# Occupational Sorting, Multidimensional Skill Mismatch, and the Child Penalty among Working Mothers

Gabrielle Pepin and Bryce VanderBerg*

October 22, 2021

## Abstract

We study the extent to which occupational sorting explains child penalties—gender gaps in labor market outcomes due to children—among working parents. Using an event-study approach and data from the National Longitudinal Surveys of Youth (NLSY) 1979 and 1997, we estimate that children generate long-run earnings gaps of over $200 per week among working parents. In the NLSY79, we find that children lead mothers to sort into lower-paying occupations in which employees tend to work fewer hours. We estimate that children increase multidimensional occupation-skill mismatch among working mothers by 0.3 standard deviations, relative both to their own levels of mismatch from before birth and to those of fathers. In the NLSY97, results suggest that improvements in labor market outcomes among fathers in response to children, rather than a worsening of labor market outcomes among mothers, seem to drive child penalties.

Keywords: Child penalty, occupational sorting, multidimensional occupation-skill mismatch, gender gap, event study

JEL Codes: J16, J13, J24

---

*Pepin: W.E. Upjohn Institute for Employment Research (Pepin@upjohn.org). VanderBerg: Michigan State University (vande853@msu.edu).

## I.  Introduction

Despite considerable gender convergence in education and labor market experience during the late 1900s, substantial gender gaps in labor market outcomes remain in the United States and other industrialized nations (Blau and Kahn 2017; Goldin 2014; Goldin, Katz, and Kuziemko 2006). Parenthood plays a large role in explaining gender gaps, as researchers document large and persistent child penalties—gender gaps in labor market outcomes due to children—across countries and family structures.[1] We are the first to study the extent to which occupational sorting explains child penalties and gender gaps among working parents. As an increasing number of active labor market policies target working families with children, understanding determinants of gender inequality among working parents is particularly important.

Children likely affect labor market outcomes via occupational sorting, as they limit the amount of time and energy parents may devote to their jobs. As women tend to bear the majority of child-rearing responsibilities, even when they work full-time, children may disproportionately lead mothers to sort into less-demanding, lower-paying occupations. Furthermore, sorting into such "family-friendly" occupations may worsen the match between women's skills and occupational requirements, which has been shown to affect wages.[2]

In this paper, we first use the event-study approach proposed by Kleven, Landais, and Søgaard (2019) and data from the National Longitudinal Surveys of Youth (NLSY) to document child penalties among individuals born in the United States during the 1950s and 1960s and during the 1980s who later become working parents. By leveraging variation in the timing of first births, we

---

[1]See Angelov, Johansson, and Lindahl (2016), Chung et al. (2017), Cortes and Pan (2020), Kleven et al. (2019), Kleven, Landais, and Søgaard (2019, 2021), and Sieppi and Pehkonen (2019).

[2]See Bacolod and Blum (2010), Deming (2017), Guvenen et al. (2020), Heckman, Stixrud, and Urzua (2006), and Lise and Postel-Vinay (2020).

allow fertility decisions to be endogenous but assume that any unobservable determinants of labor market outcomes evolve smoothly around childbirth. Comparing effects across working mothers and fathers, we show that children cause long-run earnings gaps of over $200 per week in both the NLSY79 and NLSY97. While decreases in wages and hours worked among mothers generate child penalties in the NLSY79, increases in fathers' wages and hours generate penalties in the NLSY97.

We then use the event-study framework to estimate effects of children on occupational sorting. We show that in the NLSY79, children lead mothers to sort into lower-paying occupations in which employees tend to work fewer hours. While we do not observe this phenomenon among mothers in the NLSY97, occupational sorting still appears to play a large role in explaining the child penalty, as fathers sort into higher-paying occupations with higher average hours worked. Next, as occupational sorting likely affects the degree of complementarity between parents' skills and occupational requirements, we estimate effects of children on multidimensional occupation-skill mismatch. In other words, we estimate effects on the extent to which parents' math, verbal, science and mechanical, and social, or noncognitive, skills differ from those required by their occupations.

To examine mismatch, we proxy for skills using NLSY respondents' test scores on the Armed Services Vocational Aptitude Battery (ASVAB), along with information on sociability and extracurricular participation during childhood. We link these skill measures to O*NET, which documents occupational task content, and measure the relatedness of occupation tasks to skill categories using the United States Department of Defense's Defense Manpower Data Center crosswalk. Finally, we measure the distance between individuals' multidimensional skills and the importance of those skills in their occupations. We find that, among mothers in the NLSY79, children in-

crease multidimensional occupation-skill mismatch by about 0.3 standard deviations, relative both to their own levels of mismatch from before birth and to those of fathers. In contrast, mothers in the NLSY97 exhibit decreases in skill mismatch post-childbirth. Nonetheless, large child penalties remain.

Our work falls at the intersection of existing literatures on the child penalty and occupational sorting. In terms of child penalties, there is extensive work on penalties in Europe, where earnings penalties range from about 20 percent in Denmark to about 60 percent in Germany.[3] Three papers study child penalties in the United States. Cortes and Pan (2020) and Kleven et al. (2019) use data from the Panel Survey of Income Dynamics on parents who had first children during the late 1900s and early 2000s to document long-run penalties in earnings between 30 and 40 percent. Chung et al. (2017) find similar results using data from the Survey of Income and Program Participation linked to earnings records of parents whose first children were born between 1978 and 2011. They also document decreases in the size of the child penalty over time.

Although the existence of the child penalty is well-established, evidence on its determinants is more limited. Gender norms seem to matter, as Kleven, Landais, and Søgaard (2019) find that girls who grew up in families with traditional divisions of labor incur larger penalties when they become mothers. Additionally, researchers find that child penalties are larger among couples in which the father has more education, which is line with the theory of comparative advantage (Angelov, Johansson, and Lindahl 2016; Chung et al. 2017). We contribute to the literature on understanding the sources of the child penalty by investigating the role of occupational sorting. Changes in workplace attributes due to children suggest that occupational sorting likely plays an

---

[3]See Angelov, Johansson, and Lindahl (2016), Chung et al. (2017), Kleven, Landais, and Søgaard (2019, 2021), and Sieppi and Pehkonen (2019).

important role in explaining child penalties. In particular, Kleven, Landais, and Søgaard (2019) find that, in Denmark, children lead women to move into workplaces that employ larger shares of women with children. They also find that children cause women to move from private into public sector employment, where work hours tend to be more flexible.

