Title
Three essays on maternity leave policies, utilization and consequences

Author
Goodman, Julia Marie

Publication Date
2015

Peer reviewed|Thesis/dissertation

Three essays on maternity leave policies, utilization and consequences

By

Julia Marie Goodman

A dissertation submitted in partial satisfaction of the requirements for the degree of

Doctor of Philosophy

in

Health Services and Policy Analysis

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Ralph A. Catalano, Chair

Professor Sylvia R. Guendelman

Professor William H. Dow

Professor Malcolm Potts

Summer 2015

Abstract

Three essays on maternity leave policies, utilization and consequences

by

Julia Marie Goodman

Doctor of Philosophy in Health Services and Policy Analysis

University of California, Berkeley

Professor Ralph A. Catalano, Chair

This dissertation examines maternity leave policies, utilization and consequences in three separate, but related, papers. In the Introduction, I describe the ways state-level leave policies influence, but do not perfectly predict, utilization and then discuss the heterogeneity of women's experiences during pregnancy and post-partum, and how that heterogeneity informs interpretation of results in this field.

In Chapter 2 ("Laboring until labor: the prevalence and correlates of antenatal maternity leave in the United States"), I use a national survey of English-speaking women to examine maternity leave taken in the final weeks of pregnancy. I describe individual-, employer-, and policy-level correlates of antenatal leave (ANL), focusing in particular on variation in state temporary disability insurance (TDI) laws. I find that two-thirds of women in this sample stopped working more than a week before their due date, and that state policies significantly influenced leave-taking, even after controlling for characteristics of women and their jobs. While certain individual-level characteristics of women's work were important, employer policies were not significantly associated with ANL use or duration. The literature does not yet include a national analysis of antenatal leave and its correlates. This paper fills that gap and sets up the following chapter on the consequences of taking ANL.

In Chapter 3 ("Antenatal maternity leave and childbirth using the First Baby Study: a propensity score analysis"), I use survey data from a prospective cohort in Pennsylvania to test whether women who take maternity leave at the end of pregnancy have better labor and delivery outcomes. In this sample of women giving birth for the first time, fully half of the sample did not stop working before delivery. Using propensity score matching, I find that women who did stop working at least two days prior to delivery experienced more negative delivery outcomes, including an increased likelihood of unplanned Cesarean section. This paper highlights the strong selection into leave-taking, particularly in a context of limited leave availability.

In Chapter 4 ("The impact of California's Paid Family Leave law on maternal time use"), I shift focus to postnatal maternity leave and use the American Time Use Survey, a nationally-representative dataset collected by the U.S. Census Bureau, to examine whether the country's first paid family leave (PFL) law increased the amount of time exposed women spent on childcare and decreased their time in paid work. Using a difference-in-

difference-in-difference approach with variation in state, time, and age of youngest child, I find that after PFL, women in California significantly increased the time they spent with children in their care and slightly reduced their time spent working. Exploiting a natural experiment, I am able to avoid some of the selection concerns present in the previous chapter, but the daily diary nature of the data do not illuminate whether time use changed due to leave-taking, schedule changes, or some other factor.

The results of these studies will inform future research on maternity leave and maternal health, and guide policymaking with regards to targeting and promoting maternity leave policies.

Table of Contents

Acknowledgements

Thank you to my parents, Dorothy and James Goodman, for their unwavering support and words of encouragement through this long educational journey, and for taking such great care of their grandson while his mommy wrote this dissertation.

Thank you to my research group—Deborah Karasek, April Falconi, and Alison Gemmill—for listening to me present countless versions of these papers, for compassionately providing critical feedback, and for their inspirational brilliance.

Thank you to my advisor, Ray Catalano, the best mentor I could imagine, for instilling in me a love of scholarship; for giving me insight into the world of academia; for encouraging my writing and instructing me in E-Prime; and for his thoughtful guidance and support on every major academic decision I've faced.

Thank you to Sylvia Guendelman for encouraging my interest in maternity leave and maternal health, and for developing so much of this literature to which I hope to contribute.

Thank you to Will Dow for shaping me into a more critical and careful scholar.

Thank you to Malcolm Potts for guiding me through the process of writing a dissertation and encouraging my daily writing practice.

Thank you to Kristen Kjerulff for generously sharing her First Baby Study data and for her enthusiastic support for my project.

Thank you to my smiling son, Owen, who grows and changes so much every day, for reminding me of the importance of family leave.

Finally, thank you to Matt Easterlin for being the rock at the center of our household, for listening patiently as I considered dissertation topics, for giving me brilliant feedback on five years of drafts and presentations, for encouraging me to work when I was feeling productive and rest when I was not, and for caring for our family so deeply. I could not have finished this without you.

Chapter 1
Introduction

Family, or maternity, leave policies enable workers to take time from work in order to care for themselves or for family members, making a career and caregiving more compatible. Societies develop leave policies for many reasons, including but not limited to promoting maternal and child health and keeping women close to the labor market. Importantly, these policies expand women's choices around work and caregiving.

This dissertation explores the relationship between maternity leave and women's experiences during and after pregnancy from three perspectives, each addressing the role of policy. In this Introduction, I will describe the ways state-level leave policies influence, but do not perfectly predict, utilization. I will then discuss the heterogeneity of women's experiences during pregnancy and post-partum, and how that heterogeneity informs interpretation of results in this field. Finally, I will describe my dissertation research.

The existence of state and federal leave laws does not directly translate into leave utilization. Rather, leave laws influence the availability of leave and type of leave offered at the firm level, as well as eligibility for leave. In turn, employer and individual characteristics influence whether a covered woman uses available leave. Figure 1 illustrates this multilevel framework that whittles down the population of women "exposed" to a law by virtue of residing in a covered state to the population of women who actually take leave around childbirth.

As I describe in Chapter 2, pregnant and post-partum women in the United States face a web of local, state and federal laws that address job protection and income replacement during periods of leave. This includes the federal Family and Medical Leave Act of 1993 (FMLA), which provides up to 12 weeks of unpaid, job-protected leave for one's own illness, the illness of a qualified family member, or to care for a newborn or newly adopted baby for eligible employees. Five states (California, Hawaii, Rhode Island, New Jersey, and New York) have temporary disability insurance (TDI) laws providing most private-sector workers in those states with paid short-term disability insurance. Recently, three of the five states with TDI laws (California, New Jersey, and Rhode Island) expanded their programs to include leave for caregiving and/or bonding with a new child, as well as increasing leave for one's own disability. Paid Family Leave (PFL) laws focus on bonding and caregiving, rather than disability, so they cover only leave taken after childbirth. Other states and municipalities have also passed paid leave laws. In 2007, Washington passed a paid leave law, but failed to identify a funding mechanism. The program has been indefinitely put on hold. As of March 2015, Ohio, Virginia, and Illinois provide paid leave to state employees and Washington, D.C.; St. Paul and Brooklyn Park, Minnesota; St. Petersburg, Florida; San Francisco, California; Chicago, Illinois; and Austin, Texas provide paid leave to municipal employees (U.S. Dept. of Labor, 2015).

Eligibility and coverage restrictions limit the number of workers and firms with leave. Most notably, FMLA only covers employees at firms with more than 50 employees within a 75-mile radius and who have worked at least 1,250 hours in the past year. This results in coverage of just over half (59%) of all U.S. workers (Klerman, Daley, & Pozniak,

2012). TDI and PFL laws cover far more employees in covered states, usually anyone who paid into the state's disability insurance program through baseline earnings. Though more prevalent in the five states with state-level laws, TDI plans cover some employees in other states if their employers offer TDI (the Pregnancy Discrimination Act of 1978 mandates that TDI plans must include pregnancy as a covered condition). Nationally, 75% of employers offer temporary disability insurance to their employees, more commonly among large than small employers (90% vs. 69%) (Matos & Galinsky, 2014).

Even among covered firms, failure to comply with relevant laws could limit leave availability. At the same time, employers may provide leave benefits, either formally or informally, that go beyond legal requirements. The 2014 National Study of Employers reports that among worksites that must legally comply with FMLA (i.e., more than 50 employees within a 75-mile radius), 28% provided more than 12 weeks of maternity leave (on average, 13.8 weeks). Overall, 58% offered some replacement pay to women on leave, with three-quarters of firms funding this through temporary disability insurance (Matos & Galinsky, 2014). Finally, women must know about available leave in order to take advantage of it, but evidence suggests that lack of public awareness, even among those who had experienced a life event that leave laws were designed to cover, prevents many women from using leave (Appelbaum & Milkman, 2011).

Still, the availability and awareness of leave does not perfectly predict leave-taking behavior. As this dissertation suggests, many women forgo available leave during pregnancy and after childbirth. One important characteristic of leave is whether or not women receive pay. For example, FMLA does not require employers to offer paid leave, and evidence suggests that many eligible employees either do not take leave or return to work sooner than desired because they can not afford unpaid leave (Klerman, Daley, & Pozniak, 2012).

Characteristics of the worker and workplace influence utilization of leave, conditional on leave availability. On the employer side, a workplace culture that supports leave-taking might make women feel more comfortable taking advantage of available leave. In the absence of such a supportive culture, a woman may not feel comfortable using all available leave for fear of being perceived as less committed to her job, even if she has legal protection for the right to return to her job. Other formal and informal workplace policies like the availability of reduced work hours, the option to work from home, or modified tasks for pregnant workers could influence leave-taking. On the one hand, women with flexible, accommodating jobs may be less likely to take leave because their needs are met without fully stepping out of the job (particularly for antenatal leave). On the other hand, these related policies might help women extend limited or unpaid leave, as when employers offer the option of returning to work part-time or working flexibly from home.

Finally, workers differ in their need for and preferences toward leave-taking. Women with healthy pregnancies and comfortable jobs they enjoy may choose to work until delivery, even if this means they forgo disability insurance they have contributed to and are eligible for. Women with less healthy pregnancies, with post-partum complications and/or a sick infant may use all available leave (or quit, if leave is insufficient). In addition, household finances determining the possibility of taking unpaid or partially paid leave and job commitment should impact leave-taking behavior.

Studying leave policies and their impact on maternal health and well-being contributes to both policy and research discussions because of the potential protective

effect of leave-taking. They may promote maternal and child health, keep women close to the labor market, and expand women's choices around work and caregiving. However, these policies mean very different things to different women, potentially masking their impact. When we as researchers have incomplete information about, for example, women's health or job stress or career orientation, we can't fully understand why some women take leave and others do not. If, as I suggest in Chapter 3, only relatively unhealthy women take leave, this results in the impression that leave-taking causes negative outcomes when, in fact, these women were likely already on a trajectory toward negative outcomes.

We may also expect effect modification, where leave has a differential impact depending on women's context. For example, we might expect that a woman without workplace accommodations like modified tasks and who spends her workday standing would benefit from leave-taking enough to have a measureable impact on perinatal health, while a professional woman who works from home may experience no gain from leave. This makes detecting what is likely a small effect even more difficult.

Understanding the multiple levels that separate public policies from utilization provides context for the results in each of my dissertation papers. The overarching aims of my dissertation research were as follows: 1) to map the prevalence and correlates of maternity leave, particularly during pregnancy; 2) to examine the impact of maternity leave on maternal health and well-being; and 3) to examine the impact of a specific maternity leave policy change— the implementation of California's Paid Family Leave law.

My dissertation research examines maternity leave policies, utilization and consequences in three papers. In Chapter 2 ("Laboring until labor: the prevalence and correlates of antenatal maternity leave in the United States"), I use a national survey of English-speaking women to examine maternity leave taken in the final weeks of pregnancy. I describe individual-, employer-, and policy-level correlates of antenatal leave (ANL), focusing in particular on variation in state TDI laws. I find that two-thirds of women in this sample stopped working more than a week before their due date, and that state policies significantly influenced leave-taking, even after controlling for characteristics of women and their jobs. While certain individual-level characteristics of women's work were important, employer policies were not significantly associated with ANL use or duration. The literature does not yet include a national analysis of antenatal leave and its correlates. This paper fills that gap and sets up the following chapter on the consequences of taking ANL.

In Chapter 3 ("Antenatal maternity leave and childbirth using the First Baby Study: a propensity score analysis"), I use survey data from a prospective cohort in Pennsylvania to test whether women who take maternity leave at the end of pregnancy have better labor and delivery outcomes. In this sample of women giving birth for the first time, fully half of the sample did not stop working before delivery. Using propensity score matching, I find that women who did stop working at least two days prior to delivery experienced more negative delivery outcomes, including an increased likelihood of unplanned Cesarean section. This paper highlights the strong selection into leave-taking, particularly in a context of limited leave availability.

In Chapter 4 ("The impact of California's Paid Family Leave law on maternal time use"), I shift focus to postnatal maternity leave and use the American Time Use Survey, a nationally-representative dataset collected by the U.S. Census Bureau, to examine whether the country's first paid family leave law increased the amount of time exposed women

spent on childcare and decreased their time in paid work. Using a difference-in-difference-in-difference approach with state, time, and age of youngest child variation, I find that after PFL, women in California significantly increased the time they spent with children in their care and slightly reduced their time spent working. Exploiting a natural experiment, I am able to avoid some of the selection concerns present in the previous chapter, but the daily diary nature of the data do not illuminate whether time use changed due to leave-taking, schedule changes, or some other factor.

The results of these studies will inform future research on maternity leave and maternal health, and guide policymaking with regards to targeting and promoting maternity leave policies.

References

Appelbaum, E., & Milkman, R. (2011). Leave that pay: employer and worker experiences with paid family leave in California: Center for Economic and Policy Research.

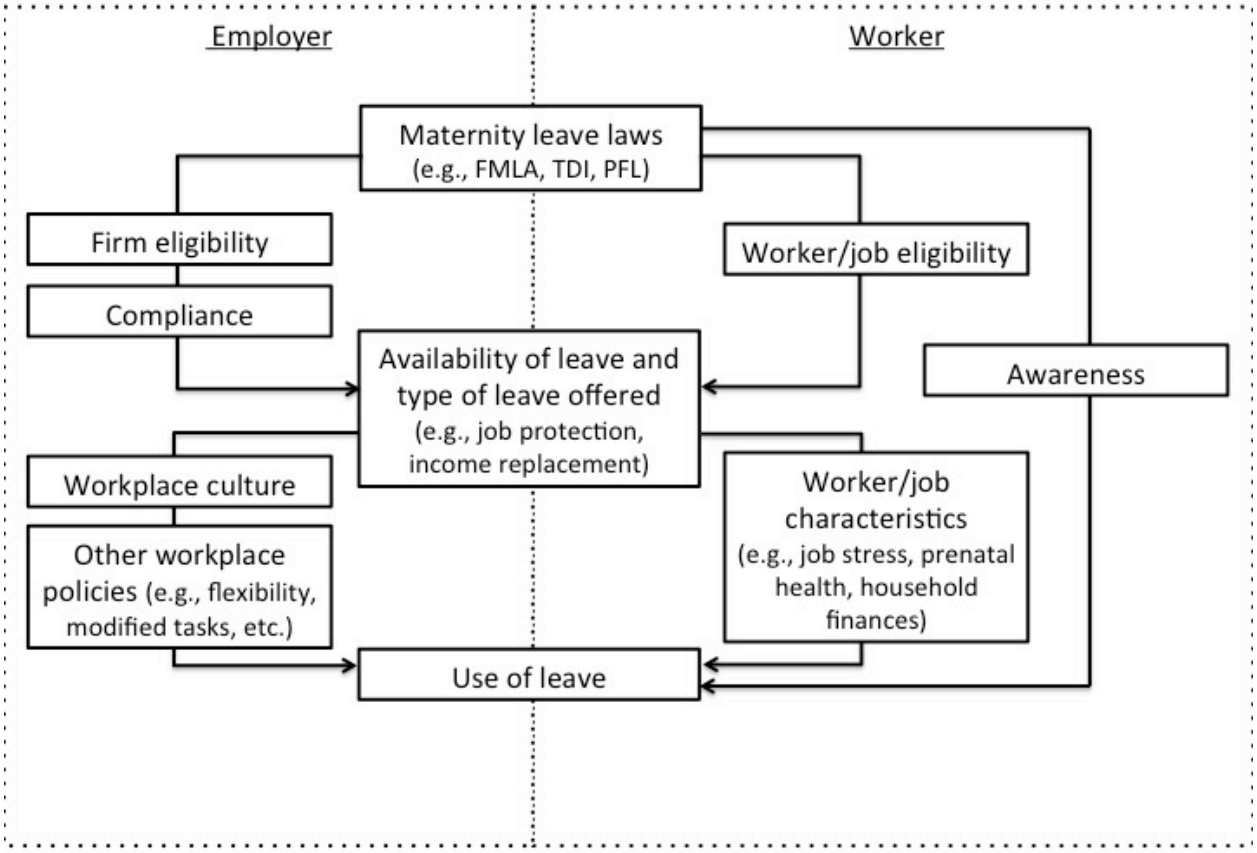
Klerman, J. A., Daley, K., & Pozniak, A. (2012). Family and medical leave in 2012: Technical report. Cambridge, MA: Abt Associates Inc.

U.S. Department of Labor (2015). DOL Factsheet: Paid Family and Medical Leave.

Matos, K., & Galinsky, E. (2014). National study of employers. New York: Families and Work Institute.

Figures

Figure 1: Multi-level framework of the pathways to maternity leave usage.



Abbreviations: Family and Medical Leave Act of 1993 (FMLA), Temporary Disability Insurance (TDI), Paid Family Leave (PFL)

Chapter 2
Laboring until labor: the prevalence and correlates of antenatal maternity leave in the United States

Abstract

A majority of women in the United States is employed during pregnancy, and increasingly work until the month they give birth. Yet, published research does not include a national analysis of the prevalence of maternity leave taken during pregnancy, or antenatal leave (ANL), nor the correlates of such leave. This study uses two waves of Listening to Mothers, a national sample of English-speaking women, to examine whether state Temporary Disability Insurance (TDI) and Paid Family Leave (PFL) laws, as well as employer policies and individual characteristics, increase the likelihood that women will stop working at least one week before their due date. Sixty-three percent of employed women in this sample stopped working at least one week before their due date. State policies significantly influenced ANL, even after controlling for characteristics of women and their jobs: living in a state with only a TDI law increased the probability of ANL by .14 ($p<.01$) and living in a state with both TDI and PFL laws increased the probability of ANL by .23 ($p<.01$) relative to women living in states with no leave laws. Women who were employed full-time or self-employed were significantly less likely to take ANL, after controlling for covariates. Employer policies were not significantly associated with ANL use or duration. This paper sheds light on an understudied policy area that has potential impacts on maternal labor market attachment and perinatal health.

Introduction

Women's labor force participation during pregnancy has changed dramatically over the past half century. American women are increasingly likely to be employed during pregnancy and, among those who are employed, to work later into their pregnancy. Two-thirds of first-time mothers worked during pregnancy in 2006-2008, up from 44% in 1961-1965, with a majority of those now working full-time (Laughlin, 2011). Among women who worked while pregnant, 82% worked until the last month of pregnancy (Laughlin, 2011).

Prior research suggests that working throughout pregnancy may have an impact on the health of both mothers and their children. Certain characteristics of the work environment have been linked to preterm delivery (PTD), small-for-gestational-age (SGA), and hypertension or preeclampsia (Mozurkewich, Luke, Avni et al., 2000). Individual-level epidemiology studies examining whether antenatal maternity leave (i.e., leave from work during pregnancy) has an impact on maternal and child health suggest that women who take antenatal leave are less likely to experience a Cesarean or otherwise difficult delivery (Guendelman, Pearl, Graham et al., 2009; Xu, Séguin, & Goulet, 2002). Two papers in the economics literature showing that expansions in maternity leave policies increase birthweight and reduce preterm delivery and infant mortality suggest that part of the observed result derives from leave taken during pregnancy, though neither examines antenatal leave directly (Rossin, 2011; Tanaka, 2005).

Despite the increase in prenatal employment and possible impact on health, the literature contains little description of the prevalence or correlates of antenatal leave taking. U.S. women who were employed during pregnancy use a variety of leave arrangements before and after birth. According to 2006-2008 data from the U.S. Census Bureau's Survey of Income and Program Participation (SIPP), 21% of women who worked during pregnancy used paid leave before birth (maternity leave, sick or vacation leave); 19% used unpaid leave before birth; 3% used disability leave before birth; and 19% quit or were fired during pregnancy (Laughlin, 2011). This distribution reflects the wide range of maternity leave arrangements American women make, piecing together federal, state, and employer-level policies in the context of their individual work environments.

American women lack many of the pregnancy accommodations and maternity leave allowances that women in other advanced economies enjoy. The U.S. does not have a federal paid leave program; instead, a web of local, state and federal laws addresses job protection and income replacement for pregnant (and postpartum) workers. The Family and Medical Leave Act of 1993 (FMLA) enables many women in the U.S. to take unpaid leave during (and after) pregnancy, without risking their jobs. FMLA provides up to 12 weeks of unpaid, job-protected leave for one's own illness, the illness of a qualified family member, or to care for a newborn or newly adopted baby. Only employees who work for firms with more than 50 employees within a 75-mile radius and who have worked at least 1,250 hours in the past year are eligible, covering just over half (59%) of all U.S. workers (Klerman, Daley, & Pozniak, 2012). FMLA covers even fewer new mothers, as this group is less likely to be employed in the year before birth, meet the hours requirement and work for a covered firm (Ruhm, 1997). Importantly, FMLA does not require employers to offer paid leave, and evidence suggests that many eligible employees either do not take leave or

return to work sooner than desired because they can not afford unpaid leave (Klerman, Daley, & Pozniak, 2012).

Though FMLA provides unpaid leave, employees taking leave during pregnancy may qualify for income replacement through short-term, or temporary, disability insurance programs. The Pregnancy Discrimination Act of 1978 mandated that employers who offer short-term disability insurance to their employees must include pregnancy-related disability as a covered condition, enabling pregnant women to receive income replacement if they are temporarily unable to work. Five states (California, Hawaii, Rhode Island, New Jersey, and New York) have temporary disability insurance (TDI) laws at the state level and most private-sector workers in those states have access to paid short-term disability insurance. With the exception of Hawaii, public-sector workers are not covered but states may cover these workers through their own private plan, or by opting into the state TDI plan (Stearns, 2015). The eligibility and benefits of these laws vary slightly by state, but in each case, workers who paid into the state's disability insurance program through baseline earnings can draw partial income replacement (55-66% of prior weekly wages, up to a maximum) for the duration of their leave. For a normal pregnancy, a worker in California or New Jersey can receive benefits for four to six weeks before her due date and six weeks after childbirth (eight for a Cesarean delivery). Workers in Hawaii, New York, and Rhode Island have six to eight weeks of leave available for use on either side of childbirth (Stearns, 2015). Though more prevalent in those five states, TDI plans cover some employees in other states. Nationally, 75% of employers offer temporary disability insurance to their employees, more commonly among large than small employers (90% vs. 69%) (Matos & Galinsky, 2014).

Recently, three of the five states with TDI laws have expanded their programs to include leave for caregiving and/or bonding with a new child, as well as increasing leave for one's own disability, resulting in longer post-partum leave availability. In 2004, California became the first U.S. state to institute a paid family leave (PFL) law, followed in 2009 by New Jersey and in 2014 by Rhode Island. In each case, PFL income replacement rates follow those for the states' TDI program (i.e., 55-66% of prior weekly wages, up to a maximum). Because PFL laws focus on bonding and caregiving, rather than disability, they cover only leave taken after childbirth.

This patchwork of leave policies has made it difficult to quantify the actual availability of maternity leave benefits. One thing that is clear is that the availability of paid maternity leave is not random in the population, but rather varies along sociodemographic lines. A recent nationally-representative survey of women delivering in U.S. hospitals in 2005 found that women who were non-Hispanic black, working full-time or from high income families were more likely to report receiving paid maternity leave, for a longer duration, and at higher levels of wage replacement than other women after controlling for age, education, health insurance status, and partner status (Shepherd-Banigan & Bell, 2014).

The presence of leave policies – whether at the employer or public policy level – does not directly translate into *use* of maternity leave. Figure 1 illustrates factors at the employer and worker levels that influence the decision to take ANL. One disconnect between policies, availability of leave and use of leave is lack of awareness. For example, evidence shows that California's landmark Paid Family Leave law suffers from limited public awareness. In a survey of employed Californians who had experienced a life event

that the program was designed to cover, fewer than half knew the program existed, making lack of awareness the primary reason respondents did not use PFL (Appelbaum & Milkman, 2011). Another issue is that not all workplaces and jobs are covered by all of the existing leave-related policies, potentially impeding leave taking in unanticipated ways. A woman could be eligible for income replacement (e.g., through TDI) while on leave, but her job and benefits are not protected (e.g., through FMLA, which has many eligibility restrictions). In this case, she may not use available paid leave out of concern for losing her job or benefits. More commonly, women are eligible for job-protected leave through FMLA, but do not live in a state with an income replacement scheme. Women who cannot afford to forgo income may take less than their available leave, if any. A recent study of the relationship between FMLA and infant health found a significant positive impact on those whose mothers were married and college-educated, but found no such relationship for the children of mothers who were unmarried and who had less than a college education (Rossin, 2011). The author suggests that this was because unmarried women with lower educational attainment were less likely to be able to afford unpaid time off of work.

Even when legally entitled to both pay and job protection, a woman may not feel comfortable using all available leave for fear of being perceived as less committed to her job. This would result from a workplace culture that does not promote or support leave-taking. Antenatal leave may be particularly subject to this type of concern because its use and availability appear to be relatively less common than postnatal leave. It may be perceived as less necessary, since there is not yet an infant to care for and women, especially those experiencing relatively healthy pregnancies, may be or appear able to conduct their jobs without modification. This pressure to forgo available leave does not necessarily come from employers. Women may want to wrap up existing work projects before taking leave or to maximize earnings before a period without pay. Women may also want to preserve limited available leave for use after the baby's birth (Frazier, Golbeck, & Lipscomb, 2001). A study of Census data suggests that women use paid and unpaid leave more often after delivery than during pregnancy, which the author suggests is due to women wanting to work as long as possible into their pregnancy in order to have more leave available after their child is born (Laughlin, 2011). In some cases, this may mean women are leaving money on the table: state TDI programs in California and New Jersey have separate allowances for antenatal and postnatal leave, meaning that women who do not take advantage of the standard 4-6 weeks before delivery can not apply this to their postnatal leave allowance, thus choosing to forgo earned benefits.

Within these policy constraints, characteristics of individual women and their jobs influence the likelihood of taking ANL. One study of women employed at least 20 hours a week during pregnancy in California found that working the night shift, lack of fulfillment in one's job, having at least a high school education, having children under age 5, being stressed and/or tired were positively associated with taking ANL (Guendelman, Pearl, Graham et al., 2006). Women without work flexibility and with short work tenure were less likely to take ANL in that study.

Women's health during pregnancy appears to predict ANL. A study of employed pregnant women in Georgia found that more than a quarter (27.7%) were advised by a healthcare provider to stop working during pregnancy; this was most likely among women who had been hospitalized during pregnancy or had previously delivered preterm (Frazier,

Golbeck, & Lipscomb, 2001). Being advised to stop working significantly decreased the likelihood that women continued working into their ninth month of pregnancy.

Published research does not include a national analysis of ANL use and its correlates. Studies of individual-level correlates of ANL have been done within states (i.e., California and Georgia) and do not allow examination of state-level policy variation. Furthermore, existing studies examine ANL use, but not the duration of leave taken. Understanding these issues is an important first step toward examining the importance of this social benefit and its effect on maternal labor force participation and perinatal outcomes.

Using a national sample of English-speaking women giving birth in U.S. hospitals at two time periods (2005 and 2011-2012), this study aims to provide a detailed description of a) the prevalence and duration of antenatal maternity leave and b) the correlates of antenatal maternity leave. I test the hypothesis that state policy (TDI only or TDI and PFL) contributes to increased use of ANL and the duration of ANL.

Methods

Data

Listening to Mothers (LTM), a national survey of English-speaking women who gave birth in U.S. hospitals in 2005 (wave II) and 2011-2012 (wave III), includes detailed questions about women's experiences during pregnancy, labor and delivery, and postpartum. Each wave includes a baseline survey conducted 1.5-17 months postpartum and a follow-up survey conducted 7-20.5 months postpartum. In wave III, employment and maternity leave questions were only asked in the follow-up survey. The combined sample of wave II and wave III follow-up consists of 2,645 women. Eligible women had been employed during pregnancy, were between 18 and 45 years of age, were able to complete the survey in English, and had given birth to a singleton child who was alive at the time of interview. After excluding women who reported not being employed during pregnancy (N=938) and women missing antenatal leave data (N=2), the analytic sample includes 1,705 women.

Harris Interactive, on behalf of Childbirth Connection, conducted the surveys. The interviews took place between January 20 and February 21, 2006 (wave II) or January 29 and April 15, 2013 (wave III postpartum) and were conducted in English.
Potential online respondents were drawn from the Harris Poll Online panel of more than 6 million active U.S. members who had been recruited from a variety of sources, including offerings made in conjunction with a number of organizations. Online-only surveys risk limited participation by groups of women without internet access. To improve representation of a broader population of black non-Hispanic and Hispanic women in the sample, a telephone over-sample was conducted to supplement the online sample using a list of households with a baby provided by Survey Sampling International. Female interviewers made up to six attempts over a four-week period to complete an interview with each potential participant. Full reports from all waves of the Listening to Mothers survey and questionnaires are available online (http://childbirthconnection.org).

Outcome variable

Antenatal leave refers to maternity leave taken prior to delivery, derived from the following survey question: how many weeks before your due date did you stop working? Women who stopped working less than 1 week before their due date were coded as taking 0 weeks (no leave).

Predictive variables

My primary predictor of leave is whether a woman lived in a state with TDI only (Hawaii, New York, or Rhode Island in both years; New Jersey in 2005) or TDI and PFL laws (California in both years, New Jersey in 2011-2012). I separately examine employer-level policies, which should be influenced by state-level policy (i.e., employers in states with PFL laws are much more likely to offer paid leave to their employees than employers in states without such laws). Employer factors include women's report of whether her company offered a paid maternity leave benefit and the generosity of maternity leave offered (number of weeks of paid leave multiplied by the wage replacement ratio). Of note, employer policies are self-reported and women may not have a complete understanding of their companies' maternity leave policies. The employer-level variables may suffer from measurement error if women who took ANL have more accurate knowledge of relevant policies than women who did not take leave (e.g., women who did not take any leave do not know how much of their salary would have been paid had they taken leave, whereas women who took leave do).

Individual factors include: full-time vs. part-time work; self-employed vs. working for other; parity; pregnancy intention; and maternal sociodemographic characteristics (age, race/ethnicity, marital status, educational attainment, household income, maternity care payment source). To estimate maternal prenatal health, I include use of fertility treatments, pre-pregnancy obesity (BMI > 30 kg/m^2), gestational age at delivery, and the number of ultrasounds she received during pregnancy (more than 3 ultrasounds suggests possible pregnancy complications). The more recent survey wave includes additional questions about women's health: whether she used medication for high blood pressure and/or depression before pregnancy; whether she had a pre-pregnancy diagnosis of diabetes; and whether she was diagnosed with gestational diabetes. I use these additional health conditions to test for interaction with state policy. Finally, I control for survey wave and time to interview (lag between birth and interview).

Analysis

I conducted statistical analyses using Stata/IC version 11.2 (StataCorp, College Station, Texas) and applied survey weights so that results more accurately reflect the target population of English-speaking women delivering singletons in U.S. hospitals in the study years. Demographic variables used for weighting include educational attainment, age, race/ethnicity, geographic region, household income, and time elapsed since last giving birth, as well as a score reflecting the respondents' propensity to be online, using data from the March 2005 and March 2011 Supplements of the U.S. Census Bureau's Current Population Survey and national natality data.

I first examined the frequency of ANL and mean duration among women who took any ANL. I then tested for differences in predictive variables between employed women who took ANL and those who did not using chi-square tests. I also tested for differences in mean duration of ANL using ANOVA among women who took ANL. Finally, I used multivariable linear regression analyses to examine the contribution of state-level policies

(residence in a TDI only state or TDI+PFL state) or employer policies (paid leave offered and generosity of leave offered) on the likelihood of taking any ANL and the duration of ANL, controlling for individual characteristics.

$$ANL_i = \alpha + \beta_1 TDI_i + \beta_2 PFL_i + \beta_3 \bar{X}_i + \beta_4 Wave_i + \beta_5 Lag_i + u_i \quad (1)$$

Equation (1) shows the hypothesized relationship between state-level policies and ANL, conditional on individual characteristics, survey wave, and lag time to interview. TDI_i is a binary indicator for residence in a state with a TDI law, but no PFL law (i.e., Hawaii, New York, or Rhode Island in both years; New Jersey in 2005). PFL_i is a binary indicator for residence in a state with both TDI and PFL laws (California in both years, New Jersey in 2011-2012). $\bar{X}_i$ is a vector of individual level characteristics. $Wave_i$ is 0 for wave II (2005) and 1 for wave III (2011-2012) and Lag_i is a categorical indicator for the number of months postpartum the interview took place.

$$ANL_i = \alpha + \gamma_1 PaidLeaveOffered_i + \gamma_2 > 6FTEweeks_i + \gamma_3 \bar{X}_i + \gamma_4 Wave_i + \gamma_5 Lag_i + u_i \quad (2)$$

Equation (2) shows the hypothesized relationship between employer-level policies and ANL, conditional on individual characteristics, survey wave, and lag time to interview. $PaidLeaveOffered_i$ is a binary indicator for whether a woman reported that her employer offered any paid leave and $> 6FTEweeks_i$ is an indicator for whether women were offered the equivalent of 6 weeks of fully paid leave.

