

Children and Gender Inequality: Evidence from Denmark*

Henrik Jacobsen Kleven, London School of Economics

Camille Landais, London School of Economics

Jakob Egholt Søgaard, University of Copenhagen

October 2016

Abstract

Despite considerable gender convergence over time, substantial gender inequality persists in all countries. Using Danish administrative data from 1980-2013 and an event study approach, we show that most of the remaining gender inequality in earnings is due to children. The arrival of children creates a gender gap in earnings of around 20% in the long run, driven in roughly equal proportions by labor force participation, hours of work, and wage rates. Underlying these “child penalties”, we find clear dynamic impacts on occupation, promotion to manager, sector, and the family friendliness of the firm for women relative to men. Based on a dynamic decomposition framework, we show that the fraction of gender inequality caused by child penalties has increased dramatically over time, from about 40% in 1980 to about 80% in 2013. As a possible explanation for the persistence of child penalties, we show that they are transmitted through generations, from parents to daughters (but not sons), consistent with an influence of childhood environment in the formation of women’s preferences over family and career.

*Kleven: h.j.kleven@lse.ac.uk; Landais: c.landais@lse.ac.uk; Søgaard: jes@econ.ku.dk. We thank Oriana Bandiera, Tim Besley, Raj Chetty, Claudia Goldin, Lawrence Katz, Gilat Levy, Marjorie McElroy, Suresh Naidu, Torsten Persson, and numerous seminar participants for helpful comments and discussions.

1 Introduction

Despite considerable gender convergence over the last century, gender inequality in earnings and wage rates continues to be substantial in all countries and the process of convergence has slowed down. The early literature on gender inequality in the labor market focused on the role of human capital accumulation and discrimination (Altonji & Blank 1999), but the disappearance of gender differences in education and the implementation of anti-discrimination policies suggest that the explanation for the remaining gender gap lies elsewhere. Based on administrative data for the full population in Denmark since 1980, we provide a simple explanation for the persistence of gender inequality: the effects of children on the careers of women relative to men are large and have not fallen over time. As a result, almost all of the remaining gender inequality can be attributed to children.

To provide context, Figure 1A shows the evolution of the gender gap in earnings for full-time workers in different countries. It is striking that the cross-country differences in gender inequality have largely disappeared over time. For example, while gender inequality in Denmark was dramatically lower than in the United States around 1980, today the gender pay gap is between 15-20% in both countries and appears to have plateaued at that level. That is, gender convergence happened earlier in Scandinavia than elsewhere, but the process also slowed down earlier in Scandinavia allowing other countries to catch up. So even though these countries feature different public policies and labor markets, they are no longer very different in terms of overall gender inequality.

To estimate the impact of children on the labor market trajectories of women relative to men, we adopt a quasi-experimental approach based event studies around the birth of the first child. For a range of labor market outcomes, we find large and sharp effects of children: women and men evolve in parallel until the birth of their first child, diverge sharply immediately after child birth, and do not converge again. Defining the “child penalty” as the percentage by which women fall behind men due to children, we find that the long-run child penalty in earnings equals about 20% over the period 1980-2013. This should be interpreted as a total penalty including the costs of children born after the first one, and we show that the penalty is increasing in the number of children. The earnings penalty can come from three margins — labor force participation, hours of

work, and the wage rate — and we find sharp effects on all three margins that are roughly equal in size.¹ We are not the first to provide event study evidence on the effect of parenthood on earnings and labor supply, but we contribute methodologically by comparing basic event study estimates to more sophisticated event study approaches that use control groups or instruments for child birth. These analyses suggest that the event study approach, once we control non-parametrically for age and time trends, does a good job of identifying child penalties even in the long run.

We then use the event study approach to shed light on a range of underlying mechanisms that may drive the earnings and wage rate impacts. We show that children affect the job characteristics of women relative to men in a way that favor family amenities over pecuniary rewards. Specifically, just after the birth of the first child, women start falling behind men in terms of their occupational rank (as ordered by earnings or wage rate levels) and their probability of becoming manager. Moreover, women switch jobs to firms that are more “family friendly” as measured by the share of women with young children in the firm’s workforce, or by an indicator for being in the public sector which is known to provide flexible working conditions for parents. The importance of the family friendliness of occupations and firms for gender equality has been much discussed in recent work (e.g. [Goldin 2014](#); [Goldin & Katz 2016](#)), but here we provide clean event study evidence that these qualitative dimensions directly respond to the arrival children.

Having estimated child penalties in the full population — allowing them to vary both across event time and across birth cohorts — we are able to decompose aggregate gender inequality over time into child-related inequality and residual inequality. We show that the fraction of total earnings inequality caused by child penalties has doubled over time, from about 40% in 1980 to about 80% in 2013. This dramatic change reflects a combination of two underlying changes: (i) child-related gender inequality in earnings has increased from about 18% to 20%, and (ii) total gender inequality in earnings has fallen from about 46% to 24%.² To understand the first effect, note that although the percentage child penalty has fallen somewhat over time, the penalty operates on a larger base due to the general increase in the earnings of women relative to men. While the literature already highlights the importance of parenthood for gender gaps, our decomposition analysis goes further by quantifying just how much of aggregate gender inequality is due to children and how this has evolved over time. This exercise is demanding as it requires estimates of the impact

¹Our ability to precisely estimate hours and wage rate effects relies on a unique administrative measure of working hours that is available for the full population.

²These gender gaps are larger than those reported in Figure 1A discussed above. This is because the cross-country evidence in Figure 1A is based on *median* earnings for *full-time* workers, whereas we are now considering *mean* earnings for *all* workers as shown in Figure 1B.

of children at each different event time and for each birth cohort, but the Danish data enable us to do this. The striking finding from this exercise is that child penalties are gradually taking over as the key driver of gender inequality.

It is worth highlighting that our dynamic decomposition analysis represents a departure from standard gender gap decompositions (see [Blau & Kahn 2016](#) for a review) both in terms of the variation used (within-person time variation as opposed to cross-sectional variation) and in terms of the question asked. Standard decomposition approaches focus on the extent to which men and women receive unequal pay for equal work: the unexplained gender gap after controlling for human capital and labor market variables, but not children. By contrast, our goal is to study the impact of children on gender inequality, not controlling for labor market variables (such as occupation and firm choices) that are transmission mechanisms for children. This is a conceptually different question: Even with perfectly equal pay for equal work — a zero gap in standard decompositions — our analysis would still show large child-related gender inequality as equal work is in practice not an option for most women with children.

Why are female child penalties so persistent after decades of effort to create gender equality through equal opportunity legislation, child care policies, and job-protected parental leave? While fully answering this question is beyond the scope of our paper, we provide evidence on one possible explanation: child penalties are transmitted through generations. We estimate the intergenerational transmission of child penalties by exploiting that our administrative measure of hours worked goes back to 1964, allowing us to relate the estimated child penalties between 1980-2013 to the within-family work history one generation before. We find that female child penalties are strongly related to the labor supply history of the maternal grandparents, but not the paternal grandparents, even after controlling for a rich set of family characteristics. For example, in traditional families where the mother works very little compared to the father, their daughter incurs a larger child penalty when she eventually becomes a mother herself. Our findings are consistent with an influence of nurture in the formation of women's preferences over family and career. This analysis is related to the work by [Fernandez *et al.* \(2004\)](#), but focusing on the intergenerational transmission of child penalties (as opposed to labor supply levels) between parents and their daughters (as opposed to daughters-in-law). The analysis is also related to recent work on the importance of gender identity norms for labor market outcomes ([Bertrand 2011](#); [Bertrand *et al.* 2013](#)). Our findings suggest that female gender identity is formed during a girl's childhood based on the gender roles of her parents.

Our paper contributes primarily to two literatures. First and foremost, we contribute to the enormous literature on gender inequality in the labor market as reviewed by for example [Altonji & Blank \(1999\)](#), [Bertrand \(2011\)](#), [Blau & Kahn \(2016\)](#) and [Olivetti & Petrongolo \(2016\)](#). Much of this literature has focused on the role of human capital, occupation and discrimination in explaining gender gaps, but there is also a sizeable amount of work on the role of parenthood. This includes papers by [Waldfogel \(1998\)](#), [Lundberg & Rose \(2000\)](#), [Paull \(2008\)](#), [Bertrand *et al.* \(2010\)](#), [Wilde *et al.* \(2010\)](#), [Fernandez-Kranz *et al.* \(2013\)](#), [Fitzenberger *et al.* \(2013\)](#), [Goldin \(2014\)](#), [Adda *et al.* \(2015\)](#), [Angelov *et al.* \(2016\)](#), and [Goldin & Katz \(2016\)](#). Our paper is most closely related to the case study of MBA graduates from Chicago Booth School of Business by [Bertrand *et al.* \(2010\)](#), and to the paper by [Angelov *et al.* \(2016\)](#) who estimate child penalties in annual earnings and full-time equivalent monthly wages using Swedish administrative data. Starting from related event studies on earnings and wage rates, we provide evidence on mechanisms (the dynamics of occupation, sector, and firm choices), investigate causal identification using different approaches, develop a dynamic decomposition approach to estimating long-run trends in child-related gender inequality, and finally study the intergenerational transmission of child penalties.

Second, we contribute to the literature on children and family labor supply. This literature has focused on the potential endogeneity of children (e.g. [Browning 1992](#)), proposing instruments for the number of children such as twin births ([Rosenzweig & Wolpin 1980](#); [Bronars & Grogger 1994](#)) and sibling sex mix ([Angrist & Evans 1998](#)). While our primary objective is to study gender inequality, our analysis also yields estimates of the impact of children on hours worked and labor force participation, separately for males and females. Our event study approach relies on a different source of variation and captures a different impact than the existing IV-approaches. As we clarify, the event study approach has the potential to capture the global treatment effect of all children in the population, as opposed to only the local treatment effect of a second child or a third child obtained from the twin or sibling sex mix instruments.³ We lay out the identification assumptions of our approach and investigate their validity, providing two compelling identification checks. First, we compare the local treatment effect of a third child based on a sex-mix IV to the local treatment effect of a third child that can be obtained from our event study approach. We show that the IV estimates and the event study estimates are almost perfectly aligned through event

³We focus on the twin and sex-mix instruments as these have been the most influential approaches, but other instruments have been proposed that might help uncover the effect of the first child as well. These instruments include miscarriages ([Hotz *et al.* 2005](#)), infertility shocks ([Aguero & Marks 2008](#)), and IVF treatment outcomes ([Lundborg *et al.* 2014](#)).

time. Second, we provide a difference-in-differences extension of our event study, using those who never have children as a control group. The difference-in-differences event study produces impacts of children that are very similar to our baseline results.

Finally, we note that children may have two conceptually different effects on labor market outcomes. One is a pre-child effect of anticipated fertility: women may invest less in education or select family friendly career paths in anticipation of motherhood. The other is a post-child effect of realized fertility: women changing their hours worked, occupation, sector, firm, etc., in response to actual motherhood. The event study approach cannot capture pre-child effects; it is designed to identify post-child effects conditional on pre-child choices. If women invest less in education and career in anticipation of motherhood, then our estimated child penalties represent lower bounds on the total lifetime impacts of children. The fact that these child penalties can account for most of the current gender inequality leaves relatively little room for pre-child effects to operate today, but it is possible that such effects were more important at the beginning of the period we study. In general, identifying anticipation effects of children is difficult without making strong structural assumptions and we make no such attempt here.⁴ We provide descriptive evidence that, while the effect of child penalties on gender inequality has been growing over time, the effect of pre-child human capital investments has fallen. This is consistent with a shift from pre-effects to post-effects of children as a society develops: while women used to pay the career cost of children upfront, they now invest in education and careers almost at the level of men, but still end up paying the child penalty after motherhood.

The paper is organized as follows. Section 2 describes the institutional background and data, section 3 lays out the event study methodology, estimates the impacts of children and investigates identification, section 4 presents our dynamic decomposition of gender inequality, section 5 analyzes intergenerational transmission, and section 6 concludes.

2 Institutional Background and Data

2.1 Institutional Background

Scandinavian countries have been praised for offering better opportunities for women to balance career and family than most other countries. This view is based on the presence of generous family

⁴Taking a structural approach, [Adda et al. \(2015\)](#) estimate that occupational choices due to anticipated fertility represent a very small fraction (less than 5%) of the total earnings loss from children.

policies — job-protected parental leave and public provision of child care — and a perception that gender norms are comparatively egalitarian in Scandinavia. Consistent with this view, Denmark has one of the highest female labor force participation rates in the world, currently around 80% as opposed to around 70% in the United States, and it has almost no remaining gender gap in participation rates. However, Figure 1 shows that this is far from the full story. The cross-country comparisons in Panel A were discussed above and shows that Denmark is no longer a strong outlier in terms of the gender gap in earnings for full-time workers. Panel B focuses on Denmark alone and shows gender gaps in different labor market outcomes for all workers. We see that the gender gap in participation has gradually disappeared over the last three decades and that the gender gap in hours worked has fallen substantially, but that large gaps persist in earnings and wage rates (defined as earnings/hours worked among those who are working). The earnings gap is now around 22% and is created mostly by differences in wage rates and to a smaller degree by differences in hours worked.⁵

Figure 2 probes the idea that gender norms are more egalitarian in Scandinavia than elsewhere. The evidence in the figure is based on questions from the International Social Survey Program (ISSP) regarding the attitudes that people have towards market work by women with and without children. Specifically the survey asks participants whether they think women should work outside the home full-time, part-time or not at all when they have no children (Panel A), have children under school age (Panel B), have children in school (Panel C), and have children who have left home (Panel D). Two striking insights emerge from the figure: one is that gender attitudes are still quite traditional — essentially that women should work full-time before having children and after the children have left home, while they should work only part-time or not at all when they have children living at home — and the other is that different countries are very similar in holding this view. The only noticeable cross-country difference is that the Scandinavian populations are somewhat more open to the idea that women with young children work part time (rather than staying at home entirely) compared to the US and UK populations, but overall the similarities in gender attitudes stand out much more than the differences. The figure is based on samples that include both men and women, but interestingly there is very little difference in these gender attitudes between men and women. Overall and in contrast to common wisdom, the evidence presented in Figures 1 and 2 raises doubts about the degree to which Scandinavian countries are

⁵As we describe below, the way we measure hours worked means that we understate somewhat the gender hours gap and by implication overstate the gender wage rate gap. That is, the decomposition in Figure 1B of the earnings gap into the underlying margins is tilted somewhat from hours worked to wage rates.

positive outliers in terms of gender equality in the labor market.

The policy environment in Denmark is one which combines large tax-transfer distortions (which may affect the gender gap due to differential labor supply elasticities between men and women) and generous family policies intended to support female labor supply. As shown by Kleven (2014), the effective tax rate on labor earnings is exceptionally large in Denmark, but so are the implicit subsidies to labor supply through publicly provided child care and public spending on other goods that are complementary to working (transportation, elder care, education, etc.). Over the period we consider, public child care is universally provided at a heavily subsidized price from around 6-12 months after birth. Until the child reaches the age where public child care becomes available, there is job-protected and paid maternity and parental leave. Up until 2001, parents were offered 14 weeks of maternity leave followed by 10 weeks of parental leave to be shared between the mother and father. Since 2002 this has been extended to 18 weeks of maternity leave and 32 weeks of parental leave. Hence, throughout the period we consider, parents were covered first by paid leave and then by public child care, with no gap between the two.

2.2 Data

The analysis is based on administrative data for the full population in Denmark between 1980-2013. For the study of intergenerational transmission we exploit additional administrative data going back to 1964. The Danish data combines several administrative registers (linked at the individual level via personal identification numbers) and contains rich information on children, earnings, labor supply, occupation, firm, education, and many other variables. Crucially, the data allows us to link family members, generations, and workers to firms.

Our main event study analysis is based on *first* child births where the parents are observed every year between 5 years before having a child and 10 years after. We are thus focusing on first child births between 1985-2003 where the parents are known, alive and reside in Denmark throughout a 15-year window around the birth. We do not impose restrictions on the relationship status of the parents: we include all individuals who have a child in a given year and follow them through the 15-year window whether or not they are married, cohabiting, separated, divorced, or have not yet formed a couple in any given year. This leaves us with a core estimation sample of around 470,000 births or 15,040,000 individual-year observations.

We estimate child penalties in earnings, labor force participation, hours worked, and wage rates (earnings/hours worked for those who are working). Our ability to estimate child penalties

in hours worked and wage rates using sharp event studies relies on a unique administrative and third-party reported measure of hours worked that is available for the full population. This measure comes from a mandated pension scheme introduced in 1964 — *Arbejdsmarkedets Tillægspension* (ATP) — that require all employers to contribute on behalf of their employees based on individual hours worked.⁶ The pension contribution is a function of hours worked in discrete steps, namely four bins of weekly hours (0-8, 9-17, 18-26, 27-) for someone paid weekly or four bins of monthly hours (0-38, 39-77, 78-116, 117-) for someone paid monthly, with the latter being much more common. Hence the annual pension contribution for someone paid monthly depends on $\sum_{i=1}^{12} h_i$ where monthly hours h_i has 4 steps, which gives an annual hours measure in 37 steps ($= 4 \times 12 - 12 + 1$). Our measure of the wage rate is defined as earnings divided by this ATP hours measure.

Because the ATP hours measure is capped, it does not capture marginal hours adjustments for those working every month of the year in the highest hours bin. For a given child penalty in earnings, this will make us underestimate the penalty in hours worked and correspondingly overestimate the penalty in wage rates. The hours measure does capture larger labor supply adjustments such as switches to different levels of part-time work and work interruptions within the year, which are important adjustments for women with children. The key advantage of our measure is that it is precisely measured for the full population over a very long time period, unlike labor market surveys that have considerable measurement error and small samples.