While we are the first to estimate effects of occupational sorting *due to children* on the gender gap, research on the extent to which occupational sorting explains gender gaps more generally dates back to the decompositional analyses of Blinder (1973) and Oaxaca (1973). More recently, Blau and Kahn (2017) estimate that occupational sorting explained about 10 percent of the gender wage gap in the United States in 1980 and about 30 percent of it in 2010. We improve upon traditional gender gap decompositions because we do not control for labor market choices likely affected by children, such as industry, in our analyses. Additionally, whereas children largely do not factor into traditional gender gap decompositions by construction, given similar numbers of children across men and women, we explicitly focus on the role of children in explaining gender inequality.

Finally, in studying effects of children on the degree of complementarity between parents' skills and occupations, we contribute to a growing literature on multidimensional occupation-skill mismatch (Addison, Chen, and Ozturk 2020; Guvenen et al. 2020; Lise and Postel-Vinay 2020; Speer 2017). While much of the existing literature on multidimensional skills focuses exclusively on men, Addison, Chen, and Ozturk (2020), who study gender differences in skill mismatch over the lifecycle, is a notable exception. Using data from the NLSYs, the authors show that mothers exhibit greater mismatch in their occupations than men and childless women, though differences in mismatch have decreased over time. They also find that mismatch is relatively large among college-educated women and parents in occupations with more job flexibility. We build on Addison, Chen,

and Ozturk (2020) by parsing the causal effects of children on mismatch from differences across mothers, fathers, and childless workers that may be unobservable to researchers.[4]

In the following section, we provide institutional details about access to parental leave and child care in the U.S. In Section III, we describe the data. In Section IV, we estimate effects of children on labor supply and occupational sorting. In Section V, we conclude.

## II. Institutional Setting

Family policies in the U.S. are notoriously ungenerous. Before the enactment of the Family and Medical Leave Act of 1993 (FMLA), workers were not guaranteed any parental leave.[5] As of 1991, some 37 percent of female full-time workers in private-sector firms with at least 100 employees had access to unpaid leave, and only 2 percent had access to paid leave. Full-time male workers in similar firms were less likely to have access to leave: some 26 percent had access to unpaid leave, and 1 percent had access to paid leave (U.S. Bureau of Labor Statistics 2020).

Since 1993, firms with at least 50 employees must offer eligible employees 12 weeks of job-protected unpaid leave for childbirth or adoption.[6] To be eligible for leave, employees must have worked at the firm for at least 12 months and have accumulated at least 1,250 work hours. FMLA increased leave coverage substantially. In 1994, some 84 percent of full-time workers in private-

---

[4]Though the authors describe their work as taking a "descriptive approach rather than establishing definitive causal associations," in some specifications, Addison, Chen, and Ozturk (2020) use the age of the NLSY respondent's sibling at the time of the sibling's first birth as an instrument for the timing of the respondent's first birth to estimate effects of children on mismatch. We do not believe that sibling age at first birth passes the monotonicity assumption required of instruments and view Addison, Chen, and Ozturk (2020) as a valuable descriptive contribution to the literature.

[5]Some states mandated some type of parental leave before 1993, including five states that offer temporary disability insurance. The Pregnancy Discrimination Act of 1978 mandated that employers in states with temporary disability insurance treat pregnancy as a short-term disability, which allows mothers to receive partial earnings without job protection for around six weeks.

[6]Employers with at least 50 employees within 75 miles of the worksite for at least 20 weeks of the last year must offer 12 weeks of unpaid leave, though some states have lower firm size thresholds or require longer leave lengths. Leave also may be taken if the employee is in poor health or cares for a close relative who is in poor health. Employers may refuse job protection for their highest-paid 10 percent of employees if leave would generate economic harm.

sector firms with at least 100 employees had access to unpaid leave. Still, only 2 percent of full-time workers in similar firms had access to paid leave, and many workers were not covered by the law. Less than 50 percent of both part-time workers in firms with at least 100 employees and full-time workers in smaller firms had access to any leave benefits (U.S. Bureau of Labor Statistics 2020).

While the U.S. still does not mandate paid parental leave, between 2004 and 2017, California, New Jersey, and Rhode Island implemented their own paid leave mandates. California and New Jersey offer six and Rhode Island offers four weeks of paid parental leave.[7] In each of these states, workers who meet a given work history requirement receive partial wage replacement up to a maximum weekly benefit, though only Rhode Island offers job protection beyond what is covered under FMLA. California and Rhode Island extend benefits to workers at small firms; New Jersey does not. As states enacted paid leave mandates, access to benefits increased: between 2005 and 2017, the proportion of private-sector workers with access to paid leave increased from 0.07 to 0.13. At the same time, the proportion with access to unpaid leave increased from 0.81 to 0.87 (U.S. Bureau of Labor Statistics 2020). Hence, while access to paid leave has increased in recent years, most workers still do not receive such benefits. Even mothers who have access to paid leave must return to work relatively shortly after childbirth to be guaranteed job protection.

In addition to limited parental leave benefits, the U.S. does not offer universal child care or pre-kindergarten, though universal schooling is available for children beginning at age five. Several states, however, operate their own universal pre-kindergarten programs for four- and, in some cases, three-year-old children. In particular, between 1995 and 2008, Florida, Georgia, Iowa, Oklahoma, Vermont, West Virginia, and Wisconsin implemented universal pre-kindergarten programs.

---

[7]California increased the maximum leave length to eight weeks in 2020.

Additionally, children aged three and four in families with incomes at or below the federal poverty level can participate in Head Start, a means-tested federal preschool program that was rolled out across the U.S. between 1965 and 1980. Head Start's objective is to promote children's cognitive and interpersonal development and school readiness through education, health and nutrition interventions, and family partnerships. An additional branch of Head Start, Early Head Start, was established in 1994 to serve pregnant women and children younger than age three who meet Head Start's income-eligibility criteria. Early Head Start consists of center-based care or home visits and is especially focused on nurturing healthy relationships between children and their caregivers.

Among households without access to free services, income support for early care and education is limited, though state and federal governments administer some cash benefits to families through the tax code. For instance, the Child and Dependent Care Credit (CDCC), which was introduced in 1976, subsidizes child care costs for working families. Between 2003 and 2020, households could claim up to $3,000 in child care expenses per child for up to two children and receive CDCC benefits worth up to 35 percent of those expenses, or $1,050. In addition, about 40 percent of workers can access dependent care flexible spending accounts (FSA) that their employers offer (U.S. Bureau of Labor Statistics 2020). Since 1986, employees who receive FSAs from their employers have been able to set aside up to $5,000 of earnings before taxes for dependent care expenses. The employer deducts this income from employees' paychecks, but employees are reimbursed for child care expenditures. Additional tax benefits for families with children, but not explicitly for child care, include the Child Tax Credit (CTC) and Earned Income Tax Credit (EITC). The CTC was introduced as a child benefit in 1997, and between 2003 and 2020, families could receive benefits of up to $1,000 per child. The EITC is an earnings subsidy targeted at low- and moderate-income families with children. As of 2020, maximum benefits for one-, two-, and three-child households

were $3,584, $5,920, and $6,660, respectively.

