ARTICLE IN PRESS

Journal of Economic Dynamics & Control xxx (xxxx) xxx

Contents lists available at ScienceDirect

Journal of Economic Dynamics and Control

journal homepage: www.elsevier.com/locate/jedc

The parental pay gap over the life cycle: Children, jobs, and labor supply

Claudia Goldin [a], Sari Pekkala Kerr [b,*], Claudia Olivetti [c]

[a] Harvard University and NBER, United States
[b] Wellesley College, United States
[c] Dartmouth College and NBER, United States

ARTICLE INFO

Keywords:
Gender gaps
Motherhood penalty
Fatherhood premium

ABSTRACT

Women earn less than men, and that is especially true of mothers relative to fathers. Much of the widening occurs after family formation when mothers reduce their hours of work. But what happens when the kids grow up? To answer that question, we estimate three earning gaps: the "motherhood penalty," the "price of being female," and the "fatherhood premium." When added together these three produce the "parental gender gap," defined as the difference in earnings between mothers and fathers. We estimate (log) earnings gaps for college graduates born around 1960 using longitudinal data from the NLSY79 and from the LEHD-Census that track respondents from their twenties to their fifties. As the children grow up and as women work more hours, the motherhood penalty is greatly reduced. But women, especially mothers, seem willing throughout their working lives to trade lower pay for various amenities, such as working in firms with management practices that are less penalizing of career interruptions or of shorter work schedules. Fathers, however, manage to expand their relative earnings gains as their children age, particularly among those working in time-intensive jobs, irrespective of work hours or firm fixed effects. The parental gender gap in earnings remains substantial over the family lifecycle.

Part of the journey of life is having a meaningful career while nurturing a family. These often occupy the same time slots and, for most employments, that creates conflict (Goldin, 2021). Mothers often reduce their hours at work and occasionally leave employment for some time or shift into less time-intensive jobs and firms. Those who plan to reduce their hours may invest in careers that impose lower penalties for work with fewer and less demanding hours; or they can move to firms that do not greatly penalize employment interruptions but have lower pay.

These realities are the main parts of an important and well-explored reason why women, especially those with a college degree, earn less than men in the decade or more following the birth of their first child. Less well-explored is what happens to women's careers when the children grow up, require less parental attention, and eventually leave home. That is the subject of this paper.

A large and internationally diverse literature has demonstrated that men and women have divergent earnings growth paths after the birth of a child, even when they were previously on the same career trajectory. That conclusion holds within couples and also when

* Corresponding author.
E-mail address: skerr3@wellesley.edu (S.P. Kerr).

https://doi.org/10.1016/j.jedc.2024.104963
Received 29 July 2024; Accepted 30 August 2024
Available online 4 September 2024

comparing mothers to fathers more generally.[1] Moving to family-friendly workplaces reduces the widening in parental earnings, but a large gap remains.[2] There is also evidence that the possibility of motherhood impacts career choice and educational investment, and also that the career cost of children influences the timing of the first birth.[3]

Much of the initial divergence between male and female earnings after a birth is due to a reduction in the hours of paid work of mothers. But a cascading often follows. Fewer hours at work when young result in less lucrative clients, fewer published papers, a lower probability of promotion, and reduced odds of making partner or obtaining tenure, to provide a few examples.[4] Moving to a lower-paying firm for family-friendly benefits and management practices may improve the chance of climbing the career ladder, but at the cost of lower earnings.[5] In addition, with fewer fully-active years of experience, even a static human capital model would predict lower future earnings.

Thus, career trajectories between mothers and fathers, and between women with and without children, diverge. Gender differences in earnings widen for some time after a birth for human capital reasons as well as for those involving various forms of discrimination.[6] In addition, disproportionate demands on women's time relative to men's may continue long after the children are grown, and aging parents frequently add to caring demands that often fall on women.[7]

But, there is a moment when childcare demands lessen and women can assume greater career and workplace challenges and shift into the more demanding jobs and firms. One obvious change, observed in most data sets, is that mothers eventually increase their weekly hours of paid work. Older female physicians (> 44 years), for example, work more hours by medical specialty than their younger (< 45 years) female colleagues, whereas older male physicians work fewer hours than their younger male colleagues.[8]

To understand what happens to the labor supply and earnings of mothers and fathers as their children grow and leave home or, at the very least, need less oversight requires extensive longitudinal data. We use two extraordinary datasets in this pursuit for almost identical birth cohorts: the NLSY79 and the LEHD-Census.

We find, as have many others, that differences in earnings between mothers and fathers are substantial and that a large portion of the gap comes from hours differences. We add to those findings by partitioning the parental gender gap into three effects (the motherhood penalty, the fatherhood premium, and the price of being female) and showing how these evolve as the children grow up.

As mothers work more hours, the motherhood penalty greatly lessens. But the fatherhood premium expands and the penalty for being a woman increases a bit. We find that women select into firms that have lower earnings but probably greater amenities. Fathers and non-fathers, however, are equally mobile between firms. We explore these differences further by parsing individuals by their earliest occupations and categorizing occupations by their time-intensiveness. We find that men who began in time intensive occupations have the greatest fatherhood premium, whereas women in those occupations have the greatest motherhood penalty.

In sum, the initial motherhood penalty is somewhat diminished with time but the fatherhood premium increases. Thus, the parental earnings gap remains large even after the children have grown up.

1. NLSY79 and the LEHD-Census

The NLSY79 allows us to observe men and women born from 1957 to 1964, as they advance into their mid-fifties and, as parents, until their youngest child has graduated from high school. It contains thick information on individuals, their families, and their employment in a longitudinal context that traverses 35 years. But the number of observations is small, especially when restricted to

[1] Angelov, Johansson, and Lindahl (2016) use administrative data from Sweden on couples. Other careful event study estimates of the impact of childbirth on female labor supply and the gender gap in earnings include Kleven, Landais, and Søgaard (2019), who use administrative data from Denmark, and Kleven, Landais, Posch, Steinhauer, and Zweimüller (2019) using a similar methodology for several countries. Cortés and Pan, 2020 use Kleven's methodology and the PSID to track the gender gap in earnings for cohorts having a first birth from the mid-1970s to the 2010s. Kuziemko, Pan, Shen, and Washington (2020) also use the method to shed light on whether women anticipate the career costs of children. Juhn and McCue (2017) find substantial motherhood penalties using the PSID-Gold Standard. Goldin and Mitchell (2017) use US administrative data and find evidence on the impact of births on labor force participation.

[2] Using the NLSY79, Yu and Hara (2021) emphasize the firm as determining part of the parental earnings gap. But the NLSY79 provides scant information on jobs and firms and has a limited sample size. Using administrative US data, Almond, Cheng, and Machado (2023) find a large child penalty for mothers even in female-led firms and firms where the majority of employees are women. Using data for Sweden, Hotz, Johansson, and Karimi (2017) show that moving to family-friendly workplaces improves the earnings of mothers, while reducing the earnings of fathers, but the impact on the parental gender earnings gap is small.

[3] Adda, Dustmann, and Stevens (2015), Herr (2015), and Wilde, Batchelder, and Ellwood (2010) all explore the role of birth timing on human capital investment, career choice, and earnings.

[4] On the critical role of early promotions, see Bronson and Thoursie (2020) who use Swedish data.

[5] In previous work, we showed that the widening of the gender earnings gap around childbirth is due in part to the shift of women to lower-paying firms and in part to their lower career prospects within a firm. See Barth, Kerr, and Olivetti (2021) and Goldin, Kerr, Olivetti, and Barth (2017).

[6] Other reasons for the so-called "motherhood penalty," are both overt and unintentional discrimination by employers, managers, and supervisors. Mothers could, deliberately and discriminatorily, be passed over for promotions, or their direct supervisors could be guilty of "sexist paternalism" that serves to protect the individual but actually harms them.

[7] Fahle and McGarry (2018) document relationships between caregiving and women's employment using data from the Health and Retirement Study (HRS). Shen (2021) finds that Medicaid policies facilitating formal paid home care led to an increase in daughters' hours of paid work.

[8] These findings are from the Community Tracking Survey (restricted use version). For example, in internal medicine, older (> 44 years) female doctors work 51.4 hours/week whereas younger (< 45 years) ones work fewer hours, 48.5. But older (> 44 years) male doctors in internal medicine work 56.2 hours whereas younger (< 45 years) ones work more, 58.5 hours.

college graduates, who are the focus of this work. In addition, the data reveal almost nothing about the firms in which these individuals were employed.

To remedy both shortcomings, we also use the Longitudinal Employer Household Dynamics (LEHD), the US linked employer-employee database. We leverage the unique matches between employees in the LEHD and respondents to the long form of the 2000 Decennial Census. The combination of these two large data sets allows us to follow men and women into their early fifties, due to the longitudinal nature of the LEHD. Our LEHD-Census sample was selected to consist of college graduate men and women born from 1960 to 1962, closely mimicking the demographics of the NLSY79 cohorts.

The LEHD-Census provides a large sample with firm-level information. The NLSY79 has fine-grained detail on education, marriage, children, and work. Both datasets are longitudinal and extend well into the part of the family lifecycle when children become independent. They each allow us to analyze what happens to parental employment and earnings when children grow up.[9]

The NLSY79 (US Department of Labor, BLS 2019) is an extraordinary longitudinal sample now in its 43 year.[10] It began in 1979 with around 13,000 14- to 22-year old male and female respondents born from 1957 to 1964. These respondents have been followed until today, with some attrition and sample changes. To have as complete a work history as possible, we employ a reasonably balanced panel of individuals whose last interview was 2018. In this work, we analyze data for only those who earned a four-year college degree by age 35.[11] Male and female respondents are included even if they never became parents for the duration of the longitudinal sample.

Our college graduate sample includes 42,880 person-year observations for those aged 25 to 59 at some point in the survey; 22,297 are for women. The sample has 1321 individuals (683 women and 638 men). College graduates are about a quarter of all NLSY individuals.

Respondents become part of our analysis sample after they had worked at least 20 h per week on average for 26 weeks per year during two consecutive years. They remain in the sample if they are equivalently employed for at least 20% of the time remaining to 2018. These sample restrictions eliminate periods with intermittent and brief employment spells and reduce the sample by just 7% for women and 2% for men.[12]

Given the sample selection criteria, the total number of person-year observations in the regression sample is 36,458, of which 17,741 are for women. The sample has 1260 individuals (635 women and 625 men). About 72% of the college graduate women had at least one birth by the end of our sample and the median age of their first birth was 29 years.[13] Almost 76% of the college graduate men became fathers in the duration of the survey and their median age at the birth of their first child was 31 years.

The LEHD is a quarterly linked employer-employee database covering more than 95% of employment in the US and containing information on employment and earnings, as well as firm characteristics. But the LEHD contains only scant person and household characteristics. However, individual person identifiers allow its linkage to the 2000 Decennial Census long-form file (constructed as a 1-in-6 household survey), which provides rich detail. It is worth noting that the details on household composition, education, occupation, and other person traits pertain only to year 2000.

Our LEHD-Census sample consists of college men and women in the 1960–62 birth cohorts who were, therefore, aged 38 to 40 in year 2000.[14] These are the central cohorts in the NLSY79 and thus these two datasets cover a nearly identical group. We obtain from the 2000 Decennial Census, information on their years of education, number and years of birth of their children, and their occupation. The birth years of the children present in the 2000 household are used to generate a longitudinal household composition while the longitudinal earnings, on a quarterly basis, are obtained from the LEHD and include the years 1991 to 2014.[15] Logarithmic earnings are given by the average quarterly earnings across all jobs held by an individual in the calendar year.

We begin following these college graduates when they are 29 to 31 years. We exclude observations with very low income (we impose a cutoff of annual earnings less than that corresponding to working 13.33 h per week for 52 weeks per year at the prevailing federal minimum wage).[16] This limitation mimics the NLSY79 sampling criteria since the LEHD does not contain information on hours worked.

[9] There had been limited research on the impact of children's ages on parental earnings due to the paucity of extensive longitudinal datasets for the US. Several researchers have used earlier waves of the NLSY79, including Herr (2015) and Wilde, Batchelder, and Ellwood (2010). See also the recent piece by Almond, Cheng, and Machado (2023), which also uses the LEHD linked to the ACS.

[10] The full name of the survey is National Longitudinal Survey of Youth 1979. See U.S. Department of Labor, Bureau of Labor Statistics (2019). Sample design information is available at: https://www.nlsinfo.org/content/cohorts/nlsy79/intro-to-the-sample/sample-design-screening-process.

[11] See Goldin, Kerr, and Olivetti (2022) for a similar analysis that includes non-college graduates.

[12] We also truncate hours of paid employment at 84 per week and remove observations for which annual earnings are less than half the contemporaneous federal minimum wage for full-time workers. Further details about data construction can be found in Appendix 2.

[13] Aggregate data from the CPS June Fertility Supplements have a somewhat higher fraction of college graduate women with at least one birth (74% for those born in 1958 increasing to 76% for those born in 1964). There may be selective attrition in the NLSY79 data or these differences may be due to sampling error.

[14] Based on aggregate data from the CPS June Fertility supplement, we capture 89% of college educated women born 1960-1962 who ever had a child. Of this birth cohort of college graduates, 33% had no births by 2000. By age 40-44 the percentage drops to 25%. To the extent that some of our "non-mothers" had a first birth after 2000, our estimates would overstate the catching up of mothers to "non-mothers" after age 40.

[15] See Appendix 2 for more detail on the construction of the LEHD-Census sample.

[16] For example, in 2000 the threshold would be $3,562, given a federal minimum wage of $5.15.

In total, we have 1.89 million person-year observations from the eight LEHD states where data coverage began by 1991 (CA, CO, IL, IN, KS, MD, PA, and WA). All individuals are followed until 2014 to the extent that they remain working in one of the eight states.[17] Women form 47% of the sample (see descriptive statistics in Appendix Table 2).[18] The sample has around 134,000 unique individuals (70,000 men and 64,000 women).[19]

The college graduate sample includes all who report having at least a BA in 2000. About a quarter of the Census respondents in the focal cohorts had a college degree in 2000. The time span allows us to follow the sample from ages 29–31 to ages 52–54. In the Census year 2000, 67% of the women and 64% of men had at least one child at home. We include men and women who had children living at home in 2000, as well as those who did not. For this LEHD-Census sample we observe fewer parents prior to first birth. By the time we first start tracking them, 52% of women and 50% of men already had at least one child.

