Baby bumps in the road: The impact of parenthood on job performance and career advancement

Olivia Healy
Northwestern University

Jennifer Heissel
Naval Postgraduate School

December 14, 2020

(Click for Latest Version)

Abstract

We evaluate the impact of parenthood on men and women's job performance and career advancement using detailed data from the U.S. Marines. For parents who remain employed after having a baby, disruptions in home life and health may spillover into their performance at work. Using monthly data from 2010 to 2019, we exploit variation in the precise timing of first births to identify impacts on health-dependent measures of worker performance. We then compare parents' promotion trajectories to similar non-parents' trajectories, using a machine learning approach that assigns non-parents to "placebo births." We find negative impacts on parents' employer-assessed physical fitness and supervisor-rated job performance, concentrated mainly among women. Consistent with these findings, women's promotion trajectories slow down in response to childbirth but men's do not. In a complementary analysis, we exploit sudden policy changes to the length of paid maternity leave to compare impacts. Longer leaves exacerbate declines in women's job-related physical fitness and slow their promotion trajectories to a greater degree. Results suggest longer periods away from work due to maternity leave may have the unintended effect of eroding job-specific skills and delaying career advancement for new mothers.

JEL Classification: J24, J16, J18, J45
Keywords: Childbirth, Job Performance, Parental Leave

1 Introduction

Almost four million babies are born in the U.S. every year (Hamilton et al., 2019), dramatically changing parents' daily lives. Parents with newborns experience sleep deprivation, increased emotional stress, and changes in their neurobiology (Saxbe et al., 2018). For women in particular, pregnancy and childbirth are major medical events. Women experience immediate risk of infection, complications and postpartum depression after having a baby (Memon and Handa, 2013; O'Hara and Swain, 1996), and physiological recovery takes on average one year (Melchiorre et al., 2016).

Despite the ubiquity of parenthood, we have limited evidence on how the mental and physical strain of having a baby implicates parents at work, particularly in terms of their on-the-job performance. Prior research has necessarily focused on new parents' labor force attachment, hours worked, or wage. These are often the only measures available, and mounting evidence shows women increasingly leave the labor market, cut back on hours worked, and earn lower wages after having a child (Agüero and Marks, 2011; Angrist and Evans, 1998; Bronars and Grogger, 1994; Cáceres-Delpiano, 2006; Cools et al., 2017; Cruces and Galiani, 2007; Jacobsen et al., 1999). These changes drive persistent earnings penalties that accrue to mothers, but not fathers, over decades (Angelov et al., 2016; Barth et al., 2017; Bertrand et al., 2010; Kleven et al., 2019a,b). One possible explanation for these findings is that mothers prefer to spend more time on caregiving or face more societal pressure to provide care than fathers. Another possible explanation is that job demands become overly taxing after having a baby, perhaps especially so for women such that they opt to stay home or move into lower paid, more flexible positions in response to motherhood.

In the current paper, we explore this second possibility, considering whether new parents become less physically able to perform well at work after having a child. We then consider whether women and men who become parents suffer slower career advancement due to having a child. Last, we investigate whether access to longer maternity leave helps offset any negative consequences of parenthood on mothers' health-related job performance and career advancement. We focus on workers who are largely required to stay on-the-job after becoming parents, limiting selection out of employment due to the demands of parenthood.

The primary advantage we have in addressing our research questions is access to detailed, consistently measured, longitudinal data for individuals. Data come from Department of Defense (DoD) administrative records on service members in the U.S. Marine Corps between 2010 and 2019. Records provide high-frequency, repeated measures of worker performance, including job-related physical fitness tests and supervisor evaluations of job proficiency. Our analytic strategy attempts to isolate the causal effect of childbirth on parents' job performance, separately for men and women. We use an event-study framework and leverage the precise timing of childbirth as an exogenous shock to parents' work performance outcomes. Next, we compare first-time parents' promotion trajectories to similar non-parents' trajectories, using a machine learning approach that assigns non-parents to "placebo births." We then examine whether gaps in career advancement open up across the birth vs. placebo birth event. We estimate the impact of maternity leave length on women's job-related physical fitness and promotion trajectories after childbirth by relying on both the event-study and machine learning matching strategies. Two major changes to DoD paid maternity leave policies over our study window create exogenous variation in the length of paid leave available to women.

While the military differs in obvious ways from many professions, aspects of the DoD context make it especially suitable for this analysis. First, we capture a diverse group of workers. Individuals who enlist in the Marines work in a range of occupations, including ones common to the civilian workforce, such as food service, traffic management, information technology, and more. Second, effective job performance in the Marines is tightly linked to physical health. Regardless of their occupational specialty, Marines need to be physically ready to support combat missions at any time. Although this professional requirement may seem unique, nearly half (45%) of jobs in the civilian labor market require at least medium physical strength, defined as work that involves frequent lifting or carrying of objects weighing up to 25 pounds (Bureau of Labor Statistics, 2017a). Moreover, the average civilian worker stands or walks for 60% of the day (Bureau of Labor Statistics, 2017b), requiring a baseline level of physical health. As such, the Marine context offers an opportunity to learn how parenthood impacts work performance for those in jobs that require both mental and physical acuity. Third, Marine Corps job performance measures are standardized across all occupation types. Whether a Marine works in food service or as a lawyer, he or she is regularly rated on the same performance metrics to determine promotion. Fourth, Marines generally commit to three- or four-year contracts, which limits selection out of work after having a child, especially compared to other employment contexts.

A final benefit of the DoD context is the opportunity to study universally-accessible, fully-paid parental leave, generally considered the gold standard by policymakers and firms but not widely available in civilian contexts. In the U.S., state-level paid leave programs generally cap wage replacement (Rossin-Slater et al., 2013), and among private firms about half of leave is fully paid, a third is partially paid, and the remainder is unpaid (Donovan, 2019). In cases where leave is not fully paid, the decision to return to work may be driven by financial need, particularly for lower-income women (Rossin-Slater et al., 2013). In our study context, fully-paid leave means take up rates are high. Also in our context, the length of paid maternity leave ranges from 6 to 18 weeks, well within the scope of expansions considered by policymakers at the federal- and state-levels in the U.S.

Using the event-study approach, we find small but meaningful impacts of the transition to parenthood on job performance, concentrated mainly among women. First, we show women's physical performance on the job is substantially lower after having a baby. Lower physical performance persists through the end of our follow up period, two years after birth. Second, mothers receive progressively lower scores on supervisor-rated performance evaluations with each month that passes after childbirth. Declines are small in magnitude, decreasing at a rate of 0.02 standard deviations per month relative to pre-pregnancy performance ratings.

In contrast, we observe minimal impacts of parenthood on fathers' job performance. The birth of a child leads to short-lived declines in men's physical performance. Fathers score 0.1 standard deviations below their pre-pregnancy levels on job-related fitness tests one month after the baby arrives. By the child's first birthday men's fitness performance is back to pre-pregnancy levels. Findings show fathers' supervisor-rated performance improves during the mother's pregnancy but improvements begin to fade once the child is born. By the child's second birthday men's job performance is rated the same as it was before the mother's pregnancy.

Consistent with the persistent negative impacts of parenthood on women's job performance, we find evidence that women's promotion trajectories slow as a result of having a child. Men's promotion trajectories are unaffected by the transition to parenthood.

Our last main finding is that longer paid maternity leave exacerbates performance declines in physical fitness and delays in career advancement for new women after having a child.[1] The longer a woman remains out of work after childbirth, the longer she struggles in terms of assessed physical ability on the job when she returns. We show that women who receive 6 weeks of maternity leave recover to their pre-pregnancy levels of job-related physical fitness within two years of childbirth. In contrast, women who receive more than 6 weeks of maternity leave perform below their pre-pregnancy fitness levels even two years after childbirth, on the order of magnitude of 0.2 to 0.4 SD lower (depending on the length of extended leave). In line with these findings, we document the largest gaps in promotion between mothers and non-mothers when maternity leave is longest. Results suggest that longer periods away from work may have the unintended effect of eroding job-specific skills and slowing career advancement.[2]

Our study contributes to a longstanding literature on the impact of fertility on parents' employment outcomes.[3] We build on this work in a few key ways. First, we directly measure actual performance on the job—a key precursor to worker output and productivity, as well as to any changes in employment, hours worked, and wage.[4] Second, while research has largely focused on female labor supply and

[1]Due to data limitations, we estimate the impact of leave length only on physical fitness performance and not on supervisor-rated job proficiency evaluations.

[2]While longer maternity leave exacerbates physical performance declines and slows promotion trajectories for women in our sample, this does not rule out that longer leave might have positive effects in other domains, such as mothers' health or children's well-being. For example, Balser et al. (2020) find similar maternity leave extensions for women in the Army and Air Force reduce postpartum depression diagnoses among mothers. In other U.S. and international contexts, evidence shows access to paid parental leave improves mother's physical and mental health (Butikofer et al., 2018; Bullinger, 2019), duration of breastfeeding (Pac et al., 2019), infant health (Bullinger, 2019), and time spent with children upon return to work (Trajkovski et al., 2019; Bailey et al., 2019).

[3]Much of the prior research in this area exploits variation in family size due to twin births or third births that result from having two prior children of the same sex. These studies find that having an additional child reduces mothers' employment and hours worked, among mothers in the U.S. and other countries (Agüero and Marks, 2011; Angrist and Evans, 1998; Bronars and Grogger, 1994; Cáceres-Delpiano, 2006; Cools et al., 2017; Cruces and Galiani, 2007; Jacobsen et al., 1999). Studies in this literature tend to find that fathers' labor force attachment is either unaffected by additional children (Angrist and Evans, 1998; Cools et al., 2017) or they do not study fathers' outcomes (Agüero and Marks, 2011; Bronars and Grogger, 1994; Cáceres-Delpiano, 2006; Cruces and Galiani, 2007; Jacobsen et al., 1999). A more recent literature reveals a large "child penalty" in earnings that accrues to mothers but not fathers in response to childbearing (Angelov et al., 2016; Barth et al., 2017; Bertrand et al., 2010; Kleven et al., 2019a,b).

[4]Most similar in spirit to our paper is a set of descriptive studies on changes in mothers' and fathers' job productivity across the transition to parenthood. For example, Azmat and Ferrer (2017) find that female lawyers with young children are less productive compared to male lawyers with young children in terms of hours billed annually, a key productivity measure in the legal profession. Similarly, Gallen (2018) explores how firm output in Denmark varies by the gender and parenthood

highlighted motherhood as a turning point in women's careers, we explore employment impacts on both women *and* men who work in the same setting and remain employed after having a child. Many papers consider fathers as a comparison for mothers when estimating child wage penalties (e.g., Angelov et al. 2016; Bertrand et al. 2010; Kleven et al. 2019a) and family policy impacts on mothers (e.g., Balser et al. 2020). If fathers are uniquely impacted by the transition to parenthood, such estimates may over- or under-estimate the effect of parenthood or policy, depending on how fathers are affected.

Our third key contribution is the ability to examine employment consequences of childbirth month-by-month and trace the dynamic responses of men and women *within* the first and second years of work after becoming parents. Prior papers that estimate child penalties use annual earnings or income and therefore cannot detect immediate, within-year impacts of childbirth (Angelov et al., 2016; Barth et al., 2017; Bertrand et al., 2010; Kleven et al., 2019a,b). Our findings on month-by-month performance impacts *within* the first year after birth point to one mechanism through which child penalties may arise. Persistent declines in health-related ability to perform at work for mothers may lead mothers – more so than fathers – to exit the labor market, cut back hours worked, or receive lower wages by the time these outcomes are measured one year post-birth in other papers. Our findings highlight the immediate period after birth as a possible critical window that gives rise to long-term child penalties.

Last, our paper adds to a growing literature on the impact of paid maternity leave on women's labor market outcomes. We contribute by isolating the impact of paid leave on a new dimension of employment, on-the-job work performance. Past research on the employment consequences of paid leave is mixed. Early studies in select U.S. states find that the introduction of paid leave improves women's labor force attachment and wages immediately following the birth (Baum and Ruhm, 2016; Byker, 2016; Rossin-Slater et al., 2013). However, more recent research finds negative effects of paid maternity leave on women's employment and wages in the long term (Bailey et al., 2019; Timpe, 2019). Among first-time mothers specifically, Bailey et al. (2019) show than an additional 6 weeks of paid leave reduces a

status of employees. She finds mothers are substantially less productive, as measured according to firm output, than other workers (non-mothers, fathers, and non-fathers), particularly during their childbearing years. Kim and Moser (2020) also find female scientists studied in the 1950s patented less during their childbearing years. They posit lower patenting productivity drives documented lower rates and slower speed of promotion to tenure for female scientists with children as compared to fathers and other women without children. These papers tend to focus on elite professions, while our results provide evidence on more diverse group of workers not currently represented in the literature.

new mother's likelihood of returning to work and her annual wages in *both* the short- and long-run. The authors' innovative research design exploits month of birth during the roll out of California's paid leave policy to compare women who were more or less likely to have access to leave directly after having a baby. Our paper adds to the current understanding of how paid leave helps or hurts women's career trajectories by considering the impact of longer paid time off on women's performance when they return to work and subsequent career advancement.[5]

The rest of the paper proceeds as follows. Section 2 provides institutional background on the DoD and details on employment, parenthood, and paid family leave in the Marines. Section 3 provides details on our data. Section 4 describes our empirical strategy, including estimation and identifying assumptions. We describe our results in Section 5 and conclude in Section 6.

