
Canadian Labour Economics Forum

*WORKING PAPER SERIES*

# Who Pays the Child Penalty? Evidence from the Panel Study of Income Dynamics

**Jamie M. Emery**
(Vancouver School of Economics)

**CLEF Award 2022 – Runner up**

**WP #43**

# Who Pays the Child Penalty? Evidence From the Panel Study of Income Dynamics*

Jamie M. Emery

University of British Columbia

Supervised by: Thomas Lemieux

February 26, 2022

**Abstract**

The impact of children on the labour market outcomes of women relative to men (the "child penalty") is well-documented, yet there is a paucity of evidence on the mechanisms behind it. I use 50 years of data from the Panel Study of Income Dynamics (PSID) to explore the importance of gendered norms and preferences and present three main findings. First, there is extensive heterogeneity in the penalty across groups of women, but not men, based on their race, marital status, and birth cohort. Second, there is a strong link between the penalty and individual-level gender-related beliefs. Third, women who grew up in households with a less traditional division of labour exhibit a smaller penalty when they have children. Taken together, my findings demonstrate that gender norms are a key driver of the penalty.

*Jamie Emery: jemery@g.harvard.edu. I would first and foremost like to thank Professor Thomas Lemieux for providing invaluable insight and guidance throughout the project. I am also extremely grateful to Professor Matt Lowe for extensive feedback and helpful conversations. Finally, I would like to thank Professor Nicole Fortin for comments on an earlier version of this paper, and whose course in labour economics was the inspiration for this project. I am grateful to the Social Science and Humanities Research Council of Canada (SSHRC) for financial support.

# 1 Introduction

Gender gaps in labour market outcomes have converged over the past century, but progress has stalled and differences in pay, employment levels, and participation within sectors persist (Olivetti and Petrongolo, 2016). Recent research has shown that the remaining gaps in high-income countries are attributed to the differential impact of children (Cortes and Pan, 2020; Kleven et al., 2019a). In particular, women experience a large, sudden, and persistent decline in earnings after having a child, while men's career trajectories are essentially unaffected. The so-called "child penalty" incurred by women is driven by reduced rates of labour force participation, hours of work, wage rates, and career progression.[1] Moreover, the labour market retreat of mothers is comparable in magnitude to a retirement effect (Kuziemko et al., 2018).

Family policies could then serve as a key tool to reduce the child penalty and the resultant inequality in the labour market. However, in practice, decades of large-scale parental leave and childcare expansions are found to have had no detectable impact on the penalty or convergence (Kleven et al., 2020). Does this null effect represent a failure in policy design, or is the child penalty simply driven by factors that are outside of the government's control? A deeper understanding of the mechanisms that drive the penalty is necessary to answer this question and inform the development of policies aimed to reduce it.

In this paper, I use 50 years of data from the U.S. Panel Study of Income Dynamics (PSID) and an event study design to explore the importance of two possible mechanisms that the government may have little scope in influencing: gendered social norms and preferences. I explore these channels by examining heterogeneity in the penalty across groups of women who may have differ in either respect. I focus on differences by demographic characteristics including race, marital status, household income, educational attainment, and birth cohort. I also use data from the PSID's Child Development Supplement (CDS) to investigate heterogeneity based on individual-level agreement with statements related to gender role attitudes and parenthood. Finally, I use the intergenerational component of the survey to examine the relationship between a mother's post-birth labour market trajectory and the relative division of labour in the household she grew up in.

Many new studies have explored heterogeneity in the penalty as the direction of the literature shifts from *documenting* the child penalty to *understanding* the mechanisms behind it. For example, Kleven et al. (2021) compare biological and adoptive mothers and find that they incur the same persistent penalty which suggests that biological aspects of motherhood are not important drivers.[2] Comparative advantage is also unlikely to rationalize the penalty as women with higher or lower earnings potential than their spouse pre-birth display the same post-birth labour market detachment (Cortes and Pan, 2020; Kleven et al., 2021).

The most likely explanation is then gendered preferences or norms. These channels are distinct yet difficult to disentangle, which leaves the welfare implications of the penalty somewhat ambiguous. On the preference front, women may desire a greater work-life balance after having children and the penalty is a reflection of that choice. Alternatively, women's labour market decisions could be influenced by traditional gender norms which prescribe that men should serve as the family's primary breadwinner and women should stay in the home. These norms become more salient during parenthood as a career bolsters the notion of a "good father" and interferes with that of a "good mother". Working then violates the gender identity

[1] Mothers fall behind in occupational rank as they are less likely to stay in jobs with long hours and greater career advancement possibilities, and more likely to transition to family-friendly firms with greater flexibility. E.g. See Bertrand et al. (2010); Casarico and Lattanzio (2021); Kleven et al. (2019a).

[2] Consistent with this, Andresen and Nix (2019) find no long-run differences in the penalty between parents in lesbian partnerships in which only one of the women is biologically linked to the child.

of women but not men, and mothers may detach from the labour market accordingly (Akerlof and Kranton, 2000; Bénabou and Tirole, 2011).

I contribute to this discussion with three main sets of findings, respectively covering (i) the overall child penalty, (ii) heterogeneity by demographic characteristics, and (iii) heterogeneity by self-reported gender-related attitudes and the relative labour supply of the grandparents.

First, I replicate the overall child penalty on the extensive (labour force participation) and intensive (logarithm of hours worked) margins in the PSID. I find that women's participation falls by 30 percentage points in the first five years of motherhood and is still 26 percentage points below pre-child participation rates in the 6-15 years post-birth.[3] In contrast, men's participation rates are essentially unaffected by children. The same qualitative pattern emerges on the intensive margin.

Second, I find stark differences in the penalty across groups of women, but little to no heterogeneity across groups of men. To begin, I look at differences by race and find that Black women retreat from the labour market at half the rate of white women post-birth. There are numerous child-penalty relevant differences in how these groups of women experience the world,[4] including the fact that Black women in the sample are much more likely to be single mothers. Accordingly, I next consider differences by marital status and find that never-married women have a penalty which is roughly two-fifths the size of married women. Never-married women may work substantially more post-child because (a) they do not have the non-labour income necessary to exercise the choice to leave the labour market (preferences) or (b) they are less pressured to embrace the traditional role of homemaker (norms).

To test (a) directly I look at differences by household income. I find that women in the first and fourth income quartiles do not have statistically different penalties which suggests (b) might be more important. One way to explore (b) indirectly is to compare women with and without a college degree. College educated women may feel less constrained by traditional gender norms as they spent time immersed in progressive environments and are perhaps more likely to marry a spouse who supports their career ambitions. However, I find women with and without a college degree incur the same penalty. This is surprising given the first group invested heavily in their human capital, but it could be that women's educational attainment has progressed quicker than the societal norms they are surrounded by. Thus, women might have invested in their education under the incorrect belief that they can one day "have it all"– a family and a career (Bertrand, 2013; Goldin, 2021; Kuziemko et al., 2018).

As norms gradually shift (Fortin, 2005) and women have greater access to things like higher education (Goldin et al., 2006) and contraceptives (Bailey, 2006; Goldin and Katz, 2002), the penalty should then improve across birth cohorts. I find this is true as women born in the 1970s and 80s have a penalty which is substantially smaller and less persistent than their counterparts born in the 50s. This prompts the question of whether the penalty has shifted to men. If women in younger cohorts are less likely to embrace the role of homemaker, are men any less likely to abide by their prescribed role as breadwinners? I find no evidence of this across birth cohorts of fathers. In fact, while I document large heterogeneity across women, the post-birth labour market trajectories of men are essentially unaffected across all groups I examine. This suggests that gender norms around parenthood might be sticky for women and stuck for men.

Third, I report on more direct tests of the norm and preference channels. I first look at differences in the penalty based on a mother's agreement to statements that assert (i) women are happier staying home with their children (to test preference), (ii) it is better if a father earns and mother stays home (to test norms), and (iii) mothers should not work when their child is

[3] These estimates are very similar in magnitude to Kuziemko et al. (2018) who use the PSID to document a 27 percentage point decrease in participation in the 10 years post-birth.

[4] Thus, highlighting the importance of an intersectional approach to studying the child penalty.

young (to test "mother's guilt"). I find that agreement with all three statements is associated with a larger decrease in hours worked in the short-run; however, only (iii) impacts participation rates and has a persistent effect on hours worked. This suggests that the child penalty is likely driven by a woman's inner conflict between her competing identities as a mother and career women. Finally, I estimate the intergenerational transmission of the penalty and find women who grew up in households with a less traditional division of labour have a smaller penalty when they have children.

Taken together, my findings demonstrate that gender norms are a key driver of the child penalty. The penalty is then consistent with the predictions of Akerlof and Kranton (2000), as gendered expectations around parenthood challenge the identities of women but not men. Accordingly, parenthood causes women to reduce their labour supply[5] as deviation decreases their utility in the form of mother's guilt.

This study contributes to the child penalty literature in three ways: (i) I explore heterogeneity across demographic characteristics that were previously unexplored; (ii) I link the penalty to individual-level responses to gender-related questions rather than country (Kleven et al., 2019b) or region-level (Casarico and Lattanzio, 2021) average responses; and (iii) I document the intergenerational transmission of the penalty in the United States, as done for Denmark (Kleven et al., 2019a).[6] In doing so, I also contribute mother-specific findings to studies that examine differences in participation across all women by race (Fortin, 2015), educational attainment (Bertrand et al., 2010; Kuziemko et al., 2018), birth cohorts (Fortin, 2019; Goldin, 2006; Goldin and Mitchell, 2017), intergenerational labour supply (Farré and Vella, 2013; Fernández et al., 2004), and gender-related attitudes (Fortin, 2005, 2015). Finally, my work can also be placed amongst studies highlighting the importance of gender identity for labour supply decisions (Bertrand, 2011; Fortin, 2015).

The remainder of the paper is organized as follows. Section 2 describes the data, Section 3 introduces the event study methodology and estimates the impacts of children, Section 4 presents my analysis of heterogeneity based on demographic characteristics and gender-related attitudes, Section 5 analyzes intergenerational transmission, and Section 6 concludes.

## 2 Data

In this section I provide a brief summary of the data, describe how I construct the analysis sample, and present summary statistics. Appendix A.3 provides a more in-depth description of the variables used throughout the analysis.

### 2.1 The Panel Study of Income Dynamics (PSID)

Estimating the child penalty requires individual-level panel data with detailed information on labour market outcomes and birth histories. The U.S. Panel Study of Income Dynamics (PSID) is a publicly available data source with such advantages. Beginning in 1968, the PSID is the longest running longitudinal household survey in the world. Respondents enter the panel from three sources: As children or grandchildren of the original 1968 sample, the 1997 and 2017 immigrant refresher sample, and marriages and births/adoptions into existing PSID families. The survey was conducted annually until 1997 and biennially thereafter. A computer-assisted

[5] On the labour demand side, employers may also internalize these norms and view a new father as dependable and committed to their work and a mother in the opposite way.