2. Labor supply and gender (log) earnings gaps from the NLSY79 and the LEHD-Census

We motivate our central questions with Fig. 1, part A and B, giving the gender ratio of earnings by age. These graphs are based on simple OLS regressions where the dependent variable is log(annual earnings), for a pooled sample of males and females. The graphs give the coefficients from the interaction of the age groups and gender, plus the constant term on female.[20]

Part A compares the results for college graduates and non-college graduates in the NLSY79. It also includes data from the LEHD-Census college sample. The first key observation of importance is that the evolution of gender earnings ratios by age for college graduates in the Census-LEHD sample (dotted line) and NLSY79 (solid line) sample is remarkably similar.

In addition, the NLSY79 comparison by education shows that the gender gap of earnings is initially larger for those without a college degree (dashed line) than for those with a college degree (solid line). But college graduate women quickly lose out relative to college graduate men, and by their early forties, the college graduate gender earnings ratio is greater than for those without a college degree. The main point is that college women do less well over time relative to college men and also in comparison with how non-college women do relative to non-college men. Of course, the men and the women in the non-college group both earn considerably less absolutely than in the college group.

Another important point is that the earnings gap by gender is substantial, even correcting for hours and weeks worked, particularly beyond the late thirties. Compare the two lines in Part B. Relative earnings level off by the early forties with time controls (dashed line) as well as in the fixed effects version with time controls (solid line). But the ratio does not increase.[21] The fixed-effects estimates with time controls, we should point out, result in somewhat larger gender earnings gaps than the OLS results with time controls. The reason, we will see, is that women with children are disproportionately selected from among those with higher earnings.

The findings in Part B are similar to those of many other researchers who find that college-graduate women earn amounts that are close to those of their male colleagues, friends, and even spouses in their first jobs out of college or professional school, but that they lose out considerably with time. That is especially true after having children and also as highly-educated men earn substantial salaries by working long hours. Less-educated women, however, start out in jobs that pay far less relative to the men with comparable education, but they do not lose out as much relative to their male peers in large part because male earnings do not increase much with age.

We use these insights as a jumping off point to investigate the role of children in the gender earnings gap and how differences in earnings between mothers and non-mothers and between mothers and fathers evolve as the children grow up. Our contribution is to understand what parts of the gender gap in earnings across the lifecycle are due to family and what parts are due to differences between men and women regardless of childbearing.

3. Regression specification and decomposition

Our goal is to understand what happens to the gender earnings gap over the lifecycle for women and men with children and to compare their outcomes, including earnings, with those of women and men without children. We produce estimates that allow us to partition the parental earnings gap into three parts: that due to the motherhood penalty, that due to the fatherhood premium, and that from being a woman.

To investigate the different components of the parental earnings gap, we estimate panel regressions that include age dummies and, for those who have children, the age of the (currently) youngest child. The samples for all regressions are pooled, containing both males and females. Because the regressions are highly saturated there are only minor differences in the coefficients of interest from those in identical regressions estimated separately by gender.

Our main specification is a variant of eq. (1) where y_{it} is an outcome, such as log annual income or average weekly hours, for

[17] Individuals in our sample may exit the data if they shift their place of employment outside the group of eight LEHD states in our sample. To check whether selective out-migration might bias our results we re-estimated all models for the sub-sample of individuals who are present in the data for a high share (50%, 80%, etc.) of all available years. The results are remarkably robust across the different `presence thresholds.'

[18] The distribution by number of children and age of the youngest child is similar between the NLSY79 and LEHD-Census samples.

[19] All numbers are rounded according to the disclosure avoidance requirements and therefore counts may not sum exactly.

[20] The annual unemployment rate is included since year fixed effects cannot be due to the limited number of cohorts. Controls are not included for number and ages of children.

[21] A related analysis using synthetic cohorts is in Goldin (2014); see also Juhn and McCue (2017), who construct approximately the same synthetic cohort analysis.

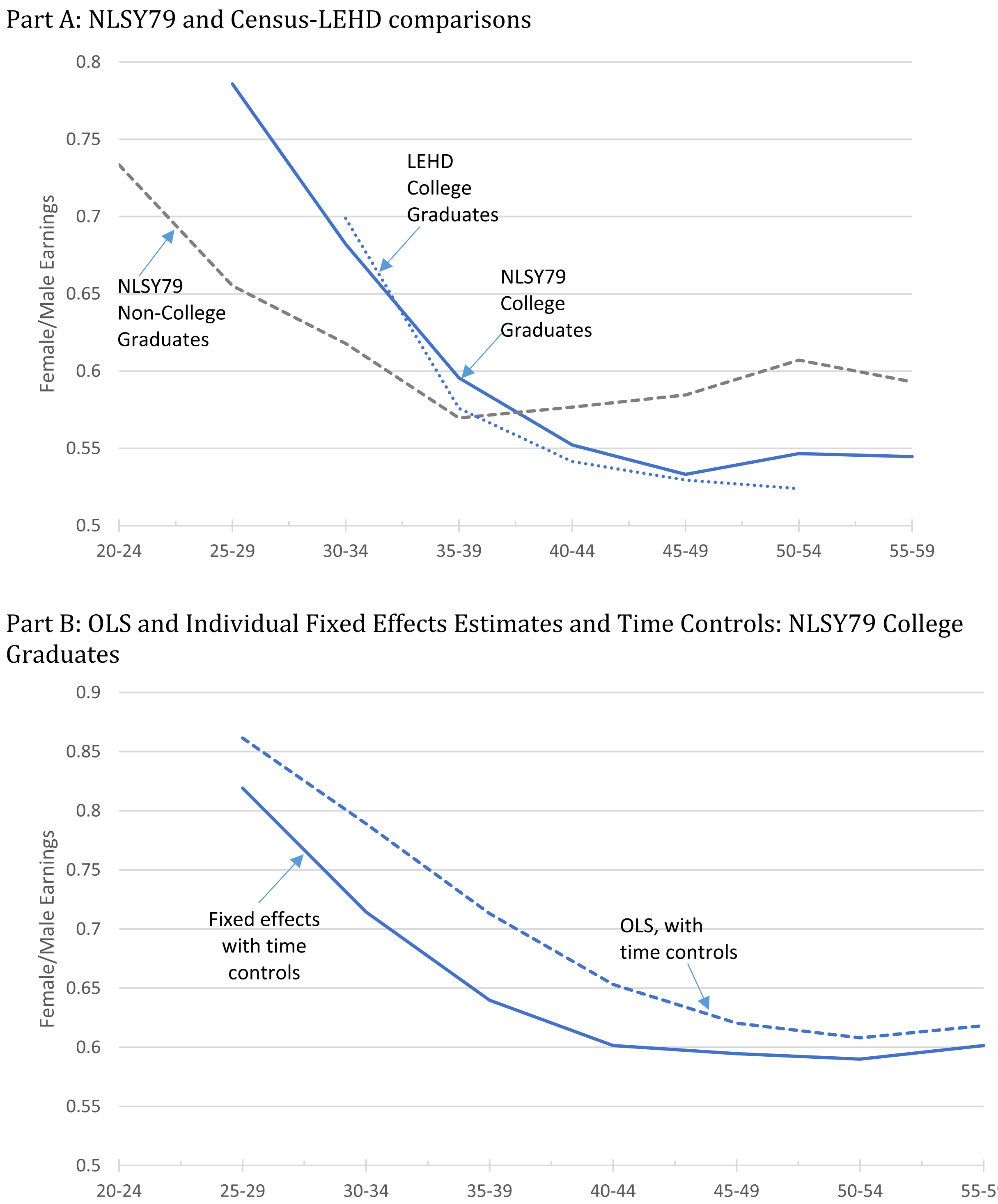

Fig. 1. Female/Male annual earnings. Part A: NLSY79 and Census-LEHD comparisons. Part B: OLS and Individual Fixed Effects Estimates and Time Controls: NLSY79 College Graduates.
Sources: Part A: Tables 2 and Table 4 col. (1), part B: Table 2, cols. (1) and (2) and Table 3, col (2).

individual i in year t.

$$y_{it} = \phi_0 + \phi_1 F_i + \phi_2 A'_{it} + \phi_3 \left(A'_{it} \cdot F_i\right) + \boldsymbol{\alpha}_1(\mathbb{K}'_{it}) + \boldsymbol{\alpha}_2(\mathbb{K}'_{it} \cdot F_i) + \psi U_t + \varepsilon_{it} \tag{1}$$

Included in all OLS estimations are: a female dummy (F_i); a vector of five-year age groupings $\left(A'_{it}\right)$ and their interaction with the female dummy, $(A'_{it} \cdot F_i)$. Also included are a vector of child variables ($\mathbb{K}'_{it}$), that contains the total number of (biological) children born up to that year, and the relevant child age bin of the youngest child at that point ($0 < 3$; $3 < 6$; $6 < 12$; $12<18$; 18 plus). The child age bins reflect a variety of milestones that impact childcare (e.g., end of diapers; entrance to elementary school; high school graduation).[22] Child variables are also interacted with female ($\mathbb{K}'_{it} \cdot F_i$). The coefficients in bold are vectors, not scalars.

When the outcome y_{it} in eq. (1) is log(earnings) we add the term $\boldsymbol{\beta} \cdot \mathbb{Z}'_{it} + \boldsymbol{\delta} \cdot \mathbb{X}'_{it}$. The term includes a vector of time variables ($\mathbb{Z}'_{it}$) and, in the vector $\mathbb{X}'_{it}$, experience and advanced degree. The time variables $\mathbb{Z}'_{it}$ are hours and weeks (in logs) in the NLSY79, and quarters worked in the LEHD-Census. In the NLSY79 earnings regressions, we add a measure of work experience and a dummy variable for whether the individual earned an advanced degree above the bachelor's, for the college graduate sample ($\mathbb{X}'_{it}$).[23] In the LEHD-Census regression we also include firm size measured as the log of employment in the dominant firm j where individual i works at time t.[24]

In all regressions, we add the national unemployment rate (U_t) in year t to account for the impact of the macroeconomy. Due to the limited number of cohorts in the NLSY79 and in the LEHD-Census sample (by construction), we cannot include year effects because of collinearity among cohort, age, and year. Finally, we add state fixed effects (S_j) in the LEHD-Census regressions.[25]

All regressions in (1) are estimated as both cross sections (OLS) and with individual fixed effects. For the OLS regressions, the error term (ε_{it}) is assumed to be i.i.d. The individual fixed effects estimations use the same variables as in the OLS, except that the female dummy is dropped and the error term is $\varepsilon_{it} = \nu_i + \epsilon_{it}$, where ν_i are individual fixed effects and ϵ_{it} is assumed to be i.i.d. For mothers and fathers this specification is identified off differences in the number and age composition of the youngest child over the family cycle.[26] Because we have establishment level information and individuals move across establishments and firms over time, we can also add firm fixed effects in the LEHD-Census regressions. In that case, the error term becomes $\varepsilon_{it} = \nu_i + f_j + u_{it}$, where f_j is the firm fixed effect and u_{it} is i.i.d. Firm fixed effects capture the impact of firm traits for an individual's earnings growth coming from both job-to-job transitions and from within-firm career opportunities. We will emphasize the individual (and firm) fixed effects results.

To investigate the different components of the parental earnings gap, we decompose predicted outcomes by age bins into three components that compare mothers to non-mothers, fathers to non-fathers and non-mothers to non-fathers.[27] Formally:

$$\left(\widehat{y}_t^{Moms} - \widehat{y}_t^{Dads}\right) = \left(\widehat{y}_t^{Moms} - \widehat{y}_t^{NonMoms}\right) - \left(\widehat{y}_t^{Dads} - \widehat{y}_t^{NonDads}\right) + \left(\widehat{y}_t^{NonMoms} - \widehat{y}_t^{NonDads}\right) \tag{2}$$

4. Motherhood penalty fatherhood premium "Price of being female"

The decomposition uses the parameter estimates from (1) and the family structure of the average parent in the data by parent's age group (Appendix Table 1, part A). The main parameters of interest for our decomposition are the estimated OLS or fixed effects coefficients on the adult age category (interacted with female)—the ϕs—and the coefficients on the children's variables—the αs. Denoting by $\overline{\mathbb{K}}_t'$ the child variables for the average parent in our regression sample at age t, the predicted outcomes by age category and parental status are given by:

$$\begin{aligned}
\widehat{y}_t^{Moms} &= (\widehat{\phi}_0 + \widehat{\phi}_1) + (\widehat{\phi}_2 + \widehat{\phi}_3)A'_t + (\widehat{\alpha}_1 + \widehat{\alpha}_2)\left(\overline{\mathbb{K}}_t'\right) \\
\widehat{y}_t^{Dads} &= \widehat{\phi}_0 + \widehat{\phi}_2 A'_t + \widehat{\alpha}_1\left(\overline{\mathbb{K}}_t'\right) \\
\widehat{y}_t^{NonMoms} &= (\widehat{\phi}_0 + \widehat{\phi}_1) + (\widehat{\phi}_2 + \widehat{\phi}_3)A'_t \\
\widehat{y}_t^{NonDads} &= \widehat{\phi}_0 + \widehat{\phi}_2 A'_t
\end{aligned} \tag{3}$$

Using the notation in eqs. (3), the differential impact of children for mothers relative to fathers is given by: $\widehat{y}_t^{Moms} - \widehat{y}_t^{Dads} = \widehat{\phi}_1 + \widehat{\phi}_3 A'_t + \widehat{\alpha}_2\left(\overline{\mathbb{K}}_t'\right)$, conflating gender and the gendered impact of children. The motherhood effect (the first term on the RHS) nets out the gender component because all are women: $\widehat{y}_t^{Moms} - \widehat{y}_t^{NonMoms} = (\widehat{\alpha}_1 + \widehat{\alpha}_2)\overline{\mathbb{K}}_t'$. The last term, the "price of being female," differences the motherhood and fatherhood effects and includes only the net impact of gender: $\widehat{y}_t^{NonMoms} - \widehat{y}_t^{NonDads} = \widehat{\phi}_1 + \widehat{\phi}_3 A'_t$.