2 Institutional Background

The DoD is the world's largest employer, with a total of 1.3 million active-duty service members across four branches – the Army, Navy, Air Force, and Marine Corps. Each branch plays a unique role in maintaining U.S. security and peace. In this paper, we focus on service members in the Marines Corps, where administrative records on job performance are readily available. The Marine Corps is an immediate response force, ready to deploy quickly to support combat missions on sea or land. Marine Corps service members make up nearly 15 percent of active-duty forces and have roughly 185,000 active-duty service members (Department of Defense, 2018).

2.1 Jobs and associated work in the U.S. Marines

Individuals begin in the Marine Corps either as a junior enlisted, akin to an entry-level worker, or as an officer, akin to a manager. Enlisted service members must have a high school degree and be between 18 and 29 years old when they begin. Officers must have at least a bachelor's degree upon entry. Service

[5] Using data on academic economists from 1980 to 2005, Antecol et al. (2018) show that family-oriented university policies can exacerbate gender-gaps in publication output and time to promotion. In their setting, gender-neutral tenure-clock stopping policies, designed to accommodate lower productivity among academics in the year after childbirth, increase men's likelihood of getting tenure while decreasing women's.

members are ranked on a numerical scale from E1 to E9 (for enlisted) and O1 to O10 (for officers). All promotions move members up in rank on the scale. Roughly 89% of Marine Corps Service members are enlisted.

In total, there are over 35 career fields in the Marines and dozens of specializations, referred to as Military Occupational Specialities, within each. Some career fields specific are specific to the military. These include infantry, field artillery, and terminal attack control (in charge of communicating with aircrafts to guide offensive air operations). Other career fields are also present in the civilian labor market. These include fields such as food services, financial management, military police and corrections, legal services, and even music (e.g., the Marine Corps band). An individual can enlist in the Marines under a career field, but he or she is not guaranteed a specific job specialization within that field. A service member's occupational specialty along with their assigned unit determines their day-to-day work environment. Generally, Marines work Monday through Friday. Each day typically begins with early morning physical training (as early as 5:30am), followed by work assignments through the evening. Most service members live and work on or near a military base and are stationed in the U.S. (83% of Service members; Department of Defense 2018), though some are stationed abroad. Marines are moved by central command roughly every 3 years for training or job assignments.

Marines sign a legally binding contract that outlines their required length of service. For enlisted service members, contracts typically require 4 years of active-duty service, while for officers the commitment is typically 3 years.[6] At the end of each contract, Marines can decide whether to re-enlist, which involves another contract with a time commitment. Importantly for our purposes, these binding contracts limit the extent to which Marines can exit the labor force after they have a baby.

All active-duty Marines work full-time. However, each military branch has a Reserve component where individuals can work part-time. Part-time work as a Marine reserve requires (1) participating in training drills one weekend per month and (2) attending a two-week work program each year. Reservists can be called upon for active-duty deployment in times of war or national emergency. All other times, Reservists typically work in civilian careers or are enrolled in higher education while they fulfill part-time

[6]Contracts can stipulate additional service beyond these minimums, requiring additional years of service in the Marine Corps Reserves, which allows service members to work part-time.

Marine Corps reserve duties. Our sample includes data on both active-duty and reserve Marines.

Effective job performance in the Marines requires both mental and physical acuity. The Marine Corps uses a standardized set of measures to evaluate performance among both active-duty and reserve Marines. Performance measures include supervisor-scored job proficiency evaluations, physical ability assessments, and a rifle marksmanship test.

Supervisors evaluate Marines for performance purposes using one of two scales, depending on the Marine's rank. Junior enlisted receive proficiency and conduct marks ("ProCons") at least twice per year, and senior enlisted and officers receive Fitness Report ("FitRep") scores at least annually. Both assessments require supervisors to rate a Marine's performance across a range of professional domains.[7] For physical ability performance assessments, Marines take two tests per year: the physical fitness test (timed running, crunches, and upper body strength) in the first half of the year and a combat fitness test (timed running, a combat-related obstacle course, and upper body strength) in the second half of the year. Last, rifle marksmanship assessments occur annually.

Scores on each of these measures give rise to a composite performance score for certain Marines, and composite scores are updated every quarter to incorporate new performance assessments.[8] Of these three measures, supervisor performance ratings influence promotion outcomes most heavily. Marines can clearly determine what they need to advance and, therefore, incentives to perform well on measured performance assessments are especially high.

[7] Domains include mission accomplishment (job-specific aptitude, competence, technical knowledge, and practical skills), character (courage, effectiveness under stress, and initiative), leadership (setting the example, communication skills), and intellect and wisdom (professional military education, decision-making ability, and judgement), among others.

[8] The composite score determines promotion for all ranks below E5, conditional on meeting requirements for minimum time in service and minimum time in the current job level to be eligible for promotion. (Promotions at lower ranks, E1 through E3, are relatively automatic after a given number of months in service and months in rank.) For promotion at ranks above E5, the same performance metrics, along with supervisor ratings, are reviewed by an evaluation board to determine promotion. As such, the Marine Corps promotion system is similar to private sector promotions based on work performance but arguably more exact. Marines can receive bonus points on their composite score if they serve as drill instructors, recruiters, or Marine security guards. They also receive bonus points as an E3 or E4 if they effectively recruit someone into the Marines. Last, Marines can receive bonus points for additional military or civilian education.

2.2 Parenthood and Paid Leave in the U.S. Marines

The DoD provides military parents with a number of family-friendly benefits, including fully-paid parental leave. We focus on policy changes to paid leave for primary caregivers (most often women) and refer to this leave as maternity leave.[9]

Prior to 2015, all DoD branches provided active-duty women with 6 weeks of paid maternity leave. In July 2015, the Secretary of the Navy announced that primary caregivers (most commonly women) in the Navy and Marine Corps would be entitled to 18 weeks of leave. Women who had given birth earlier in the year (as of January 2015) could retroactively take up the 18-week leave policy. For women who had already taken 6 weeks' leave and returned to work, they tended to use the additional 12 weeks of paid time off discontinuously. For women who were on leave or pregnant at the time of the announcement of expanded leave, the majority took the additional leave consecutively, as did mothers who became pregnant after the announcement (Bacolod et al., 2020). We analyze these two groups separately. In early 2016, the Secretary of Defense standardized maternity leave to 12 weeks for all services.[10]

3 Data

We draw data from the Marine's Total Force Data Warehouse and obtain records on all active-duty and reserve Marines who served at any point during January 2010 through December 2019. Our preferred sample includes first-time parents who were active-duty 10 months before the birth, in other words right before the pregnancy began. To ensure our results are not driven by selective attrition, we require first-time parents remain in our sample for 12 months prior to the birth and 24 months after. We also include a group of "non-parents" in our sample, defined as Marines with at least 36 months total of service who

[9]Changes to paternity leave (i.e., leave for "secondary caregivers") also occurred during this time. However, paternity leave is limited in scale (lasting 2-3 weeks, depending on the service branch), and expansion of paternity leave in the Marines was from 10 to 14 days. We do not focus in this paper on paternity leave impacts, given the small magnitude of the change.

[10]For Marine Corps mothers, this meant their paid maternity leave was reduced from 18 to 12 weeks. However, the policy change for Marines only applied to pregnancies that began 31 days after the announcement (i.e., to pregnancies that began on March 3, 2016 or later, per doctor estimation). As a result, Marine women who became pregnant after the policy announcement were aware they would get 12 rather than 18 weeks of paid leave. Given that only one month's notice was given for the policy change, many of the women who became pregnant at the beginning of the 12-week policy change likely made the decision to become pregnant before the policy announcement.

do not experience a birth during the study window. Table A1 shows how the characteristics of the sample change based on a variety of possible sample restrictions.

Our data include basic descriptive information on service member characteristics (age, gender, race/ethnicity, education status, and AFQT scores – a measure of intelligence), dependent characteristics for spouses and children (exact date of birth, gender, race/ethnicity, and whether a spouse is in the military), and job characteristics of the service member (job type, rank, time in service, time remaining in job commitment, and unit location). Our outcomes include two of the three primary measures of job performance used for promotion and retention decisions: supervisor ratings of job proficiency and physical ability assessments. We do not have data on the third measure (rifle marksmanship assessments).

For our first outcome, supervisor ratings of job proficiency, we standardize scores by year, gender and test (ProCons vs. FitRep). We then combine the Z-scores of the two tests into one outcome we call supervisor performance ratings. Note, we are missing ProCons for junior enlisted who left the Marines before October 2017. However, we observe the full history of performance ratings (including ratings prior to October 2017) for any service member who was active as of October 2017. For our sample with complete ratings, we observe scores at least twice per year among junior enlisted, once in the first half and once in the second half. For senior enlisted and officers, we observe supervisor ratings a minimum of once per year. If a Marine is transferred, discharged, or promoted, or if their supervisor changes, they will receive additional performance ratings. Marines are moved every few years, as are their designated supervisors. Decisions on movement are made from a central location, which prevents Marines from manipulating their scores by selecting their supervisors (Cunha et al., 2018). Nonetheless, the subjective nature of these assessments means we cannot distinguish true changes in job performance from supervisors' *perceptions* of changes in performance using this measure.

Our second outcome, physical performance, serves as an objective measure of job proficiency among Marines. Physical performance scores based on fitness tests are awarded on a 300-point scale, which is adjusted for age and gender such that women do not need to do as many push ups as men and older service members do not need to run as fast as younger ones to achieve the same score. We standardize raw physical fitness scores by year, gender and test type (one in the first half of the year and the other in the second). We combine the Z-scores for the two tests into one measure, generally observed twice per year per Marine. We measure physical performance outcomes for mothers up until 10 months before the birth and starting 8 months after the birth. Due to the physically demanding nature of the assessment, women are not required to take the test when pregnant, and they are exempt from tests for 6 months after giving birth. Due to concerns that commanding officers may allow some women whose test dates fall 7 months after birth to skip the test during that assessment round, we resume measurement at 8 months post birth.

Finally, using information on Marines' job rank over time, we can track promotion outcomes for Marines in our sample. We count the number of promotions a Marine receives relative to 10 months before they have a child. A value of one on the variable indicates one promotion achieved since the time point before pregnancy (t=-10).

We present characteristics of first-time parents in our preferred sample in Table 1, alongside characteristics of first-time civilian parents who are employed and have a child under the age of one. First-time parents in the Marines are younger than their civilian counterparts and have much lower rates of college-going and college completion. First-time Marine mothers and fathers also identify as Black or Hispanic at higher rates than first-time civilian parents. In contrast, marriage rates are generally similar: 86% of Marine vs. 83% of civilian fathers are married when they have their first child; and 71% of Marine vs. 78% of civilian mothers are married at first birth. We rely on the Standard Occupational Classification (SOC) system, a federal standard used to classify workers into occupational categories, to explore the distribution of job types among Marines in our sample relative to civilians. We crosswalk Marine job codes to SOC codes and find that – outside of military-specific occupations – the largest share of first-time Marine fathers work in natural resources, construction, or maintenance, while the largest share of first-time Marine mothers work in sales or office roles.[11] Civilians who have a first child and stay in the workforce tend to be employed in management, business, science, or arts.

Only a small share of first-time Marines parents in our sample are officers (akin to mangers): 14% and 9% of Marine mothers and fathers, respectively. As such, the vast majority of first births occur to

[11]Note that the vast majority of these jobs in the Marine Corps are office jobs but categorized under the umbrella of sales and office.

enlisted Marines. Finally, new Marine parents score just above average on military-specific intelligence tests, including the AFQT and GCT.

Based on descriptive differences between Marine and civilian first-time parents, results from our analyses may generalize best to younger, less educated, and more racially diverse workers.

4 Empirical Approach

4.1 Primary Research Design and Estimation

The ideal experiment to isolate the causal effect of fertility on men and women's work performance would randomly assign pregnancy and parenthood to workers. Random assignment would ensure that – on average – differences in work performance were not driven by underlying characteristics of the types of men and women who chose to become parents but rather by the transition to parenthood itself. Of course, random assignment of childbirth/pregnancy to individuals in the workforce is both unethical and infeasible. Yet, a simple post-hoc comparison of parents' relative to non-parents' is unlikely to recover a causal estimate of the effect of having a child. Those who opt into parenthood likely differ from non-parents in ways that might also correlate with work performance.

