence matters for the mother's burden. We also model differences in changes in preferences following birth, in particular for markers of adherence to gender norms. Finally, we consider whether differences in labour market penalties reflect mental health constraints.

### 4.3.1 Home production

**Housework** The UKHLS asks respondents two relevant questions. One is the hours they spent per week on housework in an average week over the last year, including time spent cooking, cleaning and doing the laundry. On average, fathers report spending approximately 6 hours per week on these tasks, and mothers report spending around 8 hours per week. In addition, for each of six specified chores—grocery shopping, cooking, cleaning, washing/ironing, gardening, and DIY jobs—respondents are asked whether they primarily did the chore on their own, whether they shared it with their partner, or whether their partner primarily did it. We code this as an indicator that takes the value of 1 if the respondent did the chore alone and zero otherwise, and we study the share of the six chores that respondents report doing alone. On average, fathers report doing 34% of chores on their own and, for mothers, the reported share is 36%. Note that when referring to the partner, this is the partner of the mother or father of the child (we consider both). Most of the time this will be the biological co-parent but the question pertains to the partner living with the parent at the time they were surveyed. In analyses discussed here we consider the outcome measuring whether chores were done by a parent themselves. Results for the outcome measuring the total time spent on chores—which are qualitatively similar—are presented in Appendix Results.

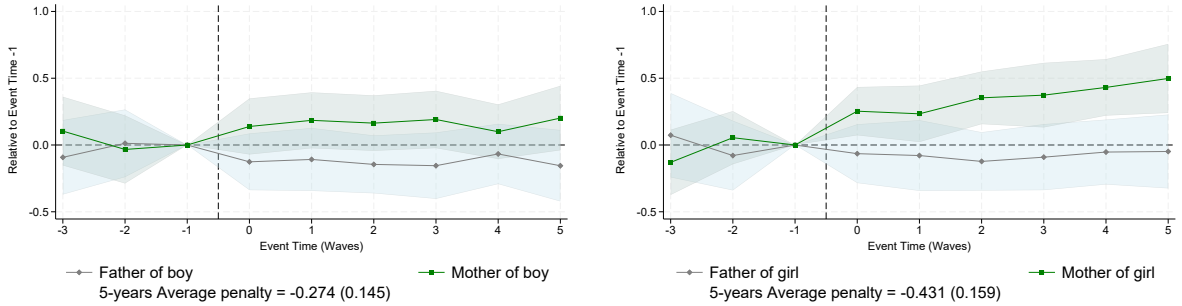
Figure 4: The child penalty in housework by gender of the firstborn child

*Fraction of chores done yourself*




5 year average daughter penalty = -0.120 (0.271)

(a) Differential responses by child gender: Father *(grey)* and Mother *(green)*




(b) Penalty: Boy and Girl

*Notes:* Top panel (a) presents estimates of $\widehat{\alpha}_j^{Girl,g}$ following equation (3) examining differences in responses to girl children for mothers (green) and fathers (grey). Panel (b) presents estimates of $\widehat{\alpha}_j^{g}$ following equation (1) and (2) for parents of boys (left hand panel) and parents of girls (right hand panels). Estimates are presented separately for mothers (green) and fathers (grey). In all cases, estimation is based on a sample of parents who had their first child between 2009 and 2019. Shaded blue and green areas or error bars represent 95% confidence intervals based on robust standard errors.

Figure 4 sheds light on whether the burden of household chores falling upon the mother of the child is different when her first child is a son vs. a daughter. The top panel of the plot presents the daughter penalty in chores, while the bottom panel presents estimates of the impact of child birth on the fraction of chores separately surrounding the birth of boy children and girl children. What is striking in Figure 4 and has not been previously documented as far as we know is that mothers of *girls* carry a larger burden of household chores than mothers of boys. This is observed both when considering the fraction of chores which a mothers does by herself (Figure 4) as well as the total time spent on chores (Appendix Figure C3). Figure 4a plot formal tests of the girl-boy *difference* in housework for each parent. The event study estimates are noisy but we still see a very clear pattern whereby, following the birth of a daughter, the mother spends more time on housework. These effects are large. Averaged over the 5 years following birth,

we estimate that mothers of girls spend around 15% *more* time on chores than mothers of boys. No similar pattern is observed for fathers where following both the birth of sons or the birth of daughters we see, if anything, slight declines in the portion of chores they report completing themselves. We do not see differential responses by child gender among fathers.

Averaged over the five years after birth, the child penalty in the share of chores done alone is 43% following a daughter and 27% following a son. These penalties are even larger when considering total time spent on household chores (Appendix Figure C3) at approximately 79% following a daughter and 59% after a son. A striking pattern is that the increase in the mother's domestic burden after a child (of either sex) is born increases with the age of the child. The child penalty is persistent and it broadly follows this pattern in child age. Our finding that mothers of daughters bear a larger penalty in housework than mothers of sons is consistent with our finding that they suffer a larger labour market penalty.

**Childcare** We define an indicator as 1 when the respondent parent says they are mainly responsible for looking after the children, and 0 when they either say that their partner is mainly responsible or that the responsibility is shared (recall that these questions are put to mothers and fathers). Of all mothers, 67% say they are primarily responsible, compared with 4% of fathers, see Table A2.

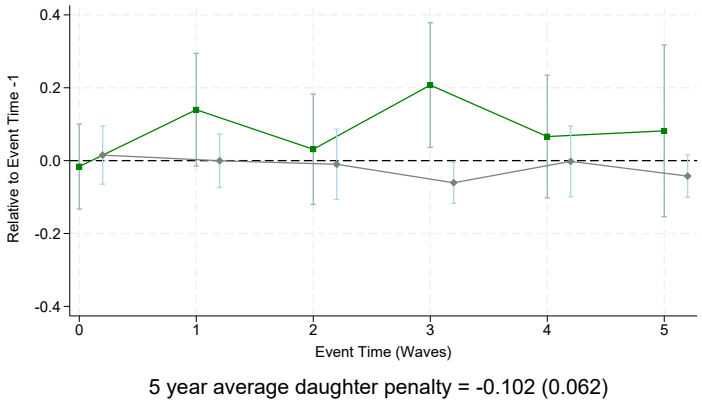
Figure 5a document formal test of equality of responses to child gender by mothers and fathers. We observe that mothers are significantly more likely to be the main caregiver following girls than following boys. Averaged over the 5 waves following birth, mothers are 10.2% more likely to report being the main caregiver following the birth of a girl than fathers. Figure 5b shows how caring responsibilities for children varies with the sex of the first child. We observe that mothers of daughters are more likely to be primarily responsible than mothers of sons. Averaged over the first five years of the child, the penalty is 67% when the child is a girl and 57% when it is a boy, a difference of about 10 percentage points which, on average, is statistically significant.

**Fertility continuation** Son preference in fertility is not unique to India and China. For instance, Dahl et al. (2008) show that shotgun marriages in the United States are more likely when the mother is carrying a son, indicating son preference among men. There is evidence from Denmark that fathers (not mothers) are a lot happier after a son is born (Kohler and Mencarini, 2016). If parents (fathers) desire sons then they will be more likely to continue fertility after a firstborn daughter than after a firstborn son. This could contribute to explaining our finding that firstborn daughters are associated with larger declines in the employment and earnings of mothers, and to mothers carrying a larger burden of household chores and childcare.

We define an indicator for having at least one more child after the first child in the sample

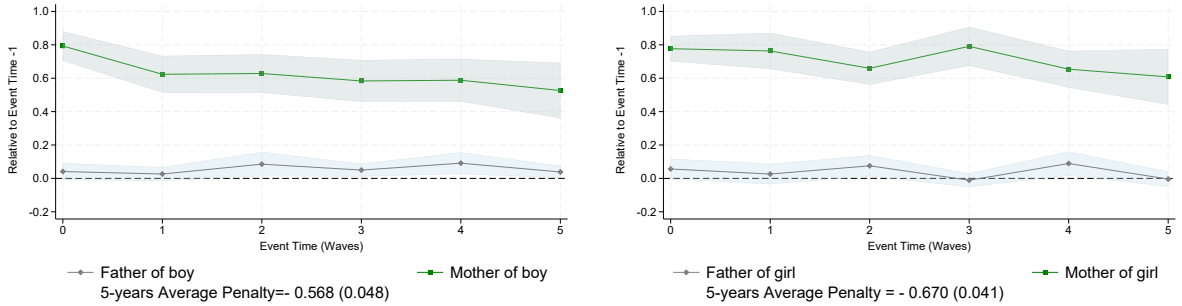
Figure 5: The child penalty in childcare responsibilities by gender of the firstborn child

*In charge of childcare yourself*




5 year average daughter penalty = -0.102 (0.062)

(a) Differential responses by child gender: Father *(grey)* and Mother *(green)*



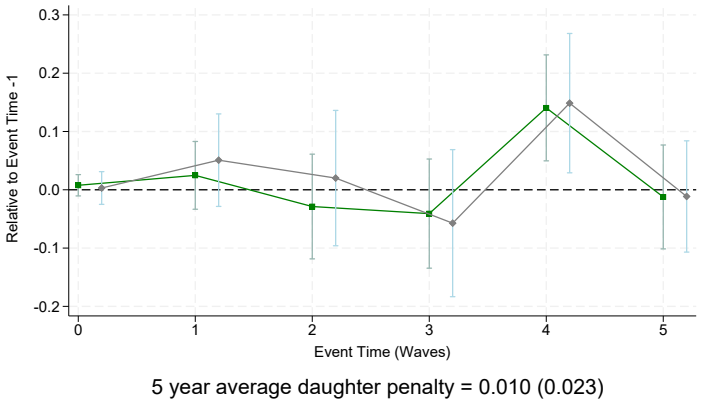

(b) Boy and Girl

*Notes:* Refer to Notes to Figure 4. Identical results are presented, however here considering as dependent variable an indicator of an individual reporting being the main carer for their child.

of parents with at least one child and refer to the indicator as future fertility. Once we take the son-daughter difference for each parent in panels (a), we see a clearer pattern—for both fathers and mothers, the probability of a subsequent child after the birth of their first is statistically significantly larger (in the fourth year from the first birth) if their first was a girl. The effect sizes are large, with around a 10pp increase in the likelihood of having another child at this point following a girl. Figures 6(b) and (c) suggest that results are quite similar for both mothers and fathers. In both cases, we observe future births principally in waves 2 and 3 after their indicator birth. The fact that both mothers and fathers have a greater likelihood of an additional birth following girls provides a potential explanation for labour market responses, particularly when combined with previous results suggesting that a greater burden of childbearing and housework responsibilities fall on mothers rather than fathers.

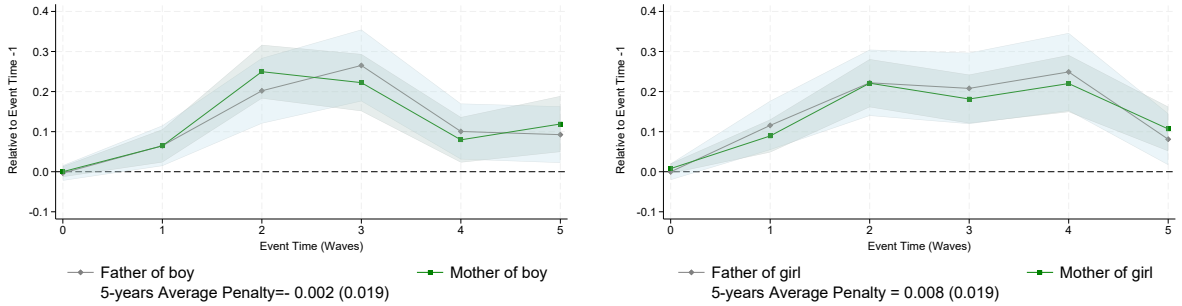
Figure 6: The child penalty in future fertility by gender of the firstborn child

*Having an additional child*




5 year average daughter penalty = 0.010 (0.023)

(a) Differential responses by child gender: Father *(grey)* and Mother *(green)*




(b) Penalty: Boy and Girl

*Notes:* Refer to Notes to Figure 4. Identical results are presented, however here considering as dependent variable an indicator of whether a future birth occurs following the first birth.

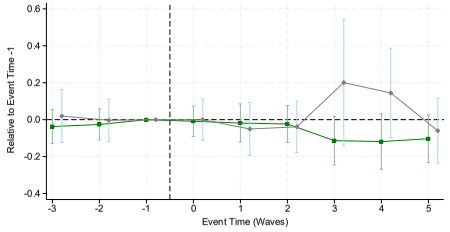
### 4.3.2 Progressive behaviours and attitudes

We now investigate how gender and political attitudes of mothers and fathers evolve after the birth of a daughter vs. a son, see Figure 7.

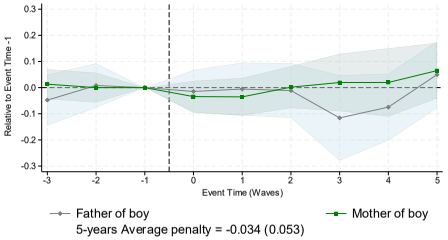
For both parents, self-reported views on gender equality (Figure 7, Panel A) start to diverge from their pre-birth levels two years after birth. These estimates are not significantly different between fathers and mothers, but it is striking that fathers and mothers responses move in opposite directions. When the child is a daughter rather than a son, fathers report more gender-equal attitudes and mothers report less gender-equal attitudes, and a gap of 0.082 percent (the daughter penalty) opens up.

Figure 7: The child penalty in parental views on progressive gender and socio-political attitudes by gender of the firstborn child

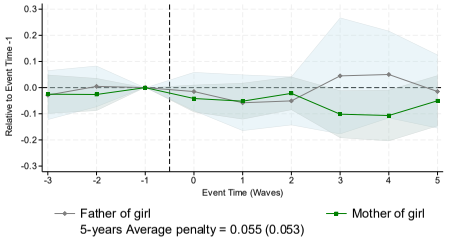
**Panel A: Gender Equality Index**




Differential responses by child gender

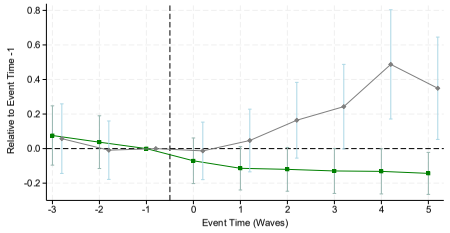



Penalty for boy

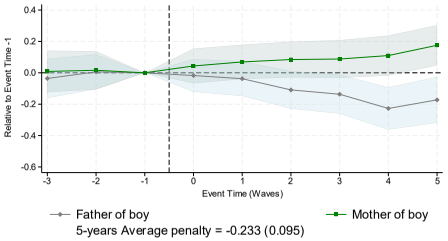



Penalty for girl

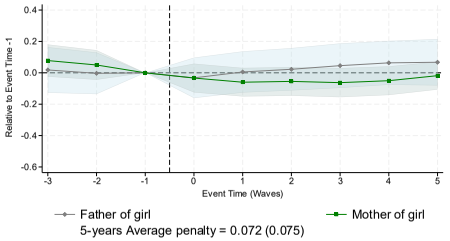
**Panel B: Progressive newspaper readership**




Differential responses by child gender

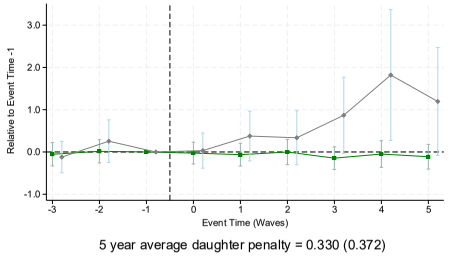



Penalty for boy

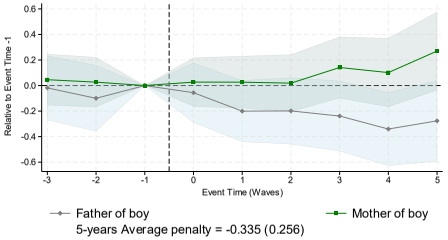



Penalty for girl

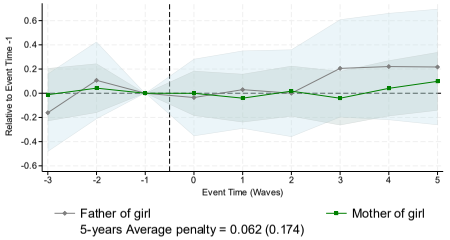
**Panel C: Progressive voting**




Differential responses by child gender




Penalty for boy




Penalty for girl

*Notes:* Refer to Notes to Figure 4. Identical results are presented, however here considering as dependent variable parental gender equality (top row), whether an individual reads progressive news (middle row), and whether an individual votes for parties classified as left wing (bottom row).

We look at political attitudes with two distinct measures, leveraging the fact that the UKHLS gathers data on the newspapers (including online sources) that people read, and on the political parties they vote for. Both measures show that fathers become more progressive after a daughter rather than a son (Figure 7, Panels B and C). This is driven by fathers becoming less progressive after a son is born relative to their pre-birth levels. What is striking, similar to the case for gender attitudes, is that mothers react in the opposite direction, becoming more progressive after a son rather than a daughter is born, relative to their pre-birth attitudes. Unlike the case of the gender equality index, in the case of these behavioural measures, differences are statistically significant for news consumption (overall and at various lags after birth) and voting behaviour (from wave 3 post-birth onwards).[18]

Overall, across three different indicators of progressive attitudes, the consistent tendency for mothers to become less progressive when a daughter is born (relative to when a son is born) is in line with our findings that women are less likely to be employed, earn less, carry a larger burden of household chores and are more likely to have another child. These results suggest that preferences play a role.

#### 4.3.3 Relationship quality and mental health

In this section we consider relationship quality and mental health, as summary markers of how parents feel after a daughter vs. a son is born.

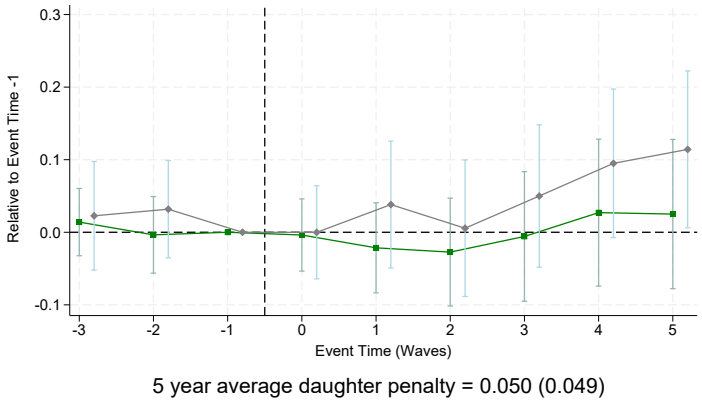
**Relationship satisfaction** Parenthood introduces profound shifts in relationship dynamics, altering happiness, stress levels, and long-term satisfaction between partners. Building on the evidence that financial constraints and domestic responsibilities affect the quality of relationships post-birth (Bertrand et al., 2013; Lundberg et al., 2016), we evaluate whether child gender plays a role in these dynamics. In particular, we focus on two relationship-specific questions: (i) whether the parent is considering divorce/separation or (ii) whether the parent regrets getting married/living together. Figures 8 and 9 suggest that fathers of daughters are more satisfied with relationships than fathers of sons, while the mother's perception of relationship stability does not seem to depend on the child's gender. When we look at the point estimates comparing mothers and fathers of boys and mothers and fathers of girls, we can see that fathers of girls tend to never regret marriage/living together or never consider divorce/separation if they have a daughter. This pattern broadly lines up with results documented above about the division of housework within the family. Mothers of girls face an increased burden of household chores and childcare with a corresponding lower demand on fathers, potentially making the early years

[18] There is already some evidence that fathers become more likely to vote progressively after a daughter is born, see Powdthavee and Oswald (2010); Washington (2008). However, Costa et al. (2019); Green et al. (2023) argue that this does not hold in more recent years. This may be because perceptions of what the Labour vs. the Conservative party offer in the UK have changed over time, with many voters crossing over. In our data, we see the tendency that the earlier studies found.

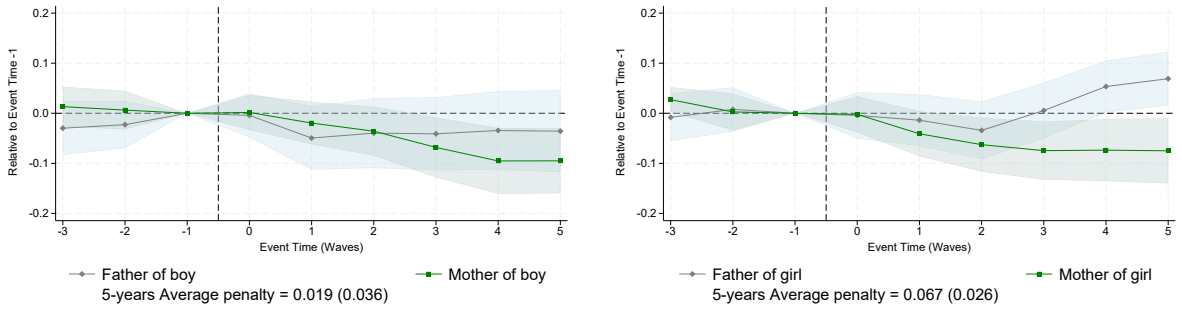
of parenthood less stressful for fathers. It is unclear how this pattern evolves later in life, since broader evidence has suggested that couples do indeed divorce at higher rates following the birth of daughters compared with sons (see e.g. Dahl et al., 2008). However Kabátek and Ribar (2020) have documented with both Dutch and US microdata that such an effect generally occurs during a child's teenage years rather than early in life.

Figure 8: The relationship stability difference by whether the firstborn child is a daughter

*Never considering divorce/separation*




(a) Differential responses by child gender: Father *(grey)* and Mother *(green)*




(b) Penalty: Boy and Girl

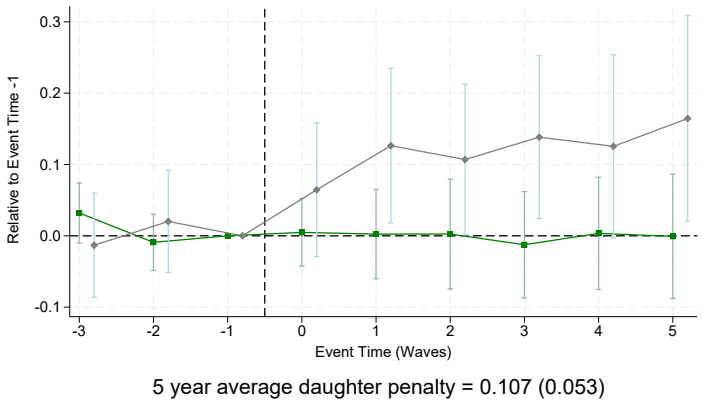
*Notes:* Refer to Notes to Figure 4. Identical results are presented, however, here considering as dependent variable an indicator of whether a parent is considering a divorce or separation.