Taken together, lack of access to family leave and universal early care and education services generates very high costs of child-rearing in the U.S. Under such constraints, the arrival of children may lead parents to move into less-demanding occupations that could negatively impact their long-run labor market outcomes. In light of this, we study effects of children on parents' labor supply and occupation choices in the following sections.

## III.  Data

To examine child penalties and occupational sorting, we link individual-level data on worker skills, occupations, and labor market outcomes from the NLSYs to occupational task content from O*NET.

### A.  *NLSY79 and NLSY97*

The NLSY79 and NLSY97 are nationally-representative panel surveys of individuals living in the United States aged 14 to 22 in 1979 and 12 to 16 in 1997, respectively. Biennial interviews of each cohort continue through the present, though interviews were conducted annually from 1979 to 1994 for the 1979 cohort and from 1997 to 2011 for the 1997 cohort. The NLSYs contain extensive information on individuals' demographics, family backgrounds, educational experiences, and labor market outcomes. Importantly for our study, the data document respondents' census occupation codes and months of birth for their children. Detailed information on individual characteristics and the long-panel nature of the surveys make the NLSYs well-suited to estimate long-run effects of children on labor market outcomes.

Another key advantage of the NLSYs is their inclusion of ASVAB test scores, which were

administered to NLSY79 respondents in 1981 and to NLSY97 respondents in 1999. The ASVAB, which measures cognitive skills in ten subjects, was developed by the United States military in 1968, and since 1976, all military branches have used it to determine eligibility for military occupations. For example, to be eligible for a position as an electronics technician in the United States Navy, a recruit's composite score from the arithmetic reasoning, mathematics knowledge, electronics information, and general science sections of the ASVAB must exceed a certain threshold.[8] ASVAB scores likely are good proxies for individuals' cognitive skills, as military researchers show that scores on sections required for occupations predict job performance (Sims and Hiatt 2001; Welsh, Kucinkas, and Curran 1990).

To study child penalties and occupational sorting, we restrict the NLSY samples to working parents who fully transitioned into the labor market before having children.[9] Similarly to Farber and Gibbons (1996), Schönberg (2007), and Speer (2017), we designate an individual as having made a full transition into the labor market if they have not been enrolled in school for two consecutive years and have worked for at least 30 hours per week in at least half of the weeks during those two years. We exclude from the sample individuals with fewer than two years of labor market experience before childbirth, individuals with missing ASVAB scores, active-duty military and veterans, and the self-employed. We classify an individual as a working parent if they worked for at least ten hours per week at a main employer for some period during the year of their first childbirth, the year immediately preceding childbirth, the year immediately following childbirth, and at

---

[8]See https://www.military.com/join-armed-forces/asvab for information on ASVAB score requirements for military occupations.

[9]We include evidence on child penalties among all parents in the online appendix. In constructing the sample, we include the NLSY79 and NLSY97 Black and Hispanic oversamples but exclude the NLSY79 economically disadvantaged oversample, which was discontinued in 1991. Results from analyses in which we exclude the Black and Hispanic oversamples are similar and available upon request.

least half of the following four years.[10] As in Cortes and Pan (2020), we require that individuals complete at least one interview before childbirth, at least one interview after childbirth, and at least four interviews during the sample period.

## B. O*NET and the Defense Manpower Data Center Crosswalk

O*NET, a database maintained by the United States Department of Labor, documents knowledge, skills, and abilities—henceforth, "tasks"—required of occupations. Specifically, for each occupation in the Standard Occupational Classification (SOC) system, expert job analysts, job supervisors, or job incumbents assign scores for the importance of 277 tasks.[11] As in Addison, Chen, and Ozturk (2020) and Guvenen et al. (2020), we use the Defense Manpower Data Center (DMDC) crosswalk, which was created by the United States Department of Defense, to measure the relatedness of O*NET tasks to ASVAB section scores. The crosswalk includes information on twenty-six O*NET tasks to which personnel research psychologists assign "at least a moderately strong probability" of being related to at least one of the following ASVAB section tests: word knowledge, paragraph comprehension, arithmetic reasoning, mathematics knowledge, general science, mechanical comprehension, and electronics information (ASVAB Career Exploration Program 2011). For each of the tasks, psychologists and psychometricians assign relatedness scores to each of the aforementioned ASVAB section tests.

## C. Creating Multidimensional Occupation-Skill Mismatch Measures

Figure 1 illustrates the procedure for creating the multidimensional occupation-skill mismatch measure for a given individual and occupation. First, we create measures of the relevance of

---

[10]Results are robust to alternative definitions of working parents and are available upon request.
[11]We map the SOC occupations into occupation categories included in the NLSYs using census occupation codes.

ASVAB section scores to occupational requirements. This allows us to link occupational requirements to individuals' bundles of skills, as demonstrated on the ASVAB test. To create measures of ASVAB score relevance to a given occupation, we use the measures of task importance within the occupation from O*NET and the measures of task relatedness to ASVAB section test scores from the DMDC crosswalk.[12] Focusing on the twenty-six tasks included in the crosswalk, we multiply the vector of occupational task importance measures by the matrix of ASVAB task-relatedness measures. This yields measures of ASVAB section score relevance to the given occupation, which we normalize to have standard deviations of 1.

Next, as each military branch combines ASVAB section scores in different ways to determine recruits' suitability for occupations, we follow Addison, Chen, and Ozturk (2020) in creating four skill categories—math, verbal, science/mechanical, and social.[13] The first three categories correspond to sections of the ASVAB test. Specifically, the mathematics knowledge and arithmetic reasoning sections of the ASVAB correspond to the math category; the word knowledge and paragraph comprehension sections correspond to verbal category; and the general science, mechanical comprehension, and electronics information sections correspond to science/mechanical category. Because multiple ASVAB sections comprise each of the skill categories, as shown in Figure 1, we follow Guvenen et al. (2020) and apply principal component analysis (PCA) to both the vector of occupational relevance measures and to the vector of the individual's ASVAB section scores.[14]

---

[12]Because the scale of the DMDC relatedness score is somewhat arbitrary, we rescale each ASVAB section's twenty-six task-relatedness scores to sum to 1.

[13]Results are robust to using alternative skill categories proposed by Addison, Chen, and Ozturk (2020), Guvenen et al. (2020), Lise and Postel-Vinay (2020), and Speer (2017) and are available upon request.

