Child Health in Elementary School following California's Paid Family Leave Program

Shirlee Lichtman-Sadot*　　Neryvia Pillay Bell†

May 2017

Abstract

We evaluate changes in elementary school children health outcomes following the introduction of California's Paid Family Leave (PFL) program, which provided parents with paid time off following the birth of a child. Our health outcomes - hearing-related problems, overweight and ADHD - are characterized by diagnosis rates which only pick up during early elementary school. Moreover, our health outcomes have been found to be negatively linked with many potential implications of extended maternity leave - increased breastfeeding, prompt medical checkups at infancy, reduced prenatal stress, and reduced non-parental care during infancy. Using the Early Childhood Longitudinal Studies (ECLS) within a difference-in-differences framework, our results suggest improvements in health outcomes among California elementary school children following PFL's introduction. Furthermore, the improvements are driven by children from less advantaged backgrounds, which is consistent with the notion that California's PFL had the greatest effect on leave-taking duration after childbirth mostly for less advantaged mothers who previously could not afford to take unpaid leave.

JEL Classifications: I14, I18, J88. **Keywords:** Maternity leave; Child health; Health Disparities

We are grateful to discussions and comments from Caroline Hoxby, Oren Rigbi, Ity Shurtz, Sally Clark Stearns, Danny Cohen-Zada, seminar participants at Ben-Gurion University's Brown Bag Seminar, and attendees at the 6th Bi-Annual American Society for Health Economists Conference.

*Dept. of Economics, Ben-Gurion University of the Negev, Beer-Sheva 84105, Israel. E-mail shirlees@bgu.ac.il.

†Dept. of Economics, Stanford University, Stanford 94305, California. E-mail neryvia@stanford.edu.

1 Introduction

In July 2004, California became the first state in the U.S. to offer paid family leave. California's Paid Family Leave (PFL) program provides employees up to 6 weeks of paid time off in order to care for their newborn, newly-adopted child, or a sick family member. Prior to PFL, 6-8 weeks of paid leave were available to California mothers immediately following the birth of a child through the state's disability insurance program. Thus, California's PFL increased paid leave eligibility for new mothers by 75-100%.

Recent studies show that the introduction of PFL increased California mothers' leave-taking duration after childbirth (Rossin-Slater et al. (2013); Baum and Ruhm (2016)). We assess whether California's PFL had an effect on children's health outcomes in early elementary school - mostly kindergarten. We examine six health outcomes: overweight, attention deficit/hyperactivity disorder (ADHD), a quantitative measure of a child's general health condition, and three hearing-related outcomes (hearing problems, communication problems, and a history of frequent ear infections).[1]

We base our hypothesis of extended parental leave having an effect on childhood health outcomes - in particular the health outcomes examined in this study - on ample existing literature exhibiting that overweight, ADHD and ear infections are negatively linked with: increased breastfeeding duration and initiation, increased parental care at infancy, and reduced prenatal stress, all of which are very plausible implications of increased parental leave following the birth of a child. We note specifically that Huang and Yang (2015) and Appelbaum and Milkman (2011) provide evidence that California's PFL increased breastfeeding duration and initiation among mothers. Furthermore, mothers' knowledge of the expected financial security while taking time off following the birth of a child may reduce maternal stress during the prenatal period. There is even evidence that PFL increased time off from work for California mothers *prior* to giving birth by more than 1 week (Baum and Ruhm (2016)), which could also contribute to reducing maternal stress during the prenatal period. By focusing on early elementary school outcomes we are able to more accurately evaluate the prevalence of these adverse health conditions since for some of the health outcomes we evaluate - particularly overweight, ADHD and communication problems - diagnosis generally increases substantially in early elementary school.[2]

We utilize data from the Early Childhood Longitudinal Studies (ECLS) for cohorts of children born in 1992-1993, 2001, and 2004-2005, thus providing us with a large sample of children born before and after the introduction of California's PFL program from nearly all U.S. states. This allows us to use California's PFL policy reform as a natural experiment and perform our analysis within a difference-in-differences (DID) framework in order to estimate changes in health outcomes for children born before and after the introduction of PFL in California, while controlling for trends in health outcomes for children born in other states

[1] Although hearing-related problems are some of the main causes behind communication problems, other causes are also prevalent, including cognitive disability, autism, and stuttering.

[2] Obesity rates among U.S. children in 2011-2012 were 8.4, 17.7 and 20.5 percent for children 2-5, 6-11, and 12-19 years old, respectively (http://www.cdc.gov/NCH's/data/hestat/obesity_child_11_12/obesity_child_11_12.pdf). Thus, obesity rates more than double during early elementary school, but increase only slightly thereafter. The average age of ADHD diagnosis is 7 years old (http://www.cdc.gov/ncbddd/adhd/data.html).

during the same period. Our results suggest an improvement in all health outcomes observed, with the exception of communication problems, for which the results are not robust to some checks or alternative specifications. For frequent ear infections, data limitations do not allow us to confirm that the results are robust to the existence of pre-trends in improvements in California or to other specifications.

The improvements we observe in health outcomes following California's PFL are primarily driven by children from less advantaged backgrounds - i.e. with a lower socioeconomic ranking or with mothers having lower educational attainment. This result is consistent with the notion that California's PFL was most substantial for less advantaged mothers, who could not afford to take unpaid time off to care for a newborn child prior to PFL's introduction.[3]

To the best of our knowledge, our study is the first to inquire about the relationship between longer maternity leave mandates and child health outcomes during early elementary school, with an emphasis on health outcomes that are more established in terms of diagnosis rates at this age - overweight, ADHD, hearing problems, communication problems, and a history of frequent ear infections. As advocates of parental leave policies continuously cite potential health benefits for mother and child as arguments in favor of these policies, it is extremely important to provide evidence on the actual health effects of these policies. Despite this need for research on child health following parental leave reforms, the existing research on this subject is quite limited and mostly focuses on breastfeeding initiation and duration (Berger et al. (2005); Baker and Milligan (2008); Huang and Yang (2015)), infant mortality (Ruhm (2000); Rossin (2011)), immunization (Berger et al. (2005); Tanaka (2005)), and birth outcomes (Rossin (2011); Tanaka (2005); Ruhm (2000)), rather than longer-term child health outcomes, which are crucial for evaluating the overall advantages gained from parental leave policies. These longer-term outcomes are particularly valuable to track, as their prevalence frequently lasts through adulthood and can affect many aspects of adult well-being, such as educational attainment, earnings or even longevity. Thus, while these health outcomes during childhood can have contemporaneous economic consequences - such as health care costs or utilization - they likely also have a long-lasting effect on children and society as a whole.