In the absence of a feasible experiment, we rely on variation in the precise timing of births to identify the effect of childbirth on first-time parents' physical health and work performance outcomes. If the transition to parenthood has an impact on health and performance, then the birth should generate a sharp change in these outcomes directly after it occurs. We can attribute any discontinuity in the outcomes at the time of the birth to the birth itself if we assume that other factors that shape job performance do not also undergo a sharp change in the same month as childbirth. In other words, while the choice to have a child may be endogenous, our event study approach exploits the exact timing of conception and subsequent childbirth as an exogenous shock to the outcomes of interest.

Our modeling approach follows first-time parents over time, examining how their outcomes change during pregnancy, the immediate post-birth period, and up to 24 months post-birth, relative to the pre-pregnancy period. We begin by estimating a fully flexible, dynamic specification in order to observe how the effect of parenthood changes with time since pregnancy and childbirth. We estimate the following model, separately for men and women:

$$Y_{it} = \alpha_i + \phi_t + \theta_{it} + \sum_{r=k_{min}}^{k_{max}} \mathbb{1}(t = t_i^* + r)\beta_r + \varepsilon_{it} \tag{1}$$

where β_r represents the effect of a birth in month t_i^* on outcomes r months later (or r months before, if $r < 0$). Effects are measured relative to month $r = -10$, which corresponds to 10 months prior to the birth and approximately one month before the start of the pregnancy. In other words, β_{12} would represent the average outcome 12 months after the birth, relative to $r = -10$ (the month before pregnancy). We censor r at $k_{min} = -18$, or $k_{min} = -24$, depending on the outcome.[12] We include α_i, an individual fixed effect to account for stable individual differences; ϕ_t, a month by year fixed effect to account for general changes over time in the outcome; and θ_{it}, an age fixed effect to account for life cycle trends in the outcome. ε_{it} is the error term. The estimation of month by year fixed effects is assisted by the inclusion of non-parents in the data, who provide an estimate of universal time-patterned changes to the outcome that affect all Marines similarly (e.g., due to changes in a fitness test standards in a particular year). We run all models separately by gender.

Individual month coefficients β_r are not of particular interest; rather, we would like to generally identify any declines (or improvements) in performance during pregnancy, any drops immediately following birth, and any recovery following the immediate impact of birth. Similar to Lafortune et al. (2018), we create a more parsimonious model of performance changes over time relative to before the pregnancy by defining:

$M_{it}^{pregnancy} = t - (t_i^* - 9)$, if person i has a baby at time t_i^* and $t_i^* - 9 \leqslant t < t_i^*$, and $M_{it}^{pregnancy} = 0$ otherwise (for monthly trends during pregnancy)

$M_{it}^{drop} = 1$, if person i has a baby at time t_i^* and $t > t_i^*$ and $M_{it}^{drop} = 0$ otherwise (for postnatal drops following birth)

[12] We examine physical performance scores 24 months prior to childbirth but measure supervisor ratings of job performance only 18 months prior to childbirth. Supervisor ratings cover roughly a 6 month retrospective period, therefore any rating awarded 18 months before the birth implicitly covers an earlier time period of performance.

$M_{it}^{recovery} = t - (t_i^* + q)$, if person i has a baby at time t_i^* and $t > t_i^* + q$, where q is time point when the person is eligible to be observed for the given outcome after the birth, and $M_i t^{recovery} = 0$ otherwise (for monthly trends above and beyond the drop in level), and

$M_{it}^{\Delta recovery} = t - (t_i^* + 12)$, if person i has a baby at time t_i^* and $t > t_i^* + 12$, and $M_{it}^{\Delta recovery} = 0$ otherwise (for any change to the monthly recovery rate that begins 13 months post-birth).

Using these different time frames, we then estimate a semi-dynamic specification that fits linear splines to portions of the data, as follows:

$$Y_{it} = \alpha_i + \phi_t + X_{it}\theta + M_{it}^{pregnancy}\beta_1 + M_{it}^{drop}\beta_2 + M_{it}^{recovery}\beta_3 + M_{it}^{\Delta recovery}\beta_4 + \varepsilon_{it} \quad (2)$$

Here, β_1 captures the monthly linear change in the outcome during the pregnancy period ($t = [-9, -1]$), relative to the pre-pregnancy period average ($t \leqslant -10$). The effect captured by β_2 represents the acute postnatal drop (if any) in the outcome in the first month parents are again assessed after childbirth, relative to the pre-pregnancy period average ($t \leqslant -10$). Then, β_3 captures the monthly linear recovery in the outcome following that initial drop, and β_4 captures any change in the monthly linear recovery rate in the child's second year of life ($t = [13, 24]$). We present a diagram of this model in Figure 1. We use this semi-dynamic spline specification as our main model to estimate the magnitude of and confidence internal around the impact of birth on health and job performance outcomes. We present parameters from this model and their implications for effects at 12 and 24 months post-birth in subsequent tables of results.

4.1.1 Identifying Assumptions

Our event study approach is a form of a difference-in-differences model. Two key assumptions must hold for us to recover plausible causal estimates of the impact of childbirth under this approach. First, it must be the case that if a service member did not have a child, his or her outcomes would have evolved similarly to the outcomes of the other same-gender adults in the sample who did not have a child in that month. If this assumption is plausible, we would expect differences in outcomes do not diverge in the period before the pregnancy. We test this assumption empirically, finding reasonable confirmation of parallel trends before pregnancy, particularly among women for whom we document the most substantial effects. Second, it must be the case that no other simultaneous shock occurs to parents in the same month as the pregnancy or the birth that would affect the outcomes of interest. While we cannot test this second assumption, we do not believe it likely that other simultaneous shocks occur in the exact month of birth across differently timed pregnancies and births for *only* the parents in our sample.

4.2 Impacts of Parenthood on Career Advancement

Beyond changes in job performance, we explore whether having a child impacts men and women's career advancement. Unlike with our job performance measures, we cannot study promotions using an event-study framework. Promotions are not high frequency, repeat outcomes, and one's propensity to be promoted is a function of how recently she or he received their last promotion. As a result, to investigate differences in promotion outcomes due to having a child, we create a comparison group of non-parents where we assign placebo births to those in our sample who did not have a child during the study window. We then compare promotion outcomes across the "birth event" for Marines with actual births vs. those with placebo births to trace whether gaps in outcomes arise during pregnancy or after the birth.

We assign placebo births to non-parents in a two step process. First, we identify non-parents who are similar to first-time parents in terms of age, race/ethnicity, military entrance exam scores (AFQT scores), marital status (including whether a spouse is also in the military), level of education, occupational field, and most recent physical performance score, measured for first-time parents 10 months before the birth.[13] We then identify non-parents who have the same job rank and number of months in service as future first-time parents, measured 10 months prior to having their first child. Of these non-parents with exact matches on rank and time in service, we select five who show the highest similarity on background characteristics as calculated in step one. We assign each of these five non-parents to a placebo birth 10 months after the time of the match, which aligns with the timing of the birth for the first-time parent to

[13]The identification process begins by using an adaptive ridge least absolute shrinkage and selection operator (LASSO) model with 10-fold validation to predict who will have a baby among the parents we observe having a baby and the set of individuals who make up the potential set of placebos. We run this analysis separately for females and males.

whom they match.[14] Our exact match on job rank and number of months in service ensures that Marines without children have similar past promotion histories as first-time parents before the pregnancy. We then compare changes in outcomes for first-time parents to the average change in outcomes for the five non-parents to whom they match across the birth event.

To lend confidence to this research design, we use it to re-estimate the impact of childbirth on Marines' supervisor ratings and physical performance. If we observe consistent findings across the event-study and placebo birth strategies, we increase our confidence in obtaining plausible causal estimates of promotion impacts using the placebo birth approach.

4.3 Maternity Leave Policy Impacts

Finally, we investigate whether access to longer paid maternity leave helps or hurts women's work performance across the transition to parenthood. We group first-time mothers into four policy categories, P_{it}^j, based on the amount of paid leave available at the time of birth. We then estimate:

$$\begin{aligned} Y_{itmy} = &\alpha_i + \gamma_t + \theta_{it} + P_{it}^j(M_{it}^{pregnancy}\beta_{j1} + M_{it}^{drop}\beta_{2j} + M_{it}^{recovery}\beta_{3j} \\ &+ M_{it}^{\Delta recovery}\beta_{4j}) + \varepsilon_{itmy} \end{aligned} \quad (3)$$

where P_{it}^1 is the initial (baseline) six-week policy. P_{it}^2 is the retroactive 18-week policy, where parents expected six weeks of leave but actually received an additional 12 weeks of leave that could be used discontinuously once back at work. P_{it}^3 is the 18-week policy where mothers could take the 18-weeks of leave continuously. Some of those who fell under P_{it}^3 received the extension in leave once already pregnant, while others - as time went on - knew about the 18 weeks of leave before deciding to become pregnant. The last policy, P_{it}^4, changed leave to 12 weeks continuous paid time off. All of those who fell under P_{it}^4 knew they would receive 12 weeks of leave before conception. These latter two policies, P_{it}^3 and P_{it}^4, could allow some manipulation by parents in anticipation of the more generous leave length. We discuss this further below.

Each coefficient from Eq. 3 represents the pregnancy trend β_{1j}, postnatal drop β_{2j}, recovery β_{3j}, and

[14]Each parent receives a weight of one in the analysis, while each match receives a weight of 0.2 per match-month.

change to recovery β_{4j} for each policy. As in the main analysis, the timing of the postnatal drop and the start of recovery time trends will vary by outcome, given that certain outcomes are not observed during pregnancy and for a period after childbirth (e.g., women's physical fitness test scores).

Naïve evaluations of parental leave policies may not be causal because in most settings the quantity of leave is not randomly assigned. In many settings, for instance, certain types of mothers (e.g., more advantaged mothers) may make choices or work at firms that provide more leave, but these mothers may have higher (or lower) work performance for reasons unrelated to the amount parental leave provided. To interpret our policy impact estimates as causal, it must be the case that women's potential job performance outcomes are not correlated with the length of their maternity leave.[15] We explore empirical evidence regarding selection into maternity leave length and present findings in our results section.[16]

5 Results

5.1 Main Impacts

We begin by examining evidence on our identifying assumptions. Figure 2 presents results from our flexible event-study model estimated using Eq. 1. The figure includes point estimates for pre-pregnancy effects, displayed in shaded blue on each graph. We conduct F-tests to assess whether the individual month coefficients β_r from Eq. 1 jointly differ from zero for months $r \leqslant -10$. The significance value for each F-test is presented in Figure 2 below each graph. We also present the slope parameter for

[15]We considered regression discontinuity (RD) as an alternative strategy for the leave analysis, given that unexpected policy changes apply to births on either side of precise birth date cut points. However, the cut points happen to fall near the end of the fiscal and calendar year – which may affect outcomes for other reasons. To address this, one option would be to embed an RD approach within a difference-in-differences strategy, comparing outcomes 3 months before and after the unexpected policy change in the year of the reform to years where no reform occurs for example, following Persson and Rossin-Slater (2019). However, we are concerned about statistical power, particularly for analyses focused on women's outcomes, given the small number of women in our dataset.

[16]Many Marine women learned about their eligibility for longer paid leave only after they returned to work or after they were pregnant, limiting women's ability to select into longer paid leave. For those women who got pregnant soon after a policy change (e.g., those women who received 12 weeks of paid leave already aware that this would be the new policy when they became pregnant), we must assume they do not select into or out of parenthood in response to the policy change. In addition, we restrict our policy impact analyses to women who received extended leave unexpectedly (once already pregnant or having given birth), mitigating concerns of selection into parenthood due to leave length. We compare whether our the results from Eq. 3 are qualitatively similar to this more restricted sample.

the pre-trend estimate alongside the standard error and significance of the estimated slope.[17] Overall, the parallel trends assumption appears to be reasonable satisfied, especially in our sample of mothers. Among fathers, there is some evidence that physical performance scores and supervisor ratings are trending positively before the pregnancy and that average pre-pregnancy outcomes are statistically different from zero. We interpret our findings on fathers taking into account these patterns.[18]

The event-study estimates of the impact of parenthood displayed in Figure 2 Panel A show women and men perform more poorly on job-related physical performance assessments when initially assessed after having a child. For women, performance declines are large and persistent. Even 24 months after giving birth, women continue to struggle to perform as well physically as they did before becoming pregnant. For men, performance declines begin during the mother's pregnancy and reach their lowest point one month after the child's birth. Declines are short-lived. By 12 months after the birth, men appear to perform at their pre-pregnancy levels on physical fitness tests. Figure 2 Panel B shows some evidence of lower supervisor ratings for women in the two years after having a child, though estimates are noisy. There appears to be no impact of having a child on fathers' supervisor-rated job performance.