[14]We follow Altonji, Bharadwaj, and Lange (2012) in standardizing ASVAB section scores to account for differences in age at the time the test was administered and test format, as the NLSY79 cohort took a pencil and paper version of the ASVAB, and the NLSY97 cohort took a computer-assisted version of the test. To adjust for test format differences, we use a crosswalk based on scores of individuals randomly assigned to one of the two test formats (Segall 1997). We then perform an equipercentile mapping to age 16 separately for each NLSY cohort. In other words, we assign test scores of those who took the test at age $a$ and scored in the $q$th percentile among age $a$ test takers the corresponding $q$th-percentile score of those who took the test at age 16. In doing so, we assume that the relative ranking

In doing so, for each skill category, we create a measure equal to the first principle component of the pertinent relevance measures or ASVAB section scores. For example, the verbal relevance measure is the first principle component of the relevance of the word knowledge and paragraph comprehension ASVAB sections to the given occupation. Analogously, the verbal skill measure is the first principle component of the individual's scores on the word knowledge and paragraph comprehension sections of the ASVAB. We then scale the occupational skill relevance measures and the individual's skill measures into percentile ranks among occupations and NLSY respondents in their cohort, respectively.[15]

In addition, we create a social skill category. To construct a measure of the relevance of social skills to a given occupation, we follow Addison, Chen, and Ozturk (2020), Guvenen et al. (2020), and Deming (2017) and rely on the following occupational task-relatedness measures from O*NET: social perceptiveness, coordination, persuasion, negotiation, instructing, and service orientation. To construct a social skill measure for each individual, we again follow Deming (2017) and use information on self-reported sociability during childhood and adolescence and club and sport participation during high school available in the NLSYs. For both the task-relatedness and skill measures, we scale the standard deviation of each of the components to equal one and apply PCA. As with the cognitive skill measures, we then convert the social occupational task-relatedness and skill measures into percentile ranks.

Finally, we compare individuals' bundles of skills to those required by their occupations. We

---

of an individual's score in their cohort's score distribution does not depend on when the cohort took the test. We also assume that the level of skill associated with a score in the $q$th percentile of the age $a$ score distribution is the same as that associated with a $q$th-percentile score in the age 16 score distribution. We do not restrict scores across NLSY cohorts and normalize them to have a mean of 0 and standard deviation of 1 in 1979.

[15]When we scale occupational skill relevance measures, we weight each occupation by the number of individuals engaged in it in the individual's NLSY cohort.

follow Guvenen et al. (2020) and define multidimensional occupation-skill mismatch as follows:

$$(1) \qquad m_{ic} = \sum_{l=1}^{4} w_l * |q(a_{il}) - q(r_{cl})|,$$

where $a_{il}$ is the skill measure for individual $i$ in skill category $l$. $r_{cl}$ represents the relevance of skill $l$ within occupation $c$. $q(a_{il})$ and $q(r_{cl})$ denote the corresponding percentile ranks of individual skill and occupational relevance. $w_l$ is the first principle component of $\sum_{l=1}^{5} |q(a_{il}) - q(r_{cl})|$. For ease of interpretation, we standardize $m_{ic}$ to have a standard deviation of 1.

Furthermore, we use data on individuals' skills and those required by their occupations to study skill overmatch and undermatch, or the extent to which individuals are overqualified or underqualified for their occupations, respectively. We define skill overmatch as follows:

$$(2) \qquad om_{ic} = \sum_{l=1}^{4} \mathbb{1}[(q(a_{il}) - q(r_{cl})) > 0](w_l * (q(a_{il}) - q(r_{cl}))),$$

where the variables are the same as those listed in Equation (1). Equation (2) implies that over-match increases with the difference between an individual's skills and those required by their occupation. If none of an individual's skill measures exceed those required by their occupation, then the individual is not overqualified for the occupation, and $om_{ic}$ equals 0. Analogously, we define skill undermatch:

$$(3) \qquad um_{ic} = \sum_{l=1}^{4} \mathbb{1}[(q(a_{il}) - q(r_{cl})) < 0](w_l * (q(a_{il}) - q(r_{cl}))).$$

We standardize both $om_{ic}$ and $um_{ic}$ to have standard deviations of 1.

*D.    Current Population Survey Outgoing Rotation Groups*

Finally, we use data from the Current Population Survey (CPS) to estimate average weekly earnings and hours worked by occupation and decade. The CPS, a nationally-representative monthly survey of over 65,000 households, is designed to measure employment. Households in the survey are interviewed for four months, ignored for eight months, then interviewed for four more months. Since 1982, respondents have documented labor force status, occupation, and usual hours worked per week during each month in the survey. We use data from the fourth and eighth outgoing interviews, during which respondents also report usual weekly earnings. Specifically, to create measures of average earnings and hours worked within occupations held by NLSY respondents, we estimate regressions of the given labor market outcome on census occupation dummies using sample weights. To allow for changes in average earnings by occupation over time, we estimate separate regressions for the 1980s, 1990s, 2000s, and 2010s.

*E.    Summary Statistics*

Table 1 displays summary statistics for working mothers and fathers separately by NLSY cohort as of the year before their first childbirth. At that time, working parents in the NLSY79 tend to be about 25 years old; working parents in the NLSY97 are closer to 24 years old on average. About 50 percent of parents in the NLSY79 are married during the year preceding childbirth, whereas only about 40 percent of parents in the NLSY97 are married. Across both cohorts, pre-childbirth average weekly earnings (2010 dollars) tend to be higher for fathers than for mothers, but the gender gap shrinks across cohorts, as gender differences in average weekly earnings decrease from $139 in the NLSY79 to $60 in the NLSY97. Both increases in mothers' average earnings and

decreases in fathers' average earnings drive the shrinking earnings gap. Similarly, fathers' average hourly wages decrease from \$15.72 to \$15.51 across cohorts while mothers' average hourly wages increase from \$13.84 to \$14.47. Fathers tend to work more hours per week than mothers, and average weekly hours worked decrease across cohorts for both genders: fathers' average hours decrease from 43 to 40, and mothers' from 39 to 38.

Turning to occupation choice, Table 1 shows that fathers tend to sort into occupations in which, on average, employees work more hours and garner more earnings, compared to mothers. In the NLSY79, average earnings within occupation are about \$700 and \$600 for working fathers and mothers, respectively. Average earnings within occupation are about \$600 for fathers and \$500 for mothers in the NLSY97. Despite sorting into lower-paying occupations, mothers tend to exhibit better occupation-skill matches. Average mismatch measures, denoted by $m_{ic}$ in Equation (1), are 1.75 and 2.00 for mothers in the NLSY79 and NLSY97, respectively. This compares to 2.03 for fathers in the NLSY79 and 2.06 for fathers in the NLSY97. Mothers also exhibit less skill overmatch and more skill undermatch compared to fathers. Measures of overmatch and undermatch indicate that, based on their skills and those required by their occupations, both mothers and fathers exhibit lower degrees of overqualification and higher degrees of underqualification for their occupations over time.