[6] My work is most closely related to Kuziemko et al. (2018) who also examine heterogeneity in the penalty using the PSID. I build on their work through the inclusion additional demographic groups (race, marital status, cohorts, and men), the intensive margin, and data from the CDS in my analysis.

telephone interview (CATI) is used to collect detailed information about individuals designated as reference persons and their spouse, as well as basic information on other members of the family unit.

The intergenerational component of the PSID allows me to connect adults in the sample to the household that they grew up in. I exploit this feature of the survey to estimate the intergenerational transmission of the child penalty. In particular, I link individuals with children, for whom I estimate the penalty for, to the work history and household labour division of their parents.

Other studies have used the PSID to study child penalties in the American context.[7] I extend on their work by utilizing the intergenerational component of the survey and integrating data from the Child Development Supplement (CDS) into the analysis.

### 2.2 The Child Development Supplement (CDS)

In 1997, the first wave of the CDS was introduced to supplement the main data collection with information on children aged 0 to 12 and their parents. Importantly for my analysis, the CDS asks children's caregivers a variety of questions related to parenthood, gender role attitudes, and work values. The initial cohort of CDS respondents was interviewed again in 2002 and 2007, and a new cohort completed the supplement in 2014.

My analysis includes mothers[8] from the 1997 and 2014 waves of the CDS. I do not include follow-up responses from the first cohort of respondents in the 2002 and 2007 waves because I am interested in differences in child penalties across individuals conditional on a single response to the questions of interest. I do not examine changes in attitudes over time, as explored by Kuziemko et al. (2018).

### 2.3 Sample Selection and Summary Statistics

**Sample Selection.** I estimate the child penalty using the 1968-2017 waves of the PSID. The sample includes all individuals who have a birth record for their first child during those years and can be followed over time. Ideally, I would impose a requirement that parents must be observed each year over the 15-year event study window. This was possible with the Danish administrative data used by Kleven et al. (2019a), but is too restrictive in the context of the survey-based American data. Instead, I restrict the sample to include individuals I observe at least once before and after the birth of their first child and at least eight times total over the event study window. I further require respondents to be over the age of 18 when they are observed for the first time in the data. Summary statistics for the corresponding sample are reported in Table 1.

**Descriptive Statistics.** As reported in Table 1, the sample restrictions described above leaves a total of 4,024 women and 3,810 men in the estimation sample. Mothers in the sample are less likely to participate in the labour force, work fewer hours, and are more likely to stay at home compared to men. Women also have their first child at a younger age (Figure A1) and are less likely to be married. Mothers and fathers have similar racial composition, educational attainment, and distribution across birth cohorts. I observe both men and women 16 times on average throughout the 21 year event window, and observe 20% of parents every year.

[7] For example, see Cortes and Pan (2020); Kleven et al. (2019b); Kuziemko et al. (2018). I compare my estimates of the child penalty to the estimates obtained in other studies that use the PSID in Section 3.2.

[8] Primary caregivers (PCGs) in the CDS are almost always women. PCGs are defined in order of precedence as the biological, step, foster, or adoptive mother of the child, the wife of a PSID reference person who is the child's father, followed by men in the same order. Due to small sample size, I do not include responses from male PCGs in the analysis.

# 3 The Impact of Parenthood on Labour Market Outcomes

## 3.1 Estimating Equation and Identification Assumptions

Following Kleven et al. (2019a), I examine changes in labour market outcomes associated with parenthood separately for men and women with an event study approach based around first childbirth. I motivate the use of an event study design in this context and compare it to alternative approaches used in the literature in Appendix A.4. I model a given labour market outcome $y$ for person $i$ in year $t$ with the following specification:

$$y_{it} = \sum_{\tau=-5,\neq-1}^{15} \beta_\tau \cdot \mathbb{1}[\tau = t - c^i] + \sum_a \gamma_a \cdot \mathbb{1}[a = \text{Age}_{it}] + \delta_t + \varepsilon_{it} \tag{1}$$

The labour market outcomes of interest, $y_{it}$, include labour force participation and the logarithm of total annual hours worked to capture the extensive and intensive margin responses respectively. In alternate specifications, I examine the probability that the individual reports their employment status as "homemaker". I discuss how each outcome variable is defined in Appendix A.3. $c^i$ is the year in which person $i$ had their first child. The event time indicator $[\tau = t - c^i]$ then represents person $i$ in year $t$ having had their first child $\tau$ years ago. I omit the year before first birth (i.e. $\tau = -1$), which implies that the event time coefficients ($\beta_\tau$) measure the change in the outcome of interest relative to the year before first childbirth.[9]

The specification includes age and year fixed effects to control non-parametrically for life-cycle and time trends. Estimates of $\beta_\tau$ are then the effect of parenthood on labour market outcomes relative to a no-child counterfactual predicted by age and year. The effects of all three sets of dummies are identified because, conditional on age and year, there is individual-level variation in the timing of first births. Appendix figure A1 depicts this key source of variation as seen in data from the PSID. Finally, standard errors are clustered at the person-level.

I present the event study results graphically by plotting the estimated values of $\beta_\tau$.[10] The child penalty is then based off a comparison of the event time coefficients from equation (1) run seperately for women and men.

The key assumptions under which $\beta_\tau$ identifies the effect of parenthood $\tau$ years after first childbirth, and the extent to which we should believe them, are laid out in Kleven et al. (2019a). To summarize: A parent's labour market trajectory must evolve *smoothly* in the absence of children. The idea is that any sharp changes in the labour market outcome of interest following the arrival of a first child are arguably orthogonal to other unobserved determinants of the outcome which should evolve smoothly over $\tau$=0. While fundamentally untestable, the presence of parallel pre-trends lends support to this assumption.

Parallel pre-trends become less informative each year post birth. Thus, smoothness is necessary but no longer sufficient to identify the long-run impact of the first child. First, an individual may have more children. The long-run coefficients then capture the effect of *total* fertility rather than just the impact of the first child. Second, there are likely other changes to non-child related labour market outcome determinants which occur over a long post-birth event window. Inclusion of the non-parametric age and year controls in (1) alleviate this concern but do not rule out the possibility of bias. To that aim, the event study estimates in the years immediately following childbirth more plausibly hold a causal interpretation. However,

[9] Kuziemko et al. (2018) normalize the event time $\tau = -2$ to zero to avoid labour market adjustments associated with pregnancy. I include this as a robustness check and find no substantive differences in the results.

[10] To summarize the results across dependent variables more concisely, I include a table which reports the results from a pooled version of Equation (1). In this specification, I pool the post birth event time dummies into two variables: and indicator for 0 to 5 and 6 to 15 years post birth.

the documented longer-term impacts of children are still illuminating and Kleven et al. (2019a) provide two identification checks which support the validity of the event study approach throughout the entire window studied.[11]

### 3.2 Results

Figure 1 depicts the impact of parenthood for women and men in the Panel Study of Income Dynamics (PSID). Panel (a) plots labour force participation of men and women with and without children over the life cycle. The figure shows differences in base level participation rates and trends across groups.

Participation rates of women with and without children are very different throughout their prime working ages in terms of both levels and evolution.[12] Mothers exhibit far lower labour market attachment in the years that they raise their children. As their children age they gradually close the gap, and by their late 40s both groups of women participate at similar rates. In contrast, father's trajectories closely mirror that of men without children and they participate at much higher rates than both groups of women, particularly during the years of family formation.[13] Consistent with the findings of Goldin et al. (2021), Figure 1 highlights that women with children pay a "motherhood penalty" (when compared to women without children) and a "parenthood penalty" (when compared to men with children).

The descriptive findings presented above reflect sample selection in addition to the actual impact of children on participation rates. To isolate the latter, I estimate Equation (1) and plot the event time coefficients for men and women separately in Figure 1 panel (b). The well-documented pattern from the child penalty literature emerges: after adjusting for lifecycle and time trends, participation rates of mothers and fathers evolve in tandem until first childbirth, at which point the mother's participation diverges and never converges thereafter.

Within the sample, women experience a large and immediate 23 percentage point decline in labour force participation relative to the year before the arrival of their first child (at which point the pre-child participation rate was 88 percent). Table A1 reports pooled event time coefficients from the 0 to 5 and 6 to 10 years following first birth. Women experience an average labour force participation decline of 30 percentage points throughout the first 5 years of motherhood. This labour market detachment persists even once children reach school age. Women are still 26 percentage points less likely to be in the labour force 6-15 years post-birth relative to the year before they gave birth to their first child. However, recall from section 3.1 that this longer-run estimate reflects the impact of *total* fertility, not just the impact of the first child.[14] In contrast, father's participation rates are essentially unaffected by children in the short and long-run. If anything, men see a small but statistically significant increase in participation after parenthood.

These estimates are similar in magnitude to other participation rate penalties estimated for the United States using the PSID. Kuziemko et al. (2018) find an average decline of 27 percentage points in the 10 years post birth, and Kleven et al. (2019b) document a 40%[15] decline over the same period.

[11] The first extends the event study framework to include a formal control group of individuals who never have children. The corresponding estimates are closely aligned with the estimates from (1). Second, they compare the LATE from the sex-mix IV to a local effect of having a third child from the event study. The two sets of estimates are nearly perfectly aligned.

[12] The latter motivates why I do not use women who never have children as a control group. I am skeptical the parallel trends assumption would hold as the groups of women appear to be on different trajectories in the years before the average age mothers in the sample has their first child.

[13] Women without children may work less than fathers for a host of reasons including other care responsibilities (E.g. of a disabled or elderly family member (Truskinovsky and Maestas, 2018)).

[14] Women in the sample have two children on average.

[15] For comparison, my estimate represents a 34% decline in participation relative to the year before first childbirth.

Figure A2 and Table A1 present the same set of results for log hours worked. I do not discuss the intensive margin response in detail as the findings are similar to the ones presented above.[16] Finally, as a robustness check, the final column of Table A1 presents the results with the probability the respondent reports that they stay at home as the outcome of interest. The results are essentially the reciprocal of the first column as women that left the labour force transitioned to staying at home.

# 4 Heterogeneity in the Child Penalty

In this section I examine heterogeneity in the child penalty across groups of women who may have differ in their preferences (or ability to exercise said preference) and in the extent to which they are pressured to embrace the traditional role of homemaker.

## 4.1 Estimating Equations

**Heterogeneity by demographic characteristics.** I examine heterogeneity in the post-birth labour market trajectories of mothers and fathers based on membership to different groups. I modify a model proposed by Kuziemko et al. (2018) to allow for comparison across multiple groups, rather than a single group relative to another, and estimate the following equation:

$$\begin{aligned} y_{it} &= \beta \text{Post}_t + \sum_a \gamma_a \cdot \mathbb{1}[a = \text{Age}_{it}] + \delta_t \\ &+ \mathbb{1}[G = g] \cdot \left[ \beta^{G=g} \text{Post}_t + \sum_a \gamma_a^{G=g} \cdot \mathbb{1}[a = \text{Age}_{it}] + \delta_t^{G=g} \right] + \varepsilon_{it}^{G=g} \end{aligned} \tag{2}$$

Where $G$ is a category of demographic characteristics and $g$ is a dummy variable for a distinct and mutually exclusive group within each category. I omit one subgroup within $G$ such that the estimates for all other groups $g$ are relative to that group.[17] For example, in my analysis by race, $g$ is a dummy variable for whether the repondent is Black and all coefficients with the superscript $G$ are relative to white parents. The other groups I consider are based on marital status (previously married and never married relative to married parents), household income (second, third and fourth quartiles relative to first), educational attainment (parents with college degrees relative to those without), and birth cohort (1960, 1970, and 1980 birth cohorts relative to parents born in 1950).