In the analysis that follows we will discuss our results for the decomposition and study how the different components vary once we

[22] We use only biological children because of the difficulty of determining the precise birth year of adopted children.

[23] Experience is defined as the fraction of the past five years that the individual was employed fewer than 20 hours per week on average per year. Advanced degrees, for the college graduate group, include all above the bachelor's (e.g., MDs, JDs, MBAs) earned to that year.

[24] For individuals with multiple jobs, the dominant firm is the one that paid the most earnings during the calendar year. Firm size (number of employees) is measured as of March 12.

[25] State fixed effects are identified off individuals who move across the eight states in our sample. We include state fixed effects out of concern that factors like the gender pay gap or the share of new mothers who exit the labor force might vary across states. However, including or excluding state fixed effects does not significantly impact our results.

[26] Because we are interested in comparing men and women with and without children the event study framework in Kleven, Landais, and Søgaard (2019) and Kleven and Landais (2017) is not ideal, since it identifies parental earnings differential by using the heterogeneity in the timing of the first birth in a sample of mothers and fathers. Also, because our time frame is long, we would clearly violate a key assumption of that framework, that women who have their first child at an early age and those who have their first later do not differ in unobservable ways.

[27] All non-mothers are female and all non-fathers are male.

control for time or firm variables, depending on the data sample.

5. Changes in labor supply in the NLSY79: weeks and hours worked

Many researchers have discovered that the role of children in impacting women's earnings across the lifecycle, and thus the gender gap in earnings, is largely determined by changes in labor supply. Hours of work decrease after a birth and stay low for some time. Weeks per year, a measure of labor force participation, also decrease, but the primary labor supply response is at the intensive margin of hours.

We estimate two regressions to understand lifecycle labor supply. The first is the number of weeks worked in the year, and the other is weekly hours, excluding zeros. Cols. (1) and (2) of Table 1 use the data in cross section (OLS) and cols. (3) and (4) employ individual fixed effects. Given that the LEHD-Census sample does not include time variables, the analysis is focused on the NLSY79 sample.

Larger and more persistent labor supply responses occur at the intensive margin of hours, especially when there are young children. The impact of children on hours of work can be seen using eq. (2), which decomposes parental gaps into three parts: that due to the motherhood penalty, that due to the fatherhood premium, and that from being a woman. We follow individuals as they advance through the five-year age groups from 25–29 to 55–59. The predictions by age, gender, and presence of children use the family structure of the average parents in the data for each age group (Appendix Table 1, Part A) and the estimated regression coefficients from Table 1.

Fig. 2 displays a stacked bar graph of the overall parental hours gap ($\widehat{y}_t^{Moms} - \widehat{y}_t^{Dads}$) and its components as described in eq. (2). Mothers work about ten fewer hours than do fathers (the bar height), or about one and a half days fewer a week, from their late twenties to late thirties, when the youngest child of most is still in pre-school. They work around eight fewer hours when they are 45 to 49 years old and the youngest is predominantly in middle and high school. When mothers are in their fifties and their youngest has likely graduated high school (60% of the parents aged 50–54 and 88% of the parents aged 55–59 have a youngest child older than 17 years), they still work almost six fewer hours than do fathers.

What drives the persistent difference in work hours between mothers and fathers? Fig. 2 decomposes the parental hours gap as shown in eq. (2). The light blue area represents the motherhood premium, the dark blue area is the negative of the fatherhood premium, and the grey area is the female effect. Relative to women without children, mothers reduce their weekly hours by about seven—or about one day a week—from their late twenties to late thirties. When they are in their early forties, and their youngest is already in elementary school (79% of those mothers have a youngest child older than six years) the difference is reduced to around five hours and is diminished further to about two hours when they are in their early fifties and the youngest has graduated high school. In the oldest age group, the difference between mothers and non-mothers is under two hours. Hours of fathers, however, are about the same as those of non-fathers in the fixed effects estimation, difference of less than one hour per week by the time they are in their fifties.[28]

Thus, about 70 percent of the six hours difference in work hours between mothers and fathers is because women work fewer hours than do men (about four fewer hours when they are in their late fifties), irrespective of children. Why all women work fewer hours than men, may have something to do with non-child care responsibilities or their planning for children who never materialized, or other preferences regarding labor supply. Note that, comparing the two panels, the hours differences are generally somewhat larger in the fixed-effects estimations, except for the comparison between fathers and non-fathers when the difference shrinks substantially, indicating that fathers typically work more hours than non-fathers.

The main point from this estimation, is that hours of paid work initially plummet with motherhood. Hours stay lower for mothers than non-mothers but increase as the youngest child begins school and eventually exits high school. Since hours exclude the zeros, one can also add in the impact from zero weeks worked during the year, although that will be small relative to the hours decline conditional on working. Mothers, therefore, do increase their work time as the children grow up but they are still far behind fathers. What all that means for the earnings of mothers relative to other women, and in comparison, to fathers, are the next items to consider.

6. The parental (log) gender gap in earnings over the lifecycle: NLSY79

The OLS estimations of log annual earnings are given in Table 2 and the related individual fixed-effects estimations are in Table 3. The estimation of log(annual earnings) in col. (1) of each table includes the age group variables and their interaction with gender, as well as the main gender effect in the cross-section estimation. Col. (2) adds the time dimension (weeks and hours in logs). Col. (3) excludes time but adds the child effects in the same manner as in the labor supply regressions of Table 1. Col. (4) adds back in the time dimension (hours and weeks in logs), and, finally, col. (5) includes a measure of low- or no-work experience and whether the respondent earned a degree beyond the bachelor's.[29]

Fig. 3 gives the impact of children on earnings for the three main fixed effects models that include the child variables: without the time variables (Table 3, col. 3), with the time variables (Table 3, col. 4), and with time, previous five years' work experience, and advanced degrees (Table 3, col. 5). The calculations assume, as we did before, that parents have children at the mean rate for women in the sample.

[28] In the OLS results, men with children put in more paid work hours, especially starting in their early forties.

[29] Recall that low- or no-work experience is defined as the fraction of each of the previous five years that the individual worked on average 20 hours or less per week. About 12% of college graduate women in their thirties and forties had low- or no-work experience in the previous five years.

Table 1
Weeks worked and weekly hours: NLSY79 college graduates.

	OLS		Individual Fixed Effects	
	(1) # Weeks Worked	(2) Weekly Hours (non-zero)	(3) # Weeks Worked	(4) Weekly Hours (non-zero)
Female[a]	−0.0727	−2.709***	0.036	−3.141
	(0.251)	(0.311)		
Age Groups				
30–34	1.221***	1.786***	1.340***	2.169***
	(0.220)	(0.321)	(0.264)	(0.385)
35–39	1.996***	2.178***	2.209***	2.874***
	(0.211)	(0.356)	(0.290)	(0.474)
40–44	1.802***	1.053***	2.063***	2.020***
	(0.222)	(0.356)	(0.349)	(0.566)
45–49	2.066***	1.008***	2.324***	1.994***
	(0.223)	(0.371)	(0.374)	(0.663)
50–54	2.071***	0.426	2.442***	1.613**
	(0.236)	(0.397)	(0.409)	(0.747)
55–59	1.218***	−0.436	1.643***	0.520
	(0.277)	(0.500)	(0.477)	(0.861)
$F\times$ Age Groups				
$F\times$ 30–34	0.0617	0.575	−0.0172	0.165
	(0.323)	(0.452)	(0.395)	(0.594)
$F\times$ 35–39	0.189	0.0389	0.0421	−0.702
	(0.311)	(0.504)	(0.447)	(0.765)
$F\times$ 40–44	−0.255	0.538	−0.456	−0.553
	(0.332)	(0.495)	(0.518)	(0.825)
$F\times$ 45–49	−0.448	0.734	−0.678	−0.302
	(0.336)	(0.514)	(0.568)	(0.946)
$F\times$ 50–54	−0.0395	1.754***	−0.564	0.530
	(0.341)	(0.562)	(0.589)	(1.056)
$F\times$ 55–59	0.0754	0.714	−0.526	0.0209
	(0.403)	(0.713)	(0.687)	(1.223)
Children (age of youngest)				
# children	−0.00211	1.307***	−0.104	0.0794
	(0.0768)	(0.148)	(0.203)	(0.321)
Ch $0<3$	1.094***	−0.763*	0.980**	−0.0834
	(0.219)	(0.397)	(0.384)	(0.664)
Ch $3<6$	0.841***	−0.892**	0.781*	−0.0906
	(0.236)	(0.444)	(0.437)	(0.716)
Ch $6<12$	0.635***	−0.0987	0.487	0.746
	(0.223)	(0.430)	(0.476)	(0.817)
Ch 12<18	0.591***	−0.400	0.294	0.119
	(0.218)	(0.429)	(0.515)	(0.864)
Ch 18+	0.218	0.361	−0.219	0.625
	(0.222)	(0.458)	(0.544)	(0.947)
$F\times$ children				
$F\times$ # children	−0.116	−2.357***	0.171	−1.362**
	(0.143)	(0.245)	(0.359)	(0.588)
F × Ch $0<3$	−3.111***	−4.506***	−4.062***	−5.507***
	(0.403)	(0.614)	(0.718)	(1.160)
F × Ch $3<6$	−0.754*	−3.452***	−1.793**	−5.071***
	(0.402)	(0.701)	(0.772)	(1.224)
F × Ch $6<12$	−0.765**	−2.828***	−1.568*	−3.985***
	(0.381)	(0.658)	(0.808)	(1.349)
F × Ch 12<18	−0.498	−0.399	−0.937	−0.737
	(0.381)	(0.668)	(0.860)	(1.450)
F × Ch 18+	−0.279	0.820	−0.506	0.537
	(0.380)	(0.693)	(0.901)	(1.577)
Unemployment rate in year t	−0.0282	−0.0934*	−0.0375	−0.112**
	(0.0286)	(0.0505)	(0.0318)	(0.0518)
Constant	48.73***	45.35***	48.91***	44.66***
	(0.256)	(0.393)	(0.261)	(0.412)
Observations	36,458	36,458	36,458	36,458
R-squared	0.024	0.091	0.021	0.031
# Individuals			1260	1260

[a] The female main effects in the fixed-effects estimation were recovered.
Source: NSLY79 (U.S. Department of Labor, Bureau of Labor Statistics, 2019.) For sample description, see text.
Notes: Omitted age group is 25–29 years. For other variable definitions and sample selection details, see notes to Table 2.
*** $p<0.01$.
** $p<0.05$.

* $p < 0.1$.

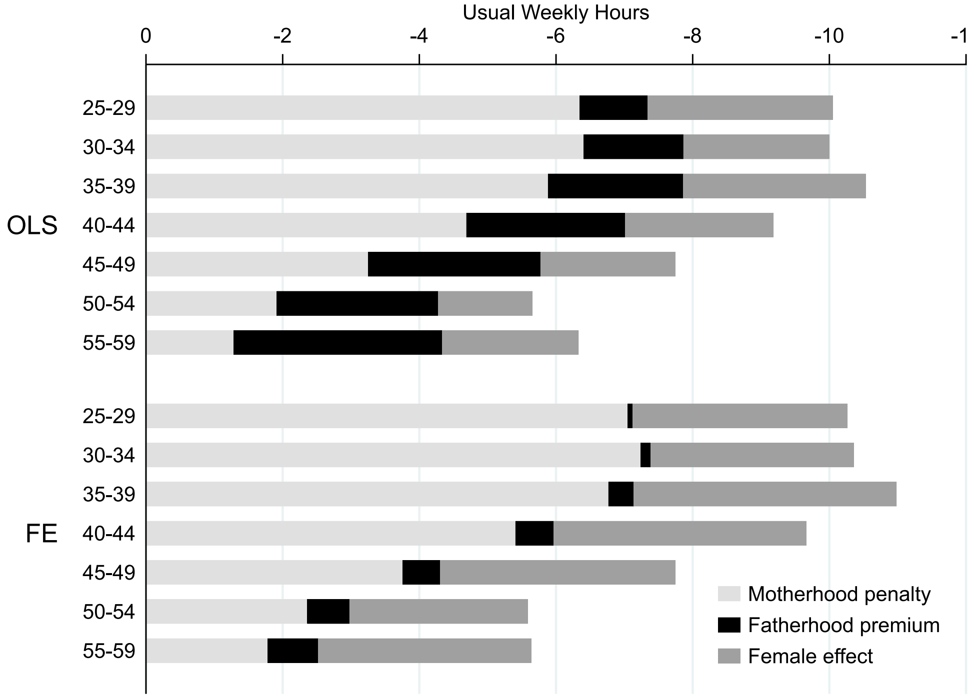

Fig. 2. The impact of children on paid work: NLSY79.
Source: Table 1, col. (2), OLS (the left panel), and col. (4), individual fixed effects (the right panel). *Notes*: The decomposition uses the mean number and age distribution of children by age of the mother. See Appendix Table 1, part A. The bar represents the overall parental gap in hours worked. Each bar decomposes the gap into the part from children (number and ages) for women (the *motherhood penalty*) and men (the negative of the *fatherhood premium*) and the part from gender differences in hours profiles by age (the *female effect*).

As in Fig. 2 the stacked bar graph decomposes the overall parental gap. The bars give the estimated impact of children on mothers' versus fathers' annual earnings. This is partitioned into the three components given in eq. (2): the *motherhood penalty* (light blue), (minus) the *fatherhood premium* (dark blue) and the *price of being female* (gray), which is often termed the gender wage residual since it is unexplained by the included variables.

As shown in Fig. 3, left panel, the log earnings differences without the time variables are extremely large, bottoming out around the late thirties when most parents have a pre-school and slightly older child. Holding hours and weeks constant, as in the middle panel, the parental gender gap is reduced in magnitude but is still substantial. When parents are in their late thirties to early forties, a college graduate mother earns less than 60 cents on a similar father's dollar ($e^{-0.58} = 0.56$). Adding the experience and advanced degree variables (the dark blue bar) increases her relative earnings, but only slightly. The income gains to having the youngest child graduate from high school are minimal.