Using only the most recent survey wave, I also test for interaction between state-level policy and maternal health measured by the number of the following conditions she reported: use of medication for high blood pressure before pregnancy; use of medication for depression before pregnancy; pre-pregnancy diagnosis of diabetes; and diagnosis of gestational diabetes.

I conduct several sensitivity analyses focusing on women most likely to have planned ANL, as opposed to women who took leave because of serious pregnancy complications or who stopped working before their due date because of early labor. First, I stratified on whether women delivered before or after their due date. Second, in separate analyses, I excluded women who delivered by planned C-section, women who likely quit or were fired during pregnancy, and women who stopped working more than 4 and 8 weeks before their due date. Finally, to minimize recall bias, I exclude women whose interview took place >18 months postpartum.

Results

Thirty-seven percent of women worked until the week their baby was due (Figure 2). Of the 63% who took ANL, most stopped between one and four weeks before their due date. In bivariate analyses, residence in a TDI only state was not associated with an increased likelihood of taking ANL, but residence in a state with both TDI and PFL was (Table 1). Eighty-three percent of women who lived in TDI+PFL states took ANL, compared to 61% of women who lived in other states (p<.001). Employer-level factors were not associated with ANL in bivariate analyses. At the individual-level, low maternal educational attainment, using Medicaid to pay for maternity care and receiving fertility treatment

significantly increased the likelihood of taking ANL, while being employed full-time or self-employed significantly decreased the likelihood of taking ANL.

Women who stopped working did so an average of 5.52 weeks (S.E. = 0.28 weeks) before their due date (Table 1). State-level policy did not significantly increase the mean duration of ANL. Women who received the equivalent of >6 weeks of fully paid leave on average took about a week shorter ANL ($p = .04$) than women with less generous leave. Being employed full-time decreased mean ANL by half a week ($p = .03$). Women who received >3 ultrasounds during pregnancy took slightly longer ANL (mean difference = .58 weeks; $p = .08$), though the difference did not reach statistical significance. Women took shorter ANL in the more recent wave (births in 2011-2012) than in the earlier wave (births in 2005).

Multivariable results

Table 2 shows the unadjusted and adjusted linear probability coefficients and 95% confidence intervals for ANL use. The first two columns show state-level policy models; the last two columns show employer-level policy models. In the unadjusted state policy model, living in a TDI only or TDI+PFL state increased the probability of taking ANL by .14 or .23, respectively (both $p < .01$). These coefficients did not change substantially when controlling for individual-level covariates and remained statistically significant at the 1% level. In these multivariable state policy models, women who were employed full-time or self-employed were significantly less likely to take ANL. The relationship between household income and ANL uptake appears complicated. Relative to women in the lowest income quintile, women with incomes in the next lowest income group (i.e., 2nd quintile) and women in the highest income group (i.e., 5th quintile) were significantly more likely to take any ANL. The employer-level policy models, both unadjusted and adjusted, indicate that whether an employer offered paid leave and the generosity of paid leave offered did not significantly predict ANL use.

Table 3 shows the unadjusted and adjusted linear regression coefficients and 95% confidence intervals for ANL duration. Again, the first two columns show state-level policy models; the last two columns show employer-level policy models. Residence in a TDI only state was not significantly associated with the duration of ANL taken in either unadjusted or adjusted models. However, women who live in states with both TDI and PFL stopped working one and a half weeks longer before their due date than women in other states ($p<.01$). After controlling for individual-level covariates, this estimate increased from 1.46 to 1.84 additional weeks and remained significant at the 1% level. Full-time work predicted shorter leave duration. Compared with women who delivered at full term (i.e., 39 or 40 completed weeks), women who delivered preterm (i.e., <37 completed weeks) stopped working approximately 3 weeks earlier ($p<.01$). In the employer-level policy models, whether an employer offered paid leave was not related to ANL duration; however, generous maternity leave predicted shorter leave duration, though this lost statistical significance in the full model.

Using only data from 2011-2012 births, I found a significant interaction between state-level policy and maternal health measured by the number of the following conditions reported: use of medication for high blood pressure or depression before pregnancy, pre-pregnancy diagnosis of diabetes, and diagnosis of gestational diabetes on ANL uptake. These data were only available in the most recent survey wave. Figure 3 illustrates the relationship between the likelihood of taking ANL and the number of maternal health

conditions by state policy. In states with no leave policies, experiencing more health conditions increased the likelihood of taking any ANL, while in states with leave policies, the slope was flat (TDI only states) or slightly negative (TDI+PFL states).

Sensitivity analyses suggest a robust relationship between state policy and ANL uptake (Appendix Table A1). Women who lived in a TDI only state exhibited increased likelihood of ANL use by .08-.15 whether they delivered before or after their due dates; when excluding women who delivered by planned C-section (n = 189); whether excluding or restricting to women who likely quit or were fired during pregnancy (n = 509 likely quitters); when excluding women who stopped working more than 4 (n = 331) or 8 (n = 162) weeks before their due date; and when excluding women whose interview took place >18 months postpartum (n = 127). Similarly, women who lived in TDI+PFL states exhibited increased likelihood of ANL use by .22-.27 in all sensitivity analyses.

Sensitivity analyses for the ANL duration models reveal a less clear relationship (Appendix Table A2). The relationship between residence in TDI only states and ANL duration fluctuated in the sensitivity tests, perhaps due to small sample sizes; however, the relationship never reached statistical significance. This reinforces the main ANL duration findings. Sensitivity analyses support the main finding that women in TDI + PFL states took significantly longer ANL (approximately one to two weeks longer) than women in states without a leave policy, except when restricting to women who delivered after their due date. In that case, I observe no significant relationship, though a small sub-sample size (n = 269) may prevent detection of a significant effect.

Discussion

This study presents the first national description of ANL and its correlates. The majority of women in this nationally-representative English-speaking U.S. sample took ANL, but almost 40% worked until the week their baby was due. Importantly, ANL appears to respond to policy: women living in states with a TDI law (i.e., Hawaii, New York, or Rhode Island in both years; New Jersey in 2005) or both TDI and PFL laws (California in both years, New Jersey in 2011-2012) were significantly more likely to take any ANL than women who lived in states without any relevant maternity leave laws. Moreover, women in states with both TDI and PFL laws took, on average, about one and a half weeks longer than women in states without any leave laws, even after controlling for characteristics of women and their jobs. The results further suggest that women in states without leave laws less commonly use ANL but also only use it only by necessity. The more health conditions women experienced before and during pregnancy, the more likely she was to take any ANL in states without relevant laws; no such relationship was observed in states with leave laws (TDI or TDI+PFL).

The results suggest a tension between job attachment and financial ability to take leave. Women in the lowest income quintile were the least likely to take leave after controlling for covariates, possibly due to an inability to forgo pay or risk losing their jobs. Full-time workers, who on average lived in higher income households, were also much less likely to take ANL. This latter result could imply that full-time employed women have greater job attachment or feel more pressure to continue working, or that full-time workers, on average, have more comfortable or flexible jobs which allow them to work throughout pregnancy.

Perhaps surprisingly, this study did not find that employer-level policies – whether paid leave was offered or the generosity of paid leave benefits – predicted ANL use or duration. One possible explanation for this null finding is that the survey did not specify whether leave offered was available before or after birth, and some employers may offer postnatal, but not antenatal, leave. Even with flexible leave, this finding supports previous observations that many women save their maternity leave for use with the baby, in which case leave generosity should predict postnatal, but not antenatal, maternity leave. In contrast, TDI laws in California and New Jersey do not allow women to "save" leave for use after delivery, but rather provide separate allowances for use ante- and post-natally. These states also have PFL programs, making them the most generous in terms of overall paid leave availability. The finding that these TDI+PFL states strongly predicted ANL may reflect this "use-it-or-lose-it" incentive structure, as well as the generosity of overall (ante- and post-natal) leave.

Employer policies were self-reported and women may not have a complete understanding of their companies' maternity leave policies. This could induce measurement error if women who took ANL have more accurate knowledge of relevant policies than women who did not take leave. For example, women who took ANL may have a detailed understanding of how much of their salary was covered while on leave, while women who took no leave might not. If women who did not take any ANL were more likely to report that their employer offered paid leave and that such leave was generous (e.g., if women who did not take ANL assumed that fully paid leave was offered, when it was not), then the relationship between employer policies and ANL would be biased toward the null. On the other hand, if women who took ANL were more likely to report that their employer offered paid leave and that such leave was generous, then the relationship between employer policies and ANL would be inflated. I find this latter explanation more probable, reinforcing my observed null finding.

This study includes women from all 50 states plus the District of Columbia over two time periods, allowing, for the first time, a national picture of ANL use. The Listening to Mothers survey is representative of English-speaking women who delivered singletons in U.S. hospitals. A major limitation of these results is that they cannot be generalized to women in the U.S. who do not speak English, or to women who delivered multiples or delivered outside a hospital. Their experiences may differ in meaningful ways and future research should examine ANL use among these groups of women. Even after applying sampling weights, Hispanic and Asian women and women born outside the U.S. are underrepresented compared to national population data derived from birth certificates (Declercq, Sakala, Corry et al., 2006; Declercq, Sakala, Corry et al., 2013).

The Listening to Mothers surveys asked women to report how many weeks before their due date they stopped working, a different measure from ANL taken relative to their actual delivery date. This fact explains the observed relationship between preterm delivery and ANL duration: women who delivered more than three weeks before their due date reported, on average, about three weeks longer ANL duration. Women who reported working until the week they were due (and who, therefore, are counted as taking no ANL), could have taken ANL if they delivered after their due date. This survey does not allow analysis of reasons for ANL, such as a healthcare provider recommending ANL. This dataset also does not include details of the woman's job tasks and work environment that could illuminate variation in leave taking.

By allowing women who need ANL to take it, antenatal maternity leave policies could have an important impact on the labor market attachment of pregnant workers. More than one-fifth of the women in this sample of employed women did not return to their pre-childbirth employer – either because they switched employers or did not return to the paid labor force at all. This figure aligns with Census Bureau estimates for the proportion of pregnant workers who quit their jobs (Laughlin, 2011). Women who otherwise would leave the labor force might remain employed if offered sufficient maternity leave.

ANL remains an understudied area of work-family policy and potential determinant of perinatal health. As women increasingly work during pregnancy, an expansive literature has found that characteristics of the prenatal work environment influence maternal and child health. Yet, almost no research examines the impact of ANL on health. The small literature that exists suggests that ANL may reduce the likelihood of delivery by Cesarean section or of reporting a difficult delivery, increase birthweight, and reduce preterm delivery, but more research in this area is needed (Guendelman, Pearl, Graham et al., 2009; Rossin, 2011; Tanaka, 2005; Xu, Séguin, & Goulet, 2002). One study of expansions in state-level unpaid leave policies found an impact on birth outcomes and attributed this to a decrease in stress through the anticipation of postpartum leave (Rossin, 2011). While it is unlikely that such anticipation would have a measureable effect on birth outcomes, living in states with postpartum leave policies could increase the likelihood of taking ANL. This may be because more generous postpartum leave availability makes women feel more comfortable using some of their flexible leave antenatally, but also because states with the most generous postpartum leave (i.e., PFL laws) all also have TDI programs that enable ANL directly.

References

Appelbaum, E., & Milkman, R. (2011). Leave that pay: employer and worker experiences with paid family leave in California: Center for Economic and Policy Research.

Declercq, E. R., Sakala, C., Corry, M. P., Applebaum, S., Connection, C., & Interactive, H. (2006). *Listening to mothers II: Report of the second national US survey of women's childbearing experiences*: Childbirth Connection.

Declercq, E. R., Sakala, C., Corry, M. P., Applebaum, S., & Herrlich, A. (2013). Listening to Mothers III: Pregnancy and Birth; Report of the Third National US Survey of Women's Childbearing Experiences. New York, NY: Childbirth Connection.

Frazier, L. M., Golbeck, A. L., & Lipscomb, L. (2001). Medically recommended cessation of employment among pregnant women in Georgia. Obstetrics & Gynecology, 97, 971-975.

Guendelman, S., Pearl, M., Graham, S., Angulo, V., & Kharrazi, M. (2006). Utilization of pay-in antenatal leave among working women in Southern California. Maternal and Child Health Journal, 10, 63-73.

Guendelman, S., Pearl, M., Graham, S., Hubbard, A., Hosang, N., & Kharrazi, M. (2009). Maternity leave in the ninth month of pregnancy and birth outcomes among working women. Women's Health Issues, 19, 30-37.

Klerman, J. A., Daley, K., & Pozniak, A. (2012). Family and medical leave in 2012: Technical report. Cambridge, MA: Abt Associates Inc.
Laughlin, L. (2011). Maternity leave and employment patterns: 2006-2008 *Current Population Reports* (pp. 70-128). Washington, D.C. : U.S. Census Bureau.
Matos, K., & Galinsky, E. (2014). National study of employers. New York: Families and Work Institute.
Mozurkewich, E., Luke, B., Avni, M., & Wolf, F. (2000). Working conditions and adverse pregnancy outcome: a meta-analysis. Obstetrics & Gynecology, 95, 623-635.
Rossin, M. (2011). The effects of maternity leave on children's birth and infant health outcomes in the United States. Journal of Health Economics, 30, 221-239.
Ruhm, C. J. (1997). Policy Watch: The Family and Medical Leave Act. Journal of Economic Perspectives, 11, 175-186.
Shepherd-Banigan, M., & Bell, J. F. (2014). Paid leave benefits among a national sample of working mothers with infants in the United States. Maternal and Child Health Journal, 18, 286-295.
Stearns, J. (2015). The Effects of Paid Maternity Leave: Evidence fromTemporary Disability Insurance. Journal of Health Economics.
Tanaka, S. (2005). Parental leave and child health across OECD countries. The Economic Journal, 115, F7-F28.
Xu, Q., Séguin, L., & Goulet, L. (2002). Effet bénéfique d'un arrêt du travail avant l'accouchement. Canadian Journal of Public Health/Revue Canadienne de Sante'e Publique, 72-77.

Tables and Figures

Table 1: ANL use and duration among employed women (N = 1,705), unadjusted estimates. Source: Listening to Mothers Survey, waves II & III.

	No ANL		Took ANL			ANL duration		
	N	Weighted means, by ANL status	N	Weighted means, by ANL status	p-value[1]	Mean	Std. Error.	p-value[2]
Total	678	36.8%	1,027	63.2%		5.52	0.28	
State policy								
None	608	40.1%	814	59.9%	<.001	5.21	0.27	0.20
TDI only[3]	32	25.9 %	77	74.1%		6.96	1.51	
TDI + PFL[4]	38	17.1%	136	82.9%		6.63	0.94	
Paid leave offered								
Yes	312	38.1%	484	61.9%	0.49	4.98	0.41	0.08
No	366	35.8%	543	64.2%		5.96	0.37	
>6 FTE weeks[5]								
Yes	142	42.8%	215	57.2%	0.09	4.42	0.58	0.04
No	536	35.3%	812	64.7%		5.78	0.31	
Work status								
Full-time	524	42.9%	612	57.1%	<.001	5.01	0.35	0.03
Part-time	154	25.9%	415	74.1%		6.25	0.44	
Self-employed								
Yes	58	50.7%	67	49.3%	0.02	6.66	1.28	0.36
No	620	35.5%	960	64.5%		5.45	0.28	
Parity								
Nulliparous	313	36.4%	501	63.6%	0.94	5.45	0.38	0.63
1 prior birth	238	37.7%	332	62.3%		5.88	0.56	
2+ prior births	127	36.5%	194	63.5%		5.15	0.51	
Pregnancy intention								
Wanted later or did not want	225	36.8%	365	63.2%	0.98	5.17	0.33	0.26
Wanted sooner or at that time	452	36.7%	662	63.3%		5.75	0.40	
Maternal age								
18-29	314	34.0%	544	66.0%	0.09	5.44	0.35	0.76
30-34	225	38.1%	307	61.9%		5.39	0.53	
35+	139	43.6%	176	56.4%		6.06	0.80	
Race/ethnicity								
Non-Hispanic white	502	40.4%	653	59.6%	0.06	5.68	0.34	0.91
Non-Hispanic black	61	33.5%	143	66.5%		5.14	0.68	
Hispanic	70	27.7%	173	72.3%		5.45	0.73	
Other	35	36.8%	49	63.2%		5.59	0.97	

Marital status								
Married/partnered	649	37.1%	970	62.9%	0.62	5.59	0.29	0.13
Unmarried/without partner	29	33.2%	57	66.9%		4.56	0.61	
Maternal education								
Less than high school	4	23.1%	23	76.9%	<.01	6.35	1.64	0.94
High school grad	58	27.2%	137	72.8%		5.62	0.58	
Some college	258	37.4%	433	62.6%		5.40	0.30	
College grad	358	46.0%	434	54.0%		5.43	0.48	
Income quintile								
1st	84	35.4%	155	64.6%	0.06	5.69	0.53	0.94
2nd	96	26.0%	182	74.0%		5.60	0.73	
3rd	121	38.6%	169	61.4%		5.82	0.70	
4th	182	41.4%	250	58.6%		5.36	0.42	
5th	160	40.2%	208	59.8%		5.11	0.66	
Maternity care payment								
Medicaid	141	29.0%	357	71.0%	<.01	5.64	0.40	0.73
Private/OOP	537	41.0%	670	59.0%		5.45	0.37	
Fertility treatment								
Yes	48	23.4%	108	76.6%	<.01	4.86	0.53	0.21
No	621	38.5%	898	61.5%		5.62	0.31	
Pre-pregnancy weight								
Obese	153	39.0%	200	61.0%	0.50	5.61	0.58	0.88
Non-obese	525	36.3%	827	63.7%		5.51	0.31	
Ultrasounds								
≤3 ultrasounds	426	38.2%	607	61.8%	0.34	5.10	0.34	0.08
>3 ultrasounds	252	34.9%	420	65.1%		6.10	0.45	
Gestational age at delivery								
Preterm	58	44.1%	66	55.9%	0.11	7.66	0.84	0.06
Early term	191	43.0%	216	57.0%		6.07	0.66	
Full term	362	34.1%	624	65.9%		5.29	0.37	
Late term	54	33.1%	90	66.9%		4.84	0.75	
Post term	7	56.1%	12	43.9%		6.60	1.19	
Wave								
II (2005)	396	38.6%	559	61.4%	0.23	6.22	0.42	<.01
III (2011-2012)	282	34.4%	468	65.6%		4.62	0.30	
Time to interview								
0-3 months	35	42.0%	49	58.0%	0.38	7.09	0.92	0.11
4-6 months	96	36.9%	144	63.1%		6.64	0.90	
7-9 months	165	36.1%	262	63.9%		5.68	0.65	
10-12 months	216	39.2%	273	60.8%		5.01	0.39	
13-15 months	56	43.7%	81	56.3%		5.72	1.22	

16-18 months	71	36.1%	113	63.9%	4.47	0.46
19-21 months	39	24.7%	105	75.3%	4.88	0.64

[1]p-value from Chi-square test

[2]p-value from F-test

[3]Includes Hawaii, New York, and Rhode Island in both waves; New Jersey in wave II

[4]Includes California in both waves; New Jersey in wave III

[5]Full-time equivalent weeks of paid leave offered = number of weeks of paid leave offered x wage replacement ratio

Table 2: ANL use among employed women in the United States (n = 1,705), adjusted estimates using survey weighted Listening to Mothers Survey, waves II & III.

	State-level policies		Employer-level policies	
	Unadjusted Coeff. [95% CI]	Adjusted Coeff. [95% CI]	Unadjusted Coeff. [95% CI]	Adjusted Coeff. [95% CI]
TDI only state[1]	0.14 [0.09 - 0.19]**	0.12 [0.07 - 0.17]**		
TDI + PFL state[2]	0.23 [0.19 - 0.27]**	0.25 [0.21 - 0.29]**		
Paid leave offered			0.00 [-0.09 - 0.09]	-0.03 [-0.13 - 0.07]
>6 FTE weeks[3]			-0.07 [-0.19 - 0.05]	-0.01 [-0.15 - 0.12]
Full-time work		-0.17 [-0.23 - -0.10]**		-0.16 [-0.22 - -0.09]**
Self-employed		-0.22 [-0.38 - -0.06]**		-0.22 [-0.39 - -0.06]**
Parity (ref: nulliparous)		–		–
1 prior birth		-0.02 [-0.10 - 0.05]		-0.04 [-0.11 - 0.04]
2+ prior births		-0.02 [-0.11 - 0.07]		-0.01 [-0.10 - 0.08]
Unintended pregnancy		0.00 [-0.07 - 0.07]		-0.01 [-0.08 - 0.06]
Maternal age (ref: 18-29)		–		–
30-34		0.01 [-0.07 - 0.09]		0.02 [-0.06 - 0.10]
35+		0.01 [-0.11 - 0.13]		0.01 [-0.11 - 0.13]
Race/ethnicity (ref: non-Hispanic white)		–		–
Non-Hispanic black		0.06 [-0.01 - 0.13]+		0.08 [-0.00 - 0.16]+
Hispanic		0.05 [-0.03 - 0.12]		0.09 [-0.00 - 0.19]+
Other		-0.07 [-0.19 - 0.05]		-0.02 [-0.18 - 0.14]
Unmarried/no partner		0.00 [-0.15 - 0.14]		0.00 [-0.14 - 0.14]
Maternal education (ref: less than high school)		–		–
High school grad		0.04 [-0.19 - 0.28]		0.04 [-0.20 - 0.27]
Some college		-0.07 [-0.30 - 0.16]		-0.07 [-0.29 - 0.15]
College grad		-0.17 [-0.41 - 0.07]		-0.16 [-0.40 - 0.08]
Income quintile (ref: 1st)		–		–
2nd		0.17 [0.03 - 0.31]*		0.18 [0.05 - 0.32]**
3rd		0.09 [-0.05 - 0.22]		0.11 [-0.03 - 0.24]
4th		0.08 [-0.06 - 0.22]		0.09 [-0.05 - 0.23]
5th		0.13 [0.01 - 0.25]*		0.15 [0.03 - 0.27]*
Medicaid payment[4]		0.08 [-0.02 - 0.19]		0.07 [-0.04 - 0.18]
Fertility help		0.07 [-0.01 - 0.16]+		0.08 [-0.01 - 0.16]+
Pre-pregnancy obesity		-0.03 [-0.10 - 0.05]		-0.03 [-0.10 - 0.05]

>3 ultrasounds		0.03 [-0.04 - 0.10]		0.04 [-0.03 - 0.12]
Gestational age at delivery (ref: full term)		–		–
Preterm		-0.11 [-0.24 - 0.03]		-0.12 [-0.26 - 0.02]+
Early term		-0.06 [-0.18 - 0.05]		-0.06 [-0.18 - 0.05]
Late term		0.03 [-0.09 - 0.15]		0.06 [-0.07 - 0.19]
Post term		0.00 [-0.33 - 0.32]		0.01 [-0.28 - 0.31]
Observations	1,705	1,537	1,705	1,537
R-squared	0.03	0.14	0.01	0.11

Robust 95% confidence intervals in brackets. Std. errors adjusted for state clustering. All models control for survey wave, time to interview, and interaction between wave and time to interview.

+ significant at 10%; * significant at 5%; ** significant at 1%

[1]Includes Hawaii, New York, and Rhode Island in both waves; New Jersey in wave II

[2]Includes California in both waves; New Jersey in wave III

[3]Full-time equivalent weeks of paid leave offered = number of weeks of paid leave offered x wage replacement ratio

[4]Medicaid was primary payment source for maternity care (versus private insurance or out-of-pocket)

Table 3: ANL duration (in weeks) among employed women who took any ANL in the United States (n = 1,027), adjusted estimates using survey weighted Listening to Mothers Survey, waves II & III.

	State-level policies		Employer-level policies	
	Unadjusted Coeff. [95% CI]	Adjusted Coeff. [95% CI]	Unadjusted Coeff. [95% CI]	Adjusted Coeff. [95% CI]
TDI only state[1]	1.49 [-1.39 - 4.36]	1.28 [-1.77 - 4.33]		
TDI + PFL state[2]	1.46 [0.53 - 2.38]**	1.84 [0.92 - 2.77]**		
Paid leave offered			0.08 [-1.43 - 1.58]	0.55 [-1.28 - 2.38]
>6 FTE weeks[3]			-1.43 [-2.84 - -0.02]*	-1.55 [-3.57 - 0.46]
Full-time work		-1.39 [-2.42 - -0.37]**		-1.19 [-2.24 - -0.13]*
Self-employed		0.25 [-2.37 - 2.88]		0.27 [-2.29 - 2.82]
Parity (ref: nulliparous)		–		–
1 prior birth		0.37 [-1.10 - 1.83]		0.24 [-1.10 - 1.59]
2+ prior births		-0.59 [-1.93 - 0.75]		-0.57 [-1.97 - 0.82]
Unintended pregnancy		-1.25 [-2.58 - 0.09]+		-1.33 [-2.74 - 0.07]+
Maternal age (ref: 18-29)		–		–
30-34		0.14 [-1.62 - 1.89]		0.23 [-1.61 - 2.07]
35+		0.63 [-0.45 - 1.71]		0.75 [-0.46 - 1.95]
Race/ethnicity (ref: non-Hispanic white)		–		–
Non-Hispanic black		-0.06 [-1.37 - 1.26]		-0.07 [-1.44 - 1.30]
Hispanic		-0.2 [-1.74 - 1.34]		0.27 [-1.35 - 1.89]
Other		-0.23 [-2.62 - 2.16]		0.48 [-1.47 - 2.42]
Unmarried/no partner		-1.24 [-3.01 - 0.52]		-1.38 [-3.08 - 0.32]
Maternal education (ref: less than high school)		–		–
High school grad		0.02 [-3.42 - 3.46]		0.03 [-3.52 - 3.58]
Some college		0.18 [-2.92 - 3.27]		0.20 [-3.05 - 3.45]
College grad		0.28 [-3.21 - 3.77]		0.51 [-3.10 - 4.12]
Income quintile (ref: 1st)		–		–
2nd		0.11 [-1.61 - 1.83]		0.03 [-1.81 - 1.88]
3rd		-0.34 [-2.30 - 1.62]		-0.24 [-2.13 - 1.65]
4th		-1.03 [-2.41 - 0.35]		-1.03 [-2.44 - 0.38]

5th		-1.38 [-3.26 - 0.50]		-1.28 [-3.10 - 0.53]
Medicaid payment[4]		0.60 [-0.34 - 1.54]		0.42 [-0.63 - 1.47]
Fertility help		-0.41 [-1.75 - 0.94]		-0.70 [-2.26 - 0.87]
Pre-pregnancy obesity		-0.15 [-1.90 - 1.60]		-0.09 [-1.82 - 1.64]
>3 ultrasounds		0.98 [-0.37 - 2.34]		1.07 [-0.32 - 2.46]
Gestational age at delivery (ref: full term)		–		–
Preterm		3.20 [1.02 - 5.38]**		3.11 [0.97 - 5.25]**
Early term		0.65 [-0.66 - 1.97]		0.67 [-0.63 - 1.97]
Late term		-0.66 [-1.89 - 0.56]		-0.41 [-1.72 - 0.91]
Post term		1.21 [-1.03 - 3.45]		1.23 [-1.06 - 3.51]
Observations	1,027	917	1,027	917
R-squared	0.03	0.09	0.03	0.09

Robust 95% confidence intervals in brackets. Std. errors adjusted for state clustering. All models control for survey wave, time to interview, and interaction between wave and time to interview.

+ significant at 10%; * significant at 5%; ** significant at 1%

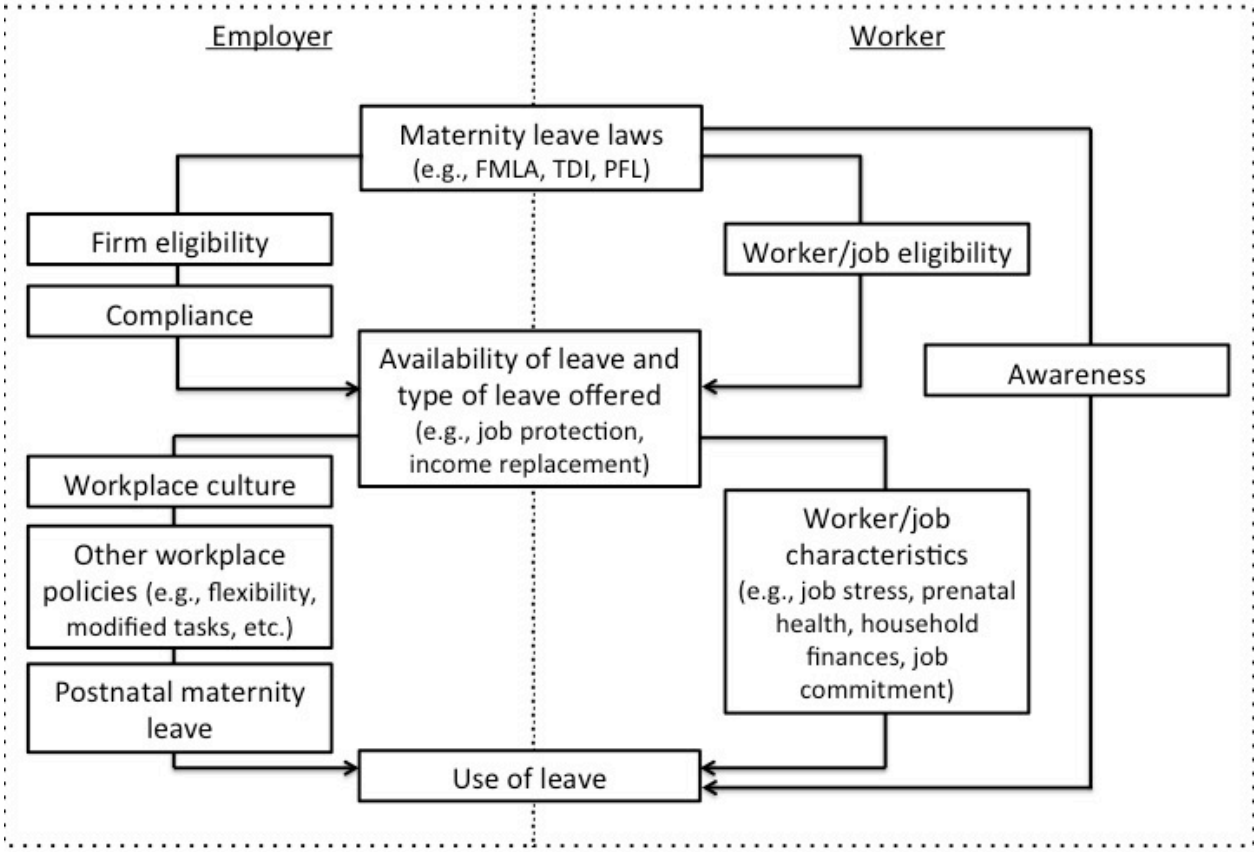
[1]Includes Hawaii, New York, and Rhode Island in both waves; New Jersey in wave II

[2]Includes California in both waves; New Jersey in wave III

[3]Full-time equivalent weeks of paid leave offered = number of weeks of paid leave offered x wage replacement ratio

[4]Medicaid was primary payment source for maternity care (versus private insurance or out-of-pocket)

Figure 1: Multi-level framework of the pathways to ANL usage.



Abbreviations: Family and Medical Leave Act of 1993 (FMLA), Temporary Disability Insurance (TDI), Paid Family Leave (PFL)

Figure 2: Distribution of ANL among employed women (in weeks) (N = 1,705). Source: Listening to Mothers Survey, waves II & III.