This equation differs from the event study specification presented in equation (1) in two key ways. First, the event time variables are pooled into one dummy variable which equals one in the 0 to 15 years following first birth and zero in the 5 years prior. Second, it fully interacts a dummy variable for group $g$ with the event dummy, and age and year fixed effects. This allows for individuals who belong to a given group to have unrestricted differences across these dimensions.

**Heterogeneity by gender role attitudes.** Sample size is an issue when examining heterogeneity in the child penalty by gender role attitudes as only a subsample of women are interviewed in the Child Development Supplement (CDS). Accordingly, I remove the sample restrictions

[16] One key difference to note: the coefficient associated with year of first birth is attenuated due to the use of a calendar-year measure of hours worked. This means that a woman who worked zero hours following childbirth may have positive hours recorded for this variable from working earlier in the year.

[17] The effect of children on the labour market outcomes is given by $\beta$ for parents in the omitted group and $\beta + \beta^{G=g}$ for parents in group $g$.

described in section 2.3 and include all mothers interviewed in the 1997 and 2014 waves of the CDS that have a birth record for their first born child in the estimation sample. I then pool the post-birth event-time coefficients from equation (1) into two periods: 0 to 5 years and 6 to 15 years after first birth. The former captures the average short-run child penalty before the child reaches school age and the latter captures the average long-run penalty once the child enters school.

To test for the importance of preferences and gender role attitudes I estimate an equation similar to (2). The only difference is in how I split up the event time[18] and replace the dummy variable indicating membership to group $g$ with a dummy variable for whether mother $i$ agrees with statement $s$:

$$\begin{aligned} y_{it} = \; & \beta_1 \, 0 \text{ to } 5 \text{ years post}_t + \beta_2 \, 6 \text{ to } 15 \text{ years post}_t \\ & + \beta_3 \, (0 \text{ to } 5 \text{ years post}_t \times \text{Agree}_i^s) + \beta_4 \, (6 \text{ to } 15 \text{ years post}_t \times \text{Agree}_i^s) \\ & + \sum_a \gamma_a \cdot \mathbb{1}[a = \text{Age}_{it}] + \delta_t + \text{Agree}_i^s \cdot \left[ \sum_a \gamma_a^s \cdot \mathbb{1}[a = \text{Age}_{it}] + \delta_t^s \right] + u_{it}^s \end{aligned} \tag{3}$$

The statements $s$ considered in the analysis include: (1) Women are much happier if they stay at home and take care of their children; (2) It is much better for everyone if the man earns the main living and the woman takes care of the home and family; and (3) Mothers should not work full time if their child is younger than 5 years old. These statements aim to capture the extent to which agreement with statements related to preferences, social norms, and mother's guilt is predictive of the labour market impact of children.

The corresponding results reflect the correlation between a mother's stated beliefs and her post-child labour market trajectory. $\beta_3$ and $\beta_4$ capture the difference in short and long-run child penalties for women that agree with the statement relative to their counterparts who disagree. Agreement with all three statements is consistent with a reason that a woman may feel pressured to detach from the labour market after having a child. Accordingly, if the channel matters, I would expect $\hat{\beta}_3 < 0$ and $\hat{\beta}_4 < 0$ – i.e. agreement with the statement is associated with a steeper labour market penalty post-birth.

## 4.2 Results and Discussion

### 4.2.1 Heterogeneity by demographic characteristics

Table 2 presents the estimates of equation (2) for mothers and fathers based on membership within demographic category $G$. Figure A3 plots the extensive margin results from the full event study framework for mothers.

**Race.** Table 2 Panel A columns (1) and (3) shows differences in the post birth labour market trajectories of white and Black women. On average, Black women incur a labour market penalty that is 11.5 percentage points smaller on average compared to white women in the 15 years post first birth. As seen in Figure A3 panel (a), This difference is starker in the short-run – white women in the sample detach from the labour market at roughly double the rate of Black women (29.4 vs. 14.6 percentage points) in the first five years post birth. White mothers gradually re-enter the workforce as their child ages, until the trajectories of the two groups converge 12-15 years post birth.

[18] I split the event time into two post periods rather than one because gendered parenthood beliefs may bind more in the short run when the child is younger and not all children are in school. Accordingly, heterogeneity based on such beliefs may differ in the post-birth period before and after age 5.

This observational difference between white and Black women highlights the importance of an intersectional approach when studying child penalties. While there are many differences in the experiences of white and Black mother's in the Unites States, part of the child penalty gap may reflect the preference or gender norm channels. On the preference front, it could be that the child penalty reflects a woman's *choice* to have greater work-life balance after having a child – a choice that some women may disproportionately have the privilege to exercise. Exiting the labour market could be interpreted as a privilege as it is only feasible if a woman and her child have sufficient resources to sustain them. The necessary resources typically come from a second earner, and Black women in the sample are 24 percentage points less likely to have a spouse (54% vs. 78%). Accordingly, detaching from the labour market may not be feasible. Another interpretation is that Black women may feel less pressure to embrace the traditional role of homemaker. Consistent with this, Fortin (2015) finds that the link between gender roles attitudes and labour force participation of Black women is weaker compared to the relationship documented for white women.

In light of this discussion, I next examine heterogeneity in the penalty by marital status and household income.

**Marital status.** Table 2 Panel B columns (1) and (3) document heterogeneous effects of children based on the marital status of the mother. The results show that the experience of unmarried mothers is very different depending on whether they were previously married (i.e. divorced, separated, or widowed) or never married. A previous marriage mutes the post-birth labour market detachment by 6 percentage points relative to married women. This difference is statistically significant but modest, especially in comparison to single women who were never married. This group of unmarried mothers detaches from the labour market 17.9 percentage points less than married women which means their labour market detachment is roughly two-fifths the size. The difference between both groups of unmarried women relative to married women is likely driven (in part) by differences in non-labour income.

The steepest labour market decline is exhibited by married women who have access to non-labour income via their spouse's earnings. If the steeper decline reflects a choice which is enabled by non-labour income, then a negative shock to non-labour income post-birth should generate an income effect which increases labour supply and reverses some of the penalty (assuming leisure is a normal good). Accordingly, women who get divorced during the post-birth period provide a natural test of the intuition provided by the neo-classical labour supply model. Sure enough, the predictions of the model are consistent with the muted penalty observed for women who get divorced after their first birth.

Never married women on the other hand never had access to non-labour income from a spouse. Accordingly, their budget line likely lies well below that of married women. Indeed, women who indicate they were never married report half the household income of married women on average. This means that the labour-leisure choice made by married mothers likely dominates the choice of single mothers, but it is not in their choice set. In other words, these women might prefer to detach from the labour market at similar rates to married women, but consistent with the findings for this group, it is simply not feasible. This interpretation is consistent with the findings of Bertrand et al. (2010), who show the child penalty incurred by female MBAs is highly dependent on spousal income as mothers in better-off households slow down much more after the arrival of children. With that said, the social norm channel is also consistent with the dampened penalty for never married women as the stay-at-home norm is not feasible for many single heads of household.

In summary, never-married women may work substantially more post-child because (i) they do not have the non-labour income necessary to exercise the choice to leave the labour

market (preferences) or (ii) they are less pressured to embrace the traditional role of homemaker (norms). To test (i) directly I next look at differences by household income.

**Household Income.** Table 2 Panel C columns (1) and (3) shows heterogeneous effects of children across household income quartiles with individual's in the first quartile as the reference group. I do not find any evidence of heterogeneity in the child penalty across this dimension. In particular, there is no statistically significant difference between mothers in the lowest and highest household income quartiles. This is evidence against the preference channel which I have argued binds when women have sufficient resources to exercise a preference for greater work-life balance. This suggests that social norms may play a more important role. One way to explore this possibility is to compare women with and without a college degree as I would expect college educated women to be relatively progressive and career oriented.

**College.** Table 2 Panel D columns (1) and (3) shows heterogeneous effects of children based on whether or not the mother has a college degree. I find the child penalty is modestly dampened amongst college educated women (who have 5 percentage point smaller employment decline), but this difference is not statistically significant. In other words, highly educated women also exhibit substantial and persistent declines in labour market attachment after they have children.[19]

The labour market withdrawal of college educated mothers post-children is known as the "opting out" phenomenon, and is regarded as a possible symptom of the end of the ideological "Gender Revolution" (Cotter et al., 2011). Indeed, it is puzzling that women who invest heavily in their human capital do not maximize the return on that investment relative to their male classmates. Again, preferences and gender norms likely reconcile this result.

Kuziemko et al. (2018) hypothesize that women underestimate the child penalty and therefore make human capital investment decisions with incorrect beliefs. Accordingly, women invest in their education under the expectation they can one day "have it all" – a family and career. In reality, many women are not able to achieve both (Coslett, 2005), and those that attempt to pay a price in life statisfaction (Bertrand, 2013). Women cannot update their beliefs until *after* they invest in their education and have a child. Accordingly, Kuziemko et al. (2018) show that agreement with statements regarding gender roles and work-family balance change discontinuously after the birth of their first child. These attitudes appear to exhibit life-cycle effects in which working women adopt more traditional attitudes after becoming mothers, and this belief updating is most prominent amongst college educated women. One interpretation of this finding is that a woman's preferences over family and career change after children.[20] Thus, it could be that women's educational attainment has progressed quicker than the societal norms they are surrounded by.

To reiterate the discussion above with some economic jargon: a child at home raises the opportunity cost of spending time in the office. This cost is proportional to the relative value placed on home life relative to work. Depending on the strength of a preference for a greater work-life balance, women may choose to drop out of the labour market entirely, reduce their hours, or transition to family friendly firms – all of which are responses to children observed empirically (Bertrand et al., 2010; Casarico and Lattanzio, 2021; Kleven et al., 2019a). Mothers that maintain relatively high career drive should detach from the labour market less. Consistent with this, Casarico and Lattanzio (2021) find that the penalty is larger and more persistent for women that work in less productive firms and firms with a higher share of female employees

[19] This withdrawal is observed even amongst women with MBAs from a top institution in the United States (Bertrand et al., 2010).

[20] Indeed, Kuziemko et al. (2018) show that women who become mothers are more likely to downgrade the importance they placed on career as teenagers relative to women without children.

relative to women that work in more productive or male-dominated workplaces, and likely place more weight on their career.