3 Impacts of Children

3.1 Event Study Methodology

The ideal experiment for studying the impact of children would be to randomize fertility. Absent such an experiment, researchers have proposed instruments for the number of children such as twin births (Rosenzweig & Wolpin 1980; Bronars & Grogger 1994) and sibling sex mix (Angrist & Evans 1998). Such instruments can provide insights on the local effect of a second child or a third child, but they cannot provide estimates of the total effect of children — and in particular of the first child — in the population. This limits the usefulness of such approaches for our agenda, which is to understand the full implications of children for gender inequality. To investigate this question, we adopt an event study approach based on sharp changes around the birth of the first child for mothers relative to fathers. Although fertility choices are not exogenous, the event of having

⁶The scheme also covers self-employed individuals who contribute on their own behalf.

a first child generates sharp changes in labor market outcomes that are arguably orthogonal to unobserved determinants of those outcomes as they should evolve smoothly over time. The event study approach has the additional advantages of tracing out the full dynamic trajectory of the effects, and of being very precise as it exploits individual-level variation in the timing of first births. We spell out the identification assumptions of our approach and investigate their validity in detail in sections 3.3 and 3.4, but here we start by describing the baseline specification and presenting a set of striking results.

For each parent in the data we denote by $t = 0$ the year in which the individual has his/her first child and index all years relative to that year. Our baseline specification considers a balanced panel of parents who we observe every year between 5 years before having their first child and 10 years after, and so event time t runs from -5 to $+10$. To investigate the very long run, we also present results for a balanced panel of parents who we observe up to 20 years after the birth of their first child. We study the evolution of wide set of labor market outcomes as a function of event time. Specifically, denoting by Y_{ist}^g the outcome of interest for individual i of gender g in year s and at event time t , we run the following regression separately for men and women

$$Y_{ist}^g = \sum_{j \neq -1} \alpha_j^g \cdot \mathbf{I}[j = t] + \sum_k \beta_k^g \cdot \mathbf{I}[k = \text{age}_{is}] + \sum_y \gamma_y^g \cdot \mathbf{I}[y = s] + \nu_{ist}^g, \quad (1)$$

where we include a full set of event time dummies (first term on the right-hand side), age dummies (second term) and year dummies (third term). We omit the event time dummy at $t = -1$, implying that the event time coefficients measure the impact of children relative to the year just before the first child birth. If we did not include age and year dummies, the estimated event coefficients $\hat{\alpha}_t^g$ would correspond simply to the mean value of the outcome at event time t , relative to the year before birth. By including a full set of age dummies we control non-parametrically for underlying life-cycle trends, and by including a full set of year dummies we control non-parametrically for time trends such as wage inflation and business cycles. We are able to identify the effects of all three sets of dummies because, conditional on age and year, there is variation in event time driven by variation in the age at which individuals have their first child. While the age and year dummies have no large effect on any of our results, the inclusion of age dummies improves the comparison between men and women as women are on average a couple of years younger than men when having their first child.

We specify equation (1) in levels rather than in logs to be able to keep the zeros in the data

(due to non-participation). We convert the estimated level effects into percentages by calculating $P_t^g \equiv \hat{\alpha}_t^g / \text{E} [\tilde{Y}_{ist}^g | t]$ where \tilde{Y}_{ist}^g is the predicted outcome when omitting the contribution of the event dummies, i.e. $\tilde{Y}_{ist}^g \equiv \sum_k \hat{\beta}_k^g \cdot \mathbf{I}[k = \text{age}_{is}] + \sum_y \hat{\gamma}_y^g \cdot \mathbf{I}[y = s]$. Hence, P_t^g captures the year- t effect of children as a percentage of the counterfactual outcome absent children. Estimating percentage effects based on a level-specification (rather than a log-specification) may raise the concern that the estimates are mainly driven by the effects at the top of the distribution, especially when considering earnings as the outcome. We will present quantile regressions of (1) that rule out this concern.

Having estimated the impacts of children on women and men separately, we define a “child penalty” on women relative to men at event time t as

$$P_t \equiv \frac{\hat{\alpha}_t^m - \hat{\alpha}_t^w}{\text{E} [\tilde{Y}_{ist}^w | t]}. \quad (2)$$

This child penalty measures the percentage by which women are falling behind men due to children at event time t . Two points are worth noting about this measure. First, while the identification of short-term child penalties (say P_0 or P_1) rely primarily on a smoothness assumption common to all event studies, the identification of long-term penalties (say P_{10} or P_{20}) requires stronger assumptions and may call for the use of a control group (such as men and women who do not have children) or an instrument. We verify our event study estimates using these alternative strategies later. Second, while our approach is based on the event of having the first child, long-run child penalties will include the impact of children born after to the first one, unless of course we condition the sample on having only one child in total. Hence, long-run child penalties have the potential to capture the total effect of children on gender inequality.

3.2 Estimating the Impacts of Children

In this section we present estimates of the impacts of children on the trajectory of a wide range of labor market outcomes for men and women. We start by showing impacts on earnings, labor supply and wage rates, and then turn to the underlying anatomy of these impacts by showing how occupation, firm, and sector choices by men and women respond to children.

Earnings, Hours, Participation, and Wage Rates:

Figure 3 plots the gender-specific impacts of children P_t^m, P_t^w across event time. As defined above, these are outcomes at event time t relative to the year before the first child birth ($t = -1$), having

controlled non-parametrically for age and time trends. The figure includes 95% confidence bands around the event coefficients, although these are not always clearly visible due to the high precision of the administrative data. Panel A starts by showing total earnings before taxes and transfers. We see that, once life-cycle and time trends are taken out, the earnings of men and women evolve in a strikingly parallel way until they become parents. But at the precise moment the first child arrives, the earnings paths of men and women diverge: women experience an immediate drop in gross earnings of almost 30%, while men experience no visible change in their earnings. Importantly, in the years following the initial drop, the earnings of women never converge back to their original level. Ten years after the birth of a first child, female earnings have plateaued around 20% below its level just before child birth, whereas male earnings are essentially unaffected by children. As shown in the figure, the long-run child penalty in the earnings of women relative to men 10 years after the first child (i.e., P_{10} defined in equation 2) is equal to 19.4%.

These earnings impacts can come from three margins: hours worked, labor force participation, and the wage rate. Panels B-D of Figure 3 show that all three margins are in play. For each of these outcomes, the trajectories of men and women are almost exactly parallel prior to having children and then diverge sharply immediately after the arrival of the first child. The gaps that open up in labor supply and wage rates are driven entirely by negative impacts on women, while men are unaffected. For all outcomes, the gender gaps are very stable from around event time 3 and women show no sign of labor market recovery 10 years after the first child. Interestingly, the estimated long-run penalties in hours, participation and wage rates are similar in magnitude, implying that these margins are roughly equally important for the earnings impacts.⁷ While our primary goal in this paper is to understand gender inequality, we note that the event studies of working hours and participation contribute to the large literature on family labor supply and fertility by providing clean estimates of the labor supply implications of the first child.

Our baseline specification is based on first child births between 1985-2003 and an event study window that includes 10 years after birth. It is of course interesting to study how the career trajectories of mothers and fathers evolve in the very long run, which is feasible to do with our data. Hence, in Figure 4 we expand the event study horizon to include 20 years after the birth of the first child. For this exercise, we consider all parents who have their first child after 1985 and keep them in the sample for the longest possible period. This sample is therefore unbalanced across event

⁷The child penalties in panels B-D of Figure 3 are unconditional penalties: when estimating the effect of parenthood on one particular margin, we are not controlling for the other two margins in the regression. This explains why the long-run penalties on the three margins do not sum up to the overall earnings penalty.

time; only the parents of birth cohorts 1985-1993 are observed all the way to event time 20. The figure is otherwise similar to Figure 3 and shows results for earnings, hours worked, participation, and wage rates. The figure shows how strikingly persistent the effects of children are. In fact, the earnings impact 20 years after child birth is almost the same as the impact 10 years after. The only qualitative difference that emerges from considering the very long run is that hours worked do eventually begin to converge, while at the same time wage rates keep diverging. The combination of the narrowing hours gap and the widening wage rate gap produces a constant earnings gap.

We now consider three extensions and robustness checks of the baseline analysis. First, while our event study approach uses the birth of the first child, the evidence presented so far is based on the full population of parents, irrespective of the total number of children they end up having. This implies that the dynamic patterns we document include the effects of children born after the first one, and the estimated long-run impacts should be interpreted as capturing the total impact of all children. To explore the implications of multiple children, we replicate the earnings event study from Figure 3A in subsamples that condition on the total number of children the woman ends up having (1, 2, 3 or 4 children). The results are presented in appendix Figure A.I. The impact of children is very sharp in all four family types and the long-run child penalty increases by roughly 10 percentage points per child. The short-run impact of the first child is about the same across families with 1, 2 or 3 children (a gender gap of 25-30% at event times 0 and 1), and only gradually do the different family sizes diverge as more children arrive in the larger families. In families with 4 children, on the other hand, the impact is larger from the outset, suggesting that these families anticipate that they will have many children and divide gender roles accordingly right after the first one.⁸

Second, the fact that we estimate percentage effects based on a level-specification (rather than a log-specification) implies that the estimates put more weight at the top of the distribution. If there is lots of heterogeneity in the impacts of children across the distribution, the mean impacts can be very different from the impacts further down the distribution. To address this concern, Figure A.II shows median impacts of children on earnings and total hours worked. These quantile regressions are based on a 1/7 subsample, which makes the confidence bands somewhat larger. The results in the figure show that the median impacts are roughly similar to mean impacts, ruling out the concern that that our results are representative mainly of the top of the distribution.

⁸Note that the event study graph for families with two children in Panel B of Figure A.I looks very similar to the graph for all families in Figure 3. This is natural given that the average completed fertility, conditional on having children, is close to two in Denmark.

Finally, in the event graphs presented so far, the drop in earnings and labor supply in event year 0 is not much larger than the drop in subsequent event years. While this may seem surprising, note that the use of calendar-year measures of earnings and labor supply create attenuation bias in the first-year dip compared to a continuous time representation: as women give birth some time during year 0, some of the earnings and hours in calendar-year 0 were realized prior to child birth. To investigate this point, we reproduce Figure 3 on a sample restricted to child births in January for which the definition of calendar years and event years coincide. The results are shown in Figure A.III from which three insights emerge. First, when focusing on January births alone we do see a pronounced dip in event year 0 as one would expect. This dip reflects the extra time taken out of the labor market immediately following delivery. Second, focusing on January births also reveal a small drop in labor market outcomes in event year -1, which can be explained by sick leave and parental leave during pregnancy.⁹ Third, the long-run penalty estimates are a bit smaller for January births than for all births. This is because these penalties are measured relative to event year -1, which includes a significant pregnancy effect for January births and therefore biases down the January penalties. If we were to measure penalties relative to event time -2, then the January penalties would not be smaller.

Occupation, Sector, and Firm:

We have seen that motherhood is associated with large and persistent effects on earnings driven in roughly equal proportions by participation, hours of work, and wage rates. These empirical patterns, and especially the large wage rate effects, beg the question of what are the underlying mechanisms? A possibility is that women, once they become mothers, make career choices in qualitative dimensions (occupations, sectors, firms, etc.) that favor family amenities over pecuniary rewards. The importance of such effects has been much discussed (see e.g., Goldin 2014), and there is plenty of cross-sectional evidence showing that women with children work in different occupations and industries than women without children or men. Still, we are not aware of any direct evidence that these qualitative dimensions respond to children. We provide such evidence in this section.

The results are presented in Figure 5, which is based on the same event study approach as the one used above. Panel A considers occupational rank in five levels: unskilled labor, skilled labor, white-collar work (low level), white-collar work (high level), and top manager. This ordering of occupations is consistent with an ordering based on average earnings or average wage rates in

⁹In Denmark, women are eligible for parental leave during the last four weeks of pregnancy.

each occupation. This panel shows that men and women are on identical trends in terms of their occupational rank prior to becoming parents (controlling non-parametrically for age effects), but that women start falling behind men soon after parenthood. Note that the occupation graphs for men and women begin to diverge in event year 1 rather than in event year 0. This is natural given that women are giving birth during year 0, so that this year consists partly of a pre-birth period and partly of a period covered by job-protected parental leave. Hence, women do not have a strong incentive to change occupation within year 0, but can wait until year 1 when they return to work. Panel B also explores occupational rank, but focuses specifically on the probability of being top manager. It is in general harder to uncover an impact of children on this margin, because relatively few individuals have risen to the managerial level prior to becoming parents. Nevertheless, the graph suggests that parenthood has a negative effect on women's prospects of becoming managers. While the male and female trends are not perfectly similar prior to child birth, they do begin to diverge at a much faster pace following child birth.

The bottom panels turn to the choice of work environment and in particular its family friendliness. We first consider the link between children and the decision to work in the Danish public sector, which has a long tradition of focusing on working conditions rather than on wages. This includes flexible working hours, leave days when having sick children, and a favorable view on long parental leaves (see [Nielsen et al. 2004](#) for a detailed description). It is therefore natural to expect that mothers would be induced to move into the public sector, a hypothesis that is clearly confirmed in Panel C. Men and women are on very similar pre-child trends in terms of their probabilities of working in the public sector, but begin to diverge soon after having a child. Ten years after child birth, women have a 10pp higher probability of public sector employment than men, relative to the year before child birth. As with occupation, the divergence mainly starts in year 1 rather than in year 0, i.e. when women return to work after their parental leave.

Finally, Panel D considers a proxy for the family friendliness of a work environment that also encompasses heterogeneity across firms in the private sector. Here we take advantage of our employer-employee matched data by defining a firm's "family friendliness" as the share of women with children below 15 years of age in the firm's workforce (excluding the considered woman's own first child when it arrives). The advantage of this kind of revealed-preference measure of family friendliness is that it can capture many aspects of an otherwise complex, multidimensional concept. Having a larger share of female employees with young children may reflect that the firm offers more generous maternity leave policies, that managers are tolerant of sick days, that co-

workers are a source of understanding and advice, and that there is workplace flexibility in various dimensions. Because our measure of family friendliness is negatively related to wage rates, if women move into more family-friendly firms following motherhood, this helps explain the wage rate penalties documented above.

The specific outcome shown in Panel D is the percentile rank in the distribution of firm family friendliness for men and women, respectively, relative to the year before child birth. Although men and women are not on identical trends prior to birth (the female trend is steeper), there is a very clear break in the relative trends around the moment of having a child. Women begin to move into family-friendly firms at a much higher pace in the years following child birth, whereas the male trend is completely unaffected by child birth. The female trend is increasing somewhat already in event year -1, consistent with a small anticipation effect of motherhood. Taking the differential pre-trend into account, we estimate a long-run effect of children on the percentile rank of the distribution of firm family friendliness for women relative to men equal to 4.36. We note that this effect is driven by mothers *switching* firms as opposed to within-firm increases in family friendliness: there is no effect of children on family friendliness when we condition on staying in the same firm over time.^{10,11}

To conclude, we have shown that women’s career trajectories change sharply due to children, creating substantial gender inequality in a range of quantitative and qualitative dimensions. The results demonstrate the difficulties that women continue to face in trying to balance career and family, and are broadly consistent with the arguments by [Goldin \(2014\)](#).

3.3 Conceptual Framework and Identification

In this section we set out a simple conceptual framework to clarify what is being estimated in the event studies. We denote by $\mathbf{k}_i = (0, \dots, k_{it}, \dots, k_{iT})$ the anticipated lifetime path of fertility for individual i : the individual starts life with zero children, has k_{it} children at event time t , and ends up with k_{iT} children over the lifetime. Earnings at time t are chosen based on the number of

¹⁰In theory, within-firm increases in family friendliness could occur because of externalities in child birth across co-workers or because of increased hiring of women with children in response to employee births, but the data does not provide support for such stories.

¹¹While we have thus shown that switching to more family-friendly firms in the private or public sectors is one of the mechanisms driving the child penalty in earnings, a related question is whether being in a family-friendly firm prior to child birth serves to moderate the child penalty. In other words, what is the pattern of heterogeneity in penalties with respect to pre-parenthood firm and sector decisions? An earlier version of the paper included a detailed heterogeneity analysis, which showed that having a more family-friendly employer by the time of child birth is indeed associated with a significantly smaller penalty. This holds even after including rich controls for other factors that are correlated with firm and sector choices.

children present at time t as well as the anticipated lifetime path of fertility. Specifically,

$$\begin{aligned} Y_{it} &= F(k_{it}, \mathbf{x}_{it}, \mathbf{z}_{it}) \\ &= F(k_{it}, \mathbf{x}(k_{it}, \mathbf{k}_i, \mathbf{z}_{it}), \mathbf{z}_{it}), \end{aligned} \quad (3)$$

where Y_{it} is earnings, $\mathbf{x}_{it} = \mathbf{x}(k_{it}, \mathbf{k}_i, \mathbf{z}_{it})$ is a set of earnings determinants that are chosen based on children, and \mathbf{z}_{it} is a set of earnings determinants that do not depend on children. Compared to the empirical specification (3), we simplify notation by leaving out indexation of gender and calendar time. The elements of \mathbf{x}_{it} include variables such as hours worked, occupation, sector and firm — variables which we have seen respond to children in the event studies — while the elements of \mathbf{z}_{it} include factors such as age, ability and preferences. Hence, in this framework earnings may respond directly to children conditional on choices (e.g. the impact of being tired or distracted at work) and indirectly through labor market choices \mathbf{x}_{it} (e.g. the impact of switching to a lower-paying, but more family-friendly firm). Furthermore, we allow labor market choices \mathbf{x}_{it} to respond both to the contemporaneous number of children k_{it} and to the entire path of past and future fertility. The latter effect captures for example that some women may take less education or opt for family-friendly career tracks knowing that they will eventually have many children.

While we do not specify the demand for children, we make the assumption that children k_{it} are exogenous to the outcome variable Y_{it} conditional on the set of underlying determinants \mathbf{z}_{it} . The assumption that “the event” (in our case, child birth) is not determined by the outcome variable is fundamental to any event study analysis. The graphical evidence presented above lends support to this assumption: there is no indication that outcomes respond *prior* to child birth (or prior to pregnancy as discussed above); the sharp breaks in career trajectories always occur *just after* having children.