We observe broadly similar patterns when considering the relationship happiness index, which is presented in an Appendix. Figure C4 shows that while both parents have a decreasing trend in the relationship's happiness following child birth, fathers of daughters are happier with the relationship than fathers of sons. Mothers do not react to the child's gender up to year 4, where mothers of girls seem to be happier with the relationship, which coincides with the probability to have a second child at that year if the first one is a girl. These measures, and indeed factors discussed above, cluster together. For example, relationship quality and considerations of divorce are highly correlated (Figure A1), but more so, parents who report a greater child care

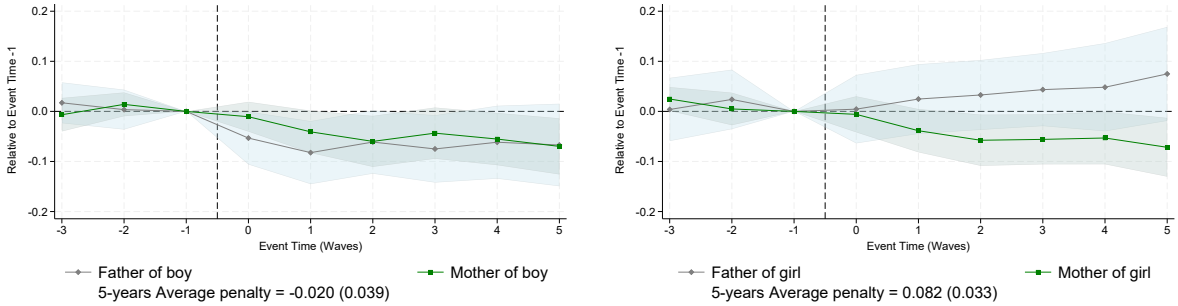
burden, and a greater chore burden also report lower satisfaction with their relationship, worse mental health, and, perhaps surprisingly, a greater likelihood to have additional children. In line with the "daughter penalties" documented here, in Figure A1 we also observe descriptively that mothers of daughters have lower salaries, report doing more chores alone, and lower relationship happiness, and that these patterns are muted among fathers. This highlights the intertwined nature of relationship quality, as well as home and labour market production.

Figure 9: The relationship stability difference by whether the firstborn child is a daughter

*Never regretting getting married/living together*




(a) Differential responses by child gender: Father *(grey)* and Mother *(green)*




(b) Penalty: Boy and Girl

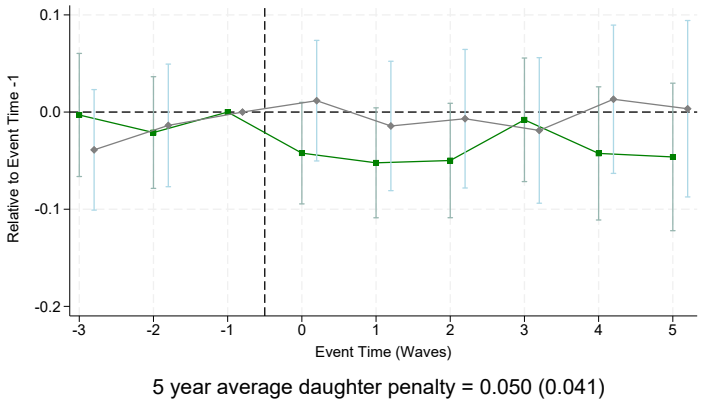
*Notes:* Refer to Notes to Figure 4. Identical results are presented, however, here considering as dependent variable an indicator of whether a parent regrets getting married/living together.

**Mental health** While most families anticipate the birth of a child as a joyous occasion, they often experience new emotional and psychological stress after the birth. These stresses are exacerbated among women—it is estimated that between 10 and 20 percent of mothers worldwide suffer perinatal depression (Gelaye et al., 2016; Baranov et al., 2020). A recent study that examines fathers alongside mothers confirms that the mental health of mothers deteriorates more than that of fathers after birth (Ahammer et al., 2023), but they proxy mental health deterioration

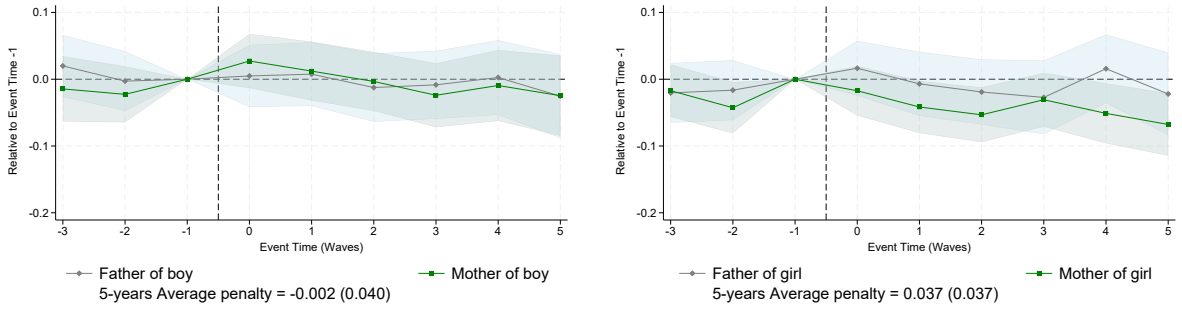
with an indicator for mental health treatment (whether or not the parent is on an antidepressant prescription). Bhalotra et al. (2025) document a similar pattern in survey data from two distinct sources. It seems plausible that this is correlated with the earnings penalty following childbirth.

Figure 10: The child penalty in mental health by gender of the firstborn child

*Parental mental health (SF-12 index)*




(a) Differential responses by child gender: Father *(grey)* and Mother *(green)*




(b) Penalty: Boy and Girl

*Notes:* Refer to Notes to Figure 4. Identical results are presented, however here considering as dependent variable parental mental health.

We provide the first investigation of the extent to which the mental health of parents varies with the sex of their firstborn child. We provide estimates for the SF-12 scale with higher scores indicating better mental health (details of this and all other variable definitions are in Appendix A.2). Our main finding is that mothers of daughters suffer a larger decline in mental health than mothers of sons. This is observed in Figure 10, and is particularly clear in the first few years postpartum, this difference is significant at the 90% level. Averaging over the initial five years postpartum, the SF-12 scale shows little divergence among parents of boys, while showing a 3.7% larger drop among mothers (relative to fathers) of girls, though confidence intervals largely overlap. Figure 10a documents the daughter penalty to be around 5% for mothers compared to fathers, at least when considering mean point estimates.

Our findings are corroborated—though similarly are often noisily estimated—by other variables in the UKHLS. We find that mothers of girls are less likely to report high overall life satisfaction and more likely to report feeling depressed and downhearted in the previous 4 weeks than fathers of boys, while the reverse is true for parents of boys. We also find that mothers of girls are more likely to report feeling worthless during the first year of parenthood than mothers of boys. However, we do not find any significant differences in the parents' self-reported ability to face problems, in their physical health, or in the number of cigarettes they smoke. See Figure C2 (life satisfaction and depression, worthlessness, physical health and health behaviours) in Appendix C.

This finding is consistent with new mothers who become homemakers in the traditional sense of the word (reducing their time on the labour market, increasing their time in home production and adopting more conservative attitudes) being happier, even if the fathers of their children appear more committed to their relationship. However, it is relevant to emphasise that we cannot identify cause and effect, only a cluster of reduced form outcomes that, to some extent hang together, and that effects of childbirth on mental health are estimated imprecisely.

## 4.4 Heterogeneity in the daughter penalty by pre-birth characteristics

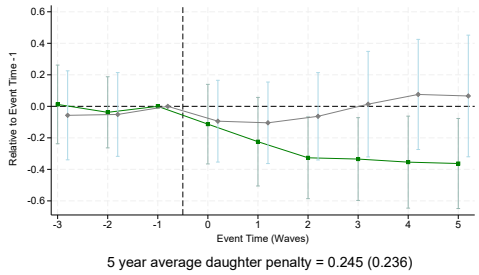
We first report results broken down by whether or not the individual (mother or father) has a university degree. We then discuss results broken down by whether or not they agree with the statement that the family suffers if the mother works full-time (henceforth, we will label this gender-progressive), using answers to this question in the wave before the focal birth. Similar to the main analysis discussed so far, these results refer to all men and all women—in other words fathers are not matched to mothers—instead, the results for fathers show how their outcomes evolve after they have a child and they are estimated independently of the results for mothers. Throughout this section we focus on earnings responses to avoid the overload created by consistently presenting earnings, employment and hours conditional on employment, but the results for employment and hours are in Appendix D Figures D1-D2.

### 4.4.1 Daughter penalty by parent's education

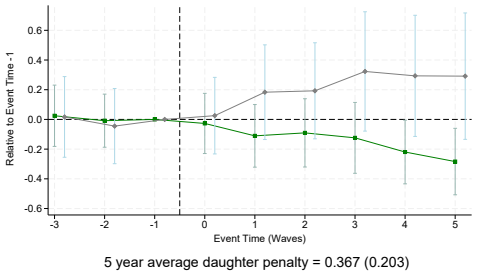
The child penalty for sons and daughters and the difference of the two (the daughter penalty) are reported in Figure 11 for two subsamples of mothers and fathers, those with a degree, and those without.

We see a daughter penalty in both groups. Observing this in distinct samples reinforces confidence in the existence of the daughter penalty. This result also confirms that the slight imbalance in parent education that we noted between parents of sons and daughters is not spuriously driving the pooled results reported earlier. What we see is interesting—although parents

Figure 11: Parents with and without degrees: child penalties in earnings

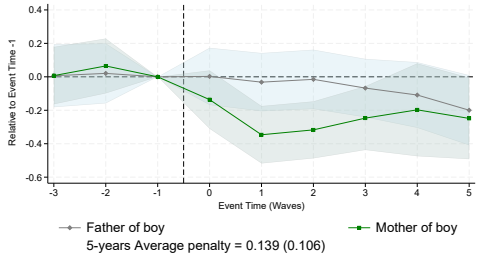



(a) Earnings (Parents without a degree)

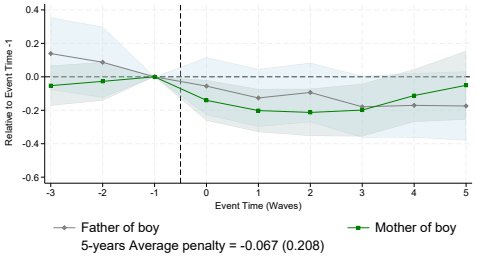



(b) Earnings (Parents with a degree)

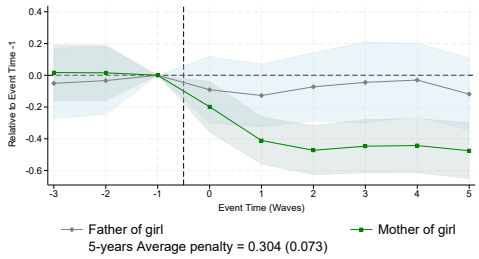
Differential responses by child gender: Father *(grey)* and Mother *(green)*



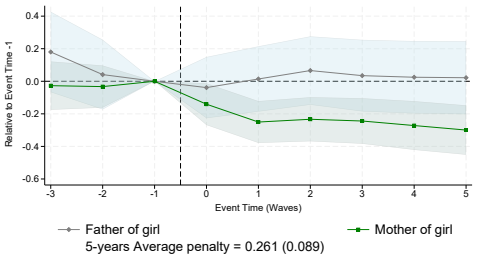

(c) Earnings and boy children (Parents without a degree)




(d) Earnings and boy children (Parents with a degree)




(e) Earnings and girl children (Parents without a degree)




(f) Earnings and girl children (Parents with a degree)

The underlying penalty in earnings by child gender and whether the parent has a degree

*Notes:* Refer to Notes to Figures 2 and 3. Identical results are presented for earnings, however here stratifying by parental education (individuals without a degree on left, and with a degree on right).

with a degree experience a *smaller child penalty* (i.e. difference between father and mother earnings after birth), they experience a *larger daughter penalty* (i.e. the child penalty differenced between parents of sons vs. daughters).[19] Unpacking the double difference underlying

[19]Figure 11 panels (a) and (b) show that the daughter earnings penalty is 36.7% for parents with a degree versus 24.5% for parents without (and that the response of fathers with a degree to a daughter vs. a son is a key driver of this). Figure 11 panels (c)-(f) break this down to show the child penalty separately for sons and daughters. For mothers without a degree, the child earnings penalty averages 30.4% after a daughter versus 13.9% after a son. For mothers with a degree, the penalty is 26.7% after a daughter but reverses to a (insignificant) 6.7% gain after a son. Employment penalties documented in Figure D1 follow a similar pattern: mothers without a degree experience a 27.6% penalty after a daughter compared to 16.7% after a son, while degree-holding mothers face penalties of 11.6% and 0.4%, respectively. For hours worked conditional on employment, there is little difference for mothers

the daughter penalty reveals that, among parents with a degree that have a son, the child penalty is small and not significantly different from zero, see Figure 11(d). Comparing with degree-parents that have a daughter (Figure 11(f)) reveals that having a son rather than a daughter leads to fathers earning less and to mothers earning more.[20]

As with the main results, we look at other correlated decisions and outcomes for young families. A striking result is that the daughter penalty in household chores and childcare is also driven by mothers with a degree taking on more of these duties after a daughter rather than a son is born, see Figures D3 and D4. The daughter penalty in housework and childcare respectively in the degree-parent group is 43.3% and 15.4%, compared to 1% and 4.9% among less educated parents.[21] This is also true for future fertility—it is driven by the degree parents, see Figure D5—in fact the pattern is reversed in less educated families. These patterns are consistent with the earnings results, pointing to an increase in specialisation after a daughter vs. a son, driven by parents with a degree education. The daughter penalty on mental health is driven by non-degree parents, see Figure D6. This, is consistent with the relatively educated women that move toward household specialisation feeling comfortable with this decision.

### 4.4.2 The daughter penalty by gender norms

This section examines daughter penalties based on a particular measure of gender norms equality, which is whether the parent agrees with the statement that the family suffers if the mother works full time. Figure 12 shows that the daughter penalty emerges from parents who are gender progressive (similar patterns are observed for employment, rather than earnings in Figure D7). Examining the underlying child penalties for sons and daughters separately, it is clear that this arises because there is essentially no child penalty when the parents have progressive gender norms and the child is a boy (and there is a child penalty when they have a girl). In contrast, among parents with regressive gender norms, there is a (similar sized) child penalty for sons and daughters, and thus only a small daughter penalty.[22]

---

without a degree. However, mothers with a degree show a smaller penalty for daughters (16.1%) than for sons (28%).

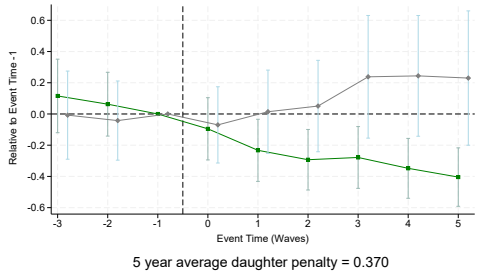
[20] It is hard to isolate why this is the case. There is outside evidence that parents with a degree tend to invest more in the human capital of their children (Chevalier et al., 2005). In our UK sample we see that parents with a degree possibly have higher joint earnings when they have a son (which they might use to send the son to childcare—we do not observe this), and more maternal time with the child when they have a daughter. As income and maternal time are both investments in children, it is hard to infer preferences or make welfare-style statements.

[21] Degree-educated mothers of girls perform 72.8% more housework than fathers, whereas those with boys perform 18.3% more. Among educated mothers, those with daughters are 70.3% more likely to bear sole childcare responsibility, compared to 54.9% for those with sons. Among less educated mothers, this is reversed- we see son penalties, though the mother-father gaps are smaller. Less educated mothers with boys face a larger housework penalty than those with girls, performing 39.1% more housework than fathers, compared to 33.2% for mothers of girls. Less educated mothers of girls also experience a higher childcare penalty (67.9%) than those with boys (62.9%).

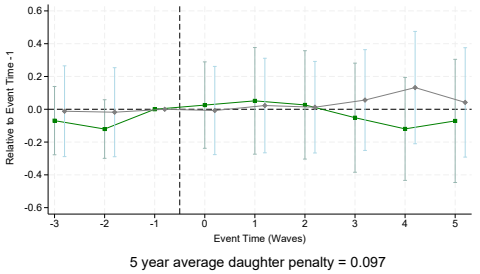
[22] Mothers of daughters who disagree with the statement that the family suffers if the mother works full time had larger child penalties in earnings, at 37% compared to only 10% among those who reported to be in agreement with the statement that families suffer if mothers work full time.

The larger daughter penalty among progressive parents lines up nicely with the larger daughter penalty among degree-educated parents, and both reflect a small, if any, child penalty following the birth of a son. This is a striking pattern—and it is notable that the share of parents reporting the progressive gender norm is only slightly higher in the degree-educated group—so these are distinct albeit overlapping predictors of a daughter penalty.

Figure 12: Daughter Penalty by Gender Norms

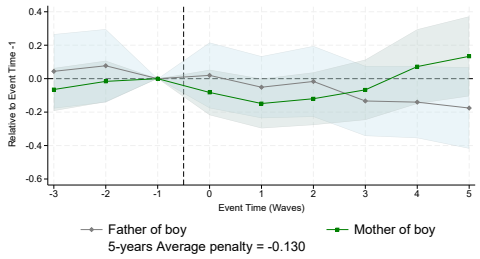



(a) Earnings: Parent disagrees with the statement: "family suffers if mother works full-time" (progressive norms)

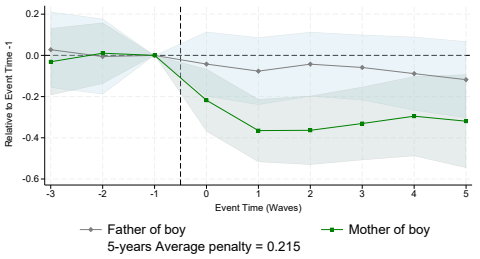



(b) Earnings: Parent agrees with the statement: "family suffers if mother works full-time" (regressive norms)

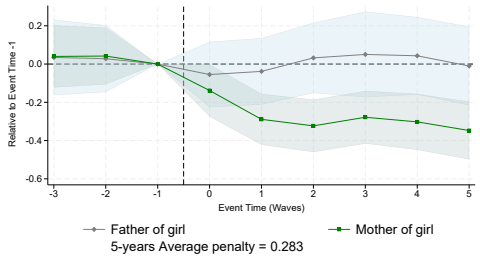
Differential responses by child gender: Father *(grey)* and Mother *(green)*



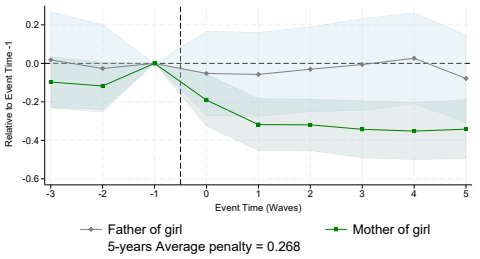

(c) Parent disagrees with the statement: "family suffers if mother works full-time"




(d) Parent agrees with the statement: "family suffers if mother works full-time"




(e) Parent disagrees with the statement: "family suffers if mother works full-time"




(f) Parent agrees with the statement: "family suffers if mother works full-time"

The underlying penalty in earnings by child gender and gender sentiments

*Notes:* Refer to Notes to Figures 2 and 3. Identical results are presented for earnings, however here stratifying by parents views on gender norms (individuals disagreeing with the statement that the "family suffers if mothers work full-time" on the left, while those agreeing with the statement on the right).

Figures D9-D12 show associated outcomes separately by beliefs about gender norms. While noisily estimated, consistent with the earnings result, couples with progressive gender norms

show a larger daughter penalty in housework (38% vs 9.4%), although a smaller daughter penalty in childcare (8.7% vs 14.9%). Fertility responses are observed to be larger among respondents with less gender progressive views (Figure D11). Indeed, among gender progressives respondents, no substantial fertility effects are observed, and the fact that a daughter penalty is observed in this group suggests that the daughter penalty cannot be *only* explained by differential fertility responses to sons and daughters. The daughter penalty on mental health is larger in progressive couples (5.6% vs 1.6%), though again this is not precisely estimated. If, however, this is viewed as a measure of the woman's stress or satisfaction with specialisation then this presents a contrast to the results for degree.

# 5 Conclusion

This study identifies a daughter penalty which is essentially a larger child penalty on earnings when the child is a daughter rather than a son. Using household survey data, it 'gets inside the household', showing that the daughter penalty in earnings is mirrored (to some extent) in a daughter penalty in housework and childcare, and that it is associated with future fertility and, suggestively, with a tax on the mental health of mothers relative to fathers. Our findings suggest that the birth of a daughter reinforces traditional gender roles within households, leading to a more unequal division of labour. Although the child penalty is smaller among parents with a college degree, the daughter penalty is larger. It is also larger among parents who have more progressive gender norms.

Our analysis leverages the random assignment of child gender, highlighting how this can have enduring economic and social consequences. In particular, daughters are disproportionately more likely to be raised in families where the division of duties is traditional. To the extent that parental role models are sex-specific, this could act as a mechanism for the perpetuation of gender norms.[23] It also follows that boys growing up with first born sisters may exhibit stronger gender norms than boys growing up with first born brothers, a pattern that we are currently exploring in a different data set. By uncovering the differential impacts of child gender, this study contributes to a deeper understanding of the behavioural and normative underpinnings of the child penalty.

[23] There is evidence that role model effects are sex-specific, although the evidence tends to emphasise the influence of successful women—including inventors, professors, politicians—on younger women, see, among others, Carrell et al. (2010); Bell et al. (2019); Breda et al. (2023).

# References


Adda, J., C. Dustmann, and K. Stevens (2017): "The Career Costs of Children," *Journal of Political Economy*, 125, 293–337.

Ahammer, A., U. Glogowsky, M. Halla, and T. Hener (2023): "The Parenthood Penalty in Mental Health: Evidence from Austria and Denmark," IZA Discussion Papers 16459, Institute of Labor Economics (IZA).