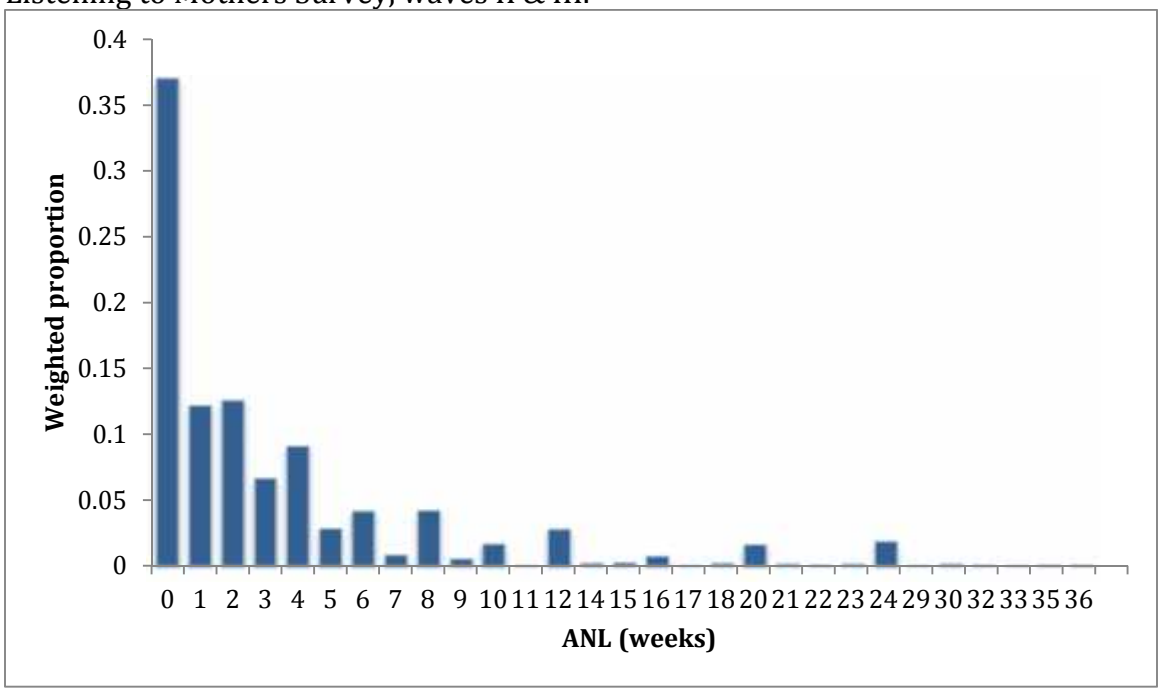
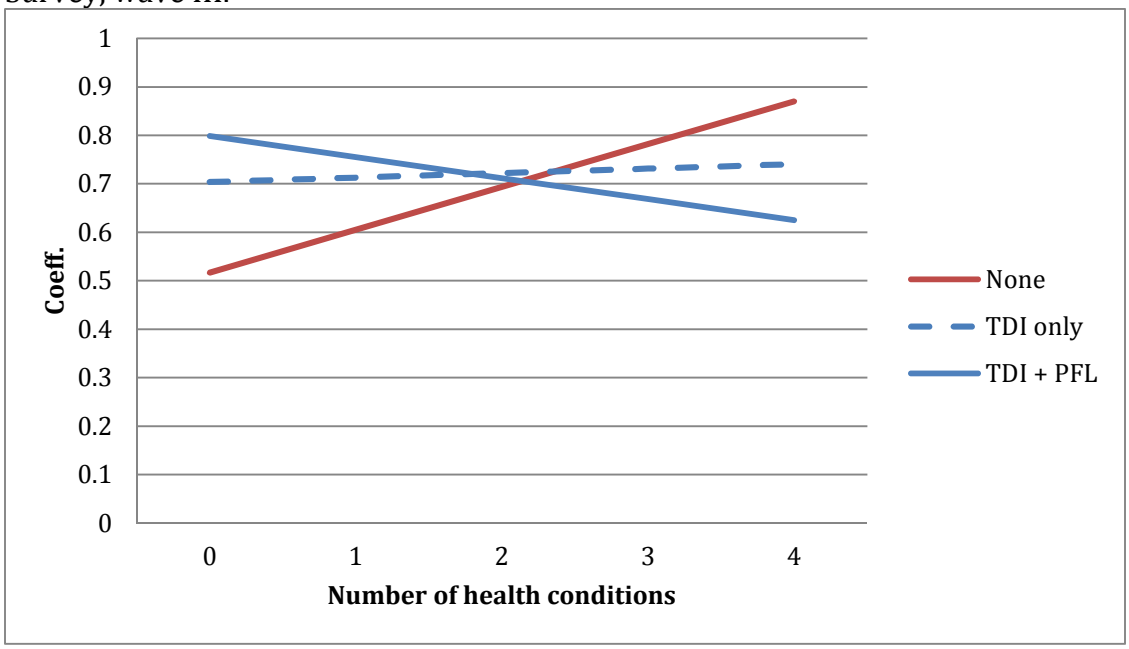


Figure 3: Relationship between ANL use and the number of pre-pregnancy health conditions experienced, by state policy group (N = 675). Source: Listening to Mothers Survey, wave III.



Notes: Standard errors in all models are adjusted for state clustering. Models are adjusted for full-time vs. part-time work; self-employed vs. working for other; parity; pregnancy intention; maternal sociodemographic characteristics (age, race/ethnicity, marital status, educational attainment, household income, maternity care payment source); use of fertility treatments; pre-pregnancy obesity (BMI > 30 kg/m^2);

gestational age at delivery; and time to interview. TDI only group includes Hawaii, New York, or Rhode Island in both years and New Jersey in 2005. TDI + PFL group includes California in both years and New Jersey in 2011-2012. Health conditions include use of medication for high blood pressure or depression before pregnancy, pre-pregnancy diagnosis of diabetes, and diagnosis of gestational diabetes.

Appendix

Table A1: Sensitivity analyses for ANL use among employed women in the United States (n = 1,705), adjusted estimates using survey weighted Listening to Mothers Survey, waves II & III.

	Delivery relative to due date		Remove if planned C-section	Quit or were fired		Censor long ANL duration		Time to interview ≤18 months
	Before due date	After due date		Remove if quit or were fired	Include only if quit or were fired	≤8 weeks only	≤4 weeks only	
	Coeff. [95% CI]	Coeff. [95% CI]	Coeff. [95% CI]	Coeff. [95% CI]	Coeff. [95% CI]	Coeff. [95% CI]	Coeff. [95% CI]	Coeff. [95% CI]
TDI only state[1]	0.10** [0.03 - 0.17]	0.15* [0.00 - 0.29]	0.11** [0.05 - 0.17]	0.13** [0.07 - 0.20]	0.08* [0.01 - 0.16]	0.14** [0.09 - 0.20]	0.13** [0.07 - 0.20]	0.14** [0.08 - 0.19]
TDI + PFL state[2]	0.25** [0.19 - 0.31]	0.22** [0.13 - 0.31]	0.25** [0.20 - 0.31]	0.24** [0.19 - 0.30]	0.22** [0.14 - 0.30]	0.25** [0.21 - 0.30]	0.27** [0.22 - 0.32]	0.25** [0.21 - 0.30]
Observations	1,114	423	1,348	1,028	509	1,375	1,206	1,410
R-squared	0.15	0.17	0.14	0.15	0.21	0.14	0.16	0.13

Notes: Robust 95% confidence intervals in brackets. Std. errors adjusted for state clustering. All models are adjusted for full-time vs. part-time work; self-employed vs. working for other; parity; pregnancy intention; maternal sociodemographic characteristics (age, race/ethnicity, marital status, educational attainment, household income, maternity care payment source); use of fertility treatments; pre-pregnancy obesity (BMI > 30 kg/m2); gestational age at delivery; survey wave, time to interview, and interaction between wave and time to interview.

* significant at 5%; ** significant at 1%

[1]Includes Hawaii, New York, and Rhode Island in both waves; New Jersey in wave II

[2]Includes California in both waves; New Jersey in wave III

Table A2: Sensitivity analyses for ANL duration among employed women in the United States (n = 1,027), adjusted estimates using survey weighted Listening to Mothers Survey, waves II & III.

	Delivery relative to due date		Remove if planned C-section	Quit or were fired		Censor long ANL duration		Time to interview ≤18 months
	Before due date	After due date		Remove if quit or were fired	Include only if quit or were fired	≤8 weeks only	≤4 weeks only	
	Coeff. [95% CI]	Coeff. [95% CI]	Coeff. [95% CI]	Coeff. [95% CI]	Coeff. [95% CI]	Coeff. [95% CI]	Coeff. [95% CI]	Coeff. [95% CI]
TDI only state[1]	-0.17 [-1.26 - 0.92]	2.96 [-2.47 - 8.38]	1.52 [-1.83 - 4.87]	0.33 [-0.52 - 1.18]	3.25 [-1.14 - 7.63]	-0.45 [-1.12 - 0.23]	-0.41+ [-0.85 - 0.04]	1.47 [-1.83 - 4.77]
TDI + PFL state[2]	2.62** [1.55 - 3.70]	-0.45 [-2.29 - 1.38]	2.00** [0.94 - 3.06]	1.61** [0.50 - 2.71]	2.22** [0.56 - 3.88]	0.75** [0.35 - 1.15]	1.12** [0.81 - 1.42]	1.94** [0.98 - 2.90]
Observations	648	269	809	526	391	755	586	827
R-squared	0.12	0.21	0.11	0.17	0.14	0.13	0.19	0.10

Notes: Robust 95% confidence intervals in brackets. Std. errors adjusted for state clustering. All models are adjusted for full-time vs. part-time work; self-employed vs. working for other; parity; pregnancy intention; maternal sociodemographic characteristics (age, race/ethnicity, marital status, educational attainment, household income, maternity care payment source); use of fertility treatments; pre-pregnancy obesity (BMI > 30 kg/m2); gestational age at delivery; survey wave, time to interview, and interaction between wave and time to interview.

+ significant at 10%; * significant at 5%; ** significant at 1%

[1]Includes Hawaii, New York, and Rhode Island in both waves; New Jersey in wave II

[2]Includes California in both waves; New Jersey in wave III

Chapter 3
Antenatal maternity leave and childbirth using the First Baby Study: a propensity score analysis

Abstract

In a prospective cohort of 1,740 employed, nulliparous women in Pennsylvania, we use propensity scores to estimate the impact of taking antenatal leave (ANL) on the probability of a negative delivery outcome (labor induction, long labor duration, unplanned C-section, and self-reported negative birth experience) and the number of negative delivery outcomes experienced. Half of the sample did not stop working before delivery. Women who stopped working at least 2 days prior to delivery experienced 16% more negative delivery outcomes, on average, than women who worked until delivery. This result appears due in part to the 25% higher predicted probability of unplanned C-section among women who took ANL. In a context of limited maternity leave availability, only relatively unhealthy women stop working before delivery.

Introduction and background

Women's labor force participation during pregnancy has changed dramatically over the past half century. American women are increasingly likely to be employed during pregnancy and, among those who are employed, to work later into their pregnancy. In 2006-2008, 66% of first-time mothers worked during pregnancy, up from 44% in 1961-1965 (Laughlin 2011). Of those, 82% worked in the month they delivered.

This increase in prenatal employment has led to great interest in the health effects of work on pregnancy. Studies examining the relationship between psychological aspects of work and health yield inconsistent results, but strong evidence suggests that physical characteristics of the work environment (e.g., hours spent standing, working the night shift, exposure to toxic chemicals) adversely affect birth outcomes (Bonzini et al. 2011; Del Bono, Ermisch, and Francesconi 2012; Mozurkewich et al. 2000). For example, a meta-analysis based on 160,988 women in 29 observational studies found that physically demanding work was significantly associated with preterm delivery (PTD), small-for-gestational-age (SGA), and hypertension or preeclampsia (Mozurkewich et al. 2000).

A small literature attempts to assess this relationship by examining the effect of taking leave from work during pregnancy ("antenatal maternity leave") on obstetric outcomes. In a sample of full-time employed California women who delivered at term, Guendelman et al. (2009) found one-fourth the odds of primary (first-time) Cesarean-section among women who took leave in the ninth month of pregnancy (OR: 0.27, CI: 0.08-0.94) compared to those who worked until delivery after adjusting for gestational age, infant sex, maternal race, parity, pre-pregnancy BMI, height and occupation. The 62 women who took antenatal leave did not differ demographically from the 385 women who did not take leave, but the statistical power was limited in this small sample.

In a larger study, Hung et al. (2002) used the Infant Feeding Practices Study of 1,194 employed women to examine characteristics of prenatal work (duration of work into pregnancy, hours worked, occupation and work intensity) and mode of delivery. Compared to women who stopped working in the first trimester, they observed no increased risk of C-section for women who worked until the last month of pregnancy or up to the date of delivery. They used women's planned stop date as a proxy for antenatal leave, which could help to deal with confounding by women who stopped working because they did not feel well (independently leading to C-sections), but this does not reflect effects of actual leave time taken. The authors also looked at work intensity, which combined the month during pregnancy she stopped working with hours per week worked during pregnancy. Women with the most intense work effort (full-time work with no plan to stop before delivery) did not have an increased risk of C-section compared with women with the least intense work effort (part-time work with plans to stop before delivery), though women with mid-level work intensity (full-time planning on stopping or part-time not planning on stopping) had lower risk of C-sec than the women with least intense work effort (OR: 0.62, 95% CI: 0.04-1.04; OR: 0.64, CI: 0.45-0.91, respectively). Like women who are not employed during pregnancy, women working part-time with plans to stop before delivery may have other unobserved characteristics that put them at greater risk of obstetric complications.

Finally, Xu, Séguin, and Goulet (2002) examined 363 women who worked at least through their first trimester and who delivered full-term infants in Montreal, Canada

during 1996. After adjusting for parity and the number of health problems during pregnancy, they found that longer work stoppage before delivery was associated with a slightly reduced odds of difficult delivery (OR: 0.96, CI: 0.93-0.99). Their outcome measure included emergency C-section, induced labor, instrumental delivery, long labor duration, maternal hemorrhage, and labor augmentation, but did not single out any of these outcomes independently.

These observed relationships could operate through two main pathways – protection and selection – linking antenatal leave (ANL) and maternal health, each leading to a different directional effect (Figure 1). Through the "protective" pathway, ANL removes women from a stressful or strenuous work environment and allows them to rest and sleep at the end of pregnancy, leading to positive maternal health outcomes. Fatigue and poor sleep have been identified as risk factors for obstetric complication (Chien and Ko 2004; Lee and Gay 2004; Mayberry et al. 1999). One study of Taiwanese women found that fatigue was significantly associated with higher odds of C-section (OR: 1.04; CI: 1.01-1.07) after controlling for maternal age, multiple gestations, history of preterm birth and abdominal operation during pregnancy (Chien and Ko 2004). A small prospective cohort study of primiparous women delivering at full-term in the United States also found that prenatal employment status (whether women were employed or not) did not predict obstetric outcomes, but sleep at the end of pregnancy did (Lee and Gay 2004). Objective measures of poor sleep quality and quantity predicted long labor duration and C-section; C-section was also more likely among women with poor self-reported sleep quality.

At the same time, the relationship between ANL and maternal health could operate through a "selective" pathway. Women who take ANL likely differ from women who do not take ANL in meaningful ways. Particularly in the United States where ANL is relatively uncommon, women who take ANL may be more likely to have suffered from serious health complications during pregnancy or hold very strenuous jobs, both of which could independently lead to worse maternal health outcomes. One study of women employed at least 20 hours a week during pregnancy in California found that working the night shift, lack of fulfillment in one's job, and being stressed and/or tired were associated with taking ANL (Guendelman et al. 2006). A study of employed pregnant women in Georgia found that more than a quarter (27.7%) were advised by a healthcare provider to stop working during pregnancy and that being so advised significantly decreased the likelihood that women continued working into their ninth month of pregnancy (Frazier, Golbeck, and Lipscomb 2001).

Our study tests whether ANL is associated with maternal perinatal health – specifically, with negative delivery outcomes. We use both propensity score matching and multivariable regression models with a rich dataset that includes detailed information on women's experiences during pregnancy, labor and delivery, and postpartum. This dataset allows us to build on prior studies, many of which lack adequate controls, like parity or maternal obesity (Chien and Ko 2004; Lee and Gay 2004; Xu et al. 2002), to better address the endogeneity of ANL. We include a novel and expansive set of covariates that capture prenatal physical and mental health more comprehensively than prior studies. We focus exclusively on nulliparous women, the group most likely to experience negative delivery outcomes. Moreover, pregnant women without other children at home may be more able to use maternity leave to rest.

We further contribute to the literature by studying ANL within the United States. Most of the work on prenatal leave-taking occurs outside the U.S. (Ceron-Mireles, Harlow, and Sanchez-Carrillo 1996; Chien and Ko 2004; Fortier, Marcoux, and Brisson 1995; Mamelle, Bertucat, and Munoz 1989; Saurel-Cubizolles et al. 2004; Sydsjö et al. 2006; Wüst 2014; Xu et al. 2002). This matters because the baseline intensity of work and duration of antenatal leave is dramatically longer outside the U.S. For example, in the study by Xu et al. (2002), women stopped working on average 7.6 weeks before delivery and only 5% worked up until delivery. Most of those women had access to paid antenatal leave. Compare that to the Guendelman et al. (2009) study where only 15% of full-time employed women took any antenatal leave and those who did mostly stopped in the last month of pregnancy. In places where long antenatal leave is the norm, detecting an effect on health may prove difficult, whereas we may observe a larger effect when women routinely work until they go into labor. Finally, our dataset allows us to more precisely measure negative delivery outcomes than prior studies could. Most studies of work and delivery outcomes in general, and C-section in particular, use an imprecise outcome variable (Chien and Ko 2004; Hung et al. 2002; Lee and Gay 2004; Xu et al. 2002). We are able to look both at the likelihood of experiencing a negative delivery outcome and examine individual negative perinatal events and experiences. Furthermore, we can single out primary, unplanned C-sections which more likely respond to the prenatal work environment as distinct from repeat or planned C-sections.

Methods

First Baby Study

We use data from the First Baby Study (FBS), a prospective interview study designed to investigate the association between mode of delivery and subsequent fertility over the course of a 3-year follow-up period after a first birth (Kjerulff et al. 2013). Baseline interviews, conducted when participants were between 30 and 42 weeks of gestation (median gestational age was 35 weeks), assessed reproductive and health history; pregnancy complications and health care utilization; mode of delivery preference; relationship factors; psychosocial factors; future birth desires and intensions; and sociodemographic factors. A one-month postpartum interview assessed factors related to labor and delivery; postpartum feelings about childbirth; in-hospital and post-discharge complications; and the health of the baby and the mother. Subsequent interviews, not used in the current study, were conducted at 6, 12, 18, 24, 30, and 36 months postpartum. Birth certificate and hospital discharge data were obtained for the women who completed baseline and one-month follow-up interviews. Eligible respondents spoke English or Spanish, were between 18 and 35, nulliparous, pregnant with a singleton and planning to deliver in a Pennsylvania hospital.

The Penn State College of Medicine IRB as well as the IRBs of participating hospitals located throughout the State of Pennsylvania approved the FBS. Study investigators recruited participants through the placement of study brochures, flyers and posters in strategic locations at a variety of venues including hospitals, obstetricians' offices, ultrasound centers, low-income clinics community health and pregnancy support centers; they sent press releases to newspapers across the state and placed ads in community

newspapers and weekly publications; and they used the internet for hospital intranet postings and webpage announcements. Study recruiters described the study and distributed brochures to potential participants attending childbirth education classes and hospital tours associated with participating hospitals. The investigators mailed brochures to potentially eligible women by a Medicaid insurer that served women across Pennsylvania. Finally, they mailed recruitment materials to women reported to fulfill eligibility requirements whose names and addresses were provided by a marketing company and compiled from information obtained from credit card companies, magazines, charities, organizations, manufacturers, and retailers (Kjerulff et al. 2013).

In order to reach target enrolment of 3,000 participants, the study authors recruited 3,080 women who consented and completed the baseline interview; 3,006 completed both the baseline and one-month postpartum interviews. The 74 women who dropped out after the baseline interview were younger, less likely to be covered by private insurance, and more likely to live in an urban area. They did not differ significantly in race/ethnicity.

Compared to a population of first, singleton births among women aged 18-36 in Pennsylvania in 2008, women in the FBS were slightly older, more likely to be White, more educated, more likely to have private insurance, and more likely to be married; they did not differ in likelihood of delivering by C-section (Kjerulff et al. 2013).

Inclusion/exclusion criteria

We excluded women who were not employed at a job for pay at any time while pregnant, assessed at the baseline interview (N = 319). We also excluded women who delivered preterm (< 37 weeks 0 days gestation) because these women did not have equal opportunity to take antenatal leave before delivery (N = 108). We further excluded women who stopped working more than 30 days before delivery (N = 5) because longer leave may indicate health problems that are independently associated with risk of C-section, women who were employed but missing the ANL variable (N = 2), and women whose ANL was estimated (N = 3; included in sensitivity analyses). Furthermore, women who stop working more than two months before delivery more likely quit or were fired than women who stop later, according to U.S. Census data from 2006-2008 (Laughlin 2011). Finally, we excluded from our main analyses women who quit or were fired during pregnancy (N = 829), though these women were included in sensitivity analyses.

After excluding women who were not employed during pregnancy, women who delivered preterm, women who stopped working more than 30 days before delivery or with missing or imputed ANL, and women who quit or were fired, our analytic sample included 1,740 women.

Definition of key variables

Independent variable – antenatal leave (ANL)

Antenatal maternity leave measures the amount of time before delivery that employed women stopped working. At the one-month postpartum interview, women who were still employed at the end of pregnancy were asked how long before their child was born they stopped working. Respondents reported the number of days before delivery they stopped working or responded that they had the baby the same day they worked.

Of the 1,743 women who were employed throughout pregnancy, 435 worked until the day they delivered their baby and 1,308 reported at least one day of leave. Responses were open-ended and were recoded into days before delivery. In most cases, re-coding was straightforward as women gave responses in number of days or weeks. However, a few cases were not as clear. We recoded "Just the weekend" and "Less than a week" as two and six days, respectively. One case referred to longer leave of uncertain duration: we coded "…stopped in June" as one month since the exact date of delivery was not available. We dropped these three cases from the main analysis, but included in a sensitivity test.

Holding no prior hypothesis for how much ANL constitutes a sufficient exposure to effect results, we focused on whether women took any leave, defined as stopping work at least two days prior to delivery. We considered women who worked until the day before they delivered to have taken no leave because they may have stopped working due to labor beginning and did not actually have any true leave from work.

Quit or fired – 829 women who reported being employed during pregnancy at the baseline interview later responded that they were no longer employed in the two weeks prior to delivery. We dropped these women from the main analysis, but included them in a sensitivity test.

Dependent variables

The primary dependent variables were whether women experienced a negative delivery outcome (induction of labor, labor lasting more than 24 hours, unplanned Cesarean delivery, or self-reported negative delivery experience) and the number of such outcomes experienced. All outcomes were measured at the one-month postpartum interview. These were cross-checked with birth certificate or medical records.

Induction of labor – The FBS survey asked women who did not have a planned C-section if a doctor or nurse in the hospital tried to cause their labor to begin by the use of drugs or another technique, i.e., did they try to induce labor. We considered women to have been induced if they both reported induced labor and not being in labor upon arrival at the hospital. We compared self-reported labor induction with measures of labor induction from the birth certificates and medical records and determined this variable to be reliable.

Labor duration – We measured labor duration as the number of hours reported from the time when contractions became regular until the baby was born. This measure excludes women who had a planned C-section, since these women did not experience labor.

Unplanned Cesarean delivery (C-section) – Mode of delivery came from the one-month postpartum survey and was verified with birth certificates (there were no inconsistencies). Mode of delivery includes four primary categories: spontaneous vaginal; instrumental vaginal (i.e., vaginal delivery with the assistance of forceps or vacuum); planned C-section (i.e., scheduled to occur before labor began); or unplanned C-section. We compared unplanned C-section to spontaneous or instrumental vaginal, excluding planned C-sections from this analysis.

Birth experience – Measured using a 16-item questionnaire, the First Baby Study Birth Experience Scale (Kinsey et al. 2013) asks mothers to rate how they felt right after having their baby (or if unconscious, after waking up) on scale of 1 (extremely) to 5 (not at all). The scale includes brief adjectives or statements from each of the four sub-dimensions of birth experience (emotional adaptation, physical discomfort, fulfillment, and negative emotional experience). Possible scores range from 16 to 80, with a high score indicating a

positive birth experience. The quintile of women with the lowest scores on the scale represents those having a negative birth experience (Elvander, Cnattingius, and Kjerulff 2013).

Other variables: covariates/potential confounders

We selected covariates that have been shown in the literature to influence ANL, primary outcomes, or both and which are unaffected by treatment status. We measured all covariates at the baseline interview, unless otherwise noted.

Employment variables: Full-time or part-time work. Job title was recoded into five major categories: professional, managerial, service, labor, and other.

Health variables: Pre-pregnancy obesity and whether the woman gained more weight than recommended during pregnancy (Yaktine and Rasmussen 2009) (both measured at one-month postpartum interview); prior miscarriage; self-reported history of diabetes or hypertension before pregnancy; presence of medical problems during current pregnancy (hypertension, high blood pressure, or pre-eclampsia that started during pregnancy; diabetes that started during pregnancy; vaginal bleeding; early or pre-term labor; bed rest or hospitalization because of premature labor); number of hospitalizations during pregnancy; number of doctor office or urgent care visits in the last month of pregnancy; whether woman received help getting pregnant; whether a provider advised a C-section during pregnancy; fear of childbirth measured with the First Baby Study Birth Anticipation Scale (BAS); preference for vaginal delivery; prenatal stress using Perceived Stress Scale (Curry, Campbell, and Christian 1994; Misra, O'Campo, and Strobino 2001), and social support (using MOS Social Support Survey, adapted from Sherbourne and Stewart (1991)). Depression at baseline was measured using the Edinburgh Postnatal Depression Scale (EPDS) and included as a covariate. Gestational age at delivery was taken from birth certificates.

Sociodemographic variables: Pregnancy intention (wanted, mistimed, or unwanted); maternal education (high school, some college, college graduate); age; race/ethnicity (Hispanic, non-Hispanic white, non-Hispanic black or other); whether married or living with partner; and insurance status measured post-partum (public; private/out-of-pocket). Poverty status was measured using the US Census Bureau classification system incorporating both household income and family composition. Women with household incomes <100% of the poverty threshold are classified as “poor”; those with household incomes between 100-200% of the poverty threshold are classified as “near poor”; and those with household incomes at or above 200% of the poverty threshold are classified as “not poor”.

Analytic approach

Propensity scores: rationale

An aim of this study is to understand to what extent women who take antenatal maternity leave differ in meaningful ways from women who work until delivery, and whether this difference consistently relates to the probability of taking ANL. Propensity scores build on traditional matching methods, which allow comparison of women with similar observable characteristics that influence the outcomes, but who differ on treatment.

However, in the presence of many covariates, finding a comparison control unit to match a given treatment unit becomes increasingly difficult. Propensity scores allow inclusion of many covariates without succumbing to the "curse of dimensionality" present with traditional matching methods.

The propensity score method allows conditioning on the probability of treatment, i.e., the likelihood of taking antenatal maternity leave. Having the same distribution of propensity scores in treatment and control groups implies that they have the same distribution of all observed covariates, as if individuals are randomly assigned (Rubin 2001). This strong assumption that the distribution of covariates is the same in treatment and control groups relies on every important factor that determines treatment being observed—unlikely the case with real data. In reality, we can ensure balance between treatment and control groups on observed covariates, but not unobserved covariates (as in randomized experiments). This study benefits from a rich dataset, which allows us to reduce bias with propensity scores; however, we can not ensure balance on unobserved covariates and, therefore, cannot eliminate bias.

Propensity scores can be used in several ways to create balanced treatment and control groups, most commonly: matching, subclassification, weighting, or some combination of these (Rubin 2001).

Statistical analysis

We conducted statistical analyses using Stata/IC version 11.2 (StataCorp, College Station, Texas).

Bivariable analyses – We tested for differences in employment, health, and sociodemographic characteristics between leave-takers and non-leave-takers using t-tests for continuous variables and chi-square tests for categorical variables. We similarly tested for unadjusted differences in outcomes.

Propensity score analysis – We used the following equation to estimate the probability of taking ANL:

$$p(X_i) = E[ANL_i | \bar{X}_i]$$

where $\bar{X}_i$ is a vector of all covariates that we expected to differ by ANL and that preceded the decision to take ANL during this pregnancy. This included: employment status; occupation; pre-pregnancy obesity; weight gain during pregnancy; history of miscarriage; pre-pregnancy hypertension or diabetes; serious health condition during pregnancy; number of hospitalizations; doctor office or urgent care visits in last month of pregnancy; help getting pregnant; fear of childbirth; preference for vaginal birth; maternal stress; social support; baseline depression; gestational age at delivery; pregnancy intention; maternal education; maternal age; maternal race/ethnicity; partnered status; insurance status; and poverty status. The literature on propensity scores contains conflicting rules for covariate selection, namely, whether to include only those variables that simultaneously predict treatment and outcomes (Caliendo and Kopeinig 2005; Caliendo and Kopeinig 2008) or to include all variables except those that are unrelated to the outcome or can not be measured before the assignment of treatments (Rubin and Thomas 1996). We chose to follow the latter and erred on the side of overinclusion.

We estimated the probability of ANL (i.e., the propensity score) using logistic regression. We looked for common support and dropped units with propensity scores outside this region. We then used these scores to match treated women (those taking ANL) with untreated women using radius matching[1]. We tested a range of covariate specifications for estimating the propensity score and radius levels to achieve the best covariate balance in matched samples.

Using matched groups, we used linear probability models to assess the impact of ANL on the likelihood of experiencing a negative delivery outcome. We used Poisson regression to examine the number of negative delivery outcomes since the distribution of outcomes skewed leftward. In all models, we included weights to account for the number of treated units each control observation was matched with (all treated units received a weight of 1). We also included the propensity score as a covariate in addition to other covariates that remained significantly different across treatment groups after matching, or which have been identified in the literature as strong potential confounders. We tested models with and without interactions between ANL and full-time work status and job category, separately.

$$Y_{ni} = \alpha + \tau ANL_i + \beta_1 Pscore_i + \beta_2 \bar{X}_i + u_i \quad (1)$$

$$Y_{ni} = \alpha + \tau ANL_i + \beta_1 Pscore_i + \beta_2 \bar{X}_i + \beta_3 ANL_i * FT_i + u_i \quad (2)$$

$$Y_{ni} = \alpha + \tau ANL_i + \beta_1 Pscore_i + \beta_2 \bar{X}_i + \beta_3 ANL_i * jobtype_i + u_i \quad (3)$$

where Y_{ni} is outcome *n* for person *i*. ANL_i is a dummy variable for whether women stopped working at least two days before delivery (ANL = 1) versus working until the day before or the day of delivery (ANL = 0). τ is the coefficient of interest in each model. $Pscore_i$ is the estimated probability of ANL for each individual *i*. $\bar{X}_i$ is a vector of covariates thought to be potential confounders, even after controlling for the propensity score. β_3 in equations (2) and (3) is the coefficient on each interaction.

Multivariable regression analysis with statistical controls

As a comparison, we estimated the relationship between ANL and our outcomes using standard multivariate models to statistically control for potential confounders. Potential confounders include all of the variables included in the propensity score analyses described above. We used linear probability models to examine whether or not women experienced any negative delivery outcome and Poisson regression for the number of negative delivery outcomes. We tested models with and without interactions between ANL and full-time work status and job category, separately.

[1] E. Leuven and B. Sianesi. (2003). "PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing". http://ideas.repec.org/c/boc/bocode/s432001.html.

Results

Descriptive statistics

875 women (50%) took more than one day of leave; 343 (20%) took one week or more. Overall, the mean ANL was 2.9 (SD 3.5) days and median ANL was 2 days. Among women who took leave, the mean ANL duration was 5.2 (SD 3.5) days and median ANL was 4 days. Figure 2 shows the distribution of ANL for all employed women in the sample.

55% of women experienced a negative delivery outcome (labor induction, labor lasting >24 hours, unplanned C-section, or negative birth experience). Women experienced a mean of .82 (SD .91) negative outcomes and a maximum of 4.

22% of women delivered early term (i.e., 37w0d to 38w6d), 64% delivered full term (i.e., 39w0d to 40w6d), 13% delivered late term (i.e., 41w0d to 41w6d), and 1% delivered post term (i.e., >42w0d). We excluded women who delivered preterm.

Propensity score estimation

Table 1 shows results of the logit models used to estimate the propensity scores. The region of common support ranged from 0.17 to 0.96. Figure 3 shows the distribution of propensity scores by ANL status before and after matching.