Although data limitations of our study do not enable us to determine the exact channels through which longer maternity leave mandates may affect child health outcomes in elementary school, we are able to rely on existing studies and propose very plausible potential mechanisms - namely breastfeeding, greater parental care during infancy, and reduced prenatal stress, as discussed above.

We also use our data to examine patterns in maternal employment at kindergarten following PFL's introduction and find suggestive evidence that decreases in overweight may be partially driven by reduced labor force participation among higher socioeconomically-ranked mothers, also in accordance with existing literature on this subject (Anderson et al. (2003); Ruhm (2008)).

The paper proceeds as follows. Section 2 discusses related literature in two aspects: First, we provide an overview of the existing literature concerning the effects of parental leave policies on various child out-

[3] See Rossin-Slater et al. (2013) for empirical evidence on differences in increased leave-taking duration following California's PFL based on mothers' socioeconomic status.

comes, highlighting this study's contribution; Second, we elaborate on the literature related to potential mechanisms that may drive a link between extended parental leave following the birth of a child and the child health outcomes examined. Section 3 provides background on California's Paid Family Leave, its introduction in July 2004, and the maternity leave benefits available to California mothers prior to its introduction. Section 4 discusses the ECLS data used for the analysis. In Section 5, we describe our empirical strategy and underlying identification assumptions. Results are presented in Section 6, followed by robustness checks in Section 7. In Section 8 we discuss the magnitude of our estimated effect in the context of magnitudes found in the existing epidemiological literature. We also present suggestive evidence from our data on changes in maternal employment at kindergarten following PFL's introduction which could explain some of the changes we observe in child overweight following California's PFL. Finally, Section 9 concludes.

2 Related Literature

2.1 Child Outcomes following Parental Leave Reforms

Our study is an important extension of the existing literature on the effects of parental leave policies, which facilitate parents' ability to bond with their children after birth. The majority of the existing literature focuses on the effects of extended maternity leave policies on cognitive outcomes during childhood and adolescence, as well as adult outcomes such as wages, graduation or unemployment. Many of these studies fail to find an effect of longer maternity leave mandates on these outcomes (Baker and Milligan (2010); Dustmann and Schonberg (2012); Rasmussen (2010)). Some studies find positive long-term cognitive effects only on children with highly-educated mothers (Liu and Nordstrom (2009); Danzer and Lavy (2017)).

Carneiro et al. (2015) is one of the few papers evaluating an *introduction* of paid maternity leave benefits, as opposed to an *extension*. The reform evaluated increased paid leave entitlements in Norway from 0 to 4 months. Moreover, the authors used data that allowed them to identify children of mothers who were eligible for the reform, based on their employment history. They find that the introduction of paid maternity leave benefits decreased dropout rates and increased earnings during adulthood. Baum (2003) is a related study that does not evaluate extended maternity leave, but does examine the effect of maternal labor force participation in the first months of a child's life. The findings suggest a decrease in cognitive measures at age 5 as a result of mothers returning to work before the child is 3 months old.

The mixed results in the above studies may be due to differences in initial levels of maternity leave durations when reforms to extend them took place - in particular, the marginal benefit of extending maternity leave duration may be declining, and as such, the effect of maternity leave reforms may vary depending on how many weeks were offered to begin with. Moreover, the discrepancies between the studies' results highlight the importance of identifying precisely the population of mothers actually treated - i.e. those who ended up taking longer leaves from their work following the birth of a child.

Our study does not have the ability to precisely identify children of mothers who were treated and extended their leave duration following the birth of a child due to California's PFL, and our estimates are intent-to-treat (ITT) estimates - i.e. we estimate changes in health outcomes among all California children in our sample following PFL's introduction, although some of these children's mothers did not necessarily increase their leave-taking duration after childbirth due to California's PFL. However, our setting does have the advantage of evaluating a change in paid leave duration from extremely low initial levels. Paid maternity leave duration eligibility in California before PFL's introduction was six to eight weeks long and the mean take-up was three weeks (Rossin-Slater et al. (2013)). Following PFL's introduction, the mean take-up increased to roughly eight weeks (Baum and Ruhm (2016); Rossin-Slater et al. (2013)). Thus, if the marginal benefit of greater leave duration is decreasing for health outcomes - as it seems to be for cognitive and other human capital accumulation measures - then our setting has the advantage of greater ability to detect this benefit, even if our treatment group is not precisely identified.

Studies on the effect of extended maternity leave benefits on children's health outcomes are fewer than those evaluating cognitive, human capital accumulation or labor market outcomes during the children's lives. Ruhm (2000) and Tanaka (2005) find that longer maternity leave mandates in European countries decreased the incidence of low birth weight, as well as infant and child mortality rates, and increased child immunization. Rossin (2011) evaluates the introduction of job-protected maternity leave in the U.S. with the introduction of the Family and Medical Leave Act (FMLA) in 1993, and finds an increase in birth weight and decreases in premature birth and infant and child mortality rates among children of higher-educated and married mothers, who were more able to take advantage of the unpaid leave option. Berger et al. (2005) do not evaluate maternity leave policies directly, but do compare health outcomes - such as breastfeeding, immunizations, and externalization of behavior problems - between children whose mothers returned to work within 12 weeks of giving birth and children of mothers with longer leave from work after birth. Their findings suggest longer breastfeeding durations and an improvement in the above-mentioned health outcomes among children whose mothers returned later to work.

Our work differs in that it examines health outcomes that can only be evaluated later in a child's life when diagnosis rates stabilize. Furthermore, our health outcomes can have a profound effect on important long-term adult outcomes, such as wages and educational attainment (see Currie et al. (2010) and Fletcher (2014) on ADHD and Cawley (2010) on childhood obesity). Thus, our work is related to a large body of research linking early childhood conditions to long-term adult outcomes and well-being. As this relationship has been more clearly established in recent years, more studies have emerged to investigate policies that can potentially improve early childhood conditions, including family policies. These policies have frequently been found to benefit primarily children from disadvantaged populations, who are often more vulnerable to adverse conditions than more advantaged children, and these disparities can have a long-standing effect on the gaps between advantaged and disadvantaged children. Currie and Rossin-Slater (2015) provide an overview of this literature and these policies.

2.2 Potential Mechanisms Linking Child Health Outcomes and Extended Parental Leave

Existing studies link improved child health outcomes to: increased breastfeeding duration and initiation, greater maternal presence in the house following the birth of a child, and lower maternal stress levels during pregnancy. All of these can be direct implications of policies providing longer paid leave following the birth of a child, and hence are potential mechanisms through which we might expect the introduction of PFL in California to result in a decrease in the adverse health outcomes we examine - overweight, ADHD, hearing problems, and ear infections - among elementary school children.