Almond, D. and L. Edlund (2004): "Son-biased sex ratios in the 2000 United States Census," *Proceedings of the National Academy of Sciences*, 101, 11763–11767.

Andresen, M. E. and E. Nix (2022a): "Can the child penalty be reduced?. Evaluating multiple policy interventions," Discussion Papers 983, Statistics Norway, Research Department.

——— (2022b): "What Causes the Child Penalty? Evidence from Adopting and Same-Sex Couples," *Journal of Labor Economics*, 40, 971–1004.

Baranov, V., S. Bhalotra, P. Biroli, and J. Maselko (2020): "Maternal Depression, Women's Empowerment, and Parental Investment: Evidence from a Randomized Controlled Trial," *American Economic Review*, 110, 824–59.

Bedard, K. and O. Deschênes (2005): "Sex Preferences, Marital Dissolution, and the Economic Status of Women," *The Journal of Human Resources*, 40, 411–434.

Bell, A., R. Chetty, X. Jaravel, N. Petkova, and J. Van Reenen (2019): "Who Becomes an Inventor in America? The Importance of Exposure to Innovation," *The Quarterly Journal of Economics*, 134, 647–713.

Bensnes, S., I. Huitfeldt, and E. Leuven (2023): "Reconciling Estimates of the Long-Term Earnings Effect of Fertility," IZA Discussion Papers 16174, Institute of Labor Economics (IZA).

Berniell, I., L. Berniell, D. De la Mata, M. Edo, Y. Fawaz, M. P. Machado, and M. Marchionni (2021): "Motherhood and the Allocation of Talent," IZA Discussion Papers 14491, Institute of Labor Economics (IZA).

Berniell, I., L. Berniell, D. de la Mata, M. Edo, M. Machado, and M. Marchionni (2024): "Motherhood and the Allocation of Talent," *Applied Economics*, 1–17.

Bertrand, J., and Pan (2013): "The Trouble with Boys: Social Influences and the Gender Gap in Disruptive Behavior," in *M. Rutter, A. Caspi, D. Fergusson, L. Horwood, R. Goodman, B. Maughan, T. Mosot*.

Bhalotra, S. and D. Brown (2018): "The Gender of the First Child and Parental Labour Supply in the UK," PhD Chapter.

Bhalotra, S., M. Daysal, J. Hirani, P. Majumdar, M. Trandafir, M. Wüst, and T. Zohar (2025): "Perinatal Depression Increases the Motherhood Penalty," Mimeo. Universities of Copenhagen and Warwick.

Bhalotra, S. and A. van Soest (2008): "Birth-spacing, fertility and neonatal mortality in India: Dynamics, frailty, and fecundity," *Journal of Econometrics*, 143, 274–290.

Bhalotra, S. R. and T. Cochrane (2010): "Where Have All the Young Girls Gone? Identification of Sex Selection in India," Tech. Rep. 5381, IZA Institute of Labor Economics.

Borusyak, K., X. Jaravel, and J. Spiess (2024): "Revisiting Event-Study Designs: Robust and Efficient Estimation," *The Review of Economic Studies*, rdae007.

Breda, T., J. Grenet, and M. Monnet (2023): "How Effective Are Female Role Models in Steering Girls Towards STEM? Evidence from French High Schools," *The Economic Journal*, 133, 1773–1811.

Callaway, B. and P. H. Sant'Anna (2021): "Difference-in-Differences with multiple time periods," *Journal of Econometrics*, 225, 200–230, themed Issue: Treatment Effect 1.

Carrell, S. E., M. E. Page, and J. E. West (2010): "Sex and Science: How Professor Gender Perpetuates the Gender Gap," *The Quarterly Journal of Economics*, 125, 1101–1144.

Chevalier, A., C. Harmon, V. O'Sullivan, and I. Walker (2005): "The Impact of Parental Income and Education on the Schooling of their Children," Studies in Economics 0504, School of Economics, University of Kent.

Chow, G. C. (1960): "Tests of Equality Between Sets of Coefficients in Two Linear Regressions," *Econometrica*, 28, 591–605.

Cortés, P. and J. Pan (2023): "Children and the Remaining Gender Gaps in the Labor Market," *Journal of Economic Literature*, 61, 1359–1409.

Costa, M., J. S. Greenlee, T. Nteta, J. H. Rhodes, and E. A. Sharrow (2019): "Family Ties? The Limits of Fathering Daughters on Congressional Behavior," *American Politics Research*, 47, 471–493.

Dahl, E., and Moretti (2008): "The Demand for Sons," *The Review of Economic Studies*, 75, 1085–1120.

Das Gupta, M., J. Zhenghua, L. Bohua, X. Zhenming, W. Chung, and B. Hwa-Ok (2003): "Why is Son Preference so Persistent in East and South Asia? A Cross Country Study of China, India, and the Republic of Korea," *The Journal of Development Studies*, 40, 153–187.

Dustmann, C. and R. Landersø (2018): "Child's Gender, Young Fathers' Crime, and Spillover Effects in Criminal Behavior," .

Fernández, R., A. Fogli, and C. Olivetti (2004): "Mothers and Sons: Preference Formation and Female Labor Force Dynamics*," *The Quarterly Journal of Economics*, 119, 1249–1299.

Fiva, J. H. and M.-E. M. King (2024): "Child Penalties in Politics," *The Economic Journal*, 134, 648–670.

Gelaye, B., M. B. Rondon, R. Araya, and M. A. Williams (2016): "Epidemiology of maternal depression, risk factors, and child outcomes in low-income and middle-income countries," *The Lancet Psychiatry*, 3, 973–982.

Goldin, C. (2015): "Hours Flexibility and the Gender Gap in Pay," *Center for American Progress*.

Green, D. P., O. Hyman-Metzger, G. Sood, and M. A. Zee (2023): "Revisiting a Natural Experiment: Do Legislators with Daughters Vote More Liberally on Women's Issues?" *Journal of Political Economy Microeconomics*, 1, 506–516.

Hancock, K., J. Lafortune, and C. Low (2024): "Winning the Bread and Baking it Too: Gendered Frictions in the Allocation of Home Production," Preliminary Draft.

Hotz, V. J., P. Johansson, and A. Karimi (2018): "Parenthood, Family Friendly Workplaces, and the Gender Gaps in Early Work Careers," *Journal of Labor Economics*, 36, 1123–1163.

Jayachandran, S. (2017): "Fertility Decline and Missing Women," *American Economic Journal: Applied Economics*, 9, 118–139.

Kabátek, J. and D. C. Ribar (2020): "Daughters and Divorce," *The Economic Journal*, 131, 2144–2170.

Kaushal, N. and F. M. Muchomba (2018): "Missing time with parents: son preference among Asians in the USA," *Journal of Population Economics*, 31, 397–427.

Kleven, H. (2022): "The Geography of Child Penalties and Gender Norms: A Pseudo-Event Study Approach," Working Paper 30176, National Bureau of Economic Research.

Kleven, H., C. Landais, and G. Leite-Mariante (2023): "The Child Penalty Atlas," Working Paper 31649, National Bureau of Economic Research.

Kleven, H., C. Landais, J. Posch, A. Steinhauer, and J. Zweimüller (2019a): "Child Penalties across Countries: Evidence and Explanations," *AEA Papers and Proceedings*, 109, 122–126.

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

Kohler, H.-P. and L. Mencarini (2016): "The Parenthood Happiness Puzzle: An Introduction to Special Issue," *European Journal of Population*, 32, 327–338.

Korenman, S. and D. Neumark (1992): "Marriage, Motherhood, and Wages," *The Journal of Human Resources*, 27, 233–255.

Leveson, L. (2012): "An Inquiry into the Culture, Practices and Ethics of the Press," National Archives.

Li, J. (2021): "Racial Difference in Child Penalty," The Warwick Economics Research Paper Series (TWERPS) 1382, University of Warwick, Department of Economics.

Lundberg, S., R. Pollak, and J. Stearns (2016): "Family Inequality: Diverging Patterns in Marriage, Cohabitation, and Childbearing," Tech. Rep. 22078.

Lundberg, S. and E. Rose (2002): "The Effects of Sons and Daughters on Men's Labor Supply and Wages," *The Review of Economics and Statistics*, 84, 251–268.

Lundborg, P., E. Plug, and A. W. Rasmussen (2017): "Can Women Have Children and a Career? IV Evidence from IVF Treatments," *American Economic Review*, 107, 1611–37.

Melentyeva, V. and L. Riedel (2023): "Child Penalty Estimation and Mothers' Age at First Birth," ECONtribute Discussion Papers Series 266, University of Bonn and University of Cologne, Germany.

Office for National Statistics (2021): "Families and the Labour Market, England: 2021," Tech. rep., Office for National Statistics, accessed: 2024-11-25.

Pollard, M. S. and P. Morgan (2002): "Emerging Parental Gender Indifference? Sex Composition of Children and the Third Birth," *American Sociological Review*, 67, 600–613.

Powdthavee, N. and A. Oswald (2010): "Daughters and Left-Wing Voting," *The Review of Economics and Statistics*, 92, 213–227.

Ronchi, M. and N. Smith (2024): "Daddy's Girl: Daughters, Managerial Decisions, and Gender Inequality," Working Paper. Revise and Resubmit, *Review of Economic Studies*.

Sun, L. and S. Abraham (2021): "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects," *Journal of Econometrics*, 225, 175–199, themed Issue: Treatment Effect 1.

Trivers, R. L. and D. E. Willard (1973): "Natural selection of parental ability to vary the sex ratio of offspring," *Science*, 179, 90–92.

van Lent, M. (2022): "Fathering Daughters and Personality," Discussion Paper 15012, IZA - Institute of Labor Economics.

Waldfogel, J. (1997): "The Effect of Children on Women's Wages," *American Sociological Review*, 62, 209–217.

Ware, J. E., M. Kosinski, D. M. Turner-Bowker, and B. Gandek (2001): "How to Score Version 2 of the SF-12 Health Survey (With a Supplement Documenting Version 1)," *Quality Metric Review*.

Washington, E. L. (2008): "Female Socialization: How Daughters Affect Their Legislator Fathers' Voting on Women's Issues," *The American Economic Review*, 98, 311–332.

**Appendices for: The Daughter Penalty**

Sonia Bhalotra, Damian Clarke, Angelina Nazarova

Online Only. Not for print.

# A Data Appendix

## A.1 Summary Statistics and Descriptive Figures

Table A1: Survey Waves and Birth Cohorts

| | Survey Waves | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Birth cohort | **2009** | **2010** | **2011** | **2012** | **2013** | **2014** | **2015** | **2016** | **2017** | **2018** | **2019** |
| 2010 cohort | -1 | 0 | +1 | +2 | +3 | +4 | +5 | +6 | +7 | +8 | +9 |
| 2011 cohort | -2 | -1 | 0 | +1 | +2 | +3 | +4 | +5 | +6 | +7 | +8 |
| 2012 cohort | -3 | -2 | -1 | 0 | +1 | +2 | +3 | +4 | +5 | +6 | +7 |
| 2013 cohort | -4 | -3 | -2 | -1 | 0 | +1 | +2 | +3 | +4 | +5 | +6 |
| 2014 cohort | -5 | -4 | -3 | -2 | -1 | 0 | +1 | +2 | +3 | +4 | +5 |
| 2015 cohort | -6 | -5 | -4 | -3 | -2 | -1 | 0 | +1 | +2 | +3 | +4 |
| 2016 cohort | -7 | -6 | -5 | -4 | -3 | -2 | -1 | 0 | +1 | +2 | +3 |
| 2017 cohort | -8 | -7 | -6 | -5 | -4 | -3 | -2 | -1 | 0 | +1 | +2 |
| 2018 cohort | -9 | -8 | -7 | -6 | -5 | -4 | -3 | -2 | -1 | 0 | +1 |

*Notes:* Lags and leads to birth timing are documented by birth cohort, with survey waves indicated in columns. Colours indicate pre-birth periods (blue), post birth periods (red) and birth years. Cohorts refer to the survey wave in which birth is observed, and survey waves are indicated in table columns. The "balanced sample" used in principal analyses consists of member of any cohort observed across each survey wave – ie are balanced across columns. Results are robust to working with the larger unbalanced sample of parents with any pre and post-birth waves observed. If a sample balanced in each of the 3 waves pre-birth and 5 years post-birth were used, this would cover only cohorts 2012-2014, and for this reason, our balanced sample focuses on balance by cohort across waves.

Table A2: Descriptive Statistics

| | Father | | | | Mother | | | |
|---|---|---|---|---|---|---|---|---|
| | Mean | SD | Min | Max | Mean | SD | Min | Max |
| Age | 27.02 | 5.68 | 16 | 48 | 25.05 | 5.18 | 16 | 45 |
| Married | 0.93 | 0.25 | 0 | 1 | 0.83 | 0.37 | 0 | 1 |
| Urban | 0.82 | 0.39 | 0 | 1 | 0.81 | 0.39 | 0 | 1 |
| Has a degree | 0.35 | 0.48 | 0 | 1 | 0.38 | 0.49 | 0 | 1 |
| **Labour market outcomes** | | | | | | | | |
| Monthly earnings | 2385.04 | 1436.64 | 33.68 | 9237.88 | 1682.52 | 1122.87 | 25.98 | 7698.23 |
| Weekly hours | 35.27 | 17.83 | 0 | 91 | 32.10 | 16.45 | 0 | 90 |
| Weekly hours c. on employment | 41.66 | 10.39 | 3 | 91 | 35.74 | 13.07 | 0 | 90 |
| Employed | 0.86 | 0.35 | 0 | 1 | 0.90 | 0.30 | 0 | 1 |
| Weekly hours w. self-employed | 35.40 | 17.24 | 0 | 91 | 32.13 | 16.35 | 0 | 90 |
| Part-time employment | 0.08 | 0.27 | 0 | 1 | 0.21 | 0.40 | 0 | 1 |
| Net monthly earnings | 1542.78 | 798.79 | 29.17 | 5300 | 1142.23 | 653.92 | 65 | 4196.99 |
| Occupation change | 0.11 | 0.32 | 0 | 1 | 0.11 | 0.31 | 0 | 1 |
| Job change | 0.12 | 0.33 | 0 | 1 | 0.13 | 0.34 | 0 | 1 |
| **Mental health** | | | | | | | | |
| Mental health (SF12) | 52.67 | 7.81 | 17.42 | 64.76 | 49.67 | 9.56 | 14.26 | 64.93 |
| Feeling worthless | 1.30 | 0.66 | 1 | 4 | 1.37 | 0.65 | 1 | 4 |
| Ability to face problems | 1.94 | 0.48 | 1 | 4 | 2.01 | 0.46 | 1 | 4 |
| Life satisfaction | 5.48 | 1.12 | 1 | 7 | 5.61 | 1.20 | 1 | 7 |
| Depression | 1.79 | 0.90 | 1 | 5 | 1.89 | 0.96 | 1 | 5 |
| **Home production** | | | | | | | | |
| Fraction of chores alone | 0.34 | 0.19 | 0 | 0.83 | 0.36 | 0.23 | 0 | 1 |
| Time spent on housework | 5.97 | 5.51 | 0 | 30 | 7.91 | 5.06 | 0 | 30 |
| Childcare mostly on self | 0.04 | 0.20 | 0 | 1 | 0.67 | 0.47 | 0 | 1 |
| **Family formation** | | | | | | | | |
| Having an additional child | 0.11 | 0.31 | 0 | 1 | 0.11 | 0.32 | 0 | 1 |
| Never consider divorce/separation | 0.97 | 0.17 | 0 | 1 | 0.99 | 0.11 | 0 | 1 |
| Never regret marriage/cohabitation | 0.97 | 0.17 | 0 | 1 | 0.99 | 0.10 | 0 | 1 |
| Relationship happiness index | 4.34 | 1.04 | 0 | 5 | 4.32 | 0.96 | 0 | 5 |
| **Parental values and attitudes** | | | | | | | | |
| Progressive news source | 0.75 | 0.44 | 0 | 1 | 0.72 | 0.45 | 0 | 1 |
| Progressive voting | 0.49 | 0.50 | 0 | 1 | 0.67 | 0.47 | 0 | 1 |
| Gender equality | 18.73 | 2.70 | 10 | 25 | 19.13 | 3.01 | 13 | 25 |
| **Health** | | | | | | | | |
| Physical health (SF12) | 54.49 | 5.94 | 27.51 | 67.74 | 53.86 | 6.90 | 25.73 | 69.56 |
| Number of cigarettes | 9.80 | 5.60 | 0 | 20 | 11.15 | 6.01 | 0 | 20 |

*Notes:* Characteristics are measured in the first wave of the survey which, by construction, is before the first child is born. Certain characteristics are not available in the first wave either by data availability (fraction of chores alone, housework, gender index) or by construction (job change and occupation change), and are measured in the wave prior to birth. Also, characteristics available only in the post-birth period (childcare, having an additional child) are measured across five post-birth periods. Sample excludes ethnic minorities. Longitudinal weight from wave 10 *j_indinus_lw* was used.

Table A2: Descriptive Statistics: balance test by child gender

| | Fathers | | | Mothers | | |
|---|---|---|---|---|---|---|
| | Daughter | Son | p-value ($H_0$: Diff=0) | Daughter | Son | p-value ($H_0$: Diff=0) |
| Age | 27.13 | 26.92 | 0.774 | 25.02 | 25.08 | 0.913 |
| Married | 0.93 | 0.93 | 0.837 | 0.85 | 0.81 | 0.310 |
| Urban | 0.81 | 0.83 | 0.766 | 0.82 | 0.80 | 0.578 |
| Has a degree | 0.30 | 0.40 | 0.078 | 0.34 | 0.43 | 0.063 |
| **Labour market outcomes** | | | | | | |
| Monthly earnings | 2422.97 | 2348.08 | 0.707 | 1697.23 | 1665.68 | 0.791 |
| Weekly hours | 35.41 | 35.13 | 0.905 | 31.41 | 32.91 | 0.372 |
| Weekly hours c. on employment | 41.85 | 41.48 | 0.798 | 35.53 | 35.99 | 0.737 |
| Employed | 0.86 | 0.87 | 0.821 | 0.89 | 0.92 | 0.346 |
| Weekly hours w. self-employed | 35.87 | 34.98 | 0.681 | 31.49 | 32.88 | 0.400 |
| Part-time employment | 0.06 | 0.10 | 0.189 | 0.22 | 0.19 | 0.595 |
| Net monthly earnings | 1580.64 | 1505.90 | 0.500 | 1152.45 | 1130.52 | 0.752 |
| Occupation change | 0.09 | 0.13 | 0.452 | 0.13 | 0.09 | 0.235 |
| Job change | 0.09 | 0.14 | 0.284 | 0.17 | 0.09 | 0.063 |
| **Mental health** | | | | | | |
| Mental health (SF12) | 52.29 | 53.01 | 0.464 | 50.38 | 48.85 | 0.113 |
| Feeling worthless | 1.26 | 1.30 | 0.527 | 1.28 | 1.33 | 0.433 |
| Ability to face problems | 1.89 | 1.96 | 0.155 | 1.98 | 1.99 | 0.833 |
| Life satisfaction | 5.40 | 5.55 | 0.304 | 5.70 | 5.51 | 0.113 |
| Depression | 1.85 | 1.74 | 0.402 | 1.85 | 1.94 | 0.375 |
| **Home production** | | | | | | |
| Fraction of chores alone | 0.35 | 0.33 | 0.595 | 0.38 | 0.34 | 0.295 |
| Time spent on housework | 5.28 | 6.37 | 0.298 | 7.66 | 8.11 | 0.526 |
| Childcare mostly on self | 0.05 | 0.03 | 0.208 | 0.69 | 0.65 | 0.189 |
| **Family formation** | | | | | | |
| Having an additional child | 0.11 | 0.10 | 0.595 | 0.12 | 0.11 | 0.380 |
| Never consider divorce/separation | 0.96 | 0.98 | 0.493 | 0.99 | 0.99 | 0.994 |
| Never regret marriage/cohabitation | 0.96 | 0.98 | 0.410 | 0.99 | 0.99 | 0.422 |
| Relationship happiness index | 4.28 | 4.39 | 0.416 | 4.25 | 4.41 | 0.081 |
| **Parental values and attitudes** | | | | | | |
| Progressive news source | 0.71 | 0.78 | 0.211 | 0.76 | 0.67 | 0.140 |
| Progressive voting | 0.54 | 0.44 | 0.209 | 0.65 | 0.68 | 0.631 |
| Gender equality | 19.04 | 18.47 | 0.402 | 19.46 | 18.89 | 0.334 |
| **Health** | | | | | | |
| Physical health (SF12) | 54.24 | 54.71 | 0.521 | 53.86 | 53.87 | 0.988 |
| Number of cigarettes | 8.61 | 11.33 | 0.306 | 10.96 | 11.45 | 0.882 |

*Notes:* Characteristics are measured in the first wave of the survey which, by construction, is before the first child is born. Certain characteristics are not available in the first wave either by data availability (fraction of chores alone, housework, gender index) or by construction (job change and occupation change), and are measured in the wave prior to birth. Being married is also measured in the wave prior to birth. Also, characteristics available only in the post-birth period (childcare, time with kid, dinner with kid, having an additional child) are measured across five post-birth periods. Sample excludes ethnic minorities. Longitudinal weight from wave 10 *j_indinus_lw* was used.