To quantify the size of the impacts and draw statistical inference, we next a more parsimonious, semi-dynamic parametric model that corresponds to patterns in our flexible models (diagrammed in Figure 1). The results from the semi-dynamic specification are summarized in Table 2. The parameters in Table 2 include: a pregnancy trend, which captures any change in the outcome by month of pregnancy; a post-birth drop, which captures any level shift in the outcome for all months observed after the birth through 24 months post; a recovery trend, which captures any monthly changes in the outcome during the post-birth period through 24 months post-birth; and a Δ recovery trend, which reflects any change in the recovery slope during months 13 to 24 after the birth. All parameters are relative to the pre-pregnancy period.

[17] We define $M_{it}^{pretrend} = t-(t_i^* - 10)$ if person i has a baby at time t_i^* and $t < t_i^* - 9$ (and 0 otherwise); we then estimate:

$$Y_{it} = \alpha_i + \phi_t + X_{it} + M_{it}^{pretrend}\beta_1 + M_{it}^{pregnancy}\beta_2 + M_{it}^{drop}\beta_3 + M_{it}^{recovery}\beta_4 + M_{it}^{\Delta recovery}\beta_5 + \varepsilon_{it} \quad (4)$$

This matches Eq. 2 except that it adds a linear trend for all of the pre-pregnancy months other than $r = -10$, which serves as the reference. We expect the event of pregnancy/birth to drive performance differences only *after* the pregnancy has taken place. Therefore, we expect to find no evidence of differences in trends before the pregnancy.

[18] Sharp discontinuities after the birth in fathers' physical performance scores suggest effects are not merely a continuation of preexisting pre-trends or level differences for this outcome. For supervisor ratings, despite some evidence of pre-trends among fathers before the pregnancy, we do not see large deviations in outcomes after the birth.

Recall, women are not subject to physical fitness tests while pregnant, so we do not estimate a pregnancy trend for women's physical performance. After the birth, we once again measure women's physical performance scores starting 8 months post-birth and women's supervisor ratings starting 6 months post-birth due to exemptions for women from testing and to account for maternity leave. Also recall that parents in our sample remain on the job as Marines for 24 months following the birth.

Columns (1) and (2) in Table 2 present results of the impact of parenthood on women's physical fitness performance scores and supervisor-evaluated job proficiency ratings. Beginning at 8 months post-birth when women are generally required to take fitness tests again, mothers perform 0.40 standard deviations below their pre-pregnancy average. Mothers recover from this initial drop at a rate of about 0.06 standard deviations per month in the first year of the child's life. By 12 months post-birth, women's expected physical performance scores are 0.18 standard deviations below their pre-pregnancy levels (p-value[12-month effect]=0.000). Mothers' physical fitness recovery slows to nearly zero in the child's second year of life (where recovery is the combination of the main recovery trend and the additional Δ recovery trend during the 13 to 24 month post-birth). Two years after having a baby, mothers' predicted physical performance remains 0.15 standard deviations lower than before pregnancy (p-value[24-month effect]=0.000).

Impacts of childbirth on women's supervisor performance ratings are small but accrue over time. As indicated by the pregnancy trend and post-birth coefficient in Table 2, Column (2), there is no change in supervisor-rated job performance during pregnancy and no immediate drop in supervisor ratings six months after the birth when we again measure ratings for mothers. However, mothers' supervisor ratings steadily decline during the post-birth period through 24 months post-birth at a small but statistically significant rate of -0.02 standard deviations per month. Two years after becoming a parent, women are rated 0.09 standard deviations below their pre-pregnancy levels of job performance (p-value[24-month effect]=0.048).[19]

Next, we turn to the outcomes for new fathers, shown in Table 2 columns (3) and (4). The pattern of effects on fathers' job-relevant physical performance is consistent with mothers but smaller in mag-

[19]Though supervisors are trained for consistency on job performance evaluations, we cannot distinguish true differences in performance from supervisors' *perceptions* of differences in performance among workers after they have a child.

nitude. Marine men's physical fitness scores decrease at a rate of 0.005 standard deviations per month during the mother's pregnancy. There is a substantial drop in physical performance in the month immediately following the birth to 0.07 standard deviations below pre-pregnancy performance levels. Fathers on average recover by the child's first birthday, such that new fathers' physical fitness actually exceeds their pre-pregnancy levels by 0.04 standard deviations 12 months after their child's birth (p-value[12-month effect]= 0.000). Fathers' rate of recovery in physical fitness performance flattens in the second year post birth (reflected by the combination of the recovery and Δ recovery trends). Twenty-four months after having a child, fathers' physical ability scores are 0.02 standard deviations higher than they were before the pregnancy. Though we observe small statistically significant improvements in fathers' physical performance at 12 and 24 months post birth, we interpret this conservatively as evidence of a return to normal, given the pre-trends observed in the flexible models (see Figure 2).

The impacts of the transition to fatherhood on men's supervisor-rated work performance in the overall sample are minimal, and again, we interpret findings with caution given evidence of pre-trends in the outcome before pregnancy. Table 2, Column (4) shows small positive trends in men's supervisor ratings during the mother's pregnancy, with ratings increasing by 0.003 standard deviations per month up until childbirth. The post-birth parameter is not statistically significant, though the recovery trend suggests small improvements in ratings during the first year of the child's life. During the child's second year of life (months 13 to 24) recovery slows to nearly zero (reflect in the combination of the recovery and recovery trends). Fathers' predicted supervisor-evaluated performance ratings are 0.05 standard deviations higher than before the pregnancy (p-value=0.000) 12 months after having a child, and 0.03 standard deviation higher 24 months after having a child (p=value=0.045). These small effect sizes combined with largely null results shown in Figure 2 lead us to conclude there are negligible effects of parenthood on fathers' supervisor-rated job performance.

5.1.1 Robustness of Main Impacts

Tables A2-A3 show the sensitivity of our results to various sample restrictions. Across columns (1) to (6) in each table, we progressively limit the sample to minimize attrition over the study window among

Marines with births and without births. This allows us to explore the extent to which our findings are driven by compositional differences in who remains in the Marines after having a child. Our preferred sample used to estimate the main impacts requires first-time parents remain in the sample through 24 months after having a child. This sample is presented in Table 2 under Column (5). Column (6) limits our preferred sample to Marines who remain due to contract requirements. In other words, Column (6) restricts to first-time parents who had a child with at least 24 months remaining in their contract.

Across sample specifications, Table A2 shows the impact of childbirth on women's physical fitness and supervisor-rated performance is largely consistent. One exception occurs among first-time mothers who have a child when at least 2 years remains in their contract. Patterns of effects on mothers' supervisor rated performance appear to diverge for this group.

In terms of effects across samples among men, Table A3 shows that the direction and magnitude of the coefficients are generally consistent. One exception is that father's predicted fitness levels 12 months after birth (the 12-month effect) appear negative for results in Columns (1) to (3) and positive for results in Columns (4) to (6). This pattern suggests that new fathers who do not stay in the sample past 12 months after having a child (those omitted from specifications [4] through [6]) may be more negatively affected in terms of physical fitness after having a child. Results on men's supervisor job performance ratings also vary slightly across samples, however we caution readers to interpret the impact of fatherhood on supervisor ratings as largely null, in part due to this variation.

5.2 Impacts of Parenthood on Career Advancement

Before we turn to the impacts of childbirth on first-time parents' career advancement, we first replicate our results on supervisor ratings and physical performance using our research design that assigns placebo births to non-parents. Figure 3 shows patterns of job performance outcomes between parents and non-parents assigned placebo births. Gaps in outcomes between parents and non-parents emerge only after childbirth for both men and women. Results are generally consistent with findings from our event-study models, suggesting the placebo birth strategy is a reasonable approximation of our event-study design.

Figure 4 presents promotion trajectories of parents and non-parents across childbirth, separately for father and mothers. We measure the outcome as the number of promotions relative to t=-10, the time period before the pregnancy began. Among fathers and non-fathers, promotion trajectories evolve nearly identically. By 24 months after having a child, fathers on average have achieved one additional promotion beyond their job rank as measured in the month prior to the pregnancy, as have non-fathers. In contrast, gaps in career advance emerge between mothers and non-mothers starting during women's pregnancies. These gaps continue to widen through 24 months post birth, such that first-time mothers have received 0.1 fewer promotions than counterfactual non-mothers, relative to their job rank before the birth. Importantly, trends in mothers and non-mothers' outcomes overlapped before the pregnancy began, indicating that the two groups were on similar career trajectories prior to the pregnancy that only mothers experienced.

5.3 Maternity Leave Policy Impacts

5.3.1 Women's Physical Performance

We next turn to the evaluation of maternity leave length and its effect on new mothers' job-related physical fitness tests. Results are presented in Table 3. We include the overall estimated effect as the first row of the table; the bottom section of rows show estimates of effects by paid maternity leave length. We do not explore the impact of variation in length of maternity on supervisor-rated job performance outcomes. We only have data on supervisor evaluations for enlisted Marines who were in still in service as of October 2017, and therefore we lack sufficient observations of supervisor ratings to study policy changes that took place in 2015 and 2016 on this outcome.

Among mothers across all maternity leave policies, physical performance takes a hit when women return to work eight months after having a baby. All of the drops at 8 months post-birth statistically differ from zero, regardless of the maternity leave policy in place. However, women who had longer leave suffered larger physical fitness performance declines when back at work 8 months post-birth. Moreover, an F-test indicates the drops across leave policies differ from each other (p-value(diff)=0.014). Put differently, we find evidence that the degree of decline in performance is statistically different based on the length of maternity leave mothers received. Next, we test whether the size of the effects vary among women who had leaves longer than 6 weeks (i.e., 6 + 12 flexible weeks vs. 18 weeks vs. 12 weeks) and find meaningful statistically significant differences (displayed in the row labeled "p(diff), >6 weeks of leave)."

Results at 12 months after birth show that mothers surprised with a retroactive 12 weeks of leave after they had given birth (those who fall under "6 weeks + 12 flex") are the only mothers for whom physical performance recovers by the child's first birthday. Notably, these mothers were eligible to use the additional 12 weeks of leave anytime before the child's first birthday. Bacolod et al. (2020) show mothers used this leave as flexible time off during the first year after child birth. It is possible for these women that flexible time off assisted with performance recovery. For all other women, physical performance is still below pre-pregnancy levels at 12 months post-birth, regardless of the length of maternity leave received. Differences in magnitude of the effects at 12 months post-birth are not statistically significant.

By the child's 2nd birthday, at 24 months post-birth, mothers continue to struggle on employer-assessed physical fitness. Women who received the shortest time off – 6 weeks – recover closest to their pre-pregnancy levels by 24 months after giving birth, performing 0.07 standard deviations below pre-pregnancy levels. By contrast, women who received longer leaves continue to display poorer physical performance than before they became pregnant on the order of 0.17 to 0.43 standard deviations lower. Notably, physical performance declines once again among mothers who received 12 weeks of flexible paid time off to take once back at work. During the child's second year of life, these mothers no longer can take advantage of flexible time off from work. The difference in the magnitude of performance declines at 24 months after birth is statistically significant across leave policies (p(diff)=0.001), and is marginally statistically different between the policies longer than 6 weeks (p(diff), >6 weeks of leave=0.075).

Due to concerns that mothers may opt into (or out of) pregnancy once they learn of a new extended (or reduced) paid leave policy, we re-estimate our main models. We split mothers who received 18 weeks of leave into two groups: (a) those who were already pregnant or on leave at the time of the policy change and thus received additional weeks of leave unexpectedly, and (b) those who became pregnant after the policy announcement and thus expected to receive the longer leave policy. We also present descriptive characteristics for mothers who fall under each policy grouping in Table A4. Mothers appear largely similar across categories. Statistically significant differences emerge in terms of age and AFQT scores (a measure of intelligence), but differences appear practically small.

We re-estimate the impact of leave length on mothers' physical performance after childbirth, splitting mothers who received 18 weeks of maternity leave into groups who did or did not expect the 18 weeks. Table A5 presents results. For mothers who were already pregnant at the time of the policy change ("18 weeks unexp."), we find physical performance drops 0.60 standard deviations below their pre-pregnancy levels by 8 months post birth. For those women who could anticipate their true policy of 18 weeks ("18 weeks exp."), their drop in physical performance at 8 months post-birth is 0.36 standard deviations. All of the 8-month post birth drops statistically differ from zero. Moreover, an F-test indicates the magnitude of the drop between mothers who received 6 vs. 6 + 12 flexible weeks of leave and between mothers who received 18 weeks of leave unexpectedly vs. expectedly also differs. This pattern generally holds across time points after the birth. If anything, results suggest women who unexpectedly receive 18 weeks of paid leave fare worse than women who are aware at the start of their pregnancy that they will receive 18 weeks of leave. If selection into policy periods drives our results, it underestimates the negative impacts of leave length on mothers' physical ability to perform when back on the job.