This framework allows for two conceptually different effects of children on earnings. One is a *pre-child effect* of future children, conditional on the current number of children k_{it} , which operates through the dependence of labor market choices \mathbf{x}_{it} on anticipated lifetime fertility \mathbf{k}_i . The other is a *post-child effect* of current children, conditional on anticipated lifetime fertility \mathbf{k}_i , which operates through both the direct effect of k_{it} and the effect of k_{it} on labor market choices \mathbf{x}_{it} . An obvious but important point is that the event studies cannot capture pre-child effects — these are incorporated in the pre-event levels that are differenced out — and is designed to identify only post-child effects. If women are investing less in education and career in anticipation of

motherhood (as the child penalty sharply reduces the return to such investments), then the pre-child effect on female earnings is negative and the event study provides a lower bound on the total effect.¹²

Under what conditions do the event studies correctly identify the post-child impacts? It is important to distinguish between short-run and long-run impacts. The short-run impact is estimated by comparing event times *just* before and after time zero. Denoting these event times by t_- , t_+ and using equation (3), the short-run event study estimates capture

$$E[Y_{it_+} - Y_{it_-}] = E[F(1, \mathbf{x}(1, \mathbf{k}_i, \mathbf{z}_{it_+}), \mathbf{z}_{it_+})] - E[F(0, \mathbf{x}(0, \mathbf{k}_i, \mathbf{z}_{it_-}), \mathbf{z}_{it_-})], \quad (4)$$

when we do not directly control for elements for \mathbf{z}_{it} through for example age and year dummies. Assuming smoothness of the average non-child earnings path, i.e. $E[F(0, \mathbf{x}(0, \mathbf{k}_i, \mathbf{z}_{it_-}), \mathbf{z}_{it_-})] \approx E[F(0, \mathbf{x}(0, \mathbf{k}_i, \mathbf{z}_{it_+}), \mathbf{z}_{it_+})]$, equation (4) identifies the short-run effect of the first child conditional on \mathbf{z}_{it_+} . With direct controls for \mathbf{z}_{it} , the smoothness assumption can be relaxed.

The long-run impact is obtained by considering an event time t_{++} long after time zero, i.e.

$$E[Y_{it_{++}} - Y_{it_-}] = E[F(k_{iT}, \mathbf{x}(k_{iT}, \mathbf{k}_i, \mathbf{z}_{it_{++}}), \mathbf{z}_{it_{++}})] - E[F(0, \mathbf{x}(0, \mathbf{k}_i, \mathbf{z}_{it_-}), \mathbf{z}_{it_-})]. \quad (5)$$

There are two differences between this impact measure and the previous one. The first difference is that the long-run impact captures the effect of total lifetime fertility k_{iT} as opposed to the effect of only the first child. The second difference is that the smoothness assumption is no longer sufficient for identification as we can still have large changes in non-child earnings components over a long event time window. Hence, if we are not fully controlling for \mathbf{z}_{it} , then the long-run child penalty may be a biased estimate of the true post-child impact. Allowing for non-parametric age and year controls as we do in specification (1) may go a long way in alleviating this problem, but we cannot be certain that there is no remaining bias. There are two potential solutions to the problem. One is to use a control group — naturally women and men who never have children — to account for the non-child earnings trend in a difference-in-differences design. The other is to leverage an instrument for child birth within our event study approach. In the next section we consider both approaches.

¹²On the other hand, if women are engaging in intertemporal substitution of work effort around the event of having a child, then the pre-child effect could be positive. However, our event graphs feature very stable pre-trends that are identical for men and women, indicating that no significant intertemporal substitution is taking place.

3.4 Identification Checks

Using Men and Women Without Children as Controls:

We first consider a difference-in-differences event study design that uses men and women who never have children as controls. The design is based on assigning placebo births to individuals who never have children, drawing from the observed distribution of age at first child among those who do have children (within cells of cohort and education).

The full details of our design are described in Appendix A.1, but here we outline two technical issues that arise when defining the control group and assigning placebo births. First, individuals observed without children include two types: those who will never have children and those who have not had children yet. The first group is the cleanest possible control group, but we face a truncation issue in identifying them. Taking age 40 as the latest age at which people have their first child (as only small fractions of both men and women have their first child after that age), the fertility of cohorts born after 1973 is truncated as they are younger than 40 when we last see them in 2013. Hence, for individuals observed without children in the later cohorts, we select those who are most likely never to have children based on a linear probability model of zero lifetime fertility ($k_{iT} = 0$) as a function of observables, estimated on the (non-truncated) cohorts born between 1955–1973. In each of the later cohorts, the number of individuals we select as having zero lifetime fertility is such that the probability of never having children is the same after 1973 as the average between 1955–1973 (as this probability has been quite stable during this time). Our control group consists of these selected individuals from the post-1973 cohorts along with all individuals without children from earlier cohorts.¹³

Second, we need to allocate placebo births to those in our control group. Here we also have to distinguish between truncated cohorts born after 1973 and non-truncated cohorts born before that time. For the older cohorts, the distribution of age at first child A is approximated by a log-normal distribution within cells of birth cohort c and education e . That is, we assume $A_{c,e} \sim \mathcal{LN}(\hat{\mu}_{c,e}, \hat{\sigma}_{c,e}^2)$ where the mean $\hat{\mu}_{c,e}$ and variance $\hat{\sigma}_{c,e}^2$ are obtained from the actual distributions within each cohort-education cell. Individuals in older cohorts without children gets a random draw from this distribution. For the younger cohorts, we draw a random age at first child from $\mathcal{LN}(\tilde{\mu}_{c,e}, \hat{\sigma}_{c,e}^2)$ where the mean $\tilde{\mu}_{c,e}$ is the predicted average age at first child obtained by estimating a linear trend on the older cohorts. That is, consistent with the stylized pattern observed for the older cohorts,

¹³In practice, because our event studies are based on first child births between 1985-2003, the later cohorts represent a small fraction of the estimation sample as most of them will be assigned a placebo birth after 2003.

we allow for an upward linear drift in the age at first child while keeping the variance constant.

With these preliminaries, we are able to implement event studies that compare our treatment group (a balanced panel of those who have their first child between 1985–2003 and are observed in a 15-year window around the first child birth) to our control group (a balanced panel of those who never have children, but have been assigned a placebo birth between 1985–2003 and are observed in a 15-year window around the placebo). The impact of children will be estimated as a difference-in-differences, i.e.

$$E [Y_{i,t>0} - Y_{i,t<0} | k_{iT} > 0] - E [Y_{i,t>0} - Y_{i,t<0} | k_{iT} = 0]. \quad (6)$$

The identification assumption is a standard parallel trends assumption, which in the notation established above implies $E [\Delta F (0, \mathbf{x} (0, \mathbf{k}_i, \mathbf{z}_{it}), \mathbf{z}_{it}) | k_{iT} > 0] = E [\Delta F (0, \mathbf{x} (0, \mathbf{0}, \mathbf{z}_{it}), \mathbf{z}_{it}) | k_{iT} = 0]$. Given the parallel trends assumption — the validity of which we can verify from the pre-trends — it is not necessary to introduce controls for \mathbf{z}_{it} . Hence, we leave out the age and year dummies in this design and simply plot the raw event coefficients.

Figure 6 shows the earnings impacts of children in this difference-in-differences design. Panel A shows women while Panel B shows men. The event studies are very sharp and confirms the key qualitative findings from the baseline specification. Women with children and women without children are on identical pre-trends, diverge sharply at the time of the first child birth, and the impact is very stable over time. The impact of children 10 years after is about 23%, slightly larger than the impact of about 21% obtained from the baseline specification as in Figure 3A. The baseline graph already suggested that we were slightly underestimating the career cost of children on women: it showed a weak upward pre-trend for women (compared to men) that the child penalty estimate did not take into account. Using women without children as a control group accounts for this trend. Finally, and consistent with our previous conclusions, Panel B shows that men are completely unaffected by children in terms of their earnings.

Using Sibling Sex Mix and Twin Births as Instruments:

As a second identification check, we compare our event study approach to an IV approach using the sex mix of the first two children as an instrument for having a third child (Angrist & Evans 1998). The idea is that parental preferences for variety make it more likely to have a third child when the first two have the same sex, while the children’s sex should have no independent im-

impact on labor market outcomes and thus satisfy the exclusion restriction.¹⁴ As the sibling sex mix instrument gives the local average treatment effect of a third child, we have to modify our event study approach to also provide the local impact of the third child in order to compare the two approaches.

We consider the following event study specification for estimating the effect of a third child:

$$\begin{aligned}
Y_{istt'} &= \sum_{j \neq -1} \alpha_j \cdot \mathbf{I}[j = t] + \sum_k \beta_k \cdot \mathbf{I}[k = \text{age}_{is}] + \sum_y \gamma_y \cdot \mathbf{I}[y = s] \\
&+ \sum_{n \neq -1} \delta_n \cdot \mathbf{I}[n = t'] + \eta_i + \nu_{istt'},
\end{aligned} \tag{7}$$

where the index t still denotes event time with respect to the first child, while the new index t' denotes event time with respect to the third child. In this specification, the top row is exactly the same as the baseline specification (1): it gives the effect of the first child, controlling for a full set of age and year dummies. The bottom row includes two new terms: a set of event time dummies around the birth of the third child (omitting $t' = -1$) and an individual fixed effect η_i .

We run the specification on a sample of women who had their first child between 1985-2003 (as before) and who have completed fertility of two or three when we last observe them in 2013.¹⁵ In order to capture the long-run effect of a third child, we expand from the previously balanced panel that included a 10-year window after the first child birth to an unbalanced panel that includes the longest possible window for each individual. For example, a woman who had her first child in 1985 and is observed until 2013 is included in the sample with event times up to $t = 28$ with respect to that child, allowing for the longest possible period after the third child.

Two points are worth noting about specification (7). First, the reason we include an individual fixed effect η_i is because we are estimating the effect of a third child in a sample that includes women who only have two children. Without the fixed effect, the event time dummies δ_n would absorb both the dynamic effect of a third child *and* any fixed differences between the types of women who have two vs three children. Second, although we are interested only in the effect of the third child in this specification, we control for past child dynamics through the event time

¹⁴The validity of this exclusion restriction can be verified based on our event study approach: we find no differences in the impacts of the first child depending on whether it is a boy or a girl, suggesting that child gender is not important for parental labor market outcomes. While the exclusion restriction is thus compelling, there might be a problem with the assumption of no defiers underlying LATE. Some parents may have preferences for a specific sex — typically boys — and are therefore less likely to have a third child if they start out with two boys. The presence of defiers due to boy-bias is arguably less of an issue in Denmark than in more traditional societies.

¹⁵As we do not include men here, we have dropped the superscript g in specification (7).

dummies around the first child birth. This is because it may matter for earnings dynamics whether the three children come in quick succession or are spread out over a longer time interval. The reason why we can separately identify event time dummies for the first child and the third child, while simultaneously controlling for a full set of age and year dummies, is that there is enough independent variation in the ages at which women have their first and third child, respectively.

The IV-specification is the same as equation (7), except that we drop the individual fixed effect and instead instrument the event time dummies for the third child birth. Specifically, we instrument each dummy $\mathbf{I}[n = t']$ by the interaction $\mathbf{I}[n = t'] \times \mathbf{I}[\text{same sex siblings}]$, which takes the value of one when the woman is at event time t' with respect to the third child and her first two children have the same sex. In contrast to previous implementations of such IV approaches, this specification traces out the full dynamic pattern of the effects of the third child.

The results are presented in Figure 7, which shows the earnings impacts of a third child obtained from the event study specification (black series) and the IV specification (red series) as a function of years since the birth of the third child. Analogous to the previous graphs, the statistic shown here is $P_{t'}^w \equiv \hat{\delta}_{t'}^w / E[\tilde{Y}_{istt'}^w | t']$ where $\tilde{Y}_{istt'}^w$ is the counterfactual outcome absent the effect of the third child, but not the other children, i.e. $\tilde{Y}_{istt'}^w \equiv \sum_j \hat{\alpha}_j \cdot \mathbf{I}[j = t] + \sum_k \hat{\beta}_k \cdot \mathbf{I}[k = \text{age}_{is}] + \sum_y \hat{\gamma}_y \cdot \mathbf{I}[y = s] + \hat{\eta}_i$.

The following key insights emerge from the figure. First and foremost, the event study estimates and the IV estimates are almost perfectly aligned through event time, providing strong support for our empirical approach. The fact that the pre-event coefficients are very similar (and close to zero) for the two approaches suggests that anticipation effects of children are not important. The fact that the post-event coefficients are very similar shows that the impacts on women who have a third child because of the sex mix of the first two (the IV-compliers) are the same as the impacts on all treated women. Second, the short-run effect of a third child is similar to the short-run effect of the first child, an earnings reduction of 20-30%. Third, the long-run effect of a third child is about 5%. This is half the size of the long-run effect of the first child among those who only have one child (see Figure A.I), which suggests that the marginal effect of children is declining in the number of children. This last conclusion is only suggestive as it is based on a cross-sectional comparison between samples of women with different lifetime fertility.

Besides sibling sex mix, a number of studies have considered the occurrence of twins at first birth (e.g. [Rosenzweig & Wolpin 1980](#); [Bronars & Grogger 1994](#)) or at second birth (e.g. [Angrist & Evans 1998](#)) as instruments for children. As an alternative to the strategies discussed above, we

may consider twins at second birth as an instrument for the third child. Figure 7 therefore compares the average impacts obtained from the event study and same-sex IV to the average impact obtained from a twin IV.¹⁶ We see that the twin estimate is considerably smaller than the event study and same-sex estimates. Why this difference? A natural interpretation is that twins represent a more efficient child production technology (e.g. due to economies of scale) and therefore impose smaller penalties on women. This implies that even though the occurrence of twin births is an exogenous event, it may not be a valid instrument for having an extra child in the standard sequential way: it does not satisfy the exclusion restriction if it changes aspects of the child care technology that have their own direct impact on earnings.¹⁷ Interestingly, our results are consistent with those of Angrist & Evans (1998), who also find smaller effects when using the twin instrument than when using the same-sex instrument.¹⁸

4 Decomposing Gender Inequality Over Time

4.1 Dynamic Decomposition Framework

In this section we decompose gender inequality into what can be attributed to children and what can be attributed to other factors, showing how this composition has evolved over a long time period. We adopt a standard Oaxaca-Blinder decomposition framework, but innovate on existing gender gap decompositions by leveraging the event study variation in order to estimate the impact of children on gender inequality. In particular, while standard decomposition approaches in the gender literature (see e.g., Blau & Kahn 2016) have been based on cross-sectional variation in education and labor market variables *not including children*, our decomposition approach focuses precisely on children and exploits within-person time variation in the presence of children. The goals of these two decomposition exercises are different: while the traditional goal has been to

¹⁶These average impacts are based on a simplified version of equation (7) in which the full set of event time dummies around the third child is replaced by a single dummy for having had the third child, i.e. $\mathbf{I} [t' \geq 0]$. For the IV-strategies, $\mathbf{I} [t' \geq 0]$ is instrumented by either $\mathbf{I} [\text{same sex siblings}]$ or by $\mathbf{I} [\text{twins in second birth}]$.

¹⁷To be clear, the exogeneity of twins implies that it does give the causal effect of a twin birth relative to a singular birth (the reduced-form impact of a twin birth dummy), but it does not give the effect of increasing the number of children through the standard sequential birth technology.

¹⁸Angrist & Evans (1998) argue that, besides economies of scale, the smaller size of the twin estimates could be driven by differences in the age of the third child: in cross-sectional comparisons, a third child triggered by twins will be older than a third child triggered by same sex siblings. Our dynamic IV approach allows us to separate the age and economies-of-scale hypotheses by estimating the impact of the third child at each event time after birth. Such an exercise reveals that the twin estimates are smaller conditional on event time, which cannot be explained by age and suggests that economies of scale are important.

estimate the wage gap between observationally equivalent males and females — the unexplained gap often interpreted as “discrimination”, although it could also reflect children — our goal is to delve into the unexplained part of traditional decomposition analysis by estimating the impact of children. The impact we estimate could operate through both the unexplained and explained parts of standard cross-sectional decomposition analyses, because those analyses include variables such as occupation, industry and experience, which represent some of the mechanisms responsible for the impact of children as we have seen above.

Two points are worth highlighting from the outset. First, provided that the child penalties are correctly identified based on the event-study variation — a point we analyzed in detail in the previous section — our decomposition into child-related gender inequality and residual gender inequality should be viewed as causal rather than purely correlational. Second, since child penalties by construction capture only the post-effects of actual fertility and not the pre-effects of anticipated fertility, residual gender inequality includes potential pre-effects of children. For example, education choices made prior to having children may reflect anticipated fertility.

We focus on gender inequality in earnings. To capture changes over time in the impact of children on gender inequality, we extend the baseline specification (1) to allow for year-specific coefficients on event time. Specifically we consider the following specification:

$$Y_{ist}^g = \sum_y \sum_{j \neq -1} \alpha_{yj}^g \cdot \mathbf{I}[j = t] \cdot \mathbf{I}[y = s] + \sum_k \beta_k^g X_{kis}^g + \nu_{ist}^g, \quad (8)$$

where we interact the event time dummies with year dummies in order to estimate year-specific event coefficients α_{yj}^g . We note that estimating event coefficients by calendar year s and event year t amounts to estimating event coefficients by birth cohort $c = s - t$. The second term on the right-hand side of equation (8) includes covariates indexed by k that may vary across individuals i and calendar time s . As in our baseline specification (1), we begin by including a full set of age dummies and year dummies in the set of covariates, but we will also consider an extended specification that includes a rich set of education dummies. An obvious but important point is that the X s should not include any earnings determinants that directly respond to the event of child birth as such covariates would bias the estimated event coefficients. That is, while it may be legitimate to control for education choices made prior to having children as we will do, controlling for factors such as occupation and firm choices (which we have seen respond to child birth) would lead to bias.