Table A2: Descriptive Statistics: balance test by parent gender

| | Daughter | | | Son | | |
|---|---|---|---|---|---|---|
| | Father | Mother | p-value ($H_0$: Diff=0) | Father | Mother | p-value ($H_0$: Diff=0) |
| Age | 27.13 | 25.02 | 0.000 | 26.92 | 25.08 | 0.003 |
| Married | 0.93 | 0.85 | 0.033 | 0.93 | 0.81 | 0.001 |
| Urban | 0.81 | 0.82 | 0.765 | 0.83 | 0.80 | 0.592 |
| Has a degree | 0.30 | 0.34 | 0.387 | 0.40 | 0.43 | 0.578 |
| **Labour market outcomes** | | | | | | |
| Monthly earnings | 2422.97 | 1697.23 | 0.000 | 2348.08 | 1665.68 | 0.000 |
| Weekly hours | 35.41 | 31.41 | 0.046 | 35.13 | 32.91 | 0.256 |
| Weekly hours c. on employment | 41.85 | 35.53 | 0.000 | 41.48 | 35.99 | 0.000 |
| Employed | 0.86 | 0.89 | 0.394 | 0.87 | 0.92 | 0.148 |
| Weekly hours w. self-employed | 35.87 | 31.49 | 0.024 | 34.98 | 32.88 | 0.245 |
| Part-time employment | 0.06 | 0.22 | 0.000 | 0.10 | 0.19 | 0.030 |
| Net monthly earnings | 1580.64 | 1152.45 | 0.000 | 1505.90 | 1130.52 | 0.000 |
| Occupation change | 0.09 | 0.13 | 0.425 | 0.13 | 0.09 | 0.258 |
| Job change | 0.09 | 0.17 | 0.102 | 0.14 | 0.09 | 0.221 |
| **Mental health** | | | | | | |
| Mental health (SF12) | 52.29 | 50.38 | 0.039 | 53.01 | 48.85 | 0.000 |
| Feeling worthless | 1.26 | 1.28 | 0.652 | 1.30 | 1.33 | 0.632 |
| Ability to face problems | 1.89 | 1.98 | 0.088 | 1.96 | 1.99 | 0.542 |
| Life satisfaction | 5.40 | 5.70 | 0.024 | 5.55 | 5.51 | 0.743 |
| Depression | 1.85 | 1.85 | 0.994 | 1.74 | 1.94 | 0.087 |
| **Home production** | | | | | | |
| Fraction of chores alone | 0.35 | 0.38 | 0.454 | 0.33 | 0.34 | 0.685 |
| Time spent on housework | 5.28 | 7.66 | 0.002 | 6.37 | 8.11 | 0.045 |
| Childcare mostly on self | 0.05 | 0.69 | 0.000 | 0.03 | 0.65 | 0.000 |
| **Family formation** | | | | | | |
| Having an additional child | 0.11 | 0.12 | 0.821 | 0.10 | 0.11 | 0.964 |
| Never consider divorce/separation | 0.96 | 0.99 | 0.145 | 0.98 | 0.99 | 0.511 |
| Never regret marriage/cohabitation | 0.96 | 0.99 | 0.031 | 0.98 | 0.99 | 0.591 |
| Relationship happiness index | 4.28 | 4.25 | 0.743 | 4.39 | 4.41 | 0.828 |
| **Parental values and attitudes** | | | | | | |
| Progressive news source | 0.71 | 0.76 | 0.316 | 0.78 | 0.67 | 0.025 |
| Progressive voting | 0.54 | 0.65 | 0.068 | 0.44 | 0.68 | 0.000 |
| Gender equality | 19.04 | 19.46 | 0.530 | 18.47 | 18.89 | 0.496 |
| **Health** | | | | | | |
| Physical health (SF12) | 54.24 | 53.86 | 0.605 | 54.71 | 53.87 | 0.245 |
| Number of cigarettes | 8.61 | 10.96 | 0.401 | 11.33 | 11.45 | 0.967 |

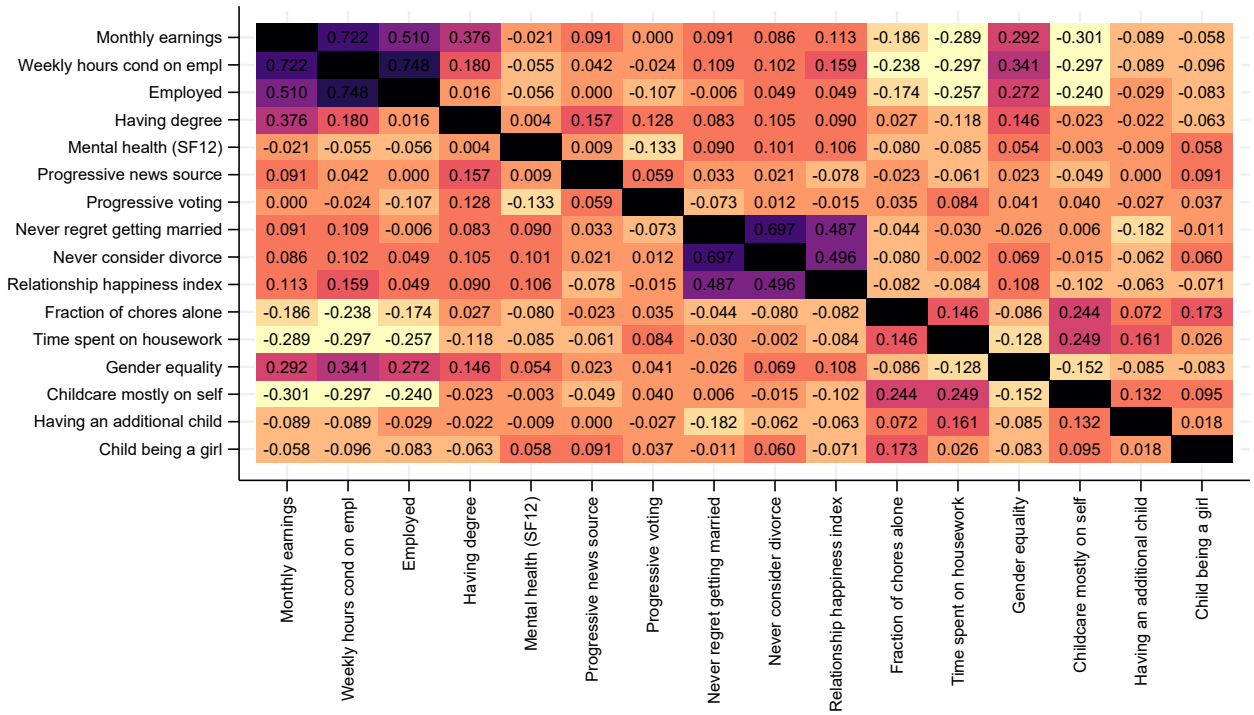
*Notes:* Characteristics are measured in the first wave of the survey which, by construction, is before the first child is born. Certain characteristics are not available in the first wave either by data availability (fraction of chores alone, housework, gender index) or by construction (job change and occupation change), and are measured in the wave prior to birth. Being married is measured in the wave prior to birth. Also, characteristics available only in the post-birth period (childcare, time with kid, dinner with kid, having an additional child) are measured across five post-birth periods. Sample excludes ethnic minorities. Longitudinal weight from wave 10 *j_indinus_lw* was used.

Table A3: Summary Statistics by sample

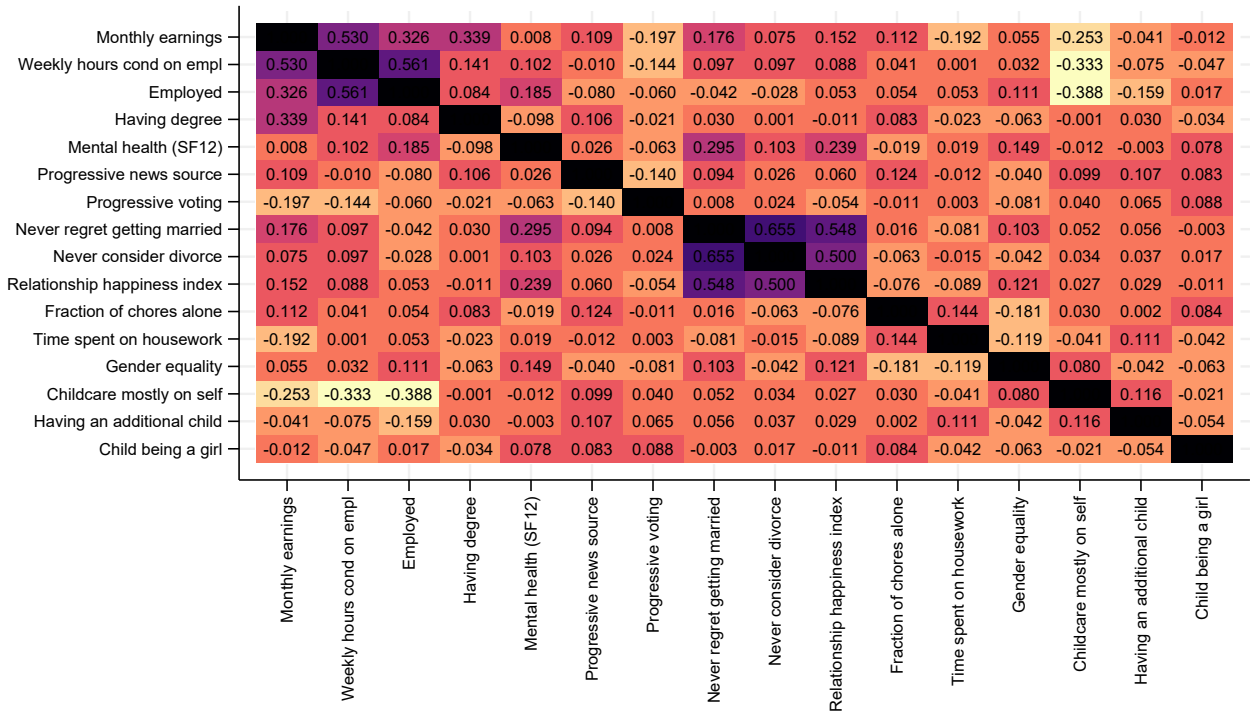
| | Full Sample | | Analysis Sample | | |
|---|---|---|---|---|---|
| | Mean | SD | Mean | SD | p-value |
| **Panel A: Total** | | | | | |
| Age | 25.94 | 5.44 | 25.93 | 5.49 | 0.969 |
| Urban | 0.81 | 0.39 | 0.82 | 0.39 | 0.906 |
| Married | 0.58 | 0.49 | 0.59 | 0.49 | 0.854 |
| Has a degree | 0.36 | 0.48 | 0.37 | 0.48 | 0.764 |
| Monthly earnings | 1968.80 | 1301.94 | 1972.40 | 1307.17 | 0.963 |
| Weekly hours | 33.64 | 16.98 | 33.45 | 17.11 | 0.848 |
| Employed | 0.89 | 0.32 | 0.88 | 0.32 | 0.876 |
| Observations | 1848 | | 748 | | |
| **Panel B: Fathers** | | | | | |
| Age | 27.09 | 5.53 | 27.02 | 5.68 | 0.890 |
| Urban | 0.81 | 0.39 | 0.82 | 0.39 | 0.856 |
| Married | 0.60 | 0.49 | 0.60 | 0.49 | 0.946 |
| Has a degree | 0.35 | 0.48 | 0.35 | 0.48 | 0.938 |
| Monthly earnings | 2364.50 | 1424.20 | 2385.04 | 1436.64 | 0.882 |
| Weekly hours | 36.23 | 17.01 | 35.27 | 17.83 | 0.549 |
| Employed | 0.88 | 0.32 | 0.86 | 0.35 | 0.508 |
| Observations | 820 | | 307 | | |
| **Panel C: Mothers** | | | | | |
| Age | 24.99 | 5.18 | 25.05 | 5.18 | 0.881 |
| Urban | 0.82 | 0.39 | 0.81 | 0.39 | 0.988 |
| Married | 0.57 | 0.50 | 0.58 | 0.49 | 0.740 |
| Has a degree | 0.37 | 0.48 | 0.38 | 0.49 | 0.738 |
| Monthly earnings | 1675.13 | 1118.23 | 1682.52 | 1122.87 | 0.929 |
| Weekly hours | 31.67 | 16.72 | 32.10 | 16.45 | 0.716 |
| Employed | 0.89 | 0.31 | 0.90 | 0.30 | 0.627 |
| Observations | 1028 | | 441 | | |

*Notes:* Summary statistics are for wave 1 only when everyone is childless. The samples excludes ethnic minorities. The reported p-value refers to a test of equality of means across samples. Longitudinal weight from wave 10 *j_indinus_lw* was used in both samples.

Figure A1: Correlations between key measures




(a) Mother




(b) Father

## A.2 Variable definitions

**Child (*lprnt, relationship_dv, alwstat*).** We use the lprnt variable in the first wave of the dataset, which asks whether an individual has ever had a biological child before the start of the sample period. We restrict the sample to individuals who have never had a biological child. There is no variable in the UKHLS that cumulatively tracks the total number of children an individual has ever had. We must create this from the household grid file 'egoalt'. This file contains an observation for every pair of individuals within each household. The variable relationship_dv tells us the relationship of the respondent ("ego") to the any other person in the household ("alter"). We use this to select only natural children of the respondent. In each wave we calculate the total number of new natural children of that individual added to the household, using the variable alwstat. We count only children who are 'newborn' or 'new entrant' to the sample according to the alwstat variable. In each wave we add the number of new biological children added in that wave to the total number of biological children in the previous wave, to create a cumulative total number of children that individual has ever had. By only counting children who are 'newborn' or 'new entrant', we ensure that we do not double count children across waves. A new entrant is a new born child who has never appeared in the dataset who is older than 0 years of age. We call this new cumulative measure 'numchild'. Once we have created 'numchild', we define a dummy variable Child that takes value 1 if an individual has ever had a child and 0 if they have not.

**Girl (*sex*).** We merge information on child gender from the child file to the main file of adults. We ensure that we keep only the children who are counted in the cumulative measure of biological children ever had, 'numchild', that we created above (see 'Child'). We must make this adjustment because the child file is constructed on the basis of the 'relationship' variable (rather than the corrected 'relationship_dv' variable, which we used to create 'numchild') and so incorrectly refers to some individuals as biological children when they are not. After dealing with this error, we create a variable that takes the value 1 if the first born child is a Girl (sex==2) and 0 otherwise (sex==1).

**Married (*marstat_dv*).** We use the variable marstat_dv, which captures de facto marital status, to create a variable that tells us an individual's current marital status. We create a dummy variable that takes value 1 if the individual is married or in civil partnership or is living as couple and 0 otherwise.

**Ethnic Minority (*racel_dv*).** We use the variable racel_dv, which captures an individual's race, to define dummy variables for Black, Asian, and Other Race. For Black, we include individuals who are 'White and Black Caribbean', 'White and Black African', 'Caribbean', 'African', and 'any other Black background'. For Asian, we include individuals who are 'White and Asian', 'Indian', 'Pakistani', 'Bangladeshi', 'Chinese', and 'any other Asian background'. For Other Race, we include individuals who are 'any other mixed background', 'Arab', 'any other ethnic background', and 'Gypsy or Irish traveller'.

**Degree (*hiqual_dv*).** We create a dummy variable that takes value 1 if an individual has degree level education, i.e. if hiqual_dv = 1.

**Partner (*ppno, hhid, ppid*).** To account for having a partner in the household, we use ppno>0 as an indicator, since ppno=0 stands for "no partner in the household". When we use partner's

information, we merge the data from indresp file based on partner's unique identifier ppid.

## A.3 Labour market outcomes

**Employed (*jbhas, jboff, jbstat*).** Dummy variable taking value 1 if individual is employed and 0 if they are not. We follow the ONS definition of employment. An individual is employed if they either did paid work last week (jbhas = 1), or if they did not do paid work last week but have a paid job (jboff = 1). An individual is also employed if they are on a government training scheme, or are doing unpaid work for a family business (jbstat = 9 or 10).

**Weekly Hours (*jbhrs, jbot, jbsemp*).** We sum the weekly hours an individual is expected to work in a normal week (jbhrs) and the number of overtime hours they work in a normal week (jbot). Both variables are defined only for those in paid employment. Therefore, our weekly hours variable takes value 0 if an individual is not in paid employment, a positive value if they are in paid employment, and missing values for the self-employed.

**Weekly Hours including self-employed (*jbhrs, jbot, jbsemp*).** We sum the weekly hours an individual is expected to work in a normal week (jbhrs) and the number of overtime hours they work in a normal week (jbot). We add hours indicated for self-employed. Therefore, our weekly hours variable takes value 0 if an individual is in neither in paid employment nor self-employed, a positive value if they are in paid employment or the self-employed.

**Monthly Earnings (*paygu_dv, All Items CPI Index (ONS)*).** We take gross monthly wages (paygu_dv) and convert into real wages (base year 2015) using the All Items CPI Index. This variable is missing for any individual who is not in paid employment.

**Net Monthly earnings (*paynu_dv*).** We use paynu_dv to construct the measure for net earnings. This variable is taken from the Understanding Society dataset and captures usual net pay per month, where 'net' refers to net of taxes on earnings and national insurance contributions.

**Part-time employment (*jbft_dv*).** This variable is derived from the variable 'full or part-time employee', where employed full time means working more than 30 hours per week. This measure is based on total hours, i.e. including both normal and overtime hours. It is computed for both employees and the self employed. Inapplicable to proxy respondents due to missing information on overtime, and respondents who do not have a paid job.

**Occupation or Industry change (*jbsoc00_cc, jbsic07*).** We create an impulse dummy variable that takes value 1 if the occupation code in the current wave differs from the occupation code in the previous wave. This tells us whether the individual has change occupation since the last wave. We create an impulse dummy variable that takes value 1 if the industry code in the current wave differs from the industry code in the previous wave. This tells us whether the individual has change industry since the last wave.

**Job change (*jbsoc00_cc, jbsic07*).** We create an impulse dummy variable that takes value 1 if either of the variables change occupation or change industry take value 1. This is a dummy variable that tells us whether an individual changed occupation or industry since the previous wave.

## A.4 Parental mental health

**Mental Health (*sf12mcs_dv*).** We create a measure of the mental health of the parent. Mental health variable is a continuous variable and it is measured at SF-12 scale for mental component summary. A higher score means a better state of mental health.

**Life satisfaction (*sclfsato*).** We define life satisfaction following the question whether the respondent is satisfied with their life overall. The scale of answers consist of 7 choices from completely dissatisfied to completely satisfied. The variable is discrete, ranging from 1 to 7.

**Depression (*scsf6c*).** Another measure of life satisfaction and mental health is to which extent the respondent is feeling depressed. We use the question "Have you felt downhearted and depressed?". We reverse the original variable and create 5-scale variable ranging from none of the time (1) to all of the time (5). This measure covers the period of the last 4 weeks before the respondent is asked.

**Feeling worthless (*scghqk*).** We use the following question to construct a variable indicating whether parent is feeling worthless: "Have you recently been thinking of yourself as a worthless person?". Answer is categorized in 4 scales: "Not at all", "No more than usual", "Rather more than usual", "Much more than usual". We keep the original scaling from 1 to 4, where the highest number indicates feeling worthless.

**Ability to face problems (*scghqh*).** We use the following question to construct a variable indicating whether parent's ability to face problems deteriorated after birth: "Have you recently been able to face up to problems?" Answer is categorized in 4 scales: "More so than usual", "Same as usual", "Less able than usual", "Much less able". We keep the original scaling from 1 to 4, where the highest number indicates being less able to face up problems.

## A.5 Home production

**Hours per week spent on housework (*howlng*).** We define housework hours using the following question: "About how many hours do you spend on housework in an average week, such as time spent cooking, cleaning and doing the laundry?"

**Fraction Chores (*hubuys, hufrys, humops, huiron, hupots, hudiy*).** We create the fraction of six household chores that a father or mother does mostly him/herself. The UKHLS asks individuals with partners who mostly does each of six household chores: grocery shopping, cooking, cleaning, washing/ironing, gardening and DIY jobs. The individual may respond: mostly self, mostly spouse/partner, shared, paid help, other. We create a dummy variable that takes the value 1 if the parent responds: mostly self. We then calculate the fraction of these six chores that he does mostly him/herself.

**Childcare (*husits*).** We define three measures of childcare using the question "Who is mainly responsible for looking after the children?" as follows: (1) Shared - a dummy variable taking 1 if respondent and her/his partner are both mainly responsible for looking after the children, and 0 otherwise, (2) Mostly on him/herself - a dummy variable taking 1 if respondent says he/she is mainly responsible for looking after the children, and 0 otherwise, (3) Mostly on the partner

- a dummy variable taking 1 if respondent says that his/her partner is mainly responsible for looking after the children, and 0 otherwise.

## A.6 Family formation

**Never considering divorce/separation (*screlpards*).** We construct a dummy variable based on the Understanding Society variable *screlpards*, which measures whether the respondent has considered divorce or separation: "(Considering your relationship with your partner or spouse...) How often do you discuss or consider divorce, separation or terminating your relationship?". This variable is available for people who have partners in the household. We also impute the values from wave 1 to wave 2 and so on, since this variable is available at odd waves only by construction. The dummy variable takes the value of 1 if the respondent reports a value of 5 or higher on this scale, indicating that they have never considered divorce. If the respondent reports a value below 5, the dummy variable takes the value of 0.

**Never regretting getting married/living together (*screlparrg*).** We create a dummy variable using the Understanding Society variable *screlparrg*, which assesses whether the respondent regrets getting married or living together: "(Considering your relationship with your partner or spouse...) Do you ever regret that you married or lived together?". This variable is available for people who have partners in the household. We also impute the values from wave 1 to wave 2 and so on, since this variable is available at odd waves only by construction. The dummy variable is assigned a value of 1 if the respondent reports a value of 5 or higher, indicating that they do not regret getting married. A value of 0 is assigned if the respondent reports a value below 5.

**Relationship happiness index (*screlparrg, screlparir, screlpards, screlhappy, screlparar*).** We construct an index that measures overall relationship happiness by summing five binary variables derived from the Understanding Society dataset. These variables include: *screlparrg* (regret getting married/living together), *screlparir* (getting on each other's nerves), *screlpards* (considering divorce/separation), *screlhappy* (degree of happiness in the relationship), and *screlparar* (quarreling with the partner). Each of these variables is converted into a dummy taking a value of 1 if the respondent reports a value of 5 or higher and 0 otherwise. We also impute the values from wave 1 to wave 2 and so on, since this variable is available at odd waves only by construction. The final index is the sum of these dummies, providing a composite measure of relationship happiness.

**Having an additional child (*addedchild*).** We create a dummy which takes value of 1 if respondent has a second child in the following years since the birth of the first one, and 0 otherwise.

## A.7 Parental values and attitude

**Progressive news (*paperm2, tvm2, netm2*).** We create a measure of the progressiveness by establishing news sources. The UKHLS asks which sources provide you personally with the most information about news and current affairs. If the respondent mentions TV, newspapers (national or local), or the internet, they are asked a follow up question about which one TV

channel, newspaper or website is the source they use most frequently. We create a dummy variable that takes the value 0 if the respondent says that the newspaper or website version of The Sun, The Daily Mail, The Daily Express or The Daily Star is their most frequently used news source. The dummy variable takes the value 1 if the respondent uses either TV, newspapers or the internet to receive their news and current affairs, and for no medium mentions one of these four organisations as their main source.

**Progressive political attitudes (*vote3, vote4*).** We create another dummy, which takes a value of 0 if a respondent voted for the conservative party and 1 if otherwise. We look at the parties represented in the House of Commons and those parties without representation in the House of Commons, but with representation in other UK legislatures. This information comes from the party registries of the Electoral Commission. Conservative parties in the UKHLS include conservative, Ulster unionist, SDLP, alliance party, democratic unionist, UK independence party, British national party.