5.3.2 Women's Career Advancement

Last, we explore whether maternity leave length exacerbates delays in mothers' promotions trajectories after having a child. We rely on the placebo birth matching strategy that compares mothers to similar non-mothers, broken out by the maternity leave policy at the time of birth. Figure 5 displays the results, which show gaps in promotion trajectories emerge between mothers and non-mothers across all maternity leave policies, but these gaps are largest when women receive longer leave. Specifically, mothers who were awarded an unexpected 12 weeks of flexible time off once they had returned to work display the greatest slow down in career advancement. The next largest gap in promotion occurs under the 18-week leave policy, followed by the 12-week leave policy. Put differently, the longer a mother was awarded for maternity leave, the slower she advanced in her career during the two years after having the child. These results largely mirror our findings on leave length and women's physical performance declines.

6 Summary and Conclusions

Using repeated, direct measures of work performance for service members in the U.S. Marine Corps and an event-study approach based on the precise month of birth, we find both men and women's job performance responds to the transition to parenthood. However, women experience large and persistent declines in job performance, while Marine men experience only short-lived declines in physical ability that fade by their child's first birth. Women suffer small declines in supervisor-rated job performance in the years after having a child as well, while men do not.

Documented changes in job performance concentrated among women are consistent with results that women's promotion trajectories slow while men's do not. Relying on a machine learning strategy to assign non-parents to placebo births, we find delays in career advancement for mothers after having a child as compared to non-mothers that occur largely in the child's second year of life. Promotion trajectories between fathers and non-fathers look almost identical over the two years following a first child.

Last, and perhaps surprisingly, we show that longer paid maternity leave drives persistent declines in women's health-related physical performance and delays women's career advancement. Overall, this evidence suggests longer periods away from work may have the unintended effect of eroding job-specific skills and career progression for women. Results do not appear to be driven by strategic fertility choices in response to extended leave policies. When we focus on women who were pregnant when the DoD extended leave from 6 to 18 weeks, we uncover a similar pattern of results.

Our findings provide a new angle on the longstanding literature in economics that shows parenthood reduces mothers' employment, hours worked and wages, but has no effect on fathers. The immediate impacts of having a child on job performance – especially in our sample of Marines who largely return to their jobs after childbirth – point to the need for policymakers and firms to increase support for recent parents. However, findings also highlight potential unintended consequences of such efforts. In our setting, additional parental leave exacerbates gender disparities in job performance and career advancement across the transition to parenthood. Exploring whether alternate family support policies, such as increased access to affordable child care, have fewer unintended consequences is an important area for future research.

References

Agüero, J. M. and Marks, M. S. (2011). Motherhood and Female Labor Supply in the Developing World: Evidence from Infertility Shocks. *The Journal of Human Resources*, 46(4):800–826.

Angelov, N., Johansson, P., and Lindahl, E. (2016). Parenthood and the Gender Gap in Pay. *Journal of Labor Economics*, 34(3):545–579.

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

Antecol, H., Bedard, K., and Stearns, J. (2018). Equal but inequitable: Who benefits from gender-neutral tenure clock stopping policies? *American Economic Review*, 108(9):2420–41.

Azmat, G. and Ferrer, R. (2017). Gender gaps in performance: Evidence from young lawyers. *Journal of Political Economy*, 125(5):1306–1355.

Bacolod, M., Heissel, J. A., and Sullivan, R. (2020). Mothers in the Military: The Effect of Department of Defense Maternity Policy on Leave Uptake and Substitution. Manuscript submitted for publication.

Bailey, M., Byker, T., Patel, E., and Ramnath, S. (2019). The Long-Term Effects of California's 2004 Paid Family Leave Act on Women's Careers: Evidence from U.S. Tax Data. Technical Report w26416, National Bureau of Economic Research, Cambridge, MA.

Balser, C., Hall, C., and Bukowinski, A. (2020). The Effect of Expanding Paid Maternity Leave on Maternal Health: Evidence From the United States Air Force & Army. SSRN Scholarly Paper ID 3601018, Social Science Research Network, Rochester, NY.

Barth, E., Kerr, S. P., and Olivetti, C. (2017). The Dynamics of Gender Earnings Differentials: Evidence from Establishment Data. Technical Report w23381, National Bureau of Economic Research, Cambridge, MA.

Baum, C. L. and Ruhm, C. J. (2016). The effects of paid family leave in california on labor market outcomes. *Journal of Policy Analysis and Management*, 35(2):333–356.

Bertrand, M., Goldin, C., and Katz, L. F. (2010). Dynamics of the Gender Gap for Young Professionals in the Financial and Corporate Sectors. *American Economic Journal: Applied Economics*, 2(3):228–255.

Bronars, S. G. and Grogger, J. (1994). The Economic Consequences of Unwed Motherhood: Using Twin Births as a Natural Experiment. *The American Economic Review*, 84(5):1141–1156.

Bullinger, L. R. (2019). The effect of paid family leave on infant and parental health in the united states. *Journal of health economics*, 66:101–116.

Bureau of Labor Statistics (2017a). Physical strength required for jobs in different occupations in 2016. Bureau of Labor Statistics (BLS) and United States Department of Labor.

Bureau of Labor Statistics (2017b). Standing or walking versus sitting on the job in 2016. Bureau of Labor Statistics (BLS) and United States Department of Labor.

Butikofer, A., Riise, J., and Skira, M. (2018). The impact of paid maternity leave on maternal health. *NHH Dept. of Economics Discussion Paper*, (04).

Byker, T. S. (2016). Paid Parental Leave Laws in the United States: Does Short-Duration Leave Affect Women's Labor-Force Attachment? *American Economic Review*, 106(5):242–246.

Cools, S., Markussen, S., and Strøm, M. (2017). Children and Careers: How Family Size Affects Parents' Labor Market Outcomes in the Long Run. *Demography; Silver Spring*, 54(5):1773–1793.

Cruces, G. and Galiani, S. (2007). Fertility and female labor supply in Latin America: New causal evidence. *Labour Economics*, 14(3):565–573.

Cunha, J. M., Shen, Y.-C., and Burke, Z. R. (2018). Contrasting the impacts of combat and humanitarian assistance/disaster relief missions on the mental health of military service members. *Defence and Peace Economics*, 29(1):62–77.

Cáceres-Delpiano, J. (2006). The Impacts of Family Size on Investment in Child Quality. *The Journal of Human Resources*, 41(4):738–754.

Department of Defense (2018). 2018 demographics report: Profile of the military community". Technical report, Department of Defense (DoD), Office of the Deputy Assistant Secretary of Defense for Military Community and Family Policy (ODASD (MCFP)), under contract with ICF (https://www.icf.com/work/human-capital.

Donovan, S. A. (2019). Paid Family Leave in the United States. Report 44835, Congressional Research Services, Washington, DC.

Gallen, Y. (2018). Motherhood and the gender productivity gap. *Becker Friedman Institute for Research in Economics Working Paper*, (2018-41).

Hamilton, B. E., Martin, J. A., Osterman, M. J., and Rossen, L. M. (2019). Births: Provisional Data for 2018. Vital Statistics Rapid Release 007, U.S. Department of Health and Human Services, Washington, DC.

Jacobsen, J. P., Pearce, J. W., and Rosenbloom, J. L. (1999). The Effects of Childbearing on Married Women's Labor Supply and Earnings: Using Twin Births as a Natural Experiment. *The Journal of Human Resources*, 34(3):449–474.

Kim, S. and Moser, P. (2020). Women in science: Lessons from the baby boom. National Bureau of Economic Research Summer Institute (NBER SI) Session on Gender in the Economy.

Kleven, H., Landais, C., Posch, J., Steinhauer, A., and Zweimüller, J. (2019a). Child Penalties Across Countries: Evidence and Explanations. NBER Working Paper w25524, National Bureau of Economic Research, Cambridge, MA.

Kleven, H., Landais, C., and Søgaard, J. E. (2019b). Children and Gender Inequality: Evidence from Denmark. *American Economic Journal: Applied Economics*, 11(4):181–209.

Melchiorre, K., Sharma, R., Khalil, A., and Thilaganathan, B. (2016). Maternal cardiovascular function in normal pregnancy: evidence of maladaptation to chronic volume overload. *Hypertension*, 67(4):754–762.

Memon, H. U. and Handa, V. L. (2013). Vaginal childbirth and pelvic floor disorders. *Women's health*, 9(3):265–277.

O'Hara, M. W. and Swain, A. M. (1996). Rates and risk of postpartum depression—a meta-analysis. *International Review of Psychiatry*, 8(1):37–54.

Pac, J. E., Bartel, A. P., Ruhm, C. J., and Waldfogel, J. (2019). Paid family leave and breastfeeding: Evidence from california. Technical report, National Bureau of Economic Research. NBER Working Paper.

Rossin-Slater, M., Ruhm, C. J., and Waldfogel, J. (2013). The effects of California's paid family leave program on mothers' leave-taking and subsequent labor market outcomes. *Journal of Policy Analysis and Management*, 32(2):224–245.

Saxbe, D., Rossin-Slater, M., and Goldenberg, D. (2018). The transition to parenthood as a critical window for adult health. *American Psychologist*, 73(9):1190–1200.

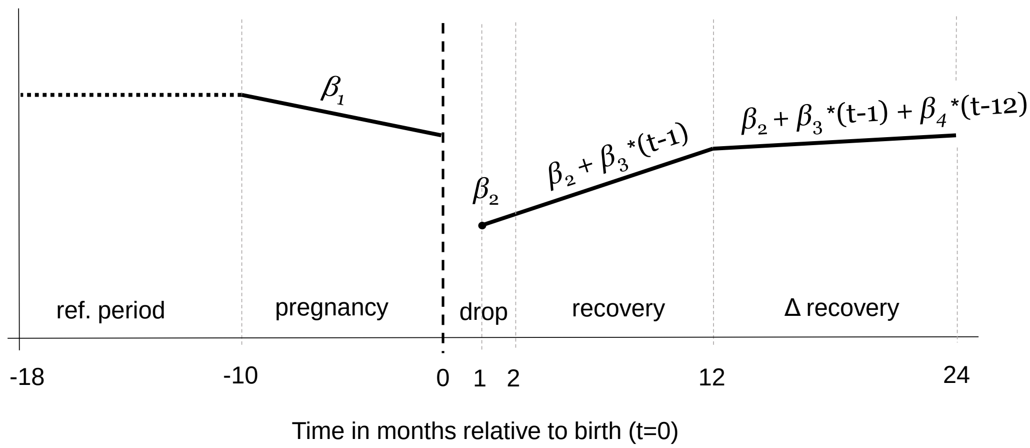
Timpe, B. (2019). The long-run effects of america's first paid maternity leave policy. Working Paper.

Trajkovski, S. et al. (2019). California paid family leave and parental time use. Technical report, Center for Policy Research, Maxwell School, Syracuse University.