Bivariate analyses

Table 2 shows characteristics of the sample by ANL status before and after matching. Before matching, women who took ANL differed from those who did not along several dimensions. Women who took ANL had more hospitalizations during pregnancy compared to women who took no ANL, though in both groups the vast majority of women were never hospitalized. Leave-takers also had more doctor office or urgent care visits in the last month of pregnancy than non-leave-takers. Women who took ANL reported higher prenatal stress and depression than women who did not take ANL, were more likely to report that the pregnancy was mistimed and less likely to report that it was wanted (no difference in the proportion reporting an unwanted pregnancy). Leave-takers were less likely to deliver in the early term period and more likely to deliver in the late term period relative to non-leave-takers. Leave-takers differed on sociodemographic characteristics: they had lower education; were younger; were less likely to be non-Hispanic White and more likely to be non-Hispanic Black (no difference in proportion reporting Hispanic ethnicity); were less likely to be married or living with a partner or privately insured; and were more likely to be near poor than non-leave-takers. Compared to women who took no leave, women who took ANL were significantly more likely to be employed full-time (vs. part-time) and less likely to work in service/labor occupations (vs. professional, managerial, or other occupations). After matching, there were no significant differences between women who took ANL and those who did not along any of these dimensions.

Table 3 shows the distribution of negative delivery outcomes by ANL status, before and after matching. Before matching, women who took ANL were significantly more likely to have any negative delivery outcome (60% vs. 50%), to have more negative delivery outcomes (0.92 vs. 0.72), to have been induced (34% vs. 27%), to have a long labor duration (14% vs. 8%), and to have had an unplanned C-section (27% vs. 20%). They also had longer mean labor duration (15.12 hours vs. 13.29 hours). Leave-takers were not significantly more likely to have reported a negative birth experience. After matching,

significant differences in the number of negative delivery outcomes, the likelihood of long labor duration and unplanned C-section, and mean labor duration remained.

Propensity score analysis

Table 4 shows results for linear probability models predicting the likelihood of experiencing a negative delivery outcome (Panel A) and Poisson regression models predicting the number of negative delivery outcomes (Panel B) using the propensity-score-matched sample with weights to account for the number of treated units each control observation was matched with (all treated units received a weight of 1). Column (1) shows unadjusted models; column (2) shows models adjusted for propensity scores; column (3) shows models adjusted for propensity scores and additional covariates (pre-pregnancy obesity, strong preference for vaginal delivery, perceived stress, employment status, and race/ethnicity); column (4) shows models that include an interaction between employment status and ANL; and column (5) shows models that include an interaction between job type and ANL. In main effects models, ANL was associated with a 4-5 percentage point increase in the likelihood of a negative delivery outcome, though these were all of marginal statistical significance. We detected a significant interaction between ANL and employment status in the model estimating the likelihood of a negative delivery outcome (Figure 4). We observed a 7 percentage point increased in the predicted probability of experiencing any negative delivery outcome among full-time employed women who took ANL, while this relationship was non-significant among part-time employed women. None of the other interactions was significant. We detected a significant increase in the number of negative delivery outcomes experienced in our main effects models. Women who took ANL experienced an average of .14 more negative outcomes than women who took no ANL ($p<.01$), even after controlling for the propensity score and additional covariates. This represents a 16% increase over the mean of .86 outcomes.

Table 5 shows linear probability and linear regression results for labor induction, labor duration, unplanned C-section, and negative birth experience in propensity-score matched groups. All models in Table 5 adjust for the propensity score and pre-pregnancy obesity, strong preference for vaginal delivery, perceived stress, employment status and race/ethnicity. Panel A shows results with only main effects; Panel B shows models with an interaction between employment status and ANL (none of these interaction terms was significant). Women who took ANL had a .05 increase in the likelihood of experiencing labor lasting longer than 24 hours ($p<.01$) and a .06 increase in the likelihood of experiencing an unplanned C-section ($p<.05$). These represent a 42% and 25% increases over the means, respectively. Their mean labor duration was 1.21 hours longer ($p<.05$) than women who did not take ANL. There were not significant associations with labor induction or self-reported negative birth experience.

Multivariable regression analysis with statistical controls

Table 6 shows traditional multivariable linear probability, linear regression, and Poisson regression results for negative delivery outcomes (any and count), labor induction, labor duration, unplanned C-section and self-reported negative birth experience in the unmatched sample, adjusting for pre-pregnancy obesity; whether the woman gained more weight than recommended during pregnancy; prior miscarriage; self-reported history of diabetes or hypertension before pregnancy; presence of medical problems during current

pregnancy (hypertension, high blood pressure, or pre-eclampsia that started during pregnancy; diabetes that started during pregnancy; vaginal bleeding; early or pre-term labor; bed rest or hospitalization because of premature labor); number of hospitalizations during pregnancy; number of doctor office or urgent care visits in the last month of pregnancy; whether woman received help getting pregnant; whether a provider advised a C-section during pregnancy; fear of childbirth measured with the First Baby Study Birth Anticipation Scale (BAS); strong preference for vaginal delivery; prenatal stress using Perceived Stress Scale; social support; baseline EPDS; gestational age at delivery; pregnancy intention; maternal education; maternal age and age squared; race/ethnicity; whether married or living with partner; insurance status and poverty status. Coefficients from these models were similar in magnitude and significance to the results from propensity score analyses, though the interaction between ANL and employment status was non-significant in all models and the main effect relationship between ANL and the likelihood of experiencing any negative delivery outcome became statistically significant ($p<.05$). ANL was also associated with a significant increase in the number of negative delivery outcomes experienced; the likelihood of a long labor duration or unplanned C-section; and an increase in the mean labor duration. Full unadjusted and adjusted results with and without interaction terms are reported in the Appendix (Tables A1a-A1d).

Sensitivity analyses

We re-ran our main models using only the subset of observations with stronger common support since the tails of the propensity score distribution may not have been well-matched (Appendix Table A2). Models in Panel A dropped 10% of the observations who took ANL at which the propensity score density of the women who did not take ANL is the lowest; models in Panel B used a 20% cut-off. As we dropped more observations, the strength of association between ANL and any negative delivery outcome, the number of negative delivery outcomes, and labor induction increased. The strength of association between ANL and both labor duration variables and unplanned C-section decreased. In these models, ANL was not significantly associated with mean labor duration.

Including the three women for whom we estimated exact ANL duration did not change the results (Appendix Table A3). We repeated all analyses excluding women who stopped working exactly two days prior to delivery because these women may have worked until labor began, even if more than one day prior to delivery. The results did not change, except that ANL significantly predicted the likelihood of any negative delivery outcome (Appendix Table A3). We then repeated all analyses excluding women who experienced labor longer than 24 hours. Most results did not change, but the relationship between ANL and labor duration disappeared (Appendix Table A3).

We re-estimated propensity scores and repeated all analyses among only women who delivered past their due date (i.e., 40 completed weeks or 280 days) (Appendix Table A4). Results did not change in this subgroup. We next estimated the duration of ANL relative to women's due date, rather than her actual delivery date, in a subgroup of women who stopped working before their due date and delivered after their due date. With this test, we attempted to identify women who had planned a period of ANL and were able to take it because they delivered after their due date. ANL relative to women's due date was not significantly associated with any outcome (Appendix Table A4).

We re-estimated propensity scores and repeated all analyses comparing women who quit or were fired during pregnancy to women who did not take any ANL. We characterized women who had reported being employed at baseline but no longer employed in the two weeks before delivery as having likely quit their jobs or been fired (N = 829). The results for these women were similar to results for women who took ANL, with some key differences (Appendix Table A5). Women who quit or were fired had a 7 percentage point increase in their predicted probability of experiencing a negative delivery outcome ($p<.01$) and an increase of .21 negative delivery outcomes ($p<.01$). The individual outcomes driving these composite effects differed from women who took ANL. Women who quit or were fired had a 10 percentage point increase in the probability of being induced compared to women who worked until delivery, but none of the associations with other outcomes reached statistical significance.

We repeated all analyses stratified on maternal health, separately estimating propensity scores for healthy and unhealthy women (Appendix Table A6). We characterized as healthy those women who reported no pre-pregnancy hypertension or diabetes, serious health conditions, or hospitalizations during pregnancy; were not advised by a provider during pregnancy to have a C-section; reported <5 doctor office visits in last month of pregnancy; and were not depressed during pregnancy. Among healthy women, ANL did not increase the likelihood of any negative delivery outcome or increase mean labor duration. Among unhealthy women, all coefficients increased, though smaller sample sizes prevent detection of statistical significance in some models. Unhealthy women who took ANL had a 10 percentage point increase in the likelihood of experiencing any negative delivery outcome ($p<.01$) and experienced .26 more negative outcomes ($p<.01$) than unhealthy women who did not take ANL. Unhealthy women who took ANL were marginally more likely to experience long labor duration ($p <.10$) and experienced an 11 percentage point increase in the likelihood of an unplanned C-section ($p<.01$).

To better understand how the duration of ANL impacts negative delivery outcomes, we examined ANL as a categorical variable in standard multivariable regression models (Appendix Table A7). We divided ANL into five categories: no ANL, two days, three to six days, one week, and more than one week ANL. We repeated all traditional multivariable regression analyses using the unmatched sample and this categorical ANL variable. These results suggest that long ANL (more than one week) did not drive the observed associations. ANL longer than one week was not significantly associated with any outcome after adjusting for covariates. Similarly, very short ANL does not appear to drive the results, with the notable exception of labor duration. Women who took two days of ANL reported significantly longer labor duration than women who took no ANL, but this could reflect women who stopped working because they were in labor, rather than taking ANL before labor began. We also observe significantly longer labor duration among women who took three to six days ANL. Since labor longer than 48 hours occurred in less than 1% of the sample, ANL likely preceded labor in this group. Women who stopped working 3-6 days before delivery were significantly more likely to experience any negative delivery outcome, labor longer than 24 hours, or unplanned C-section and experienced more negative delivery outcomes. Women who stopped working one week before delivery were significantly more likely to have been induced and experienced more negative delivery outcomes.

Discussion

In this sample of nulliparous employed women in Pennsylvania, women who took ANL experienced more negative delivery outcomes than women who took no leave. In particular, women who took ANL were significantly more likely to have an unplanned C-section and to have long labor duration. Women who took ANL had 25% higher predicted probability of unplanned C-section. ANL was significantly associated with the likelihood of a negative delivery outcome among full-time employed women. Our results are robust and held up to a range of sensitivity analyses. Finding that ANL negatively impacts maternal health suggests that even after controlling for an extensive set of observable employment, health, and sociodemographic characteristics, women who take ANL continue to differ in unobserved characteristics that lead to negative delivery outcomes.

Additional evidence for ongoing selection comes from sensitivity analyses comparing women who quit or were fired to women who did not take any ANL. These analyses yielded similar results to the main findings, indicating that women who take ANL and women who quit or were fired share characteristics that lead to negative delivery outcomes— likely unmeasured prenatal health conditions and/or difficult jobs. Stratifying on maternal health further illuminated ongoing selection. Analyses including only healthy women yielded no significant associations, while the strength of association increased when limited to unhealthy women.

Like all but a few U.S. states, Pennsylvania does not have any paid maternity leave laws. Eligible Pennsylvania women who wish to stop working during pregnancy can take advantage of the federal Family and Medical Leave Act, which provides up to 12 weeks of unpaid leave for, among other things, one's own illness. Women may choose not use this leave for two reasons: it is entirely unpaid and using it during pregnancy reduces the amount of leave available after childbirth. Pennsylvania women may take antenatal leave through temporary disability programs, where available. The Pregnancy Discrimination Act of 1978 mandates that employers treat pregnant employees the same as they would any other temporarily disabled employee; if an employer provides leave for any employee with a short-term disability, he or she must provide the same leave for a woman on maternity leave or disabled by pregnancy. Five states (California, Hawaii, New Jersey, New York, Rhode Island) and Puerto Rico have state temporary disability insurance laws that provide income replacement to employees on leave; Pennsylvania has no such state law, though individual employers may offer relevant benefits. With limited options for taking leave during (and after) pregnancy, women in Pennsylvania may only take ANL when medically necessary, as the results of this study imply.

Our study benefits from a rich dataset with extensive questions on pregnancy, labor and delivery, and postpartum. Propensity scores facilitated exploitation of this detailed information while avoiding the "curse of dimensionality" present in traditional methods. Still, the narrow set of employment characteristics limits our ability to remove bias. Furthermore, we cannot determine the reason women took leave, and whether they intended the amount of leave they took (i.e., they delivered too soon to take the leave they had planned or took leave because they delivered later than expected). The First Baby Study includes a diverse sample of nulliparous women in Pennsylvania. Results cannot

necessarily be generalized to multiparous women, nor to nulliparas nationally, though Pennsylvania's lack of antenatal leave offerings compares to that of all but a handful of U.S. states.

Primary C-sections, even among low-risk women, continue to increase in the United States, leading to much deliberation about how to curb this trend (Spong et al. 2012). In addition to understanding the institutional factors driving this increase, health professionals and social scientists need to understand what other modifiable factors influence C-sections, as well as other negative delivery outcomes. Work characteristics of pregnant women, including maternity leave-taking, remain largely unexplored as potential influences on maternal health. In the United States, 80% of first-time mothers who were employed during pregnancy worked into their last month of pregnancy (Laughlin 2011). Still, almost no research explores the health impact of antenatal leave on obstetric or other outcomes.

Not finding support for a protective effect of ANL does not imply that pregnant women should be encouraged to work until delivery or that laws and policies allowing leave do not help. In fact, the strong selection into ANL in this sample of Pennsylvania women suggests that women who take leave suffer from particularly difficult pregnancies or jobs and could benefit from expanded availability of leave. Many women in the United States do not have access to paid, job-protected leave so must forgo income or possibly risk their jobs by taking ANL. This study found an impact of ANL on negative delivery outcomes resulting from women who had to take leave, rather than women who took leave for protective reasons (i.e., to rest and prepare for delivery). The appropriate counterfactual for women who took leave should be women who needed leave but did not take it, rather than the healthy women who did not take leave. Future research should attempt to examine this relationship.

References

Bonzini, M., K. T. Palmer, D. Coggon, M. Carugno, A. Cromi, and M. M. Ferrario. 2011. "Shift work and pregnancy outcomes: a systematic review with meta-analysis of currently available epidemiological studies." *BJOG: An International Journal of Obstetrics & Gynaecology* 118(12): 1429-37.

Caliendo, M. and S. Kopeinig. 2005. "Some Practical Guidance for the Implementation of Propensity Score Matching. IZA Discussion Paper No. 1588. The Institute for the Study of Labor. Bonn."

Caliendo, M. and S. Kopeinig. 2008. "Some practical guidance for the implementation of propensity score matching." *Journal of economic surveys* 22(1): 31-72.

Ceron-Mireles, P., S. D. Harlow, and C. Sanchez-Carrillo. 1996. "The risk of prematurity and small-for-gestational-age birth in Mexico City: the effects of working conditions and antenatal leave." *American Journal of Public Health* 86(6): 825-31.

Chien, L. Y. and Y. L. Ko. 2004. "Fatigue during pregnancy predicts caesarean deliveries." *Journal of advanced nursing* 45(5): 487-94.

Curry, M. A., R. A. Campbell, and M. Christian. 1994. "Validity and reliability testing of the prenatal psychosocial profile." *Research in nursing & health* 17(2): 127-35.

Del Bono, E., J. Ermisch, and M. Francesconi. 2012. "Intrafamily resource allocations: a dynamic structural model of birth weight." *Journal of Labor Economics* 30(3): 657-706.
Elvander, C., S. Cnattingius, and K. H. Kjerulff. 2013. "Birth experience in women with low, intermediate or high levels of fear: findings from the first baby study." *Birth* 40(4): 289-96.
Fortier, I., S. Marcoux, and J. Brisson. 1995. "Maternal work during pregnancy and the risks of delivering a small-for-gestational-age or preterm infant." *Scandinavian journal of work, environment & health*: 412-18.
Frazier, L. M., A. L. Golbeck, and L. Lipscomb. 2001. "Medically recommended cessation of employment among pregnant women in Georgia." *Obstetrics & Gynecology* 97(6): 971-75.
Guendelman, S., M. Pearl, S. Graham, V. Angulo, and M. Kharrazi. 2006. "Utilization of pay-in antenatal leave among working women in Southern California." *Maternal and Child Health Journal* 10(1): 63-73.
Guendelman, S., M. Pearl, S. Graham, A. Hubbard, N. Hosang, and M. Kharrazi. 2009. "Maternity leave in the ninth month of pregnancy and birth outcomes among working women." *Women's Health Issues* 19(1): 30-37.
Hung, S., D. R. Morrison, L. A. Whittington, and S. B. Fein. 2002. "Prepartum work, job characteristics, and risk of cesarean delivery." *Birth* 29(1): 10-17.
Kinsey, C. B., K. Baptiste-Roberts, J. Zhu, and K. H. Kjerulff. 2013. "Effect of previous miscarriage on the maternal birth experience in the First Baby Study." *Journal of Obstetric, Gynecologic, & Neonatal Nursing* 42(4): 442-50.
Kjerulff, K. H., D. L. Velott, J. Zhu, C. H. Chuang, M. M. Hillemeier, I. M. Paul, and J. T. Repke. 2013. "Mode of first delivery and women's intentions for subsequent childbearing: Findings from the First Baby Study." *Paediatric and Perinatal Epidemiology* 27(1): 62-71.
Laughlin, L. 2011. "Maternity leave and employment patterns: 2006-2008." *Current Population Reports.* pp. 70-128. Washington, D.C. : U.S. Census Bureau.
Lee, K. A. and C. L. Gay. 2004. "Sleep in late pregnancy predicts length of labor and type of delivery." *American journal of obstetrics and gynecology* 191(6): 2041-46.
Mamelle, N., I. Bertucat, and F. Munoz. 1989. "Pregnant women at work: rest periods to prevent preterm birth?" *Paediatric and Perinatal Epidemiology* 3(1): 19-28.
Mayberry, L. J., S. Gennaro, L. Strange, M. Williams, and A. De. 1999. "Maternal fatigue: Implications of second stage labor nursing care." *Journal of Obstetric, Gynecologic, & Neonatal Nursing* 28(2): 175-81.
Misra, D. P., P. O'Campo, and D. Strobino. 2001. "Testing a sociomedical model for preterm delivery." *Paediatric and Perinatal Epidemiology* 15(2): 110-22.
Mozurkewich, E., B. Luke, M. Avni, and F. Wolf. 2000. "Working conditions and adverse pregnancy outcome: a meta-analysis." *Obstetrics & Gynecology* 95(4): 623-35.
Rubin, D. B. 2001. "Using propensity scores to help design observational studies: application to the tobacco litigation." *Health Services and Outcomes Research Methodology* 2(3-4): 169-88.
Rubin, D. B. and N. Thomas. 1996. "Matching using estimated propensity scores: relating theory to practice." *Biometrics*: 249-64.
Saurel-Cubizolles, M. J., J. Zeitlin, N. Lelong, E. Papiernik, G. C. Di Renzo, and G. Bréart. 2004. "Employment, working conditions, and preterm birth: results from the Europop case-control survey." *Journal of epidemiology and community health* 58(5): 395-401.

Sherbourne, C. D. and A. L. Stewart. 1991. "The MOS social support survey." *Social Science & Medicine* 32(6): 705-14.
Spong, C. Y., V. Berghella, K. D. Wenstrom, B. M. Mercer, and G. R. Saade. 2012. "Preventing the First Cesarean Delivery: Summary of a Joint Eunice Kennedy Shriver National Institute of Child Health and Human Development, Society for Maternal-Fetal Medicine, and American College of Obstetricians and Gynecologists Workshop." *Obstetrics & Gynecology* 120(5): 1181-93.
Sydsjö, A., J. Brynhildsen, K. E. Selling, A. Josefsson, and G. Sydsjö. 2006. "Influence of rest during pregnancy on birth weight in working women." *Obstetrics & Gynecology* 107(5): 991-96.
Wüst, M. 2014. "MATERNAL EMPLOYMENT DURING PREGNANCY AND BIRTH OUTCOMES: EVIDENCE FROM DANISH SIBLINGS." *Health economics*.
Xu, Q., L. Séguin, and L. Goulet. 2002. "Effet bénéfique d'un arrêt du travail avant l'accouchement." *Canadian Journal of Public Health/Revue Canadienne de Sante'e Publique*: 72-77.
Yaktine, A. L. and K. M. Rasmussen. 2009. *Weight Gain During Pregnancy:: Reexamining the Guidelines*: National Academies Press.

Tables and Figures

Table 1: Logit regression estimation of propensity scores in a sample of employed, nulliparous women in Pennsylvania (N = 1,740).

	Likelihood of ANL (propensity score) Coeff. [95% CI]
Pre-pregnancy obesity	0.02 [-0.25 - 0.29]
Gained > recommended	0.12 [-0.09 - 0.32]
Prior miscarriage	-0.05 [-0.33 - 0.23]
Pre-pregnancy hypertension or diabetes	0.57 [-0.02 - 1.16]
Serious health condition during pregnancy	-0.12 [-0.38 - 0.13]
Mean number of hospitalizations	0.15 [-0.07 - 0.38]
Doctor office or urgent care visits in last month of pregnancy (ref: 2-4 visits)	
0-1 visits	0.12 [-0.3 - 0.54]
5+ visits	0.27 [0.01 - 0.54]*
Help getting pregnant	-0.10 [-0.42 - 0.22]
Provider advised C-sec during pregnancy	0.34 [-0.05 - 0.74]
Birth Anticipation Scale	0.00 [-0.02 - 0.03]
Strong preference for vaginal birth	-0.16 [-0.46 - 0.14]
Maternal stress	0.02 [-0.01 - 0.05]
Low social support	0.00 [-0.21 - 0.21]
Baseline EPDS	0.04 [0.01 - 0.08]*
Gestational age at delivery (ref: full term)	
Early term	-0.91 [-1.17 - -0.65]**
Late term	1.07 [0.73 - 1.41]**
Post term	0.67 [-0.42 - 1.76]
Service or labor (ref: professional, managerial, or clerical)	-0.05 [-0.30 - 0.20]
Full-time employed (ref: part-time)	-0.72 [-1.08 - -0.36]**
Pregnancy intention (ref: wanted)	
Mistimed	-0.19 [-0.47 - 0.08]
Unwanted	-0.74 [-1.80 - 0.33]
Mother's age at baseline	-0.13 [-0.51 - 0.26]
Mother's age at baseline squared	0.00 [-0.01 - 0.01]
Maternal education (ref: college grad)	
High school graduate or less	-0.09 [-0.54 - 0.37]
Some college or vocational programs	-0.16 [-0.43 - 0.11]
Maternal race/ethnicity (ref: Non-Hispanic White)	
Hispanic	0.00 [-0.64 - 0.63]
Non-Hispanic Black or Other	0.17 [-0.31 - 0.65]
Unmarried or not living with partner (ref: married or living with partner)	0.01 [-0.51 - 0.54]
Public insurance (ref: private or OOP)	0.48 [-0.01 - 0.96]
Poverty status (ref: non-poor)	
Poor	-0.09 [-0.69 - 0.52]
Near poor	0.44 [-0.08 - 0.96]
Constant	2.54 [-3.00 - 8.07]
Observations	1740

Pseudo R-squared	0.09
Region of common support	[0.17, 0.96]

Abbreviations: CI, confidence interval; EPDS, Edinburgh Postnatal Depression Scale; OOP, out of pocket.

Notes: Coefficients and 95% confidence intervals. Serious health condition defined as hypertension, high blood pressure, or pre-eclampsia that started during pregnancy; diabetes that started during pregnancy; vaginal bleeding; early or pre-term labor; bed rest or hospitalization because of premature labor. The First Baby Study Birth Anticipation Scale (BAS) measures fear of childbirth. Maternal stress measured with the Perceived Stress Scale. Early term refers to gestations of 37w0d to 38w6d; full term refers to 39w0d to 40w6d, late term refers to 41w0d to 41w6d, and post term refers to >42w0d. We excluded preterm deliveries (i.e., <37w0d) from the sample. Poverty status was measured using the US Census Bureau classification system incorporating both household income and family composition. Women with household incomes <100% of the poverty threshold are classified as "poor"; those with household incomes between 100-200% of the poverty threshold are classified as "near poor"; and those with household incomes at or above 200% of the poverty threshold are classified as "not poor".

* significant at 5%; ** significant at 1%.

Table 2: Characteristics of the sample by ANL status among employed, nulliparous women in Pennsylvania, before and after matching (N = 1,740).

	Unmatched sample			Matched sample		
	ANL Mean or %	No ANL Mean or %	*p*	ANL Mean or %	No ANL Mean or %	*p*
Pre-pregnancy obesity	0.20	0.19	0.67	0.20	0.17	0.12
Gained > recommended	0.54	0.52	0.24	0.54	0.55	0.78
Prior miscarriage	0.15	0.17	0.49	0.15	0.14	0.35
Pre-pregnancy hypertension or diabetes	0.04	0.03	0.11	0.04	0.03	0.48
Serious health condition during pregnancy	0.22	0.23	0.38	0.22	0.21	0.72
Mean number of hospitalizations	0.18	0.12	0.01	0.18	0.16	0.29
Doctor office or urgent care visits in last month of pregnancy						
0-1 visits	0.07	0.06	0.60	0.07	0.08	0.23
2-4 visits	0.72	0.77	0.01	0.72	0.70	0.42
5+ visits	0.21	0.17	0.02	0.21	0.21	0.77
Help getting pregnant	0.11	0.14	0.05	0.11	0.10	0.89
Provider advised C-sec during pregnancy	0.08	0.06	0.13	0.08	0.07	0.56
Birth Anticipation Scale	16.96	16.54	0.05	16.96	16.94	0.90
Strong preference for vaginal birth	0.13	0.14	0.60	0.13	0.15	0.14
Maternal stress	18.71	17.79	<0.01	18.71	19.07	0.12
Low social support	0.42	0.39	0.19	0.42	0.41	0.72
Baseline EPDS	5.98	5.20	<0.01	5.98	5.92	0.75
Gestational age at delivery						
Early term	0.15	0.30	<0.01	0.15	0.14	0.76
Full term	0.65	0.63	0.41	0.65	0.67	0.31
Late term	0.19	0.06	<0.01	0.19	0.18	0.37
Post term	0.01	0.01	0.09	0.01	0.01	0.84
Service or labor (vs. professional, managerial, or clerical)	0.82	0.93	<0.01	0.82	0.81	0.53
Full-time employed (vs. part-time)	0.33	0.27	<0.01	0.33	0.33	0.92
Pregnancy intention						
Wanted	0.71	0.77	0.01	0.71	0.70	0.48
Mistimed	0.28	0.22	<0.01	0.28	0.29	0.46
Unwanted	0.01	0.01	0.79	0.01	0.01	0.88

Mother's age at baseline	27.82	28.83	<0.01	27.82	27.72	0.60
Maternal education						
High school graduate or less	0.10	0.06	<0.01	0.10	0.13	0.17
Some college or vocational programs	0.24	0.23	0.38	0.24	0.23	0.55
College grad	0.65	0.71	0.01	0.65	0.64	0.70
Maternal race/ethnicity						
Hispanic	0.03	0.02	0.40	0.03	0.03	0.51
Non-Hispanic White	0.90	0.93	0.01	0.90	0.89	0.45
Non-Hispanic Black or Other	0.07	0.04	0.02	0.07	0.09	0.21
Unmarried or not living with partner (vs. married or living with partner)	0.08	0.04	<0.01	0.08	0.10	0.16
Public insurance (vs. private or OOP)	0.13	0.05	<0.01	0.13	0.12	0.38
Poverty status						
Poor	0.04	0.03	0.06	0.04	0.05	0.35
Near poor	0.08	0.03	<0.01	0.08	0.08	0.94
Not poor	0.87	0.94	<0.01	0.87	0.86	0.60

Abbreviations: ANL, antenatal leave; EPDS, Edinburgh Postnatal Depression Scale; OOP, out of pocket.

Notes: p-values from t-tests. Serious health condition defined as hypertension, high blood pressure, or pre-eclampsia that started during pregnancy; diabetes that started during pregnancy; vaginal bleeding; early or pre-term labor; bed rest or hospitalization because of premature labor. The First Baby Study Birth Anticipation Scale (BAS) measures fear of childbirth. Maternal stress measured with the Perceived Stress Scale. Early term refers to gestations of 37w0d to 38w6d; full term refers to 39w0d to 40w6d, late term refers to 41w0d to 41w6d, and post term refers to >42w0d. We excluded preterm deliveries (i.e., <37w0d) from the sample. Poverty status was measured using the US Census Bureau classification system incorporating both household income and family composition. Women with household incomes <100% of the poverty threshold are classified as “poor”; those with household incomes between 100-200% of the poverty threshold are classified as “near poor”; and those with household incomes at or above 200% of the poverty threshold are classified as “not poor”. In matched models, we matched women who took ANL with one or more women who did not take ANL on estimated propensity scores within a 0.04 caliper radius.

Table 3: Distribution of negative delivery outcomes by ANL status among employed, nulliparous women in Pennsylvania, before and after matching (N = 1,740).

	Unmatched sample			Matched sample		
	ANL	No ANL	*p*	ANL	No ANL	*p*
	Mean or %	Mean or %		Mean or %	Mean or %	
Any negative delivery outcome	0.60	0.50	<0.01	0.60	0.56	0.09
Number of negative delivery outcomes	0.92	0.72	<0.01	0.92	0.80	0.01
Labor induced	0.34	0.27	<0.01	0.34	0.33	0.57
Labor >24 hours	0.14	0.08	<0.01	0.14	0.10	<0.01
Hours in labor	15.12	13.29	<0.01	15.12	13.81	0.01
Unplanned C-section	0.27	0.20	<0.01	0.27	0.21	<0.01
Negative birth experience	0.20	0.18	0.19	0.20	0.18	0.40

Abbreviations: ANL, antenatal leave; EPDS, Edinburgh Postnatal Depression Scale; OOP, out of pocket.

Notes: p-values from t-tests. Serious health condition defined as hypertension, high blood pressure, or pre-eclampsia that started during pregnancy; diabetes that started during pregnancy; vaginal bleeding; early or pre-term labor; bed rest or hospitalization because of premature labor. The First Baby Study Birth Anticipation Scale (BAS) measures fear of childbirth. Maternal stress measured with the Perceived Stress Scale. Early term refers to gestations of 37w0d to 38w6d; full term refers to 39w0d to 40w6d, late term refers to 41w0d to 41w6d, and post term refers to >42w0d. We excluded preterm deliveries (i.e., <37w0d) from the sample. Poverty status was measured using the US Census Bureau classification system incorporating both household income and family composition. Women with household incomes <100% of the poverty threshold are classified as "poor"; those with household incomes between 100-200% of the poverty threshold are classified as "near poor"; and those with household incomes at or above 200% of the poverty threshold are classified as "not poor". In matched models, we matched women who took ANL with one or more women who did not take ANL on estimated propensity scores within a 0.04 caliper radius. Negative delivery outcomes include labor induction, labor lasting more than 24 hours, unplanned C-section, and self-reported negative birth experience.

Table 4: Multivariable linear probability and Poisson regression results for negative delivery outcomes (any and count) in propensity-score matched groups among employed, nulliparous women in Pennsylvania (N = 1,740). Coefficients and 95% confidence intervals.

	(1)	(2)	(3)	(4)	(5)
Panel A	Any negative delivery outcome				
	Coeff. [95% CI]	Coeff. [95% CI]	Coeff. [95% CI]	Coeff. [95% CI]	Coeff. [95% CI]
ANL (≥2 days)	0.04 [-0.00 - 0.09]	0.05 [0.00 - 0.09]*	0.04 [-0.00 - 0.09]	-0.06 [-0.17 - 0.05]	0.05 [-0.01 - 0.10]
Full-time employed				0.01 [-0.08 - 0.10]	
Full-time employed x ANL (≥2 days)				0.13 [0.01 - 0.24]*	
Service or labor occupation					-0.06 [-0.13 - 0.01]
Service or labor x ANL (≥2 days)					-0.02 [-0.12 - 0.08]
Propensity score	N	Y	Y	Y	Y
Additional covariates	N	N	Y	Y	Y
Full-time work interaction	N	N	N	Y	N
Service or labor interaction	N	N	N	N	Y
Mean of negative delivery outcomes	0.58	0.58	0.58	0.58	0.58
Weighted observations	1747	1747	1747	1747	1747
R-squared/Pseudo R-squared	0.00	0.03	0.04	0.04	0.04
Panel B	Number of negative delivery outcomes				
	Coeff. [95% CI]	Coeff. [95% CI]	Coeff. [95% CI]	Coeff. [95% CI]	Coeff. [95% CI]
ANL (≥2 days)	0.15 [0.04 - 0.25]**	0.15 [0.05 - 0.25]**	0.14 [0.03 - 0.24]**	0.14 [-0.10 - 0.38]	0.13 [0.00 - 0.25]*
Full-time employed				0.23 [0.02 - 0.44]*	
Full-time employed x ANL (≥2 days)				0 [-0.27 - 0.27]	
Service or labor occupation					-0.13 [-0.30 - 0.04]
Service or labor x ANL (≥2 days)					0.04 [-0.18 - 0.26]
Propensity score	N	Y	Y	Y	Y
Additional covariates	N	N	Y	Y	Y
Full-time work interaction	N	N	N	Y	N
Service or labor interaction	N	N	N	N	Y
Mean of negative delivery outcomes	0.86	0.86	0.86	0.86	0.86
Weighted observations	1747	1747	1747	1747	1747

R-squared/Pseudo R-squared	0.00	0.01	0.02	0.02	0.02

Abbreviations: CI, confidence interval; ANL, antenatal leave.