**Gender equality index (*scopfama-scopfamf*).** The gender attitudes index was generated using five questions reflecting gender attitudes. We use five UKHLS questions measuring gender views. In particular, we create respective dummy variables (from 1 "strongly agree" to 5 "strongly disagree") for each of the following statements: "pre-school child suffers if the mother works", "family suffers if mother works full-time", "husband should earn, wife should stay at home", "husband and wife should contribute to household income" (we reverse the scale for this variable, i.e. 1 "strongly disagree" to 5 "strongly agree"), "employers should help mothers combine jobs and childcare" (we reverse the scale for this variable, i.e. 1 "strongly disagree" to 5 "strongly agree"). We create an additive index (5 to 25), where the higher the score, the stronger support for equal gender roles.

## A.8 Health

**SF12 Physical Component Summary (*sf12pcs_dv*).** This measure converts valid answers to the origin questions w_sf1 w_sf2a w_sf2b w_sf3a w_sf3b w_sf4a w_sf4b w_sf5 w_sf6a w_sf6b w_sf6c w_sf7 into a single physical functioning score, resulting in a continuous scale with a range of 0 (low functioning) to 100 (high functioning). For complete details on scoring methods see Ware et al. (2001).

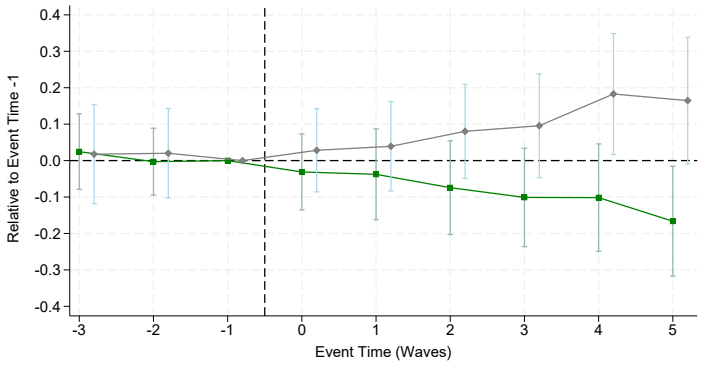
**Number of cigarettes (*ncigs*).** We use this variable to capture the usual number of cigarettes smoked per day. The exact question asked in the survey is as follows: "Approximately how many cigarettes a day do you usually smoke, including those you roll yourself?". This is a continuous variable.

# B Alternative Specifications

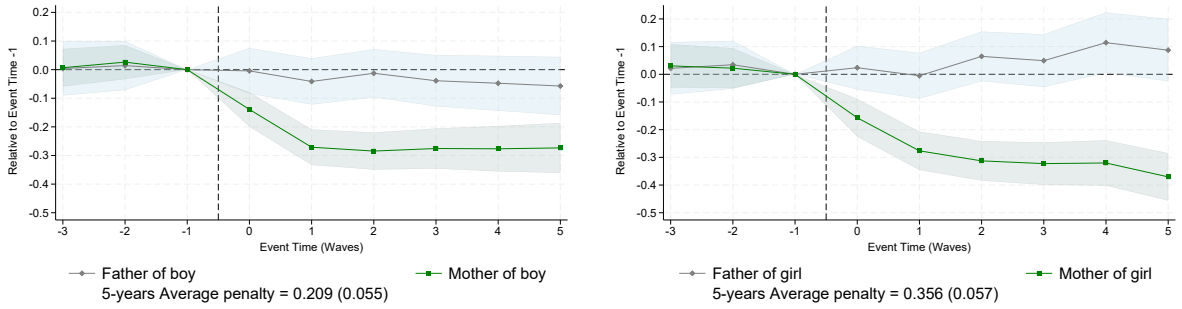
## B.1 Labour market outcomes on the full sample of parents

Figures B1-B3 present the labor market trajectories for both mothers and fathers following the birth of their first child, separated by child gender (boy or girl), run on the full sample of parents. The results support the main findings of this paper. Figure B1 illustrates the overall earnings penalty, showing that mothers experience a substantial decline in earnings after childbirth, particularly if their firstborn is a girl, while fathers' earnings remain largely unaffected if they have a boy, and increasing if they have a girl. Figure B2 depicts the employment status, highlighting that mothers face a significant drop in employment rates after childbirth, with a more pronounced effect for those with daughters. Fathers, in contrast, display no changes in employment regardless of the child's gender. Figure B3 focuses on working hours (conditional on employment), showing that mothers, especially those with daughters, tend to reduce their working hours significantly after childbirth. Fathers, on the other hand, exhibit little change in their working hours.

Figure B1: The child penalty in monthly earnings by gender of the firstborn child



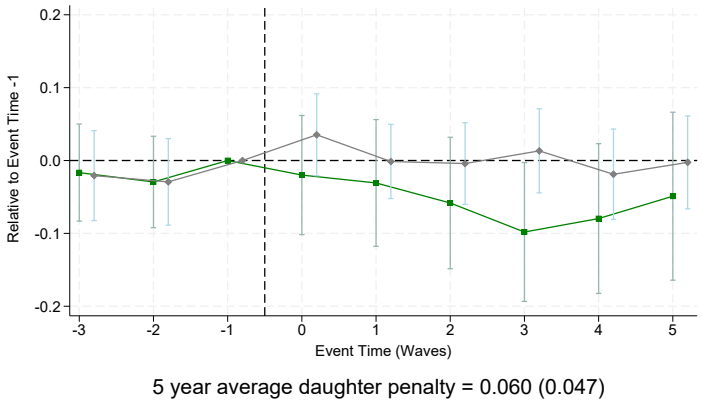

(a) Differential responses by child gender: Father and Mother



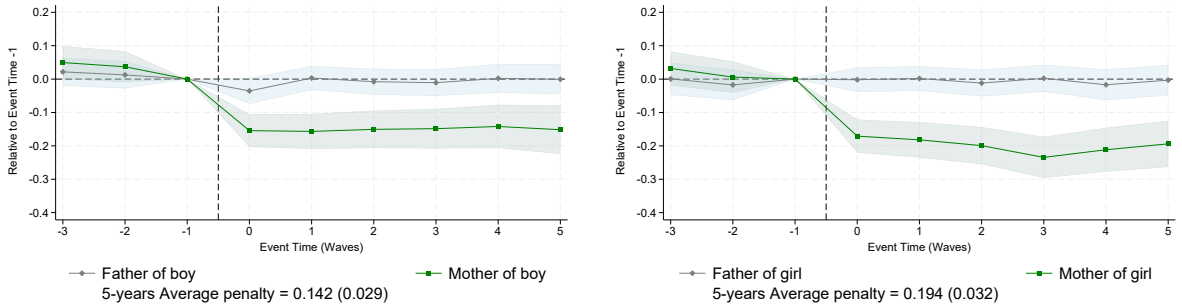

(b) Monthly earnings: Boy and Girl

*Notes:* Top panel (a) presents estimates of $\widehat{\alpha}_j^{Girl,g}$ following equation (3) examining differences in responses by girl children for mothers (green) and fathers (grey). Panel (b) presents estimates of $\widehat{\alpha}_j^{g}$ following equation (1) and (2) for parents of boy (left hand panel) and parents of girls (right hand panels). In all cases, estimation is based on a full sample of parents who had their first child between 2009 and 2019. Shaded blue and green areas or error bars represent 95% confidence intervals based on robust standard errors.

Figure B2: The child penalty in monthly earnings by gender of the firstborn child



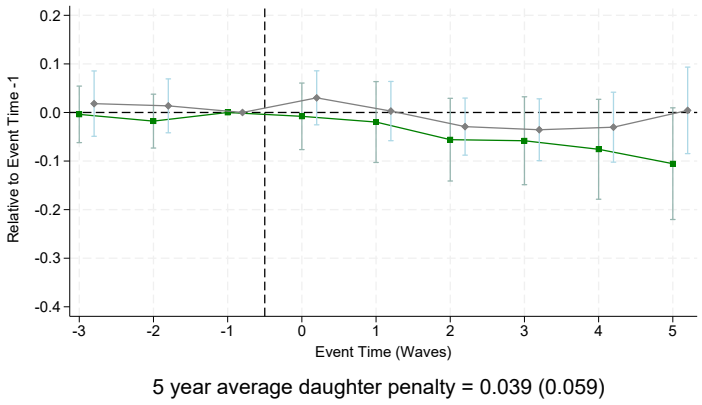

(a) Differential responses by child gender: Father and Mother



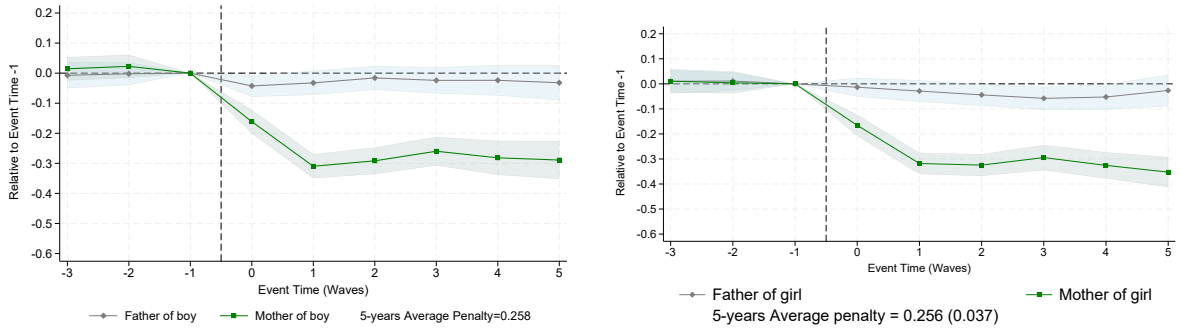

(b) Employment: Boy and Girl

*Notes:* Top panel (a) presents estimates of $\widehat{\alpha}_j^{Girl,g}$ following equation (3) examining differences in responses by girl children for mothers (green) and fathers (grey). Panel (b) presents estimates of $\widehat{\alpha}_j^{g}$ following equation (1) and (2) for parents of boy (left hand panel) and parents of girls (right hand panels). In all cases, estimation is based on a full sample of parents who had their first child between 2009 and 2019. Shaded blue and green areas or error bars represent 95% confidence intervals based on robust standard errors.

Figure B3: The child penalty in hours conditional on employment by gender of the firstborn child




(a) Differential responses by child gender: Father and Mother



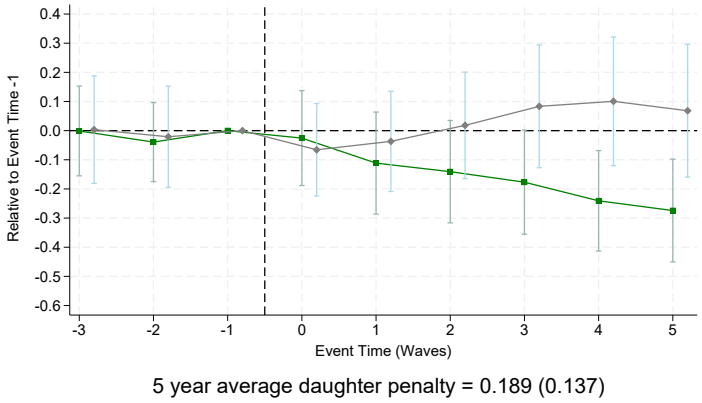

(b) Hours conditional on employment: Boy and Girl

*Notes:* Top panel (a) presents estimates of $\widehat{\alpha}_j^{Girl,g}$ following equation (3) examining differences in responses by girl children for mothers (green) and fathers (grey). Panel (b) presents estimates of $\widehat{\alpha}_j^{g}$ following equation (1) and (2) for parents of boy (left hand panel) and parents of girls (right hand panels). In all cases, estimation is based on a full sample of parents who had their first child between 2009 and 2019. Shaded blue and green areas or error bars represent 95% confidence intervals based on robust standard errors.

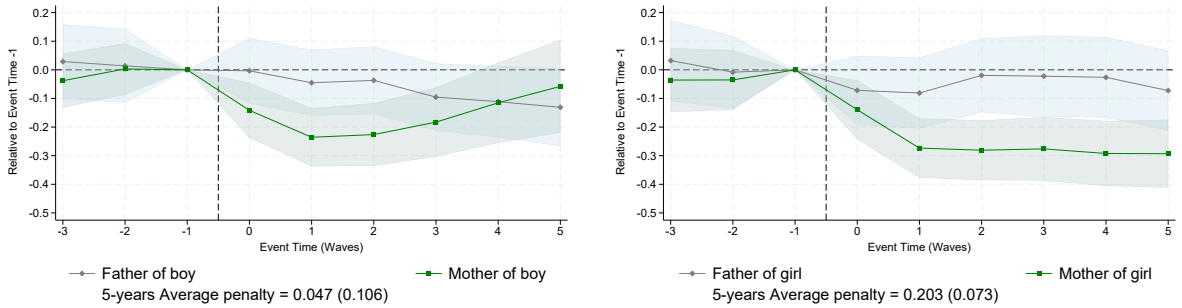
## B.2 Labour market outcomes on the analysis sample including ethnic minorities

Figures B4-B6 present the labour market trajectories for both mothers and fathers following the birth of their first child, separated by child gender (boy or girl), using the full sample of parents. These results align with the main findings of the paper. Figure B4 illustrates the overall earnings penalty, showing that mothers experience a substantial decline in earnings after childbirth, particularly if their firstborn is a girl, while fathers' earnings remain relatively stable when they have a boy and tend to increase slightly if they have a girl. Figure B5 focuses on employment status, highlighting that mothers experience a significant drop in employment rates after childbirth, especially if they have daughters, whereas fathers display no notable changes in employment, regardless of the child's gender. Figure B6 looks at working hours conditional on employment, demonstrating that mothers, particularly those with daughters, significantly reduce their working hours after childbirth. Fathers, interestingly, also work less when they have a daughter reducing the gap between mothers and fathers.

Figure B4: The child penalty in monthly earnings by gender of the firstborn child



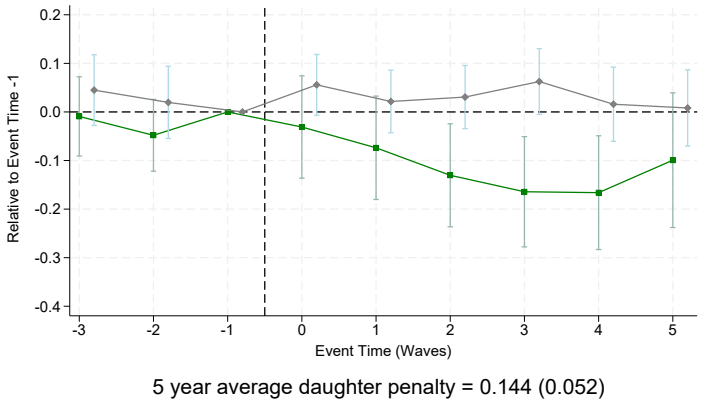

(a) Differential responses by child gender: Father and Mother



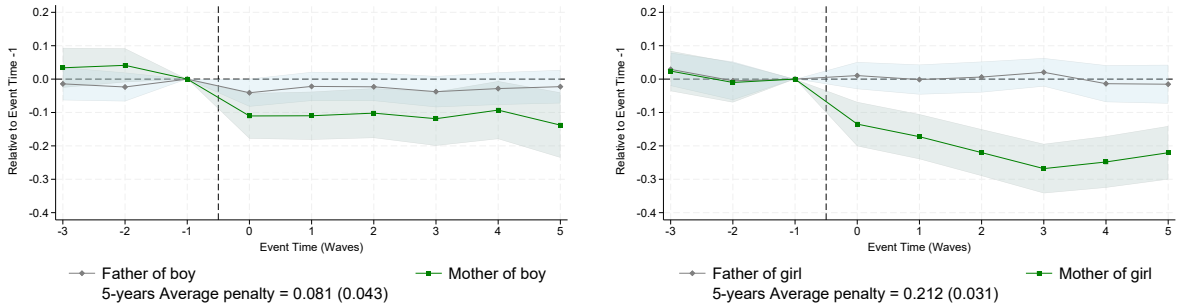

(b) Monthly earnings: Boy and Girl

*Notes:* Top panel (a) presents estimates of $\widehat{\alpha}_j^{Girl,g}$ following equation (3) examining differences in responses by girl children for mothers (green) and fathers (grey). Panel (b) presents estimates of $\widehat{\alpha}_j^{g}$ following equation (1) and (2) for parents of boy (left hand panel) and parents of girls (right hand panels). In all cases, estimation is based on the analysis sample of parents including ethnic minorities who had their first child between 2009 and 2019. Shaded blue and green areas or error bars represent 95% confidence intervals based on robust standard errors.

Figure B5: The child penalty in employment by gender of the firstborn child



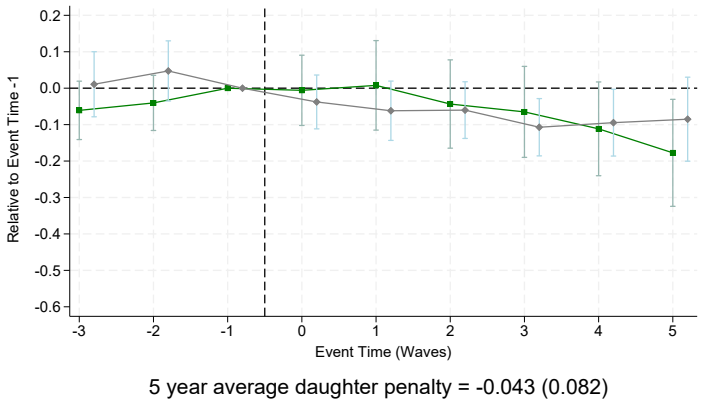

(a) Differential responses by child gender: Father and Mother



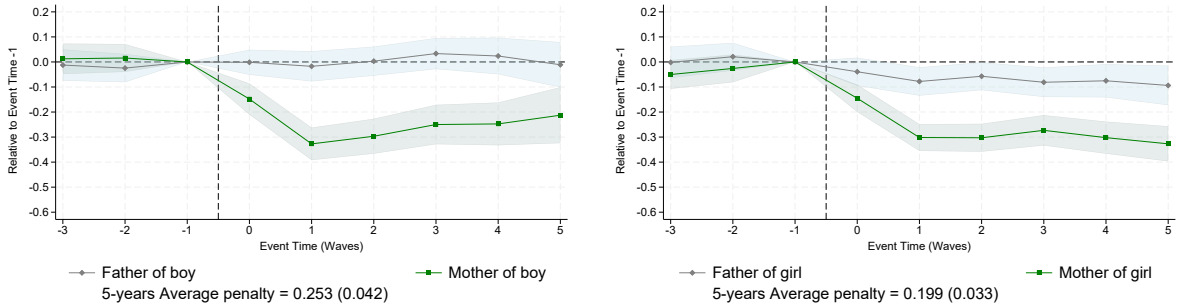

(b) Employment: Boy and Girl

*Notes:* Top panel (a) presents estimates of $\widehat{\alpha}_j^{Girl,g}$ following equation (3) examining differences in responses by girl children for mothers (green) and fathers (grey). Panel (b) presents estimates of $\widehat{\alpha}_j^{g}$ following equation (1) and (2) for parents of boy (left hand panel) and parents of girls (right hand panels). In all cases, estimation is based on the analysis sample of parents including ethnic minorities who had their first child between 2009 and 2019. Shaded blue and green areas or error bars represent 95% confidence intervals based on robust standard errors.

Figure B6: The child penalty in hours conditional on employment by gender of the firstborn child




(a) Differential responses by child gender: Father and Mother




(b) Hours conditional on employment: Boy and Girl

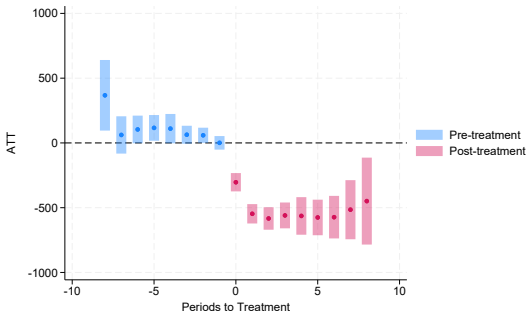
*Notes:* Top panel (a) presents estimates of $\widehat{\alpha}_j^{Girl,g}$ following equation (3) examining differences in responses by girl children for mothers (green) and fathers (grey). Panel (b) presents estimates of $\widehat{\alpha}_j^{g}$ following equation (1) and (2) for parents of boy (left hand panel) and parents of girls (right hand panels). In all cases, estimation is based on the analysis sample of parents including ethnic minorities who had their first child between 2009 and 2019. Shaded blue and green areas or error bars represent 95% confidence intervals based on robust standard errors.

## B.3 Labour market trajectories using Callaway & Sant'Anna estimator

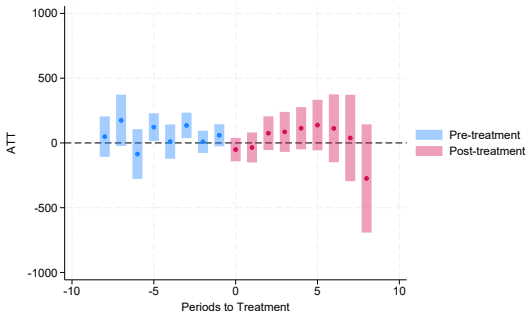
In this section we report figures capturing absolute effects of child-bearing rather than proportional changes as in (1), focused only on child gender-specific responses among mothers and fathers. We use Callaway and Sant'Anna (2021)'s estimator, which we describe in more detail in the Section 3. First, we support our findings of the general penalty for mothers across earnings, employment, and hours worked. Second, we compare average absolute changes by child gender. We find consistency with our main results, where we observe larger penalties in earnings, employment, and hours worked for mothers of girls.