Huang and Yang (2015) and Appelbaum and Milkman (2011) provide evidence that California's PFL increased breastfeeding duration among new mothers. According to Appelbaum and Milkman (2011), PFL roughly doubled the median number of weeks new mothers breastfed from 5 to 9 for mothers holding low-quality jobs, and from 5 to 11 for mothers holding high quality jobs. According to Huang and Yang (2015), the probability of breastfeeding following California's PFL increased by 10-20 percentage points at various infant ages that are less than 1 year old. Baker and Milligan (2008) show that extended maternity leave increases breastfeeding duration among new mothers in Canada. Berger et al. (2005) also show that longer maternity leave durations among 1979 National Longitudinal Survey of Youth (NLSY) mothers are associated with longer breastfeeding durations.

The epidemiological literature provides ample evidence positively linking improvements in health outcomes with increased breastfeeding duration. Ear infections are found to be associated with decreased breastfeeding durations, which in turn affect the prevalence of hearing problems (Duncan et al. (1993); Duffy et al. (1997); Paradise et al. (1997)). Arenz et al. (2004) and Owen et al. (2005) review published epidemiological studies investigating the association between breastfeeding and obesity and conclude that breastfeeding is associated with a reduced risk of obesity among children. Furthermore, the Center for Disease Control and Prevention (CDC) lists reduced breastfeeding as one of the potential contributors to childhood overweight or obesity.[4] Julvez et al. (2007) and Mimouni-Bloch et al. (2013) are two separate studies of populations in Spain and Israel, respectively, showing that longer breastfeeding durations - even among siblings - are associated with a decrease in the likelihood of being diagnosed with ADHD during childhood. Thus, one potential channel through which our health outcomes - specifically overweight, ADHD and ear infections - may improve with more generous parental leave policies is increases in breastfeeding durations or initiations, as mothers postpone their return to work.

Rossin-Slater et al. (2013), Baum and Ruhm (2016), and Bartel et al. (2015) find that California's PFL increased mothers' and fathers' leave-taking duration following the birth of a child by 3-5 weeks and 2-3 days, respectively. Although there is no direct evidence linking increased parental presence in the household following birth - without the addition of increased breastfeeding duration - with improved childhood health outcomes specifically in terms of overweight, ADHD, ear infections or hearing problems, this is

[4]Source: http://www.cdc.gov/obesity/childhood/causes.html

still a potential channel. Berger et al. (2005) present evidence that longer maternity leave durations among NLSY mothers increased the probability of the child receiving timely medical check-ups and vaccinations. Prompt medical check-ups may also have an effect on some of our long-term health outcomes, in particular the development of acute ear infections, which can have an effect on childhood hearing problems, or proper awareness of suggested children's diets to prevent overweight. In a meta-analysis of epidemiological studies, Uhari et al. (1996) find that child care outside the home significantly increases the risk of frequent ear infections, due to greater contact with other children who may be infected. By spending more time at home with a newborn, parents are likely delaying child care at a day care center, which would provide some protection against frequent ear infections.

It is also very plausible that PFL's introduction reduced stress during pregnancy, as mothers knew in advance that they would be financially secure during their time off from work to be with their newborn.[5] Furthermore, Baum and Ruhm (2016) present evidence that PFL increased time off from work for California mothers *prior* to giving birth by more than 1 week, which could also contribute to reducing maternal stress during the prenatal period. Reduced stress-levels during the prenatal period could also play a role in improving child health outcomes later in life. Higher stress levels during pregnancy can adversely affect birthweight and the probability of premature birth (Copper et al. (1996); Carlson (2015); Persson and Rossin-Slater (2015)), which are strongly linked to ADHD and overweight among children.[6]

Income effects may also play a role, although it is not clear that the household would experience an increase in its overall income during this period, as leave-taking is actually associated with a reduction in overall salaries, due to the fact that the benefit is only 55% of pay. Thus, mothers (and fathers) may be foregoing some salary if they choose to take longer leave from work as opposed to continue work at 100% of their salary. Even if income effects would play a role, it should be noted that the maximal monetary benefit from PFL as of 2004 was less than $5000 for the six weeks of paid leave. It would be very surprising if such an amount would have an effect on children's long-term health outcomes. Nevertheless, it should still be kept in mind that longer maternity leave mandates may have positive long-run implications in terms of maternal labor force participation,[7] and this could potentially increase household income in the long-run and in turn positively affect children's long-term health outcomes.

Rossin-Slater et al. (2013) and Baum and Ruhm (2016) find evidence of increased maternal employment up to 36 and 12 months after birth, respectively. We are not aware of existing studies linking maternal employment to our child health outcomes, with the exception of overweight. Anderson et al. (2003) and Ruhm (2008) both find evidence that greater maternal employment for children from higher socioeconomic

[5]In Rossin (2011), reduced prenatal stress is the suggested mechanism for findings on improved birth outcomes following the introduction of FMLA - unpaid job-protected maternity leave in the U.S.

[6]Persson and Rossin-Slater (2015) directly link maternal stress during pregnancy to ADHD in the offspring later in life.

[7]According to Ruhm (1998), increases in paid parental leave in Europe increased female employment but decreased their wages. Mark Curtis et al. (2016) examine new hires and labor market flows specifically in the context of the introduction of PFL in California. Their findings also suggest a decrease in wages among young women but an increase in employment rates. Baum and Ruhm (2016) and Rossin-Slater et al. (2013) find higher probabilities of maternal employment in California 9-36 months after the birth of a child following PFL's introduction. Schonberg and Ludsteck (2014) find very minimal long-term effects of extended maternity leave mandates on maternal labor market outcomes and Lalive and Zweimüller (2009) find no long-term effects on employment and earnings.

status while growing up increases these children's risk of overweight and obesity. Both papers find no effect on children from lower socioeconomic backgrounds.