Defining the mean gender gap in year s as $\Delta_s \equiv \{E[Y_{ist}^m|s, m] - E[Y_{ist}^w|s, w]\} / E[Y_{ist}^m|s, m]$ and using specification (8), we can rearrange terms so as to obtain

$$\hat{\Delta}_s = \underbrace{\frac{E[P_{st}\tilde{Y}_{ist}^w|s, w]}{E[\tilde{Y}_{ist}^m|s, m]}}_{\text{child penalties}} + \underbrace{\frac{\sum_k (\hat{\beta}_k^m - \hat{\beta}_k^w) E[X_{kis}^m|s, m]}{E[\tilde{Y}_{ist}^m|s, m]}}_{\text{different returns to } Xs} + \underbrace{\frac{\sum_k \hat{\beta}_k^w \{E[X_{kis}^m|s, m] - E[X_{kis}^w|s, w]\}}{E[\tilde{Y}_{ist}^m|s, m]}}_{\text{differences in } Xs} \quad (9)$$

where $P_{st} \equiv \frac{\hat{\alpha}_{st}^m - \hat{\alpha}_{st}^w}{E[\tilde{Y}_{ist}^w|s, t, w]}$ is the child penalty at event time t in calendar year s , \tilde{Y}_{ist}^w is the predicted counterfactual earnings (i.e., absent children) for women, and \tilde{Y}_{ist}^m is the predicted actual earnings for men. The first term on the right-hand side captures the impact of child penalties on gender inequality, the second term captures the impact of different coefficients on non-child covariates (such as different returns to education), while the last term captures the impact of differences in non-child covariates (such as different levels of education). In the standard language of decomposition analysis (see e.g., Fortin *et al.* 2011), the first two terms represent “unexplained” effects (different regression coefficients) while the last term represents “explained effects” (different observables).¹⁹

To decompose gender inequality over the full period 1980-2013, we expand from the previously balanced panel of parents who have their first child between 1985-2003 to the full population of parents who have their first child between 1965-2013. For first child births after 1980, we observe parents for at least one year prior to child birth and are therefore able to estimate child penalties relative to event time -1 as we have done so far. Moreover, because we keep parents in the sample for the longest possible number of years (for example, parents who have their first child in 1981 are observed until event time $t = 32$, conditional on still being alive and residing in Denmark), we are able to estimate child penalties at all event times for the post-1980 birth cohorts. On the other hand, for first child births up until 1980, we observe only post-event years and are therefore unable to directly estimate child penalties associated with these births. For these birth cohorts we therefore rely on an extrapolation of the post-1980 penalties that we describe below. The reason we include the earlier cohorts in the estimation sample when running specification (8) is that they help estimating the effect of the non-child covariates (such as education).

¹⁹For simplification purposes, equation (8) does not include an explained effect of differences in children between men and women. We leave out this term in the equation (but not in the analysis), because it is always very close to zero. In fact, in a balanced panel of men and women who have children together, the fractions of men and women at each event time in each calendar year are by construction the same. As we describe below, our decomposition analysis is based on an unbalanced panel, and so the explained effect of children will be non-zero resulting from differential attrition of men and women due to deaths and migration. However, the explained effect of children due to such differential attrition is in practice tiny.

In Figure 8 we show earnings penalties by birth cohort obtained from specification (8). Panel A focuses on short-run penalties (event times 0-10) and Panel B focuses on long-run penalties (event times 11-20). Each panel includes a linear OLS fit in order to highlight the trend. We see that there is no trend in the short-run penalties, but a linear downward trend in the long-run penalties. We use these estimated trends to extrapolate child penalties to earlier birth cohorts: penalties between event times 0-10 are assumed to be constant at the average level of the later cohorts, while penalties between event times 11-20 are assumed to follow the linear trend estimated for the later cohorts.²⁰ With these extrapolations we obtain estimates of the child penalty P_{st} for every event time and every year between 1980-2013, allowing us to decompose gender inequality into child-related inequality and non-child inequality over the full period.

4.2 Decomposition Results

The first results of our decomposition analysis are shown in Figure 9A, which is based on the specification without education controls. The blue-shaded area shows child-related gender inequality (i.e., the effect of child penalties) while the grey-shaded area captures everything else. We see that the fraction of gender inequality that can be attributed to children has increased dramatically over time, from about 40% in 1980 to about 80% in 2013. This secular increase reflects a combination of two underlying changes: (i) child-related gender inequality in earnings has increased from 18% to almost 20%, and (ii) total gender inequality in earnings has fallen from 46% to 24%. To understand the first effect, note that although the *percentage* child penalty has fallen slightly over time (as shown in Figure 8), the penalty operates on a larger base due to the general increase in the earnings of women relative to men coming from the second effect. As shown in equation (9), the impact of children on gender inequality depends both on the size of child penalties and on the relative counterfactual earnings between men and women. Hence, during a period where non-child gender inequality is falling (for example, due to convergence in education levels) while child penalties are roughly constant or falling by less, there will be a tendency for child-related gender inequality to

²⁰For event times beyond 20 for the earlier cohorts, we assume that the penalty stays constant at its level in event time 20. Such a steady state assumption can be justified by the results presented in Figure 4A, which showed that earnings penalties are quite stable from around event time 10.

go up.^{21,22}

Our findings imply that, to a first approximation, the gender inequality that remains today is all about children. The fact that our approach misses the potential pre-effects of anticipated fertility is likely to reinforce the conclusion that gender inequality is *now* all about children, but it could change the conclusion that this was not also the case 30-40 years ago. It is conceivable that, while the impact of child penalties (post-effects) has increased over time, the importance of pre-effects has fallen. Indeed, at a time where gender norms were more traditional than they are today, it would be natural if women invested less in education and careers in anticipation of motherhood. In that case, more of the impact of children would be built into pre-event outcomes. In other words, there may have been a shift over time from pre-effects to post-effects of children that make us underestimate the importance of children in the beginning of the period.

While it is in general difficult to identify such pre-effects of children, suggestive evidence may be obtained by including education choices made prior to child birth in the specification. Hence, we include dummies for different education levels: primary school, secondary school, vocational training, short post-secondary school, bachelor's degree, and master's/phd degrees. The results are presented in Figure 9B, which shows child-related gender inequality in blue and education-related gender inequality in orange. The education-related part includes both the effect of different education levels (explained effect) and the effect of different education coefficients/returns (unexplained effect), with the latter being quantitatively more important. The explained effect of education is small from the start and turns negative in the early 2000s as women overtake men in terms of education levels. This is consistent with the disappearance of the college gap in the US (Goldin *et al.* 2006) and in most other high- and middle-income countries (Kleven & Landais 2016). As for the unexplained effect of education, this absorbs the effect of men and women choosing different education fields, conditional on level. In particular, women tend to choose “softer” fields than men (such as health care and teaching as opposed to construction and engineering) that are

²¹Besides these long-run changes in the composition of gender inequality, Figure 9 shows a short-run business cycle effect: during recessions (early 1980s, early 1990s, and 2008-09) overall gender inequality falls, but child-related inequality does not, and so the fraction of gender inequality due to children increases. In fact, during the global financial crisis in 2008-09 we estimate that child-related gender inequality was more than 90% of total inequality.

²²To understand how the increasing importance of children plays out over the life-cycle, Figure A.IV decomposes the age profiles of gender inequality in earnings for two specific years, 1985 and 2013. We see that there is relatively little earnings inequality in the tails, before having children and around the age of retirement. The largest earnings inequality occurs among those aged 30-55, around the time where most families have children living at home, and this is also where we estimate the impact of children — the difference between the dashed grey and solid grey lines in the figure — to be largest. While in 1985 there is a large difference between male earnings and counterfactual female earnings (i.e. absent children) throughout the life-cycle, in 2013 there is only a small difference between male earnings and counterfactual female earnings throughout the life-cycle.

not as highly remunerated, but may offer better family amenities.

Two main insights emerge from Panel B. First, the inclusion of education controls has only a small impact on the estimation of child-related gender inequality. It is still the case that child-related inequality is close to 80% of total inequality at the end of the period. The robustness of the decomposition analysis to education controls results from the event study variation we use: the child effect is identified from within-person time variation around child birth, while the education effect is obtained from cross-sectional variation and therefore does not absorb the child effect. Of course, the fact that the education effect is based on cross-sectional variation implies that, unlike the child effect, it is only correlational. Second, while the child-related gender gap has been growing over time, the education-related gender gap has been shrinking dramatically. Education-related inequality was almost as large as child-related inequality at the beginning of the period, but has almost disappeared over time. Our findings are consistent with a secular shift in the relative importance of post-effects of children (child penalties) and pre-effects of children (as proxied by education choices).

In this section we have tried to re-orient the traditional focus of gender gap decompositions by considering the impact of children and by exploiting sharp time variation instead of cross-sectional variation. Rather than studying the extent to which men and women receive unequal pay for equal work (the unexplained gap after controlling for human capital and job characteristics), we show that men and women receive unequal pay largely because of the unequal distribution of child care responsibilities. Even with perfectly equal pay for equal work, there would still be large gender inequality in earnings as equal work is not an option for the majority of women who are faced with the lion's share of child care responsibilities.

5 Intergenerational Transmission of Child Penalties

5.1 Background

Why are child penalties on women so large and persistent? There is clearly some role for traditional economic explanations related to comparative advantage in infant child care and the associated experience effects of birth-related work interruptions. At the same time, the fact that the highly gendered effects of children persist over the entire career path of parents — and considering that women are now more educated than men on average — suggests that there is more going on than traditional comparative advantage. As supporting evidence, Figure [A.V](#) shows that there is

essentially no heterogeneity in child penalties with respect to the relative education levels of the parents: women in the top quartile of the relative education distribution incur about the same penalties as those in the bottom quartile of the distribution. This is not the empirical pattern one would expect if the comparative advantage channel was very strong.

Another possible interpretation of our findings appeals to differential preferences or norms regarding the appropriate roles of women and men who have children. Indeed, we have shown in Figure 2 that views on the appropriate gender roles in families with children remain conservative in all countries: the vast majority of both men and women hold the view that women should not work full-time as long as she has children living at home. This raises the question of where these gendered preferences are coming from? In this section we present evidence that child penalties are transmitted through generations — from parents to their daughters — consistent with an influence of nurture in the formation of female preferences over family and career.

Our analysis relates to the literature studying the importance of gender identity norms in the labor market, as reviewed by Bertrand (2011). Many have argued that gender identity is formed during childhood, and some papers have documented the existence of intergenerational correlations in gender identity norms and female labor supply. Fernandez *et al.* (2004) find that the labor force participation of married women is positively correlated with the labor force participation of their husbands' mothers, but *not* with the labor force participation of their own mothers, after controlling for various socio-economic characteristics. Their interpretation is that men growing up with working mothers develop more modern gender role attitudes and therefore have stronger preferences for working wives. Related, Farre & Vella (2013) show that mothers' gender role attitudes are correlated with their children's attitudes and labor force participation. They also find such correlations between mothers and daughters-in-law, similar to Fernandez *et al.* (2004).

Our analysis diverges from previous studies in two fundamental respects. First, we consider the intergenerational transmission of child penalties — i.e., labor supply changes around child birth for women relative to men — rather than the intergenerational transmission of labor supply levels. Working with child penalties takes care of some of the omitted variable concerns encountered when working with labor supply levels. Our analysis relies on our ability to link three generations — children, their parents, and their maternal and paternal grandparents — in the Danish administrative data. For comparability with earlier work, we also consider intergenerational correlations in labor supply levels. Second, we demonstrate the existence of a link between child penalties and the maternal grandparents, but not the paternal grandparents, in contrast to the ear-

lier findings discussed above. Our findings are consistent with the idea that women’s preferences over family and career is shaped by the gender roles she is exposed to during her childhood.

5.2 Specification

The analysis is based on our baseline event study sample of men and women who have their first child between 1985-2003 and are observed in a 15-year window around the first child birth. To link the child penalties for these parents to the past labor market behavior of the grandparents, we exploit that our administrative ATP measure of hours worked goes back to 1964. This allows us to investigate the relationship between child penalties for the 1985-2003 births and the relative labor supply of grandmothers and grandfathers during 1964-1979, distinguishing between maternal and paternal grandparents.

Denoting the cumulated labor supplies between 1964-1979 of the maternal grandmother and grandfather by h_i^{mm} and h_i^{mf} , we rank parents by quantiles of the distribution of $h_i^{mm} - h_i^{mf}$. For the paternal grandparents, we similarly rank parents by quantiles of the distribution of $h_i^{pm} - h_i^{pf}$. A higher rank in these distributions implies that the grandparents were more “modern” in terms of their gender division of labor. We base the rankings on within-cohort distributions of the grandparents in order to capture grandparental behavior relative to the norm for their generation.

We estimate how child penalties vary by these grandparental rank measures. Because here we are not primarily interested in the entire dynamic path of child penalties, we adopt a more parsimonious specification that replaces the full set of event time dummies for $t = -5, \dots, 10$ by a single dummy for being at positive event times $t \geq 0$. In other words, we are considering average child penalties over the 10-year period following child birth. When considering the effect of maternal grandparents, we write the specification as follows

$$Y_{is}^g = \sum_q \alpha_q^g \cdot \mathbf{I}[\text{after}_{is}] \cdot \mathbf{I}[\text{grand}_{iq}^m] + \sum_k \beta_k^g \cdot \mathbf{I}[k = \text{age}_{is}] + \sum_y \gamma_y^g \cdot \mathbf{I}[y = s] + \delta^g X_i^m + \nu_{is}, \quad (10)$$

where $\mathbf{I}[\text{after}_{is}]$ is an indicator for individual i having had his/her first child in year s , $\mathbf{I}[\text{grand}_{iq}^m]$ is an indicator for the maternal grandparents being in quantile q of the distribution of relative labor supply for their generation, and X_i^m is a vector of controls for the maternal grandparents (see below). As before, we run the regression separately for men and women, estimating the child penalty as $P_q \equiv (\hat{\alpha}_q^m - \hat{\alpha}_q^w) / E[\tilde{Y}_{is}^w]$ in the q th quantile of the relative labor supply distribution

of the maternal grandparents. The specification for the paternal grandparents is the same, only replacing m by p .

We allow for a rich set of controls for the characteristics of the grandparents in order to ensure that what appears as intergenerational transmission of child penalties is not just a transmission of other characteristics that are correlated with child penalties. The controls included are the following. First, we include education dummies for both the grandmother and grandfather (maternal and paternal, respectively) capturing both the length and the field of education. The length of education is divided into primary school, secondary school, vocational training, short post-secondary school, bachelor's degree, and master's/phd degrees. Above secondary school, each level is divided into different fields such that we end up with 22 education dummies for each of the grandparents. These controls ensure that the intergenerational correlation in child penalties does not reflect a transmission of educational preferences or ability. Second, we control for the wealth level of the grandparents. We use the average net wealth of the grandfather over the years 1980-90 and control for quantiles of the within-cohort wealth rank of the grandfather.²³ This ensures that the child penalties are not driven by wealth effects that are transmitted through generations. Finally, we include a full set of dummies for the birth cohort of both the grandmother and the grandfather.

5.3 Results

Our main set of results is presented in Figure 10. Each panel plots the child penalty in earnings against quintiles of the relative labor supply distribution of the grandparents. The left panels consider the maternal grandparents and the right panels consider the paternal grandparents. The top panels show intergenerational correlations without any controls for education and wealth, while the bottom panels allow for a rich set of controls as described above.

The figure shows a clear downward-sloping relationship between the child penalty on mothers and the relative labor supply of the maternal grandmother. That is, women incur smaller earnings penalties due to children if they themselves grew up in a family where the mother worked more relative to the father. The effect of going from the bottom quintile to the top quintile of the relative labor supply distribution of the maternal grandparents is about 4 percentage points. The size of this effect is roughly unaffected by including the detailed non-parametric controls for education and wealth.

²³This third-party reported wealth measure is available for the universe of Danish taxpayers as it was collected for the purpose of a wealth tax that existed until 1996.

At the same time, the figure shows a zero effect of the paternal grandparents: the relationship between the female child penalty and the relative labor supply of the paternal grandparents is completely flat, and this holds independently of whether we include grandparental controls. The differential pattern between maternal and paternal grandparents is interesting for two reasons. First, it is helpful for ruling out the threat from omitted variables that are present in both sets of grandparents, i.e. family background variables that both parents have and which affect child penalties. Second, it suggests that female child penalties are driven partly by female preferences formed during her childhood, rather than by male preferences formed during his childhood.

Our findings diverge from previous work focusing on intergenerational transmission that operate through the gender attitudes of sons rather than daughters. It is worth asking if our results are driven by looking at a different outcome — child penalties rather than labor supply levels — or by using different data. Table 1 therefore shows intergenerational correlations in labor force participation levels (columns 1-3) and earnings levels (columns 4-6) for women (Panel A) and men (Panel B). The results that are most directly comparable to the previous literature are those in columns (1)-(3) of Panel A on the labor force participation of women.

The table shows that the results for labor supply levels are different than for child penalties, and also that we do not replicate earlier findings. Two key takeaways emerge from the table. First, when considering labor supply levels, we obtain positive intergenerational correlations everywhere: between women and their own mothers and their mothers-in-law, as well as between men and their own mothers and mothers-in-law. This holds for labor force participation and earnings, and independently of the set of controls we include. Second, the intergenerational correlation with one's own parents is always stronger than with one's partner's parents. This last finding is inconsistent with earlier findings that women are correlated with their husband's parents, but not their own.

Instead of intergenerational correlations in levels, recent work has focused on correlations in rank. We therefore replicate Table 1 in a rank-rank specification in appendix Table A.I, replacing labor force participation (for which we cannot do rank) with total hours worked. This table confirms the qualitative insights described above.

To summarize, our findings are consistent with an influence of nurture in the formation of female preferences over children and career. This can be seen in the relationship between the labor supply *impacts* of children — but not in the labor supply *levels* — and the gender roles women were exposed to when growing up.

6 Conclusion

Despite considerable gender convergence over time, substantial gender inequality persists in all countries. Using full-population administrative data from Denmark and a quasi-experimental event study approach, we show that most of the remaining gender inequality can be attributed to the dynamic effects of children. We have presented three main sets of results.