**Birth Cohort.** Table 2 Panel E columns (1) and (3) shows heterogeneous effects based on the mother's birth cohort. All estimates are relative to women born in the 1950s. The results show that the penalty has improved with younger birth cohorts. the post-birth labour market retreat of women born in the 1960s is not statistically different from their counterparts born in the 1950s, but a different story emerges for women born in the 1970s and 80s. Women from these two cohorts detach from the labour market 13 and 21 percentage points less respectively relative to women born in the 1950s. For women born in the 1980s, it appears the penalty is not as permanent or even that large compared to earlier cohorts. With that said, only 400 women in the estimation sample were born in the 1980s and most were born in 1950-1970.

Nevertheless, these findings reflect fundamental changes in the labour market experiences of women across cohorts.[21] For example, trends in labour force participation vary substantially based on whether a woman was born in a cohort (mid 1950s onward) with access to the Pill (Bailey, 2006; Goldin and Katz, 2002) or exposure to the Women's Liberation Movement which popularized the identity of a "career woman" (Fortin, 2015).

Given that the penalty is starting to subside for more recent cohorts of women, did it shift to their husband? Table 2 Panel E column (2) does not show any evidence of offsetting behaviour from men in younger cohorts. This suggests that women in these more recent cohorts became reliant on alternative childcare arrangements outside of maternal or paternal care. Although Kleven et al. (2020) finds such policies had no impact on the penalty in Austria, my findings are consistent with childcare and other family policies playing an important role.

**Men.** Table 2 columns (2) and (4) reports heterogeneous effects across each group of men. White and married men have a modest child *premium* on both the extensive and intensive margins. Men with college degrees also see a bump on the intensive margin. In contrast, black men exhibit a modest 2.2 percentage point decline in participation post-children and are the only group of men to showcase a statistically significant negative change. This set of findings is consistent with that of Hodges and Budig (2010) who examine the "daddy bonus" using the 1979-2006 waves of the National Longitudinal Survey of Youth. They find that men who are white, married, in households with a traditional gender division of labour, and college graduates see the biggest post-child earnings premium. With that said, the key takeaway is that the labour market trajectories of men across all groups is largely unaffected by the arrival of children.

#### 4.2.2 Heterogeneity by Gender Role Attitudes

Table 3 shows heterogeneity in the child penalty across women based on whether or not they agree with a given statement. Figure A4 shows the corresponding event study estimates.

As seen in columns (5) and (6), agreement with the statement "*mothers should not work full time if their child is younger than 5 years old*" is the most important for the child penalty in terms of participation and hours worked. Women who indicate they agree with the statement are 14.7 percentage points more likely to exit the labour market in the first five years of motherhood relative to women that disagree. Agreement is no longer associated with a larger decline in participation in the 6-15 years post-birth, but it is is associated with a larger hours decline in both the short and long-run. This highlights that women tend to have a preference for maternal care when their children are young and may feel guilty placing the child in alternative childcare

[21] See Fortin (2019); Goldin (2021); Goldin and Mitchell (2017) for a detailed discussion.

arrangements. Once the child enters formal school, the mothers that express this concern re-enter the workforce but work fewer hours.

Consistent with this, Fortin (2005) finds that mother's guilt – measured by disagreement with the statement "*a working mother can establish just as warm and secure a relationship with her child as a mother who does not work*" – is closely associated with a woman's employment status. Thus, for women that agree with this statement, access to formal childcare arrangements is unlikely the binding constraint which keeps them out of the labour market in the early years of motherhood. This could explain why childcare subsidies have not always improved female labour force participation (Havnes and Mogstad, 2011; Kleven et al., 2020).

In contrast, agreement with the other two statements related to preferences and gender norms is only associated with a larger penalty on the intensive margin and in the short-run. As reported in columns (1) and (2), women that state they are happier at home are no less likely to exit the labour force after having a child relative to women that disagree. Combined with the first finding presented above, this suggests that a woman's preference to stay home is more likely driven out of consideration for the well-being of her child rather than herself. As reported in columns (3) and (4), women that believe a traditional division of labour is better for everyone are also no less likely to exit the labour force after having a child relative to women that disagree. This suggests that traditional values do not pressure a mother to leave the labour force altogether, but it greatly reduces her attachment.

# 5 The Intergenerational Transmission of Work Preferences

Having shown the differential impact of children on the labour market behaviours of women and men, and presented evidence on the importance of gender related attitudes as a key mechanism in section 3.2, the next question is: when are these beliefs about parenthood formed? Are they learned early in life (Kleven et al., 2019a) or developed after becoming a mother (Kuziemko et al., 2018)? The intergenerational transmission of the penalty offers a window in to the role *nurture* plays in the formation of a woman's preferences over family and career.

## 5.1 Estimating Equation

To study the intergenerational transmission of the penalty I link individual's to the household division of labour of their parents. Let $h_i^m$ and $h_i^f$ denote average annual hours worked of person $i$'s mother and father (whom I will refer to as grandparents from here on out) respectively. I rank each set of grandparent's relative labour supply by quantiles of the distribution of $h_i^m$ - $h_i^f$. A higher rank in this distribution reflects a more "modern" gender division of labour in their childhood home.

Sample size is an issue for this specification as I do not observe hours worked for each set of grandparents. Accordingly, I include all respondents for whom I observe the event of first birth in the sample and replace the full set of event time dummies with a single dummy variable which equals one for positive event times. I am then considering the *average* labour market penalty 15 years following first childbirth. Following Kleven et al. (2019a), I estimate the following equation:

$$y_{it} = \sum_{q=2}^{5} \alpha_q \cdot [\text{After}_t] \cdot \mathbb{1}[\text{Grand}_q] + \sum_a \gamma_a \cdot \mathbb{1}[a = \text{Age}_{it}] + \delta_t + \sum_{q=2}^{5} \phi_q \cdot \mathbb{1}[\text{Grand}_q] + \eta \cdot [\text{After}_t] + \zeta \cdot X_i + \epsilon_{it} \tag{4}$$

Where $\mathbb{1}[\text{After}_t]$ is an event time indicator which equals one for all $t \geq 0$, and $\mathbb{1}[\text{Grand}_q]$ is an indicator which equals one if the relative labour supply of the grandparents is the $q$th quantile. I omit $q = 1$ so all coefficients are relative to households where the grandparents had the most traditional division of labour. $X_i$ includes a full set of dummies for each grandparent's birth cohort. The cohort fixed effects are included to consider the impact of relative grandparental labour supply conditional on the norm for their cohort. In this specification standard errors are robust to an arbitrary form of heteroskedasticity.

### 5.2 Results and Discussion

The results are presented in Table 4 and show that the gender division of labour in the household they grew up matters for women but not men. In particular, women who grew up in a household with a less traditional division of labour[22] are less likely to exit the labour market relative to women who grew up in the most traditional households. In particular, their employment penalty is only three quarters the size. However, I find no evidence that this relative labour division has any impact on hours worked. In contrast, men who grew up in a household where their mother worked more are 3 percentage points less likely to work pre-child, but are no more likely to exit the labour force post-child relative to men who grew up in more traditional homes.

My findings echo that of Kleven et al. (2019a) for Denmark. They find that child penalties are transmitted from parents to their daughters, but not via their in-laws.[23] This means that women incur a larger participation penalty if they grew up in a traditional household in which their mother worked very little compared to their father.

Fernández et al. (2004) contend that a woman's mother-in-law plays an important role in her labour market outcomes as men who grew up with a working mother are more likely to have a wife that works. These men may have developed more egalitarian gender role attitudes and a preference to marry a working woman. This could play a role in the size of their spouse's child penalty because these men may believe in a more equal division of childcare and household responsibilities. I do not test this directly as I do not link respondents to their in-laws; however, I do not find any evidence that son's who grew up observing less traditional divisions of labour take on any of the penalty themselves when they become parents.

My findings speak to the importance of nurture in the formation of a woman's preferences over family and career, but not necessarily the importance of a father's reference point for household labour division. If gender norms are passed down from parents to their daughters, then more egalitarian policies could shape the child penalty for the next generation. In particular, evidence of intergenerational transmission of gender role norms is consistent with the prediction that any shock which increases a woman's labour supply while raising children could reduce the child penalty for subsequent generations. One such shock to maternal labour

[22] i.e. their parents are in the top two quintiles of the distribution of $h_i^m$ - $h_i^f$, where $m$ denotes mother (not male).

[23] I only consider the impact of an individual's own parents, not their spouses. However, this means that I also consider the role of a man's parents on his own labour supply.

supply is universal childcare subsidies, as seen in Quebec (Baker et al., 2008) and following the Lanham Act of WWII (Herbst, 2017).

# 6 Conclusions

The child penalty is well-documented and pervasive across individuals, households, countries, and over time, yet there is a paucity of evidence on the mechanisms behind it. It is important to develop a better understanding of these mechanisms in order to inform the creation and expansion of policies that are aimed to reduce it. In this paper, I explore whether the penalty is driven by gendered preferences or norms.

I use 50 years of data from the Panel Study of Income Dynamics (PSID) to document extensive heterogeneity in post-birth labour market trajectories across groups of women but not men. First, I show that mothers who are Black, single, and from younger birth cohorts are substantially less likely to detach from the labour market relative to their counterparts who are white, married, and from older cohorts. Second, I show there is a strong link between the penalty and individual-level gender-related beliefs. Agreement with a statement asserting that a woman should not work while her child is young is associated with the largest penalty, which suggests it is likely driven by a woman's inner conflict between her competing identities as a mother and career woman. Finally, I estimate the intergenerational transmission of the penalty and find women who grew up in households with a less traditional division of labour have a smaller penalty when they have children.

Taken together, my findings demonstrate that gender norms are a key driver of the child penalty. The penalty then reflects inequities and inefficiencies in the labour market that policy must strive to improve upon in order to fully embrace the talents of both men and women. Moreover, both men and women would benefit from less rigid gendered norms such that both parties can pursue both their family and career aspirations.

## References


**Akerlof, George A and Rachel E Kranton**, "Economics and identity," *The quarterly journal of economics*, 2000, *115* (3), 715–753.

**Andresen, Martin Eckhoff and Emily Nix**, "What causes the child penalty? Evidence from same sex couples and policy reforms," Technical Report, Discussion Papers 2019.

**Angrist, Joshua D and William N Evans**, "Children and Their Parents' Labor Supply: Evidence from Exogenous Variation in Family Size," *The American Economic Review*, 1998, *88* (3), 450–477.

**Bailey, Martha J**, "More power to the pill: The impact of contraceptive freedom on women's life cycle labor supply," *The quarterly journal of economics*, 2006, *121* (1), 289–320.

**Baker, Michael, Jonathan Gruber, and Kevin Milligan**, "Universal child care, maternal labor supply, and family well-being," *Journal of political Economy*, 2008, *116* (4), 709–745.

**Bénabou, Roland and Jean Tirole**, "Identity, morals, and taboos: Beliefs as assets," *The Quarterly Journal of Economics*, 2011, *126* (2), 805–855.