3 California's Paid Family Leave Program

3.1 California Maternity Leave Prior to PFL

The United States is the only industrialized nation in the world to have no legislative mandate providing women paid maternity leave. At the federal level, the Family Medical Leave Act (FMLA) of 1993 provides up to 12 weeks of job-protected unpaid leave to employees of both genders for medical and family purposes, such as: personal or family illness, military service, family military leave, pregnancy, adoption, or the foster care placement of a child. FMLA covers less than 60% of the U.S. workforce, due to eligibility criteria based on the size of the employer and the tenure of the employee.[8]

In five states, new mothers can receive partial wage replacement through state-funded temporary disability insurance (TDI) programs: California, Hawaii, New Jersey, New York, and Rhode Island. In California, the TDI program is called State Disability Insurance (SDI). With the exception of Hawaii, these programs were created in the 1940's but did not apply to pregnant women until the 1970's with the passage of the Pregnancy Discrimination Act. TDI leave is job-protected. In 2004, all five states provided at least six weeks of TDI benefits for pregnancy at various wage replacement rates, ranging from 0.5 to 0.58, with a maximum weekly benefit rate ranging from $170 in New York to $728 in California.[9] In California, eligibility for SDI payments for expectant and new mothers begins four weeks before the due date and runs six or eight weeks immediately after birth for a vaginal or cesarean delivery, respectively.

The existence of FMLA job-protected leave for 12 weeks after the birth of a child along with SDI payments for 6-8 weeks immediately after the birth of a child resulted in two possibilities concerning maternity leave for California mothers prior to PFL's introduction: Mothers covered by FMLA (less than 60% of the workforce nationally) were eligible for 12 weeks of job-protected leave, out of which 6-8 weeks were paid with partial wage replacement; Mothers not covered by FMLA were eligible for 6-8 weeks of protected paid leave immediately after the birth of a child.

Employers can provide paid family leave benefits to their employees independent of the benefits set at the state and federal levels. While there are reports indicating that the percent of workers in the U.S. covered by employer-provided paid family leave policies has risen over the last two decades, coverage is still very low and was 11 percent of workers in 2012 (U.S. Department of Labor (2012); Van-Giezen (2013)).

[8] A survey conducted by the U.S. Department of Labor in 2000 showed that only 58.3% of employees in the U.S. were covered by the FMLA. Source: http://woe.dol.gov/whd/fmla/chapter3.htm

[9] The extremely low cap for New York of $170 in weekly payments implies that in New York the actual replacement rate is much less than 0.5 for the vast majority of the TDI recipient population.

3.2 The Introduction of PFL

California's PFL Program went into effect on July 1, 2004 and was the first program in the U.S. to provide paid leave of up to six weeks for workers who take time off to care for an ill family member or for parental bonding with a newborn, beyond disability insurance payments for mothers giving birth. The program extended paid leave eligibility for new mothers in California from the 6-8 weeks already available under the state's SDI program to 12-14 weeks.

The program passed as part of California's legislation in September 2002, 22 months prior to its implementation. Thus, the program's implementation was anticipated and mothers could have planned their pregnancies accordingly, if the monetary benefit of six weeks paid leave was sufficiently valuable. Indeed, Lichtman-Sadot (2014) presents evidence of changes in birth timings in California through November 2004 which were consistent with mothers postponing their births. In the construction of our dataset, we address this concern for the slight variation in mothers' observed or unobserved characteristics based on manipulation of birth timing during 2004, as further explained in Section 4.

The PFL program is funded by an employee-paid payroll tax as part of California's state disability insurance (SDI) program. Thus, any employee who was eligible for California's SDI program is also eligible for benefits under the PFL program. Unlike FMLA, California's PFL is nearly universal in its coverage: apart from some self-employed persons, virtually all private-sector (and nonprofit-sector) workers are included, regardless of the size of their employer. California public-sector employees may be covered if the agency or unit that employs them opts into the program, but most are not eligible for PFL. Workers need not have been with their current employer for any specific period of time to be eligible for PFL; they need only to have earned at least $300 in an SDI-covered job during any quarter in the "base period," which is five to seventeen months before filing a PFL claim.

California's PFL does not provide job protection. However, mothers covered by FMLA, which provides 12 weeks of *unpaid* job-protected leave for new parents, can combine the SDI and PFL payments to receive 12 weeks of *paid* job-protected leave during the period they are eligible for FMLA benefits. New mothers not covered by FMLA can receive the SDI and PFL payments for 12-14 weeks after giving birth (as specified above) but only the period of SDI payments is job-protected.

PFL significantly increased the amount of parental time spent with the child immediately after birth. Prior to PFL, California mothers took an average of three weeks of maternity leave after childbirth, and the introduction of PFL increased this leave-taking by 3 to 5 weeks (Baum and Ruhm (2016); Rossin-Slater et al. (2013)), thus more than doubling the amount of mothers' leave from work following the birth of a child.[10] This increased leave for new mothers was mostly concentrated around the time SDI payments end (Baum and Ruhm (2016)). Further, Rossin-Slater et al. (2013) show that other leave-taking (including vacations, layoffs and labor disputes) are unaffected by PFL, indicating that PFL actually increased maternity leave

[10] Bartel et al. (2015) show that the probability of leave-taking for California fathers increased by over 46 percent of the pre-PFL mean (6 days) in California after the introduction of PFL.

rather than merely crowding out other forms of paid and unpaid leave that mothers might have previously used in the time after birth. Rossin-Slater et al. (2013) find that PFL had especially large effects on leave-taking for black, non-college educated, unmarried and Hispanic mothers. The focus of this paper is the effect of this extended maternal leave - together with increased breastfeeding and reduced maternal stress resulting from it - on child health outcomes.

Since California's path-breaking introduction of PFL, several other states have implemented paid family leave programs - New Jersey in July 2009, Rhode Island in January 2014, and New York approved a PFL policy in April 2016.

4 Data

4.1 Sample

We use data from the Early Childhood Longitudinal Study (ECLS), a survey conducted by the National Center for Education Statistics (NCES), that longitudinally follows the development, school readiness, and early school experiences of children in the U.S. We use results from the kindergarten surveys, when children are 5-6 years of age.

The ECLS-K1999 and ECLS-K2011 follow a nationally-representative sample of children from over 1,000 kindergarten programs beginning in the fall of the 1998–99 and 2010-2011 school years, respectively. The ECLS-K1999 covers a population of over 21,000 kindergartners born late 1992 through 1993, prior to the introduction of PFL in California, while the ECLS-K2011 covers over 18,000 children born late 2004 through 2005, after the introduction of PFL in California. The ECLS-B covers a nationally-representative sample of children born in 2001, before the introduction of PFL, and followed from birth through kindergarten entry - 10,688 infants are included in the first wave but by kindergarten only 7,022 remain. The time span between our pre- and post-treatment periods is less than three years, which is the time difference between the latest ECLS-B birth and the earliest ECLS-K2011 birth.