First, we have shown that the impact of children on women is large and persistent across a wide range of labor market outcomes, while at the same time men are unaffected. The female child penalty in earnings is close to 20% in the long run. Underlying this earnings penalty, we find sharp impacts of children on labor force participation, hours worked, wage rates, occupation, sector, and firm choices. Together, these findings provide a quite complete picture of the behavioral margins that adjust in response to parenthood and how strongly gendered these margins are. We have laid out the identification assumptions of our event study approach, providing two important robustness checks that support our strategy: a difference-in-differences extension of the event study and an IV-approach based on sibling sex mix.

Second, we have decomposed gender inequality into what can be attributed to children and what can be attributed to other factors. We have shown that the fraction of child-related gender inequality has increased dramatically over time, from around 40% in 1980 to around 80% in 2013. Therefore, to a first approximation, the remaining gender inequality is all about children. Our decomposition analysis represents a re-orientation of traditional gender gap decompositions: instead of studying the extent to which men and women receive unequal pay for equal work (the unexplained gap after controlling for human capital and job characteristics, but not children), we study the extent to which they receive unequal pay as a result of children (but not necessarily for equal work). The unexplained gap in traditional decomposition analyses is often labeled “discrimination”, but our analysis highlights that the unexplained gap is largely due to children. This does not rule out discrimination, but implies that potential discrimination operates through the impacts of children.

Third, we have provided evidence in favor of environmental influences in the formation of preferences over family vs career. In particular, we have shown that the female child penalty is strongly related to the work history of the maternal grandparents: women who grow up in traditional families with a male breadwinner and a female homemaker incur larger child penalties when they themselves become mothers. At the same time, the female child penalty is unrelated

to the work history of the paternal grandparents. Overall, these findings are consistent with the notion that child penalties are influenced by female gender identity formed during her childhood, as opposed to child penalties being driven by male gender identity formed during his childhood. These results diverge from previous work emphasizing preference formation transmitted from mothers to sons.

Our paper is agnostic about the potential welfare and policy implications of our findings. Although the term “child penalty” may have normative connotations, we do not draw any normative conclusions here. The previous gender literature focusing on the unexplained gender gap had a very natural normative benchmark: equal pay for equal jobs. Our paper highlights that unequal pay is due to children, which may be good or bad depending on the perspective. A traditional economic view would focus on comparative advantage in child rearing (due to innate gender differences in abilities or preferences for child care vs market work) along with gains from specialization, in which case our findings do not necessarily call for policy intervention. Another view is that the unequal effects of children are driven by environmental factors such as culture, social norms or discrimination, producing potential inequities and inefficiencies. Our findings on intergenerational transmission are consistent with — but do not conclusively prove — the existence of such environmental factors. Future work should dig deeper into the underlying mechanisms and the implied welfare implications.

References

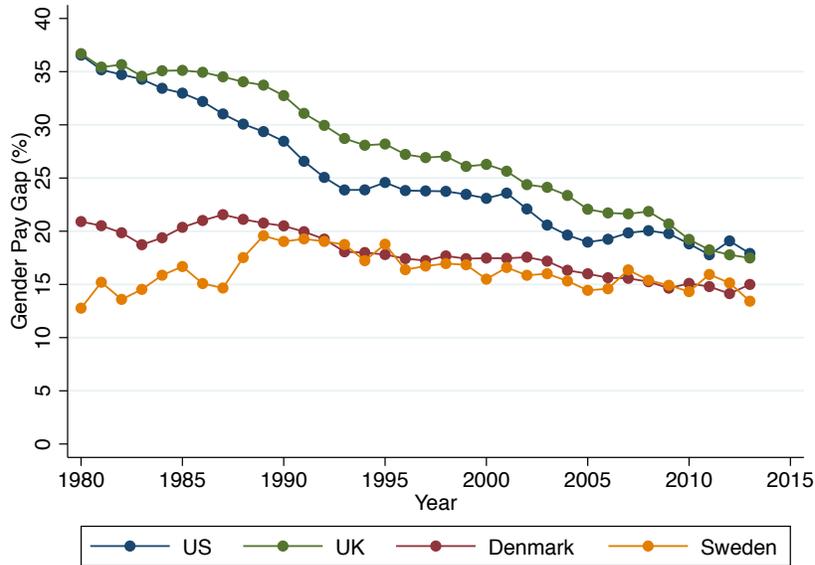
- ADDA, JEROME, DUSTMANN, CHRISTIAN, & STEVENS, KATRIEN. 2015. The Career Costs of Children. *Journal of Political Economy*, forthcoming.
- AGUERO, JORGE M., & MARKS, MINDY S. 2008. Motherhood and Female Labor Force Participation: Evidence from Infertility Shocks. *American Economic Review, Papers and Proceedings*, **98**, 500–504.
- ALTONJI, JOSEPH G., & BLANK, REBECCA M. 1999. Race and Gender in the Labor Market. *Chap. 48, pages 3143–3259 of: ASHENFELTER, O., & CARD, D. (eds), Handbook of Labor Economics, vol.3.* North Holland: Elsevier Science Publishers.
- ANGELOV, NIKOLAY, JOHANSSON, PER, & LINDAHL, ERICA. 2016. Parenthood and the Gender Gap in Pay. *Journal of Labor Economics*, **34**.
- ANGRIST, JOSHUA D., & EVANS, WILLIAM N. 1998. Children and Their Parents' Labor Supply: Evidence from Exogenous Variation in Family Size. *American Economic Review*, **88**, 450–477.
- BERTRAND, MARIANNE. 2011. New Perspectives on Gender. *Chap. 17, pages 1543–1590 of: ASHENFELTER, O., & CARD, D. (eds), Handbook of Labor Economics, vol.4b.* North Holland: Elsevier Science Publishers.
- BERTRAND, MARIANNE, GOLDIN, CLAUDIA, & KATZ, LAWRENCE F. 2010. Dynamics of the Gender Gap for Young Professionals in the Financial and Corporate Sectors. *American Economic Journal: Applied Economics*, **2**, 228–255.
- BERTRAND, MARIANNE, KAMENICA, EMIR, & PAN, JESSICA. 2013. *Gender Identity and Relative Income Within Households*. Working Paper, October 2013.
- BLAU, FRANCINE, & KAHN, LAWRENCE. 2016. The Gender Wage Gap: Extent, Trends, and Explanations. *Journal of Economic Literature*, forthcoming.
- BRONARS, STEPHEN G., & GROGGER, JEFF. 1994. The Economic Consequences of Unwed Motherhood: Using Twin Births as a Natural Experiment. *American Economic Review*, **84**, 1141–1156.
- BROWNING, MARTIN. 1992. Children and Household Economic Behavior. *Journal of Economic Literature*, **30**, 1434–1475.

- FARRE, LIDIA, & VELLA, FRANCIS. 2013. The Intergenerational Transmission of Gender Role Attitudes and its Implications for Female Labour Force Participation. *Economica*, **80**, 219–247.
- FERNANDEZ, RAQUEL, FOGLI, ALESSANDRA, & OLIVETTI, CLAUDIA. 2004. Mothers and Sons: Preference Formation and Female Labor Force Dynamics. *Quarterly Journal of Economics*, **119**, 1249–1299.
- FERNANDEZ-KRANZ, DANIEL, LACUESTA, AITOR, & RODRIGUEZ-PLANAS, NURIA. 2013. The Motherhood Earnings Dip: Evidence from Administrative Records. *Journal of Human Resources*, **48**, 169–197.
- FITZENBERGER, BERND, SOMMERFELD, KATRIN, & STEFFES, SUSANNE. 2013. Causal Effects on Employment After First Birth: A Dynamic Treatment Approach. SOEP Working Paper No. 576.
- FORTIN, NICOLE, LEMIEUX, THOMAS, & FIRPO, SERGIO. 2011. Decomposition Methods in Economics. *Pages 1–102 of: ASHENFELTER, O., & CARD, D. (eds), Handbook of Labor Economics, vol.4.* North Holland: Elsevier Science Publishers.
- GOLDIN, CLAUDIA. 2014. A Grand Gender Convergence: Its Last Chapter. *American Economic Review*, **104**, 1091–1119.
- GOLDIN, CLAUDIA, & KATZ, LAWRENCE F. 2016. A Most Egalitarian Profession: Pharmacy and the Evolution of a Family-Friendly Occupation. *Journal of Labor Economics*, **34**, 705–746.
- GOLDIN, CLAUDIA, KATZ, LAWRENCE F., & KUZIEMKO, ILYANA. 2006. The Homecoming of American College Women: The Reversal of the College Gender Gap. *Journal of Economic Perspectives*, **20**, 133–156.
- HOTZ, V. JOSEPH, MCELROY, SUSAN WILLIAMS, & SANDERS, SETH G. 2005. Teenage Childbearing and Its Life Cycle Consequences: Exploiting a Natural Experiment. *Journal of Human Resources*, **40**, 683–715.
- KLEVEN, HENRIK. 2014. How Can Scandinavians Tax So Much? *Journal of Economic Perspectives*, **28**, 77–98.
- KLEVEN, HENRIK, & LANDAIS, CAMILLE. 2016. Gender Inequality and Economic Development: Micro Evidence Across Countries. Working Paper.

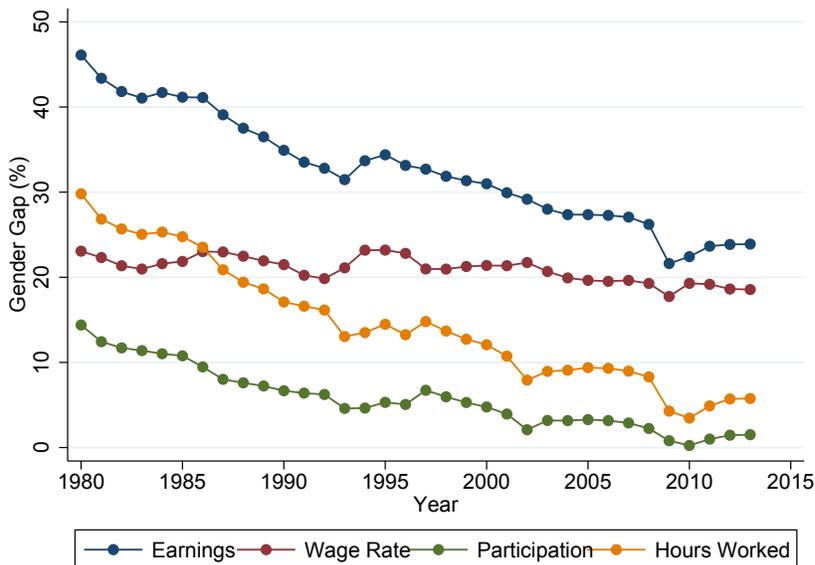
- LUNDBERG, SHELLY, & ROSE, ELAINA. 2000. Parenthood and the Earnings of Married Men and Women. *Labour Economics*, **7**, 689–710.
- LUNDBORG, PETTER, PLUG, ERIK, & RASMUSSEN, ASTRID WÄCERTZ. 2014. Fertility Effects on Female Labor Supply: IV Evidence from IVF Treatments. IZA Discussion Paper No. 8609.
- NIELSEN, HELENA SKYT, SIMONSEN, MARIANNE, & VERNER, METTE. 2004. Does the Gap in Family-Friendly Policies Drive the Family Gap? *Scandinavian Journal of Economics*, **106**, 721–744.
- OLIVETTI, CLAUDIA, & PETRONGOLO, BARBARA. 2016. The Evolution of the Gender Gap in Industrialized Countries. *Annual Review of Economics*, **8**, forthcoming.
- PAULL, GILLIAN. 2008. Children and Women’s Hours of Work. *Economic Journal*, **118**, F8–F27.
- ROSENZWEIG, MARK R., & WOLPIN, KENNETH I. 1980. Life-Cycle Labor Supply and Fertility: Causal Inferences from Household Models. *Journal of Political Economy*, **88**, 328–348.
- WALDFOGEL, JANE. 1998. Understanding the Family Gap in Pay for Women with Children. *Journal of Economic Perspectives*, **12**, 137–156.
- WILDE, ELIZABETH TY, BATCHELDER, LILY, & ELLWOOD, DAVID T. 2010. *The Mommy Track Divides: The Impact of Childbearing on Wages of Women of Differing Skill Levels*. NBER Working Paper No. 16582, December 2010.

Figure 1: Gender Gaps Across Countries 1980-2013

A: Convergence of the Gender Pay Gap Across Countries
 Median Earnings for Full-Time Workers



B: Evolution of Gender Gaps in Denmark
 Means for All Workers

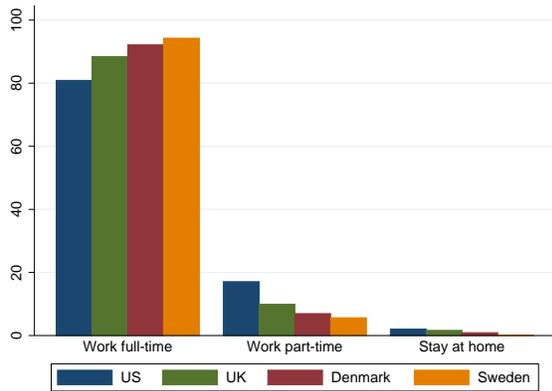


Notes: The time series in Panel A are drawn from OECD.org, except for Denmark where we use our own calculations of median earnings for full-time workers aged 16-64 (where full time is defined based on the ATP hours measure described in section 2.2). Our calculations for Denmark uses the same underlying data as the official OECD series, but is more consistent with the sample definitions used for the other countries. In Panel B the gaps in earnings and participation are calculated for the entire population aged 16-64 regardless of employment status, while the gaps in hours worked and wage rates are calculated conditional on participation.

Figure 2: Gender Norms Across Countries

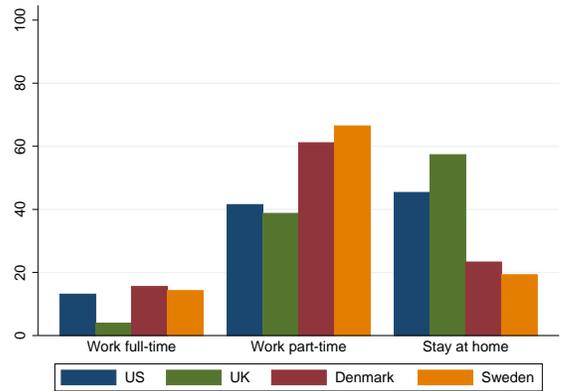
A: Women Without Children

Do you think that women should work outside the home full-time, part-time or not at all when they are married but with no children?



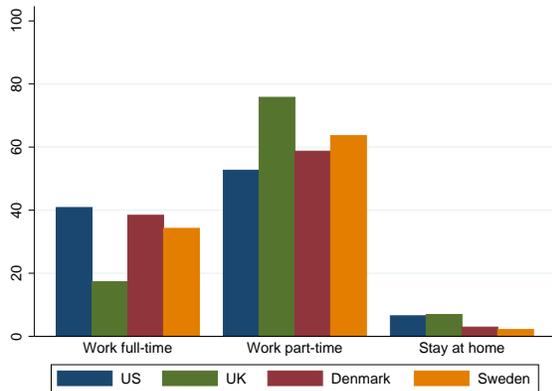
B: Women With Children Under School Age

Do you think that women should work outside the home full-time, part-time or not at all when there is a child under school age?



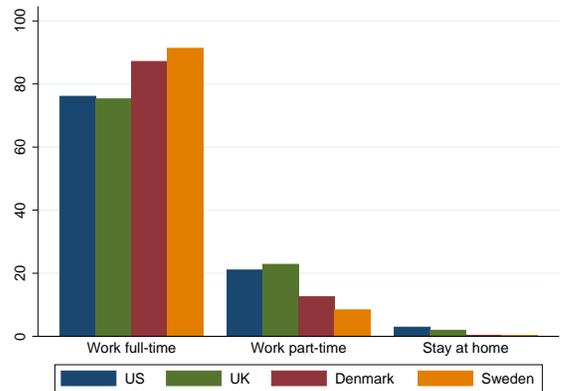
C: Women With Children In School

Do you think that women should work outside the home full-time, part-time or not at all when the youngest child is still in school?