**Bertrand, Marianne**, "New perspectives on gender," in "Handbook of labor economics," Vol. 4, Elsevier, 2011, pp. 1543–1590.

_ , "Career, family, and the well-being of college-educated women," *American Economic Review*, 2013, *103* (3), 244–50.

_ **, Claudia Goldin, and Lawrence F Katz**, "Dynamics of the gender gap for young professionals in the financial and corporate sectors," *American economic journal: applied economics*, 2010, *2* (3), 228–55.

**Browning, Martin**, "Children and household economic behavior," *Journal of Economic Literature*, 1992, *30* (3), 1434–1475.

**Casarico, Alessandra and Salvatore Lattanzio**, "Behind the Child Penalty: Understanding What Contributes to the Labour Market Costs of Motherhood," 2021.

**Cortes, Patricia and Jessica Pan**, "Children and the remaining gender gaps in the labor market," Technical Report, National Bureau of Economic Research 2020.

**Coslett, Caitlin G**, "Can Women Have it All?: Gender Differences in the Relationship Between Career and Family for Top Corporate Executives." PhD dissertation 2005.

**Cotter, David, Joan M Hermsen, and Reeve Vanneman**, "The end of the gender revolution? Gender role attitudes from 1977 to 2008," *American journal of sociology*, 2011, *117* (1), 259–89.

**Farré, Lídia and Francis Vella**, "The intergenerational transmission of gender role attitudes and its implications for female labour force participation," *Economica*, 2013, *80* (318), 219–247.

**Fernández, Raquel, Alessandra Fogli, and Claudia Olivetti**, "Mothers and sons: Preference formation and female labor force dynamics," *The Quarterly Journal of Economics*, 2004, *119* (4), 1249–1299.

**Fortin, Nicole M**, "Gender role attitudes and the labour-market outcomes of women across OECD countries," *oxford review of Economic Policy*, 2005, *21* (3), 416–438.

_ , "Gender role attitudes and women's labor market participation: Opting-out, aids, and the persistent appeal of housewifery," *Annals of Economics and Statistics/Annales d'Économie et de Statistique*, 2015, (117/118), 379–401.

_ , "Increasing earnings inequality and the gender pay gap in Canada: Prospects for convergence," *Canadian Journal of Economics/Revue canadienne d'économique*, 2019, *52* (2), 407–440.

**Goldin, Claudia**, "The quiet revolution that transformed women's employment, education, and family," *American economic review*, 2006, *96* (2), 1–21.

_ , *Career and Family: Women's Century-Long Journey toward Equity*, Princeton University Press, 2021.

_ **and Joshua Mitchell**, "The new life cycle of women's employment: Disappearing humps, sagging middles, expanding tops," *Journal of Economic Perspectives*, 2017, *31* (1), 161–82.

_ **and Lawrence F Katz**, "The power of the pill: Oral contraceptives and women's career and marriage decisions," *Journal of political Economy*, 2002, *110* (4), 730–770.

_ **,** _ **, and Ilyana Kuziemko**, "The homecoming of American college women: The reversal of the college gender gap," *Journal of Economic perspectives*, 2006, *20* (4), 133–156.

_ **, Sari Pekkala Kerr, and Claudia Olivetti**, "The other side of the mountain: women's employment and earnings over the family cycle," 2021.

**Havnes, Tarjei and Magne Mogstad**, "Money for nothing? Universal child care and maternal employment," *Journal of Public Economics*, 2011, *95* (11-12), 1455–1465.

**Herbst, Chris M**, "Universal child care, maternal employment, and children's long-run outcomes: Evidence from the US Lanham Act of 1940," *Journal of Labor Economics*, 2017, *35* (2), 519–564.

**Hodges, Melissa J and Michelle J Budig**, "Who gets the daddy bonus? Organizational hegemonic masculinity and the impact of fatherhood on earnings," *Gender & Society*, 2010, *24* (6), 717–745.

**Kleven, Henrik, Camille Landais, and Jakob Egholt Søgaard**, "Children and gender inequality: Evidence from Denmark," *American Economic Journal: Applied Economics*, 2019a, *11* (4), 181–209.

_ **,** _ **, and** _ , "Does biology drive child penalties? Evidence from biological and adoptive families," *American Economic Review: Insights*, 2021, *3* (2), 183–98.

_ **,** _ **, Johanna Posch, Andreas Steinhauer, and Josef Zweimuller**, "Child penalties across countries: Evidence and explanations," 2019b, *109*, 122–26.

_ **,** _ **,** _ **,** _ **, and Josef Zweimüller**, "Do family policies reduce gender inequality? Evidence from 60 years of policy experimentation," Technical Report, National Bureau of Economic Research 2020.

**Kuziemko, Ilyana, Jessica Pan, Jenny Shen, and Ebonya Washington**, "The mommy effect: Do women anticipate the employment effects of motherhood?," Technical Report, National Bureau of Economic Research 2018.

**Olivetti, Claudia and Barbara Petrongolo**, "The evolution of gender gaps in industrialized countries," *Annual review of Economics*, 2016, *8*, 405–434.

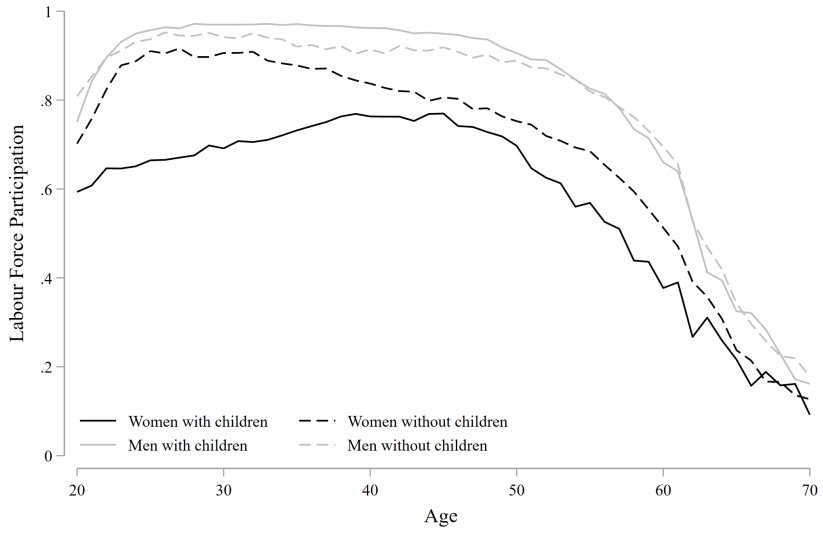
**Rosenzweig, Mark R and Kenneth I Wolpin**, "Testing the quantity-quality fertility model: The use of twins as a natural experiment," *Econometrica: journal of the Econometric Society*, 1980, pp. 227–240.

**Truskinovsky, Yulya and Nicole Maestas**, "Caregiving and labor force participation: new evidence from the american time use survey," *Innovation in Aging*, 2018, *2* (Suppl 1), 580–580.

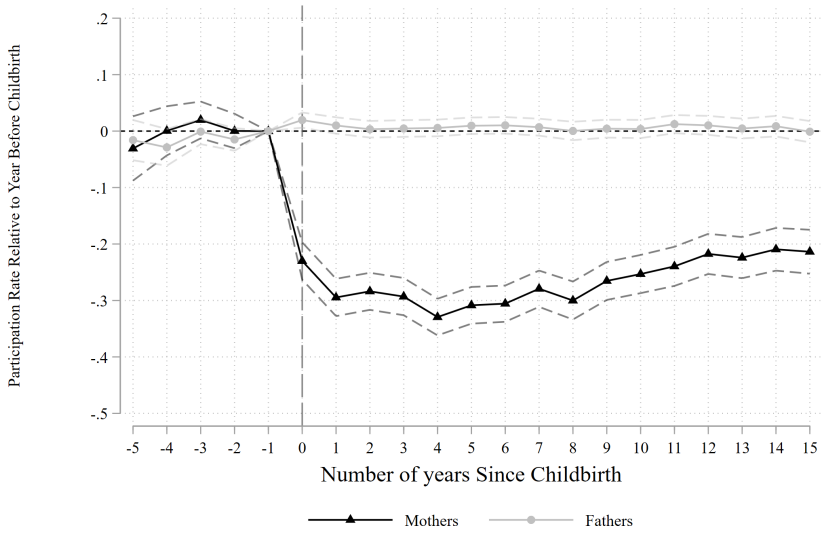
# Figures

**Figure 1:** The Impact of Parenthood on Labour Force Participation

**(a)** Descriptive Evidence




**(b)** Event Study Estimates




*Notes*: Panel (a) plots labour force participation for working age individuals with and without children. Respondents are considered in the labour force if they are working now, temporarily laid off, or unemployed and looking for work. Panel (b) plots the estimated event-time coefficients from equation (1) with labour force participation as the dependent variable. Sample includes all respondents for whom I observe at least once before and after the event of childbirth and at least eight times total over the event window. The dashed lines show 95% confidence bands based on standard errors clustered at the person level.

# Tables

**Table 1:** Summary Statistics

| | | Parent: | | |
|---|---|---|---|---|
| | All | Mother | Father | Difference |
| **Panel A: Labour Market Outcomes** | | | | |
| In labour force | 0.83 | 0.69 | 0.97 | -0.28* |
| | ( 0.37) | ( 0.46) | ( 0.17) | |
| Total annual work hours in previous year | 1658.11 | 1166.94 | 2148.45 | -981.51* |
| | ( 963.14) | ( 903.37) | ( 746.82) | |
| Homemaker | 0.14 | 0.28 | 0.00 | 0.27* |
| | ( 0.35) | ( 0.45) | ( 0.07) | |
| **Panel B: Demographic Characteristics** | | | | |
| White | 0.72 | 0.70 | 0.73 | -0.03* |
| | ( 0.45) | ( 0.46) | ( 0.44) | |
| Black | 0.24 | 0.26 | 0.23 | 0.04* |
| | ( 0.43) | ( 0.44) | ( 0.42) | |
| Has college degree | 0.40 | 0.39 | 0.41 | -0.03* |
| | ( 0.49) | ( 0.49) | ( 0.49) | |
| Married | 0.77 | 0.70 | 0.84 | -0.13* |
| | ( 0.42) | ( 0.46) | ( 0.37) | |
| Never married | 0.05 | 0.06 | 0.03 | 0.03* |
| | ( 0.21) | ( 0.24) | ( 0.18) | |
| Divorced, separated, or widowed | 0.19 | 0.24 | 0.13 | 0.10* |
| | ( 0.39) | ( 0.43) | ( 0.34) | |
| Birth cohort==1950 | 0.33 | 0.32 | 0.35 | -0.03* |
| | ( 0.47) | ( 0.47) | ( 0.48) | |
| Birth cohort==1960 | 0.25 | 0.28 | 0.22 | 0.05* |
| | ( 0.43) | ( 0.45) | ( 0.42) | |
| Birth cohort==1970 | 0.15 | 0.16 | 0.14 | 0.02* |
| | ( 0.35) | ( 0.36) | ( 0.35) | |
| Birth cohort==1980 | 0.05 | 0.07 | 0.03 | 0.04* |
| | ( 0.22) | ( 0.25) | ( 0.18) | |
| **Panel C: Number of Observations Over Event Window** | | | | |
| Total | 16.23 | 16.16 | 16.29 | -0.13* |
| | ( 4.30) | ( 4.25) | ( 4.34) | |
| Pre first birth | 4.40 | 4.38 | 4.43 | -0.05* |
| | ( 1.15) | ( 1.16) | ( 1.14) | |
| Post first birth | 12.23 | 12.19 | 12.27 | -0.09* |
| | ( 3.80) | ( 3.76) | ( 3.85) | |
| Number of individuals | 7,834 | 4,024 | 3,810 | |
| Observations | 57,909 | 28,946 | 28,963 | |

*Notes*: Table reports summary statistics for the estimation sample from the 1968-2017 waves of the Panel Study of Income Dynamics. The first three columns show variable means with standard deviations in parentheses and the last column shows differences across mothers and fathers. Asterisks denote statistical significance at 1% size in t-tests corrected for clustering at the respondent level. Panel A includes the labour market outcomes used as dependent variables throughout the analysis. Panel B includes the demographic characteristics I examine heterogeneity in the child penalty across. Panel C reports the number of times I observe individuals over the event study window.