7 Figures

Figure 1: Stylized Representation of the Semi-Dynamic Specification, Eq. 2:

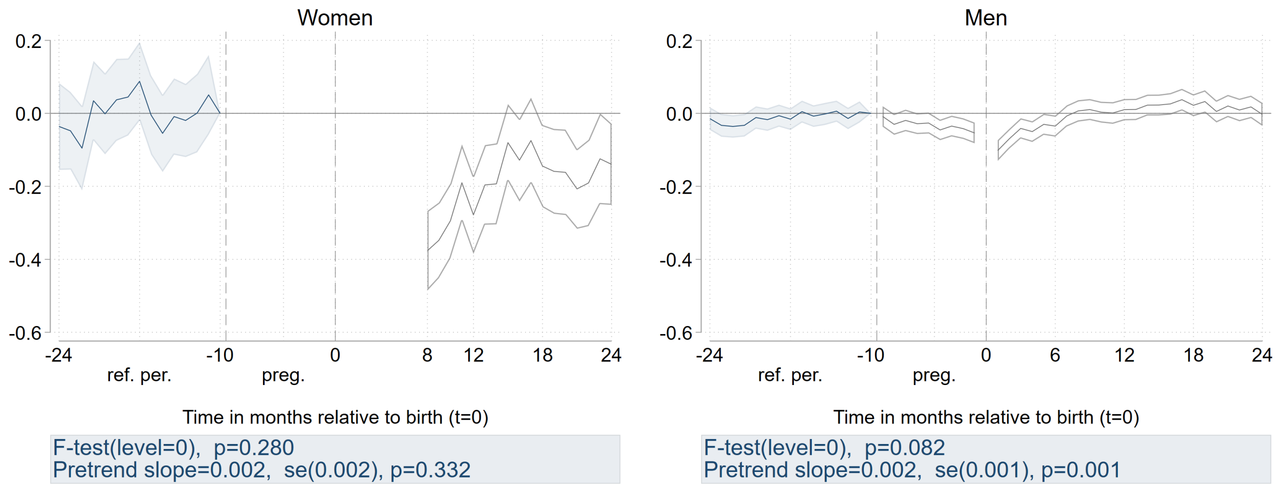
$$Y_{it} = \alpha_i + \phi_t + X_{it}\theta + M_{it}^{pregnancy}\beta_1 + M_{it}^{drop}\beta_2 + M_{it}^{recovery}\beta_3 + M_{it}^{\Delta recovery}\beta_4 + \varepsilon_{it}$$



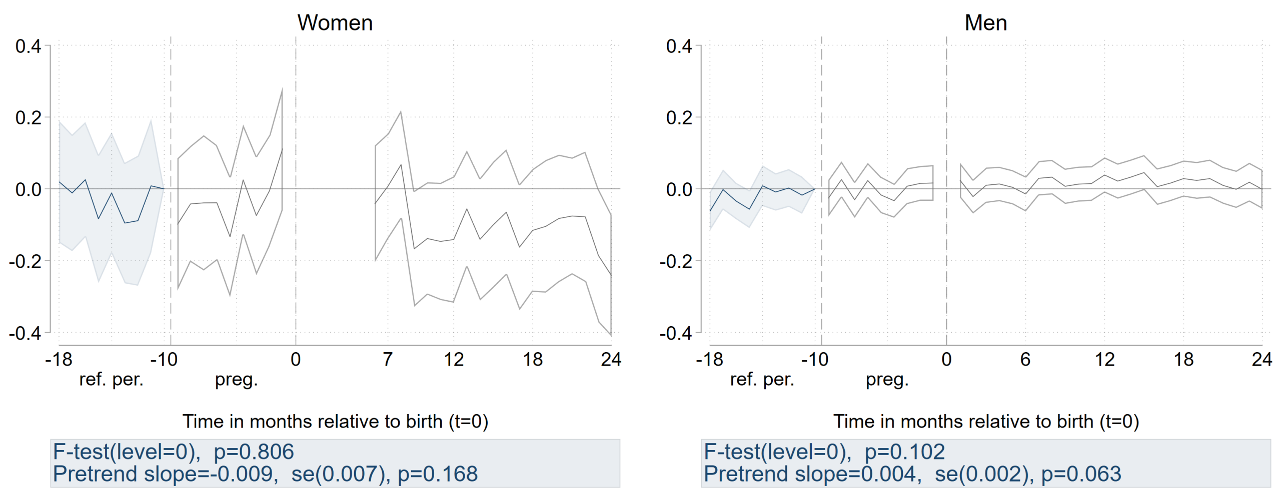
Notes: Figure displays a diagram of parameters defined in Eq. 2 where the post-birth drop (β_2) is estimated in the first month following pregnancy. This holds for all models among the sample of men. Among the sample of women, we begin measuring the post-birth drop (β_2) for physical fitness performance at 8 months and supervisor ratings at 6 months after the birth. We also cannot estimate β_1, the pregnancy trend, in the model on women's physical fitness outcomes, due to restrictions on when pregnant women's fitness is assessed.

Figure 2: Event-Study Estimates of the Impact of Birth on Job Performance

(a) Physical Performance Scores

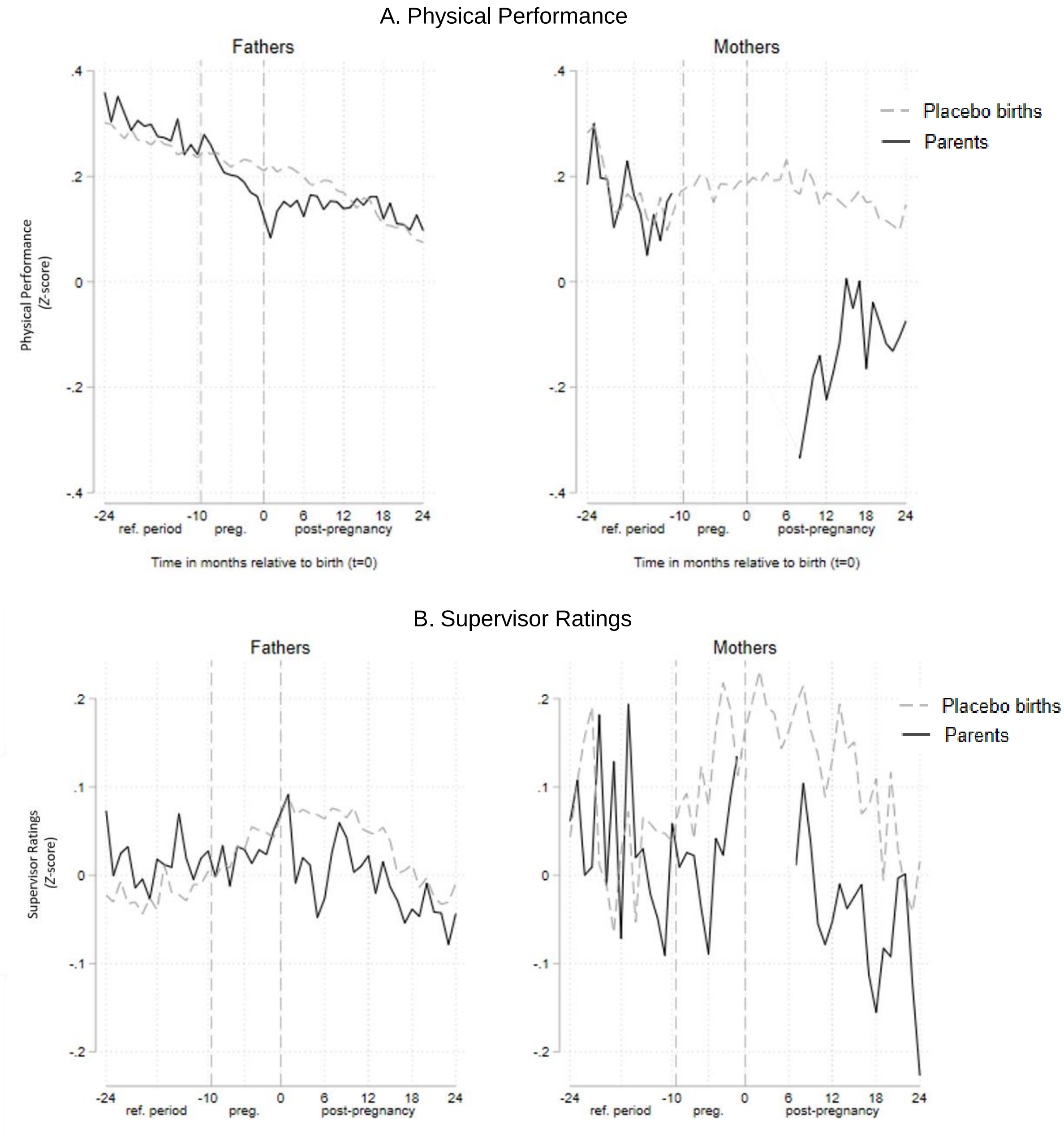


(b) Supervisor Ratings



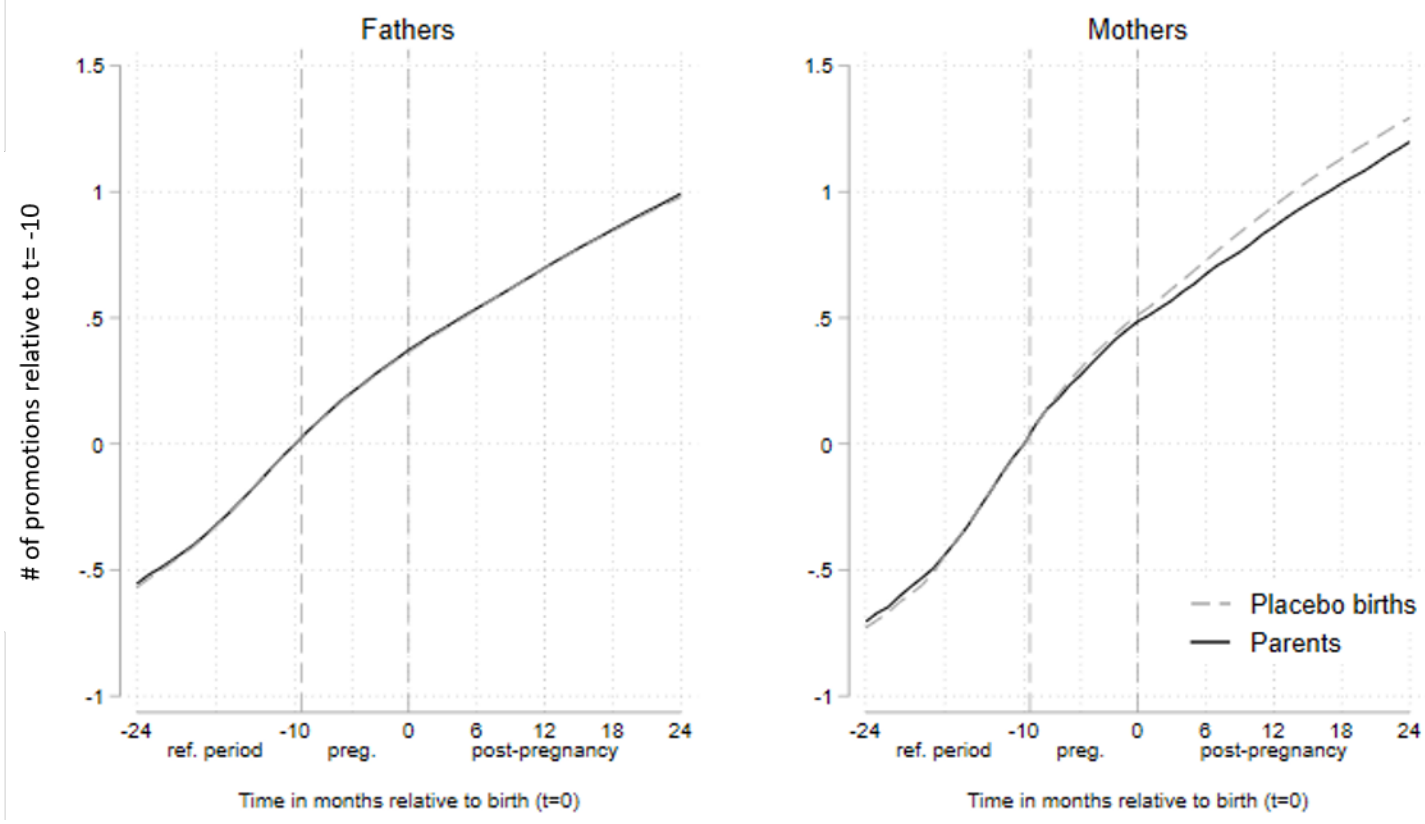
Notes: Displays coefficients from event-study regressions. Outcomes include (1) standardized scores (mean=0, SD=1) from physical/combat fitness tests and (2) standardized (mean=0, SD=1) scores from supervisor-rated job performance evaluations, measured separately for males and females. Sample includes first-time parents who remain in the sample at least 12 months before and 24 months after birth, as well as a control group of same-gender Marines who do not have a birth during the study window and remain in service 3+ years. Regressions include fixed effects for individuals, month by year, and age. The reference month is 10 months before the birth (t=-10), capturing the time point before the start of pregnancy. Vertical dotted lines reflect the start of the pregnancy (t=-9.5) and the birth (t=0). Standard errors are clustered at the individual level and included as shaded/hollowed area representing a 95% confidence interval.