Given that hours of work greatly drop for women with young children, it is not surprising that the parental gender gap is enormously large when the children are young. But, the gap remains large even as the youngest child graduates from high school. Plus, even with the same number of contemporaneous hours and previous five years' work experience, as given in the right panel mothers still earn persistently and considerably less than fathers; around 60 percent when they are in their late forties and fifties. What factors contribute to the persistent parental gender earnings gap?

The light blue area shows the pure motherhood effect (what we term the "motherhood penalty") by comparing the earnings of women with and without children, at that moment. The immediate impact of children on earnings, through the channel of fewer hours and weeks, is clear by comparing the left panel (not holding hours and weeks constant) with the middle panel (holding the time variables constant). Decreased hours and weeks account for about two-thirds of the difference in earnings of mothers relative to non-mothers when they are in their late twenties and thirties, and about a half of the difference in their early forties. Over time as mothers' work hours catch up to non-mothers', the importance of this channel declines. By the time they are in their fifties, the differences in hours and weeks worked explain about one quarter of the difference in earnings of mothers relative to non-mothers.

The right panel add variables for the previous five years of low hours or non-work experience and the presence of an advanced degree. Holding all of that constant, those with their youngest child mainly in pre-school and elementary-school ages (say mothers in their late thirties) earn 12 log points less than women without children. By the time the youngest is out of high school (say mothers in their late fifties) the difference diminishes to about 7 log points (see Table 5, col. 2) and is not statistically significant.

Table 2
Male and female pooled OLS estimations of log (Annual earnings): NLSY79 college graduates.

	Log (Annual Earnings)				
	(1)	(2)	(3)	(4)	(5)
Female (F)	−0.241***	−0.149***	−0.163***	−0.118***	−0.143***
	(0.0184)	(0.0156)	(0.0189)	(0.0160)	(0.0158)
Age Groups					
30–34	0.332***	0.269***	0.252***	0.205***	0.179***
	(0.0183)	(0.0159)	(0.0186)	(0.0164)	(0.0159)
35–39	0.598***	0.509***	0.452***	0.392***	0.354***
	(0.0194)	(0.0175)	(0.0203)	(0.0185)	(0.0181)
40–44	0.750***	0.679***	0.572***	0.536***	0.487***
	(0.0204)	(0.0190)	(0.0220)	(0.0206)	(0.0201)
45–49	0.824***	0.740***	0.640***	0.594***	0.541***
	(0.0205)	(0.0185)	(0.0230)	(0.0214)	(0.0211)
50–54	0.870***	0.798***	0.706***	0.674***	0.613***
	(0.0214)	(0.0195)	(0.0252)	(0.0237)	(0.0234)
55–59	0.809***	0.769***	0.672***	0.671***	0.607***
	(0.0281)	(0.0265)	(0.0318)	(0.0301)	(0.0294)
$F \times$ Age Groups					
$F \times$ 30–34	−0.141***	−0.0879***	0.00561	−0.00302	0.00204
	(0.0266)	(0.0229)	(0.0268)	(0.0235)	(0.0230)
$F \times$ 35–39	−0.277***	−0.189***	−0.00373	−0.00273	−0.0100
	(0.0274)	(0.0242)	(0.0285)	(0.0258)	(0.0252)
$F \times$ 40–44	−0.352***	−0.277***	−0.0267	−0.0265	−0.0336
	(0.0292)	(0.0261)	(0.0313)	(0.0283)	(0.0278)
$F \times$ 45–49	−0.388***	−0.328***	−0.0676**	−0.0605**	−0.0633**
	(0.0297)	(0.0261)	(0.0334)	(0.0301)	(0.0295)
$F \times$ 50–54	−0.362***	−0.348***	−0.0846**	−0.102***	−0.103***
	(0.0296)	(0.0263)	(0.0354)	(0.0324)	(0.0319)
$F \times$ 55–59	−0.368***	−0.332***	−0.128***	−0.114***	−0.121***
	(0.0387)	(0.0355)	(0.0446)	(0.0410)	(0.0402)
Children (age youngest)					
# children			0.0853***	0.0611***	0.0573***
			(0.00969)	(0.00945)	(0.00924)
Ch 0 < 3			0.109***	0.105***	0.0944***
			(0.0244)	(0.0233)	(0.0227)
Ch 3 < 6			0.138***	0.139***	0.137***
			(0.0275)	(0.0263)	(0.0258)
Ch 6 < 12			0.131***	0.127***	0.128***
			(0.0273)	(0.0264)	(0.0256)
Ch 12<18			0.130***	0.128***	0.132***
			(0.0294)	(0.0288)	(0.0280)
Ch 18+			0.0381	0.0292	0.0467
			(0.0310)	(0.0302)	(0.0299)
$F \times$ # children			−0.235***	−0.176***	−0.145***
			(0.0158)	(0.0144)	(0.0140)
F × Ch 0 < 3			−0.0583	0.0934***	0.0689**
			(0.0385)	(0.0348)	(0.0340)
F × Ch 3 < 6			−0.105**	−0.0172	−0.0132
			(0.0447)	(0.0402)	(0.0394)
F × Ch 6 < 12			−0.117***	−0.0644*	−0.0708*
			(0.0423)	(0.0381)	(0.0369)
F × Ch 12<18			−0.0603	−0.0624	−0.0838**
			(0.0440)	(0.0408)	(0.0398)
F × Ch 18+			0.0990**	0.0549	0.00676
			(0.0448)	(0.0414)	(0.0410)
Time					
Log hours		0.814***		0.783***	0.582***
		(0.0219)		(0.0218)	(0.0227)
Log weeks		0.551***		0.542***	0.510***
		(0.0290)		(0.0286)	(0.0272)
Education, Experience					
Advanced degree					0.266***
					(0.00874)
Frac out last 5 years					−0.734***
					(0.0349)
Unemploy. rate in year t	−0.0123***	−0.00924***	−0.0123***	−0.00923***	−0.00991***
	(0.00340)	(0.00312)	(0.00334)	(0.00307)	(0.00299)
Constant	10.89***	5.666***	10.84***	5.775***	6.668***
	(0.0255)	(0.136)	(0.0254)	(0.135)	(0.136)

(*continued on next page*)

Table 2 (*continued*)

	Log (Annual Earnings)				
	(1)	(2)	(3)	(4)	(5)
Observations	36,458	36,458	36,458	36,458	36,458
R-squared	0.202	0.333	0.230	0.348	0.384

Robust standard errors in parentheses.
*** $p < 0.01$.
** $p < 0.05$.
* $p < 0.1$.
Source: NSLY79 (U.S. Department of Labor, Bureau of Labor Statistics, 2019.) For sample description, see text.
Notes: Sample truncates hours at 84 per week and imposes minimum annual earnings of half the 2009 federal minimum wage × 1400 h per year. All earnings data are in 2019 dollars.
M, F: Male, female.
Age: Omitted age group 25–29 years. The NLSY79 became biennial after 1994. In consequence, there are fewer respondents in their mid- to late-fifties for the 2014 and 2016 waves.
Children (Ch): Children are those born to the woman (or fathered by the man) by the age given. Adopted children's year of birth was not always available. Age of child is the age of the youngest. The number of children is top-coded at three.
Adv degree: All advanced degrees above the bachelor's.
Frac out last 5 yrs: Share of the past five years not working > 20 h per week on average for the year is our measure of experience.
Unemploy. rate in year t: Unemployment rate is used instead of year dummies.
NLSY79 2018 wt are used for all years.

Thus, mothers greatly narrow the earnings gap with women who have not yet had, or will never have, children. But that occurs, in part, because they increase their hours and partly compensate for lost job experience. One can clearly see that as the children get older and become more independent, college graduate mothers make up for lost time relative to other women. There is a distinct U-shaped relationship between the age of the mother (thus the age of the youngest child) and the motherhood penalty. The earnings penalty to women from having children is large but decreases.

As large as is the motherhood penalty for women with younger children, the parental gender gap in earnings is considerably greater and remains large. As seen in Fig. 3 the parental gender earnings gap hovers around 60 percent starting when men and women are in their late thirties to late fifties. The motherhood penalty can explain 12 (in their late thirties) to 7 (in their late fifties) log points of the difference.

What accounts for the remaining 42 to 53 log points?[30] There are two primary factors that reveal the crux of why mothers earn far less than fathers. First, all women get paid less than same-age men and this difference increases over time. Women aged 35–39 get 22.6 log points less than same-age men.[31] Women in their late fifties (who never had children) earn about 30 log points less than same-age (and childless) men, even controlling for time and experience.

The remainder (now in excess of 19 log points) comes from the fatherhood premium (dark blue). Fathers enjoy increased earnings relative to non-fathers with differences in hours worked or experience contributing a tiny fraction of this differential (not surprisingly, since, as we saw in Fig. 2, fathers work less than one hour more than non-fathers throughout their, observable, lifetimes.) By their late thirties fathers' earnings are about 20 log points larger than non-fathers'. The fatherhood premium increases by an additional 3 log points by the time men are in their fifties.[32]

There may be astonishment at this finding. Not only do women lose earnings by their years of raising children, but men earn a premium that goes above and well beyond what could be explained by gender differences in work hours and labor force attachment. A curious finding, for future exploration, is the increased penalty after age fifty from being female (col. 4). Whether that is due to increased demands from aging parents that fall disproportionately on daughters is not yet clear.

7. The parental (log) gender gap in earnings over the lifecycle: LEHD-Census

In previous work we showed that firm characteristics play a role in explaining the widening of the gender earnings gap during child-bearing years (Barth et al., 2021 and Goldin et al., 2017).[33] Next, we leverage the LEHD-Census sample to investigate whether sorting of employees into high and low performing firms might explain some portion of the parental gender earnings gap. It could if women are tied stayers or tied movers and are employed in suboptimal locations.

[30] These calculations are the difference between col. (1) and col. (2) in Table 5 computed at ages 35-39 and 55-59.
[31] The (recovered) college graduate female main effect (see Table 3, col. 5) is –17.4 log points and, in addition, 35-39 year old women earn 5.31 log points less than the female base age group.
[32] See estimates for age 35-39 and 55-59 in Table 5, col. (3). The fatherhood premium at age 35-39 (see Table 3, col. 5 and Appendix Table 1) is 13 log points (= 0.0677 × 1.93) for the number of children plus about 6.5 log points from the premium to having a youngest child of the various ages.
[33] The growing literature using linked employer-employee data studies the evolution of the gender pay gap within and across firms and finds linkages, for example, to firm size and the average pay level of the firm (e.g., Card, Cardoso, and Kline, 2016; Hara, 2018; Jones and Kaya, 2023; Song, Price, Guvenen, Bloom, and von Wachter, 2019).

Table 3
Male and female pooled fixed effects estimations of log (Annual earnings): NLSY79 college graduates.

	Log(Annual Earnings)				
	(1)	(2)	(3)	(4)	(5)
Recovered Female Dummy[a]	−0.2666	−0.1995	−0.1832	−0.1506	−0.1736
Age Groups					
30–34	0.338***	0.293***	0.279***	0.239***	0.218***
	(0.0197)	(0.0169)	(0.0214)	(0.0186)	(0.0188)
35–39	0.605***	0.543***	0.493***	0.440***	0.408***
	(0.0252)	(0.0229)	(0.0296)	(0.0267)	(0.0273)
40–44	0.754***	0.705***	0.613***	0.572***	0.533***
	(0.0285)	(0.0264)	(0.0368)	(0.0342)	(0.0348)
45–49	0.809***	0.754***	0.653***	0.607***	0.567***
	(0.0302)	(0.0277)	(0.0431)	(0.0398)	(0.0402)
50–54	0.850***	0.803***	0.690***	0.649***	0.608***
	(0.0332)	(0.0306)	(0.0488)	(0.0456)	(0.0462)
55–59	0.834***	0.811***	0.676***	0.659***	0.618***
	(0.0407)	(0.0373)	(0.0554)	(0.0512)	(0.0514)
$F \times$ Age Groups					
$F \times$ 30–34	−0.180***	−0.137***	−0.0326	−0.0335	−0.0279
	(0.0302)	(0.0252)	(0.0317)	(0.0271)	(0.0266)
$F \times$ 35–39	−0.317***	−0.247***	−0.0581	−0.0485	−0.0521
	(0.0365)	(0.0316)	(0.0443)	(0.0390)	(0.0386)
$F \times$ 40–44	−0.370***	−0.309***	−0.0824	−0.0671	−0.0673
	(0.0418)	(0.0369)	(0.0536)	(0.0481)	(0.0479)
$F \times$ 45–49	−0.362***	−0.320***	−0.0996	−0.0815	−0.0800
	(0.0447)	(0.0392)	(0.0631)	(0.0564)	(0.0553)
$F \times$ 50–54	−0.338***	−0.328***	−0.119*	−0.114*	−0.113*
	(0.0458)	(0.0408)	(0.0686)	(0.0627)	(0.0620)
$F \times$ 55–59	−0.326***	−0.309***	−0.133*	−0.115	−0.126*
	(0.0556)	(0.0492)	(0.0795)	(0.0712)	(0.0707)
Children (age youngest)					
# children			0.0720***	0.0688***	0.0677***
			(0.0240)	(0.0225)	(0.0215)
Ch 0 < 3			0.0519	0.0436	0.0351
			(0.0465)	(0.0433)	(0.0411)
Ch 3 < 6			0.0846*	0.0781*	0.0757*
			(0.0502)	(0.0468)	(0.0447)
Ch 6 < 12			0.0878	0.0781	0.0756
			(0.0534)	(0.0496)	(0.0472)
Ch 12<18			0.103*	0.0989*	0.0951*
			(0.0593)	(0.0552)	(0.0530)
Ch 18+			0.0971	0.0933	0.0905
			(0.0691)	(0.0651)	(0.0630)
$F \times$ # children			−0.195***	−0.170***	−0.142***
			(0.0422)	(0.0382)	(0.0363)
F × Ch 0 < 3			−0.131*	0.00491	0.00572
			(0.0761)	(0.0666)	(0.0633)
F × Ch 3 < 6			−0.178**	−0.0773	−0.0425
			(0.0827)	(0.0730)	(0.0680)
F × Ch 6 < 12			−0.162*	−0.0932	−0.0726
			(0.0880)	(0.0777)	(0.0731)
F × Ch 12<18			−0.0696	−0.0548	−0.0587
			(0.0954)	(0.0848)	(0.0803)
F × Ch 18+			0.0429	0.0316	0.00554
			(0.106)	(0.0951)	(0.0909)
Time					
Log hours		0.542***		0.525***	0.359***
		(0.0281)		(0.0278)	(0.0258)
Log weeks		0.432***		0.431***	0.408***
		(0.0273)		(0.0271)	(0.0255)
Education, Experience					
Adv degree					0.200***
					(0.0289)
Frac out last 5 years					−0.719***
					(0.0548)
Unemploy. rate in yr t	−0.0106***	−0.00834***	−0.0104***	−0.00810***	−0.00856***
	(0.00275)	(0.00246)	(0.00272)	(0.00245)	(0.00241)
Constant	10.77***	7.068***	10.78***	7.126***	7.835***
	(0.0231)	(0.147)	(0.0241)	(0.146)	(0.138)

(continued on next page)

Table 3 (*continued*)

	Log(Annual Earnings)				
	(1)	(2)	(3)	(4)	(5)
Observations	36,458	36,458	36,458	36,458	36,458
R-squared	0.246	0.360	0.263	0.370	0.402
# individuals	1260	1260	1260	1260	1260

[a] The female main effects in the fixed effects estimation were recovered.
Sources and Notes: See Table 2.
*** $p < 0.01$.
** $p < 0.05$.
* $p < 0.1$.