**Table 2:** Heterogeneity in the Child Penalty by Demographic Characteristics

| | Participation | | Log hours | |
|---|---|---|---|---|
| sample | women | men | women | men |
| **Panel A: Race** | | | | |
| post birth | -0.305* | 0.018* | -0.429* | 0.078* |
| | ( 0.015) | ( 0.006) | ( 0.025) | ( 0.014) |
| post birth × black | 0.115* | -0.040* | 0.204* | -0.086* |
| | ( 0.032) | ( 0.010) | ( 0.054) | ( 0.031) |
| **Panel B: Marital Status** | | | | |
| post birth | -0.304* | 0.018* | -0.465* | 0.054* |
| | ( 0.013) | ( 0.005) | ( 0.024) | ( 0.012) |
| post birth × never married | 0.179* | 0.065 | 0.178 | -0.081 |
| | ( 0.076) | ( 0.045) | ( 0.193) | ( 0.167) |
| post birth × previously married | 0.066* | -0.003 | 0.094* | -0.006 |
| | ( 0.031) | ( 0.021) | ( 0.045) | ( 0.052) |
| **Panel C: Household Income** | | | | |
| post birth | -0.290* | 0.100* | -0.397* | 0.156* |
| | ( 0.056) | ( 0.041) | ( 0.059) | ( 0.041) |
| post birth × second quartile | -0.028 | -0.089* | -0.061 | -0.033 |
| | ( 0.062) | ( 0.042) | ( 0.067) | ( 0.044) |
| post birth × third quartile | 0.047 | -0.077 | 0.068 | -0.098* |
| | ( 0.059) | ( 0.042) | ( 0.067) | ( 0.043) |
| post birth × fourth quartile | 0.086 | -0.080 | 0.085 | -0.109* |
| | ( 0.058) | ( 0.042) | ( 0.063) | ( 0.043) |
| **Panel D: College** | | | | |
| post birth | -0.284* | 0.027* | -0.514* | -0.005 |
| | ( 0.021) | ( 0.009) | ( 0.037) | ( 0.021) |
| post birth × college | 0.059 | -0.020 | 0.025 | 0.089* |
| | ( 0.033) | ( 0.017) | ( 0.064) | ( 0.036) |
| **Panel E: Birth Cohort** | | | | |
| post birth | -0.340* | -0.001 | -0.459* | 0.034 |
| | ( 0.021) | ( 0.006) | ( 0.037) | ( 0.018) |
| post birth × cohort==1960 | 0.050 | 0.017 | 0.049 | -0.039 |
| | ( 0.028) | ( 0.012) | ( 0.049) | ( 0.030) |
| post birth × cohort==1970 | 0.134* | 0.066* | 0.161* | -0.036 |
| | ( 0.032) | ( 0.018) | ( 0.053) | ( 0.035) |
| post birth × cohort==1980 | 0.211* | -0.008 | 0.351 | -0.021 |
| | ( 0.071) | ( 0.011) | ( 0.208) | ( 0.148) |

*Notes.* Each column reports estimates of Equation (2) from separate regressions with the dependent variable listed in the top heading. For each subsample I report the estimated coefficients on the pooled event time dummy variable on its own and interaction with a dummy for group $g$. Each specification includes a full set of interactions between the dummy variable for group $g$ and the after indicator variable, age fixed effects, the time fixed effects. Sample includes all respondents from the 1968-2017 waves of the Panel Study of Income Dynamics for whom I observe at least once before and after the event of first childbirth and at least eight times total over the event window. This specification pools all birth event-time coefficients from equation (1) into a single dummy which equals one for $\tau \geq 0$ and zero before the first birth. Standard errors robust to clustering at the person level are reported in parentheses. Asterisks denote statistical significance at the 5% level.

**Table 3:** Heterogeneity in the Child Penalty by Gender Role Attitudes

| | Women Happier at Home | | Better if Dad Earns | | Mother Shouldn't Work | |
|---|---|---|---|---|---|---|
| Outcome | lfp | log hours | lfp | log hours | lfp | log hours |
| 0 to 5 years post | -0.242* | -0.405* | -0.243* | -0.421* | -0.197* | -0.357* |
| | ( 0.018) | ( 0.032) | ( 0.017) | ( 0.033) | ( 0.018) | ( 0.033) |
| 6 to 15 years post | -0.273* | -0.480* | -0.265* | -0.489* | -0.243* | -0.439* |
| | ( 0.023) | ( 0.046) | ( 0.023) | ( 0.047) | ( 0.022) | ( 0.046) |
| agree$^s$× (0 to 5 years post) | -0.041 | -0.197* | -0.051 | -0.147* | -0.147* | -0.294* |
| | ( 0.033) | ( 0.071) | ( 0.032) | ( 0.067) | ( 0.030) | ( 0.062) |
| agree$^s$× (6 to 15 years post) | 0.030 | -0.166 | 0.003 | -0.090 | -0.056 | -0.204* |
| | ( 0.048) | ( 0.106) | ( 0.045) | ( 0.101) | ( 0.042) | ( 0.094) |
| Proportion agree | 0.26 | 0.26 | 0.29 | 0.29 | 0.38 | 0.38 |
| Age fixed effects | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Year fixed effects | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Individuals | 1,127 | 1,082 | 1,150 | 1,105 | 1,152 | 1,106 |
| Observations | 14,225 | 10,504 | 14,549 | 10,765 | 14,591 | 10,807 |

*Notes.* Table reports estimates from Equation (3). Each column reports estimates from separate regressions. Standard errors robust to clustering at the person level are reported in parentheses. Asterisks denote statistical significance at the 5% level. Sample includes all respondents in the 1997 and 2014 waves of the PSID Child Development Supplement that has a birth record for their first born child. agree$_j$ is an indicator variable which equals one if the respondent agrees with statement $s$ and zero otherwise. The statements in the order presnted in columns are as follows: (1) Women are much happier if they stay at home and take care of their children; (2) It is much better for everyone if the man earns the main living and the woman takes care of the home and family; and (3) Mothers should not work full time if their child is younger than 5 years old. ; The event time indicator is split into 0 to 5 years and 6 to 15 years post first birth.

**Table 4:** The Intergenerational Transmission of Work Preferences

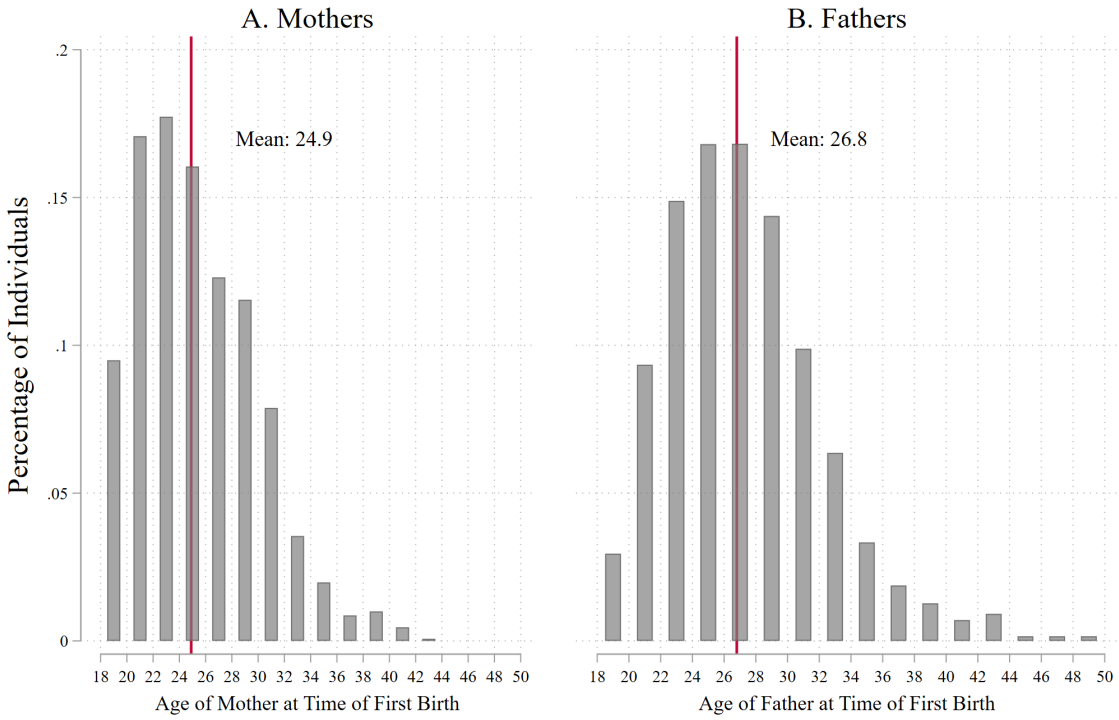
| | lfp | | log hours | |
|---|---|---|---|---|
| Sample | Women | Men | Women | Men |
| after | -0.204* | 0.012 | -0.404* | -0.001 |
| | ( 0.015) | ( 0.010) | ( 0.035) | ( 0.023) |
| $Grand_{Q2}$ | 0.038* | 0.004 | 0.076* | 0.015 |
| | ( 0.015) | ( 0.013) | ( 0.035) | ( 0.027) |
| $Grand_{Q3}$ | 0.021 | -0.002 | 0.080* | 0.003 |
| | ( 0.015) | ( 0.013) | ( 0.035) | ( 0.026) |
| $Grand_{Q4}$ | 0.041* | -0.011 | 0.077* | -0.045 |
| | ( 0.016) | ( 0.014) | ( 0.036) | ( 0.028) |
| $Grand_{Q5}$ | 0.032 | -0.033* | 0.079* | -0.067 |
| | ( 0.017) | ( 0.016) | ( 0.040) | ( 0.035) |
| $Grand_{Q2}\times$ after | 0.015 | 0.003 | -0.034 | -0.042 |
| | ( 0.018) | ( 0.014) | ( 0.043) | ( 0.031) |
| $Grand_{Q3}\times$ after | 0.035 | 0.003 | 0.015 | -0.060* |
| | ( 0.018) | ( 0.014) | ( 0.043) | ( 0.030) |
| $Grand_{Q4}\times$ after | 0.043* | 0.010 | 0.057 | -0.023 |
| | ( 0.019) | ( 0.015) | ( 0.044) | ( 0.032) |
| $Grand_{Q5}\times$ after | 0.058* | 0.021 | 0.040 | 0.001 |
| | ( 0.021) | ( 0.017) | ( 0.048) | ( 0.038) |
| Age fixed effects | ✓ | ✓ | ✓ | ✓ |
| Year fixed effects | ✓ | ✓ | ✓ | ✓ |
| Grandparent controls | ✓ | ✓ | ✓ | ✓ |
| Individuals | 2,711 | 2,182 | 2,401 | 1,910 |
| Observations | 23,814 | 18,234 | 15,423 | 13,562 |

*Notes.* Table reports estimates from Equation (4). Each column reports estimates from separate regressions. Robust standard errors are repored in parentheses. Asterisks denote statistical significance at the 5% level. The dependent variable in column (1) and (2) is labour force participation. The dependent variable in column (3) and (4) is log hours worked. Grandparent controls include a full set of dummies for the birth cohort of the respondent's mother and father.