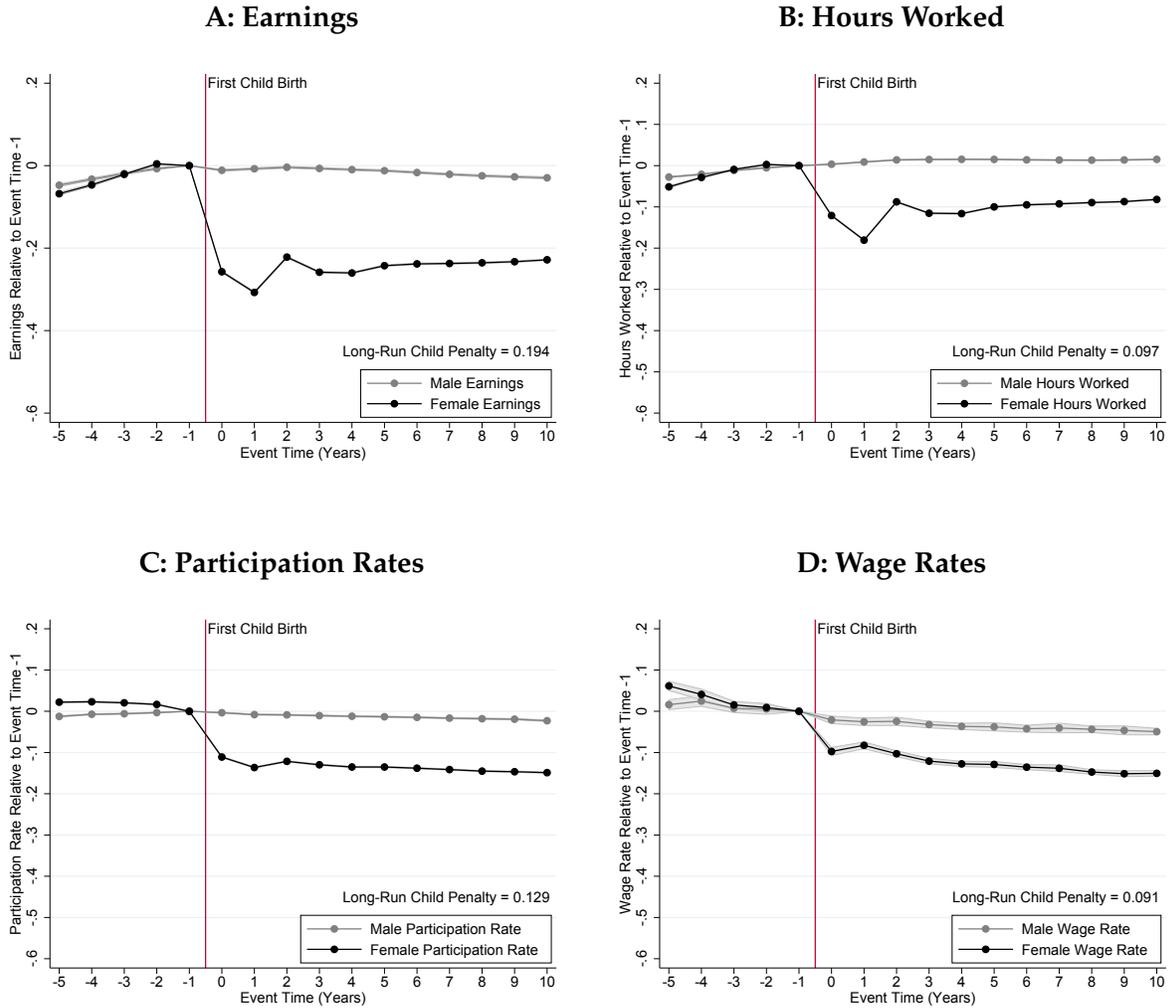
D: Women With Children Who Have Left Home

Do you think that women should work outside the home full-time, part-time or not at all when the child has left the home?



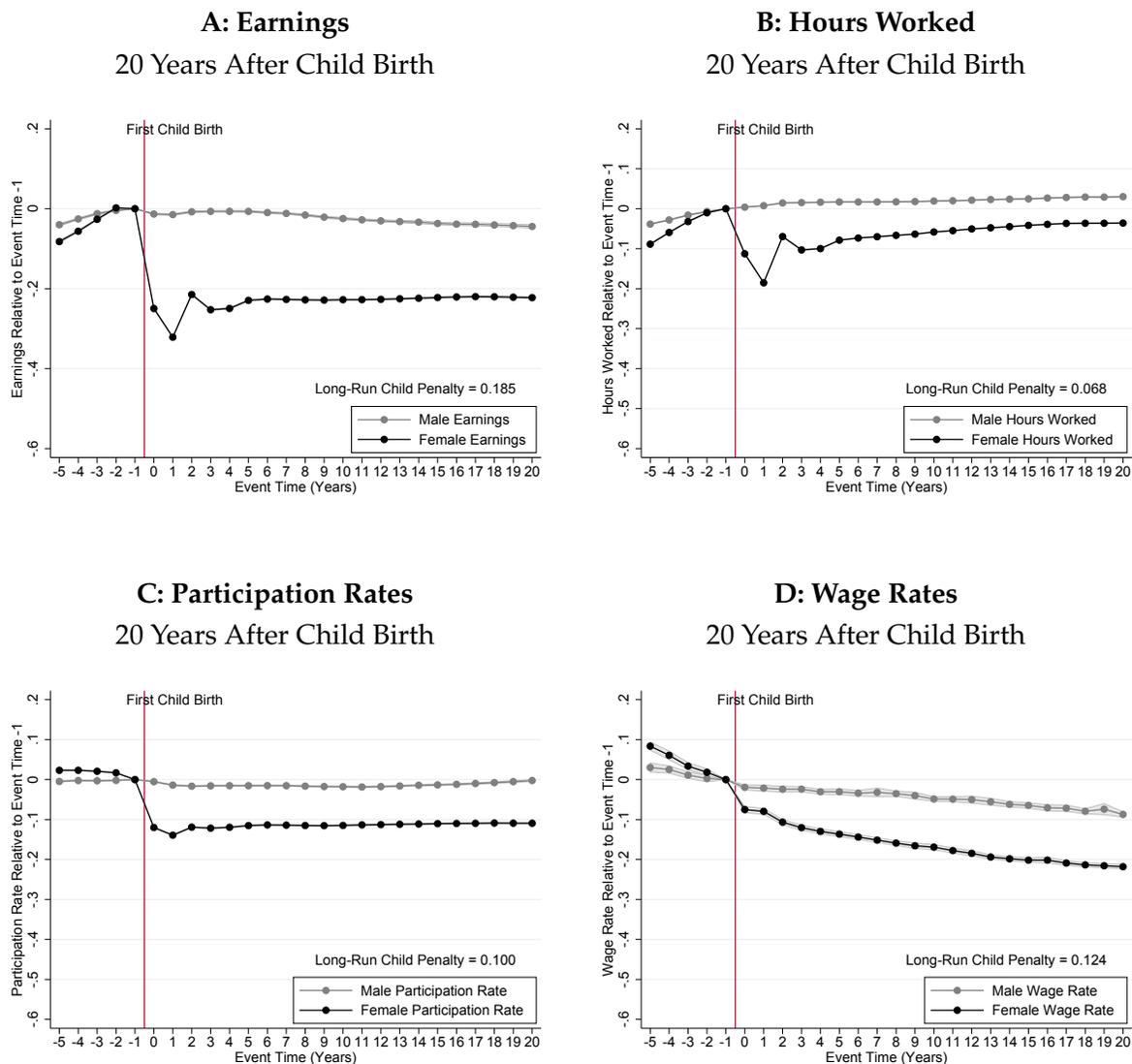
Notes: The figure is based on data from the International Social Survey Program (ISSP) in 2002. Each panel shows shares (in percent) choosing each of the 3 listed categories.

Figure 3: Impacts of Children



Notes: The graphs show event time coefficients estimated from equation (1) as a percentage of the counterfactual outcome absent children (i.e., $P_t^g \equiv \hat{\alpha}_t^g / E[\tilde{Y}_{ist}^g | t]$ as defined in section 3.1) for men and women separately and for different outcomes. Each panel also reports a “child penalty” — the percentage by which women are falling behind men due to children — defined as $P_t \equiv (\hat{\alpha}_t^m - \hat{\alpha}_t^w) / E[\tilde{Y}_{ist}^w | t]$. The long-run child penalty is measured at event time 10. All of these statistics are estimated on a balanced sample of parents, who have their first child between 1985-2003 and who are observed in the data during the entire period between 5 years before and 10 years after child birth. The effects on earnings and participation are estimated unconditional on employment status, while the effects on hours worked and wage rates are estimated conditional on participation. The shaded 95 % confidence intervals are based on robust standard errors.

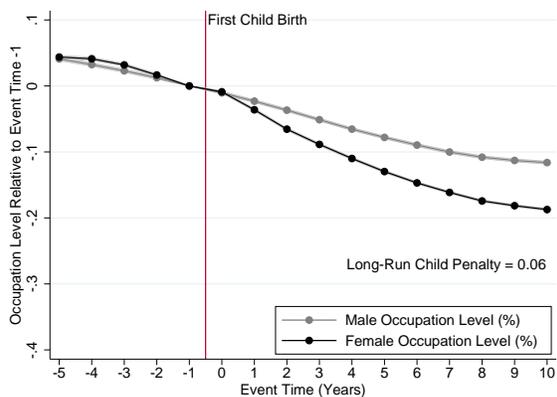
Figure 4: Impacts of Children in the Very Long Run



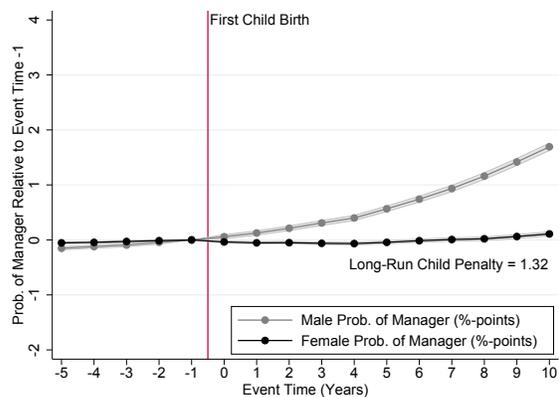
Notes: This figure is constructed in the same way as Figure 3, but expanding the event time window to include 20 years after the birth of the first child. In order to do this, we expand from the previously balanced panel of parents who have their first child between 1985-2003 to an unbalanced panel of parents who have their first child at any time after 1985. Only the birth cohorts 1985-1993 are observed all the way to event time 20. The long-run child penalty is measured at event time 20. The shaded 95 % confidence intervals are based on robust standard errors.

Figure 5: Anatomy of Child Impacts

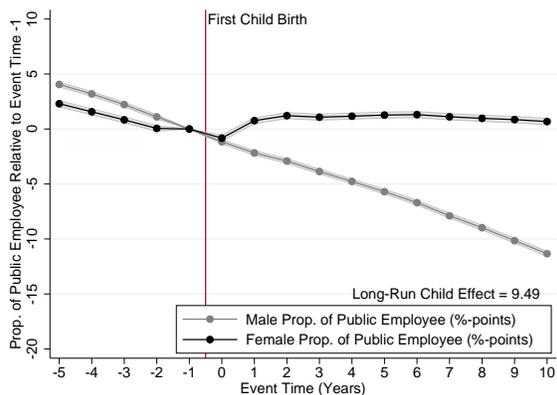
A: Occupational Rank
Levels 1-5 from Unskilled Labor to Manager



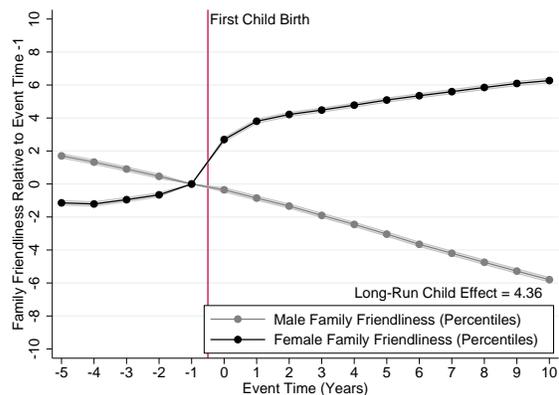
B: Probability of Being Manager
Manager Dummy



C: Probability of Public Sector Job
Public Sector Dummy



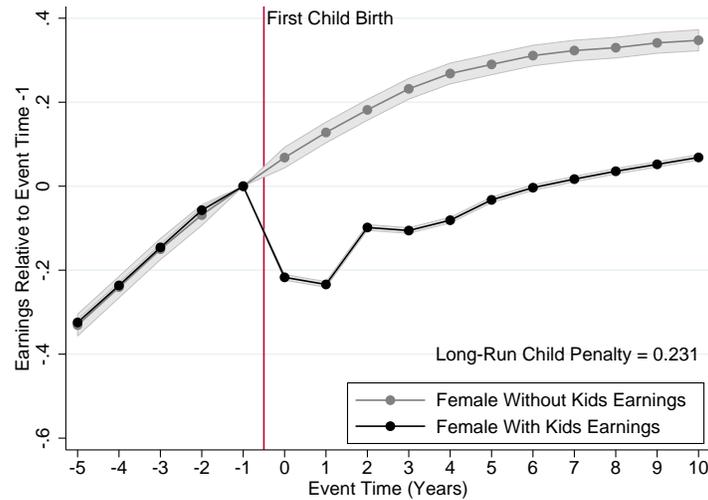
D: Family Friendliness of Firm
Share of Women with Young Children in the Firm



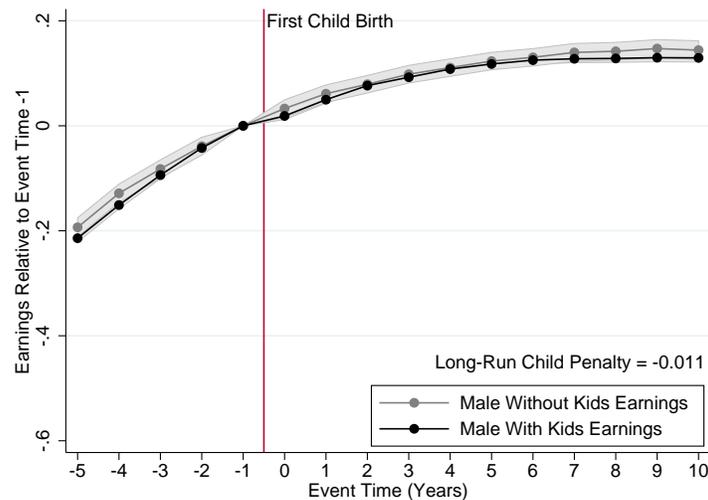
Notes: The graphs show percentage impacts of children on men and women P_t^m, P_t^w as well as the long-run child penalty P_t (at event time 10) as defined in Section 3.1. The effects on occupational rank are estimated conditional on not being self-employed or an assisting spouse. The effects on the family friendliness of firms (defined as the share of women with children below age 15 in the firm, excluding the considered woman's own child) is estimated conditional on being in a firm with at least 10 employees. The long-run effect of children on the family friendliness of firms (= 4.36) takes into account the differential pre-trend between men and women. The shaded 95 % confidence intervals are based on robust standard errors.

Figure 6: Impacts of Children in a Difference-in-Differences Event Study Design

A: Women Who Have Children vs Women Who Don't
Earnings Impact

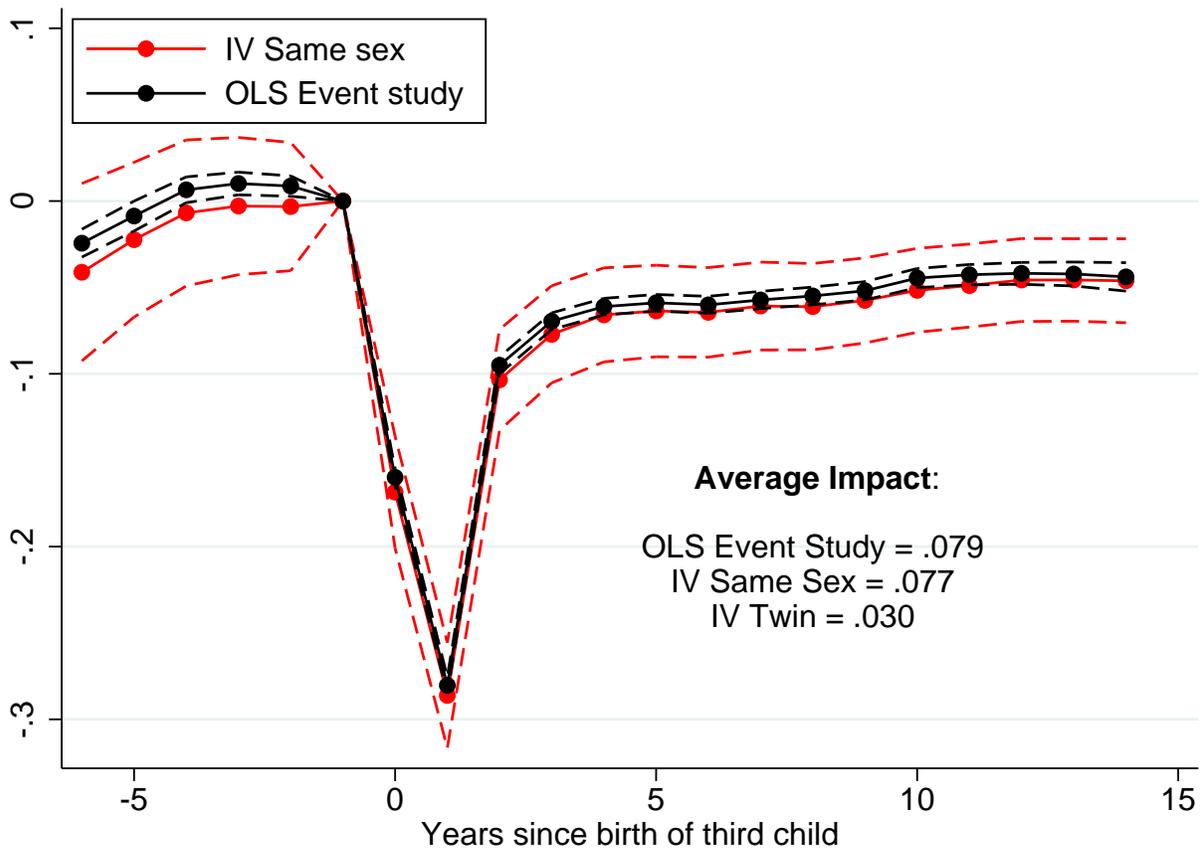


B: Men Who Have Children vs Men Who Don't
Earnings Impact



Notes: The graphs show the earnings impacts of children on women P_t^w (Panel A) and men P_t^m (Panel B). As in the previous graphs, these statistics are obtained from the baseline event study specification (1), but here we run the specification both on those who have children and those who never have children (assigning placebo births based on the observed distribution of age at first child among those who have children). The details of how we construct the control groups of men and women who never have children are described in Section 3.4 and in Appendix A.1. The figure reports long-run child penalties for men and women separately, estimated as a difference-in-differences between those who have children and those who never have children (as opposed to previous penalty measures based on comparing men and women, both of whom have children). The shaded 95 % confidence intervals are based on robust standard errors.