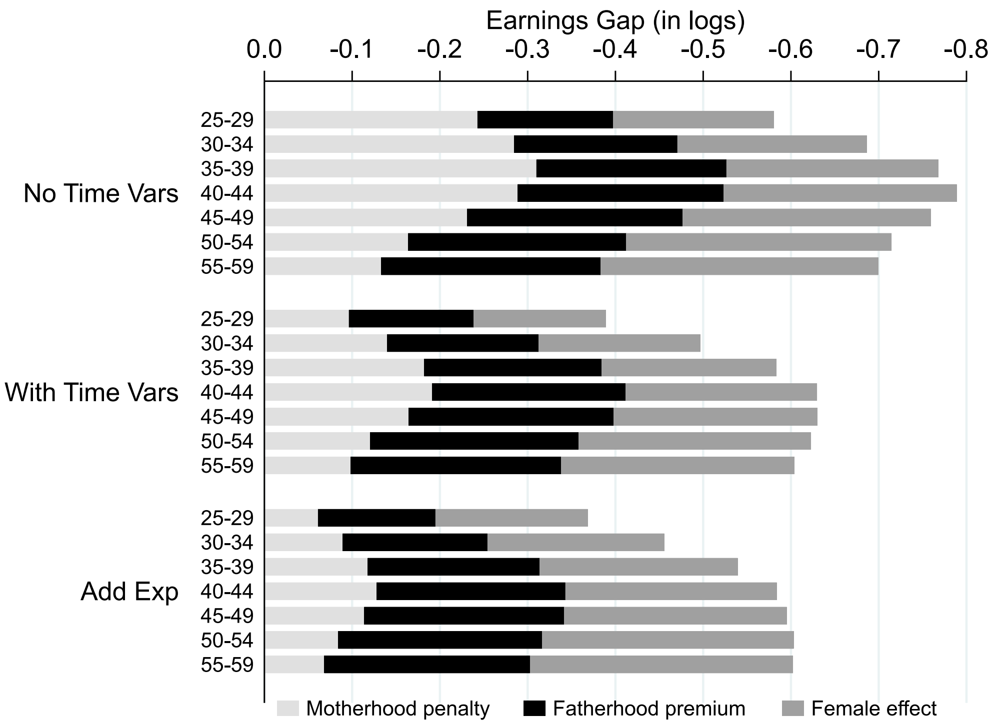

Fig. 3. The impact of children on the earnings gap: NLSY79.
Sources: Estimates are for the fixed effects regressions in Table 3, no time variables (the left panel) from col. (3); with time variables (the mid panel) from col. (4); with time variables, experience, and advanced degrees (the right panel), from col. (5). *Notes*: Time variables are log hours and log weeks. Experience is the fraction of the last five years that the individual worked > 20 h per week on average. The decomposition uses the mean number and age distribution of children by age of the mother. See Appendix Table 1, part A. The bar represents the overall parental gap in earnings. This is decomposed into the part from children (number and ages) for women (the *motherhood penalty*) and men (the negative of the *fatherhood premium*) and the part from gender differences in hours profiles by age (the *female effect*).

The dependent variable in Table 4 is log annual earnings computed as the sum of quarterly earnings across all jobs held by an individual in the calendar year. We report the decomposition results in Fig. 4, which uses the fixed effects estimates from Table 4 and the family structure of the average parents in the regression sample from Appendix Table 1, part B.

Fig. 4, part A, shows a subtle U-shaped relationship between the age of the mother (thus the age of the youngest child) and the parental gender earnings gap. The finding is similar to that from the NLSY79 without the time controls, where the parental earnings gap peaked at 40–44 and then slightly contracted when parents are in their forties and fifties (see Fig. 3, part A and Table 5, part A). In the Census-LEHD fixed effects regressions the gap widens by about 10 log points when parents are in their late thirties and early forties, but it is the same at age 30–34 (when most children are in pre-school age or younger) as it is at age 50–54 (when most children are in high school or older). Controlling for firm fixed effects reduces the gap by about 20 percent across all age groups.[34]

Fig. 4, part B decomposes the parental gender earnings gap in its three components. The first panel in part B, shows that for the

[34] Women in the LEHD-Census will be more attached to the labor force by the design of the sample, and we observe children only in 2000.

Table 4
Male and female pooled estimations of log (Annual earnings): LEHD-Census.

	Log (Annual Earnings)				
	(1) OLS	(2) OLS	(3) Person FE	(4) +Estab Traits	(5) + Estab FE
Female (F)	−0.326***	−0.166***	−0.189	−0.129	−0.151
	(0.0020)	(0.0024)			
Age Groups					
35–39	0.279***	0.199***	0.211***	0.190***	0.185***
	(0.0023)	(0.0023)	(0.0023)	(0.0022)	(0.0014)
40–44	0.439***	0.347***	0.352***	0.323***	0.327***
	(0.0023)	(0.0025)	(0.0034)	(0.0031)	(0.0017)
45–49	0.512***	0.431***	0.415***	0.387***	0.406***
	(0.0023)	(0.0027)	(0.0041)	(0.0038)	(0.0019)
50–54	0.526***	0.464***	0.427***	0.403***	0.441***
	(0.0027)	(0.0032)	(0.0048)	(0.0044)	(0.0021)
$F \times$ Age Groups					
$F \times$ 35–39	−0.154***	−0.041***	−0.050***	−0.046***	−0.040***
	(0.0031)	(0.0031)	(0.0034)	(0.0032)	(0.0019)
$F \times$ 40–44	−0.212***	−0.053***	−0.060***	−0.052***	−0.045***
	(0.0032)	(0.0035)	(0.0049)	(0.0046)	(0.0023)
$F \times$ 45–49	−0.211***	−0.057***	−0.052**	−0.042***	−0.033***
	(0.0033)	(0.0038)	(0.0059)	(0.0054)	(0.0026)
$F \times$ 50–54	−0.194***	−0.066***	−0.0055**	−0.043***	−0.034***
	(0.0037)	(0.0044)	(0.0068)	(0.0062)	(0.0030)
Children (age youngest)					
# children		0.072***	0.055***	0.054***	0.058***
		(0.0024)	(0.0031)	(0.0029)	(0.0015)
Ch 0 < 3		0.134***	0.031***	0.021***	0.013***
		(0.0032)	(0.0059)	(0.0054)	(0.0029)
Ch 3 < 6		0.129***	0.058***	0.048***	0.043***
		(0.0036)	(0.0063)	(0.0058)	(0.0031)
Ch 6 < 12		0.116***	0.059***	0.053***	0.058***
		(0.0033)	(0.0068)	(0.0063)	(0.0033)
Ch 12<18		0.085***	0.063***	0.057***	0.071***
		(0.0035)	(0.0075)	(0.0069)	(0.0035)
Ch 18+		0.011***	0.072***	0.066***	0.084***
		(0.0042)	(0.0086)	(0.0079)	(0.0040)
$F \times$ # children		−0.168***	−0.171***	−0.151***	−0.137***
		(0.0018)	(0.0055)	(0.0051)	(0.0024)
F × Ch 0 < 3		−0.063***	−0.118***	−0.0121***	−0.101***
		(0.0048)	(0.0099)	(0.0092)	(0.0043)
F × Ch 3 < 6		−0.100***	−0.166***	−0.155***	−0.123***
		(0.0053)	(0.0107)	(0.0099)	(0.0048)
F × Ch 6 < 12		−0.174***	−0.158***	−0.142***	−0.111*
		(0.0048)	(0.0114)	(0.0105)	(0.0049)
F × Ch 12<18		−0.127***	−0.058***	−0.046***	−0.029***
		(0.0050)	(0.0122)	(0.0112)	(0.0052)
F × Ch 18+		−0.026***	0.031***	0.036***	0.038***
		(0.0057)	(0.0134)	(0.0123)	(0.0058)
Time					
Quarters worked		0.132***	0.100***	0.094***	0.084***
		(0.0010)	(0.0009)	(0.0009)	(0.0005)
Firm traits					
Size				0.020***	0.010***
				(0.0006)	(0.0005)
Mean earnings				0.314***	
				(0.0028)	
Public sector				−0.016***	
				(0.0064)	
Unemploy. rate in year t	−0.0726***	−0.686***	−1.043***	−0.916***	−0.760***
	(0.0415)	(0.0410)	(0.0271)	(0.0256)	(0.0183)
Constant	9.787***	9.214***	9.321***	6.281***	N/A
	(0.0038)	(0.0051)	(0.0071)	(0.0318)	
Observations	1860,000	1860,000	1860,000	1860,000	1860,000
R-squared	0.139	0.175	0.157	0.378	0.848

Robust standard errors in parentheses.
*** $p < 0.01$.
** $p < 0.05$.
* $p < 0.1$.
Sources: LEHD and 2000 Decennial Census (U.S. Census Bureau). For sample description, see text.
Notes: Sample imposes minimum annual earnings corresponding to working 52 weeks at 13.33 h per week at the prevailing federal minimum

wage. All earnings data are in 2014 dollars. All columns control for state fixed effects. Col. (4) also controls for 2-digit NAICS fixed effects.

M, F: Male, female.

Age: Omitted age group 29–34 years.

Children (Ch): Children are those born to the woman (or fathered by the man) by the age given. Adopted children are included. Age of child is the age of the youngest. The number of children is top-coded at three.

Unemploy. rate in year t: Unemployment rate is used instead of year dummies.

Size: Log employment of the firm person works in year t.

Mean earnings: The (log) mean payroll/employment (1991–2014) of the firm the person works in year t.

Public sector: Works in a public-sector entity in year t.

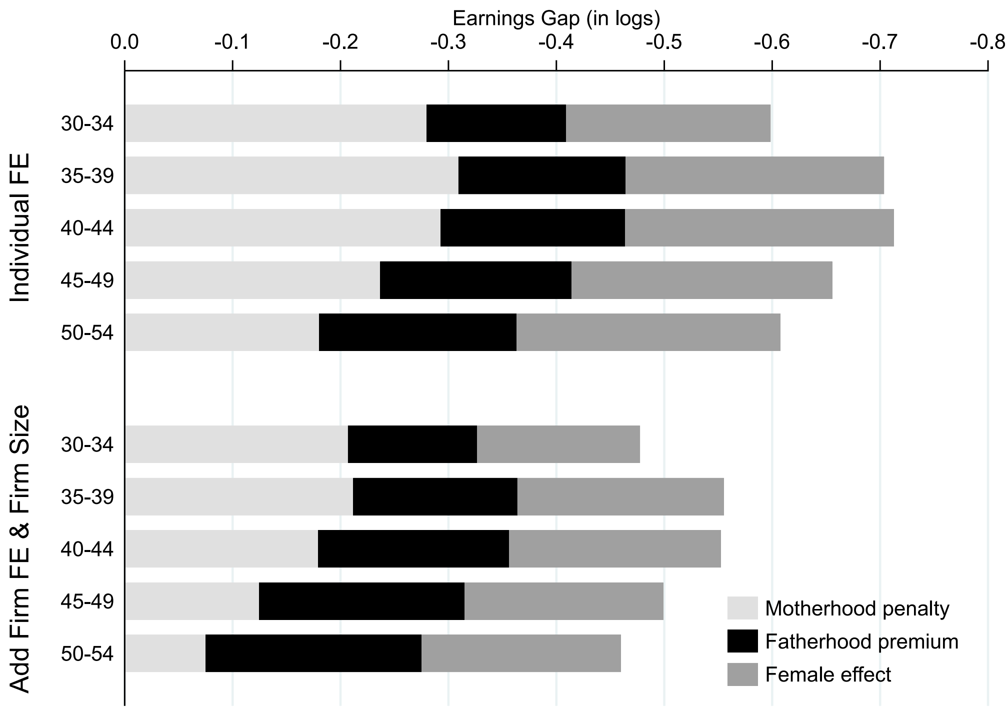

Fig. 4. The impact of children on the earnings gap: LEHD-Census.
Sources: Estimates are for the regressions in Table 4, individual fixed effects (left panel) from col. (3); with, in addition, firm fixed effects and firm size (right panel), col. (4). *Notes*: The decomposition uses the mean number and age distribution of children by age of the mother. See Appendix Table 1, part B. The bar represents the overall parental gap in earnings. This is decomposed into the part from children (number and ages) for women (the *motherhood penalty*) and men (the negative of the *fatherhood premium*) and the part from gender differences in hours profiles by age (the *female effect*).

specification with individual fixed effects (the light blue bars), the earnings gap of mothers relative to non-mothers hovers around 30 log points until age 40–44. It then declines to 18 log points by age 50–54. As we already know from the NLSY79 analysis, some of the decline comes from a reduction in the hours gap between mothers and non-mothers as they (and their children) age.

Comparing the light blue and dark blue bars in the first panel of Fig. 4, part B, suggests that firm sorting can explain part of the motherhood penalty. Firm fixed effects (and firm size) explain a significant proportion of the earnings gap between mothers and non-mothers (from 7.5 to 18 log points when they are in their early fifties). The relative importance of firm-related factors increases as the motherhood penalty shrinks. Comparing the light blue and dark blue bars, the motherhood gap declines by 26% at age 30–34, 38% at age 40–44, and 58% at age 50–54 with firm fixed effects.