## IV.  Evidence on Occupational Sorting and the Child Penalty

### A.  *Empirical Strategy*

We use the event-study method proposed by Kleven, Landais, and Søgaard (2019) to estimate effects of children on parents' labor supply and occupation choices. In doing so, we allow for

endogenous fertility but assume that unobservable determinants of labor market outcomes evolve smoothly around childbirth. Under this assumption, we attribute any discontinuity in outcomes around childbirth to effects of children. The smoothness assumption would be violated if, for instance, parents time childbirth to coincide with a job promotion. We estimate the following event-study model separately by gender and NLSY cohort, where event time $t = 0$ during the year the individual has their first child:

$$(4) \qquad Y_{ist}^g = \sum_{j \neq -1} \alpha_f^g * \mathbb{1}[j = t] + \sum_k \beta_k^g * \mathbb{1}[k = age_{is}] + \sum_y \gamma_y^g * \mathbb{1}[y = s] + \delta^g X_i^g + v_{ist}^g.$$

$Y_{ist}^g$ is the outcome of interest for individual $i$ of gender $g$ in year $s$ relative to event time $t$. We omit the indicator for $t - 1$ so that the $\hat{\alpha}_j^g$ coefficients measure effects of children relative to the year before birth. We include age and year dummies to control non-parametrically for life-cycle and time trends, such as inflation and business cycles. $X_i^g$ includes controls for education at labor market entry and race and an indicator for entering the sample after $t - 5$.[16] We cluster standard errors at the individual level.

Equation (4) leverages differences in birth timing, conditional on age, year, and individual characteristics, to estimate post-childbirth effects of children on labor market outcomes. Short-run measures of the child penalty capture effects of a first child, whereas long-run measures may capture effects of total fertility. Despite differences in interpretation of effects, for both short- and long-run penalties, $\hat{\alpha}_j^g$ only captures effects of children that are realized after childbirth. Thus, to the extent that children reduce pre-childbirth labor market investments of women relative to men,

---

[16]We include controls for individual characteristics in our models, whereas Kleven, Landais, and Søgaard (2019) generally do not, because our data exhibit incomplete overlap in event time, age, year, and education and in event time, age, year, and race. Thus, if we did not control for individual characteristics, the age and year dummies would capture effects of some but not all education levels and races.

we underestimate effects of children on the gender gap.

*B.   Results*

Figure 2 displays results among working mothers (solid blue lines) and fathers (red dashed lines) in the NLSY79. The upper-left panel of Figure 2 documents effects on weekly earnings. The panel shows that, conditional on life-cycle and time trends and time-invariant individual characteristics, pre-birth earnings trajectories are fairly similar for mothers and fathers.  Then, when children arrive, mothers' and fathers' earnings paths diverge. Mothers experience immediate decreases in earnings that continue to grow for at least ten years after childbirth.  Fathers' earnings increase initially but settle around prebirth levels by ten years post-childbirth, though 95 percent confidence bands indicate that long-run estimates are a bit noisy. Results imply that children cause long-run earnings gaps of over $200 per week between working mothers and fathers.

Similarly, the upper-middle and right panels of Figure 2 show effects on hourly wages and hours worked per week, respectively. As with earnings, working parents' wages and hours worked trend fairly similarly before birth but diverge immediately thereafter. Mothers' hours worked continue to decrease until they plateau at around 4 fewer hours per week five years post-birth. Mothers' wages continue to fall through at least ten years post-birth, when wages are nearly $5 less per hour. This constitutes a 29 percent decrease from the pre-birth mean. Meanwhile, fathers' hours worked remain relatively constant, and their wages increase slightly, at least through the medium-run.

The remaining panels of Figure 2 display effects on occupational sorting. Results imply that while fathers remain unaffected, the arrival of children immediately causes women to sort into occupations with lower average earnings in which employees tend to work fewer hours per week. Effect sizes, which remain relatively stable over time, imply that children lead women to enter

17

occupations that pay about $50 less per week and where employees work about 1 fewer hour per week on average. Turning to effects on skill mismatch, children, again, do not seem to affect fathers' outcomes. Mothers' mismatch gradually increases after the arrival of children, however. Ten years after birth, their mismatch has increased by about 0.3 standard deviations. Results from the lower panels of Figure 2 suggest that increases in mismatch are driven by mothers becoming more overqualified for their occupations.

Next, Figure 3 presents results among working parents in the NLSY97. As with results among parents in the NLSY79, the arrival of children generates gender gaps in earnings, wages, and hours worked. Unlike in the NLSY79, however, better labor market outcomes among fathers, rather than worse outcomes among mothers, tend to drive results. Specifically, fathers' earnings, wages, and hours worked continuously increase through four years post-childbirth. At the same time, mothers experience very little change in earnings and wages and work about two fewer hours per week. Thus, while the sizes of gender gaps in labor market outcomes remain similar across cohorts, fathers' responses to children drive child penalties in the more recent cohort. In line with this, children lead fathers to sort into higher-paying occupations that require more hours worked per week while average earnings and hours worked within mothers' occupations do not change. Nonetheless, children lead mothers, but not fathers, to sort into occupations that are better matches for their bundles of skills. As estimated effects on overmatch and undermatch are rather noisy, it is unclear whether children lead mothers to move into occupations for which they are overqualified or underqualified.

## V. Conclusion

Consistent with existing literature, we document sizable long-term child penalties among working parents in the NLSY79 and NLSY97.[17] Evidence suggests that occupational sorting plays an important role in explaining child penalties, as children lead mothers, but not fathers, in the NLSY79 to sort into occupations with lower average pay in which employees tend to work fewer hours per week. In the NLSY97, children cause fathers, but not mothers, to sort into higher-paying occupations with higher average hours worked. Additionally, children generate both absolute and relative increases in multidimensional occupation-skill mismatch among mothers in the NLSY79. Evidence suggests that mismatch effects are driven by mothers becoming more overqualified for their occupations. In contrast, mothers in the NLSY97 exhibit decreases in skill mismatch post-childbirth, but better occupation matches do not compensate for increases in wages and hours worked among fathers, and large child penalties remain.

In future versions of this paper, we plan to further investigate drivers of occupational sorting and their effects on gender inequality, as well as differences in outcomes over time. In doing so, we will examine heterogeneity in outcomes across demographic groups and individuals with different preferences over work and family during adolescence. We also plan to provide descriptive evidence on occupation flows and the extent to which parents sort into occupations that require similar skills.

## References

Addison, John T., Liwen Chen, and Orgul D. Ozturk. 2020. "Occupational Skill Mismatch: Differences by Gender and Cohort." *ILR Review* 73 (3): 730–767.