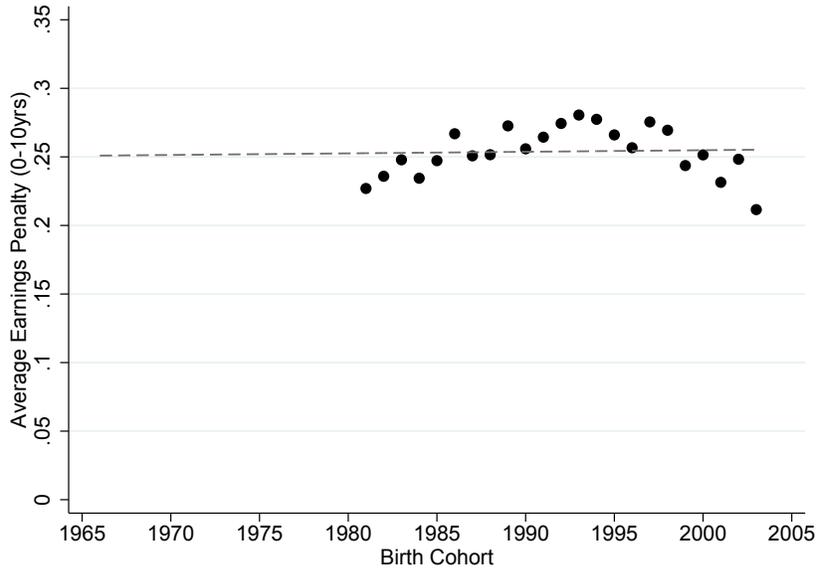
Figure 7: Earnings Impact of Third Child in Event Study vs IV Designs
 Sibling Sex Mix or Twin Births as Instruments



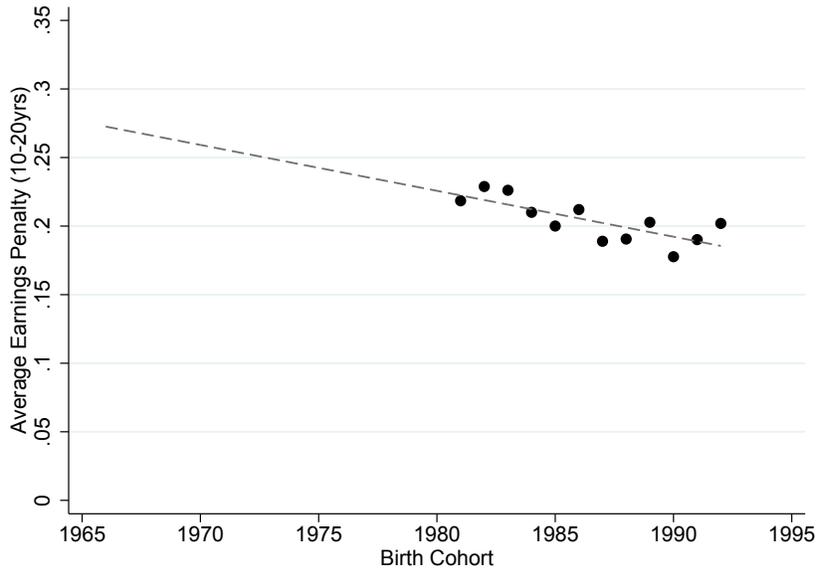
Notes: The figure shows the earnings impact of a third child on women (P_t^w defined in section 3.4) obtained from the OLS event study specification (black series) and the IV same-sex specification (red series) as a function of years since the birth of the third child. The event study estimates are based on specification (7). The IV-specification is based on the same specification, but dropping the individual fixed effect and instead instrumenting the event time dummies for the third child birth using the sex mix of the first two children. The 95 % confidence intervals are based on robust standard errors. The figure also compares the average impact (i.e., across event times 0-15) obtained from the event study, the IV using sibling sex, and an IV using twins in the second birth.

Figure 8: Earnings Penalties by Birth Cohort and Extrapolation

A: Average Earnings Penalty From Event Time 0-10
Zero Trend Across Cohorts



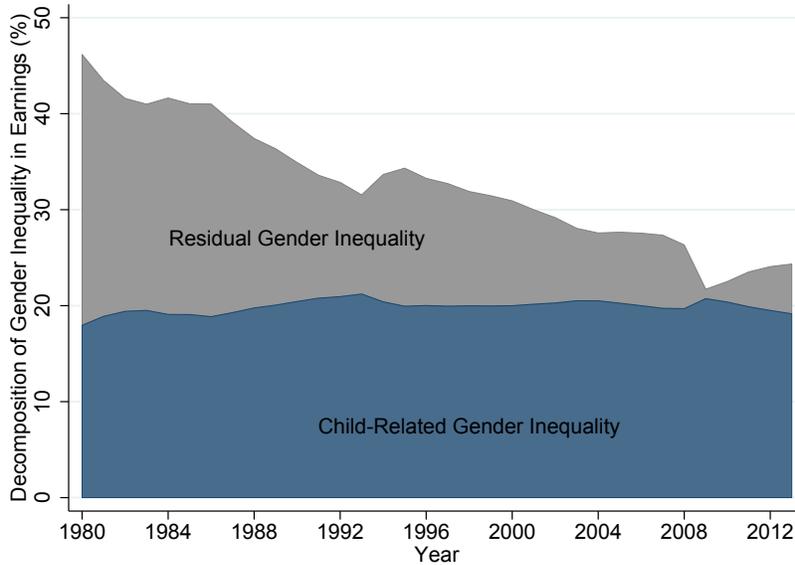
B: Average Earnings Penalty From Event Time 11-20
Downward Trend Across Cohorts



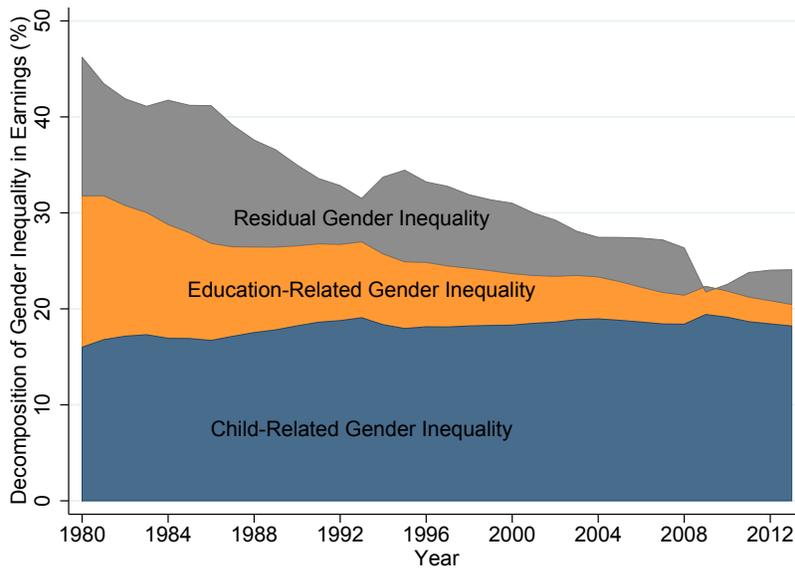
Notes: The figure shows earnings penalties by birth cohort obtained from (8). Panel A shows the average penalty across event times 0-10, while Panel B shows the average penalty across event times 11-20. Each panel also includes a linear OLS fit. There is a zero trend in the 0-10 year penalty, while there is a linear downward trend in the 11-20 year penalty. We use these linear trends to extrapolate child penalties to birth cohorts prior to 1981.

Figure 9: Decomposing Gender Inequality in Earnings

A: Child-Related Inequality vs Non-Child Inequality

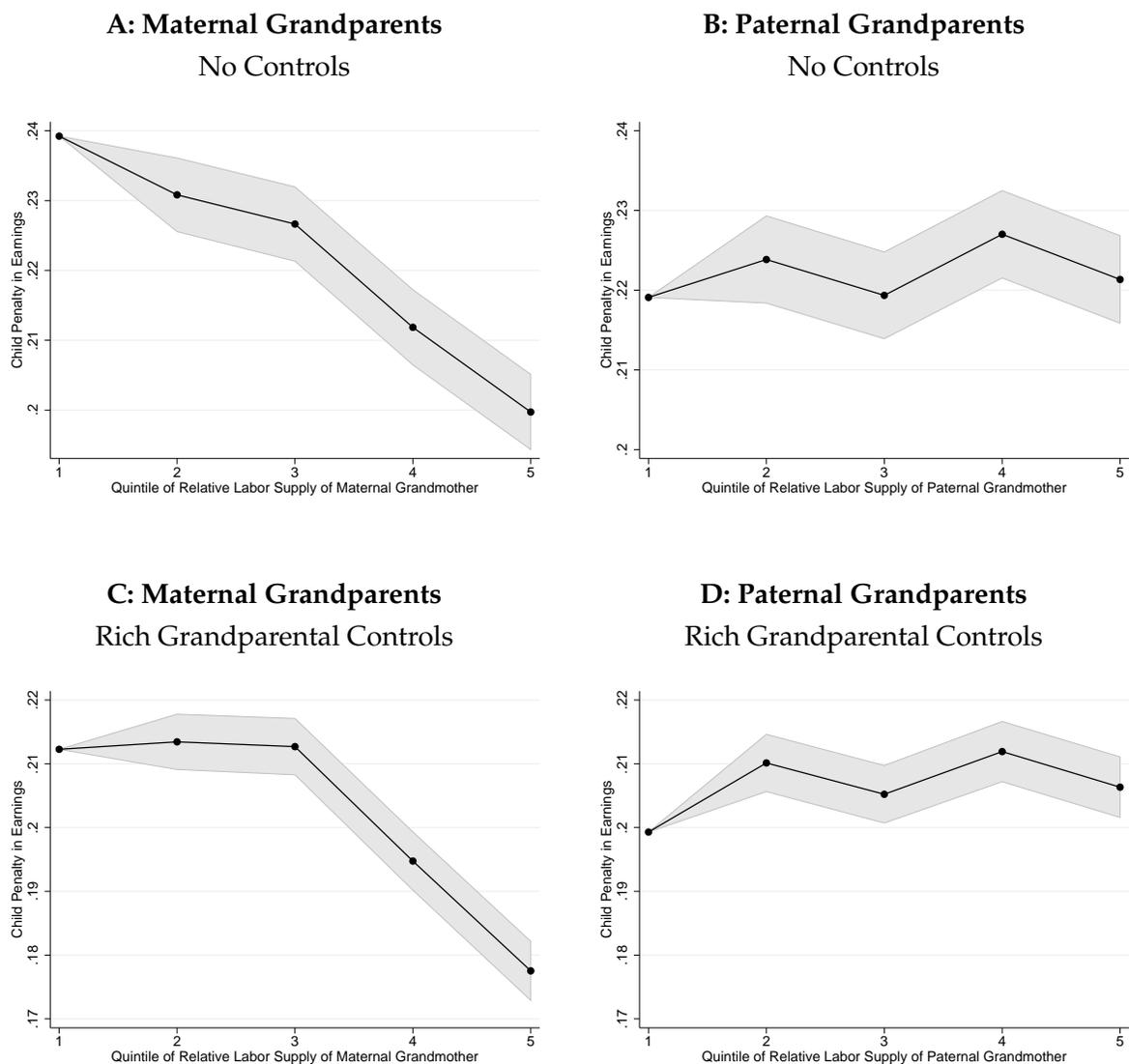


B: Child-Related Inequality vs Education-Related Inequality
(Post-Child Effects vs Pre-Child Effects)



Notes: The figure shows dynamic Oaxaca-Blinder decompositions based on equations (8) and (9). The decomposition shown in Panel A allows for year-specific event-time coefficients and control for a full set of age and year dummies. The decomposition in Panel B augments the specification with education dummies: primary school, secondary school, vocational training, short post-secondary school, bachelor's degree, and master's/phd degrees. In both panels, the child-related gender gap captures the effect of child penalties P_{st} (the "unexplained" effect of children). In Panel B, the education-related gender gap includes both the effect of different education levels (explained effect) and the effect of different education coefficients (unexplained effect). The residual gender gap represents the effects (explained and unexplained) of age and year dummies.

Figure 10: Intergenerational Transmission of Child Penalties in Earnings



Notes: Each panel shows the child penalty in earnings against quintiles of the relative labor supply distribution of the grandparents. The relative labor supply of grandparents is based on cumulated hours worked over the period 1964-79 (obtained from ATP pension contributions). The specification is equation (10) and the statistic reported is P_q as defined in Section 5.2. The left-hand panels show the effect of the maternal grandparents, while the right-hand panels show the effect of the paternal grandparents. The top panels do not include controls for grandparental characteristics, while the bottom panels allow for a rich set of non-parametric controls for the education level/fields, within-generation wealth rank, and birth cohort of the grandparents. The shaded 95 % confidence intervals are based on robust standard errors.

Table 1: Intergenerational Correlation of Labor Supply and Earnings: Level Regressions

	Labor Force Participation (LFP)			Earnings		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Women						
LFP Own Mother	0.029*** (0.0010)	0.020*** (0.0012)	0.019*** (0.0011)			
LFP Mother-in-Law	0.021*** (0.0009)	0.014*** (0.0011)	0.010*** (0.0010)			
Earnings Own Mother				0.242*** (0.0025)	0.164*** (0.0029)	0.090*** (0.0024)
Earnings Mother-in-Law				0.181*** (0.0025)	0.107*** (0.0029)	0.044*** (0.0023)
Panel B: Men						
LFP Own Mother	0.037*** (0.0010)	0.034*** (0.0012)	0.033*** (0.0012)			
LFP Mother-in-Law	0.019*** (0.0010)	0.015*** (0.0012)	0.015*** (0.0012)			
Earnings Own Mother				0.296*** (0.0042)	0.184*** (0.0048)	0.108*** (0.0042)
Earnings Mother-in-Law				0.228*** (0.0039)	0.117*** (0.0046)	0.027*** (0.0041)
<u>Controls:</u>						
Grandparental		×	×		×	×
Parental			×			×
<i>N</i>	5779928	4592396	4592396	5779928	4592396	4592396

Notes: Robust standard errors in parentheses. Grandparental controls include a full set of cohort dummies for each grandparent, 22 dummies for the education level/fields of each grandparent, and dummies for the quartiles of the wealth rank of the grandfather. Parental controls include a full set of age dummies for each parent, 22 dummies for the education level/fields of each parent, completed fertility of the household by 2013, the number of children in the household at the time of observation, the income level of the spouse, and dummies for the region of residence of the household.

A Appendix

A.1 Technical Details of DD Event Study Design

Defining the Control Group (Those Who Never Have Children):

As discussed in section 3.4, when wanting to use those who never have children as controls, we face a truncation issue. If we take age 40 as the latest age at which people typically have their first child, then first child births will be truncated for cohorts born after 1973 as they are younger than 40 when we last see them in 2013. This implies that, among individuals observed without children in these younger generations, some of them will have children later and should therefore be dropped from our control group.

To compare individuals with positive lifetime fertility to those with zero lifetime fertility, we need, for those born after 1973 and observed without children by 2013, to separate between those who will never have children and those who will have children later. We do this based on the following approach, which is implemented separately for women and men:

1. For individuals born between 1955-1973 (non-truncated cohorts), we estimate a linear probability model of zero lifetime fertility ($k_{iT} = 0$) on observable characteristics X :

$$P[k_{iT} = 0] = X'\beta, \tag{11}$$

where X includes the following dummy controls when running the specification for women (and symmetrically for the men): quartiles of the income distribution of the woman's cohort, quartiles of the wealth distribution of the woman's cohort, quartiles of the wealth distribution of the man's (spouse's) cohort, education length/degree (8 categories), decade of generation of the maternal grandmother, decade of generation of maternal grandfather, region of residence.

2. We predict the probability of zero lifetime fertility for individuals born after 1973 based on the model above: $\hat{P} = X'\hat{\beta}$.
3. For each cohort c born after 1973, we rank individuals based on their \hat{P} s. Among those from cohort c who have no children by 2013, we pick the n_c individuals with the highest \hat{P} s such that $\frac{n_c}{N_c} = P_{1955-1973}$, where N_c is the total number of individuals (men and women, respectively) in the cohort and $P_{1955-1973}$ is the average fraction of individuals with $k_{iT} =$

0 for cohorts 1955-1973. These n_c individuals are assumed to have zero lifetime fertility. The rest of cohort c observed with no children are assumed to have children later, and are therefore not included among the controls. Our analysis thus assumes that the fraction of people who never have children remains the same for cohorts born after 1973 as it was for cohorts born between 1955-1973, which can be justified by the observation that this fraction did in fact stay roughly constant across the (non-truncated) cohorts up until 1973.