The fatherhood premium (the second panel in Fig. 4, part B) is not much affected by firm controls. The premium declines by a negligible amount when men are in their thirties, and it increases by 1 to 1.7 log points when they are in their late forties and early fifties, suggesting that fathers and non-father are equally mobile. Finally, as shown in the third panel in part B, including firm fixed effects reduces the "price of being female" by about 20%, with little variation over the family cycle.

Our Census-LEHD results suggest that systematic differences in how men and women select into firms over the family cycle are important for mothers but do not impact the fatherhood premium. Some firms may be family friendly but they may also pay lower wages.

Based on our analysis it appears that both time variables and firms play an important role in explaining how earnings differentials between mothers and non-mothers evolve as children become more independent. Gender differences in hours worked, or in the types of firms for which men and women work, can explain some of the residual gender gap, but across data sets and with all the controls, the earnings of women in their late fifties are persistently and meaningfully lower (20 to 30 log points) than are those of similarly aged

Table 5
Decomposition of parental gender gap in earnings for NLSY79 and LEHD-Census.

Part A: NLSY79				
Age Group	(1) Parental Gender Gap in Earnings	(2) Motherhood Penalty	(3) Fatherhood Premium	(4) Price of Being Female
25–29	−0.368*	−0.061§	0.134*	−0.174*
30–34	−0.456*	−0.089#	0.165*	−0.201*
35–39	−0.539*	−0.118*	0.196*	−0.226*
40–44	−0.584*	−0.128*	0.215*	−0.241*
45–49	−0.595*	−0.114#	0.228*	−0.254*
50–54	−0.603*	−0.084	0.232*	−0.287*
55–59	−0.603*	−0.068	0.235*	−0.300*
Part B: LEHD-Census				
Age Group	(1) Parental Gender Gap in Earnings	(2) Motherhood Penalty	(3) Fatherhood Premium	(4) Price of Being Female
30–34	−0.478*	−0.207*	0.120*	−0.151*
35–39	−0.555*	−0.212*	0.152*	−0.191*
40–44	−0.552*	−0.179*	0.177*	−0.196*
45–49	−0.499*	−0.125*	0.190*	−0.184*
50–54	−0.460*	−0.075*	0.200*	−0.185*

* $p < 0.01$.
# $p < 0.05$.
§ $p < 0.1$.

Sources: For the NLSY79, Table 3, col. (5) and Appendix Table 1, part A. For LEHD-Census, Table 4, col. (5) and Appendix Table 1, part B.

Notes: The NLSY79 estimates use the results from the individual fixed effects estimation with log(hours), log(weeks), previous five year's work experience, and advanced degrees (for the college graduate sample). The Census-LEHD estimates use the results from the regression with individual fixed effects, firm size and firm fixed effects. All parents are assumed to have children given by the data for women in Appendix Table 1 with regard to number and age of the youngest. Cols. (1) = col. (2) - col. (3) + col. (4).

men. But fathers' earnings advantage relative to men who did not yet or will never have a child is not explained by greater work hours (see Fig. 3, panel B) or whether fathers work for "better" or higher-paying firms than non-fathers (Fig. 4, panel B).

8. The fatherhood premium

There is a large and long-standing literature regarding the fatherhood premium and the male marriage premium. The literature has assessed whether fathers (or married men) earn more because they work harder after they have children (or get married) or, alternatively, whether they become fathers (or get married) when they are earning more. Another possibility is that various principals in the labor market (e.g., supervisors) reward fathers and married men more on the basis of their conception of fairness or their personal preference.

In one of the earliest research articles using an individual fixed effects framework to assess these hypotheses, Korenman and Neumark (1991), found that the earnings profiles of men steepen after marriage and that they receive higher performance ratings from their supervisors. Both findings suggest that men work harder after marriage and that the labor market may also favor them. Studies that followed concur with their conclusion that selection into marriage is less important than the treatment effect of marriage and also of fatherhood. But the jury, according to others, is still out.[35]

The impact of having children, however, is not necessarily the same as that from being married. Some researchers have found a fatherhood premium for US men using OLS.[36] The main concern with OLS is that the premium could stem from self-selection into parenthood, if men with the highest earnings potential were more likely to have children. Several studies for the US using longitudinal data and fixed effects models confirm that possibility. Lundberg and Rose (2000), for example, find a fatherhood premium but only for couples where the wife reduces her hours of work after the birth of the first child. Budig (2014) and Killewald (2013), using the NLSY79, both find a fatherhood premium especially for married, co-resident, and biological fathers.[37][38]

[35] Killewald and Lundberg (2017) provide evidence that the relationship between marriage (and divorce) and earnings is not causal, but comes from unanticipated positive (and negative) shocks. Killewald and Gough (2013) show that even women and men without children earn a marriage premium. Killewald (2013), using the NLSY79, concludes that residential and married fathers have more interest in working for the betterment of their children, consistent with Korenman and Neumark (1991) but more nuanced.

[36] See, for example, Juhn and McCue (2017), which uses the CPS and shows an increase in the fatherhood premium for recent cohort. Although they use fixed effects models in their estimates regarding mothers, they cannot do that for fathers.

[37] A recent piece by Kunze (2020), using Norwegian data on twin brothers, finds that selection into fatherhood was far less important to their earnings than were other aspects of their backgrounds.

[38] Teasing apart the effect of marriage and children is not feasible given our data since marriage and children are highly correlated in our cohorts. Restricting the sample to couples who are continuously married suggests that our findings are not driven by changes in marital status (e.g. divorce when the children `leave the nest').

For the NLSY79 sample, we can study whether the fatherhood premium can be explained by positive self-selection into fatherhood using the Armed Forces Qualification Test (AFQT) scores available in the NLSY79.[39] This information is commonly used in the literature to measure cognitive ability. Our results indicate, across all terciles of the standardized AFQT distribution, that men in our sample are equally likely to marry and/or have children. Similar results hold for mothers. If anything, women in the bottom tercile of the AFQT distribution are less likely to have children, though the difference is not statistically significant at standard levels.

The precise reasons for the positive relationship between men's earnings and fatherhood are important. As illustrated by the OLS and fixed-effect comparison in Fig. 2, even fathers do seem to work more hours than non-fathers, on average, but the fatherhood premium persists holding hours constant. No matter how rich the longitudinal data of the NLYS79 and the Census-LEHD are, they do not allow us to disentangle precisely why college graduate fathers have earnings that rise so much more with age than any other group. We can refute that the reason concerns usual selection issues, and we know that it is not due to sorting into better-paying firms.

Among different sex couples, men are enabled to become fathers while continuing to advance in their careers because women disproportionately take care of the children. Mothers cut back on their hours, work less-demanding jobs, and earn less. But something else must be operating because the earnings gap between women and men without children also grows with age.

For men, having the children and a wife who is the caregiver is related to their earnings boost. Whether it is causal or whether marriage and children result from some exogenous boost is secondary. Put simply: the motherhood penalty becomes small as the children grow up, but the fatherhood advantage remains large and increases with age, especially among college graduates. In both samples we have used, the fatherhood premium accounts for about 40% of the parental gender gap in earnings by the time mothers and fathers are in their fifties.

9. Further exploring the fatherhood premium: time intensity of occupations

One possibility is that the size of the fatherhood premium and its widening with age come disproportionately from fathers with time-intensive jobs. Time-intensive occupations, in our two samples, tend to be found in sectors, such as in finance, law, and healthcare, that hire more highly-educated workers.

To explore the possibility we have coded the time intensity of the occupations of our NLSY79 respondents. We do this for the occupations that each had when they were fairly young. In particular, we use the occupation respondents had immediately before they had their first child, if they ever had a child. For those without children, we predict an age of a pseudo first birth and code the occupation of these individuals as we did for those who had children. We use this early occupation to get their preferred, somewhat unconstrained, occupational choice rather than the one they chose later, possibly to accommodate family demands.

Our coding of the time intensity of the occupation is based on a combination of two measures. One is the fraction of individuals in that occupation working 45 or more hours a week as given in the 1990 Census. We chose the 1990 Census because it is the most relevant data given the year our respondents had their first birth. The other measure we use is an average of five O*NET characteristics that are designed to measure the lack of time flexibility, and thus the time intensity, in the occupation (see Goldin 2014).

Time-intensive occupations are those meeting two criteria: (1) the fraction of workers in the 1990 Census working 45 or more hours per week is in the top third of all occupations for that criterion, and (2) the average of the five O*NET characteristics is also in the top third for that measure. Appendix 3 gives further details on the definition of time-intensive and not time-intensive occupations and the age of the respondents when they were employed in the occupation.

We estimate the parental gender gap in earnings using the same procedure we employed in the previous section. We now partition the sample by whether the individual had an occupation when younger that we deem as being "time intensive." Among college graduates about 31.4% of the total sample had a time-intensive occupation.[40] We use the fixed-effects estimates that include all controls (similar to Table 3, col. 5) and produce a partition of the parental gender gap into the three components, as we did before.

We show in Table 6, part A the results for those with time-intensive occupations and, in part B, for those in occupations that were not deemed time intensive. These results can be compared with the full group, given in Table 5, part A.

The differences in the partitions, as shown in Fig. 5, are striking. The fatherhood premium for the time-intensive group is generally twice the size for the non-time-intensive group.[41] As indicated in the table, the fatherhood premium is precisely estimated in both groups (the confidence intervals around the estimates for the time-intensive and non-time-intensive group barely overlap). The price of being female (essentially the residual) is substantially smaller at older ages for the time-intensive group, and the motherhood penalty is larger. The full parental gender gap in earnings is essentially unchanged.

It would appear that men who had time-intensive occupations when younger were enabled or motivated to work even harder when they had children than were men who were not fathers. Having children may have motivated them to worker longer hours or at more demanding jobs in other ways. Stay-at-home wives or those working part-time may enable their husbands to focus on their careers by easing time constraints or offering advice and motivation. Irrespective of the reasons why fathers work harder, time-intensive

[39] The NLSY79 uses results from a subset of the ten components of the Armed Services Vocational Aptitude Battery (ASVAB) to generate an AFQT score. Details on the construction of the AFQT score can be found here: https://www.nlsinfo.org/content/cohorts/nlsy79/topical-guide/education/aptitude-achievement-intelligence-scores.

[40] Our college graduate regression sample with the time-intensive occupation measure has 1,207 unique individuals or just 53 fewer than the full sample. Of the 609 women, 141 are in the time-intensive group. Of the 598 men, 238 are in the time-intensive group.

[41] The fatherhood premium for college graduates in the non-time-intensive group is the same as for non-college men (see Table 4B in Goldin, Kerr, and Olivetti, 2022).

Table 6
Decomposition of parental gender gap in earnings by the time intensity of the earliest occupation, NLSY79.

Part A: Time-Intensive Occupations				
Age Group	(1) Parental Gender Gap in Earnings	(2) Motherhood Penalty	(3) Fatherhood Premium	(4) Price of Being Female
25–29	−0.333*	0.027	0.192*	−0.168
30–34	−0.355*	−0.012	0.209*	−0.134*
35–39	−0.432*	−0.073	0.237*	−0.122§
40–44	−0.498*	−0.142	0.267*	−0.089
45–49	−0.578*	−0.165	0.286*	−0.127
50–54	−0.622*	−0.141	0.296*	−0.185§
55–59	−0.574*	−0.114	0.302*	−0.158
Part B: Not Time-Intensive Occupations				
Age Group	(1) Parental Gender Gap in Earnings	(2) Motherhood Penalty	(3) Fatherhood Premium	(4) Price of Being Female
25–29	−0.296*	−0.096*	0.068§	−0.132
30–34	−0.400*	−0.118*	0.101*	−0.181*
35–39	−0.481*	−0.130*	0.126*	−0.225*
40–44	−0.514*	−0.127*	0.137*	−0.250*
45–49	−0.511*	−0.114#	0.141*	−0.256*
50–54	−0.523*	−0.085	0.141#	−0.297*
55–59	−0.541*	−0.066	0.144#	−0.331*

* $p < 0.01$.
\# $p < 0.05$.
§ $p < 0.1$.

Source: NSLY79 (U.S. Department of Labor, Bureau of Labor Statistics, 2019).

Notes: Of the original NLSY79 sample of 1260 unique individuals, 1207 (parents and non-parents) have non-missing time-intensive occupations at an early age (and before their first birth). Time-intensive occupations and not time-intensive occupations are determined by the average share of workers with 45+ hours per week in these occupations in the 1990 Census and the average of five normalized characteristics from O*NET (see Goldin 2014). The procedure is described in Appendix 3. Of the 1207 individuals in this analysis, 828 were in occupations that were not time intensive and 379 were in a time-intensive occupation before their first birth or at an equivalently early age. These estimates use the results from the individual fixed effects estimation with log(hours), log(weeks), previous five year's work experience, and advanced degrees (for the college graduate sample). The specification only includes three dummies representing age of the youngest child: less than 6, 6 to 17 and 18 and above.

occupations have highly non-linear earnings with respect to hours worked (Goldin 2014). Extra effort exerted by these men in their twenties and thirties appears to have been disproportionately rewarded later through promotions and other career opportunities.

The same advantage was unavailable to men in the non-time-intensive occupation group.[42] These men may increase their hours when they have children relative to non-fathers, but they are not in occupations that greatly reward increased effort. Rather, they are in occupations with flatter wage-age profiles with less scope for earnings to grow dynamically through promotions and job-hopping.

We also implement a variant of the occupation analysis using the LEHD-Census. Given that the information on children and occupation is available only in 2000 we cannot observe occupation prior to having children, but we do observe detailed industry codes annually from the longitudinal LEHD.

We use that information to divide the three-digit NAICS industries into three terciles based on the share of employees within those industries who work in time-intensive occupations. For each individual we define their industrial time-intensity based on the first year they are observed in a job in the LEHD, which for most people is 1991. We then re-estimate the LEHD models separately for the industry-level terciles, using the initial industry in which each employee is observed. The results of this analysis, presented in Appendix Table 3, reinforce that using the NLSY79. The parental earnings gap is largest for individuals who started their work careers in a time-intensive industry. Both the fatherhood premium and the motherhood penalty grow more in time-intensive industries. The residual wage gap is also largest in the most time-intensive industries (about 6 log points throughout the life cycle.)