---

[17]See (Angelov, Johansson, and Lindahl 2016; Chung et al. 2017; Cortes and Pan 2020; Kleven et al. 2019; Kleven, Landais, and Søgaard 2019, 2021; Sieppi and Pehkonen 2019).

Altonji, Joseph G., Prashant Bharadwaj, and Fabian Lange. 2012. "Changes in the Characteristics of American Youth: Implications for Adult Outcomes." *Journal of Labor Economics* 30 (4): 783–828.

Angelov, Nikolay, Per Johansson, and Erica Lindahl. 2016. "Parenthood and the Gender Gap in Pay." *Journal of Labor Economics* 34 (3): 545–579.

ASVAB Career Exploration Program. 2011. *The ASVAB Career Exploration Program: Theoretical and Technical Underpinnings of the Revised Skill Composites and OCCU-Find.* Technical report.

Bacolod, Marigee P., and Bernardo S. Blum. 2010. "Two Sides of the Same Coin: Residual Inequality and the Gender Gap." *Journal of Human Resources* 45 (1): 197–242.

Blau, Francine D., and Lawrence M. Kahn. 2017. "The Gender Wage Gap: Extent, Trends, and Explanations." *Journal of Economic Literature* 55 (3): 789–865.

Blinder, Alan S. 1973. "Wage Discrimination: Reduced Form and Structural Estimates." *Journal of Human Resources* 8 (4): 436–455.

Chung, YoonKyung, Barbara Downs, Danielle H. Sandler, and Robert Sienkiewicz. 2017. "The Parental Gender Earnings Gap in the United States." *Center for Economic Studies Discussion Paper 17-68.*

Cortes, Patricia, and Jessica Pan. 2020. "Children and the Remaining Gender Gaps in the Labor Market." *NBER Working Paper 27980.*

Deming, David J. 2017. "The Growing Importance of Social Skills in the Labor Market." *Quarterly Journal of Economics* 132 (4): 1593–1640.

Farber, Henry S., and Robert Gibbons. 1996. "Learning and Wage Dynamics." *Quarterly Journal of Economics* 111 (4): 1007–1047.

Goldin, Claudia. 2014. "A Grand Gender Convergence: Its Last Chapter." *American Economic Review* 104 (4): 1091–1119.

Goldin, Claudia, and Lawrence F. Katz. 2011. "The Cost of Workplace Flexibility for High-Powered Professionals." *Annals of the American Academy of Political and Social Science* 638:45–67.

Goldin, Claudia, Lawrence F. Katz, and Ilyana Kuziemko. 2006. "The Homecoming of American College Women: The Reversal of the College Gender Gap." *Journal of Economic Perspectives* 20 (4): 133–156.

Guvenen, Faith, Burhan Kuruscu, Satoshi Tanaka, and David Wiczer. 2020. "Multidimensional Skill Mismatch." *American Economic Journal: Macroeconomics* 12 (1): 210–244.

Heckman, James J., Jora Stixrud, and Sergio Urzua. 2006. "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior." *Journal of Labor Economics* 24 (3): 411–482.

Kleven, Henrik, Camille Landais, Johanna Posch, Andreas Steinhauer, and Josef Zweimüller. 2019. "Child Penalties Across Countries: Evidence and Explanations." *AEA Papers and Proceedings* 109:122–126.

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

———. 2021. "Does Biology Drive Child Penalties? Evidence from Biological and Adoptive Families." *American Economic Review: Insights* 3 (2): 183–198.

Lise, Jeremy, and Fabien Postel-Vinay. 2020. "Multidimensional Skills, Sorting, and Human Capital Accumulation." *American Economic Review* 110 (8): 2328–2376.

Oaxaca, Ronald. 1973. "Male-Female Wage Differentials in Urban Labor Markets." *International Economic Review* 14 (3): 793–709.

Rosen, Sherwin. 1986. "The Theory of Equalizing Differences." In *Handbook of Labor Economics: Volume 1,* edited by Orley C. Ashenfelter and Richard Layard, 641–692. Amsterdam and Boston: Elsevier, North-Holland.

Schönberg, Uta. 2007. "Testing for Asymmetric Employer Learning." *Journal of Labor Economics* 25 (1): 651–691.

Segall, Daniel O. 1997. "Equating the CAT-ASVAB." In *Computerized Adaptive Testing: From Inquiry to Operation,* edited by William A. Sands, Brian K. Waters, and James R. McBride. Washington, DC: American Psychological Association.

Sieppi, Antti, and Jaakko Pehkonen. 2019. "Parenthood and Gender Inequality: Population-Based Evidence on the Child Penalty in Finland." *Economics Letters* 182:5–9.

Sims, William H., and Catherine Hiatt. 2001. *Marine Corps Selection and Classification.* Technical report. Center for Naval Analysis.

Speer, Jamin D. 2017. "Pre-Market Skills, Occupational Choice, and Career Progression." *Journal of Human Resources* 52 (1): 187–246.