# A Appendix

## A.1 Supplementary Figures

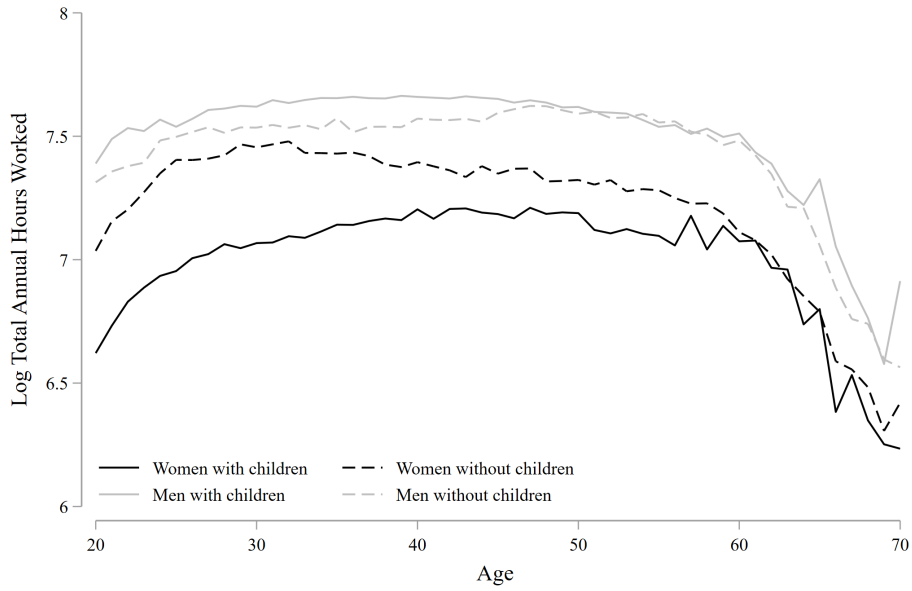
**Figure A1:** Distribution of Parental Age at First Birth



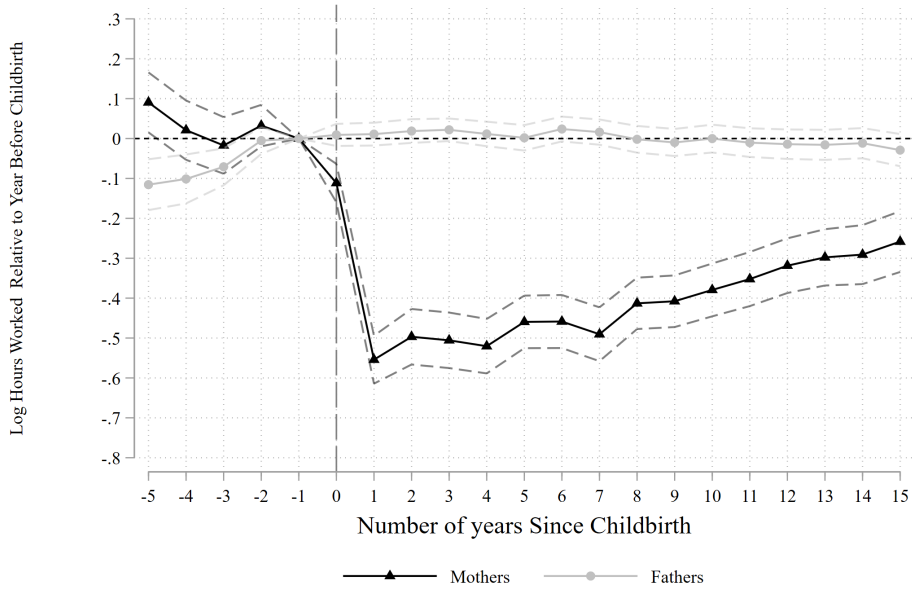

*Notes*: Figure displays the distribution of the age in which men and women had their first child in two year bins. Vertical line marks the mean.

**Figure A2:** The Impact of Parenthood on Hours Worked

**(a)** Descriptive Evidence



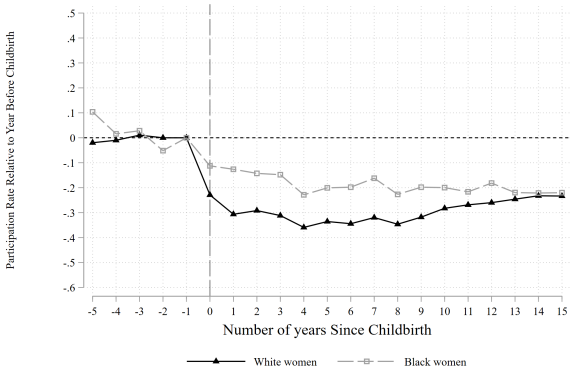

**(b)** Event Study Estimates



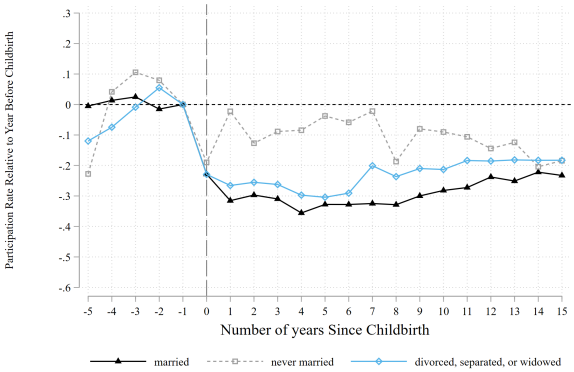

*Notes*: Panel (a) plots hours worked for working age individuals with and without children. Hours worked refers to total annual work hours on all jobs including overtime. This figure includes individuals who worked zero hours so it captures both extensive and intensive margin responses. Panel (b) only captures the intensive margin response to children as it plots the estimated event-time coefficients from equation (1) with log hours worked as the dependent variable. Sample includes all respondents for whom I observe at least once before and after the event of childbirth and at least eight times total over the event window. The dashed lines show 95% confidence bands based on standard errors clustered at the person level.