10. Summary

Many excellent and well-identified studies have demonstrated that a woman's earnings take a nose-dive directly after the birth of a child. The decrease is mainly, but not entirely, due to a reduction in hours of work. Diminished earnings have been shown in these studies to remain for some time. But what happens when the children grow up?

[42] The results are not for couples. We have assembled data on couple earnings and occupations in the NLSY79 and find that the parental gender (log) earnings gap for the college educated is greatest in families where the husband is in a time-intensive occupation, irrespective of the time intensity of the wife's occupation. The parental gap is the lowest, for all education groups, in families with both (different sex) members of the couple in non-time-intensive occupations.

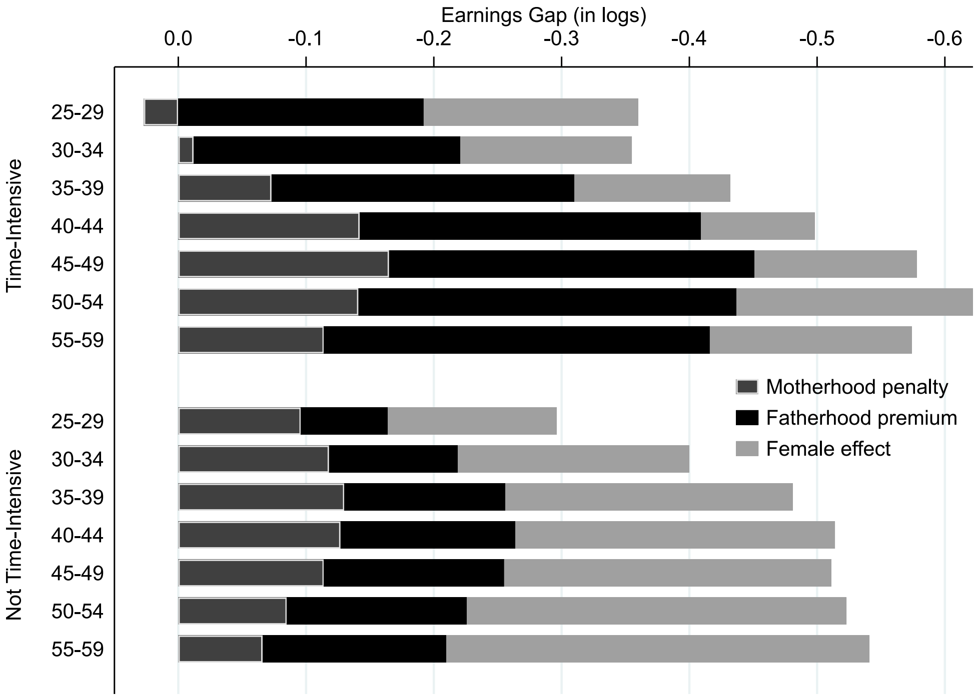

Fig. 5. The impact of children on the earnings gap by the time intensity of the earliest (pre-birth) occupation.
Source: Estimates are for the decomposition in Table 6, time-intensive occupations (left panel) from Panel A; not time-intensive occupations (right panel), panel B. *Notes*: These estimates use the results from the individual fixed effects estimation with log(hours), log(weeks), previous five year's work experience, and advanced degrees (for the college graduate sample). The decomposition uses the mean number and age distribution of children by age of the mother. See Appendix Table 1, part A. The bar represents the overall parental gap in earnings. This is decomposed into the part from children (number and ages) for women (the *motherhood penalty*) and men (the negative of the *fatherhood premium*) and the part from gender differences in hours profiles by age (the *female effect*).

Our contribution has been twofold. We have added many more years of parenthood to analyze the impact of children as they mature and become more independent. We have also analyzed the impact of parenthood by partitioning the effect into three quantities: the motherhood penalty, the fatherhood premium, and the price of being female. Unlike other studies, we define the motherhood penalty as the difference in earnings between women who are mothers and those who are not or who are not yet.

Our findings can be summarized as follows. As the youngest enters grade school and beyond, women's hours increase relative to those of non-mothers and to those of fathers. Mothers narrow the earnings gap relative to women who have not yet had, or will never have, children. College graduate women with children gain 18 log points (0.31 – 0.13) relative to those without children as they move from their late thirties to their late fifties, using the NLSY79 estimation that does not have the time variables. They gain 8 log points (0.18 – 0.10) controlling for hours.[43] They would earn 16 log points more relative to fathers if they could work the same hours and weeks as the fathers, but they don't advance on fathers given hours and weeks. They just hold their relative place.[44]

We have shown using the LEHD-Census that mothers disproportionately sort into firms that are lower paying, but that may have preferable work conditions or other amenities. Men, however, regardless of fatherhood status do not engage in the same type of sorting across firms. Therefore, some portion of the parental wage gap that we have identified using the NLSY79 data is due to the sorting of mothers into firms that are lower paying.

Another portion, we demonstrate, is due to the intensity of work as determined by the occupations individuals had prior to family formation. In the NLSY79, we used the first occupation observed for individuals before the first (actual or placebo) birth. In the LEHD-Census we used a similar design but at the three-digit industry level. In both analyses, the time intensity of the occupation was an important component of earnings differences.

We began with the notion that life is a complicated journey. Parenthood is an important part when mothers slow down, reduce their hours of work, and occasionally leave employment for some time or shift into less time-intensive jobs and firms. But there is a moment

[43] The estimate of 0.31 is from Fig. 3, part A for 35-39 years old college graduate mothers relative to non-mothers. That of 0.13 is for 55-59 years old. Both use the regression with no time variables. Using the regression with time variables, but not experience and advanced degrees, gives 0.18 for 35-39 years old and 0.10 for 55-59 years old.

[44] At age 40-44 college graduate mothers earn −0.789 relative to college graduate fathers in the estimation without time variables and −0.630 with time variables. Therefore, 16 log points = (0.789 – 0.630) is the amount mothers get for working the same number of hours and weeks. But the gender (log) earnings gap between mothers and fathers computed with time variables does not change much with age after 40-44. See Fig. 3, part A.

when childcare demands greatly lessen and women increase their hours of paid work and assume greater career challenges. We think of that moment, metaphorically, as when mothers can race ahead. But even though they increase their hours of work, they do not, on average, attain gender equality with their male counterparts. Their inability to earn the same as fathers is due to several factors: the positive relationship that children have with the earnings of men, their negative relationship with women's earnings, and the price of being female independent of motherhood status. The journey was never among equals.

Data availability

The Decennial Census - LEHD data used are confidential. The rest of the data are publicly available.

Acknowledgments

The authors thank the Russell Sage Foundation (Grant #85-18-05) and the National Science Foudation (NSF) (Grant #1823635) for providing research funding. They are grateful to Lucy Cheskin and Jennifer Walsh for exceptional research assistance, especially during the trying pandemic years. Any views expressed are those of the authors and not those of the U.S. Census Bureau. The Census Bureau has reviewed this data product to ensure appropriate access, use, and disclosure avoidance protection of the confidential source data used to produce this product. This research was performed at a Federal Statistical Research Data Center under FSRDC Project Number 1571. (CBDRB-FY23-P1571-R10417, R10631 and R10926). This research uses data from the Census Bureau's Longitudinal Employer Household Dynamics Program, which was partially the following National Science Foundation Grants SES-9978093, SES-0339191 and ITR-0427889; National Institute on Aging Grant AG018854; and grants from the Alfred P. Sloan Foundation.

Appendix 1

Appendix Table 1

Appendix Table 1
Age distribution and number of children by age of mother.

Part A: NLSY79						
		Fraction with Children by Age, among All Mothers				
Mother's age	Number of children	0 to 2	3 to 5	6 to 11	12 to 17	18+ years
25–29	1.324	0.780	0.150	0.065	0.005	0.000
30–34	1.666	0.588	0.281	0.109	0.021	0.001
35–39	1.930	0.285	0.300	0.347	0.059	0.008
40–44	2.021	0.066	0.146	0.488	0.257	0.040
45–49	2.064	0.004	0.026	0.274	0.490	0.207
50–54	2.082	0.000	0.002	0.048	0.353	0.597
55–59	2.122	0.000	0.000	0.007	0.104	0.888
Part B: LEHD-Census						
		Fraction with Children by Age, among All Mothers				
Mother's age	Number of children	0 to 2	3 to 5	6 to 11	12 to 17	18+ years
30–34	1.614	0.618	0.235	0.132	0.015	0.000
35–39	1.918	0.300	0.257	0.343	0.086	0.007
40–44	2.023	0.022	0.145	0.493	0.275	0.065
45–49	2.070	0.000	0.000	0.243	0.514	0.243
50–54	2.096	0.000	0.000	0.008	0.397	0.603

Sources and Notes: Part A: The sample is the same as that used for Tables 2 and 3. Part B: The sample is the same as that used for Table 4 and Appendix Table 3. Row numbers for the fraction of children by age and age of mother sum to 1.0. The distribution and numbers of children by the age of fathers is almost identical.

Appendix 2: Construction of the regression samples

Part A: NLSY79

The NLSY79 provides an informative work history record for each respondent through constructed week-by-week arrays. These arrays cover the entire duration of a respondent's participation in the survey, including years in which they were not interviewed, through a backfilling procedure undertaken by NLSY staff. There are two week-by-week arrays central to our analysis: the status array and the hours array. The status array reports the job number of the primary jobs worked each week, or other labor force status if applicable. In addition to the status array, the hours array provides actual hours worked each week across all jobs. As described below, this measure also assists with the interpolation of self-reported income.

We consider separately income from wages/salary alone and total income including both wage/salary and any own business/farm income.[45] The measure is collected only when the respondent is interviewed (unlike variables from the week-by-week arrays that are backfilled for missed interview years) and refer to the calendar year prior to the interview. Data are missing from skipped interviews and after 1993 when the survey switches from being annual to being biennial. To recover income in non-survey years and years in which a respondent is not interviewed but works positive hours according to the week-by-week array, we use a simple interpolation of missing values that uses the job and hour information from the week-by-week array. Although the method is imperfect, it allows us to recover a significant share of the missing incomes, particularly in the non-interview years after 1993. The imputation algorithm is described in the online appendix Goldin, Kerr and Olivetti (2021).

Once we have interpolated the income variables, we then impose certain work history and income restrictions to produce the analysis sample. We consider respondents as employed only if they have earning of at least half the amount earned by a full-time, full-year worker making the federal minimum wage applicable in that year. We begin following a respondent's work history when they have worked positive hours for at least 26 weeks per year and at least 20 h on average across the weeks with positive hours for two consecutive years. We employ that restriction to include those who are committed to being in the labor force. We continue to follow such respondents if they meet the requirements just mentioned for at least 20% of the time between their first eligible year (as defined above) and 2018.

To obtain information on the age of the youngest child in each year, we rely primarily on the dates of birth of all biological children of the respondent. Given our ability to follow families over multiple decades in the NLSY79, it is important to note that the youngest child in a given year is not necessarily the same child as the "youngest" child in the previous or subsequent years since children enter the sample and could exit. Thus, the "age of youngest child" may not always increase linearly across survey years.

The NLSY79 provides multiple measures of educational attainment. The first is the highest grade completed. This measure is collected beginning in 1979 and we rely specifically on the revised version of the variable constructed by the NLSY79 staff.[46] The second is the highest degree completed since the date of the last interview. That question appears first in the NLSY79 in 1988. We construct a highest degree completed variable that captures the highest degree earned by the respondent in each year, leveraging the information contained in both the highest degree and the highest grade completed variables. We then determine the year in which each respondent graduated college with a bachelor's degree and construct the advanced degree dummy variable.

Current marital status is determined according to the start and end dates provided for the respondent's first four marriages, if applicable. The respondent is considered "currently married" from the year the marriage began until the end of that marriage.

A fuller description of the regression sample construction is contained in an on-line appendix (Goldin, Kerr, and Olivetti 2021).

Part B: LEHD-2000 Census

To produce the LEHD-2000 Census regression sample we first select all households from the 2000 Decennial Census long-form administered to 1-in-6 households. We exclude respondents living in Group Quarters and only include individuals who are related to the household head (see below for details). The output is a representative sample covering all U.S. states.

Using the "relate" variable in each year of the Census, we assign heterosexual spouses and children, as available, to the respondent (note that in this draft we do not use the spousal information). As the child relationships are critical to our current analysis (and the spouse variable to future analyses), we limit our sample to those who are one of the following: reference person, spouse, mother, father, sister, brother, daughter-in-law, son-in-law, sister-in-law, brother-in-law, biological son or daughter, step son or daughter, adopted son or daughter, grandson, granddaughter, niece, or nephew.

Individuals can be identified using the protected identification key (PIK) that has been generated through the Census Bureau's Person Identification Validation System (PVS) to allow data linkage across the Census data infrastructure.[47] The Decennial Census does not include Social Security Numbers, but instead each respondent has been assigned a PIK through the PVS. The PVS has a very high match rate, but the system is not perfect (about 90% of the Decennial Census respondents can be assigned a PIK).[48] We exclude the small number of duplicate persons and households with the same PIK.

Using the PIK we then match individuals from the Decennial long-form records to the quarterly LEHD. For each state, the LEHD consists of several files describing the establishment, worker, and the worker's job. The employer characteristics file (ECF) provides quarterly establishment level traits such as county location, detailed industry of the establishment, employment, and payroll. The employment history file (EHF) gives for each calendar quarter individual earnings while employed in each company the person worked for. The person-level information contained in the individual characteristics file (ICF) includes the date of birth, gender, race, place of birth, and citizenship status. The ICF is not used in the current project to derive person-level detail as we have chosen, instead, to use the richer information available in the Decennial Census.

The 2014 vintage of the LEHD employed here covers the years 1991 to 2014.[49] The 2014 LEHD snapshot is described in greater technical detail in Vilhuber (2018). It is a vast database of private-sector firms and most public sector entities containing administrative quarterly earnings data from state unemployment insurance (UI) records.[50]

[45] When a respondent has both wage/salary and business/farm income, their total income is the sum.