Figure B7: Child penalty in labour market outcomes: all children

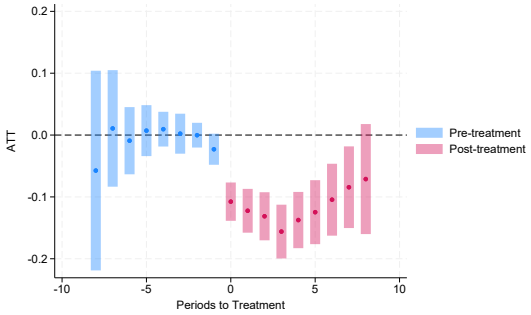
(a) Monthly earnings - Mothers



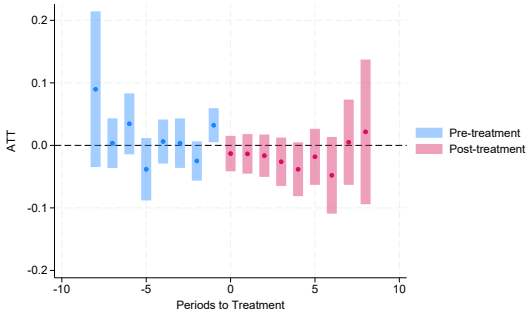

(b) Monthly earnings - Fathers



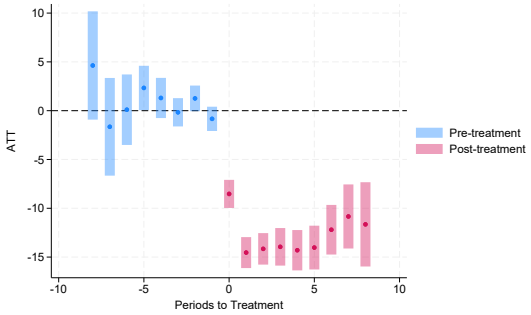

(c) Employment - Mothers



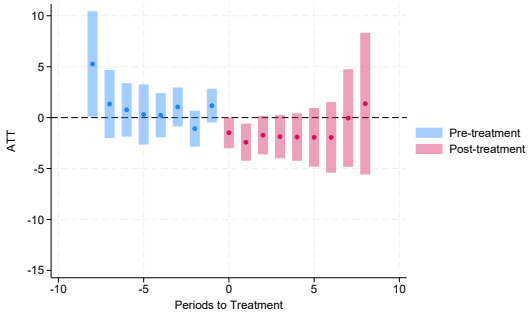

(d) Employment - Fathers




(e) Total hours - Mothers



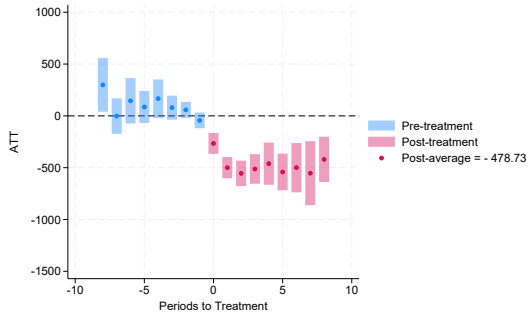

(f) Total hours - Fathers



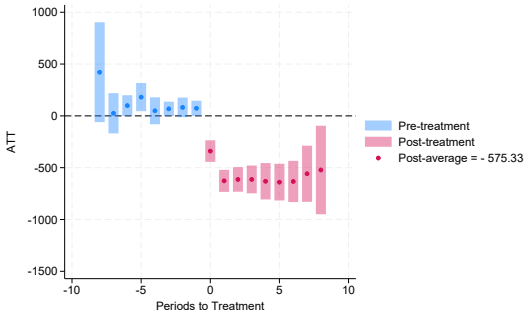

*Notes:* This graph shows the adoption of Callaway & Sant'Anna estimator using working hours. Similarly to the main results, we use not yet treated as the control group. Implemented with improved doubly robust DiD estimator based on inverse probability of tilting and weighted least squares.

Figure B8: Monthly earnings: heterogeneity by gender

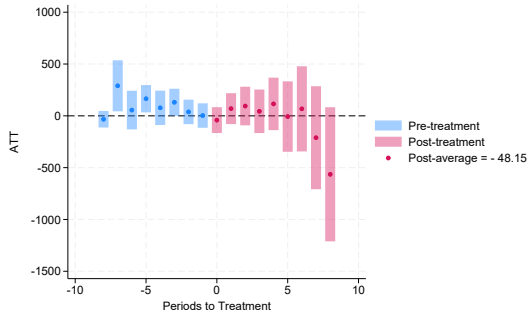
(a) Mothers with boy



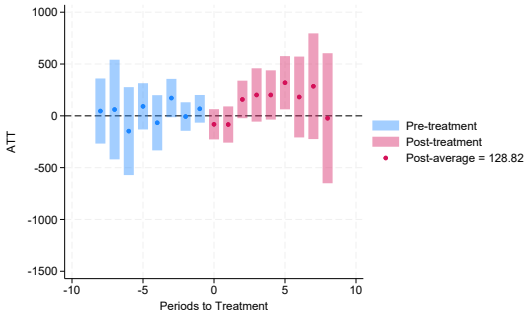

(b) Mothers with girl




(c) Fathers with boy



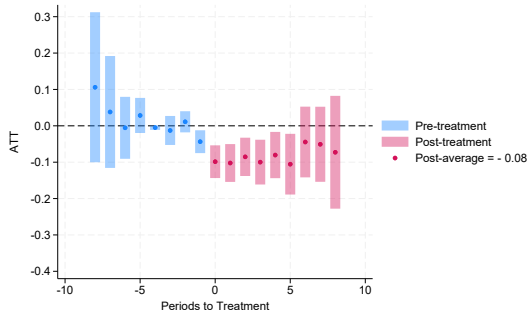

(d) Father with girl



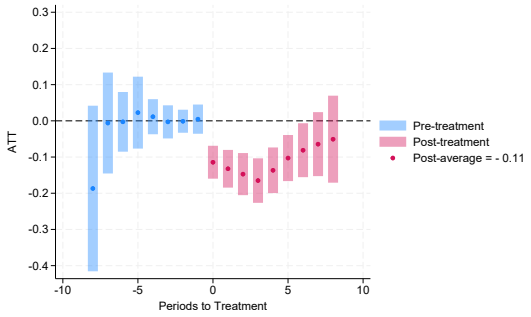

*Notes:* This graph shows the adoption of Callaway & Sant'Anna estimator using monthly earnings as an outcome. Similarly to the main results, we use not yet treated as the control group. Implemented with improved doubly robust DiD estimator based on inverse probability of tilting and weighted least squares.

Figure B9: Employment: heterogeneity by gender

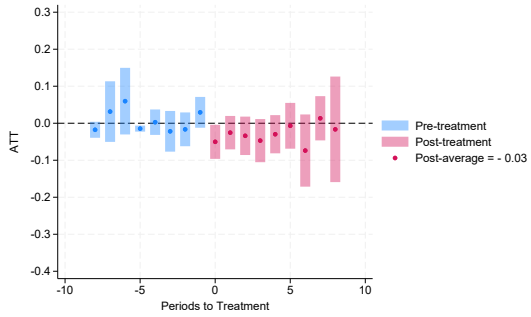
(a) Mothers with boy



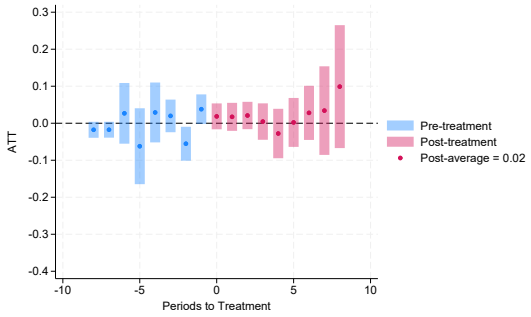

(b) Mothers with girl




(c) Fathers with boy



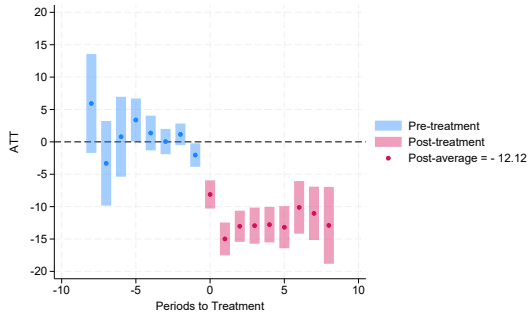

(d) Father with girl



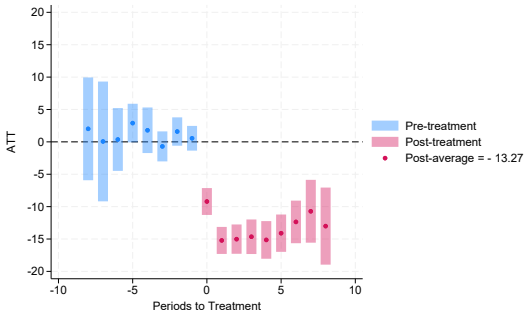

*Notes:* This graph shows the adoption of Callaway & Sant'Anna estimator using the probability to be employed as an outcome. Similarly to the main results, we use not yet treated as the control group. Implemented with improved doubly robust DiD estimator based on inverse probability of tilting and weighted least squares.

Figure B10: Total working hours: heterogeneity by gender

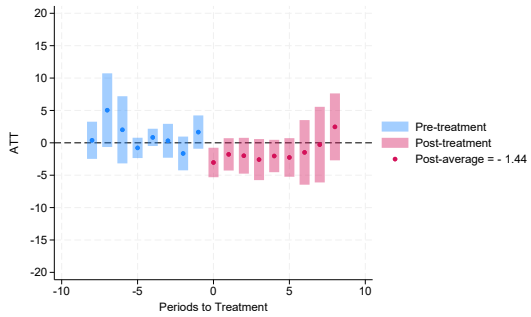
(a) Mothers with boy



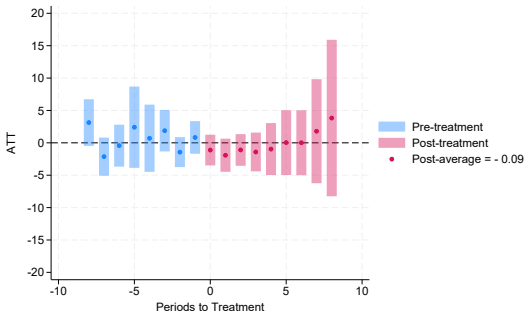

(b) Mothers with girl




(c) Fathers with boy



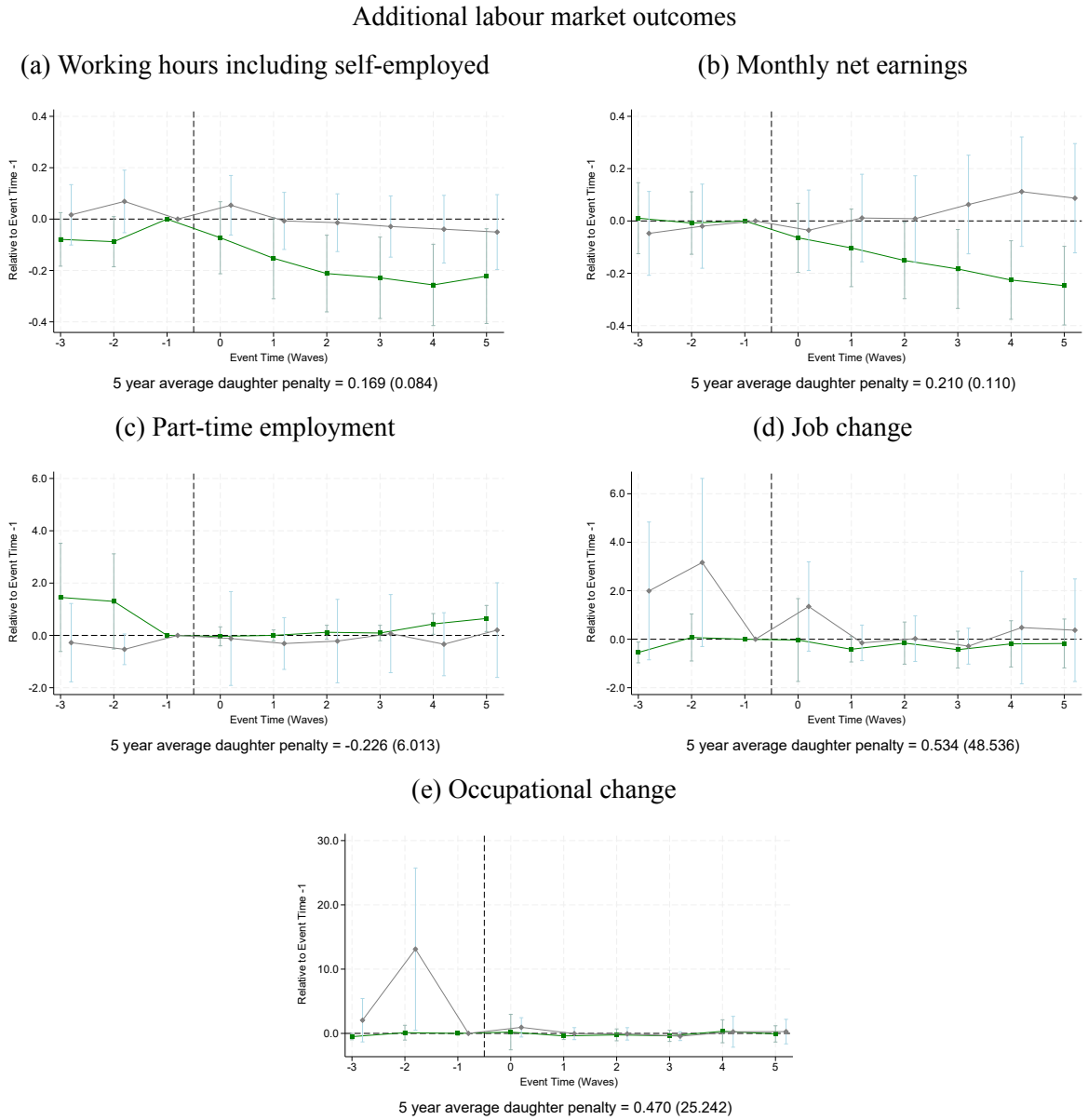

(d) Father with girl




*Notes:* This graph shows the adoption of Callaway & Sant'Anna estimator using total working hours as an outcome. Similarly to the main results, we use not yet treated as the control group. Implemented with improved doubly robust DiD estimator based on inverse probability of tilting and weighted least squares.

# C Additional Results

Figure C1: Differential responses by child gender: Father *(grey)* and Mother *(green)*

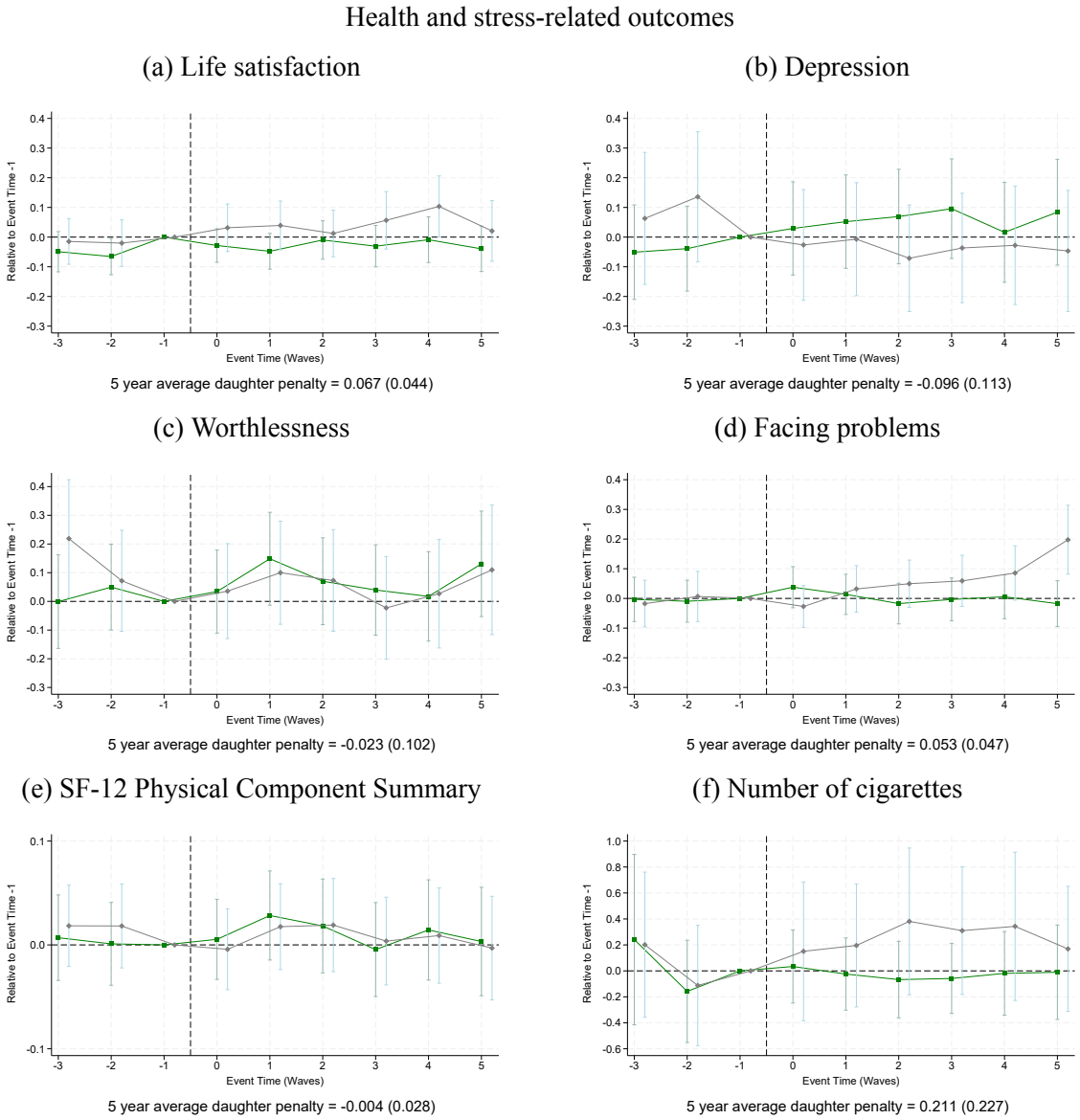



*Notes:* This figure shows the estimated impacts of parenthood $\widehat{\alpha}_t^g$ on (a) total working hours including self-employment, (b) monthly net earnings, (c) part-time employment, (e) the probability of job change, (f) the probability of occupational change before and after having the first child (in %). It focuses on mothers (green line) and fathers (grey line). We obtain the event time coefficients following Equation (3) on the analysis sample of parents. The whiskers represent 95% confidence intervals based on robust standard errors.

This Figure explores how the having a girl affects working hours including self-employment, net earnings, and employment patterns for mothers of girls compared to fathers of girls. Subfigure (a) shows that mothers with daughters work 16.9% fewer hours than respective fathers, when we include self-employed parents, supporting our main findings. Subfigure (b) reveals a similar pattern in monthly net earnings, where mothers of daughters earn 21% less than fathers. In Subfigure (c), part-time employment shows a penalty of 22.6% for mothers with daughters, though with significant uncertainty. Subfigures (d) and (e) shift the focus to job and occupational changes. In Subfigure (d), mothers of daughters are 53.4% more likely to change jobs than

fathers, and in Subfigure (e), they are 47% more likely to change occupations. However, for these last three outcomes—part-time employment, job and occupational changes—the standard errors are large, suggesting estimates on these outcomes should not be considered informative.

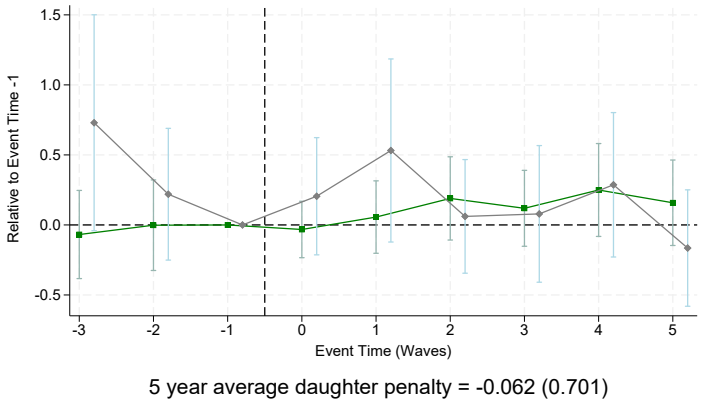
Figure C2: Differential responses by child gender: Father *(grey)* and Mother *(green)*:



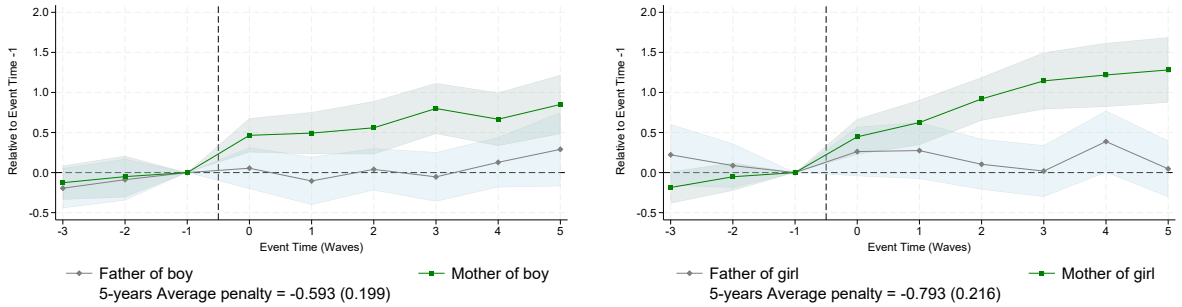

*Notes:* This figure shows the estimated impacts of parenthood $\widehat{\alpha}_t^g$ on (a) life satisfaction, (b) depression, (c) feeling worthless, (d) ability to face problems, (e) physical component summary, (f) number of cigarettes before and after having the first child (in %). It focuses on mothers (green line) and fathers (grey line). We obtain the event time coefficients following Equation (3) on analysis sample of parents. The whiskers represent 95% confidence intervals based on robust standard errors.

This Figure presents several measures of well-being, including life satisfaction, depression, feelings of worthlessness, ability to face problems, and general physical health including smoking patterns. Subfigure (a) shows a 6.7% lower life satisfaction for mothers of daughters compared to fathers. In line with this, Subfigure (b) reports higher levels of depression for mothers after the birth of a daughter. Subfigures (c) and (d) show that mothers of daughters are 2.3% more likely to feel worthless than fathers, and 5.3% less likely to be able to face up problems. Subfigure (e) analyses physical health using the SF-12 Physical Component Summary, with a minimal difference between mothers and fathers. Subfigure (f) investigates smoking behaviour, showing that fathers of daughters smoke 21.1% more than mothers.