Notes: Coefficients and 95% confidence intervals for linear probability models (Panel A) and linear regression models (Panel B). All models are weighted to account for the number of units each observation was matched with. Additional covariates include pre-pregnancy obesity, strong preference for vaginal delivery, perceived stress, employment status and race/ethnicity. Negative delivery outcomes include labor induction, labor lasting more than 24 hours, unplanned C-section, and self-reported negative birth experience.

* significant at 5%; ** significant at 1%.

Table 5: Multivariable linear probability and linear regression results for labor induction, labor duration, unplanned C-section, and negative birth experience in propensity-score matched groups among employed, nulliparous women in Pennsylvania (N = 1,740). Coefficients and 95% confidence intervals.

	Labor induced	Labor >24 hours	Hours in labor	Unplanned C-section	Negative birth experience
	Coeff. [95% CI]	Coeff. [95% CI]	Coeff. [95% CI]	Coeff. [95% CI]	Coeff. [95% CI]
Panel A: Main effects only					
ANL (≥2 days)	0.02 [-0.03 - 0.06]	0.05 [0.02 - 0.08]**	1.21 [0.25 - 2.16]*	0.06 [0.02 - 0.10]**	0.02 [-0.02 - 0.06]
Propensity score	Y	Y	Y	Y	Y
Additional covariates	Y	Y	Y	Y	Y
Mean of negative delivery outcomes	0.33	0.12	14.54	0.24	0.19
Weighted observations	1747	1618	1618	1648	1739
R-squared/Pseudo R-squared	0.06	0.01	0.01	0.03	0.02
Panel B: Full-time work interaction					
ANL (≥2 days)	-0.02 [-0.12 - 0.08]	0.03 [-0.05 - 0.10]	0.97 [-1.33 - 3.26]	0.04 [-0.06 - 0.13]	0.09 [0.00 - 0.18]*
Full-time employed	0.08 [-0.01 - 0.16]	0.03 [-0.03 - 0.10]	2.39 [0.45 - 4.32]*	0.05 [-0.03 - 0.13]	0.04 [-0.03 - 0.12]
Full-time employed x ANL (≥2 days)	0.05 [-0.07 - 0.16]	0.02 [-0.06 - 0.11]	0.29 [-2.24 - 2.83]	0.03 [-0.08 - 0.14]	-0.09 [-0.19 - 0.01]
Propensity score	Y	Y	Y	Y	Y
Additional covariates	Y	Y	Y	Y	Y
Mean of negative delivery outcomes	0.33	0.12	14.54	0.24	0.19
Weighted observations	1747	1618	1618	1648	1739
R-squared/Pseudo R-squared	0.06	0.01	0.01	0.03	0.03

Abbreviations: CI, confidence interval; ANL, antenatal leave.

Notes: Coefficients and 95% confidence intervals for linear probability models (labor induction, long labor duration, unplanned C-section, and negative birth experience) and linear regression models (hours in labor). Panel A includes only main effects; Panel B includes an interaction between ANL and full-time employment. All models are weighted to account for the number of units each observation was matched with. Additional covariates include pre-pregnancy obesity, strong preference for vaginal delivery, perceived stress, employment status and race/ethnicity.

* significant at 5%; ** significant at 1%

Table 6: Multivariable linear probability, linear regression, and Poisson regression results for negative delivery outcomes (any and count), labor induction, labor duration, unplanned C-section and self-reported negative birth experience in unmatched groups among employed, nulliparous women in Pennsylvania (N = 1,740). Coefficients and 95% confidence intervals.

	Any negative delivery outcome	Number of negative delivery outcomes	Labor induced	Labor >24 hours	Labor duration	Unplanned cesarean	Negative birth experience
	Coeff. [95% CI]	Coeff. [95% CI]	Coeff. [95% CI]	Coeff. [95% CI]	Coeff. [95% CI]	Coeff. [95% CI]	Coeff. [95% CI]
ANL (≥2 days)	0.06 [0.01 - 0.11]*	0.16 [0.04 - 0.27]**	0.03 [-0.01 - 0.08]	0.04 [0.01 - 0.08]*	1.39 [0.40 - 2.39]**	0.06 [0.02 - 0.10]**	0.01 [-0.03 - 0.05]
Mean outcome	0.55	0.82	0.31	0.11	14.20	0.24	0.19
Observations	1740	1740	1740	1606	1606	1635	1730
R-squared/ Pseudo R-squared	0.10	0.04	0.11	0.03	0.04	0.07	0.10

Abbreviations: CI, confidence interval; ANL, antenatal leave; EPDS, Edinburgh Postnatal Depression Scale; OOP, out of pocket.

Notes: Coefficients and 95% confidence intervals. All models control for pre-pregnancy obesity; whether the woman gained more weight than recommended during pregnancy; prior miscarriage; self-reported history of diabetes or hypertension before pregnancy; presence of medical problems during current pregnancy (hypertension, high blood pressure, or pre-eclampsia that started during pregnancy; diabetes that started during pregnancy; vaginal bleeding; early or pre-term labor; bed rest or hospitalization because of premature labor); number of hospitalizations during pregnancy; number of doctor office or urgent care visits in the last month of pregnancy; whether woman received help getting pregnant; whether a provider advised a C-section during pregnancy; fear of childbirth measured with the First Baby Study Birth Anticipation Scale (BAS); strong preference for vaginal delivery; prenatal stress using Perceived Stress Scale; social support; baseline EPDS; gestational age at delivery; pregnancy intention; maternal education; maternal age and age squared; race/ethnicity; whether married or living with partner; insurance and poverty status. Serious health condition defined as hypertension, high blood pressure, or pre-eclampsia that started during pregnancy; diabetes that started during pregnancy; vaginal bleeding; early or pre-term labor; bed rest or hospitalization because of premature labor. The First Baby Study Birth Anticipation Scale (BAS) measures fear of childbirth. Maternal stress measured with the Perceived Stress Scale. Early term refers to gestations of 37w0d to 38w6d; full term refers to 39w0d to 40w6d, late term refers to 41w0d to 41w6d, and post term refers to >42w0d. We excluded preterm deliveries (i.e., <37w0d) from the sample. Poverty status was measured using the US Census Bureau classification system incorporating both household income and family composition. Women with household incomes <100% of the poverty threshold are classified as "poor"; those with household incomes between 100-200% of the poverty threshold are classified as "near poor"; and those with household incomes at or above 200% of the poverty threshold are classified as "not poor". Negative delivery outcomes include labor induction, labor lasting more than 24 hours, unplanned C-section, and self-reported negative birth experience.

* significant at 5%; ** significant at 1%

Figure 1: Theoretical relationships between ANL and maternal health – protection and selection pathways.

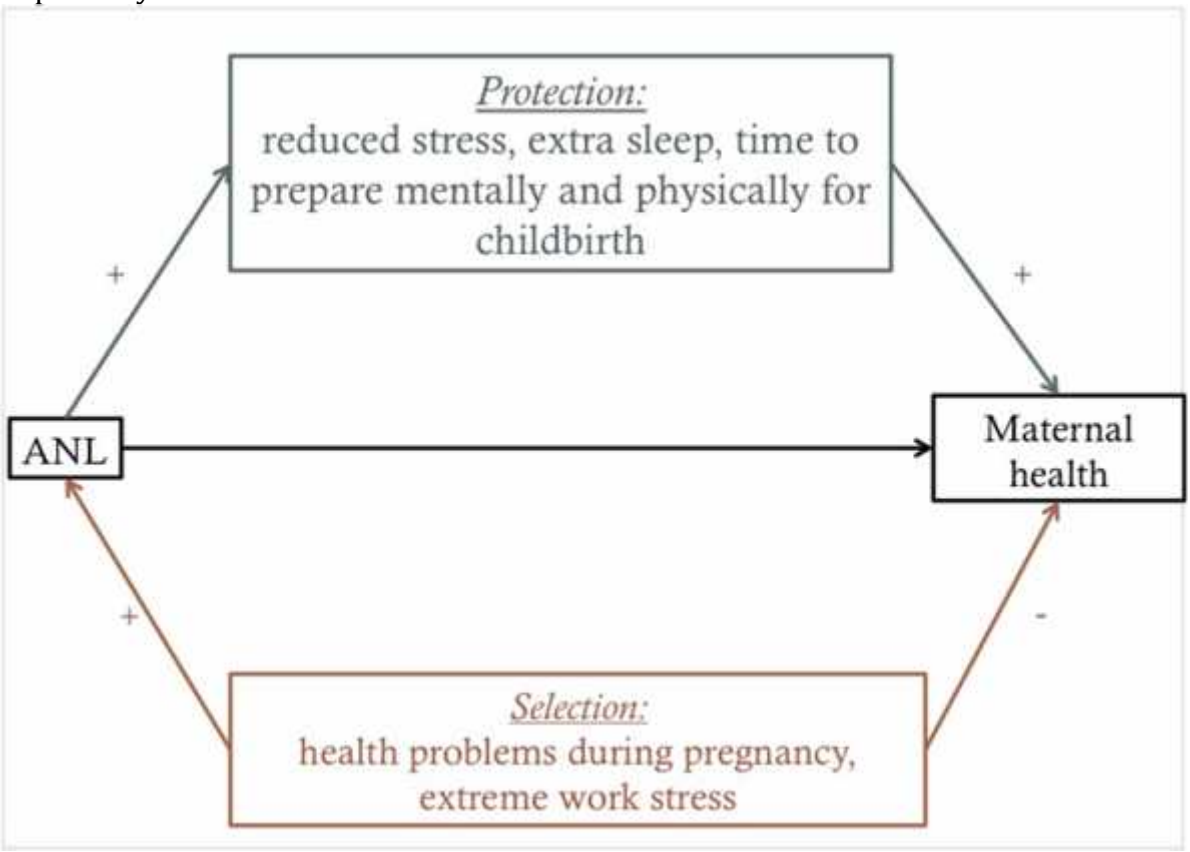
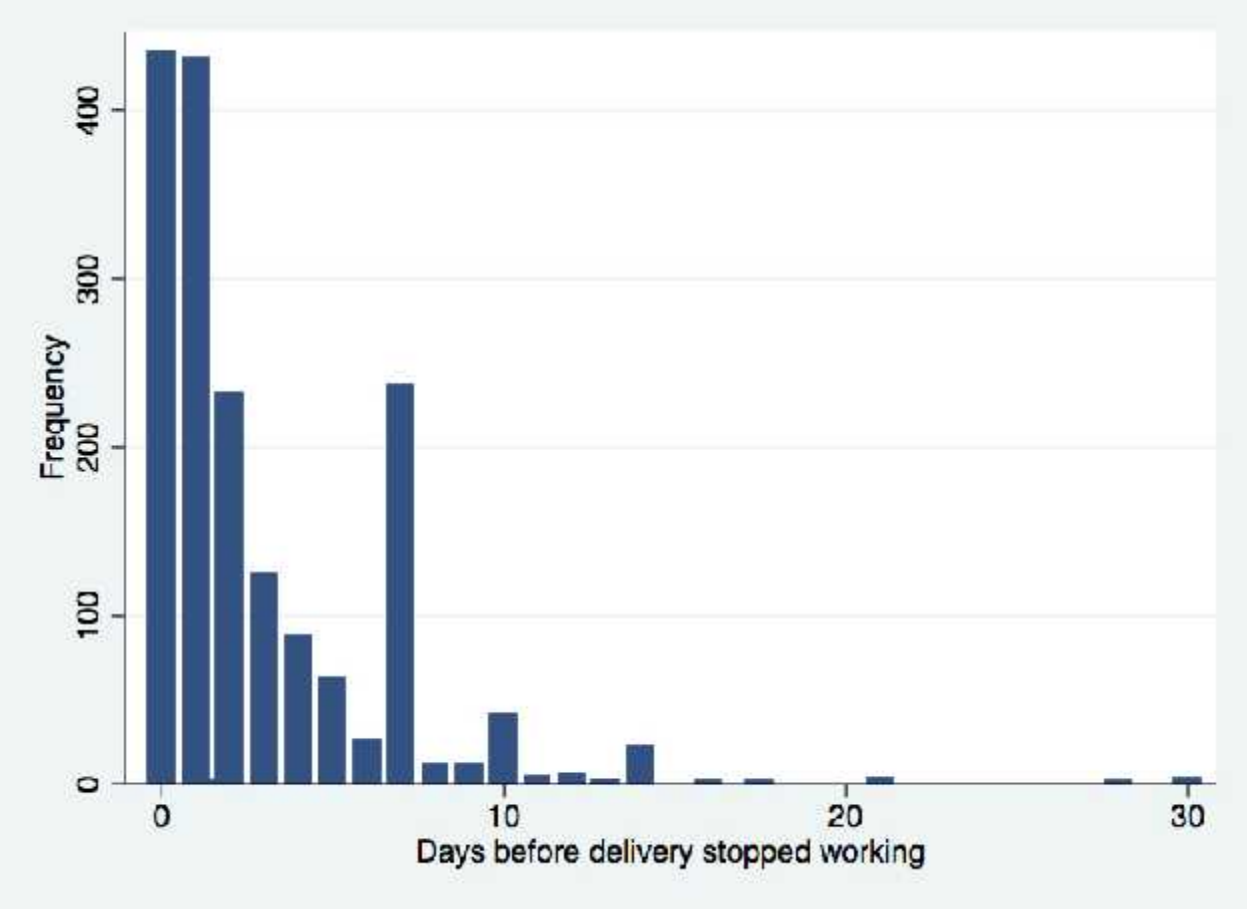


Figure 2: Distribution of ANL (in days) among employed, nulliparous women in Pennsylvania (N = 1,740).



Notes: Includes women who delivered >36w6d and stopped working <31 days before delivery.

Figure 3: Propensity score distribution by ANL before and after matching among employed, nulliparous women in Pennsylvania (N = 1,740).

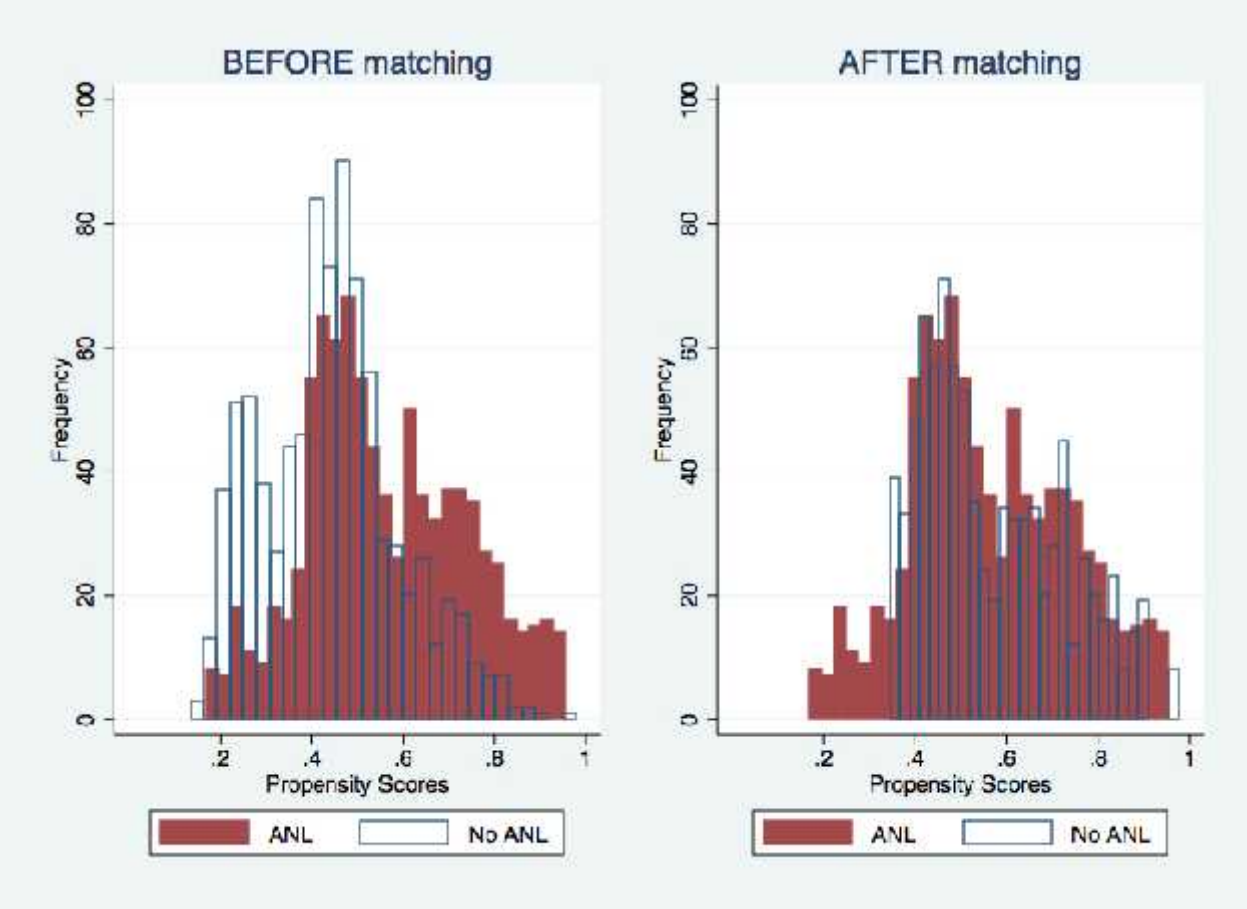
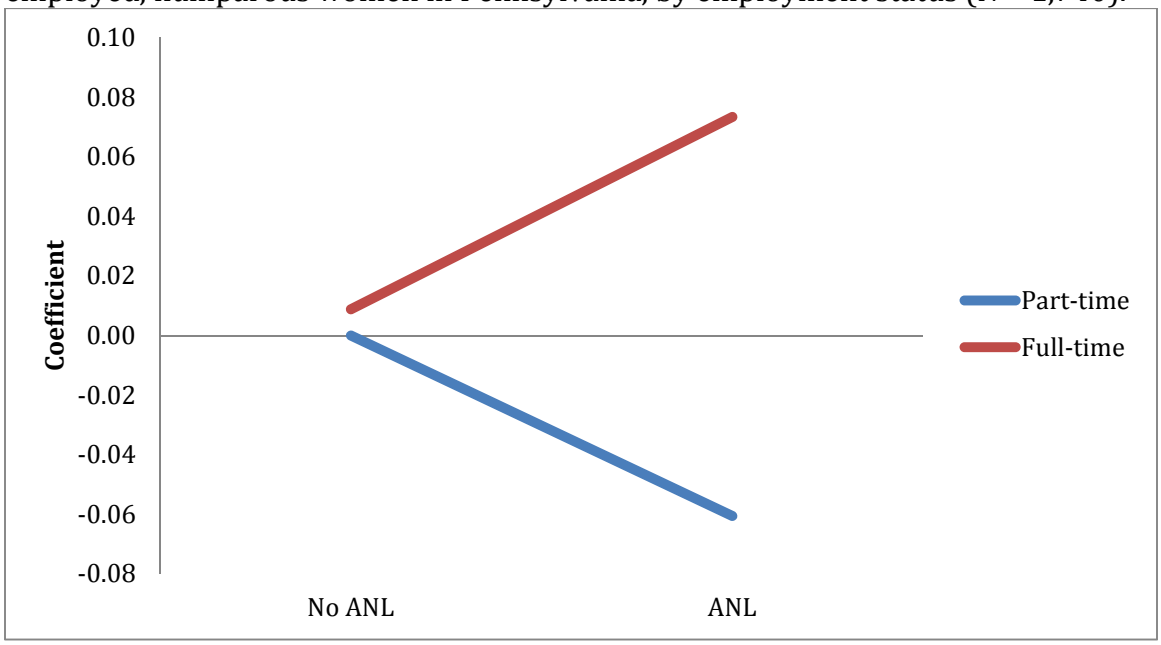


Figure 4: Results from multivariable linear probability models for the likelihood of experiencing any negative delivery outcome in propensity-score matched groups among employed, nulliparous women in Pennsylvania, by employment status (N = 1,740).



Appendix

Table A1a: Unadjusted and adjusted multivariable linear probability and Poisson regression results for negative delivery outcomes (any and count) in unmatched groups among employed, nulliparous women in Pennsylvania (N = 1,740). Coefficients and 95% confidence intervals.

	Any negative delivery outcome			Number of negative delivery outcomes		
	Unadjusted	Adjusted	Adjusted with full-time employment interaction	Unadjusted	Adjusted	Adjusted with full-time employment interaction
	Coeff. [95% CI]	Coeff. [95% CI]	Coeff. [95% CI]	Coeff. [95% CI]	Coeff. [95% CI]	Coeff. [95% CI]
ANL (≥2 days)	0.1 [0.05 - 0.14]**	0.06 [0.01 - 0.11]*	0.01 [-0.14 - 0.15]	0.25 [0.15 - 0.36]**	0.16 [0.04 - 0.27]**	0.22 [-0.13 - 0.56]
Full-time employed (ref: part-time)		-0.08 [-0.15 - 0.00]	-0.11 [-0.24 - 0.01]		-0.13 [-0.30 - 0.04]	-0.08 [-0.40 - 0.24]
Full-time employed x ANL (≥2 days)			0.06 [-0.09 - 0.21]			-0.07 [-0.42 - 0.29]
Pre-pregnancy obesity		0.06 [-0.00 - 0.12]	0.06 [0.00 - 0.12]*		0.15 [0.02 - 0.28]*	0.15 [0.02 - 0.28]*
Gained > recommended		0.04 [-0.01 - 0.08]	0.04 [-0.01 - 0.08]		0.16 [0.05 - 0.27]**	0.16 [0.05 - 0.27]**
Prior miscarriage		0.03 [-0.03 - 0.10]	0.03 [-0.03 - 0.10]		0.04 [-0.11 - 0.18]	0.04 [-0.11 - 0.18]
Pre-pregnancy hypertension or diabetes		0.15 [0.02 - 0.28]*	0.15 [0.02 - 0.28]*		0.21 [-0.05 - 0.46]	0.21 [-0.05 - 0.46]
Serious health condition during pregnancy		0.05 [-0.00 - 0.11]	0.05 [-0.00 - 0.11]		0.14 [0.01 - 0.26]*	0.14 [0.01 - 0.26]*
Mean number of hospitalizations		0.03 [-0.02 - 0.08]	0.03 [-0.02 - 0.08]		0.06 [-0.05 - 0.16]	0.06 [-0.05 - 0.16]
Doctor office or urgent care visits in last month of pregnancy (ref: 2-4 visits)						
0-1 visits		0 [-0.09 - 0.10]	0 [-0.09 - 0.10]		0.05 [-0.16 - 0.27]	0.05 [-0.16 - 0.27]
5+ visits		0.12 [0.06 - 0.18]**	0.12 [0.06 - 0.18]**		0.27 [0.14 - 0.40]**	0.27 [0.14 - 0.40]**
Help getting pregnant		0 [-0.08 - 0.07]	-0.01 [-0.08 - 0.07]		0.01 [-0.15 - 0.17]	0.01 [-0.15 - 0.18]
Provider advised C-sec during pregnancy		-0.11 [-0.20 - -0.02]*	-0.11 [-0.20 - -0.02]*		-0.24 [-0.45 - -0.03]*	-0.24 [-0.45 - -0.03]*
Birth Anticipation Scale		0.01 [0.00 - 0.01]**	0.01 [0.00 - 0.01]**		0.02 [0.01 - 0.04]**	0.02 [0.01 - 0.04]**
Strong preference for vaginal birth		-0.06 [-0.13 - 0.01]	-0.06 [-0.13 - 0.01]		-0.15 [-0.32 - 0.01]	-0.15 [-0.32 - 0.01]
Maternal stress		0 [-0.00 - 0.01]	0 [-0.00 - 0.01]		0.01 [-0.01 - 0.02]	0.01 [-0.01 - 0.02]
Low social support		0 [-0.04 - 0.05]	0 [-0.04 - 0.05]		0.03 [-0.08 - 0.14]	0.03 [-0.08 - 0.14]
Baseline EPDS		0 [-0.00 - 0.01]	0 [-0.00 - 0.01]		0 [-0.01 - 0.02]	0 [-0.01 - 0.02]
Gestational age at delivery (ref: full term)						
Early term		0.07 [0.01 - 0.12]*	0.07 [0.01 - 0.12]*		0.1 [-0.03 - 0.24]	0.1 [-0.03 - 0.24]
Late term		0.27 [0.20 - 0.34]**	0.27 [0.20 - 0.34]**		0.48 [0.34 - 0.62]**	0.48 [0.34 - 0.62]**
Post term		0.24 [0.01 - 0.47]*	0.24 [0.01 - 0.47]*		0.56 [0.15 - 0.98]**	0.56 [0.15 - 0.98]**

Service or labor (ref: professional, managerial, or clerical)		0.01 [-0.05 - 0.07]	0.01 [-0.05 - 0.07]		0.07 [-0.06 - 0.19]	0.06 [-0.06 - 0.19]
Pregnancy intention (ref: wanted)						
Mistimed		0.03 [-0.03 - 0.09]	-0.03 [-0.27 - 0.20]		-0.25 [-0.08 - 0.20]	0.06 [-0.08 - 0.20]
Unwanted		-0.04 [-0.27 - 0.20]	0.03 [-0.03 - 0.09]		0.06 [-0.88 - 0.38]	-0.25 [-0.89 - 0.38]
Mother's age at baseline		0.06 [-0.02 - 0.15]	0.06 [-0.02 - 0.15]		0.04 [-0.16 - 0.24]	0.04 [-0.16 - 0.24]
Mother's age at baseline squared		0 [-0.00 - 0.00]	0 [-0.00 - 0.00]		0 [-0.00 - 0.00]	0 [-0.00 - 0.00]
Maternal education (ref: college grad)						
High school graduate or less		0 [-0.10 - 0.10]	0 [-0.10 - 0.10]		-0.1 [-0.35 - 0.14]	-0.1 [-0.35 - 0.14]
Some college or vocational programs		-0.03 [-0.09 - 0.03]	-0.03 [-0.09 - 0.03]		-0.1 [-0.24 - 0.05]	-0.1 [-0.24 - 0.05]
Maternal race/ethnicity (ref: Non-Hispanic White)						
Hispanic		0.02 [-0.12 - 0.16]	0.02 [-0.12 - 0.16]		-0.15 [-0.50 - 0.21]	-0.15 [-0.50 - 0.21]
Non-Hispanic Black or Other		0.08 [-0.02 - 0.19]	0.08 [-0.02 - 0.19]		0.02 [-0.22 - 0.25]	0.02 [-0.22 - 0.25]
Unmarried or not living with partner (ref: married or living with partner)		-0.08 [-0.19 - 0.03]	-0.08 [-0.19 - 0.04]		-0.23 [-0.50 - 0.05]	-0.23 [-0.50 - 0.05]
Public insurance (ref: private or OOP)		-0.03 [-0.13 - 0.07]	-0.03 [-0.13 - 0.07]		-0.04 [-0.28 - 0.20]	-0.04 [-0.28 - 0.20]
Poverty status (ref: non-poor)						
Poor		0.01 [-0.13 - 0.14]	0.01 [-0.13 - 0.14]		0.1 [-0.20 - 0.39]	0.1 [-0.20 - 0.39]
Near poor		0.02 [-0.09 - 0.13]	0.02 [-0.09 - 0.13]		-0.02 [-0.28 - 0.23]	-0.02 [-0.28 - 0.23]
Mean outcome	0.55	0.55	0.55	0.82	0.82	0.82
Observations	1740	1740	1740	1740	1740	1740
R-squared/Pseudo R-squared	0.01	0.10	0.10	0.01	0.04	0.04

Abbreviations: CI, confidence interval; ANL, antenatal leave; EPDS, Edinburgh Postnatal Depression Scale; OOP, out of pocket.

Notes: Coefficients and 95% confidence intervals. Serious health condition defined as hypertension, high blood pressure, or pre-eclampsia that started during pregnancy; diabetes that started during pregnancy; vaginal bleeding; early or pre-term labor; bed rest or hospitalization because of premature labor. The First Baby Study Birth Anticipation Scale (BAS) measures fear of childbirth. Maternal stress measured with the Perceived Stress Scale. Early term refers to gestations of 37w0d to 38w6d; full term refers to 39w0d to 40w6d, late term refers to 41w0d to 41w6d, and post term refers to >42w0d. We excluded preterm deliveries (i.e., <37w0d) from the sample. Poverty status was measured using the US Census Bureau classification system incorporating both household income and family composition. Women with household incomes <100% of the poverty threshold are classified as "poor"; those with household incomes between 100-200% of the poverty threshold are classified as "near poor"; and those with household incomes at or above 200% of the poverty threshold are classified as "not poor". Negative delivery outcomes include labor induction, labor lasting more than 24 hours, unplanned C-section, and self-reported negative birth experience.

* significant at 5%; ** significant at 1%

Table A1b: Unadjusted and adjusted multivariable linear probability results for labor induction and long labor duration in unmatched groups among employed, nulliparous women in Pennsylvania (N = 1,740). Coefficients and 95% confidence intervals.