Our sample is restricted to children whose state of residence is reported, who reside with their biological mother, who were born in the United States, and who do not have a twin sibling. We exclude children not residing with their biological mother in kindergarten since the ECLS-K datasets do not indicate who was caring for the child when s/he was born, whether the care-taker was eligible or not for PFL, and whether the mother was present to be able to breastfeed the child, one of our key proposed mechanisms for health outcome improvements. Twin siblings are excluded because mothers may have very different responses in terms of their return to work and breastfeeding initiation and duration following the birth of twins, which may in turn generate a different effect for twin siblings in response to PFL's introduction.

We limit the sample of states to those covered in the ECLS-K2011 and either the ECLS-K1999 or ECLS-B. Children from the excluded states do not contribute to the analysis, as they are observed only in either the pre-treatment or the post-treatment periods and our specifications include state fixed effects. This restriction resulted in the exclusion of roughly 2,000 children from 8 states and Washington, DC, with 42 states remaining in the sample.

We further limit our sample to children with non-missing values across all the health variables we investigate. This is to ensure that any varying effects across different health outcomes are not being driven by a different sample of children. While this exclusion results in the loss of roughly 5,000 observations for some of our health outcomes, the results are not very sensitive to this.

The PFL program was announced 22 months prior to its implementation on July 1, 2004. As such, we are concerned with mothers timing their pregnancies during 2004 in order to be eligible for the PFL benefit after the birth of their child. Indeed, Lichtman-Sadot (2014) presents evidence of changes in birth timings in California through November 2004 which were consistent with mothers postponing their births. This raises the concern that treatment is assigned to some California mothers who self-selected into a pregnancy occurring after PFL's introduction. If these mothers have observable or unobservable characteristics that are different from California mothers who gave birth prior to PFL's introduction, and these characteristics are correlated with our outcomes of interest, then our identification assumptions would be jeopardized.

Due to concern for mothers selecting their birth timing in California through November 2004, we wish to restrict our sample of California children from the ECLS-K2011 to those born beginning December 2004. Because December 2 is the cutoff date for starting kindergarten in California during our entire sample period (i.e. for children starting kindergarten between Fall 1998 and Fall 2010), we restrict the age composition of all kindergarten cohorts for all states based on their local kindergarten entry cutoff date. Thus, for each state in our sample, each cohort includes only children who entered kindergarten "on time" based on their local kindergarten entry cutoff date. We exclude children who were either "redshirted" and postponed their kindergarten entry by one year or children who entered kindergarten too early.[11, 12] This results in a consistent age composition across all states and cohorts, that accommodates our need to drop California children born prior to December 2004 in the kindergarten 2011 cohort.[13] Our consistent age composition of children across all states and cohorts ensures that the results are not biased, for example, by the exclusion of children who were "redshirted" in California, but the inclusion of such children in other states. If "redshirting" children becomes more prevalent in these other states over time, this may differentially affect observed child health outcomes in these other states.

[11]Several states (CO, MA, NH, NJ, NY, PA, VT) allow the kindergarten cutoff date to be determined by the local educational authority (LEA). For these states, we assigned a cutoff month based on the pattern of birth months for children who turned 6 before January 1 of their kindergarten year.

[12]The ECLS-B samples children born between January 1, 2001 and December 31, 2001. As such, these children attended kindergarten either during the 2006-2007 school year or during the 2007-2008 school year, depending on their local kindergarten cutoff dates and parental and professional discretion. We keep in our sample children from both kindergarten school years in the ECLS-B when their birth month is according to their state's kindergarten cutoff month for that specific school year. Our specifications control for the school year ECLS-B children are surveyed in, along with ECLS-survey-specific fixed effects.

[13]The exact months of birth included depends on the kindergarten cutoff date in each state for each cohort. Thus, in California (Dec. 2 cutoff), children born in December 2004 - November 2005 are included, while in a state such as Texas where the cutoff date is September 1 we include children born September 2004 - August 2005.

4.2 Health Outcomes and Treatment Assignment

We examine six health outcomes provided in the data: overweight, attention deficit/hyperactivity disorder (ADHD), a rating of the child's overall health condition, hearing problems, communication problems and a history of frequent ear infections.[14] All health outcomes are based on parents' responses to interviews during kindergarten, with the exception of overweight, for which we use height and weight measurements of the children from the kindergarten interview in order to calculate their BMI. We construct a dummy variable equal to one if a child's BMI in kindergarten was over the 85^{th} percentile for his/her age, which indicates that the child is overweight.[15] We construct three indicator variables equal to one if the child's parent reported by kindergarten that he/she was diagnosed with ADHD, hearing problems, or communication problems by a health professional.[16] Parents are asked in the kindergarten questionnaires about frequent ear infections as the child was growing up, and we create an indicator variable equal to one if the parent answered positively to this question.[17] Parents are requested to assess their child's health on a scale of 1 (best) to 5 (worst). We standardize these scales to Z-scores based on means and standard deviations of these scales in each survey.

In some specifications with ADHD as the dependent variable we use only first grade data from the two ECLS-K surveys, as the ECLS-B only surveys children through kindergarten. We need to use first grade, rather than kindergarten, ADHD data for statistical power in the differential effects analysis in Section 6.4, which separately estimates PFL's effect on population subgroups. Diagnosis rates for ADHD increase by more than 50%, and even double for some populations, between kindergarten and first grade. Moreover, although far from established in the medical and epidemiological literature, the leading causes of ADHD are related to environments and conditions very early in a child's life, most likely prior to pre-school. Thus, first grade diagnosis rates should more precisely capture the prevalence of ADHD, without being driven by contemporaneous conditions experienced by the child between kindergarten and first grade.

Our data has the advantage of covering a large and diverse sample of children in kindergarten from across the United States, along with information regarding their mothers. Despite the rich set of variables available from this data, the ECLS-K data does not provide information on the child's state of birth. Thus, our measure of treatment (i.e. being a child in California) is based on the child's state of residence while in kindergarten, which likely results in measurement error in our treatment variable. Furthermore, in the ECLS-K, there are no questions on the mother's labor force participation prior to giving birth or length of

[14] The following health outcomes are included in the ECLS questionnaires but we do not include them in our analysis, due to very low positive reporting of these (less than 1% of the sample): learning disability, dyslexia and autism. Questions about diagnosis of vision problems are also included in the surveys, but we do not include this variable in the analysis, due to significant changes in California public school vision testing during 2005 - see http://wow.cde.ca.gov/ls/he/hn/documents/visionreport.pdf. Child cognitive assessment scores are available for each ECLS study separately, but they are not comparable across studies - see https://nces.ed.gov/ecls/comparisons2011.asp

[15] BMI percentiles by age in months and sex are available from the 2000 Center for Disease Control (CDC) Growth Charts for the United States - http://www.cdc.gov/nchs/data/series/sr_11/sr11_246.pdf

[16] The diagnosis of hearing problems include difficulties due to ear infections, ear drum problems, toxic exposure to drugs, genetic causes, and other factors. Communication problems include, among others, articulation problems and stuttering.