Figure C3: Hours spent on housework in an average week




(a) Differential responses by child gender: Father *(grey)* and Mother *(green)*

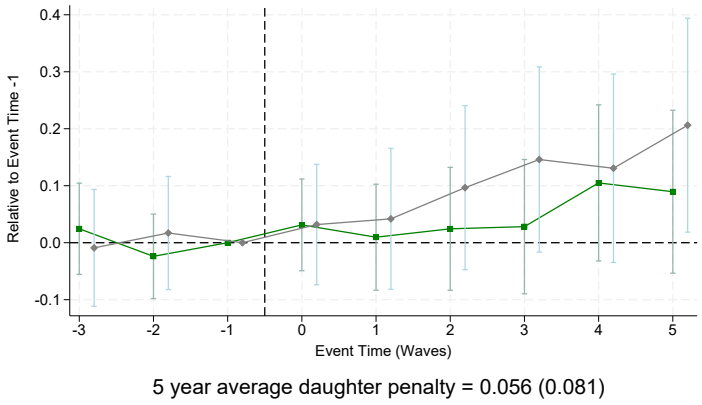



(b) The underlying penalty in time spent on housework by child gender

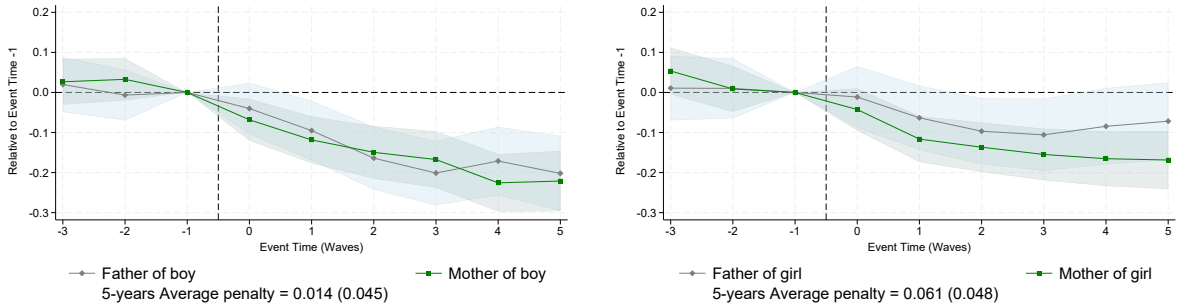
*Notes:* Refer to Notes to Figure 10. Identical results are presented, however now considering the outcome of time spent on chores per week.

This Figure reveals how the burden of housework shifts depending on whether the firstborn is a boy or a girl. In Panel (a), we show that the average "daughter penalty" in housework is around 6.2% over the five years following the child's birth, though the estimate has a high standard error, suggesting some uncertainty around the result. Panel (b) delves deeper into the time commitment to housework, comparing parents of boys and girls. For mothers and fathers of boys, the penalty over the same five-year period is estimated at 59.3%, with mothers shouldering more of the burden than fathers. When the firstborn is a girl, however, the penalty increases to 79.3%, further illustrating that mothers of girls spend significantly more time on housework compared to mothers of boys. This pattern suggests that mothers face a greater increase in domestic responsibilities, especially after the birth of a daughter.

Figure C4: Relationship happiness index




(a) Differential responses by child gender: Father *(grey)* and Mother *(green)*




(b) The underlying penalty in relationship happiness by child gender

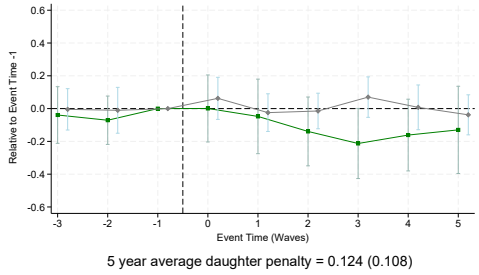
*Notes:* Refer to Notes to Figure 10. Identical results are presented, however here considering as dependent variable an index of relationship happiness

This Figure examines how relationship happiness shifts following the birth of a first child, depending on whether the child is a boy or a girl. Panel (a) shows that, on average, there is a modest "daughter penalty" in relationship happiness—around 5.6% over the five years after birth. Panel (b) breaks down this effect by child gender. For mothers of boys, the estimated penalty in relationship happiness is relatively small—about 1.4% on average—with both mothers and fathers reporting similar trends. When the firstborn is a girl, however, the penalty increases to approximately 6.1%, with mothers experiencing a steeper decline in relationship happiness compared to fathers.

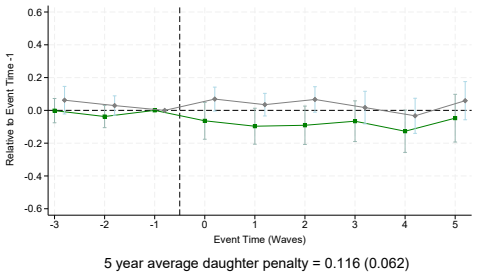
# D Heterogeneity

## D.1 Results by Parental Education Level

Figure D1: Child Penalties by Education (Employment)

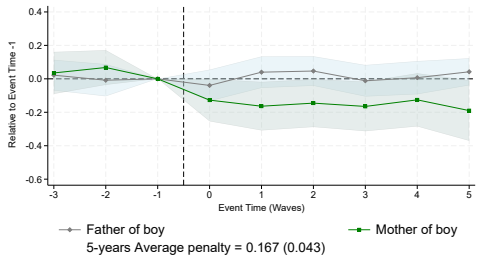



(a) Employment (Parents without a degree)

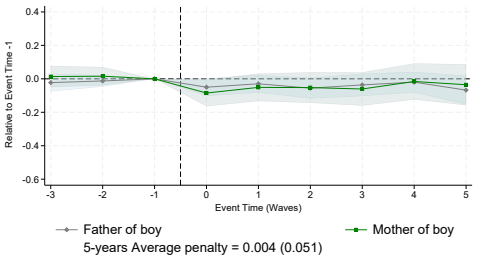



(b) Employment (Parents with a degree)

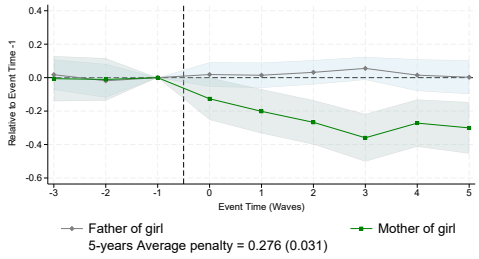
Differential responses by child gender: Father *(grey)* and Mother *(green)*



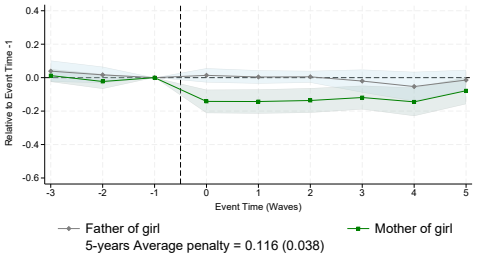

(c) Employment and boy children (Parents without a degree)




(d) Employment and boy children (Parents with a degree)




(e) Employment and girl children (Parents without a degree)

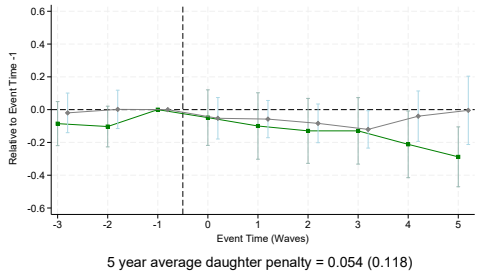



(f) Employment and girl children (Parents with a degree)

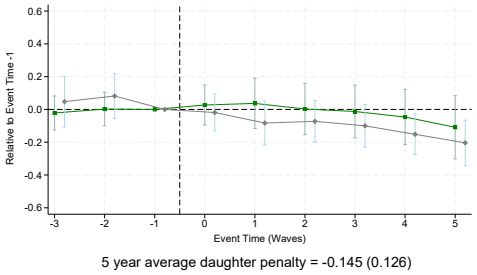
The underlying penalty in employment by child gender and whether the parent has a degree

*Notes:* Refer to Notes to Figures 2 and 3. Identical results are presented for employment, however here stratifying by parental education (individuals without a degree on left, and with a degree on right).

Figure D2: Child penalties by Education (Working hours conditional on employment)

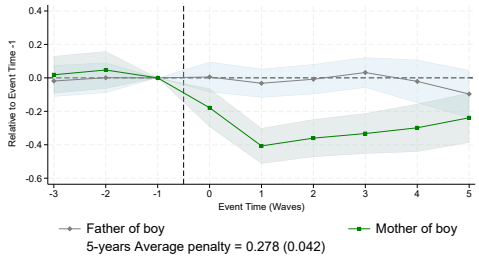



(a) Working hours (Parents without a degree)

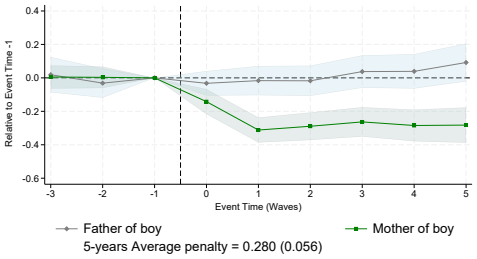



(b) Working hours (Parents with a degree)

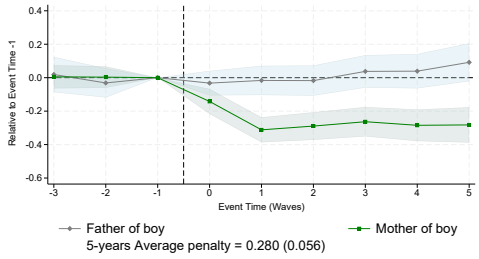
Differential responses by child gender: Father *(grey)* and Mother *(green)*



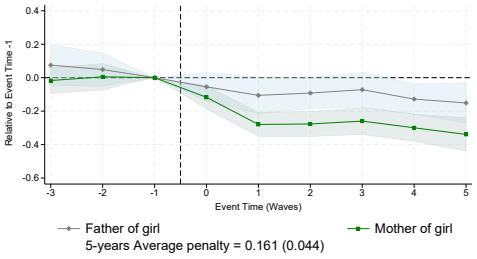

(c) Working hours and boy children (Parents without a degree)




(d) Working hours and boy children (Parents with a degree)




(e) Working hours and girl children (Parents without a degree)

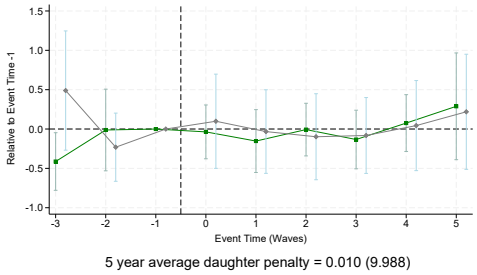



(f) Working hours and girl children (Parents with a degree)

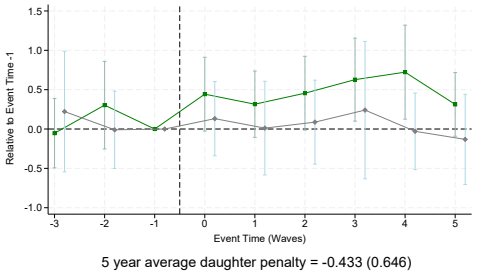
The underlying penalty in working hours by child gender and whether the parent has a degree

*Notes:* Refer to Notes to Figures 2 and 3. Identical results are presented for working hours, however here stratifying by parental education (individuals without a degree on left, and with a degree on right).

Figure D3: Fraction of chores trajectories by Education

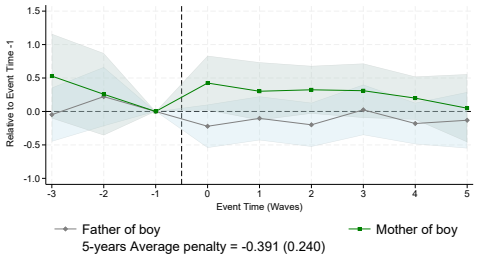



(a) Chores (Parents without a degree)

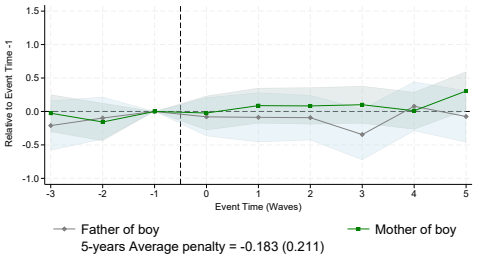



(b) Chores (Parents with a degree)

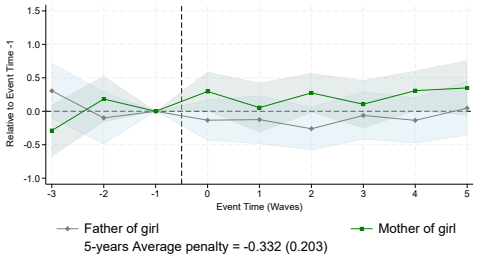
Differential responses by child gender: Father *(grey)* and Mother *(green)*



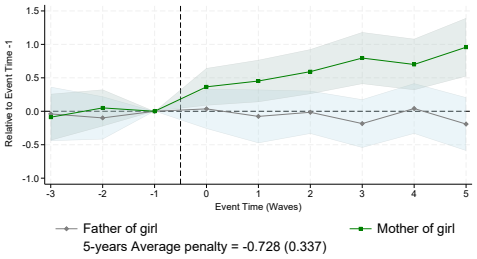

(c) Chores and boy children (Parents without a degree)




(d) Chores and boy children (Parents with a degree)




(e) Chores and girl children (Parents without a degree)

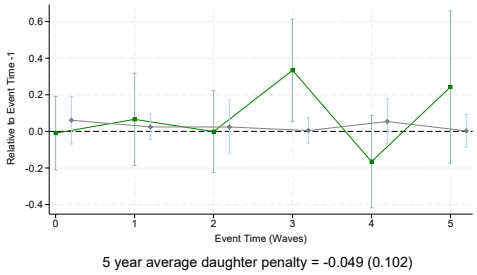



(f) Chores and girl children (Parents with a degree)

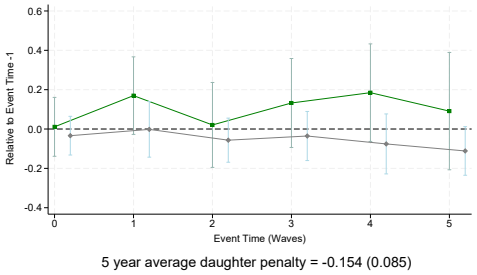
The underlying penalty in household chores by child gender and whether the parent has a degree

*Notes:* Refer to Notes to Figure 4. Identical results are presented for chores, however here stratifying by parental education (individuals without a degree on left, and with a degree on right).

Figure D4: Childcare responsibilities trajectories by Education

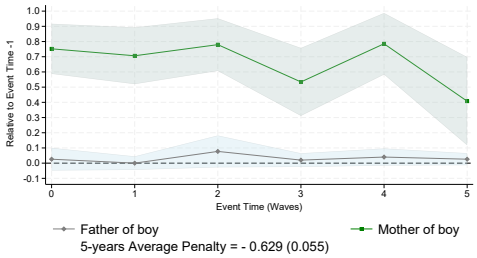



(a) Childcare (Parents without a degree)

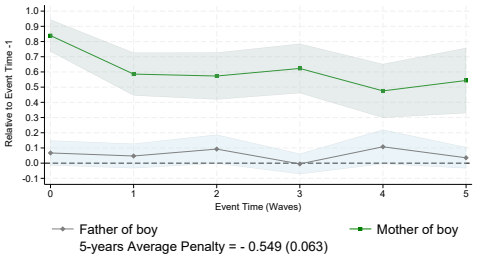



(b) Childcare (Parents with a degree)

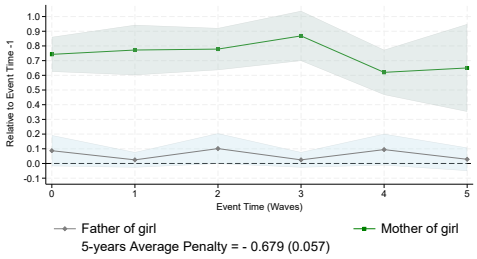
Differential responses by child gender: Father *(grey)* and Mother *(green)*



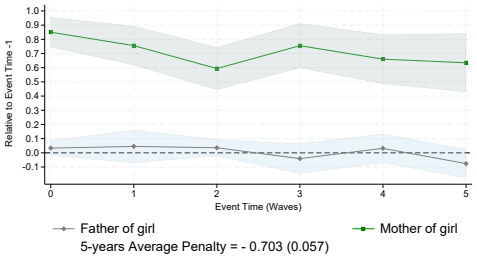

(c) Childcare and boy children (Parents without a degree)




(d) Childcare and boy children (Parents with a degree)




(e) Childcare and girl children (Parents without a degree)

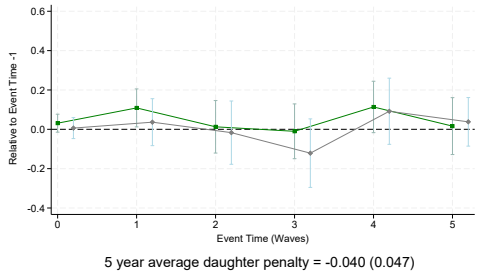



(f) Childcare and girl children (Parents with a degree)

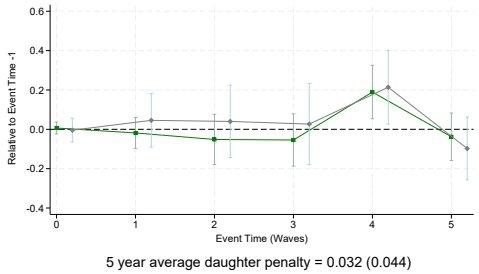
The underlying penalty in childcare responsibilities by child gender and whether the parent has a degree

*Notes:* Refer to Notes to Figure 5. Identical results are presented for childcare, however here stratifying by parental education (individuals without a degree on left, and with a degree on right).

Figure D5: Future fertility trajectories by Education

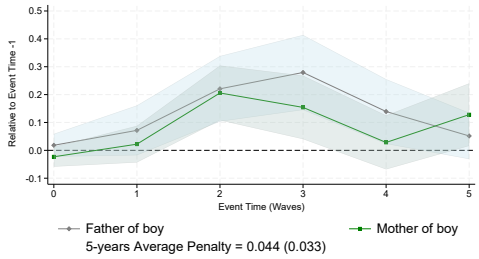



(a) Additional child (Parents without a degree)

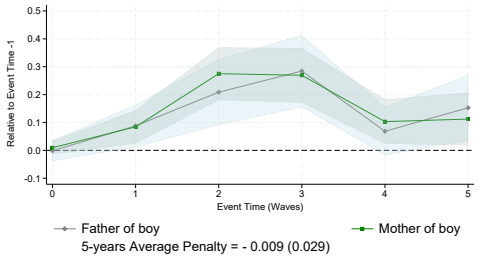



(b) Additional child (Parents with a degree)

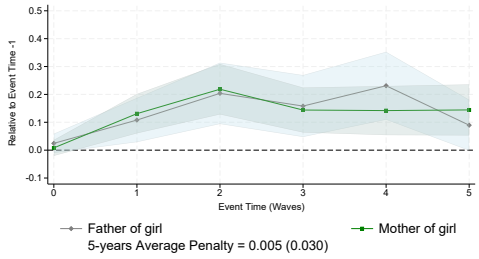
Differential responses by child gender: Father *(grey)* and Mother *(green)*



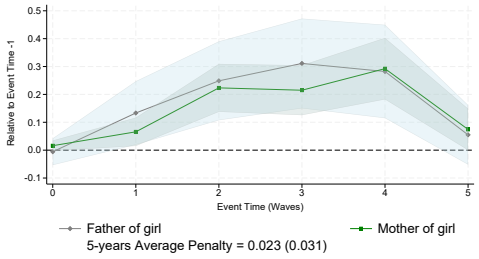

(c) Additional child and boy (Parents without a degree)




(d) Additional child and boy (Parents with a degree)




(e) Additional child and girl (Parents without a degree)

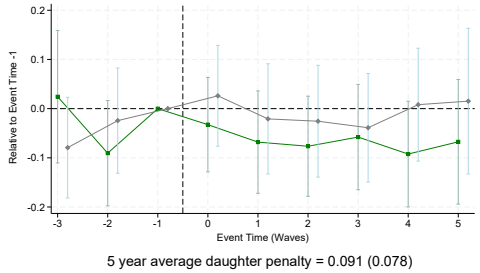



(f) Additional child and girl (Parents with a degree)

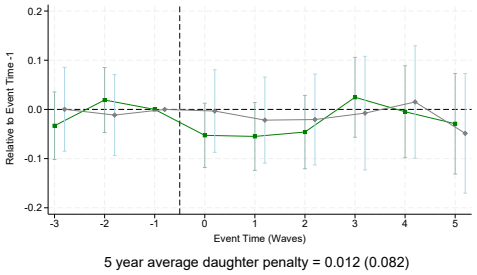
The underlying penalty in future fertility by child gender and whether the parent has a degree

*Notes:* Refer to Notes to Figure 6. Identical results are presented for future fertility, however here stratifying by parental education (individuals without a degree on left, and with a degree on right).