Figure 3: Placebo Birth DiD Estimates of the Impact of Birth on Job Performance



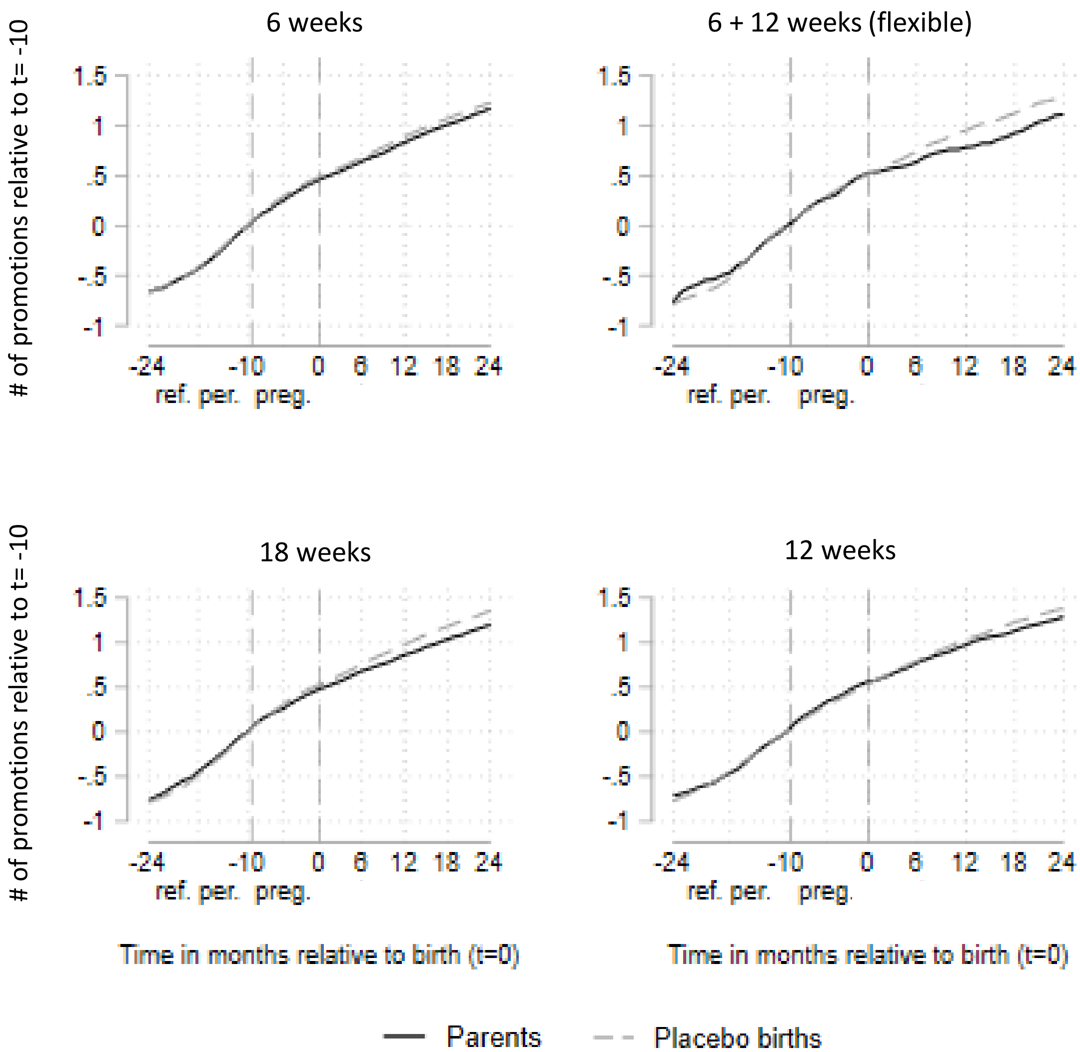
Notes: Displays differences in outcomes between first-time parents and non-parents across birth/placebo birth events for each month pre- and post-birth. Outcomes include (1) standardized scores (mean=0, SD=1) from supervisor-rated job performance evaluations conducted 1-2x per year and (2) standardized (mean=0, SD=1) scores from physical/combat fitness tests conducted 2x per year. Non-parents assigned to placebo births are limited to those whose rank and number of months in service is an exact match with parents' 10 months before the birth. Among those with an exact match on rank and months in service, each parent's outcomes are compared to the five non-parents most similar to parents in their propensity to have a child based on age, race/ethnicity, military entrance exam scores (AFQT scores), marital status (including whether a spouse is also in the military), level of education, occupational field, and physical performance scores. The reference month is 10 months before the birth (t=-10), capturing the time point before the start of pregnancy. Vertical dotted lines reflect the start of the pregnancy (t=-9.5) and the birth (t=0).

Figure 4: Placebo Birth DiD Estimates of the Impact of Birth on Promotion Trajectories



Notes: Displays differences in the number of promotions relative to t=-10 between first-time parents and non-parents across birth/placebo birth events for each month pre- and post-birth. Non-parents assigned to placebo births are limited to those whose rank and number of months in service is an exact match with parents' 10 months before the birth. Among those with an exact match on rank and months in service, each parent's outcomes are compared to the five non-parents most similar to parents in their propensity to have a child based on age, race/ethnicity, military entrance exam scores (AFQT scores), marital status (including whether a spouse is also in the military), level of education, occupational field, and physical performance scores. The reference month is 10 months before the birth (t=-10), capturing the time point before the start of pregnancy. Vertical dotted lines reflect the start of the pregnancy (t=-9.5) and the birth (t=0).

Figure 5: Placebo Birth DiD Estimates of Women's Promotion Trajectories by Maternity Leave Length



Notes: Displays promotion trajectories among women, split into subgroups based on the length of paid maternity leave a mother could receive based on the date she gave birth. Maternity leave policies include: 6 weeks of leave total, 6 weeks of leave plus 12 weeks of flexible paid time off for women who returned to work to be taken during the child's first year, 18 weeks of leave total, and 12 weeks of leave total. See notes on Figure 4 for additional details.

8 Tables

Table 1: Characteristics of First-Time Parents

	Fathers		**Mothers**	
	Marines	Civilians	Marines	Civilians
Age (mean)	25.5	31.6	23.8	29.9
Education				
Some College/Associates	5%	27%	5%	28%
Bachelor's Degree	16%	30%	12%	34%
Marital Status				
Married	86%	83%	71%	78%
Race				
Black (Non-Hispanic)	9%	5%	14%	6%
Hispanic	14%	13%	21%	11%
Job Classification				
Mngmt./business/science/arts	10%	45%	15%	58%
Service	4%	11%	7%	15%
Sales/office	12%	15%	34%	24%
Construction/maint.	29%	15%	18%	0%
Production/moving/transpo.	14%	14%	19%	3%
Military Specific Chars.				
Officer	14%	–	9%	–
AFQT score (percentile)	63.3	–	59.1	–
GCT score (m=100; sd=20)	111.5	–	104.1	–
N	25,598	59,423	2,555	49,013

Notes: Displays characteristics of first-time parents in the Marine Corps in our sample alongside characteristics of first-time civilian parents in the labor market. Time-varying characteristics of Marines in our sample (e.g., age) are measured 10 months before the birth (t=-10). Data on civilians come from the American Community Survey 1-year estimates, 2010 to 2019. We limit the civilian sample to adults who are employed in the civilian labor market and have a first child under age 1. Job categories correspond to Standard Occupational Classification (SOC) System groups applied to U.S. Marine Corps job codes and available in the American Community Service. Military specific variables include whether a Marine is ranked as an officer (akin to manager) and AFQT and GCT scores, which are measures of intelligence. We do not observe these military-specific variables in the civilian sample.

Table 2: Impacts of Childbirth on Job Performance Among First-Time Parents

	Women		Men	
	Physical Performance (1)	Supervisor Ratings (2)	Physical Performance (3)	Supervisor Ratings (4)
Model Parameters				
Pregnancy trend	–	0.005	-0.004***	0.003*
		(0.005)	(0.001)	(0.001)
Post-birth drop $(birth - 24mos.)$	-0.390***	0.034	-0.066***	0.016
	(0.031)	(0.041)	(0.006)	(0.011)
Recovery trend $(birth - 24mos.)$	0.053***	-0.018*	0.009***	0.003*
	(0.010)	(0.008)	(0.001)	(0.001)
Δ Recovery trend $(13 - 24mos.)$	-0.050***	0.016	-0.010***	-0.004*
	(0.012)	(0.011)	(0.001)	(0.002)
Estimated Effects				
12-month effect	-0.177***	-0.072	0.036***	0.046***
p-value	[0.000]	[0.069]	[0.000]	[0.000]
24-month effect	-0.142***	-0.087*	0.022**	0.026*
p-value	[0.000]	[0.049]	[0.001]	[0.040]
Pre-pregnancy mean	0.140	-0.015	0.296	0.011
N of Individuals	20,471	9,854	270,636	112,968
Observations	154,148	83,385	2,063,665	986,862
R^2	0.589	0.416	0.597	0.428

Notes: Displays coefficients from Eq.2, the semi-dynamic event-study specification. Outcomes include (1) standardized scores (mean=0, SD=1) from supervisor-rated job performance evaluations conducted 1-2x per year and (2) standardized (mean=0, SD=1) scores from physical/combat fitness tests conducted 2x per year. Women's supervisors ratings are not measured 0 to 5 months post-birth due to overlap with maternity leave. Women's physical performance scores are not measured 9 months before through 7 months after birth when women are not required to take fitness tests. Regressions include individual, month by year, and age fixed effects. The parameter "Pregnancy trend" captures trends during pregnancy, if observed. "Post-birth drop" is an indicator equal to 1 after the birth, starting in t=1 for all men's outcomes; and t=6 for women's supervisor ratings and m=8 for women's physical performance. "Recovery trend" estimates monthly changes in the outcome for the entire post-birth period. "Δ Recovery trend" estimates any change in the slope in the second year post-birth. Robust standard errors are clustered by individual, shown in parentheses. p-values that test whether 12-month and 24-month average effects differ from zero are shown in brackets. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 3: Impact of Childbirth on Women's Physical Performance, by Maternity Leave Policy

	8 months post		12 months post		24 months post		
	Effect size	*p*	Effect size	*p*	Effect size	*p*	N
Main effect:	-0.387***	0.000	-0.171***	0.000	-0.135***	0.000	153,910
Effects by paid leave length:							
6 weeks	-0.320***	0.000	-0.148***	0.000	-0.062	0.074	153,910
6 weeks + 12 flex	-0.754***	0.000	-0.034	0.732	-0.434***	0.000	
18 weeks	-0.483***	0.000	-0.209***	0.000	-0.163**	0.004	
12 weeks	-0.358***	0.000	-0.234***	0.000	-0.270***	0.000	
F-test of effects:							
p(diff), all effects	0.014		0.287		0.002		
p(diff), >6 weeks of leave	0.070		0.194		0.100		

Notes: Displays the expected value of the outcome across three time points after childbirth, by maternity leave length among women. Estimates comes from Eq.2, the semi-dynamic event-study specification, interacted with indicators for the length of paid maternity leave a mother could receive based on the date she gave birth. "p(diff), all effects" presents the p-value for an F-test on differences in the estimates across all policy periods. "p(diff), $| > 6$ weeks of leave" presents the p-value for an F-test on differences in the estimates among policy periods with greater than 6 weeks of leave. Regressions exclude women with births in November and December 2016 (including from estimates of the main effect presented here) because women with births in these months could have fallen into one of two policy periods based on their doctor-estimated date of conception, which we do not observe. See Table 2 for additional details on model specification. Robust standard errors are clustered at the individual level in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

9 Appendix Tables

Table A1: Sample Characteristics of First-Time Parents Across Samples

	Mothers			Fathers		
	All	Obs. -12/+12	Obs. -12/+24	All	Obs. -12/+12	Obs. -12/+24
Age (mean)	23.31	23.34	23.80	24.75	25.25	25.54
Education						
Some college	0.05	0.05	0.05	0.05	0.05	0.05
College	0.09	0.09	0.12	0.13	0.15	0.16
Marital Status						
Married	0.69	0.70	0.71	0.84	0.86	0.86
Race						
African American	0.14	0.14	0.14	0.09	0.09	0.09
Hispanic	0.22	0.22	0.22	0.15	0.15	0.14
Job Classification						
Mngmt./Business/Science/Arts	0.12	0.13	0.15	0.08	0.09	0.10
Service	0.07	0.07	0.07	0.04	0.04	0.04
Sales/Office	0.35	0.34	0.33	0.12	0.12	0.11
Construction/Maint.	0.19	0.19	0.18	0.30	0.30	0.29
Production/Moving/Transpo.	0.20	0.20	0.19	0.13	0.14	0.14
Military	0.07	0.07	0.08	0.32	0.31	0.31
Military Specific Chars.						
Officer	0.06	0.07	0.09	0.11	0.13	0.14
AFQT score (percentile)	58.29	58.47	59.48	61.98	62.73	63.30
GCT score (m=100; sd=20)	103.13	103.32	104.22	110.36	110.93	111.45
Observations	5,887	3,955	2,483	50,888	33,962	25,365

Notes: Displays characteristics of first-time parents across samples. ”Obs. -/+12” restricts to Marines who remain in the sample 12 months before and 12 after their first birth, while ”Obs. =-12/+24” restricts to Marines who remain in the sample for 12 months before and 24 months after the first birth. Time-varying characteristics (e.g., age) are measured 10 months before the birth (-10). Job categories correspond to Standard Occupational Classification (SOC) System groups applied to U.S. Marine Corps job codes. All sample averages are percentages with the exception of age and AFQT and GCT scores. AFQT and GCT scores are measures of intelligence, with scoring scales described.