[46] See the NLSY79 Education Topical Guide for more information.

[47] https://www.census.gov/library/working-papers/2014/adrm/carra-wp-2014-01.html.

[48] Detailed information about the PIK-ing process can be found at: https://www.norc.org/PDFs/May%202011%20Personal%20Validation%20and%20Entity%20Resolution%20Conference/PVS%20Assessment%20Report%20FINAL%20JULY%202011.pdf.

[49] States with coverage starting by 1991 include: California, Colorado, Illinois, Indiana, Kansas, Maryland, Pennsylvania, and Washington State.

[50] UI earnings capture wages, salaries, and taxable bonuses. Earnings are not top-coded. Excluded from the LEHD are self-employment earnings and employees of the Federal government.

The LEHD allows identification of single unit firms (i.e., firms that are one establishment) and the separate establishments of multi-unit firms. The state employer identification number (SEIN) is based on the state tax ID. SEINs are therefore state specific. They generally (but not always) correspond to state UI reporting entities, whereas SEIN reporting units (SEINUNIT) correspond to an establishment. The ECF also provides a federal tax ID (EIN) as well as a firm identifier (FIRMID or "alpha") that allows linkage of the establishments of multi-unit firms.[51] For certain types of firms, the linkage provides a clear mapping of the firm's footprint in the LEHD states. However, it is generally not possible to link establishments that are part of franchises using the LEHD alone. Non-employer firms and individuals filing as self-employed (e.g., contractors and sole proprietors) are not included in the LEHD.

The LEHD is rich in many employment details but is missing information relevant for our study. It does not contain specific information on job characteristics such as occupation, position held, hours worked, or hourly wages. Demographics, family composition, birth dates of children, education, and marital status are not in the LEHD, but come instead from the 2000 U.S. Decennial Census long-form.

Given that we are interested in observing as long a stretch of the family life-cycle as possible for the birth cohorts of interest we only selected the eight LEHD states where earnings data were available starting at least by 1991. We make no restrictions regarding the states in which the individuals resided during 2000, when the Decennial Census was collected. But we do limit the estimation sample to individuals who met the minimum earnings criteria explained in the text. Descriptive Statistics for the sample are below.

Appendix Table 2

Appendix Table 2
LEHD-Census sample, descriptive statistics.

	College Educated	
Panel A: Person, Year	Men	Women
Ln(Quarterly Earnings)	9.97	9.50
Age	41.2	41.4
Number of Children by Year t	1.22	1.18
Number of Quarters Worked	3.72	3.74
Ln(Mean Earnings in Establishment)	9.77	9.54
Ln(Establishment Size)	4.69	4.98
Public Sector Entity (%)	14.1	25.9
NAICS: Education Services (%)	9.9	22.7
NAICS: Health Services (%)	7.0	19.7
NAICS: Finance, Insurance, Real-Estate (%)	9.4	8.3
NAICS: Other Services (%)	25.1	20.4
NAICS: Retail Trade (%)	5.6	4.9
NAICS: Manufacturing (%)	13.3	5.4
NAICS: Construction (%)	3.3	1.2
NAICS: Other (%)	26.4	17.5
Unemployment Rate (%)	6.6	6.6
Sample Size	986,000	874,000
Panel B: Main Job Quarterly Earnings by Age		
25–29	n.a.	n.a.
30–34	$18,730	$13,090
35–39	$28,410	$16,360
40–44	$34,300	$18,570
45–49	$37,660	$19,940
50–54	$39,130	$20,500
Panel C: Establishment Means	Public Employment	Private Employment
Ln(Mean Earnings in Establishment)	9.43	9.47
Ln(Establishment Size)	6.14	4.28

Sources: LEHD 1991–2014 and 2000 Decennial Census.

Appendix 3: Construction of time-intensive occupations and industries

Part A: Construction of Time-Intensive Occupations, NLSY79

The results in Table 6, part A divide the original sample of college graduate men and women into those who had a "time-intensive occupation" at an early age and those who did not. We have produced this sub-sample in the following manner.

For those in our regression sample who eventually became parents, we consider the occupation associated with the job they had one to three years prior to the birth of their first child. For those who never had a biological child (as of 2018), we assign each respondent a predicted date of a "pseudo first birth" based on a regression of year of first birth on gender, race, age at college graduation, and respondent's birth year. Once non-parents are imputed a birth year of a "pseudo" child, we assign them a pre-birth occupation the same

[51] EINs cannot be used to link establishments to firms as the identifier is not always unique within a firm, and is designed for tax-reporting purposes. The EINs within a firm are not structured in any deterministic way.

way we did for the parents.

We determine whether the occupation is time intensive based on several factors. One is the share of workers in that occupation working 45 or more hours (based on the 1990 Census). Another is the average of the five O*NET characteristics in Goldin (2014). Occupations are deemed time intensive if they were in the top tercile of the distribution of both the share working 45 or more hours and the (normed and averaged) O*NET scores. All other occupations are considered to be "not time intensive."

The original sample included 1260 unique individuals of which just 53 did not have a first occupation that could be used in the analysis (26 women, 27 men). Of the remaining 1207 respondents (parents and non-parents) with a non-missing occupation variable before the actual or "pseudo first birth," 379 respondents (31%) had a time-intensive occupation and 828 did not.

Of the 379 with time-intensive occupations, 238 are men and 141 are women. The median age at first birth was 31 for women and 32 for men. About 8% of the men and 7% of the women never married. Of the 828 who had a "not time-intensive occupation," 360 are men and 468 are women. The median age at first birth was 29 for women and 30 for men. About 12% of men and about 10% women never married.

Part B: Construction of Time-Intensive Industries, LEHD-Census

We implement a variant of the NLSY79 occupation analysis for the LEHD-Census and present those results in Table 6, part B. Given that occupation information is available only in 2000, generally after the mothers and fathers in our sample have started their families, we instead use occupational information based on the three-digit NAICS industries provided in the LEHD.

We construct a measure of the time-intensity of the occupations in the three-digit NAICS industries based on the share of employees within those industries who work in time-intensive occupations according to the occupational scores described in Table 6, part A. We then divide the three-digit NAICS industries into three terciles based on the time-intensity of the weighted occupations for each. These estimates are used to gauge the industrial time-intensity of employed individuals in the LEHD-Census sample. We then re-estimate the LEHD-Census models separately for the industry-level terciles, using the initial industry in which each employee was observed in the LEHD, which for most people was 1991.

Appendix Table 3

Appendix Table 3

Decomposition of parental gender gap in earnings by the time intensity of the early industry, LEHD-Census.

Part A: Most Time-Intensive Industries				
Age Group	(1) Parental Gender Gap in Earnings	(2) Motherhood Penalty	(3) Fatherhood Premium	(4) Price of Being Female
30–34	−0.552*	−0.175*	0.151*	−0.227*
35–39	−0.634*	−0.178*	0.194*	−0.261*
40–44	−0.637*	−0.147*	0.227*	−0.264*
45–49	−0.582*	−0.095*	0.242*	−0.245*
50–54	−0.534*	−0.048*	0.252*	−0.234*
Part B: Least Time-Intensive Industries				
Age Group	(1) Parental Gender Gap in Earnings	(2) Motherhood Penalty	(3) Fatherhood Premium	(4) Price of Being Female
30–34	−0.436*	−0.192*	0.083*	−0.161*
35–39	−0.471*	−0.188*	0.104*	−0.179*
40–44	−0.446*	−0.149*	0.121*	−0.176*
45–49	−0.399*	−0.092*	0.127*	−0.181*
50–54	−0.365*	−0.043*	0.128*	−0.194*

* $p < 0.01$,
# $p < 0.05$,
§ $p < 0.1$

Sources: LEHD 1991–2014 and 2000 Decennial Census.

Notes: These estimates use the results from the estimation with individual fixed effects, establishment fixed effects, and establishment size.

References

Adda, Jérôme, Dustmann, Christian, Stevens, Katrien, 2015. The career costs of children. J. Polit. Econ. 125 (2), 293–337.

Almond, Douglas, Cheng, Yi, Machado, Celia, 2023. Large motherhood penalties in US administrative microdata. Proc. Natl. Acad. Sci. 120 (29).

Angelov, Nikolay, Johansson, Per, Lindahl, Erica, 2016. Parenthood and the gender gap in pay. J. Labor. Econ. 34 (3), 545–579.

Barth, Erling, Kerr, Sari Pekkala, Olivetti, Claudia, 2021. The dynamics of gender earnings differentials: evidence from establishment data. Eur. Econ. Rev. 134 (C).

Bronson, Mary Ann, Thoursie, Peter Skogman, 2020. The wage growth and within-firm mobility of men and women: New evidence and theory. Working Paper.

Budig, Michelle J., 2014. Third Way, Next Report. https://www.thirdway.org/report/the-fatherhood-bonus-and-the-motherhood-penalty-parenthood-and-the-gender-gap-in-pay.

Card, David, Cardoso, Ana Rute, Kline, Patrick, 2016. Bargaining, sorting, and the gender wage gap: quantifying the impact of firms on the relative pay of women. Q. J. Econ. 131 (2), 633–686.

Fahle, Sean, McGarry, Kathleen, 2018. Women working longer: labor market implications of providing family care. In: Goldin, C., Katz, L. (Eds.), Women Working Longer: Increased Employment at Older Ages. University of Chicago Press, Chicago, IL, pp. 157–184.

Goldin, Claudia., 2021. Career & Family: Women's Century-Long Journey toward Equity. Princeton University Press.

Goldin, Claudia., 2014. A grand gender convergence: its last chapter. Am. Econ. Rev. 104 (4), 1091–1119.

Goldin, Claudia, Mitchell, Joshua, 2017. The new lifecycle of women's employment: disappearing humps, sagging middles, expanding tops. J. Econ. Perspect. 31 (1), 161–182.

Goldin, Claudia, Kerr, Sari Pekkala, Olivetti, Claudia, Barth, Erling, 2017. The expanding gender earnings gap: evidence from LEHD-2000 Census. Am. Econ. Rev. 107 (5), 110–114.

Cortés, Patricia, and Pan, Jessica. 2020. "Children and the remaining gender gaps in the labor market." NBER Working Paper No. 27980. October.

Goldin, Claudia, Sari Pekkala Kerr, and Claudia Olivetti. 2022. "When the kids grow up: women's employment and earnings across the family cycle." National Bureau of Economic Research (NBER) Working Paper, No. 30323, August.

Goldin, Claudia, Sari Pekkala Kerr, and Claudia Olivetti. 2021. "Appendix: data extract NLSY79: the other side of the mountain: women's employment and earnings over the family cycle," https://scholar.harvard.edu/files/goldin/files/appendix_dataextract_nlsy79.pdf.

Hara, Hiromi., 2018. The gender wage gap across the wage distribution in Japan: within- and between-establishment effects. Labour. Econ. 53 (August), 213–229.

Herr, Jane Leber, 2015. The labor supply effects of delayed first birth. In: American Economic Review, Papers and Proceedings, 105, pp. 630–637.

Hotz, Joseph, Per Johansson, and Arizo Karimi. 2017. "Parenthood, family friendly workplaces, and the gender gaps in early work careers." NBER Working Paper No. 24173, December.

Jones, Melanie, Kaya, Ezgi, 2023. The UK gender pay gap: does firm size matter? Economica 90 (359), 937–952.

Juhn, Chinhui, McCue, Kristen, 2017. Specialization then and now: marriage, children, and the gender earnings gap across cohorts. J. Econ. Perspect. 31 (1), 183–204.

Killewald, Alexandra., 2013. A reconsideration of the fatherhood premium: marriage, co-residence, biology, and fathers' wages. Am. Sociol. Rev. 78, 96–116.

Killewald, Alexandra, Gough, Margaret, 2013. Does specialization explain marriage penalties and premiums? Am. Sociol. Rev. 78 (3), 477–502.

Killewald, Alexandra, Lundberg, Ian, 2017. New evidence against a causal marriage wage premium. Demography 54 (3), 1007–1028.

Kleven, Henrik, Landais, Camille, 2017. Gender inequality and economic development: fertility, education, and norms. Economica 84 (334), 180–209.

Kleven, Henrik, Landais, Camille, Posch, Johanna, Steinhauer, Andreas, Zweimüller, Josef, 2019a. Child penalties across countries: evidence and explanation. AEA Pap. Proc. 109 (May), 122–126.

Kleven, Henrik, Landais, Camille, Søgaard, Jakob Egholt, 2019b. Children and gender inequality: evidence from Denmark. AEJ: Appl. Econ. 11 (4), 181–209.

Korenman, Sanders, Neumark, David, 1991. Does marriage really make men more productive? J. Hum. Resour. 26 (2), 282–307.

Kunze, Astrid., 2020. The effect of children on male earnings and inequality. Rev. Econ. Househ. 18, 683–710.

Kuziemko, Ilyana, Jessica Pan, Jenny Shen, and Ebonya Washington. 2020. "The mommy effect: do women anticipate the employment effects of motherhood." NBER Working Paper No. 24740. December.

Lundberg, and Rose, 2000. Parenthood and the earnings of married men and women. Labour. Econ. 7, 689–710.

Shen, Karen (2021) "Who benefits from public financing of home care for low-income seniors?" Working Paper. Harvard University, May.

Song, Jae., Price, David, Guvenen, Fatih, Bloom, Nicholas, von Wachter, Till, 2019. Firming up inequality. Q. J. Econ. 134 (1), 1–50.

U.S. Department of Labor, Bureau of Labor Statistics, 2019. National longitudinal survey of youth 1979 (NLSY79) cohort, 1979-2016 (rounds 1-27). Produced and Distributed by the Center for Human Resource Research (CHRR). The Ohio State University, Columbus, OH.

Vilhuber, Lars., 2018. LEHD infrastructure S2014 files in the FSRDC. Center Econ. Stud. CES 18–27.

Wilde, Elizabeth Ty, Lily Batchelder, and David T. Ellwood. 2010. "The mommy track divides: the impact of childbearing on wages of women of differing skill levels." NBER Working Paper No. 16582. December.

Yu, Wei-Hsin, Hara, Yuko, 2021. Motherhood penalties and fatherhood premiums: effects of parenthood on earnings growth within and across firms. Demography 58 (1), 247–272.