U.S. Bureau of Labor Statistics. 2020. *Employee Benefits Survey.*

Welsh, John R., Susan K. Kucinkas, and Linda T. Curran. 1990. *Armed Services Vocational Battery (ASVAB): Integrative Review of Validity Studies.* Technical report 90-22. Brooks Air Force Base, TX: Air Force Systems Command.
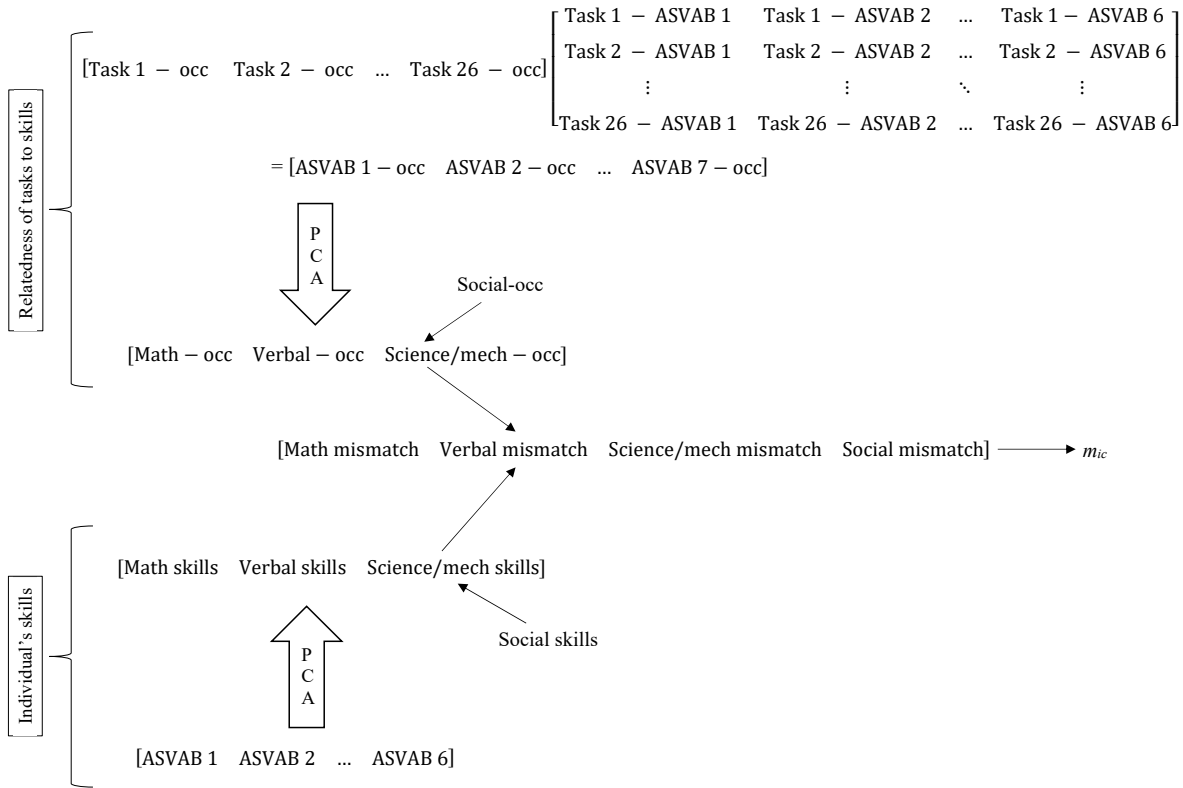
# VI. Tables and Figures

Table 1: Summary Statistics

| | NLSY79 | | NLSY97 | |
|---|---|---|---|---|
| | Fathers | Mothers | Fathers | Mothers |
| Age | 25.15 | 25.06 | 24.32 | 24.24 |
| | (3.65) | (3.79) | (3.20) | (3.38) |
| White | 0.485 | 0.477 | 0.717 | 0.721 |
| | (0.500) | (0.500) | (0.451) | (0.449) |
| Black | 0.267 | 0.275 | 0.188 | 0.152 |
| | (0.443) | (0.447) | (0.391) | (0.360) |
| Hispanic | 0.248 | 0.248 | 0.095 | 0.127 |
| | (0.432) | (0.432) | (0.294) | (0.333) |
| Married | 0.464 | 0.535 | 0.376 | 0.418 |
| | (0.499) | (0.499) | (0.485) | (0.494) |
| Weekly earnings ($) | 683 | 544 | 632 | 572 |
| | (493) | (414) | (403) | (452) |
| Hourly wage ($) | 15.72 | 13.84 | 15.51 | 14.47 |
| | (9.65) | (10.54) | (9.48) | (9.69) |
| Weekly hours | 43.15 | 38.69 | 40.22 | 37.76 |
| | (10.61) | (8.96) | (9.64) | (10.23) |
| Avg weekly earnings within occupation ($) | 721 | 599 | 594 | 492 |
| | (274) | (243) | (216) | (206) |
| Avg weekly hours within occupation | 38.55 | 35.52 | 35.84 | 32.55 |
| | (4.14) | (4.59) | (4.42) | (4.14) |
| Occupation-skill mismatch | 2.03 | 1.75 | 2.06 | 2.00 |
| | (1.076) | (0.927) | (0.986) | (0.997) |
| Overmatch | 0.719 | 0.557 | 0.413 | 0.310 |
| | (0.657) | (0.553) | (0.435) | (0.395) |
| Undermatch | –0.389 | –0.429 | –0.670 | –0.721 |
| | (0.507) | (0.497) | (0.587) | (0.584) |
| Observations | 875 | 622 | 399 | 433 |

Notes: Summary statistics among working parents in the NLSY79 and NLSY97 as of the year before their first childbirth. Standard deviations are listed in parentheses.
Source: Authors' calculations using the NLSY79, NLSY97, and Current Population Survey.

Figure 1: Procedure for Creating Multidimensional Occupation-Skill Mismatch Measure

$$[\text{Task 1} - \text{occ} \quad \text{Task 2} - \text{occ} \quad ... \quad \text{Task 26} - \text{occ}] \begin{bmatrix} \text{Task 1} - \text{ASVAB 1} & \text{Task 1} - \text{ASVAB 2} & ... & \text{Task 1} - \text{ASVAB 6} \\ \text{Task 2} - \text{ASVAB 1} & \text{Task 2} - \text{ASVAB 2} & ... & \text{Task 2} - \text{ASVAB 6} \\ \vdots & \vdots & \ddots & \vdots \\ \text{Task 26} - \text{ASVAB 1} & \text{Task 26} - \text{ASVAB 2} & ... & \text{Task 26} - \text{ASVAB 6} \end{bmatrix}$$

Relatedness of tasks to skills

$$= [\text{ASVAB 1} - \text{occ} \quad \text{ASVAB 2} - \text{occ} \quad ... \quad \text{ASVAB 7} - \text{occ}]$$

PCA

Social-occ

$$[\text{Math} - \text{occ} \quad \text{Verbal} - \text{occ} \quad \text{Science/mech} - \text{occ}]$$

$$[\text{Math mismatch} \quad \text{Verbal mismatch} \quad \text{Science/mech mismatch} \quad \text{Social mismatch}] \longrightarrow m_{ic}$$

$$[\text{Math skills} \quad \text{Verbal skills} \quad \text{Science/mech skills}]$$

Social skills

PCA

Individual's skills

$$[\text{ASVAB 1} \quad \text{ASVAB 2} \quad ... \quad \text{ASVAB 6}]$$

Notes: Procedure for creating multidimensional occupation-skill mismatch measure for a given individual and occupation. "Task – occ" is the O*NET score of the importance of a task for a given occupation. "Task – ASVAB" is the DMDC crosswalk's relatedness score of a task for a given ASVAB section test.