[17] This health outcome is not available in the ECLS-B survey so all results presented for frequent ear infections use just the two ECLS-K surveys.

maternity leave after giving birth.[18] From this perspective, our estimates of any treatment effect will be at the intent-to-treat (ITT) level.[19]

4.3 Hearing Screening Data

An important robustness check is a regression specification with hearing problems as an outcome variable while controlling for the percent of newborns screened for hearing problems in each state by their year of birth (see discussion in Section 5.2 on the need for this robustness check). Our data on hearing screenings was constructed by Directors of Speech and Hearing Programs in State Health and Welfare Agencies (DSHPSHWA) in collaboration with the Centers for Disease Control and Prevention (CDC) Early Hearing Detection and Intervention (EHDI) via an annual survey sent to state representatives.

Unfortunately, this data is available only starting in 1999. As such, this robustness check is only possible in regression specifications that use the ECLS-B and ECLS-K2011 data, as ECLS-K1999 children were born in 1993. We observe large increases in hearing screening tests during this period - from a mean of 75.04% in 2001 (year of birth for ECLS-B children) to a mean of 94.75% in 2005 (latest year of birth for ECLS-K2011 children).

5 Empirical Strategy

5.1 Regression Specifications

We wish to assess how child health outcomes vary 5-6 years after birth, depending on whether the mother was exposed to extended paid maternity leave benefits resulting from the implementation of California's PFL program. We employ a difference-in-differences (DID) strategy, which compares outcomes between children born in California prior to PFL and children born in California after PFL's implementation, while controlling for trends in these outcomes among children outside of California during both these periods.

With the exception of the child's Z-score health scale, all our dependent variables are indicator variables. We report all our baseline results when the dependent variable is an indicator variable using a non-linear probit model. Thus, when our dependent variable is the child's Z-score health scale, our baseline specification takes the following standard DID form:

[18] Recall from Section 3 that women who had not worked 5-17 months prior to giving birth are not eligible for PFL benefits.

[19] The ECLS-K's do inquire in the kindergarten questionnaire about when the child began regular child care arrangements (home day care/person other than parents/formal child care). Unfortunately, a large number of parents do not report any child care outside of the home (13 percent and 11 percent of the ECLS-K2011 and ECLS-K1999, respectively), and for the observations that do report the existence of regular child care arrangements, the average age reported for children is much larger than the general age of children when mothers in the US return to work following the birth of a child. In particular, the average age at which a child is reported to have first received care outside the home is 21 months in both ECLS-K1999 and ECLS-K2011. This contrasts sharply with an average length of maternity leave of 10.3 weeks in the U.S. over 2006-2008 for the 71% of working women who take any maternity leave at all (U.S. Department of Health and Human Services (2011)). It also contrasts with the ECLS-B data which indicate that the average maternity leave was 11 weeks for those taking any maternity leave, and that among those mothers who worked (62% of all mothers at the time the child is 9 months old), they returned to, or started, work when the child was 13.7 weeks on average.

$$Outcome_{iys} = \alpha_0 + \alpha_1 California_{is} * PostPFL_{iy} + \alpha_2 ECLSSurvey_{iy} + \alpha_3 X_{iys} + \gamma_s + \varepsilon_{iys} \quad (1)$$

When the dependent variable in an indicator variable, our baseline specification takes the following form:

$$Outcome_{iys} = \Phi(\alpha_0 + \alpha_1 California_{is} * PostPFL_{iy} + \alpha_2 ECLSSurvey_{iy} + \alpha_3 X_{iys} + \gamma_s) \quad (2)$$

where Φ is the CDF of the normal distribution.

In equations (1) and (2), the dependent variable is a measure of child i's health in kindergarten, with y representing the ECLS survey the child took part in $y\epsilon\{K1999, B, K2011\}$ and s representing his/her state of residence. Our coefficient of interest is α_1, the coefficient on the dummy variable that is equal to one if the child is observed in California and was born after PFL's implementation. α_1 measures how our outcome of interest changed for California children who were born after PFL's implementation in comparison to children outside of California. α_2 controls for differences in the outcome's levels between children surveyed in different ECLS surveys and as such also captures differences for children who were born after PFL's implementation (i.e. children who are in the ECLS-K2011 survey). Within the ECLS-B survey, we also control for differences between children surveyed in kindergarten of different school years (2006-2007 vs. 2007-2008) by splitting the indicator for the ECLS-B survey into two.

The indicator for whether a child is observed in California is absorbed in the state-level fixed effects (γ_s) in the specification, which allow us to control for non-time-varying state-level characteristics that are correlated with our outcomes of interest. We also control for child-individual characteristics (X_{iys}): mother's age during the kindergarten interview (<26, 26 to 30, 31 to 35, 36 to 40, 41 to 45, >45, and missing), child's race (non-Hispanic white, non-Hispanic black, Hispanic, other, and missing), mother's education (less than high school, high school diploma, some college, bachelor, graduate studies, and missing), child's month of birth (indicators), child's age at start of kindergarten, household's socioeconomic status rating in quartiles, whether the child's mother was married at birth, whether English is the child's second language, the child's sex, number of older siblings the child has, and a quadratic function of household income.[20]

PFL's introduction in California had the greatest effect in terms of leave-taking duration after childbirth on less-advantaged mothers (Rossin-Slater et al. (2013)), who could not afford to take *unpaid* leave already provided by FMLA. As such, we wish to assess whether changes in outcomes varied differentially following

[20] Household socioeconomic scale and income may be endogenous to our specification, if the PFL program had an effect on overall household income in the long term (e.g., through different maternal labor force participation patterns). However, we note that our measure of socioeconomic status is categorical and indicates the household's quartile in the distribution of socioeconomic scales in the survey, and our measure of household income is derived from income brackets provided in the original ECLS data files. Thus, while the PFL may have affected household income or socioeconomic ranking, we do not think the effect will be substantial enough such that households change income brackets or quartiles in their socioeconomic scale. Nevertheless, all results are robust to the exclusion of household income and socioeconomic status as control variables.

California's PFL program, based on child, household or maternal characteristics. For this purpose, we expand the dummy variable $California_{is} * PostPFL_{iy}$ in equation (1) to interact with additional dummy variables for child, household or maternal characteristics, such as socioeconomic status, mother's education level, and child's gender. All other controls in the DID specification (i.e. the $California$ and $PostPFL$ indicator variables) also now interact with additional dummy variables for child characteristics.