Figure D6: Mental health trajectories by Education

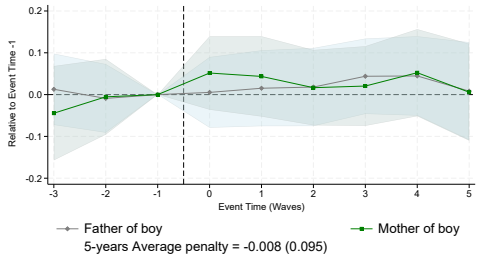



(a) Mental health (Parents without a degree)

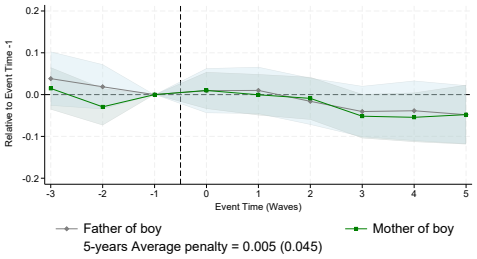



(b) Mental health (Parents with a degree)

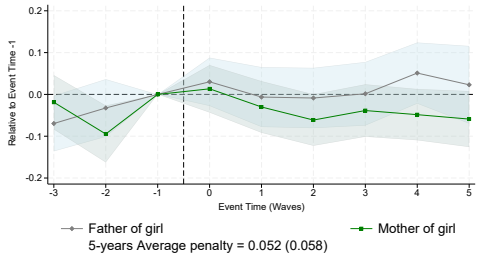
Differential responses by child gender: Father *(grey)* and Mother *(green)*



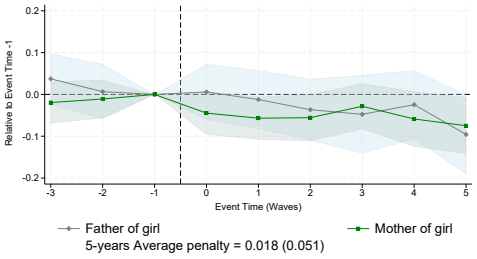

(c) Mental health and boy children (Parents without a degree)




(d) Mental health and boy children (Parents with a degree)




(e) Mental health and girl children (Parents without a degree)



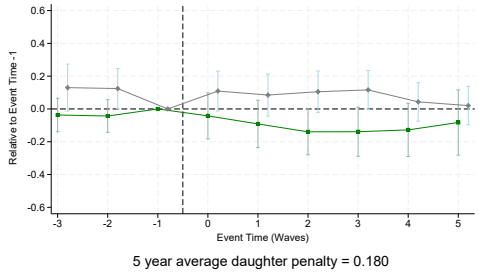

(f) Mental health and girl children (Parents with a degree)

The underlying penalty in mental health by child gender and whether the parent has a degree

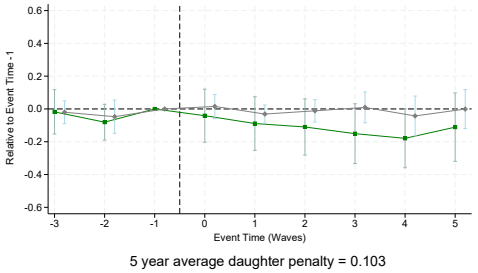
*Notes:* Refer to Notes to Figure 10. Identical results are presented for mental health, however here stratifying by parental education (individuals without a degree on left, and with a degree on right).

## D.2 Comparison between partners

Figure D7: Employment trajectories by gender equality sentiments

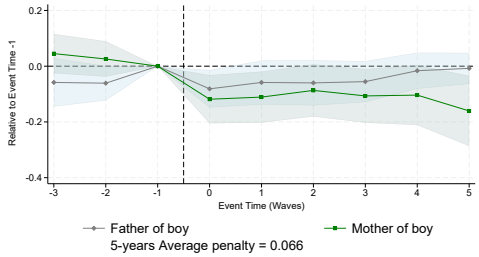



(a) Employment: Parent disagrees with the statement: "family suffers if mother works full-time" (progressive norms)

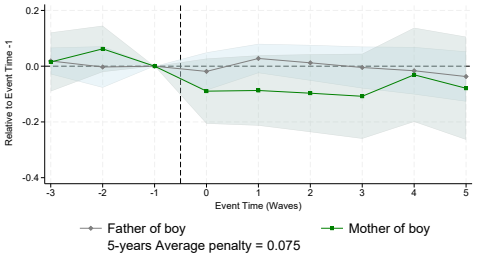



(b) Employment: Parent agrees with the statement: "family suffers if mother works full-time" (regressive norms)

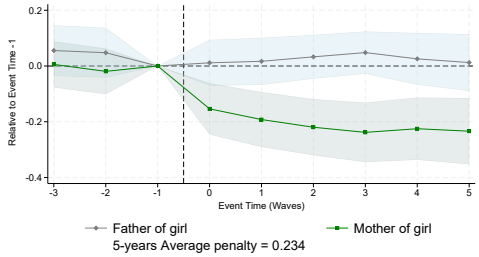
Differential responses by child gender: Father *(grey)* and Mother *(green)*



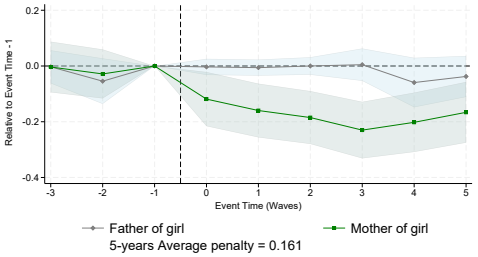

(c) Parent disagrees with the statement: "family suffers if mother works full-time"




(d) Parent agrees with the statement: "family suffers if mother works full-time"




(e) Parent disagrees with the statement: "family suffers if mother works full-time"

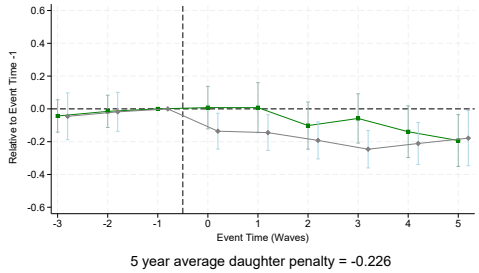



(f) Parent agrees with the statement: "family suffers if mother works full-time"

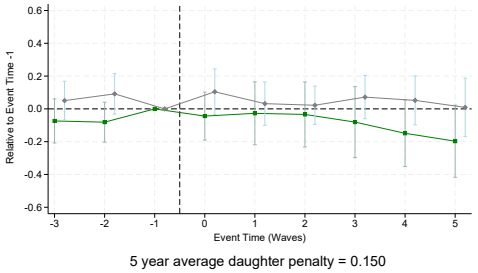
The underlying penalty in employment by child gender and gender sentiments

*Notes:* Refer to Notes to Figures 2 and 3. Identical results are presented for employment, however here stratifying by parents views on gender norms (individuals disagreeing with the statement that the "family suffers if mothers work full-time" on the left, while those agreeing with the statement on the right).

Figure D8: Working hours (conditional on employment) trajectories by gender equality sentiments

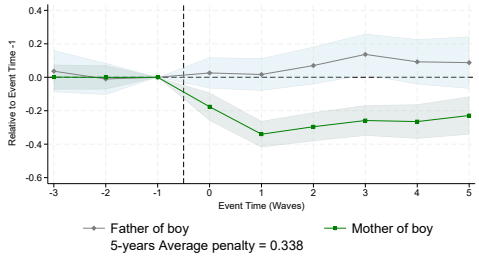



(a) Working hours: Parent disagrees with the statement: "family suffers if mother works full-time" (progressive norms)

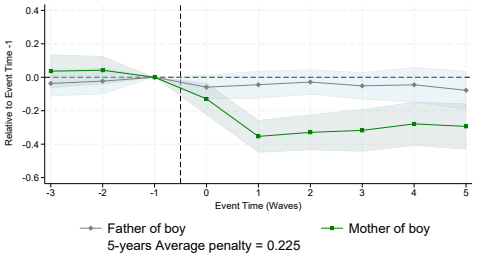



(b) Working hours: Parent agrees with the statement: "family suffers if mother works full-time" (regressive norms)

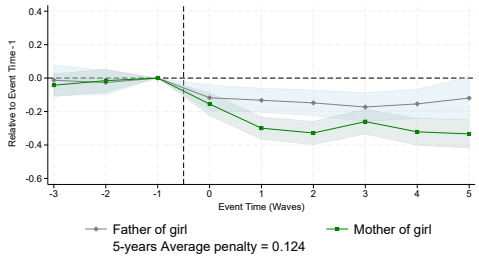
Differential responses by child gender: Father *(grey)* and Mother *(green)*



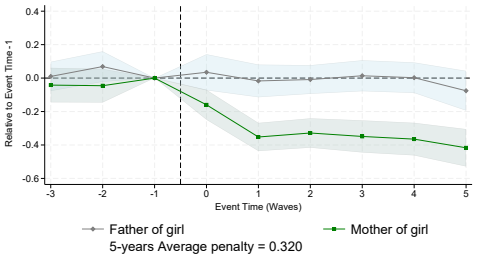

(c) Parent disagrees with the statement: "family suffers if mother works full-time"




(d) Parent agrees with the statement: "family suffers if mother works full-time"




(e) Parent disagrees with the statement: "family suffers if mother works full-time"

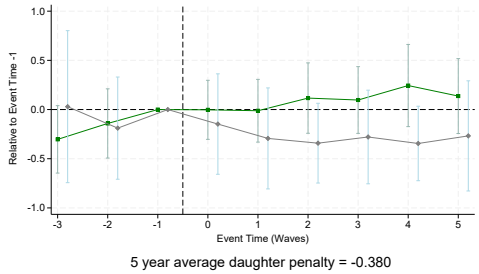



(f) Parent agrees with the statement: "family suffers if mother works full-time"

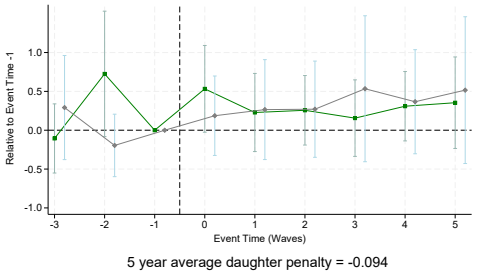
The underlying penalty in working hours by child gender and gender sentiments

*Notes:* Refer to Notes to Figures 2 and 3. Identical results are presented for working hours, however here stratifying by parents views on gender norms (individuals disagreeing with the statement that the "family suffers if mothers work full-time" on the left, while those agreeing with the statement on the right).

Figure D9: Fraction of chores trajectories by gender equality sentiments

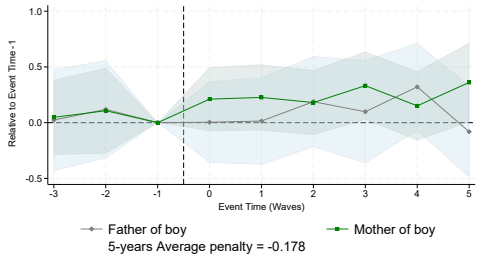



(a) Fraction of chores on self: Parent disagrees with the statement: "family suffers if mother works full-time" (progressive norms)

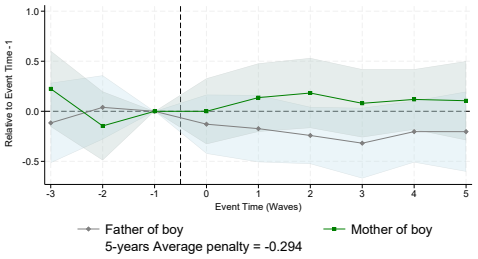



(b) Fraction of chores on self: Parent agrees with the statement: "family suffers if mother works full-time" (regressive norms)

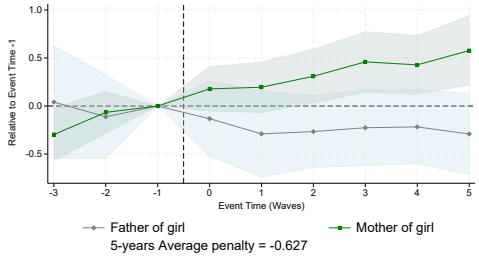
Differential responses by child gender: Father *(grey)* and Mother *(green)*



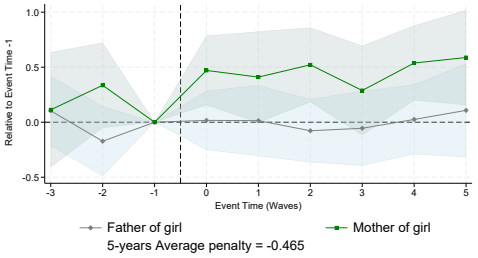

(c) Parent disagrees with the statement: "family suffers if mother works full-time"




(d) Parent agrees with the statement: "family suffers if mother works full-time"




(e) Parent disagrees with the statement: "family suffers if mother works full-time"

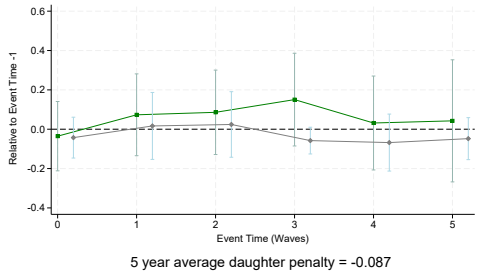



(f) Parent agrees with the statement: "family suffers if mother works full-time"

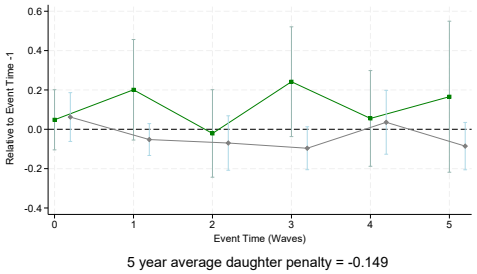
The underlying penalty in household chore division by child gender and gender sentiments

*Notes:* Refer to Notes to Figure 4. Identical results are presented for chores, however here stratifying by parents views on gender norms (individuals disagreeing with the statement that the "family suffers if mothers work full-time" on the left, while those agreeing with the statement on the right).

Figure D10: Childcare responsibilities by gender equality sentiments

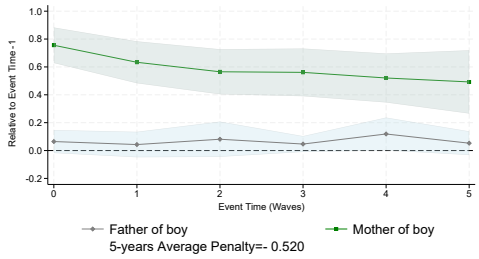



(a) Childcare on self: Parent disagrees with the statement: "family suffers if mother works full-time" (progressive norms)

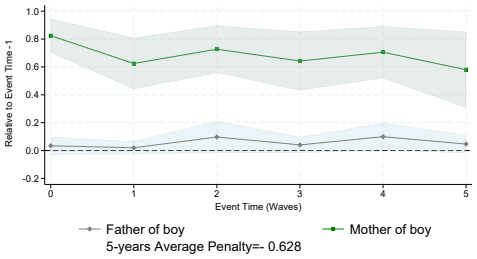



(b) Childcare on self: Parent agrees with the statement: "family suffers if mother works full-time" (regressive norms)

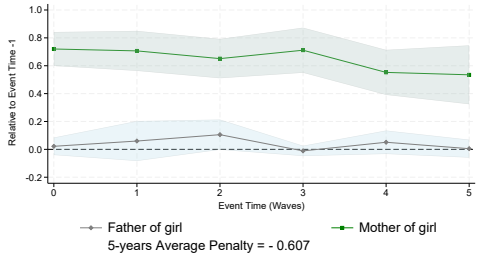
Differential responses by child gender: Father *(grey)* and Mother *(green)*



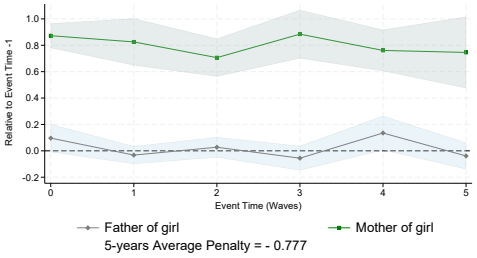

(c) Parent disagrees with the statement: "family suffers if mother works full-time"




(d) Parent agrees with the statement: "family suffers if mother works full-time"




(e) Parent disagrees with the statement: "family suffers if mother works full-time"

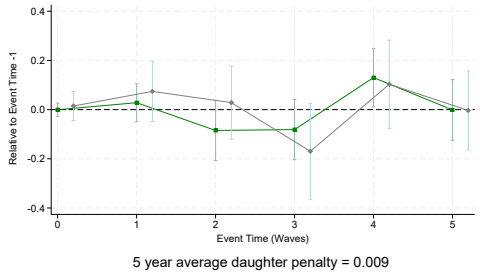



(f) Parent agrees with the statement: "family suffers if mother works full-time"

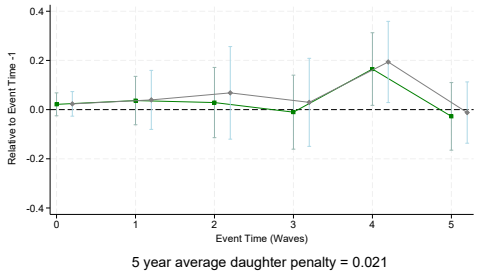
The underlying penalty in childcare provision by child gender and gender sentiments

*Notes:* Refer to Notes to Figure 5. Identical results are presented for childcare, however here stratifying by parents views on gender norms (individuals disagreeing with the statement that the "family suffers if mothers work full-time" on the left, while those agreeing with the statement on the right).

Figure D11: Future fertility trajectories by gender equality sentiments

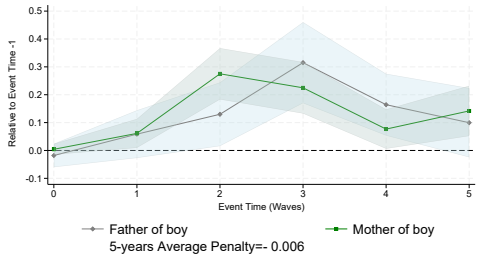



(a) Future fertility: Parent disagrees with the statement: "family suffers if mother works full-time" (progressive norms)

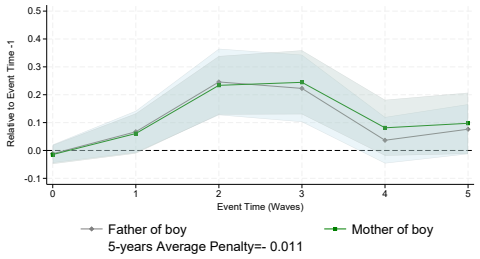



(b) Future fertility: Parent agrees with the statement: "family suffers if mother works full-time" (regressive norms)

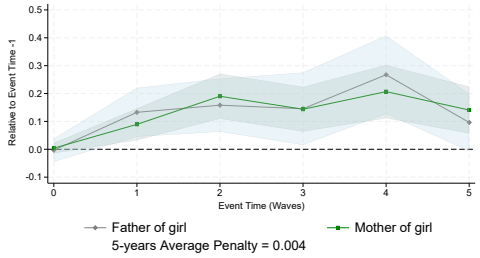
Differential responses by child gender: Father *(grey)* and Mother *(green)*



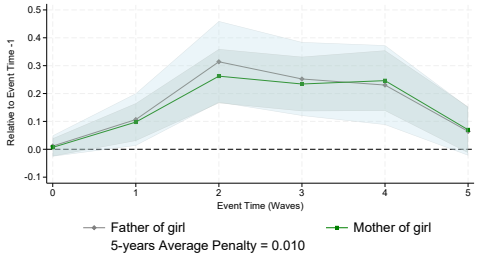

(c) Parent disagrees with the statement: "family suffers if mother works full-time"




(d) Parent agrees with the statement: "family suffers if mother works full-time"




(e) Parent disagrees with the statement: "family suffers if mother works full-time"

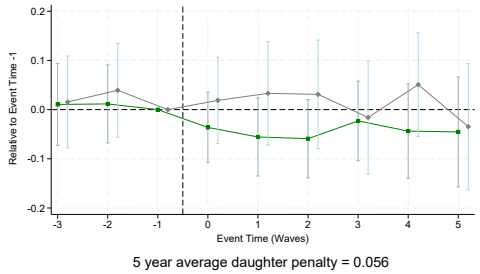



(f) Parent agrees with the statement: "family suffers if mother works full-time"

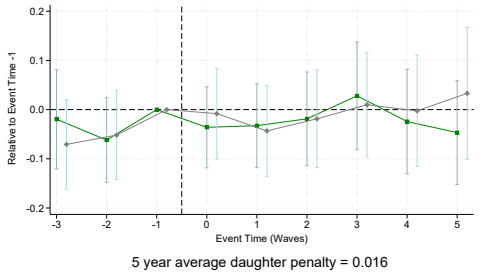
The underlying penalty in future fertility by child gender and gender sentiments

*Notes:* Refer to Notes to Figure 6. Identical results are presented for future fertility, however here stratifying by parents views on gender norms (individuals disagreeing with the statement that the "family suffers if mothers work full-time" on the left, while those agreeing with the statement on the right).

Figure D12: Mental health trajectories by gender equality sentiments

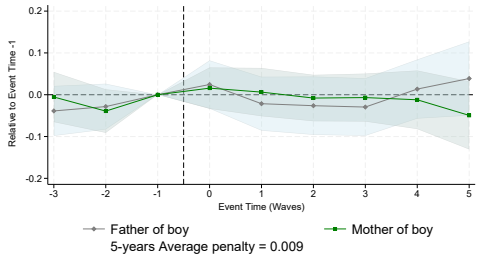



(a) Mental health: Parent disagrees with the statement: "family suffers if mother works full-time" (progressive norms)

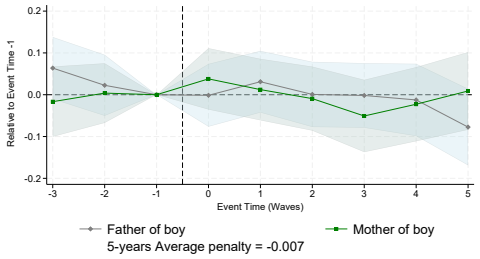



(b) Mental health: Parent agrees with the statement: "family suffers if mother works full-time" (regressive norms)

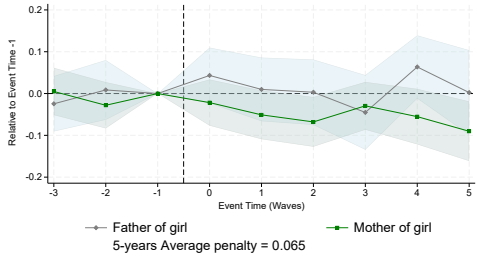
Differential responses by child gender: Father *(grey)* and Mother *(green)*



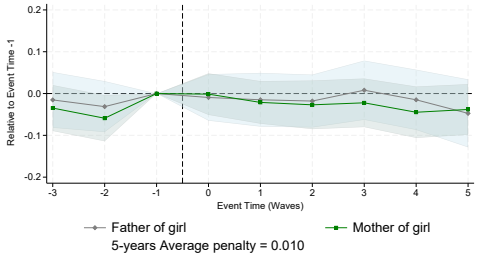

(c) Parent disagrees with the statement: "family suffers if mother works full-time"




(d) Parent agrees with the statement: "family suffers if mother works full-time"




(e) Parent disagrees with the statement: "family suffers if mother works full-time"




(f) Parent agrees with the statement: "family suffers if mother works full-time"

The underlying penalty in mental health by child gender and gender sentiments

*Notes:* Refer to Notes to Figure 10. Identical results are presented for mental health, however here stratifying by parents views on gender norms (individuals disagreeing with the statement that the "family suffers if mothers work full-time" on the left, while those agreeing with the statement on the right).