Table A2: Impact of Childbirth Across Sample Specifications: Women

First births, restricted to	–	-12/+12	-12/+12	-12/+24	-12/+24	-12/+24 contracted
No births, restricted to	–	–	⩾24 m.o.s.	–	⩾36 m.o.s.	⩾36 m.o.s.
	(1)	(2)	(3)	(4)	(5)	(6)
			A. Physical Performance			
Model Parameters						
Pregnancy trend	–	–	–	–	–	–
Post-birth drop $(birth - 24mos.)$	-0.383***	-0.392***	-0.393***	-0.386***	-0.390***	-0.360***
	(0.025)	(0.027)	(0.027)	(0.031)	(0.031)	(0.056)
Recovery trend $(birth - 24mos.)$	0.044***	0.049***	0.048***	0.053***	0.053***	0.045**
	(0.008)	(0.009)	(0.009)	(0.010)	(0.010)	(0.017)
Δ Recovery trend $(13 - 24mos.)$	-0.041***	-0.045***	-0.045***	-0.050***	-0.050***	-0.036
	(0.010)	(0.010)	(0.010)	(0.012)	(0.012)	(0.021)
Estimated Effects						
12-month effect	-0.205***	-0.197***	-0.199***	-0.172***	-0.177***	-0.179***
p-value	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
24-month effect	-0.157***	-0.153***	-0.156***	-0.136***	-0.142***	-0.063
p-value	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.152]
Pre-pregnancy mean	-0.001	0.033	0.033	0.140	0.140	0.246
N of Individuals	30,880	29,251	25,492	27,726	20,471	18,626
Observations	188,733	184,492	174,466	179,297	154,148	146,292
R^2	0.602	0.600	0.595	0.598	0.589	0.587
			B. Supervisor Ratings			
Model Parameters						
Pregnancy trend	0.008**	0.012**	0.011**	0.005	0.005	0.008
	(0.003)	(0.003)	(0.003)	(0.005)	(0.005)	(0.009)
Post-birth drop $(birth - 24mos.)$	0.071*	0.105**	0.100**	0.050	0.034	-0.006
	(0.029)	(0.030)	(0.030)	(0.041)	(0.041)	(0.083)
Recovery trend $(birth - 24mos.)$	-0.025***	-0.030***	-0.030***	-0.017*	-0.018*	0.022
	(0.006)	(0.006)	(0.006)	(0.008)	(0.008)	(0.015)

(Continued on next page)

Table A2: Impact of Childbirth Across Sample Specifications: Women *(Continued)*

First births, restricted to	–	-12/+12	-12/+12	-12/+24	-12/+24	-12/+24 contracted
No births, restricted to	–	–	⩾24 m.o.s.	–	⩾36 m.o.s.	⩾36 m.o.s.
	(1)	(2)	(3)	(4)	(5)	(6)
Δ Recovery trend $(13-24mos.)$	0.025**	0.030**	0.030**	0.015	0.016	-0.030
	(0.009)	(0.009)	(0.009)	(0.011)	(0.011)	(0.020)
Estimated Effects						
12-month effect	-0.081*	-0.076*	-0.082*	-0.053	-0.072	0.128
p-value	[0.011]	[0.022]	[0.014]	[0.177]	[0.069]	[0.070]
24-month effect	-0.088*	-0.074	-0.079	-0.085	-0.087*	0.031
p-value	[0.027]	[0.074]	[0.054]	[0.052]	[0.049]	[0.697]
Pre-pregnancy mean	-0.060	-0.084	-0.084	-0.015	-0.015	-0.123
N of Individuals	17,105	16,306	13,226	15,547	9,854	8,889
Observations	112,958	110,155	101,595	106,171	83,385	78,422
R^2	0.442	0.438	0.427	0.436	0.416	0.410

Notes: Displays coefficients from the semi-dynamic specification in Eq. 2 and the average effect at 12 months and 24 months, with p-values below in brackets, for various sample specifications. Samples vary in terms of the restrictions we place on parents with first births and same-gender Marines without a birth during the study window. The descriptor "-/+12" restricts to Marines who remain in the sample 12 months before and 12 after their first birth; "-12/+24" restricts to Marines who remain in the sample for 12 months before and 24 months after the first birth; and "-12/+24 contracted" restricted to Marines observed at least 12 months before the birth and who remained in the Marines for 24 months after as required by their contract. "⩾ 24 m.o.s." and "⩾ 36 m.o.s." restricts to Marines who do not experience a birth and have at least 24 or 36 months of service (m.o.s.) in the Marines respectively. Robust standard errors clustered by ID in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table A3: Impact of Childbirth Across Sample Specifications: Men

First births, restricted to	–	-12/+12	-12/+12	-12/+24	-12/+24	-12/+24 contracted
No births, restricted to	–	–	⩾24 m.o.s.	–	⩾36 m.o.s.	⩾36 m.o.s.
	(1)	(2)	(3)	(4)	(5)	(6)
A. Physical Performance Scores						
Model Parameters						
Pregnancy trend	-0.009***	-0.006***	-0.007***	-0.004***	-0.005***	-0.003*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Post-birth drop $(birth - 24mos.)$	-0.133***	-0.088***	-0.089***	-0.064***	-0.066***	-0.067***
	(0.005)	(0.005)	(0.005)	(0.006)	(0.006)	(0.009)
Recovery trend $(birth - 24mos.)$	0.009***	0.007***	0.007***	0.009***	0.009***	0.013***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Δ Recovery trend $(13 - 24mos.)$	-0.008***	-0.006***	-0.006***	-0.011***	-0.011***	-0.012***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Estimated Effects						
12-month effect	-0.033***	-0.016**	-0.016**	0.040***	0.037***	0.079***
p-value	[0.000]	[0.001]	[0.001]	[0.000]	[0.000]	[0.000]
24-month effect	-0.019**	-0.007	-0.007	0.023***	0.021**	0.099***
p-value	[0.001]	[0.270]	[0.261]	[0.000]	[0.001]	[0.000]
Pre-pregnancy mean	0.203	0.246	0.246	0.296	0.296	0.312
N of Individuals	367,101	350,427	312,941	342,031	271,545	255,692
Observations	2,443,274	2,373,724	2,272,459	2,325,044	2,073,770	1,964,569
R^2	0.604	0.603	0.599	0.602	0.597	0.595
B. Supervisor Job Performance Ratings						
Model Parameters						
Pregnancy trend	0.000	0.003**	0.003*	0.003*	0.003*	-0.005
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Post-birth drop $(birth - 24mos.)$	-0.021*	0.017	0.015	0.023*	0.016	-0.022
	(0.009)	(0.010)	(0.010)	(0.011)	(0.011)	(0.019)
Recovery trend $(birth - 24mos.)$	0.001	-0.001	-0.001	0.003*	0.003*	0.004*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)

(Continued on next page)

Table A3: Impact of Childbirth Across Sample Specifications: Men *(Continued)*

First births, restricted to	–	-12/+12	-12/+12	-12/+24	-12/+24	-12/+24 contracted
No births, restricted to	–	–	⩾24 m.o.s.	–	⩾36 m.o.s.	⩾36 m.o.s.
	(1)	(2)	(3)	(4)	(5)	(6)
Δ Recovery trend $(13 - 24mos.)$	0.000	0.001	0.001	-0.006**	-0.004*	-0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)
Estimated Effects						
12-month effect	-0.011	0.006	0.005	0.054***	0.046***	0.019
p-value	[0.250]	[0.579]	[0.620]	[0.000]	[0.000]	[0.304]
24-month effect	0.003	0.002	0.002	0.023	0.026*	0.050*
p-value	[0.812]	[0.895]	[0.853]	[0.078]	[0.045]	[0.012]
Pre-pregnancy mean	0.034	0.016	0.016	0.011	0.011	-0.014
N of Individuals	179,600	172,794	144,330	168,717	113,746	104,128
Observations	1,273,739	1,246,255	1,160,662	1,224,492	994,038	943,167
R^2	0.439	0.436	0.432	0.435	0.428	0.423

Notes: Displays coefficients from the semi-dynamic specification in Eq. 2 and the average effect at 12 months and 24 months, with p-values below in brackets, for various sample specifications. Samples vary in terms of the restrictions we place on parents with first births and same-gender Marines without a birth during the study window. The descriptor "-/+12" restricts to Marines who remain in the sample 12 months before and 12 after their first birth; "-12/+24" restricts to Marines who remain in the sample for 12 months before and 24 months after the first birth; and "-12/+24 contracted" restricted to Marines observed at least 12 months before the birth and who remained in the Marines for 24 months after as required by their contract. "$\geqslant$ 24 m.o.s." and "$\geqslant$ 36 m.o.s." restricts to Marines who do not experience a birth and have at least 24 or 36 months of service (m.o.s.) in the Marines respectively. Robust standard errors clustered by ID in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table A4: Sample Characteristics of First-Time Mothers Across Policy Periods

	Length of Paid Maternity Leave					
	6 weeks	6 + 12 weeks flex	18 weeks (unexp.)	18 weeks (exp.)	12 weeks	*p*(diff)
Age (mean)	23.67	24.01	24.03	24.61	23.62	0.08
Education						
Some College	0.04	0.05	0.06	0.07	0.05	0.45
College	0.11	0.13	0.13	0.16	0.12	0.26
Marital Status						
Married	0.70	0.71	0.75	0.72	0.68	0.31
Race						
Black	0.15	0.11	0.12	0.12	0.18	0.12
Hispanic	0.19	0.20	0.23	0.25	0.30	
Job Classification						
Mngmt./business/science/arts	0.14	0.17	0.14	0.18	0.16	0.60
Service	0.06	0.06	0.07	0.10	0.07	0.26
Sales/office	0.35	0.27	0.29	0.32	0.31	0.11
Construction/maint.	0.17	0.20	0.22	0.20	0.18	0.34
Production/moving/transpo.	0.19	0.21	0.18	0.15	0.23	0.21
Military	0.09	0.08	0.10	0.06	0.07	0.36
Military Specific Chars.						
Officer	0.08	0.10	0.12	0.11	0.09	0.24
AFQT score (percentile)	60.13	60.41	57.79	60.07	57.51	0.04
GCT score (m=100; sd=20)	104.42	105.37	103.90	104.97	102.91	0.12
N	1,471	121	294	177	365	2,483

Notes: Displays characteristics of first-time mothers in our preferred sample by maternity leave length. Women who received ”18 weeks (unexp.)” are those who were pregnant or already on leave at the time of the policy change from 6 to 18 weeks of paid leave. Women who receive ”18 weeks (exp.)” are those gave birth far enough after the policy change occurred that they would have known they would receive 18 weeks of leave at the time of conception. p(diff) displays the p-value of an F-test for differences across groups. See Table 1 for notes on variables included.

Table A5: Impact of Childbirth on Women's Physical Performance, by Maternity Leave Policy

	8 months post		12 months post		24 months post		
	Effect size	*p*	Effect size	*p*	Effect size	*p*	N
Main effect:	-0.387***	0.000	-0.171***	0.000	-0.135***	0.000	153,910
Effects by paid leave length:							
6 weeks	-0.320***	0.000	-0.148***	0.000	-0.063	0.074	153,910
6 + 12 flex	-0.754***	0.000	-0.034	0.725	-0.433***	0.000	
18 weeks (unexp.)	-0.601***	0.000	-0.192**	0.001	-0.094	0.202	
18 weeks (exp.)	-0.311**	0.002	-0.227**	0.001	-0.280**	0.001	
12 weeks	-0.358***	0.000	-0.235***	0.000	-0.270***	0.000	
F-test of effects:							
p(diff), all effects	0.008		0.341		0.001		
p(diff), 6 + 12 flex vs. 18 unexp.	0.001		0.384		0.000		
p(diff), 18 unexp. vs. 18 exp.	0.000		0.260		0.038		

Notes: Displays the expected value of the outcome across three time points after childbirth, by maternity leave length among women. Estimates comes from Eq.2, the semi-dynamic event-study specification, interacted with indicators for a mother's length of leave based on the date she gave birth. Birth drop corresponds to 8 months after birth. "p(diff), all effects" presents the p-value for an F-test on differences in the estimates across all policy periods. "p(diff), 6 + 12 flex vs. 18 unexp.) presents the p-value for an F-test on differences in outcomes between women with 6 weeks of leave + 12 flexible weeks' time off as compared to women who received 18 continuous paid weeks of leave unexpectedly. "p(diff), 18 unexp. vs. 18 exp." tests the difference between the effect of the 18-week policy period when women received longer leave unexpectedly compared to the effect when they received the same leave length but knew about the policy before becoming pregnant. We exclude women with births in November and December 2016 (including estimates of the main effect presented here) because women with births in these months could have fallen into one of two policy periods based on estimated date of conception, which we do not observe. See Table 2 for additional details on model specification. Robust standard errors are clustered at the individual level in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.