	Labor induced			Labor >24 hours		
	Unadjusted	Adjusted	Adjusted with full-time employment interaction	Unadjusted	Adjusted	Adjusted with full-time employment interaction
	Coeff. [95% CI]	Coeff. [95% CI]	Coeff. [95% CI]	Coeff. [95% CI]	Coeff. [95% CI]	Coeff. [95% CI]
ANL (≥2 days)	0.07 [0.03 - 0.11]**	0.03 [-0.01 - 0.08]	0.11 [-0.02 - 0.24]	0.06 [0.03 - 0.09]**	0.04 [0.01 - 0.08]*	0.01 [-0.09 - 0.11]
Full-time employed (ref: part-time)		-0.03 [-0.10 - 0.04]	0.02 [-0.09 - 0.14]		0.03 [-0.03 - 0.08]	0 [-0.09 - 0.09]
Full-time employed x ANL (≥2 days)			-0.08 [-0.22 - 0.06]			0.03 [-0.07 - 0.14]
Pre-pregnancy obesity		0.12 [0.07 - 0.18]**	0.12 [0.06 - 0.18]**		0.01 [-0.03 - 0.05]	0.01 [-0.03 - 0.05]
Gained > recommended		0.02 [-0.02 - 0.06]	0.02 [-0.02 - 0.06]		0 [-0.03 - 0.04]	0 [-0.03 - 0.04]
Prior miscarriage		-0.04 [-0.10 - 0.02]	-0.04 [-0.10 - 0.02]		0.01 [-0.04 - 0.05]	0.01 [-0.03 - 0.05]
Pre-pregnancy hypertension or diabetes		0.15 [0.03 - 0.27]*	0.14 [0.02 - 0.26]*		0.02 [-0.07 - 0.11]	0.02 [-0.07 - 0.11]
Serious health condition during pregnancy		0.07 [0.02 - 0.12]*	0.07 [0.02 - 0.12]*		0.02 [-0.02 - 0.05]	0.02 [-0.02 - 0.05]
Mean number of hospitalizations		-0.01 [-0.05 - 0.04]	-0.01 [-0.05 - 0.04]		0.01 [-0.02 - 0.04]	0.01 [-0.02 - 0.04]
Doctor office or urgent care visits in last month of pregnancy (ref: 2-4 visits)						
0-1 visits		0.01 [-0.07 - 0.10]	0.01 [-0.07 - 0.10]		0.01 [-0.06 - 0.07]	0.01 [-0.06 - 0.07]
5+ visits		0.14 [0.09 - 0.20]**	0.15 [0.09 - 0.20]**		0 [-0.04 - 0.04]	0 [-0.04 - 0.04]
Help getting pregnant		-0.01 [-0.08 - 0.05]	-0.01 [-0.08 - 0.05]		0.01 [-0.04 - 0.06]	0.01 [-0.04 - 0.06]
Provider advised C-sec during pregnancy		-0.08 [-0.16 - -0.00]*	-0.08 [-0.17 - -0.00]*		0.05 [-0.03 - 0.12]	0.05 [-0.02 - 0.12]
Birth Anticipation Scale		0 [-0.00 - 0.01]	0 [-0.00 - 0.01]		0 [-0.00 - 0.01]	0 [-0.00 - 0.01]
Strong preference for vaginal birth		-0.02 [-0.08 - 0.04]	-0.02 [-0.08 - 0.04]		0.01 [-0.04 - 0.05]	0.01 [-0.04 - 0.05]
Maternal stress		0 [-0.01 - 0.00]	0 [-0.01 - 0.00]		0 [-0.00 - 0.01]	0 [-0.00 - 0.01]
Low social support		0.01 [-0.03 - 0.05]	0.01 [-0.03 - 0.05]		-0.01 [-0.04 - 0.02]	-0.01 [-0.04 - 0.02]
Baseline EPDS		0 [-0.01 - 0.00]	0 [-0.01 - 0.00]		0 [-0.01 - 0.01]	0 [-0.01 - 0.01]
Gestational age at delivery (ref: full term)						
Early term		0.07 [0.01 - 0.12]*	0.07 [0.01 - 0.12]*		-0.03 [-0.07 - 0.01]	-0.03 [-0.07 - 0.01]
Late term		0.3 [0.24 - 0.37]**	0.3 [0.24 - 0.37]**		0.04 [-0.00 - 0.09]	0.04 [-0.00 - 0.09]
Post term		0.31 [0.10 - 0.53]**	0.31 [0.10 - 0.52]**		0.06 [-0.10 - 0.22]	0.06 [-0.10 - 0.22]
Service or labor (ref: professional, managerial, or clerical)		-0.02 [-0.07 - 0.03]	-0.02 [-0.07 - 0.03]		0.01 [-0.03 - 0.05]	0.01 [-0.03 - 0.05]
Pregnancy intention (ref: wanted)						
Mistimed		-0.02 [-0.07 - 0.04]	-0.02 [-0.31 - 0.12]		-0.03 [-0.07 - 0.02]	-0.03 [-0.16 - 0.16]
Unwanted		-0.09 [-0.31 - 0.12]	-0.1 [-0.07 - 0.04]		0 [-0.16 - 0.16]	0 [-0.07 - 0.02]

Mother's age at baseline		0.04 [-0.04 - 0.12]	0.04 [-0.04 - 0.12]		-0.03 [-0.09 - 0.02]	-0.03 [-0.09 - 0.02]
Mother's age at baseline squared		0 [-0.00 - 0.00]	0 [-0.00 - 0.00]		0 [-0.00 - 0.00]	0 [-0.00 - 0.00]
Maternal education (ref: college grad)						
High school graduate or less		0.07 [-0.03 - 0.16]	0.07 [-0.03 - 0.16]		-0.03 [-0.10 - 0.04]	-0.03 [-0.10 - 0.04]
Some college or vocational programs		0.01 [-0.05 - 0.06]	0.01 [-0.05 - 0.07]		-0.01 [-0.05 - 0.03]	-0.01 [-0.05 - 0.03]
Maternal race/ethnicity (ref: Non-Hispanic White)						
Hispanic		-0.01 [-0.14 - 0.12]	-0.01 [-0.14 - 0.12]		0.03 [-0.07 - 0.12]	0.03 [-0.07 - 0.12]
Non-Hispanic Black or Other		-0.07 [-0.17 - 0.02]	-0.07 [-0.17 - 0.02]		-0.02 [-0.09 - 0.05]	-0.02 [-0.09 - 0.05]
Unmarried or not living with partner (ref: married or living with partner)		0 [-0.11 - 0.10]	-0.01 [-0.11 - 0.10]		-0.07 [-0.14 - 0.01]	-0.07 [-0.14 - 0.01]
Public insurance (ref: private or OOP)		-0.04 [-0.14 - 0.05]	-0.04 [-0.14 - 0.05]		0.08 [0.01 - 0.15]*	0.08 [0.01 - 0.15]*
Poverty status (ref: non-poor)						
Poor		0.05 [-0.07 - 0.17]	0.05 [-0.07 - 0.17]		0 [-0.09 - 0.09]	0 [-0.09 - 0.09]
Near poor		0.04 [-0.07 - 0.14]	0.03 [-0.07 - 0.13]		0.01 [-0.07 - 0.08]	0.01 [-0.07 - 0.08]
Mean outcome	0.31	0.31	0.31	0.11	0.11	0.11
Observations	1740	1740	1740	1606	1606	1606
R-squared/Pseudo R-squared	0.01	0.11	0.11	0.01	0.03	0.03

Abbreviations: CI, confidence interval; ANL, antenatal leave; EPDS, Edinburgh Postnatal Depression Scale; OOP, out of pocket.

Notes: Coefficients and 95% confidence intervals. Serious health condition defined as hypertension, high blood pressure, or pre-eclampsia that started during pregnancy; diabetes that started during pregnancy; vaginal bleeding; early or pre-term labor; bed rest or hospitalization because of premature labor. The First Baby Study Birth Anticipation Scale (BAS) measures fear of childbirth. Maternal stress measured with the Perceived Stress Scale. Early term refers to gestations of 37w0d to 38w6d; full term refers to 39w0d to 40w6d, late term refers to 41w0d to 41w6d, and post term refers to >42w0d. We excluded preterm deliveries (i.e., <37w0d) from the sample. Poverty status was measured using the US Census Bureau classification system incorporating both household income and family composition. Women with household incomes <100% of the poverty threshold are classified as "poor"; those with household incomes between 100-200% of the poverty threshold are classified as "near poor"; and those with household incomes at or above 200% of the poverty threshold are classified as "not poor". Negative delivery outcomes include labor induction, labor lasting more than 24 hours, unplanned C-section, and self-reported negative birth experience.

* significant at 5%; ** significant at 1%

Table A1c: Unadjusted and adjusted multivariable linear probability and linear regression results for labor duration and unplanned C-section in unmatched groups among employed, nulliparous women in Pennsylvania (N = 1,740). Coefficients and 95% confidence intervals.

	Labor duration			Unplanned cesarean		
	Unadjusted	Adjusted	Adjusted with full-time employment interaction	Unadjusted	Adjusted	Adjusted with full-time employment interaction
	Coeff. [95% CI]	Coeff. [95% CI]	Coeff. [95% CI]	Coeff. [95% CI]	Coeff. [95% CI]	Coeff. [95% CI]
ANL (≥2 days)	1.84 [0.89 - 2.78]**	1.39 [0.40 - 2.39]**	-0.31 [-3.35 - 2.72]	0.07 [0.03 - 0.11]**	0.06 [0.02 - 0.10]**	0.03 [-0.09 - 0.16]
Full-time employed (ref: part-time)		1.54 [-0.08 - 3.16]	0.23 [-2.50 - 2.96]		-0.01 [-0.08 - 0.06]	-0.03 [-0.15 - 0.09]
Full-time employed x ANL (≥2 days)			1.89 [-1.28 - 5.07]			0.03 [-0.11 - 0.16]
Pre-pregnancy obesity		0.09 [-1.19 - 1.37]	0.13 [-1.15 - 1.41]		0.06 [0.00 - 0.11]*	0.06 [0.00 - 0.12]*
Gained > recommended		0.76 [-0.20 - 1.71]	0.76 [-0.19 - 1.72]		0.1 [0.06 - 0.15]**	0.1 [0.06 - 0.15]**
Prior miscarriage		0.8 [-0.52 - 2.12]	0.85 [-0.47 - 2.17]		0.05 [-0.01 - 0.11]	0.05 [-0.01 - 0.11]
Pre-pregnancy hypertension or diabetes		0.59 [-2.17 - 3.35]	0.6 [-2.16 - 3.36]		0.04 [-0.08 - 0.16]	0.04 [-0.08 - 0.16]
Serious health condition during pregnancy		0.06 [-1.12 - 1.23]	0.06 [-1.12 - 1.23]		0.02 [-0.03 - 0.07]	0.02 [-0.03 - 0.07]
Mean number of hospitalizations		0 [-1.04 - 1.04]	-0.01 [-1.05 - 1.03]		0.01 [-0.03 - 0.06]	0.01 [-0.03 - 0.06]
Doctor office or urgent care visits in last month of pregnancy (ref: 2-4 visits)						
0-1 visits		1.07 [-0.88 - 3.01]	1.06 [-0.89 - 3.00]		0.05 [-0.04 - 0.13]	0.05 [-0.04 - 0.13]
5+ visits		0.52 [-0.74 - 1.78]	0.48 [-0.78 - 1.75]		0.08 [0.02 - 0.13]**	0.08 [0.02 - 0.13]**
Help getting pregnant		0.37 [-1.13 - 1.88]	0.35 [-1.15 - 1.86]		0.05 [-0.01 - 0.12]	0.05 [-0.01 - 0.12]
Provider advised C-sec during pregnancy		1.72 [-0.46 - 3.90]	1.77 [-0.41 - 3.95]		0.03 [-0.06 - 0.13]	0.04 [-0.06 - 0.13]
Birth Anticipation Scale		0.07 [-0.05 - 0.18]	0.06 [-0.05 - 0.18]		0 [-0.01 - 0.00]	0 [-0.01 - 0.00]
Strong preference for vaginal birth		0.28 [-1.10 - 1.66]	0.25 [-1.13 - 1.64]		-0.08 [-0.14 - -0.02]**	-0.08 [-0.14 - -0.02]**
Maternal stress		0.08 [-0.07 - 0.24]	0.08 [-0.07 - 0.24]		0 [-0.01 - 0.01]	0 [-0.01 - 0.01]
Low social support		0.59 [-0.41 - 1.58]	0.58 [-0.42 - 1.57]		0.01 [-0.03 - 0.06]	0.01 [-0.03 - 0.06]
Baseline EPDS		0.04 [-0.12 - 0.21]	0.05 [-0.12 - 0.21]		0 [-0.01 - 0.01]	0 [-0.01 - 0.01]
Gestational age at delivery (ref: full term)						
Early term		-1.73 [-2.91 - -0.54]**	-1.74 [-2.93 - -0.55]**		0.01 [-0.04 - 0.06]	0.01 [-0.04 - 0.06]
Late term		0.52 [-0.93 - 1.97]	0.52 [-0.93 - 1.97]		0.1 [0.04 - 0.16]**	0.1 [0.04 - 0.16]**
Post term		3.8 [-0.98 - 8.59]	3.89 [-0.90 - 8.67]		0.11 [-0.10 - 0.32]	0.11 [-0.10 - 0.32]
Service or labor (ref: professional, managerial, or clerical)		-0.39 [-1.56 - 0.78]	-0.4 [-1.57 - 0.78]		0.01 [-0.04 - 0.06]	0.01 [-0.04 - 0.06]
Pregnancy intention (ref: wanted)						
Mistimed		-0.2 [-2.47 - 0.09]	-0.09 [-4.92 - 4.75]		-0.01 [-0.22 - 0.21]	0.06 [0.00 - 0.11]*
Unwanted		-1.19 [-5.04 - 4.63]	-1.18 [-2.46 - 0.10]		0.06 [-0.00 - 0.11]	0 [-0.21 - 0.21]
Mother's age at baseline		-0.29 [-2.05 - 1.48]	-0.29 [-2.05 - 1.48]		0.04 [-0.04 - 0.12]	0.04 [-0.04 - 0.12]
Mother's age at baseline squared		0.01 [-0.03 - 0.04]	0.01 [-0.03 - 0.04]		0 [-0.00 - 0.00]	0 [-0.00 - 0.00]

Maternal education (ref: college grad)						
High school graduate or less		0.74 [-1.39 - 2.88]	0.7 [-1.43 - 2.83]		0.01 [-0.08 - 0.10]	0.01 [-0.08 - 0.10]
Some college or vocational programs		0.72 [-0.56 - 2.00]	0.69 [-0.59 - 1.96]		0 [-0.06 - 0.05]	-0.01 [-0.06 - 0.05]
Maternal race/ethnicity (ref: Non-Hispanic White)						
Hispanic		-0.08 [-2.97 - 2.80]	-0.1 [-2.99 - 2.78]		-0.05 [-0.17 - 0.08]	-0.05 [-0.17 - 0.08]
Non-Hispanic Black or Other		-0.43 [-2.56 - 1.71]	-0.44 [-2.58 - 1.70]		0.05 [-0.04 - 0.15]	0.05 [-0.04 - 0.15]
Unmarried or not living with partner (ref: married or living with partner)		-2.18 [-4.50 - 0.13]	-2.14 [-4.45 - 0.18]		0 [-0.10 - 0.10]	0 [-0.10 - 0.10]
Public insurance (ref: private or OOP)		2.21 [0.08 - 4.35]*	2.24 [0.10 - 4.37]*		0.02 [-0.07 - 0.11]	0.02 [-0.07 - 0.11]
Poverty status (ref: non-poor)						
Poor		-1.1 [-3.79 - 1.59]	-1.09 [-3.78 - 1.60]		0 [-0.12 - 0.11]	0 [-0.12 - 0.11]
Near poor		0.98 [-1.29 - 3.26]	1.02 [-1.25 - 3.30]		0 [-0.10 - 0.10]	0 [-0.10 - 0.10]
Mean outcome	14.20	14.20	14.20	0.24	0.24	0.24
Observations	1606	1606	1606	1635	1635	1635
R-squared/Pseudo R-squared	0.01	0.04	0.04	0.01	0.07	0.07

Abbreviations: CI, confidence interval; ANL, antenatal leave; EPDS, Edinburgh Postnatal Depression Scale; OOP, out of pocket.

Notes: Coefficients and 95% confidence intervals. Serious health condition defined as hypertension, high blood pressure, or pre-eclampsia that started during pregnancy; diabetes that started during pregnancy; vaginal bleeding; early or pre-term labor; bed rest or hospitalization because of premature labor. The First Baby Study Birth Anticipation Scale (BAS) measures fear of childbirth. Maternal stress measured with the Perceived Stress Scale. Early term refers to gestations of 37w0d to 38w6d; full term refers to 39w0d to 40w6d, late term refers to 41w0d to 41w6d, and post term refers to >42w0d. We excluded preterm deliveries (i.e., <37w0d) from the sample. Poverty status was measured using the US Census Bureau classification system incorporating both household income and family composition. Women with household incomes <100% of the poverty threshold are classified as "poor"; those with household incomes between 100-200% of the poverty threshold are classified as "near poor"; and those with household incomes at or above 200% of the poverty threshold are classified as "not poor". Negative delivery outcomes include labor induction, labor lasting more than 24 hours, unplanned C-section, and self-reported negative birth experience.

* significant at 5%; ** significant at 1%

Table A1d: Unadjusted and adjusted multivariable linear probability results for self-reported negative birth experience in unmatched groups among employed, nulliparous women in Pennsylvania (N = 1,740). Coefficients and 95% confidence intervals.

	Negative birth experience		
	Unadjusted	Adjusted	Adjusted with full-time employment interaction
	Coeff. [95% CI]	Coeff. [95% CI]	Coeff. [95% CI]
ANL (≥2 days)	0.02 [-0.01 - 0.06]	0.01 [-0.03 - 0.05]	0.03 [-0.08 - 0.14]
Full-time employed (ref: part-time)		-0.11 [-0.17 - -0.05]**	-0.09 [-0.19 - 0.01]
Full-time employed x ANL (≥2 days)			-0.03 [-0.15 - 0.09]
Pre-pregnancy obesity		-0.04 [-0.09 - 0.00]	-0.04 [-0.09 - 0.00]
Gained > recommended		0.02 [-0.01 - 0.06]	0.02 [-0.01 - 0.06]
Prior miscarriage		0.02 [-0.03 - 0.07]	0.02 [-0.03 - 0.07]
Pre-pregnancy hypertension or diabetes		0.05 [-0.05 - 0.16]	0.05 [-0.05 - 0.15]
Serious health condition during pregnancy		0 [-0.04 - 0.05]	0 [-0.04 - 0.05]
Mean number of hospitalizations		0.03 [-0.01 - 0.07]	0.03 [-0.01 - 0.07]
Doctor office or urgent care visits in last month of pregnancy (ref: 2-4 visits)			
0-1 visits		-0.02 [-0.10 - 0.05]	-0.02 [-0.10 - 0.05]
5+ visits		0.03 [-0.02 - 0.08]	0.03 [-0.02 - 0.08]
Help getting pregnant		-0.05 [-0.11 - 0.01]	-0.05 [-0.11 - 0.01]
Provider advised C-sec during pregnancy		-0.07 [-0.14 - -0.00]*	-0.07 [-0.14 - -0.00]*
Birth Anticipation Scale		0.01 [0.01 - 0.02]**	0.01 [0.01 - 0.02]**
Strong preference for vaginal birth		-0.04 [-0.10 - 0.01]	-0.04 [-0.09 - 0.01]
Maternal stress		0.01 [0.00 - 0.01]**	0.01 [0.00 - 0.01]**
Low social support		0.02 [-0.01 - 0.06]	0.02 [-0.01 - 0.06]
Baseline EPDS		0.01 [0.00 - 0.01]**	0.01 [0.00 - 0.01]**
Gestational age at delivery (ref: full term)			
Early term		0.02 [-0.02 - 0.06]	0.02 [-0.02 - 0.06]
Late term		0.02 [-0.04 - 0.07]	0.02 [-0.04 - 0.07]
Post term		0.11 [-0.07 - 0.29]	0.11 [-0.07 - 0.29]
Service or labor (ref: professional, managerial, or clerical)		0.03 [-0.01 - 0.08]	0.03 [-0.01 - 0.08]
Pregnancy intention (ref: wanted)			
Mistimed		0.05 [0.00 - 0.10]*	0.05 [-0.23 - 0.14]
Unwanted		-0.04 [-0.23 - 0.14]	-0.05 [0.00 - 0.10]*
Mother's age at baseline		-0.02 [-0.08 - 0.05]	-0.02 [-0.08 - 0.05]
Mother's age at baseline squared		0 [-0.00 - 0.00]	0 [-0.00 - 0.00]
Maternal education (ref: college grad)			
High school graduate or less		-0.11 [-0.19 - -0.03]**	-0.11 [-0.19 - -0.03]**
Some college or vocational programs		-0.06 [-0.10 - -0.01]*	-0.06 [-0.10 - -0.01]*
Maternal race/ethnicity (ref: Non-Hispanic White)			
Hispanic		-0.07 [-0.19 - 0.04]	-0.07 [-0.19 - 0.04]
Non-Hispanic Black or Other		0.06 [-0.02 - 0.14]	0.06 [-0.02 - 0.14]
Unmarried or not living with partner (ref: married or living with partner)		-0.11 [-0.20 - -0.02]*	-0.11 [-0.20 - -0.02]*
Public insurance (ref: private or OOP)		-0.1 [-0.19 - -0.02]*	-0.1 [-0.18 - -0.02]*

Poverty status (ref: non-poor)			
Poor		0 [-0.10 - 0.11]	0 [-0.10 - 0.11]
Near poor		-0.07 [-0.16 - 0.02]	-0.07 [-0.16 - 0.02]
Mean outcome	0.19	0.19	0.19
Observations	1730	1730	1730
R-squared/Pseudo R-squared	0.00	0.10	0.10

Abbreviations: CI, confidence interval; ANL, antenatal leave; EPDS, Edinburgh Postnatal Depression Scale; OOP, out of pocket.

Notes: Coefficients and 95% confidence intervals. Serious health condition defined as hypertension, high blood pressure, or pre-eclampsia that started during pregnancy; diabetes that started during pregnancy; vaginal bleeding; early or pre-term labor; bed rest or hospitalization because of premature labor. The First Baby Study Birth Anticipation Scale (BAS) measures fear of childbirth. Maternal stress measured with the Perceived Stress Scale. Early term refers to gestations of 37w0d to 38w6d; full term refers to 39w0d to 40w6d, late term refers to 41w0d to 41w6d, and post term refers to >42w0d. We excluded preterm deliveries (i.e., <37w0d) from the sample. Poverty status was measured using the US Census Bureau classification system incorporating both household income and family composition. Women with household incomes <100% of the poverty threshold are classified as "poor"; those with household incomes between 100-200% of the poverty threshold are classified as "near poor"; and those with household incomes at or above 200% of the poverty threshold are classified as "not poor". Negative delivery outcomes include labor induction, labor lasting more than 24 hours, unplanned C-section, and self-reported negative birth experience.

* significant at 5%; ** significant at 1%

Table A2: Sensitivity analysis dropping observations at low propensity score densities. Multivariable linear probability, linear regression, and Poisson regression results for negative delivery outcomes (any and count), labor induction, labor duration, unplanned C-section, and negative birth experience in propensity-score matched groups among employed, nulliparous women in Pennsylvania (N = 1,740). Coefficients and 95% confidence intervals.

	Any negative delivery outcome	Number of negative delivery outcomes	Labor induced	Labor >24 hours	Hours in labor	Unplanned C-section	Negative birth experience
	Coeff. [95% CI]	Coeff. [95% CI]	Coeff. [95% CI]	Coeff. [95% CI]	Coeff. [95% CI]	Coeff. [95% CI]	Coeff. [95% CI]
Panel A: 10% drop							
ANL (≥2 days)	0.06 [0.01 - 0.11]*	0.15 [0.04 - 0.25]**	0.05 [0.00 - 0.09]*	0.04 [0.00 - 0.07]*	0.93 [-0.08 - 1.94]	0.06 [0.02 - 0.10]**	0 [-0.04 - 0.04]
Mean of negative delivery outcomes	0.56	0.84	0.31	0.11	14.59	0.24	0.19
Weighted observations	1573	1573	1573	1452	1452	1477	1567
R-squared/Pseudo R-squared	0.03	0.02	0.06	0.01	0.01	0.02	0.04
Panel B: 20% drop							
ANL (≥2 days)	0.07 [0.02 - 0.12]*	0.17 [0.05 - 0.29]**	0.06 [0.01 - 0.11]*	0.03 [0.00 - 0.07]*	0.89 [-0.16 - 1.95]	0.05 [0.00 - 0.10]*	0.01 [-0.03 - 0.05]
Mean of negative delivery outcomes	0.54	0.80	0.29	0.11	14.52	0.23	0.19
Weighted observations	1398	1398	1398	1286	1286	1308	1392
R-squared/Pseudo R-squared	0.02	0.01	0.04	0.01	0.01	0.01	0.04

Abbreviations: CI, confidence interval; ANL, antenatal leave.

Notes: Coefficients and 95% confidence intervals for linear probability models (any negative delivery outcome, labor induction, long labor duration, unplanned C-section, and negative birth experience), Poisson regression (number of negative delivery outcomes), and linear regression models (hours in labor). Panel A drops 10% of the observations who took ANL at which the propensity score density of the women who did not take ANL is the lowest. Panel B uses a 20% cut-off. All models are weighted to account for the number of units each observation was matched with and adjust for the propensity score, pre-pregnancy obesity, strong preference for vaginal delivery, perceived stress, employment status and race/ethnicity. Negative delivery outcomes include labor induction, labor lasting more than 24 hours, unplanned C-section, and self-reported negative birth experience.

* significant at 5%; ** significant at 1%

Table A3: Sensitivity analysis to assess potential misclassification. Multivariable linear probability, linear regression, and Poisson regression results for negative delivery outcomes (any and count), labor induction, labor duration, unplanned C-section, and negative birth experience in propensity-score matched groups among employed, nulliparous women in Pennsylvania (N = 1,740 – 1,743). Coefficients and 95% confidence intervals.

	Any negative delivery outcome	Number of negative delivery outcomes	Labor induced	Labor >24 hours	Hours in labor	Unplanned C-section	Negative birth experience
	Coeff. [95% CI]	Coeff. [95% CI]	Coeff. [95% CI]	Coeff. [95% CI]	Coeff. [95% CI]	Coeff. [95% CI]	Coeff. [95% CI]
Panel A: With imputed ANL							
ANL (≥2 days)	0.04 [-0.00 - 0.09]	0.13 [0.03 - 0.24]**	0.02 [-0.03 - 0.06]	0.05 [0.02 - 0.08]**	1.16 [0.20 - 2.12]*	0.06 [0.02 - 0.10]**	0.02 [-0.02 - 0.06]
Mean of negative delivery outcomes	0.58	0.86	0.33	0.12	14.55	0.24	0.19
Weighted observations	1749	1749	1749	1620	1620	1650	1741
R-squared/Pseudo R-squared	0.04	0.02	0.06	0.01	0.01	0.03	0.02
Panel B: Drop if ANL = 2 days							
ANL (≥2 days)	0.06 [0.01 - 0.11]*	0.17 [0.06 - 0.27]**	0.03 [-0.02 - 0.07]	0.05 [0.02 - 0.08]**	1.05 [0.05 - 2.06]*	0.07 [0.03 - 0.12]**	0.02 [-0.02 - 0.06]
Mean of negative delivery outcomes	0.59	0.87	0.34	0.12	14.42	0.25	0.19
Weighted observations	1515	1515	1515	1406	1406	1433	1507
R-squared/Pseudo R-squared	0.04	0.02	0.06	0.02	0.01	0.03	0.02
Panel C: Drop if labor > 24 hours							
ANL (≥2 days)	0.03 [-0.02 - 0.08]	0.12 [0.00 - 0.24]*	0.03 [-0.02 - 0.08]	NA	0.11 [-0.51 - 0.72]	0.06 [0.02 - 0.10]**	0.01 [-0.03 - 0.05]
Mean of negative delivery outcomes	0.55	0.77	0.37	--	11.89	0.22	0.18
Weighted observations	1426	1426	1426	--	1426	1426	1420
R-squared/Pseudo R-squared	0.06	0.03	0.08	--	0.02	0.04	0.02

Abbreviations: CI, confidence interval; ANL, antenatal leave.

Notes: Coefficients and 95% confidence intervals for linear probability models (any negative delivery outcome, labor induction, long labor duration, unplanned C-section, and negative birth experience), Poisson regression (number of negative delivery outcomes), and linear regression models (hours in labor). Panel A includes 2 women for whom we imputed ANL duration; Panel B drops women who stopped working exactly 2 days before delivery; Panel C drops women who were in labor for longer than 24 hours. All models are weighted to account for the number of units each observation was matched with and adjust for the propensity score, pre-pregnancy obesity, strong preference for vaginal delivery, perceived stress, employment status and race/ethnicity. Negative delivery outcomes include labor induction, labor lasting more than 24 hours, unplanned C-section, and self-reported negative birth experience.

* significant at 5%; ** significant at 1%

Table A4: Sensitivity analysis among women who delivered on or past their due date. Multivariable linear probability, linear regression, and Poisson regression results for negative delivery outcomes (any and count), labor induction, labor duration, unplanned C-section, and negative birth experience in propensity-score matched (Panel A) and unmatched (Panel B) groups among employed, nulliparous women in Pennsylvania (N = 1,190). Coefficients and 95% confidence intervals.

	Any negative delivery outcome	Number of negative delivery outcomes	Labor induced	Labor >24 hours	Hours in labor	Unplanned C-section	Negative birth experience
	Coeff. [95% CI]	Coeff. [95% CI]	Coeff. [95% CI]	Coeff. [95% CI]	Coeff. [95% CI]	Coeff. [95% CI]	Coeff. [95% CI]
Panel A: Delivery on or post due date only							
ANL (≥2 days)	0.06 [0.00 - 0.12]*	0.20 [0.07 - 0.32]**	0.06 [0.00 - 0.11]*	0.03 [-0.01 - 0.07]	1.20 [0.09 - 2.31]*	0.09 [0.04 - 0.14]**	0.00 [-0.05 - 0.04]
Mean of negative delivery outcomes	0.59	0.87	0.32	0.13	15.12	0.24	0.19
Weighted observations	1190	1190	1190	1123	1123	1148	1185
R-squared/Pseudo R-squared	0.02	0.01	0.02	0.03	0.03	0.02	0.02
Panel B: ANL relative to due date							
ANL (≥2 days)	-0.01 [-0.03 - 0.00]	-0.03 [-0.05 - 0.00]	-0.00 [-0.02 - 0.01]	-0.00 [-0.02 - 0.01]	0.04 [-0.27 - 0.35]	-0.00 [-0.02 - 0.01]	-0.01 [-0.02 - -0.00]
Mean of negative delivery outcomes	0.57	0.84	0.30	0.14	15.14	0.24	0.18
Observations	383	383	383	359	359	361	379
R-squared/Pseudo R-squared	0.05	0.07	0.06	0.01	0.00	0.06	0.08

Abbreviations: CI, confidence interval; ANL, antenatal leave.

Notes: Coefficients and 95% confidence intervals for linear probability models (any negative delivery outcome, labor induction, long labor duration, unplanned C-section, and negative birth experience), Poisson regression (number of negative delivery outcomes), and linear regression models (hours in labor). Panel A includes only women who delivered on or past their due date. Panel B includes only women who delivered past their due date and stopped working before their due date (i.e., likely planned ANL). We calculated ANL relative to women's due date rather than actual delivery date. Models in Panel A are weighted to account for the number of units each observation was matched with and adjust for the propensity score, pre-pregnancy obesity, strong preference for vaginal delivery, perceived stress, employment status and race/ethnicity. Models in Panel B are unweighted and no not adjusted for the propensity score. Negative delivery outcomes include labor induction, labor lasting more than 24 hours, unplanned C-section, and self-reported negative birth experience.

* significant at 5%; ** significant at 1%

Table A5: Sensitivity analysis among women who quit their jobs or were fired. Multivariable linear probability, linear regression, and Poisson regression results for negative delivery outcomes (any and count), labor induction, labor duration, unplanned C-section, and negative birth experience in propensity-score matched groups among employed, nulliparous women in Pennsylvania, excluding women who took ANL (N = 1,474). Coefficients and 95% confidence intervals.

	Any negative delivery outcome	Number of negative delivery outcomes	Labor induced	Labor >24 hours	Hours in labor	Unplanned C-section	Negative birth experience
	Coeff. [95% CI]	Coeff. [95% CI]	Coeff. [95% CI]	Coeff. [95% CI]	Coeff. [95% CI]	Coeff. [95% CI]	Coeff. [95% CI]
Quit or were fired (vs. no ANL)	0.07 [0.02 - 0.12]**	0.21 [0.09 - 0.33]**	0.10 [0.05 - 0.14]**	0.02 [-0.02 - 0.06]	0.21 [-0.97 - 1.40]	0.04 [-0.01 - 0.08]	0.02 [-0.02 - 0.07]
Mean of negative delivery outcomes	0.59	0.88	0.34	0.12	14.35	0.24	0.20
Weighted observations	1474	1474	1474	1364	1364	1414	1456
R-squared/Pseudo R-squared	0.04	0.04	0.15	0.03	0.01	0.06	0.03

Abbreviations: CI, confidence interval; ANL, antenatal leave.

Notes: Coefficients and 95% confidence intervals for linear probability models (any negative delivery outcome, labor induction, long labor duration, unplanned C-section, and negative birth experience), Poisson regression (number of negative delivery outcomes), and linear regression models (hours in labor). All models are weighted to account for the number of units each observation was matched with and adjust for the propensity score, pre-pregnancy obesity, strong preference for vaginal delivery, perceived stress, gestational age at delivery, employment status and race/ethnicity. Negative delivery outcomes include labor induction, labor lasting more than 24 hours, unplanned C-section, and self-reported negative birth experience. We characterize women who reported being employed during pregnancy at the baseline interview later responded that they were no longer employed in the 2 weeks prior to delivery as women who quit or were fired. We exclude women who took ANL from these analyses.

* significant at 5%; ** significant at 1%

Table A6: Sensitivity analysis, by maternal health. Multivariable linear probability, linear regression, and Poisson regression results for negative delivery outcomes (any and count), labor induction, labor duration, unplanned C-section, and negative birth experience in propensity-score matched groups among employed, nulliparous women in Pennsylvania (N = 1,740). Coefficients and 95% confidence intervals.

	Any negative delivery outcome	Number of negative delivery outcomes	Labor induced	Labor >24 hours	Hours in labor	Unplanned C-section	Negative birth experience
	Coeff. [95% CI]	Coeff. [95% CI]	Coeff. [95% CI]	Coeff. [95% CI]	Coeff. [95% CI]	Coeff. [95% CI]	Coeff. [95% CI]
Panel A: Healthy women							
ANL (≥2 days)	0.02 [-0.04 - 0.09]	0.07 [-0.09 - 0.23]	0.02 [-0.03 - 0.08]	0.03 [-0.01 - 0.07]	0.79 [-0.43 - 2.01]	0.02 [-0.03 - 0.08]	-0.01 [-0.06 - 0.04]
Mean of negative delivery outcomes	0.50	0.71	0.25	0.10	13.76	0.21	0.17
Weighted observations	945	945	945	893	893	908	940
R-squared/Pseudo R-squared	0.10	0.04	0.11	0.02	0.03	0.04	0.04
Panel B: Unhealthy women							
ANL (≥2 days)	0.10 [0.03 - 0.16]**	0.26 [0.11 - 0.41]**	0.06 [-0.01 - 0.12]	0.05 [-0.00 - 0.10]	1.47 [-0.05 - 2.99]	0.11 [0.05 - 0.17]**	0.04 [-0.01 - 0.10]
Mean of negative delivery outcomes	0.61	0.92	0.34	0.13	14.85	0.26	0.22
Weighted observations	868	868	868	781	781	798	865
R-squared/Pseudo R-squared	0.06	0.03	0.08	0.02	0.02	0.06	0.07

Abbreviations: CI, confidence interval; ANL, antenatal leave.