As such, if our sample is broken down to n mutually exclusive groups ($n \geq 2$), with each group assigned a dummy variable $Char_{iy}^{j}$ for $j \epsilon \{1, ..., n\}$, which equals one for that group and zero otherwise (e.g., four indicator variables for which quartile of the socioeconomic status ranking the child's household belongs to), then our linear regression specification (equation (1)) will now be of the form:

$$\begin{aligned} Outcome_{iys} = \; & \beta_0 + \sum_{j=1}^{n} \beta_1^j California_{is} * PostPFL_{iy} * Char_{iy}^j \\ & + \sum_{j=1}^{n} \beta_2^j California_{is} * Char_{iy}^j + \sum_{j=1}^{n} \beta_3^j PostPFL_{iy} * Char_{iy}^j \\ & + \beta_4 X_{iys} + \gamma_s + \varepsilon_{iys} \end{aligned} \tag{3}$$

It was not possible to compute the third-order interaction effects required for estimating differential changes in response to California's PFL in non-linear probit models, due to failure of Stata's optimization routine in about 50% of our desired regressions. Thus, when the dependent variable is an indicator variable, we revert to a linear probability model (LPM), despite a relatively large fraction (5-13%) of predicted values falling outside the (0,1) range when examining hearing problems, ADHD, and frequent ear infections with differential effects based on children's characteristics. We note that a LPM for our baseline DID equation, without differential effects, produced similar estimates on the main interaction term for the DID specification as those from a probit model with the interaction effects calculated in the way proposed by Ai and Norton (2003).[21]

In equation (3), X_{iys}, γ_s and ε_{iys} are as defined in equation (1). Note that X_{iys} includes $Char_{iy}$, by its definition above. The coefficients of interest in equation (3) are β_1^j for $j \epsilon \{1, ..., n\}$. Each β_1^n estimates how our outcome of interest changed for California children born after PFL's implementation who belong to group $j \epsilon \{1, ..., n\}$ within the $Char$ breakdown.

5.2 Identification

Our main identification assumption is that absent the introduction of PFL, health outcomes for children both in and outside of California would have continued on the same trend. Our identification assumption would be violated if California children experienced any other policy introductions or changes to their

[21] All our standard errors are heteroskedasticity-robust, in addition to being clustered at the state level. This eliminates concerns for inefficiency of the standard errors, which arises from using OLS without accounting for heteroskedasticity within a LPM framework.

environment between 2001 and late 2004 - besides the introduction of PFL - which affected their health outcomes or the diagnosis rates of health outcomes. We are unaware of any such policy or environmental changes during this period that would affect our outcomes of interest, with the potential exception of hearing problems.[22]

During this period many states introduced legislation promoting newborn hearing screening programs.[23] In 1993 (the year of birth of most of the ECLS-K1999 cohort), only Hawaii and Rhode Island had mandated newborn hearing screening, while currently 43 states, including California, have legislative mandates related to newborn hearing screening programs. This likely contributed to the increase in the percentage of newborn screened for hearing loss from 47% in 1999 to 92% in 2004 in the U.S. (Curry and Gaffney (2010)) In turn, the percent of children positively diagnosed with hearing problems also increased during this period. This can raise concerns for our identification strategy if California was differentially slower in adopting these policies, thus increasing its diagnosis of hearing problems to a lesser extent than other states over our sample period. To address this concern, we conduct a robustness check for the results on hearing problems by including the percentage of newborns screened for hearing in a child's state and year of birth to control for the extent of hearing screening programs. Because data on the percent of newborns screened each year for each state is only available beginning 1999, this robustness check can only be performed using ECLS-B and ECLS-K2011 data (as children in ECLS-K99 were born in 1992/3).

We note that our results show an improvement across a range of health outcomes - ADHD, overweight, hearing problems, and ear infections - and it seems unlikely that each of these changes is being driven by a different policy or environmental change. Since our results demonstrate differential effects of PFL for children from advantaged and less advantaged backgrounds, changes in California's policies or environment that threaten our identification assumption would have to be targeted at these precise populations. We further assume as part of our identification that the average trend for children outside of California was maintained during this period - i.e. no substantial policy reforms or environmental changes occurred in states outside of California (with the exception of the increases in hearing screenings), and if any such changes did occur, they cancel each other out when examining the overall trend for children in all states outside of California.

With the exception of overweight, our health outcomes are parent-reported. Moreover, many of our health outcomes are based on diagnosis of a certain health condition by health professionals. This raises concern that any effect on these health outcomes in response to California's PFL is not reflecting an actual improvement in children's health, but rather a decrease in parents' awareness of potential health conditions or a decrease in access to healthcare professionals who diagnose these conditions. While we are unable to address this concern empirically and prove otherwise, we could not think of a compelling story or mechanism for why parents who spend a few more weeks at home with their child as an infant would experience

[22]For ear infections, the time span for which California children could not have experienced any policy introductions or significant changes related to this outcome would have to be 1999-2011, as the analysis only uses the ECLS-K surveys.

[23]For details see http://www.infanthearing.org/legislation/ and http://www.ncsl.org/research/health/newborn-hearing-screening-state-laws.aspx.

decreases in their awareness of potential health conditions or access to health professionals as the child is growing up. On the contrary, it would make more sense for there to be an *increase* in awareness of health conditions or access to health professionals if parents spend more time with their child after birth. In this case our estimated effect of California's PFL would be biased downwards.

Identification is also potentially threatened if, over our sample period, there were changes in the time-varying characteristics of California children or their mothers/households, which are also potential determinants of health outcomes. Such changes could be, for example, driven by compositional changes in California's population during this period relative to population trends in states outside of California. We note that it is highly unlikely that the introduction of PFL affected mothers' migration patterns into California, as the monetary benefit - capped at less than $5,000 over a six-week period - should not have justified inter-state migration. Another potential source for differential compositional changes in our sample is due to our exclusion of children who were redshirted - if California's PFL affected child health or other outcomes, then this could potentially change the composition of children being redshirted as parents feel their child is more or less ready for entering kindergarten after PFL's introduction.

We test whether maternal and child characteristics in California changed substantially between the periods before PFL's introduction and after PFL's introduction, relative to these characteristics in all other states, by estimating a variation of our DID model in equation (1) with child, maternal or household characteristics used as dependent variables (and no control variables). We do this using either the ECLS-K1999 and ECLS-B samples as our pre-treatment samples, or just the ECLS-K1999 sample as our pre-treatment sample, as the ECLS-K1999 is more similar to the ECLS-K2011 sample in its sampling procedure. Our tests are at the state-survey level. The results - described in more detail and presented in Table 8 in the Appendix - show that coefficient estimates in 3 of the 34 regressions run with maternal or household characteristics as dependent variables are statistically significant at the 10% level or less, which is consistent with 10% of tests being significant by chance.