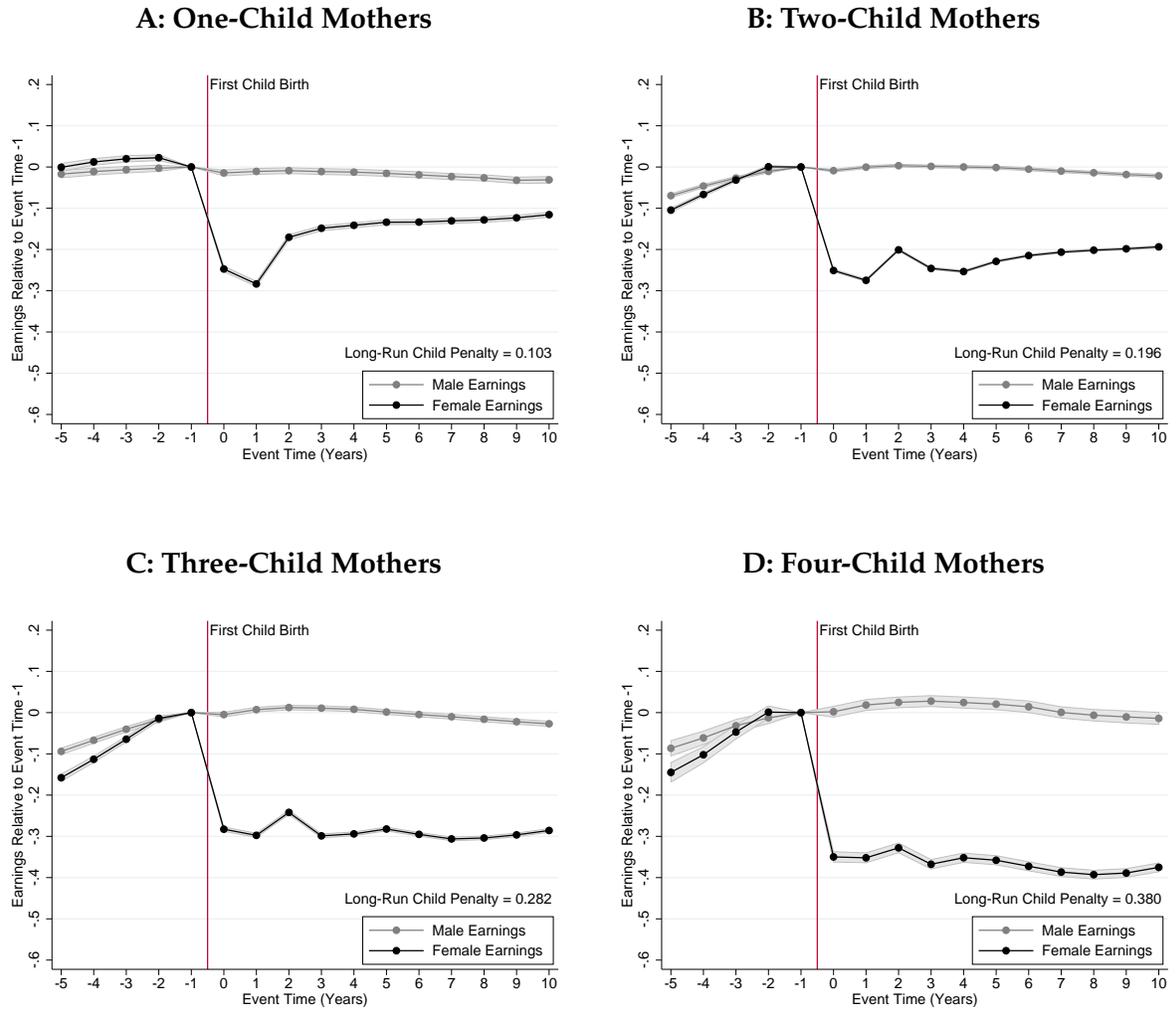
4. Our control group includes everybody without children from cohorts $c \leq 1973$ and, using the procedure above, the selected n_c individuals without children from cohorts $c > 1973$.

Assigning Placebo Births to Controls:

We then need to assign placebo births to those in our control group. This is done as follows:

1. For the older, non-truncated cohorts $c \leq 1973$, the distribution of age at first child is approximated by a log-normal distribution within cells of birth cohort and education. Denoting age at first child conditional on cohort and education by $A_{c,e}$, we assume $A_{c,e} \sim \mathcal{LN}(\hat{\mu}_{c,e}, \hat{\sigma}_{c,e}^2)$ where the mean $\hat{\mu}_{c,e}$ and variance $\hat{\sigma}_{c,e}^2$ are obtained from the actual distributions within each cohort-education cell. For the cells of cohort and education, we divide birth cohorts into decades (1920s, 1930s, etc.) and define education as highest level achieved in four cells (less than high school, high school, short higher education, long higher education). For individuals belonging to cohorts $c \leq 1973$ observed with $k_{iT} = 0$, we draw a random age at first child from $\mathcal{LN}(\hat{\mu}_{c,e}, \hat{\sigma}_{c,e}^2)$.
2. For the younger, truncated cohorts $c \geq 1974$ with imputed $k_{iT} = 0$, we draw a random age at first child from $\mathcal{LN}(\tilde{\mu}_{c,e}, \hat{\sigma}_{c,e}^2)$ where the mean $\tilde{\mu}_{c,e}$ is the predicted average age at first child obtained by estimating a linear trend on the older non-truncated cohorts. That is, consistent with the stylized pattern observed across the older cohorts, we allow for an upward linear drift in the age at first child while keeping the variance constant.

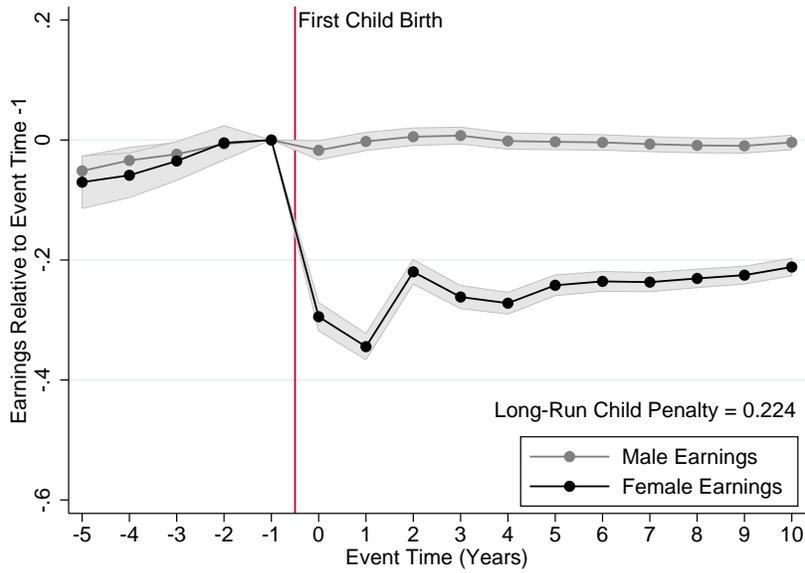
Figure A.I: Earnings Impacts by Number of Children



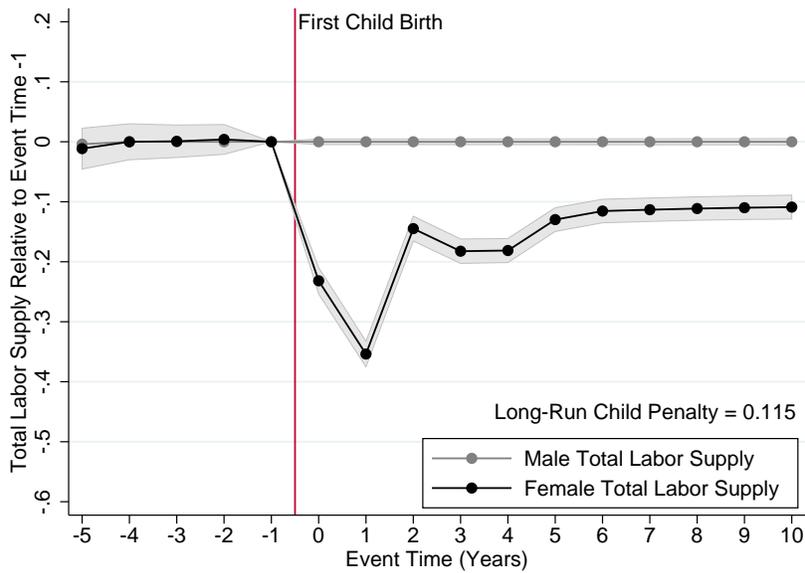
Notes: The figure shows the impact of children on earnings exactly as in Figure 3A, but splitting the sample by the woman's total number of children as of 2013 (1, 2, 3, or 4 children).

Figure A.II: Median Impacts of Children

A: Earnings

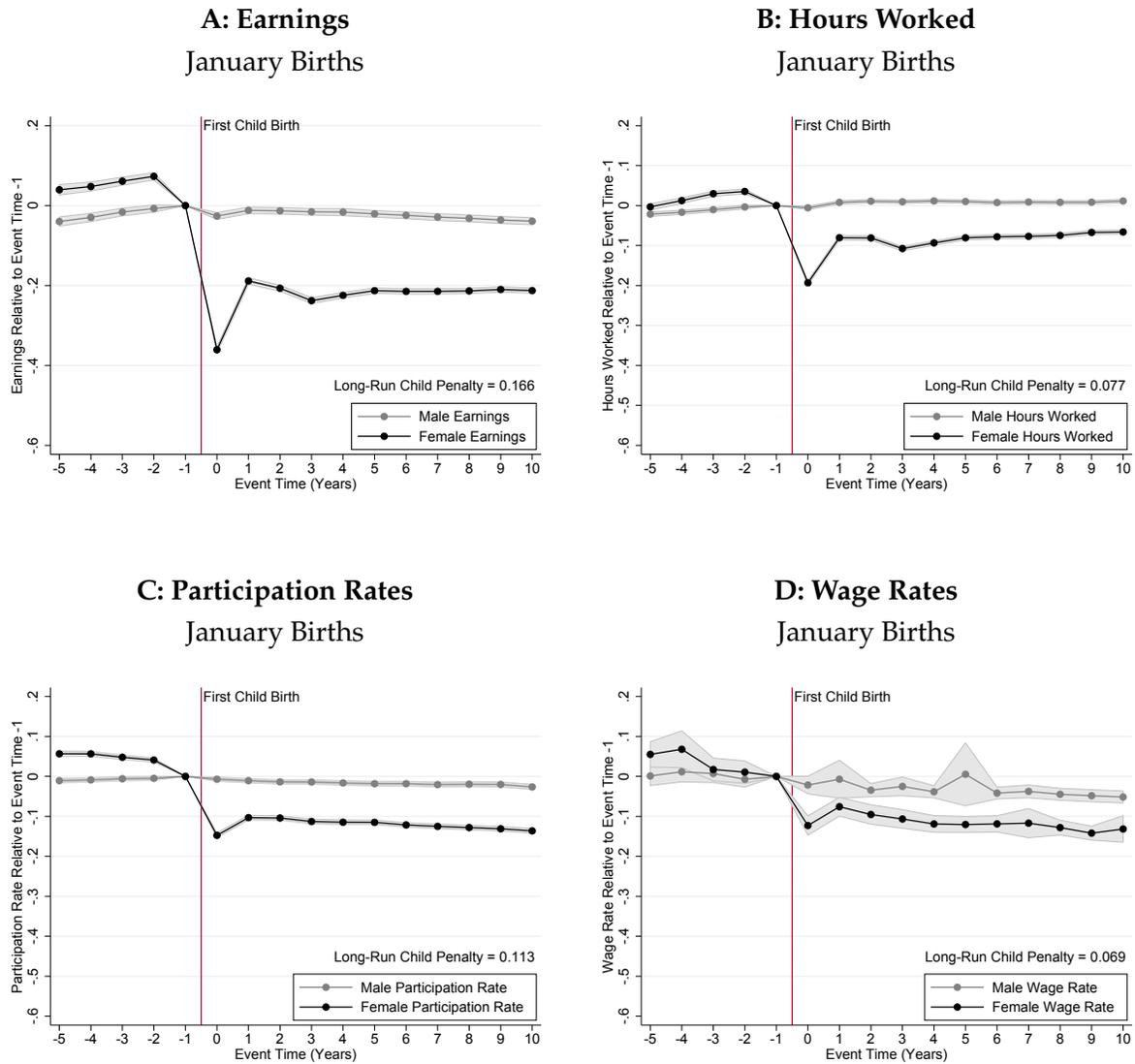


B: Total Labor Supply



Notes: The panels are constructed in the same way as in Figure 3, but showing median impacts on earnings and total hours worked (including the zeros, thereby combining the intensive and extensive margins). These quantile regressions are based on a 1/7 subsample, which makes the confidence bands somewhat larger.

Figure A.III: Impacts of Children Born in January



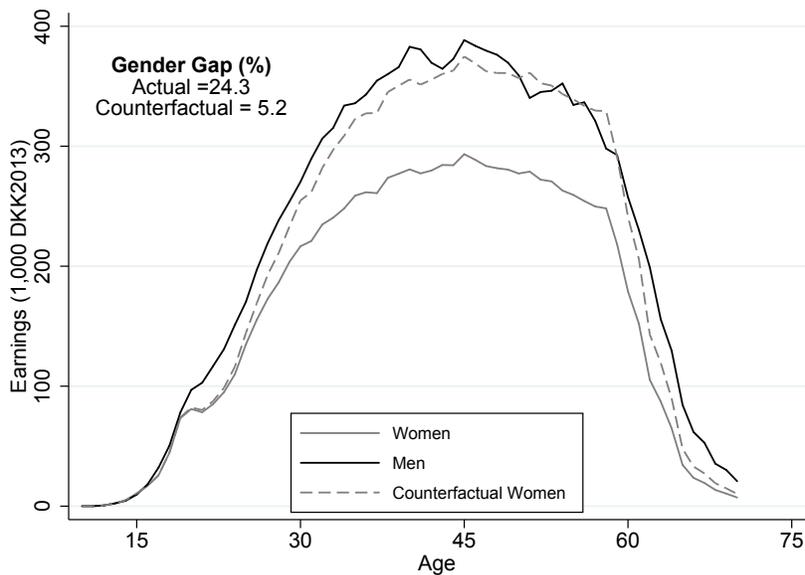
Notes: This figure is constructed exactly as Figure 3, but estimated on the subsample of individuals who have their first child in January.

Figure A.IV: Decomposing the Age Profiles of Gender Inequality in Earnings

A: 1985-Profile Decomposed



B: 2013-Profile Decomposed

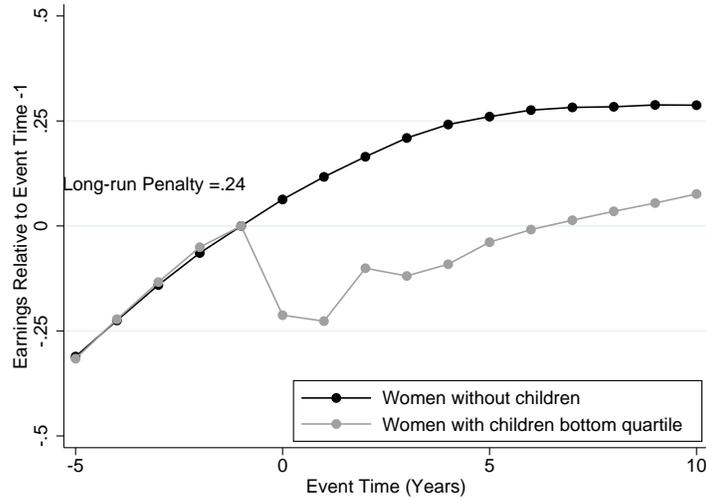


Notes: Based on the child penalties estimated from specification (8), this figure decomposes the within-year age profiles of gender inequality in earnings for 1985 and 2013. The difference between actual male earnings (solid black) and counterfactual female earnings (dashed grey) corresponds to non-child inequality, while the difference between counterfactual female earnings (dashed grey) and actual female earnings (solid grey) corresponds to child-related inequality. These within-year age profiles aggregate to the averages shown in Figure 9A for those two years.

Figure A.V: Earnings Impact of Children by Relative Parental Education

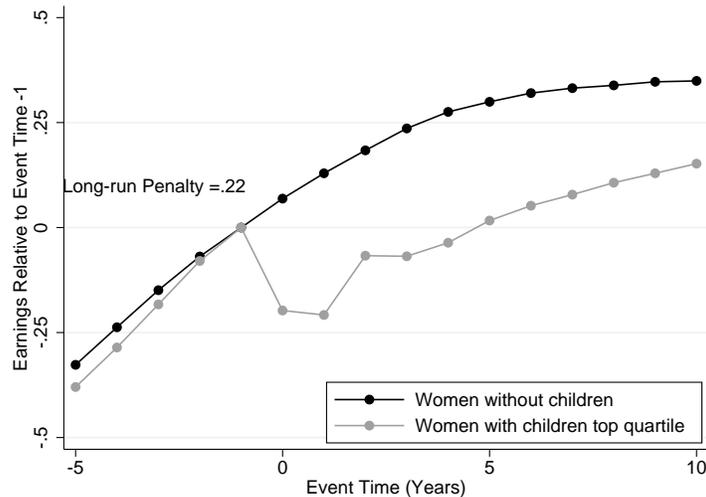
A: Bottom Quartile of Relative Education of Parents

Women Who Have Children vs Women Who Don't



B: Top Quartile of Relative Education of Parents

Women Who Have Children vs Women Who Don't



Notes: The figure shows heterogeneity in the earnings impact of children depending on the relative education of the parents. The analysis is based on the difference-in-differences event study approach of Figure 6, comparing women who have children to women who never have children. Education is divided into the following six levels: primary school, secondary school, vocational training, short post-secondary school, bachelor's degree, and master's/phd degrees. Panel A shows impacts on women in the bottom quartile of the distribution of relative parental education, while Panel B shows impacts on women in the top quartile of this distribution. There is almost no heterogeneity in the child penalty by relative education. The shaded 95 % confidence intervals are based on robust standard errors.

Table A.I: Intergenerational Correlation of Labor Supply and Earnings: Rank-Rank Regressions

	Total Hours Rank			Earnings Rank		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Women						
Hours Rank Own Mother	0.068*** (0.0011)	0.066*** (0.0012)	0.051*** (0.0012)			
Hours Rank Mother-in-Law	0.046*** (0.0010)	0.042*** (0.0012)	0.029*** (0.0012)			
Earnings Rank Own Mother				0.106*** (0.0012)	0.077*** (0.0014)	0.049*** (0.0011)
Earnings Rank Mother-in-Law				0.078*** (0.0012)	0.050*** (0.0014)	0.025*** (0.0011)
Panel B: Men						
Hours Rank Own Mother	0.050*** (0.0010)	0.049*** (0.0012)	0.043*** (0.0012)			
Hours Rank Mother-in-Law	0.028*** (0.0010)	0.026*** (0.0012)	0.017*** (0.0012)			
Earnings Rank Own Mother				0.091*** (0.0012)	0.064*** (0.0014)	0.044*** (0.0012)
Earnings Rank Mother-in-Law				0.067*** (0.0012)	0.036*** (0.0014)	0.013*** (0.0012)
<u>Controls:</u>						
Grandparental		×	×		×	×
Parental			×			×
<i>N</i>	5779928	4592396	4592396	5779928	4592396	4592396

Notes: Robust standard errors in parentheses. Grandparental controls include a full set of cohort dummies for each grandparent, 22 dummies for the education level/fields of each grandparent, and dummies for the quartiles of the wealth rank of the grandfather. Parental controls include a full set of age dummies for each parent, 22 dummies for the education level/fields of each parent, completed fertility of the household by 2013, the number of children in the household at the time of observation, the income level of the spouse, and dummies for the region of residence of the household.