Notes: Coefficients and 95% confidence intervals for linear probability models (any negative delivery outcome, labor induction, long labor duration, unplanned C-section, and negative birth experience), Poisson regression (number of negative delivery outcomes), and linear regression models (hours in labor). Panel A (healthy women) includes only women who reported no pre-pregnancy hypertension or diabetes, serious health conditions, or hospitalizations during pregnancy; were not advised by a provider during pregnancy to have a C-section; reported <5 doctor office visits in last month of pregnancy; and were not depressed during pregnancy. Panel B (unhealthy women) includes only women with at least one of those conditions. All models are weighted to account for the number of units each observation was matched with and adjust for the propensity score, pre-pregnancy obesity, strong preference for vaginal delivery, perceived stress, gestational age at delivery, employment status, maternal age, (maternal age)^2, and race/ethnicity. Negative delivery outcomes include labor induction, labor lasting more than 24 hours, unplanned C-section, and self-reported negative birth experience.

* significant at 5%; ** significant at 1%

Table A7: Sensitivity analysis with categorical ANL variable. Multivariable linear probability, linear regression, and Poisson regression results for negative delivery outcomes (any and count), labor induction, labor duration, unplanned C-section, and negative birth experience in unmatched groups among employed, nulliparous women in Pennsylvania (N = 1,740). Coefficients and 95% confidence intervals.

	Any negative delivery outcomes		Number of negative delivery outcomes		Labor induced	
	Unadjusted Coeff. [95% CI]	Adjusted Coeff. [95% CI]	Unadjusted Coeff. [95% CI]	Adjusted Coeff. [95% CI]	Unadjusted Coeff. [95% CI]	Adjusted Coeff. [95% CI]
ANL (ref: no ANL)						
2 days	0.02 [-0.05 - 0.10]	0.01 [-0.06 - 0.08]	0.11 [-0.06 - 0.27]	0.06 [-0.10 - 0.23]	0.01 [-0.06 - 0.07]	-0.01 [-0.07 - 0.06]
3-6 days	0.12 [0.05 - 0.18]**	0.09 [0.02 - 0.15]**	0.3 [0.16 - 0.44]**	0.21 [0.06 - 0.35]**	0.07 [0.00 - 0.13]*	0.04 [-0.02 - 0.10]
7 days	0.12 [0.05 - 0.19]**	0.07 [-0.00 - 0.14]	0.3 [0.15 - 0.46]**	0.19 [0.03 - 0.35]*	0.11 [0.05 - 0.18]**	0.07 [0.00 - 0.14]*
>7 days	0.14 [0.04 - 0.24]**	0.06 [-0.04 - 0.16]	0.31 [0.10 - 0.51]**	0.14 [-0.08 - 0.36]	0.11 [0.01 - 0.20]*	0.02 [-0.07 - 0.11]
Mean of negative delivery outcomes	0.55	0.55	0.82	0.82	0.31	0.31
Observations	1740	1740	1740	1740	1740	1740
R-squared/Pseudo R-squared	0.01	0.10	0.01	0.04	0.01	0.11

Abbreviations: CI, confidence interval; ANL, antenatal leave.

Notes: Coefficients and 95% confidence intervals for linear probability models (any negative delivery outcome, labor induction, long labor duration, unplanned C-section, and negative birth experience), Poisson regression (number of negative delivery outcomes), and linear regression models (hours in labor). Adjusted models control for pre-pregnancy obesity; whether the woman gained more weight than recommended during pregnancy; prior miscarriage; self-reported history of diabetes or hypertension before pregnancy; presence of medical problems during current pregnancy (hypertension, high blood pressure, or pre-eclampsia that started during pregnancy; diabetes that started during pregnancy; vaginal bleeding; early or pre-term labor; bed rest or hospitalization because of premature labor); number of hospitalizations during pregnancy; number of doctor office or urgent care visits in the last month of pregnancy; whether woman received help getting pregnant; whether a provider advised a C-section during pregnancy; fear of childbirth measured with the First Baby Study Birth Anticipation Scale (BAS); strong preference for vaginal delivery; prenatal stress using Perceived Stress Scale; social support; baseline EPDS; gestational age at delivery; pregnancy intention; maternal education; maternal age and age squared; race/ethnicity; whether married or living with partner; insurance and poverty status. Negative delivery outcomes include labor induction, labor lasting more than 24 hours, unplanned C-section, and self-reported negative birth experience.

* significant at 5%; ** significant at 1%

Labor >24 hours		Hours in labor		Unplanned C-section		Negative birth experience	
Unadjusted Coeff. [95% CI]	Adjusted Coeff. [95% CI]	Unadjusted Coeff. [95% CI]	Adjusted Coeff. [95% CI]	Unadjusted Coeff. [95% CI]	Adjusted Coeff. [95% CI]	Unadjusted Coeff. [95% CI]	Adjusted Coeff. [95% CI]
0.05 [-0.00 - 0.09]	0.03 [-0.02 - 0.08]	2.17 [0.71 - 3.63]**	1.69 [0.21 - 3.16]*	0.02 [-0.04 - 0.08]	0.02 [-0.04 - 0.08]	0.02 [-0.04 - 0.08]	0.01 [-0.05 - 0.06]
0.09 [0.05 - 0.13]**	0.08 [0.03 - 0.12]**	2.2 [0.88 - 3.52]**	1.79 [0.43 - 3.14]**	0.09 [0.03 - 0.15]**	0.09 [0.03 - 0.15]**	0.03 [-0.03 - 0.08]	0.01 [-0.04 - 0.06]
0.04 [-0.00 - 0.09]	0.02 [-0.03 - 0.07]	1.56 [0.12 - 3.00]*	1.03 [-0.48 - 2.54]	0.09 [0.02 - 0.15]**	0.06 [-0.00 - 0.13]	0.02 [-0.03 - 0.08]	0.01 [-0.05 - 0.06]
0.04 [-0.02 - 0.11]	0.02 [-0.05 - 0.09]	0.65 [-1.38 - 2.68]	0.07 [-2.04 - 2.19]	0.08 [-0.01 - 0.17]	0.06 [-0.03 - 0.15]	0.04 [-0.04 - 0.12]	0.02 [-0.06 - 0.10]
0.11	0.11	14.20	14.20	0.24	0.24	0.19	0.19
1606	1606	1606	1606	1635	1635	1730	1730
0.01	0.03	0.01	0.04	0.01	0.07	0.00	0.10

Chapter 4
The impact of California's Paid Family Leave law on maternal time use

Abstract

In 2004, California became the first U.S. state to institute a paid family leave (PFL) law. I exploit this natural experiment to examine how access to paid leave affected maternal time use. Using triple difference estimation on data from the American Time Use Survey, I test whether after PFL implementation, mothers of infants (children less than 1 year of age) living in California reduced their time spent working and increased time spent on childcare relative to mothers of older children and to all mothers outside California. The results suggest that a significant association exists between PFL in California and how mothers of young children spend their time. After PFL, women in California increased the time they spent with children in their care by approximately three hours per day ($p<.01$) and reduced their time spent working ($p<.10$). All observed associations were limited to mothers of infants, lending support to the argument that the observed changes in mothers of young children were related to PFL.

Keywords: paid family leave; work; childcare; time use; female labor supply

Introduction

Family leave policies aim to enable workers to take time from work in order to care for themselves or for family members, making a career and childrearing more compatible. Much evidence suggests that the availability of paid leave has positive impacts on child health (Baum, 2003; Berger, Hill, & Waldfogel, 2005; Brooks-Gunn, Han, & Waldfogel, 2002; Tanaka, 2005), maternal mental and physical health (Pat McGovern, Dowd, Gjerdingen et al., 2007; P McGovern, Dowd, Gjerdingen et al., 1997), and increases breastfeeding (Berger, Hill, & Waldfogel, 2005; Guendelman, Kosa, Pearl et al., 2009; Kimbro, 2006; Kirkland & Fein, 2003). Studies outside the United States have shown that increasing paid family leave mandates results in increased leave-taking among mothers of infants (Baker & Milligan, 2008; Kluve & Tamm, 2009).

The United States stands apart as one of the few countries in the world that lacks a national paid leave law and, until recently, any state paid leave laws. The federal Family and Medical Leave Act of 1993 (FMLA) enables many women in the U.S. to take unpaid leave with job security. FMLA provides up to 12 weeks of unpaid, job-protected leave for one's own illness, the illness of a qualified family member, or to care for a newborn or newly adopted baby. Only employers with more than 50 employees and employees who have worked at least 1,250 hours in the past year are covered.

Studies of FMLA suggest a limited effect on leave-taking, plausibly due to strict eligibility criteria and because it is entirely unpaid (Baum, 2003). Eligible employees must work for firms with more than 50 employees within a 75-mile radius and must have worked at least 1,250 hours in the past year, resulting in just over half (59%) of all U.S. workers being covered (Klerman, Daley, & Pozniak, 2012). FMLA covers even fewer new mothers, as this group is less likely to be employed in the year before birth, meet the hours requirement and work for a covered firm (Ruhm, 1997). For those who are eligible, many find it impossible to go three months without pay and do not use available leave (Klerman, Daley, & Pozniak, 2012).

American women use various strategies for piecing together maternity leave. Some employees may qualify for paid leave during and just after pregnancy if their employers offer short-term disability insurance. According to the Pregnancy Discrimination Act of 1978, employers that offer short-term disability insurance to their employees must include pregnancy-related disability as a covered condition. Five states (California, Hawaii, Rhode Island, New Jersey, and New York) have temporary disability insurance laws and most, if not all, workers in those states have access to paid short-term disability insurance. In California, women can use this short-term insurance to receive partial income replacement for up to six weeks after childbirth (eight weeks, if delivery was by Cesarean section). In addition to these policies, women may have maternity leave offered by their employers, either formally or informally. Finally, some women quit their jobs in order to stay home with very young children.

In 2004, California became the first U.S. state to institute a paid family leave law. California's Paid Family Leave (PFL) law entitles any worker who pays into the State Disability Insurance (SDI) fund to six weeks of leave with income replacement up to 55 percent of prior wages. PFL can be used to care for a seriously ill child, spouse, parent, or registered domestic partner or to bond with a newborn or newly adopted child. For bonding, PFL must be used within a year of the birth or adoption of a child (Employment

Development Department, 2013a). This program builds on the short-term disability leave previously available in California, doubling the amount of post-partum maternity leave available to birth mothers. Unlike FMLA, PFL covers almost all private sector workers in California.

As the first state in the country to pass a paid family leave law, California provides a natural experiment in which to examine the relationship between increased paid family leave and how parents of very young children spend their time. While it appears that unpaid leave mandates have little effect on the general population, we can exploit the timing of PFL in California to examine whether partial income replacement allows more women to spend time caring for infants. I hypothesized that in the years after PFL implementation, mothers of infants (children less than 1 year of age) living in California reduced their time spent in paid work and increased time spent on childcare relative to mothers of older children and to all mothers outside California.

Data

I used data from the American Time Use Survey (ATUS) to estimate changes in how mothers of very young children spent their time in response to California's Paid Family Leave law (PFL). The ATUS data are collected and processed by the U.S. Census Bureau (American Time Use Survey User's Guide, 2010). Households that have completed their final month of the Current Population Survey (CPS) can be contacted to participate in the ATUS. One person who is at least age 15 is randomly selected from the household and asked questions about his or her time use on a given day. Data are collected through computer-assisted telephone interviewing (CATI), available in English or Spanish. The ATUS sampling was randomized by day of the week with half of the sample reporting about weekdays and half reporting about weekend days. Data files from the ATUS were linked with CPS data files. Response rates for the ATUS ranged from a high of 57.8 percent in 2003 to a low of 52.5 percent in 2007.

My primary analysis included employed adult women with a child under age six (N=4,586). Employment status corresponded to whether the respondent reported being employed (either at work or absent) at the final CPS interview which took place between two and five months before the ATUS interview. I applied sampling weights to account for oversampling of certain groups, as well as differential response rates by demographic characteristics and days of the week. Additionally, the ATUS is not uniformly distributed across the day of the week; unweighted estimates will overestimate time spent on weekend activities and underestimate time spent on weekday activities. The final weights indicate the number of person-days each respondent represents. In order to use geographic, demographic, employment and time use data, I combined the Respondent, Activity summary, and ATUS-CPS files.

California's Paid Family Leave law (PFL) was signed into law in 2002, but did not take effect until July 2004. To examine differences in time use before and after PFL implementation, I merged ATUS data files from 2003, 2004 (January to June only), 2006, 2007 and 2008. I defined post-treatment years starting in 2006 in order to allow time for Californians to learn about the new law. I created a pre-/post-PFL dummy variable using 2006 to 2008 data as the post-PFL period. As California was the only state with a PFL law

during this time period, I created a dummy for residence in California versus all other states. Finally, I created a dummy variable for women with children under age 1 (those most likely affected by the law) versus women with children ages 1 to 6.

Time use responses are coded into 17 major categories, each with two additional levels of detail. I defined work as the total number of hours spent on work and work-related activities, including working at a job, work-related activities, engaging in other income-generating activities, and searching and interviewing for jobs. I defined primary childcare as any activity related to caring for household children, activities related to children's education, and activities related to children's health. Respondents separately reported secondary childcare—care given while engaging in other activities (e.g., respondent indicates that children were under their supervision while primary activity was preparing dinner). Secondary childcare is not recorded when the primary activity is childcare. Total childcare is simply the sum of primary and secondary childcare, or the total amount of time spent caring for household children. To illustrate the types of activities respondents reported within these categories and the distribution of their time use among these activities, Table 1 presents a detailed list of primary childcare and work-related activities and the weighted mean number of minutes that women in the 2003 cohort spent on each activity in the survey day.

I included a standard set of demographic covariates, including the number of household children under age 18; marital/partner status (whether or not the woman has a spouse or partner living in the home); educational attainment (less than high school, high school graduate, some college, college graduate); Hispanic ethnicity; and maternal age.

Methods

I used a difference-in-difference-in-difference (DDD) analysis to examine time use among mothers of very young children in California compared to mothers of older children and to all mothers outside California. This analysis used variation in time (pre-PFL vs. post-PFL), place (California vs. other states), and target age group (mothers of infants vs. mothers of children aged 1 to 6) to estimate changes in time use in the population expected to respond to the law: Californian mothers of infants after 2004. I exploited the fact that California was the only state to introduce paid family leave during the time period specified. My pre-treatment group was all employed adult women with youngest child under age six in the 2003 and 2004 (January through June) files. Post-treatment respondents were employed adult women with youngest child under age six in the 2006 to 2008 files. Labor force participation trends for mothers of young children may differ between California and other states over the study period, so I included a third comparison of mothers of children under age 1—who are among the primary targets of PFL—and mothers of children aged 1 to 6, who should be less likely to take advantage of PFL but can be expected to have similar state-level labor market trends as mothers of younger children.

I estimated the following equation:

$$Y_{isya} = \alpha + \beta_1 Post_y + \beta_2 CA_s + \beta_3 Infant_a + \beta_4 Post_y x CA_s + \beta_5 Post_y x Infant_a + \beta_6 CA_s x Infant_a + \tau CA_s x Infant_a x Post_y + \varepsilon_{isya}$$

for each individual *i* in state *s* in year *y* for age group a. Y_{isya} is either hours spent caring for household children or hours spent on work and work-related activities in the day of interview. $Post_y$ is an indicator equal to 1 if the interview took place after PFL was implemented (2006-2008), and 0 otherwise (2003-June 2004). CA_s is an indicator equal to 1 if the respondent lived in CA, and 0 otherwise. $Infant_a$ is an indicator equal to 1 if the respondent had a child under age 1, and 0 otherwise. ε_{isya} is an individual-specific error term. The coefficient of interest is τ, which measures the difference-in-difference-in-difference estimate of the effect of PFL on time use in California among parents of very young children. I further estimated the equation in the presence of demographic controls, including number of children under 18 and whether or not a spouse or partner was present in the household; maternal education; Hispanic ethnicity; and age. Respondents could have come from the same household; however, clustering standard errors at the household level did not significantly change the results (Appendix, Table A1).

As a sensitivity test, I repeated all analyses among fathers of young children. Though fathers increasingly use PFL in California, they remain far less likely than mothers to take parental leave. Approximately 88 percent of PFL claims filed between 2004 and 2009 were for bonding and, of those, more than three-quarters were filed by females (Employment Development Department, 2013b).

Results

Table 2 shows means and standard deviations for each outcome and covariate in each of the eight time-place-age groups (pre-PFL vs. post-PFL, California vs. other states, and mothers of infants vs. mothers of children aged 1 to 6), as well as the first, second, and third differences. The first difference (D) compares women in California to women in other states on each variable. The difference of interest is within the post-PFL period among mothers of infants; other first differences illustrate the appropriateness of each comparison population. The second difference (DD) examines how the difference between California and other states has changed over time for each age group. The comparison of interest is whether California differed from other states in the post-PFL period after subtracting the difference between California and other states in the pre-PFL period. Finally, the third difference (DDD) subtracts the trend in the differences among mothers of older children who should not be affected by PFL, but who otherwise might follow similar demographic and time use trends as mothers of infants. For example, among mothers of infants in the post-PFL period, Californians were significantly more likely to be Hispanic than non-Californians. This difference was larger in the post-PFL period than it was in the pre-PFL period. However, this trend of an increasing difference between California and other states was the same among mothers of older children, so the increasing share of Hispanic women in California over time should not bias my estimates in the DDD models. Of the covariates, only maternal education remained statistically significantly different in the third difference. Mothers of infants post-PFL in California were significantly more likely to have less than a high school degree than comparable women in other states, even after removing the difference in the pre-PFL period and the change in the difference among mothers of older children.

Table 3 presents the estimates for time (in hours) spent on work and work-related activities. Columns (1) and (2) show difference-in-difference (DD) estimates for women with children under age 1 and women with older children, respectively. Residing in California after PFL was not significantly associated with a change in time spent on work in either of these models, though the estimate approached statistical significance among mothers of infants (-1.90 hours; $p<.10$). In column (3), these models are combined into a difference-in-difference-in-difference (DDD) model. The coefficient on the triple interaction term did not reach statistical significance, but suggests a decrease of approximately 1.77 hours worked. Small cell sizes may have made it difficult to detect significance in the DDD model. Columns (4) through (6) present these same models, but include demographic controls. After controlling for the number of children under 18, whether or not a spouse or partner was present in the household, maternal education, maternal ethnicity, and maternal age, women in California after PFL worked 2.65 hours less than women in California before PFL and women in other states ($p<.05$). As expected, there was no relationship among mothers of older children. Accounting for the trend among mothers of older children in the DDD model, the relationship between PFL and work persisted but was of marginal statistical significance.

Table 4 presents results for time spent on primary childcare. In models without additional controls, PFL was not significantly associated with primary childcare. After controlling for covariates, PFL was associated with an increase in time spent on primary childcare among mothers of infants ($p<.05$), but this relationship went away after controlling for the trend among mothers of older children in the DDD model.

Table 5 shows results for time spent on secondary childcare. PFL was significantly associated with secondary childcare among mothers of infants, both with and without adjusting for potential confounders. Among mothers of infants, PFL was associated with an increase of 3.29-3.80 hours per day spent on secondary childcare ($p<.01$). After controlling for trends among mothers of older children, these estimates dropped to 2.96-3.14 hours per day in the DDD models, but remained statistically significant.

I observed a similar pattern in the results for total childcare, which sums primary and secondary childcare (Table 5). Among mothers of infants, PFL was associated with an increase of 4.05-5.19 hours per day spent on total childcare ($p<.01$). This dropped to an increase of 3.56-4.18 hours per day in the DDD models, but remained statistically significant ($p<.01$).

As a sensitivity analysis, I repeated all analyses among fathers, a group much less likely to take PFL than mothers. As expected, PFL was not associated with time spent on work or childcare among fathers (Tables 7-10).

Discussion

This is one of the first studies to examine how California's landmark Paid Family Leave law affected time use among mothers of young children. The results suggest that there is a significant association between PFL in California and how mothers of young children spend their time. After PFL, women in California increased on average the time they spent with children in their care by approximately three hours per day. Their total time spent on childcare activities increased on average by over four hours per day. This

may have been possible through a reduction in hours spent working, though small sample sizes made detection of such an effect difficult. Women in California spent about two fewer hours working than their counterparts outside the state with a 95 percent confidence interval that includes negative four hours. Though not statistically significant, the results indicating some reduction in work hours are consistent with prior studies on paid leave in California. Rossin-Slater, Ruhm, and Waldfogel (2012) used CPS data to examine maternity leave-taking and employment levels among California woman and found evidence that PFL doubled the overall use of maternity leave.

The estimate of four hours per day may be implausibly large. In 2006, approximately 553,000 California women gave birth and 132,000 filed PFL claims for bonding for an average of 5.37 weeks (Centers for Disease Control and Prevention, 2015; Employment Development Department, 2013b). This implies that in any given week in 2006, only 2.5% of California mothers of infants should be using PFL. Moreover, most PFL claims are filed in the first 3 months after childbirth. The true relationship between PFL and childcare may fall closer to the bottom of the confidence interval— an increase of a little more than one hour per day.

Interestingly, I found that PFL increased secondary childcare while having only a small impact on primary childcare activities. This indicates that while mothers of young children did not increase the amount of time they dedicated exclusively to childcare, they did spend significantly more time with children in their care. If the time needed for non-childcare activities (e.g., housework, eating, personal hygiene) is relatively fixed but PLF allowed more mothers of infants to stay home with them, it is plausible that some fraction of the activities previously done while children were either at daycare or asleep may have occurred while children were present and therefore would be recorded as an increase in secondary, not primary, childcare.

All observed associations were limited to mothers of infants. Mothers of older children and fathers did not significantly change their time use after PFL, lending support to the argument that the observed changes in mothers of young children were related to legislation that targeted them (i.e., PFL).

This study does not examine maternity leave directly. The ATUS provides important information about how time was actually spent, but we do not know whether women were on maternity leave, were working reduced hours in order to care for their children, or were unwillingly working reduced hours due to furloughs or cutbacks as the economy began to contract. We might also see a reduction in average hours worked if some fraction of the population were unemployed, potentially as a result of PFL. Further examination of why women outside California increased their time spent on work and work-related activities over the study period are needed. Interpretation of these results is limited by small sample sizes in some subpopulations. While my overall sample size was 4,586 women, the number of mothers of infants living in California in the pre-treatment period was quite small. The ATUS began collecting data in 2003—a year and a half before PFL implementation—so the pre-treatment group cannot be expanded. Other studies using triple-difference models with a state comparison suffer from similar small size concerns (Bruckner & Nobles, 2013).

ATUS and linked CPS data did not include data on maternal or child health conditions, so I was unable to account for potential differences along these dimensions. Women who had difficult pregnancies or childbirth experiences, or those whose children

had health problems, may have been much more likely to take advantage of PFL. Further study is needed to examine the effects of PFL on maternity leave taken by using direct measures of paid and unpaid time off of work after the birth of a child. It will also be important to see whether women in states with more recently enacted paid family leave laws (i.e., New Jersey and Rhode Island) respond similarly. Further research should be conducted once sufficient post-treatment data become available in those states.

Acknowledgements
Grant sponsor: National Institute of Child Health & Human Development; Grant Number: T32 HD07275.

References

American Time Use Survey User's Guide. (2010). U.S. Census Bureau and U.S. Bureau of Labor Statistics.

Baker, M., & Milligan, K. (2008). How does job protected maternity leave affect mothers' employment? Journal of Labor Economics, 26, 655-692.

Baum, C. (2003). The effects of maternity leave legislation on mothers' labor supply after childbirth. Southern Economic Journal, 69, 772-799.

Berger, L. M., Hill, J., & Waldfogel, J. (2005). Maternity leave, early maternal employment and child health and development in the US*. The Economic Journal, 115, F29-F47.

Brooks-Gunn, J., Han, W., & Waldfogel, J. (2002). Maternal employment and child cognitive outcomes in the first three years of life: The NICHD study of early child care." National Institute of Child Health and Human Development. Child Development, 73, 1052-1072.

Bruckner, T. A., & Nobles, J. (2013). Intrauterine stress and male cohort quality: The case of September 11, 2001. Social Science & Medicine, 76, 107-114.

Centers for Disease Control and Prevention. National Center for Health Statistics. VitalStats. Retrieved from: /nchs/vitalstats.htm. Access date: July 22, 2015.

Employment Development Department. (2013a). Paid Family Leave (PFL) Fact Sheet. Sacramento, CA: State of California.

Employment Development Department. (2013b). Paid Family Leave Program Statistics: State of California.

Guendelman, S., Kosa, J. L., Pearl, M., Graham, S., Goodman, J., & Kharrazi, M. (2009). Juggling work and breastfeeding: effects of maternity leave and occupational characteristics. Pediatrics, 123, e38-e46.

Kimbro, R. (2006). On-the-job moms: work and breastfeeding initiation and duration for a sample of low-income women. Maternal and Child Health Journal, 10, 19-26.

Kirkland, V., & Fein, S. (2003). Characterizing reasons for breastfeeding cessation throughout the first year postpartum using the construct of thriving. Journal of Human Lactation, 19, 278-285.

Klerman, J. A., Daley, K., & Pozniak, A. (2012). Family and medical leave in 2012: Technical report. Cambridge, MA: Abt Associates Inc.

Kluve, J., & Tamm, M. (2009). *Now Daddy's Changing Diapers and Mommy's Making Her Career: Evaluating a Generous Parental Leave Regulation Using a Natural Experiment*. Paper presented at the IZA, Bonn.

McGovern, P., Dowd, B., Gjerdingen, D., Dagher, R., Ukestad, L., McCaffrey, D., & Lundberg, U. (2007). Mothers' health and work-related factors at 11 weeks postpartum. The Annals of Family Medicine, 5, 519-527.

McGovern, P., Dowd, B., Gjerdingen, D., Moscovice, I., Kochevar, L., & Lohman, W. (1997). Time off work and the postpartum health of employed women. Medical Care, 35, 507-521.

Rossin‐Slater, M., Ruhm, C. J., & Waldfogel, J. (2012). The Effects of California's Paid Family Leave Program on Mothers' Leave‐Taking and Subsequent Labor Market Outcomes. Journal of Policy Analysis and Management.

Ruhm, C. J. (1997). Policy Watch: The Family and Medical Leave Act. Journal of Economic Perspectives, 11, 175-186.

Tanaka, S. (2005). Parental leave and child health across OECD countries. The Economic Journal, 115, F7-F28.

Tables

Table 1: Average time spent on each work and primary childcare activity in 2003 among employed adult women with youngest child under age 6, weighted means (N = 1,382). ATUS, 2003.

Primary childcare		Work	
Activity	Minutes/day	Activity	Minutes/day
Physical care for household children	61.68 (2.40)	Work, main job	238.00 (7.98)
Reading to/with household children	4.58 (0.39)	Work, other job(s)	4.25 (1.56)
Playing with household children, not sports	27.93 (1.92)	Security procedures related to work	0 (0)
Arts and crafts with household children	0.16 (0.07)	Working, n.e.c.	0 (0)
Playing sports with household children	0.66 (0.33)	Socializing, relaxing, and leisure as part of job	0 (0)
Talking with/listening to household children	3.07 (0.54)	Eating and drinking as part of job	0.52 (0.27)
Helping/teaching household children (not rel. to educ.)	0.85 (0.17)	Sports and exercise as part of job	0.02 (0.02)
Organization & planning for household children	1.16 (0.23)	Security procedures as part of job	0 (0)
Looking after household children (as a primary activity)	5.27 (0.76)	Work-related activities, n.e.c.	0.20 (0.07)
Attending household children's events	2.93 (0.54)	Income-generating hobbies, crafts, and food	0 (0)
Waiting for/with household children	0.81 (0.16)	Income-generating performances	0 (0)
Picking up/dropping off household children	5.84 (0.39)	Income-generating services	0.76 (0.50)
Caring for & helping household children, n.e.c.	1.77 (0.93)	Income-generating rental property activities	0 (0)
Homework (household children)	4.15 (0.57)	Other income-generating activities, n.e.c.	0.91 (0.55)
Meetings and school conferences (household children)	0.50 (0.17)	Active job search	0.23 (0.23)
Homeschooling of household children	0.22 (0.15)	Other job search activities	0.02 (0.02)
Waiting associated with household children's education	0 (0)	Job interviewing	0.18 (0.14)
Activities related to household child's education, n.e.c	0.11 (0.06)	Waiting associated with job search or interview	0 (0)
Providing medical care to household children	2.25 (0.55)	Job search and Interviewing, n.e.c.	0 (0)
Obtaining medical care for household children	0.96 (0.26)	Work and work-related activities, n.e.c.	0 (0)
Waiting associated with household children's health	0.99 (0.31)		
Activities related to household child's health, n.e.c.	0.09 (0.06)		

Abbreviations: n.e.c., not elsewhere classified

Table 2: Summary statistics. ATUS, 2003-2008.

Panel A: Mothers of children under 1	Pre-PFL (2003-2004)					Post-PFL (2006-2008)					
	CA		Other states		D[1]	CA		Other states		D[1]	DD[2]
Characteristic	Mean	SD	Mean	SD		Mean	SD	Mean	SD		
Work (hours)	4.72	0.74	3.34	0.32	1.38 +	3.57	0.72	4.09	0.23	-0.52	-1.90 +
% with 0 work hours	0.32	0.10	0.51	0.04	-0.19 +	0.40	0.11	0.39	0.03	0.00	0.19
Primary caregiving (hours)	2.28	0.44	3.37	0.20	-1.09 *	2.47	0.62	2.80	0.14	-0.33	0.76
% with 0 primary caregiving hours	0.01	0.01	0.07	0.02	-0.05 **	0.11	0.06	0.10	0.02	0.00	0.05
Secondary caregiving (hours)	3.65	0.77	6.26	0.32	-2.62 **	6.30	0.66	5.63	0.29	0.67	3.29 **
% with 0 secondary caregiving hours	0.15	0.13	0.03	0.01	0.12	0.05	0.03	0.08	0.02	-0.03	-0.15
Total caregiving (hours)	5.92	1.07	9.63	0.39	-3.71 **	8.77	0.96	8.43	0.33	0.34	4.05 **
% with 0 total caregiving hours	0.00	0.00	0.01	0.01	-0.01 **	0.03	0.02	0.05	0.02	-0.02	0.00
Number of children <18	2.36	0.24	2.27	0.12	0.09	2.66	0.74	2.01	0.08	0.65	0.55
One household child	0.26	0.10	0.35	0.04	-0.09	0.33	0.10	0.41	0.03	-0.09	0.01
No spouse/partner present	0.33	0.13	0.14	0.02	0.19	0.45	0.14	0.23	0.03	0.21	0.02
Maternal education											
<HS education	0.08	0.06	0.11	0.03	-0.03	0.37	0.15	0.13	0.02	0.25	0.27
HS grad	0.14	0.07	0.31	0.04	-0.17 *	0.10	0.06	0.27	0.03	-0.17 **	0.01
Some college	0.26	0.12	0.17	0.03	0.09	0.20	0.08	0.18	0.02	0.02	-0.07
College grad	0.52	0.12	0.42	0.04	0.11	0.32	0.10	0.43	0.03	-0.10	-0.21
Hispanic ethnicity	0.42	0.12	0.17	0.03	0.25 *	0.68	0.10	0.15	0.02	0.54 **	0.29 +
Age	29.93	1.58	30.77	0.64	-0.84	32.89	1.54	30.17	0.52	2.72 +	3.56
N	31		291			40		472			

Panel B: Mothers of children 1 to 6	Pre-PFL (2003-2004)					Post-PFL (2006-2008)						
	CA		Other states		D[1]	CA		Other states		D[1]	DD[2]	DDD[3]
Characteristic	Mean	SD	Mean	SD		Mean	SD	Mean	SD			
Work (hours)	4.21	0.42	4.17	0.15	0.05	4.55	0.33	4.63	0.11	-0.08	-0.12	-1.77
% with 0 work hours	0.39	0.05	0.39	0.02	0.00	0.35	0.04	0.34	0.01	0.01	0.00	0.19
Primary caregiving (hours)	1.44	0.13	1.84	0.06	-0.40 **	1.48	0.16	1.72	0.05	-0.24	0.16	0.60
% with 0 primary caregiving hours	0.23	0.05	0.16	0.02	0.07	0.26	0.04	0.16	0.01	0.10 **	0.03	0.03
Secondary caregiving (hours)	5.14	0.40	5.68	0.15	-0.54	5.45	0.44	5.67	0.12	-0.21	0.33	2.96 **
% with 0 secondary caregiving hours	0.12	0.04	0.08	0.01	0.04	0.17	0.04	0.09	0.01	0.08 **	0.05	-0.20
Total caregiving (hours)	6.58	0.47	7.52	0.18	-0.94 +	6.94	0.50	7.39	0.13	-0.45	0.49	3.56 **
% with 0 total caregiving hours	0.10	0.04	0.06	0.01	0.04	0.10	0.04	0.10	0.04	0.00	-0.04	0.03
Number of children <18	2.09	0.11	1.97	0.04	0.12	2.06	0.08	1.98	0.03	0.08	-0.03	0.59
One household child	0.36	0.05	0.37	0.02	-0.01	0.33	0.04	0.36	0.01	-0.04	-0.03	0.04
No spouse/partner present	0.36	0.05	0.26	0.02	0.09 +	0.30	0.05	0.28	0.01	0.03	-0.07	0.09
Maternal education												
<HS education	0.19	0.04	0.10	0.01	0.09 +	0.15	0.03	0.09	0.01	0.06 +	-0.03	0.30 +
HS grad	0.22	0.04	0.28	0.02	-0.06	0.23	0.04	0.27	0.01	-0.04	0.02	-0.01
Some college	0.22	0.05	0.21	0.02	0.01	0.20	0.03	0.18	0.01	0.02	0.02	-0.09
College grad	0.38	0.05	0.41	0.02	-0.03	0.43	0.04	0.47	0.01	-0.04	-0.01	-0.21
Hispanic ethnicity	0.41	0.05	0.16	0.01	0.25 **	0.44	0.04	0.15	0.01	0.29 **	0.03	0.25
Age	32.64	0.95	33.03	0.37	-0.39	34.48	0.83	32.83	0.24	1.65 +	2.04	1.52
N	142		1376			218		2016				

Notes: Includes employed adult women with youngest child under age 6. California's Paid Family Leave (PFL) program began July 1, 2004. Pre-PFL period extends through June 2004.