Figure 2: Effects of Children in the NLSY79
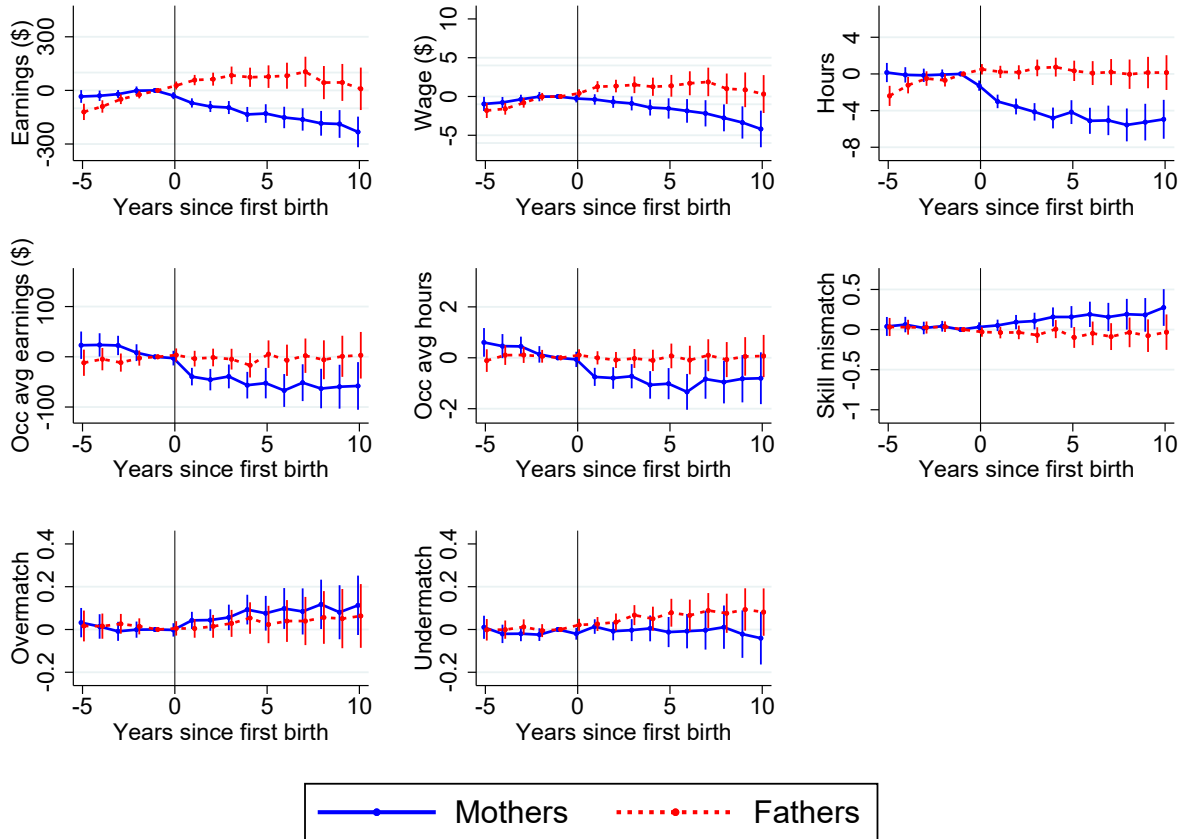
Notes: Effects of children on weekly earnings; hourly wages; weekly hours worked; average weekly earnings within occupation; average weekly hours worked within occupation; and standard deviations of multidimensional occupation-skill mismatch, skill overmatch, and skill undermatch among working mothers (blue solid line) and fathers (red dashed line) in the NLSY79. Vertical lines denote 95% confidence bands.

Figure 3: Effects of Children in the NLSY97
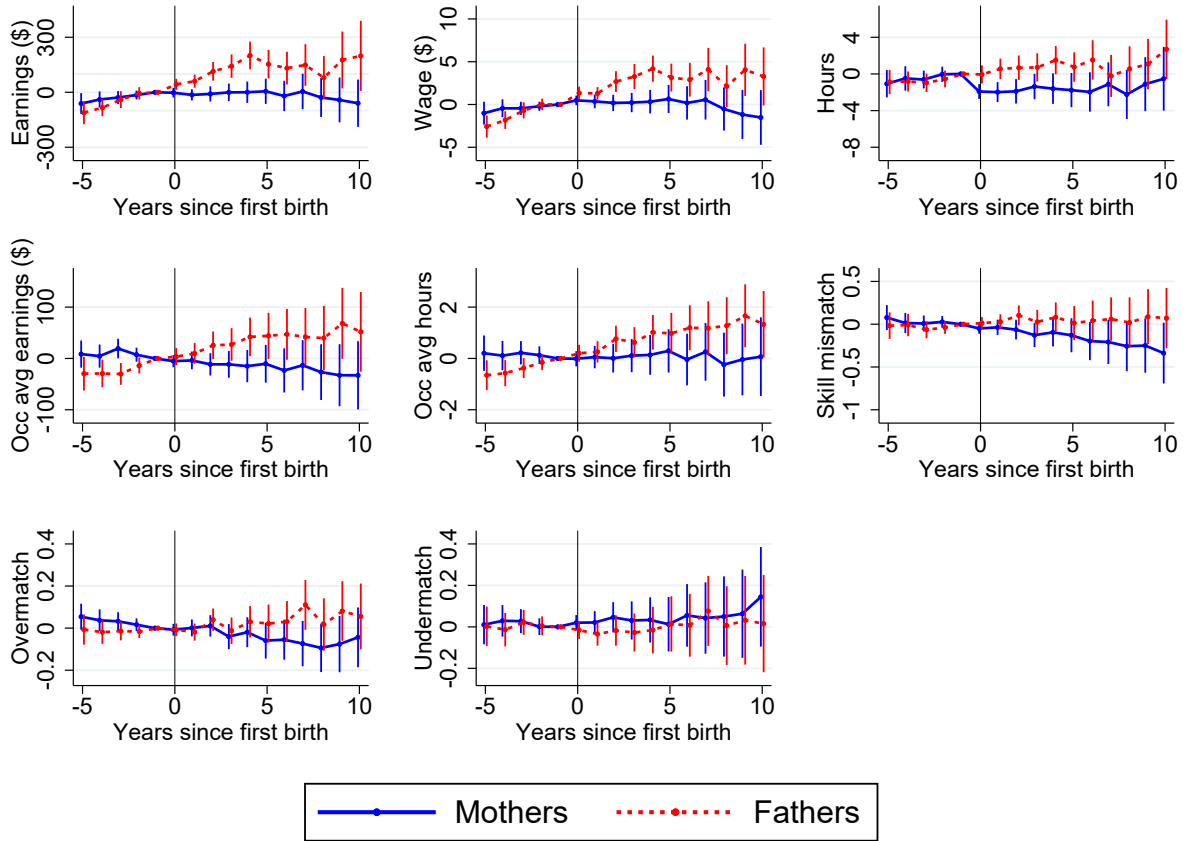


Notes: Effects of children on weekly earnings; hourly wages; weekly hours worked; average weekly earnings within occupation; average weekly hours worked within occupation; and standard deviations of multidimensional occupation-skill mismatch, skill overmatch, and skill undermatch among working mothers (blue solid line) and fathers (red dashed line) in the NLSY97. Vertical lines denote 95% confidence bands.