**Figure A3:** Heterogeneity in the Child Penalty by Demographic Characteristics

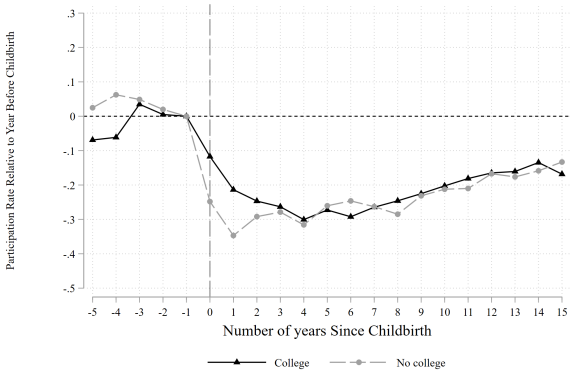
**(a)** Race



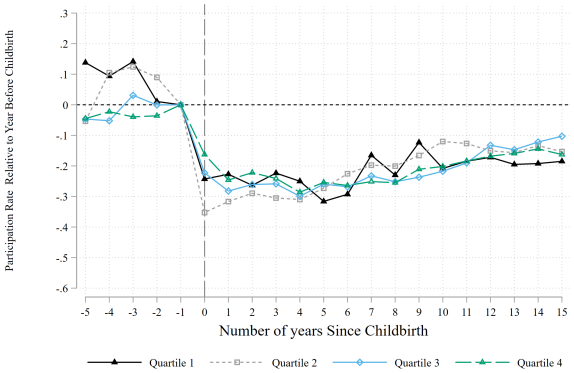

**(b)** Marital status



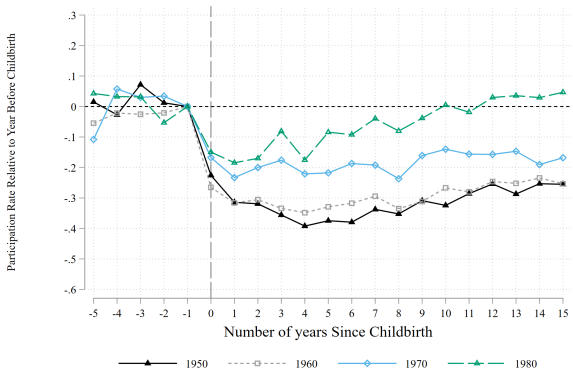

**(c)** College




**(d)** Household Income

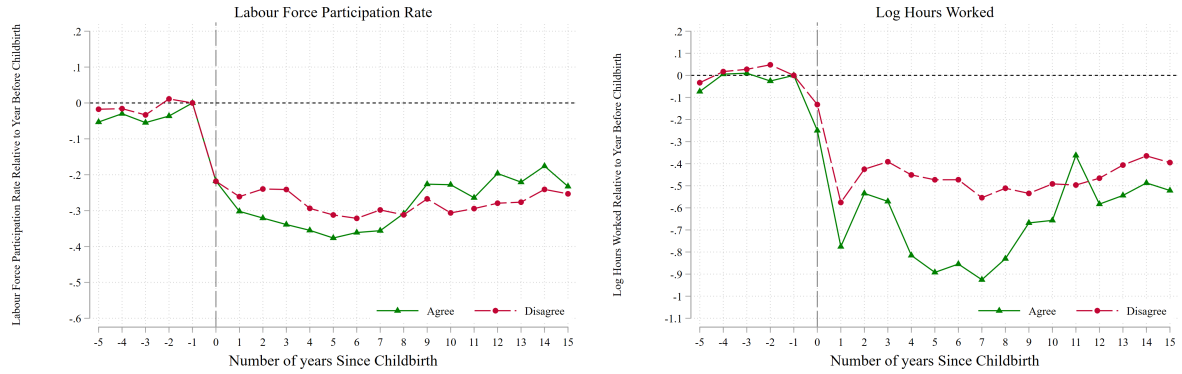



**(e)** Birth cohort

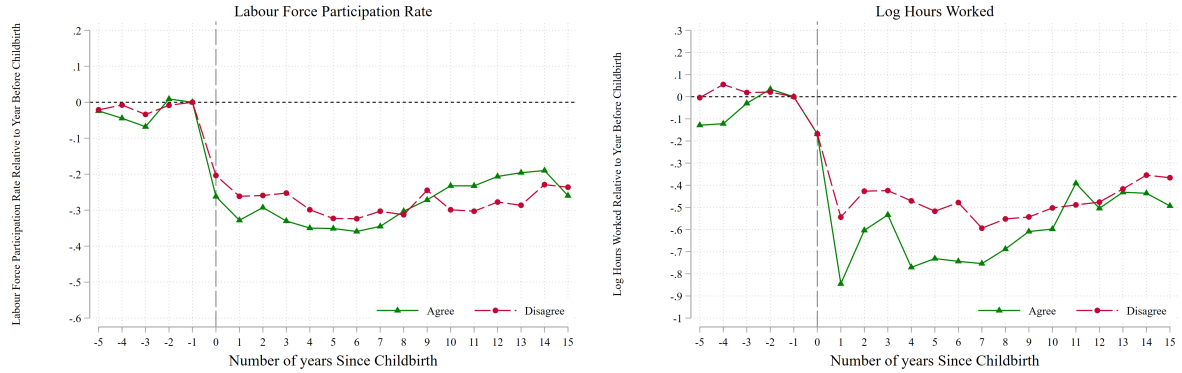



*Notes*: Figure plots the estimated event-time coefficients from equation (1) for different demographic groups with labour force participation as the dependent variable. Sample includes all respondents for whom I observe at least once before and after the event of childbirth and at least eight times total over the event window. The dashed lines show 95% confidence bands based on standard errors clustered at the person level.

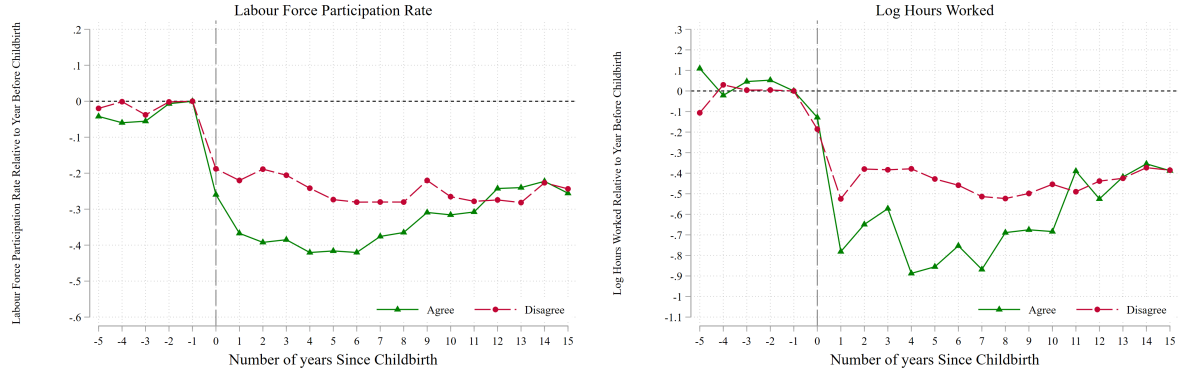
**Figure A4:** Heterogeneity in the Child Penalty by Gender Role Attitudes




"Women are much happier if they stay at home and take care of their children."




"It is much better for everyone if the man earns the main living and the woman takes care of the home and family."




"Mothers should not work full time if their child is younger than 5 years old."

## A.2 Supplementary Tables

**Table A1:** The Impact of Parenthood on Labour Market Outcomes

| | Participation | | Log hours | | Probability homemaker | |
|---|---|---|---|---|---|---|
| sample | women | men | women | men | women | men |
| 0 to 5 years post | -0.297* | 0.017* | -0.458* | 0.050* | 0.314* | -0.000 |
| | ( 0.011) | ( 0.005) | ( 0.020) | ( 0.012) | ( 0.010) | ( 0.001) |
| 6 to 15 years post | -0.261* | 0.015* | -0.400* | 0.037* | 0.260* | -0.003 |
| | ( 0.013) | ( 0.006) | ( 0.025) | ( 0.014) | ( 0.012) | ( 0.002) |
| Pre-child mean | 0.944 | 0.956 | 7.413 | 7.598 | 0.019 | 0.007 |
| Age fixed effects | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Year fixed effects | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Individuals | 4,024 | 3,810 | 3,839 | 3,839 | 4,024 | 3,810 |
| Observations | 28,944 | 28,963 | 27,593 | 36,029 | 28,944 | 28,963 |

*Notes.* Each column reports estimates from separate regressions with the dependent variable listed in the top heading. Sample includes all respondents from the 1968-2017 waves of the Panel Study of Income Dynamics for whom I observe at least once before and after the event of first childbirth and at least eight times total over the event window. This specification pools the post birth event-time coefficients from equation (1) into two periods: 0 to 5 years and 6 to 15 years after first birth. Standard errors robust to clustering at the person level are reported in parentheses. Asterisks denote statistical significance at the 5% level.

## A.3 Variable Definitions

### A.3.1 Outcomes of interest

Labour force participation and status as a homemaker is available from 1979 onwards. Both outcomes are dummy variables and based on answers to a question asking whether the respondent is working now, temporarily laid off, unemployed and looking for work, retired, keeping house, or a student. Respondents are coded as "in labour force" if they are working now, temporarily laid off, or unemployed and looking for work. Those who report that they are retired, keeping house, or a student are coded as "not in labour force". All other values are coded as missing. Respondents are coded as "staying at home" if they report keeping house, "not staying home" if they belong to the other specified groups, and missing otherwise.

Work hours are available in all years for reference persons and their spouse. The variable represents total annual work hours on all jobs including overtime in a given year. I use the logarithm of this variable as the left-hand side variable of equation (1). This removes individuals who worked zero hours that year to capture the intensive margin response following childbirth. As a robustness check to the self-reported employment status discussed above, I capture the extensive margin with a dummy variable for which the respondent is "working" if they report a positive number of hours worked and "not working" if they report zero hours worked.

### A.3.2 Event time

The event time variable is calculated as the number of months between the date of birth of the respondent's first child and the current interview date divided by 12 and rounded to the closest year. Date of birth of an individual's first child is obtained from the PSID's 1985-2019 Childbirth and Adoption History File. Date of birth is only known up to the month, season, or year for some children. In these cases, I apply the assumptions used in Kuziemko et al. (2018) and Kleven et al. (2020): (1) if the birth month is missing the child is recorded as born in September; (2) If the child is born in winter, spring, summer, or fall the child is assumed to be born January, April, July, and October respectively; (3) if the birth year is missing the date of birth is recorded as missing not included in the sample.

### A.3.3 Demographic information

In each survey wave, the PSID records the age, race, household income and number of years of education a respondent has received. The respondent is coded as having a college degree if they complete 16 or more years of schooling.

### A.3.4 Questions from the Child Development Supplement

I create dummy variables based on individual's responses to a set of questions which ask them to indicate their level of agreement with a given statement. The variable equals 1 if the respondent "agrees" – i.e. if they indicate that they agree/strongly agree with the statement– and 0 if they disagree/strongly disagree. The statements used in my analysis are as follows:

- It is much better for everyone if the man earns the main living and the woman takes care of the home and family.
- Mothers should not work full time if their child is younger than 5 years old.
- Women are much happier if they stay at home and take care of their children.

## A.4 Empirical Challenge

An extensive literature reports estimates of the relationship between children and family labour supply. The vast majority of these studies focus on the link between fertility and female labour supply and find a negative correlation between the two. However, the interpretation of this correlation is unclear due to the endogeneity of children (Browning, 1992).

One such challenge is that the causal relationship between fertility and female labour market outcomes could run in either direction. As pointed out by Angrist and Evans (1998), this endogeneity issue is reflected in the academic research agenda in this area. To formalize this point, consider a model which can be generically expressed as,

$$y_1 = \theta y_2 + X\beta + u \tag{5}$$

where $y_i$ $(i = 1, 2)$ is either a fertility variable or a measure of labour market attachment, $X$ denotes a set of controls, and $u$ is an error term. One branch of the literature designates fertility as $y_1$ to measure the effect of labour-force attachment on fertility, whereas others place fertility as $y_2$ to estimate a labour supply equation. Since fertility cannot be simultaneously dependent and exogenous, neither equation will likely hold a causal interpretation but both are consistent with the documented negative correlation between the two variables.

Second, women's labour supply and fertility choices are jointly determined by factors which may not be controlled for with $X$ in equation (5). For example, innovation in contraception led women to increase career investment (Goldin and Katz, 2002) and decrease fertility. The "Pill Revolution" also coincided with other important societal shifts in gender role attitudes (Fortin, 2015) which would impact labour force attachment and fertility in the same direction. These types of confounding factor would lead the estimate of $\theta$ to overstate the negative relationship between $y_1$ and $y_2$.

The ideal experiment designed to uncover $\theta$ would randomize fertility. Absent such an experiment, researchers have found creative quasi-natural experiments in which some women are randomly nudged have an additional child. For instance, there is plausibly exogenous variation in the number of children a woman has generated by twin births (Rosenzweig and Wolpin, 1980) and the sex mix of the first two children (Angrist and Evans, 1998). This variation can be exploited to uncover the local average treatment effect (LATE) of having a second or third child.

While interesting, more recent studies have focused on estimating the impact of *total* fertility on labour market outcomes. These studies use an event study approach based around the arrival of a first child proposed by Kleven et al. (2019a). While fertility choices are not exogenous, the sharp discontinuity in labour market outcomes around the timing of first births is arguably orthogonal to confounding factors which should evolve smoothly over time. I discuss the validity of this assumption in greater depth in section 3.1.

The event study approach has many desireable features. First, the estimates return the effect of the first child on labour market outcomes rather than the LATE associated with having a second (in the case of twins) or third (in the case of sibling sex mix). Second, it allows the researcher to trace out the dynamic trajectory of the effect of parenthood on labour market outcomes as the child ages. This feature is important for assessing the impacts of family policies such as paternal leave and childcare programs. Third, while the approach is based around the event of first childbirth, women may have additional children throughout the event study window so estimates of long-run child penalties capture the effect of *total* fertility. Finally, the effect is estimated with greater precision than possible with an instrumental variables approach. See Kleven et al. (2019a) for a more detailed discussion comparing the two strategies.[24]

[24]The authors estimate the effect of parenthood using both the event study and two instrumental variables approaches and compare the results.