[1]Difference: mean outside California subtracted from the mean in California. [2]Difference-in-difference: mean difference pre-PFL subtracted from mean difference post-PFL This shows whether the difference between women in California and women in other states changed from the pre- to post-PFL period.. [3]Difference-in-difference-in-difference: mean difference-in-difference among mothers of older children subtracted from mean difference-in-difference among mothers of infants. This shows whether the change in the difference between women in California and women in other states before and after PFL was the same for mothers of older children as it was for mothers of infants.

+ significant at p<.10; * significant at p<.05; ** significant at p<.01

Table 3. DD and DDD estimates of time in hours spent on work per day among employed adult women with youngest child under age 6 (N = 4,586). ATUS, 2003-2008.

	(1)	(2)	(3)	(4)	(5)	(6)
	DD		DDD	DD with controls		DDD with controls
	Mothers of infants	Mothers of children 1-6	All mothers	Mothers of infants	Mothers of children 1-6	All mothers
	Coeff. [S.E.]	Coeff. [S.E.]	Coeff. [S.E.]	Coeff. [S.E.]	Coeff. [S.E.]	Coeff. [S.E.]
Post-PFL (ref: pre-PFL)	0.75+	0.46*	0.46*	0.69+	0.45*	0.46*
	[0.40]	[0.19]	[0.19]	[0.39]	[0.18]	[0.18]
CA (ref: other states)	1.38+	0.05	0.05	1.34	-0.17	-0.19
	[0.81]	[0.44]	[0.44]	[0.86]	[0.43]	[0.44]
Post-PFL x CA	-1.90+	-0.12	-0.12	-2.65*	-0.16	-0.15
	[1.11]	[0.57]	[0.57]	[1.09]	[0.55]	[0.55]
Mother of infant			-0.83*			-0.72*
			[0.35]			[0.35]
Post-PFL x mother of infant			0.28			0.23
			[0.44]			[0.43]
CA x mother of infant			1.33			1.38
			[0.92]			[0.95]
Post-PFL x CA x mother of infant			-1.77			-2.09+
			[1.24]			[1.24]
Number of household children < 18				-0.20	-0.16+	-0.15+
				[0.17]	[0.09]	[0.08]
No spouse/partner present				0.15	0.39+	0.36+
				[0.46]	[0.21]	[0.19]
Maternal education (ref: college grad)						
<HS education				2.17**	-0.09	0.45
				[0.65]	[0.36]	[0.31]
HS grad				0.71	-0.17	0.01
				[0.48]	[0.22]	[0.20]
Some college				0.36	-0.39	-0.24
				[0.47]	[0.24]	[0.22]
Hispanic ethnicity				0.77	0.80**	0.76**
				[0.53]	[0.25]	[0.22]
Age				0.02	0.02	0.02+
				[0.02]	[0.01]	[0.01]
Constant	3.34**	4.17**	4.17**	2.55**	3.78**	3.62**
	[0.32]	[0.15]	[0.15]	[0.88]	[0.46]	[0.42]
Observations	834	3752	4586	834	3752	4586
R-squared	0.01	0	0.01	0.05	0.01	0.02

Notes: Coefficients and robust standard errors in brackets. All models include sampling weights to account for oversampling of certain groups, as well as differential response rates by demographic characteristics and days of the week. Work was defined as the total number of hours spent on work and work-related activities, including working at a job, work-related activities, engaging in other income-generating activities, and searching and interviewing for jobs. DD models among mothers of children aged 1 to 6 (columns 2 and 5) are falsification tests since PFL should not have affected time use in this group.

+ significant at 10%; * significant at 5%; ** significant at 1%
Abbreviations: DD: difference-in-difference; DDD: difference-in-difference-in-difference; S.E.: standard error; PFL: California's Paid Family Leave program which was implemented in July 2004; CA: California.

Table 4. DD and DDD estimates of time in hours spent on primary childcare per day among employed adult women with youngest child under age 6 (N = 4,586). ATUS, 2003-2008.

	(1)	(2)	(3)	(4)	(5)	(6)
	DD		DDD	DD with controls		DDD with controls
	Mothers of infants	Mothers of children 1-6	All mothers	Mothers of infants	Mothers of children 1-6	All mothers
	Coeff. [S.E.]	Coeff. [S.E.]	Coeff. [S.E.]	Coeff. [S.E.]	Coeff. [S.E.]	Coeff. [S.E.]
Post-PFL (ref: pre-PFL)	-0.57*	-0.12	-0.12	-0.59**	-0.16*	-0.16*
	[0.24]	[0.08]	[0.08]	[0.22]	[0.08]	[0.08]
CA (ref: other states)	-1.09*	-0.40**	-0.40**	-1.13*	-0.23	-0.21
	[0.48]	[0.15]	[0.15]	[0.52]	[0.14]	[0.14]
Post-PFL x CA	0.76	0.16	0.16	1.38*	0.18	0.17
	[0.80]	[0.23]	[0.23]	[0.63]	[0.21]	[0.21]
Mother of infant			1.53**			1.47**
			[0.21]			[0.20]
Post-PFL x mother of infant			-0.45+			-0.39
			[0.26]			[0.24]
CA x mother of infant			-0.69			-0.84+
			[0.50]			[0.49]
Post-PFL x CA x mother of infant			0.6			1.04
			[0.83]			[0.71]
Number of household children < 18				-0.07	0.11**	0.04
				[0.08]	[0.04]	[0.04]
No spouse/partner present				-0.34	-0.24**	-0.27**
				[0.26]	[0.08]	[0.08]
Maternal education (ref: college grad)						
<HS education				-1.67**	-1.04**	-1.19**
				[0.31]	[0.12]	[0.12]
HS grad				-1.31**	-0.46**	-0.64**
				[0.25]	[0.10]	[0.10]
Some college				-0.36	-0.53**	-0.49**
				[0.33]	[0.09]	[0.10]
Hispanic ethnicity				-0.50*	-0.40**	-0.42**
				[0.24]	[0.09]	[0.09]
Age				0	-0.02**	-0.01**
				[0.01]	[0.00]	[0.00]
Constant	3.37**	1.84**	1.84**	4.21**	2.62**	2.68**
	[0.20]	[0.06]	[0.06]	[0.50]	[0.17]	[0.18]
Observations	834	3752	4586	834	3752	4586
R-squared	0.02	0	0.07	0.14	0.06	0.13

Notes: Coefficients and robust standard errors in brackets. All models include sampling weights to account for oversampling of certain groups, as well as differential response rates by demographic characteristics and days of the week. Primary childcare was defined as any activity related to caring for household children, activities related to children's education, and activities related to children's health. DD models among mothers of children aged 1 to 6 (columns 2 and 5) are falsification tests since PFL should not have affected time use in this group.

+ significant at 10%; * significant at 5%; ** significant at 1%
Abbreviations: DD: difference-in-difference; DDD: difference-in-difference-in-difference; S.E.: standard error; PFL: California's Paid Family Leave program which was implemented in July 2004; CA: California.

Table 5. DD and DDD estimates of time in hours spent on secondary childcare per day among employed adult women with youngest child under age 6 (N = 4,586). ATUS, 2003-2008.

	(1)	(2)	(3)	(4)	(5)	(6)
	DD		DDD	DD with controls		DDD with controls
	Mothers of infants	Mothers of children 1-6	All mothers	Mothers of infants	Mothers of children 1-6	All mothers
	Coeff.	Coeff.	Coeff.	Coeff.	Coeff.	Coeff.
	[S.E.]	[S.E.]	[S.E.]	[S.E.]	[S.E.]	[S.E.]
Post-PFL (ref: pre-PFL)	-0.64	-0.02	-0.02	-0.37	-0.01	-0.02
	[0.43]	[0.19]	[0.19]	[0.43]	[0.19]	[0.19]
CA (ref: other states)	-2.62**	-0.54	-0.54	-2.37**	-0.35	-0.33
	[0.84]	[0.43]	[0.43]	[0.83]	[0.42]	[0.42]
Post-PFL x CA	3.29**	0.33	0.33	3.80**	0.3	0.31
	[1.10]	[0.63]	[0.63]	[1.08]	[0.61]	[0.61]
Mother of infant			0.58			0.17
			[0.36]			[0.34]
Post-PFL x mother of infant			-0.62			-0.37
			[0.47]			[0.45]
CA x mother of infant			-2.07*			-1.97*
			[0.94]			[0.87]
Post-PFL x CA x mother of infant			2.96*			3.14**
			[1.27]			[1.20]
Number of household children < 18				0.51*	0.44**	0.44**
				[0.21]	[0.09]	[0.09]
No spouse/partner present				-1.94**	-1.47**	-1.55**
				[0.54]	[0.23]	[0.21]
Maternal education (ref: college grad)						
<HS education				-1.41*	-0.3	-0.56+
				[0.62]	[0.39]	[0.33]
HS grad				-0.47	0.3	0.14
				[0.58]	[0.23]	[0.22]
Some college				-0.23	-0.13	-0.16
				[0.48]	[0.26]	[0.23]
Hispanic ethnicity				-0.38	-0.29	-0.3
				[0.56]	[0.27]	[0.24]
Age				-0.08**	-0.03*	-0.04**
				[0.02]	[0.01]	[0.01]
Constant	6.26**	5.68**	5.68**	8.10**	6.18**	6.58**
	[0.32]	[0.15]	[0.15]	[0.86]	[0.48]	[0.43]
Observations	834	3752	4586	834	3752	4586
R-squared	0.02	0	0	0.09	0.04	0.04

Notes: Coefficients and robust standard errors in brackets. All models include sampling weights to account for oversampling of certain groups, as well as differential response rates by demographic characteristics and days of the week. Secondary childcare is care given while engaging in other activities (e.g., respondent indicates that children were under their supervision while primary activity was preparing dinner). Secondary childcare is not recorded when the primary activity is childcare. DD models among mothers of children aged 1 to 6 (columns 2 and 5) are falsification tests since PFL should not have affected time use in this group.

+ significant at 10%; * significant at 5%; ** significant at 1%
Abbreviations: DD: difference-in-difference; DDD: difference-in-difference-in-difference; S.E.: standard error; PFL: California's Paid Family Leave program which was implemented in July 2004; CA: California.

Table 6. DD and DDD estimates of time in hours spent on total childcare per day among employed adult women with youngest child under age 6 (N = 4,586). ATUS, 2003-2008.

	(1)	(2)	(3)	(4)	(5)	(6)
	DD		DDD	DD with controls		DDD with controls
	Mothers of infants	Mothers of children 1-6	All mothers	Mothers of infants	Mothers of children 1-6	All mothers
	Coeff. [S.E.]	Coeff. [S.E.]	Coeff. [S.E.]	Coeff. [S.E.]	Coeff. [S.E.]	Coeff. [S.E.]
Post-PFL (ref: pre-PFL)	-1.21*	-0.14	-0.14	-0.96+	-0.17	-0.18
	[0.51]	[0.22]	[0.22]	[0.49]	[0.21]	[0.21]
CA (ref: other states)	-3.71**	-0.94+	-0.94+	-3.50**	-0.58	-0.53
	[1.14]	[0.50]	[0.50]	[1.15]	[0.47]	[0.47]
Post-PFL x CA	4.05**	0.49	0.49	5.19**	0.48	0.48
	[1.53]	[0.72]	[0.72]	[1.40]	[0.68]	[0.67]
Mother of infant			2.11**			1.64**
			[0.43]			[0.42]
Post-PFL x mother of infant			-1.07+			-0.77
			[0.56]			[0.53]
CA x mother of infant			-2.77*			-2.81*
			[1.24]			[1.17]
Post-PFL x CA x mother of infant			3.56*			4.18**
			[1.69]			[1.56]
Number of household children < 18				0.44*	0.55**	0.48**
				[0.22]	[0.10]	[0.10]
No spouse/partner present				-2.28**	-1.71**	-1.82**
				[0.61]	[0.25]	[0.23]
Maternal education (ref: college grad)						
<HS education				-3.08**	-1.33**	-1.76**
				[0.77]	[0.42]	[0.37]
HS grad				-1.79**	-0.16	-0.50*
				[0.63]	[0.25]	[0.24]
Some college				-0.59	-0.66*	-0.65*
				[0.58]	[0.28]	[0.26]
Hispanic ethnicity				-0.87	-0.69*	-0.72**
				[0.67]	[0.29]	[0.27]
Age				-0.07*	-0.04**	-0.05**
				[0.03]	[0.01]	[0.01]
Constant	9.63**	7.52**	7.52**	12.30**	8.80**	9.26**
	[0.40]	[0.18]	[0.18]	[1.10]	[0.54]	[0.50]
Observations	834	3752	4586	834	3752	4586
R-squared	0.02	0	0.02	0.14	0.06	0.08

Notes: Coefficients and robust standard errors in brackets. All models include sampling weights to account for oversampling of certain groups, as well as differential response rates by demographic characteristics and days of the week. Total childcare is the sum of primary and secondary childcare. Primary childcare was defined as any activity related to caring for household children, activities related to children's education, and activities related to children's health. Secondary childcare is care given while engaging in other activities (e.g., respondent indicates that children were under their supervision while primary activity was preparing dinner). Secondary childcare is not recorded when the primary activity is childcare. DD models among mothers of children aged 1 to 6 (columns 2 and 5) are falsification tests since PFL should not have affected time use in this group.

+ significant at 10%; * significant at 5%; ** significant at 1%
Abbreviations: DD: difference-in-difference; DDD: difference-in-difference-in-difference; S.E.: standard error; PFL: California's Paid Family Leave program which was implemented in July 2004; CA: California.

Table 7. DD and DDD estimates of time in hours spent on work per day among employed adult men with youngest child under age 6 (N = 4,844). ATUS, 2003-2008.

	(1)	(2)	(3)	(4)	(5)	(6)
	DD		DDD	DD with controls		DDD with controls
	Fathers of infants	Fathers of children 1-6	All fathers	Fathers of infants	Fathers of children 1-6	All fathers
	Coeff. [S.E.]	Coeff. [S.E.]	Coeff. [S.E.]	Coeff. [S.E.]	Coeff. [S.E.]	Coeff. [S.E.]
Post-PFL (ref: pre-PFL)	0.34	0.04	0.04	0.38	0.06	0.06
	[0.39]	[0.19]	[0.19]	[0.40]	[0.18]	[0.18]
CA (ref: other states)	0.23	0	0	0.45	0.02	0.03
	[1.22]	[0.42]	[0.42]	[1.17]	[0.42]	[0.42]
Post-PFL x CA	-0.97	-0.48	-0.48	-1.19	-0.44	-0.43
	[1.37]	[0.61]	[0.61]	[1.36]	[0.60]	[0.60]
Father of infant			-0.43			-0.51
			[0.33]			[0.33]
Post-PFL x father of infant			0.3			0.33
			[0.43]			[0.43]
CA x father of infant			0.23			0.42
			[1.29]			[1.26]
Post-PFL x CA x father of infant			-0.48			-0.73
			[1.49]			[1.47]
Number of household children < 18				-1.01	-0.75*	-0.82*
				[0.88]	[0.38]	[0.35]
No spouse/partner present				-0.1	-0.07	-0.08
				[0.16]	[0.11]	[0.09]
Paternal education (ref: college grad)						
<HS education				0.66	-0.22	-0.03
				[0.78]	[0.33]	[0.31]
HS grad				0.17	0.40+	0.35+
				[0.46]	[0.22]	[0.20]
Some college				0.18	-0.43	-0.32
				[0.56]	[0.27]	[0.24]
Hispanic ethnicity				-0.36	0.11	0
				[0.55]	[0.27]	[0.24]
Age				-0.01	-0.01	-0.01
				[0.02]	[0.01]	[0.01]
Constant	5.63**	6.05**	6.05**	6.18**	6.51**	6.55**
	[0.30]	[0.14]	[0.14]	[0.84]	[0.47]	[0.42]
Observations	960	3884	4844	960	3884	4844
R-squared	0	0	0	0.01	0.01	0.01

Notes: Coefficients and robust standard errors in brackets. All models include sampling weights to account for oversampling of certain groups, as well as differential response rates by demographic characteristics and days of the week. Work was defined as the total number of hours spent on work and work-related activities, including working at a job, work-related activities, engaging in other income-generating activities, and searching and interviewing for jobs.

+ significant at 10%; * significant at 5%; ** significant at 1%
Abbreviations: DD: difference-in-difference; DDD: difference-in-difference-in-difference; S.E.: standard error; PFL: California's Paid Family Leave program which was implemented in July 2004; CA: California.

Table 8. DD and DDD estimates of time in hours spent on primary childcare per day among employed adult men with youngest child under age 6 (N = 4,844). ATUS, 2003-2008.

	(1)	(2)	(3)	(4)	(5)	(6)
	DD		DDD	DD with controls		DDD with controls
	Fathers of infants	Fathers of children 1-6	All fathers	Fathers of infants	Fathers of children 1-6	All fathers
	Coeff. [S.E.]	Coeff. [S.E.]	Coeff. [S.E.]	Coeff. [S.E.]	Coeff. [S.E.]	Coeff. [S.E.]
Post-PFL (ref: pre-PFL)	0.13	0.03	0.03	0.12	0.04	0.04
	[0.15]	[0.06]	[0.06]	[0.14]	[0.05]	[0.05]
CA (ref: other states)	0.39	0.02	0.02	0.37	0.18	0.18
	[0.44]	[0.18]	[0.18]	[0.42]	[0.17]	[0.17]
Post-PFL x CA	-0.90+	0.04	0.04	-0.61	0.12	0.11
	[0.47]	[0.32]	[0.32]	[0.45]	[0.32]	[0.32]
Father of infant			0.25+			0.23+
			[0.13]			[0.13]
Post-PFL x father of infant			0.1			0.09
			[0.16]			[0.15]
CA x father of infant			0.37			0.29
			[0.47]			[0.45]
Post-PFL x CA x father of infant			-0.94+			-0.87
			[0.57]			[0.56]
Number of household children < 18				-0.43*	-0.55**	-0.52**
				[0.20]	[0.09]	[0.08]
No spouse/partner present				0.06	0.09	0.08
				[0.05]	[0.08]	[0.06]
Paternal education (ref: college grad)						
<HS education				-0.86**	-0.50**	-0.58**
				[0.20]	[0.12]	[0.10]
HS grad				-0.58**	-0.35**	-0.40**
				[0.16]	[0.07]	[0.06]
Some college				-0.15	-0.24**	-0.23**
				[0.23]	[0.09]	[0.08]
Hispanic ethnicity				-0.34*	-0.44**	-0.41**
				[0.17]	[0.08]	[0.07]
Age				0	-0.01	-0.01
				[0.01]	[0.01]	[0.00]
Constant	1.26**	1.01**	1.01**	1.40**	1.46**	1.38**
	[0.12]	[0.04]	[0.04]	[0.40]	[0.14]	[0.15]
Observations	960	3884	4844	960	3884	4844
R-squared	0.01	0	0.01	0.08	0.06	0.06

Notes: Coefficients and robust standard errors in brackets. All models include sampling weights to account for oversampling of certain groups, as well as differential response rates by demographic characteristics and days of the week. Primary childcare was defined as any activity related to caring for household children, activities related to children's education, and activities related to children's health.

+ significant at 10%; * significant at 5%; ** significant at 1%

Abbreviations: DD: difference-in-difference; DDD: difference-in-difference-in-difference; S.E.: standard error; PFL: California's Paid Family Leave program which was implemented in July 2004; CA: California.

Table 9. DD and DDD estimates of time in hours spent on secondary childcare per day among employed adult men with youngest child under age 6 (N = 4,844). ATUS, 2003-2008.

	(1)	(2)	(3)	(4)	(5)	(6)
	DD		DDD	DD with controls		DDD with controls
	Fathers of infants	Fathers of children 1-6	All fathers	Fathers of infants	Fathers of children 1-6	All fathers
	Coeff.	Coeff.	Coeff.	Coeff.	Coeff.	Coeff.
	[S.E.]	[S.E.]	[S.E.]	[S.E.]	[S.E.]	[S.E.]
Post-PFL (ref: pre-PFL)	-0.04	-0.04	-0.04	0.05	0.01	0.01
	[0.36]	[0.17]	[0.17]	[0.35]	[0.16]	[0.16]
CA (ref: other states)	-0.78	-0.12	-0.12	-0.61	0	-0.02
	[0.89]	[0.38]	[0.38]	[0.78]	[0.38]	[0.38]
Post-PFL x CA	0.46	0.14	0.14	0.48	0.28	0.3
	[1.08]	[0.52]	[0.52]	[0.98]	[0.51]	[0.50]
Father of infant			0.29			0.19
			[0.30]			[0.30]
Post-PFL x father of infant			0			0.01
			[0.39]			[0.39]
CA x father of infant			-0.66			-0.56
			[0.97]			[0.89]
Post-PFL x CA x father of infant			0.32			0.13
			[1.20]			[1.12]
Number of household children < 18				-2.29**	-1.77**	-1.84**
				[0.72]	[0.46]	[0.41]
No spouse/partner present				0.21	0.18*	0.19*
				[0.15]	[0.08]	[0.07]
Paternal education (ref: college grad)						
<HS education				-1.65*	-0.3	-0.57+
				[0.70]	[0.37]	[0.33]
HS grad				-0.48	-0.47**	-0.47**
				[0.41]	[0.18]	[0.17]
Some college				-0.26	-0.08	-0.13
				[0.47]	[0.21]	[0.19]
Hispanic ethnicity				0.18	-0.56*	-0.41
				[0.55]	[0.27]	[0.25]
Age				-0.04*	-0.02*	-0.03**
				[0.02]	[0.01]	[0.01]
Constant	4.37**	4.09**	4.09**	5.81**	4.93**	5.10**
	[0.28]	[0.12]	[0.12]	[0.70]	[0.39]	[0.35]
Observations	960	3884	4844	960	3884	4844
R-squared	0	0	0	0.04	0.02	0.03

Notes: Coefficients and robust standard errors in brackets. All models include sampling weights to account for oversampling of certain groups, as well as differential response rates by demographic characteristics and days of the week. Secondary childcare is care given while engaging in other activities (e.g., respondent indicates that children were under their supervision while primary activity was preparing dinner). Secondary childcare is not recorded when the primary activity is childcare.

+ significant at 10%; * significant at 5%; ** significant at 1%
Abbreviations: DD: difference-in-difference; DDD: difference-in-difference-in-difference; S.E.: standard error; PFL: California's Paid Family Leave program which was implemented in July 2004; CA: California.

Table 10. DD and DDD estimates of time in hours spent on total childcare per day among employed adult men with youngest child under age 6 (N = 4,844). ATUS, 2003-2008.

	(1)	(2)	(3)	(4)	(5)	(6)
	DD		DDD	DD with controls		DDD with controls
	Fathers of infants	Fathers of children 1-6	All fathers	Fathers of infants	Fathers of children 1-6	All fathers
	Coeff.	Coeff.	Coeff.	Coeff.	Coeff.	Coeff.
	[S.E.]	[S.E.]	[S.E.]	[S.E.]	[S.E.]	[S.E.]
Post-PFL (ref: pre-PFL)	0.09	-0.01	-0.01	0.17	0.05	0.05
	[0.40]	[0.18]	[0.18]	[0.39]	[0.17]	[0.17]
CA (ref: other states)	-0.39	-0.11	-0.11	-0.25	0.17	0.16
	[1.14]	[0.45]	[0.45]	[0.99]	[0.44]	[0.44]
Post-PFL x CA	-0.44	0.18	0.18	-0.13	0.4	0.41
	[1.31]	[0.72]	[0.72]	[1.17]	[0.68]	[0.68]
Father of infant			0.54			0.43
			[0.34]			[0.33]
Post-PFL x father of infant			0.1			0.1
			[0.44]			[0.42]
CA x father of infant			-0.29			-0.27
			[1.22]			[1.09]
Post-PFL x CA x father of infant			-0.62			-0.74
			[1.49]			[1.38]
Number of household children < 18				-2.72**	-2.32**	-2.36**
				[0.79]	[0.48]	[0.43]
No spouse/partner present				0.28+	0.27+	0.27*
				[0.17]	[0.14]	[0.11]
Paternal education (ref: college grad)						
<HS education				-2.51**	-0.80*	-1.15**
				[0.78]	[0.41]	[0.36]
HS grad				-1.06*	-0.82**	-0.87**
				[0.45]	[0.20]	[0.18]
Some college				-0.41	-0.32	-0.35
				[0.57]	[0.24]	[0.22]
Hispanic ethnicity				-0.16	-0.99**	-0.82**
				[0.62]	[0.28]	[0.27]
Age				-0.04+	-0.03*	-0.03**
				[0.02]	[0.01]	[0.01]
Constant	5.64**	5.10**	5.10**	7.22**	6.38**	6.48**
	[0.31]	[0.13]	[0.13]	[0.84]	[0.43]	[0.40]
Observations	960	3884	4844	960	3884	4844
R-squared	0	0	0	0.06	0.05	0.05

Notes: Coefficients and robust standard errors in brackets. All models include sampling weights to account for oversampling of certain groups, as well as differential response rates by demographic characteristics and days of the week. Total childcare is the sum of primary and secondary childcare. Primary childcare was defined as any activity related to caring for household children, activities related to children's education, and activities related to children's health. Secondary childcare is care given while engaging in other activities (e.g., respondent indicates that children were under their supervision while primary activity was preparing dinner). Secondary childcare is not recorded when the primary activity is childcare.

+ significant at 10%; * significant at 5%; ** significant at 1%

Abbreviations: DD: difference-in-difference; DDD: difference-in-difference-in-difference; S.E.: standard error; PFL: California's Paid Family Leave program which was implemented in July 2004; CA: California.

Appendix

Table A1. Sensitivity tests clustering standard errors at the household level. DD and DDD estimates of time in hours spent on work and total childcare per day among employed adult women with youngest child under age 6 (N = 4,586). ATUS, 2003-2008.

	(1)	(2)	(3)	(4)	(5)	(6)
	Hours spent on work			Hours spend on total childcare		
	DD with controls		DDD with controls	DD with controls		DDD with controls
	Mothers of infants	Mothers of children 1-6	All mothers	Mothers of infants	Mothers of children 1-6	All mothers
	Coeff. [S.E.]	Coeff. [S.E.]	Coeff. [S.E.]	Coeff. [S.E.]	Coeff. [S.E.]	Coeff. [S.E.]
Post-PFL (ref: pre-PFL)	0.69+	0.45*	0.46*	-0.96+	-0.17	-0.18
	[0.39]	[0.18]	[0.18]	[0.49]	[0.21]	[0.21]
CA (ref: other states)	1.34	-0.17	-0.19	-3.50**	-0.58	-0.53
	[0.86]	[0.44]	[0.44]	[1.15]	[0.47]	[0.47]
Post-PFL x CA	-2.65*	-0.16	-0.15	5.19**	0.48	0.48
	[1.09]	[0.56]	[0.56]	[1.40]	[0.67]	[0.67]
Mother of infant			-0.72*			1.64**
			[0.35]			[0.41]
Post-PFL x mother of infant			0.23			-0.77
			[0.43]			[0.53]
CA x mother of infant			1.38			-2.81*
			[0.95]			[1.17]
Post-PFL x CA x mother of infant			-2.09+			4.18**
			[1.24]			[1.56]
Number of household children < 18	-0.2	-0.16+	-0.15+	0.44*	0.55**	0.48**
	[0.17]	[0.09]	[0.08]	[0.22]	[0.10]	[0.10]
No spouse/partner present	0.15	0.39+	0.36+	-2.28**	-1.71**	-1.82**
	[0.46]	[0.21]	[0.19]	[0.61]	[0.25]	[0.23]
Maternal education (ref: college grad)						
<HS education	2.17**	-0.09	0.45	-3.08**	-1.33**	-1.76**
	[0.65]	[0.35]	[0.31]	[0.77]	[0.42]	[0.37]
HS grad	0.71	-0.17	0.01	-1.79**	-0.16	-0.50*
	[0.48]	[0.22]	[0.20]	[0.63]	[0.25]	[0.24]
Some college	0.36	-0.39	-0.24	-0.59	-0.66*	-0.65*
	[0.47]	[0.24]	[0.22]	[0.58]	[0.28]	[0.26]
Hispanic ethnicity	0.77	0.80**	0.76**	-0.87	-0.69*	-0.72**
	[0.53]	[0.25]	[0.22]	[0.67]	[0.29]	[0.27]
Age	0.02	0.02	0.02+	-0.07*	-0.04**	-0.05**
	[0.02]	[0.01]	[0.01]	[0.03]	[0.01]	[0.01]
Constant	2.55**	3.78**	3.62**	12.30**	8.80**	9.26**
	[0.88]	[0.46]	[0.41]	[1.10]	[0.54]	[0.50]
Observations	834	3752	4586	834	3752	4586
R-squared	0.05	0.01	0.02	0.14	0.06	0.08

Notes: Coefficients and clustered standard errors in brackets. Standard errors are clustered on respondents'' household. All models include sampling weights to account for oversampling of certain groups, as well as differential response rates by demographic characteristics and days of the week. Work was defined as the total number of hours spent on work and work-related activities, including working at a job, work-related activities, engaging in other income-generating activities, and searching and interviewing for jobs. Total childcare is the sum of primary and secondary childcare. Primary childcare was defined as any activity related to caring for household children, activities related to children's education, and activities related to children's health. Secondary childcare is care given while engaging in other activities (e.g., respondent indicates that children were under their supervision while primary activity was preparing dinner). Secondary childcare is not recorded when the primary activity is childcare. DD models among mothers of children aged 1 to 6 (columns 2 and 5) are falsification tests since PFL should not have affected time use in this group.

+ significant at 10%; * significant at 5%; ** significant at 1%

Abbreviations: DD: difference-in-difference; DDD: difference-in-difference-in-difference; S.E.: standard error; PFL: California's Paid Family Leave program which was implemented in July 2004; CA: California.

Chapter 5
Conclusion

This dissertation examined maternity leave policies, utilization and consequences in three papers, each providing a separate, but related, conclusion. The first paper provides, for the first time, a picture of antenatal maternity leave prevalence and correlates in the United States. I concluded that state-level policies that provide income replacement increase the likelihood that women will take antenatal leave, but other factors, like characteristics of women's work and prenatal health conditions, are also important. The second paper, in which I examine the relationship between antenatal leave taken and negative delivery outcomes, highlights the challenges with using observational data to analyze policy impacts. Despite a rich dataset with detailed pre- and post-natal health information, I concluded that my finding that antenatal leave was associated with an increase in negative delivery outcomes resulted from ongoing selection, rather than a causal effect. This strong selection is informative in its own way: I suggest that in a context of limited leave availability (like the United States as a whole, and Pennsylvania in particular), only relatively unhealthy women take antenatal leave. This presents challenges for studying leave, but also informs targeting and promotion of leave policies. Finally, the third paper uses a natural experiment to examine the impact of the nation's first paid family leave program. I conclude that women exposed to the law changed the way they spent their time—increasing the amount of time with children in their care and marginally decreasing their time spent in paid work.

The overarching theme of this dissertation is that leave policies appear to have a small but significant impact on leave utilization, but measuring the impact of this leave utilization proves challenging. In particular, women under the same policy constraints differ in their need for and preferences toward leave, leading to a non-random sample of leave-takers. Future work that better addresses selection concerns through a quasi-experimental design will move the field forward.