We further address concerns regarding our identification assumption in several ways. We check for pre-existing trends using the ECLS-K1999 and ECLS-B samples within a DID framework, where ECLS-B children are assigned the $PostPFL$ indicator. We also conduct the analysis using a more comparable set of control states constructed through a variation of the synthetic control method specified in Abadie et al. (2010) and Abadie and Gardeazabal (2003). We elaborate on the synthetic control method and how it is used in this study in Section 7.1.

6 Results

6.1 Summary Statistics

Table 1 presents summary statistics, broken down by the period before treatment (ECLS-K1999 and ECLS-B) and the period after treatment (K2011) and for children in California and children outside of California. As can be seen, there are significant differences between California and other states in the racial composition and other socioeconomic variables (such as education or socio-economic status). This can raise concern that differences in trends in our health outcomes are being driven by either these observed differences or unobservable differences correlated with these variables. We will address this concern in two ways: First, we will show regression results both with and without controlling for child, mother and household characteristics. If the coefficient estimates between the two specifications are similar, then this suggests that our results are not driven by differences in either observed or unobservable characteristics. Second, we will implement a variation of the synthetic control method. This method will limit the states in our control group to those most comparable to California in terms of the outcome and characteristic trends over time prior to treatment. For further details, see Section 7.1.

We can use the means in Table 1 to get an initial sense of PFL's effect on California children by computing the differences between California and the other states in the pre-treatment (K1999 and B) and post-treatment (K2011) differences. The changes in the means over the two periods are given in the last column of Table 1, and suggest that PFL had a positive effect on the health outcomes of California children: there is a reduction of 4.3, 0.6, 2.2, 0.5, and 2.5 percentage points in the probability of overweight, ADHD, hearing problems, communication problems, and frequent ear infections, respectively. Post PFL, parents of California children report a reduction in the health scale of 0.056 of a standard deviation for the Z-score constructed for this measure, which represents an improvement in the health scale.

At first glance, it appears from Table 1 that the improvements in ADHD and hearing problems are being driven by a worsening of these health outcomes in other states rather than an improvement in California. However, the US as a whole experienced increasing ADHD rates over this period, and the increases in hearing problems are likely due to state legislative changes concerning hearing screenings. Pastor, Duran and Reuben (2015) show that ADHD diagnosis rates among US children aged 5-17 years increased from 7% in 1997-1999 to 10.2% in 2012-2014.[24] This increase in ADHD rates may be due to many factors including changes in diagnostic criteria, increased availability of checklists on the internet, and advocacy groups (Conrad and Bergey (2014)). The increase in hearing problem diagnosis in other states is likely due to the introduction of programs promoting newborn hearing screening, as discussed in Section 5.2 where we also address the potential concerns for identification raised by these increases in hearing screenings.

Table 1 also shows a significant decline in frequent ear infections across California and other states that mirrors the trend experienced for the US as a whole over this period (Marom et al. (2014)). Possible reasons for the decline in frequent ear infections over this period include increased breastfeeding, reduced smoking, and pneumococcal and influenza virus vaccines (Chonmaitree et al. (2016)).

[24] The numbers in Table 1 are lower than these overall diagnosis rates since we are only looking at diagnosis by the time the child is in kindergarten.

Table 1: Summary Statistics

	ECLS-K:1999 & ECLS-B				ECLS-K:2011				
	CA		Other States		CA		Other States		DID
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Estimate
Child health outcomes									
Overweight, indicator	0.312	0.464	0.279	0.449	0.284	0.451	0.293	0.455	-0.043
ADHD, indicator	0.006	0.080	0.014	0.116	0.007	0.083	0.020	0.141	-0.006
Health scale, z score	0.148	1.048	-0.024	0.987	0.088	1.032	-0.028	0.983	-0.056
Hearing problems, indicator	0.026	0.158	0.033	0.178	0.022	0.145	0.051	0.220	-0.022
Communication problems, indicator	0.050	0.219	0.081	0.272	0.045	0.207	0.081	0.272	-0.005
Frequent ear infections, indicator	0.234	0.424	0.325	0.468	0.078	0.269	0.193	0.395	-0.025
Child characteristics									
Age at start of kindergarten, months	63.790	3.398	65.940	3.623	63.740	3.364	65.980	3.541	-0.086
Female, indicator	0.498	0.500	0.501	0.500	0.506	0.500	0.495	0.500	0.014
Home language is not English, indicator	0.394	0.489	0.117	0.322	0.406	0.491	0.132	0.339	-0.003
White, indicator	0.233	0.423	0.562	0.496	0.225	0.418	0.569	0.495	-0.015
Black, indicator	0.070	0.256	0.167	0.373	0.018	0.133	0.126	0.332	-0.011
Hispanic, indicator	0.438	0.496	0.147	0.354	0.474	0.500	0.203	0.402	-0.020
Other race, indicator	0.258	0.438	0.124	0.329	0.282	0.450	0.101	0.302	0.047
Mother characteristics									
Age, years	33.590	6.147	32.890	5.975	34.970	6.106	33.640	5.942	0.627
Education: Grade 12 or less, indicator	0.207	0.405	0.115	0.319	0.217	0.412	0.110	0.313	0.015
Education: High school diploma, indicator	0.249	0.433	0.295	0.456	0.191	0.393	0.202	0.402	0.034
Education: Some college, indicator	0.290	0.454	0.327	0.469	0.264	0.441	0.335	0.472	-0.034
Education: Bachelor degree, indicator	0.163	0.370	0.169	0.374	0.204	0.404	0.214	0.410	-0.004
Education: Graduate degree, indicator	0.091	0.287	0.094	0.292	0.124	0.330	0.139	0.346	-0.011
Married at birth, indicator	0.703	0.457	0.711	0.453	0.713	0.453	0.700	0.458	0.021
Household characteristics									
Continuous SES measure	-0.018	0.848	0.065	0.792	-0.066	0.874	0.020	0.804	-0.002
Household income, constant 1999 dollars	56,257.1	58,514.1	53,929.1	52,264.7	55,090.7	50,232.2	51,956.2	44,009.3	806.5
Number of older siblings	0.932	1.124	0.881	1.037	0.877	0.947	0.907	1.017	-0.08
Observations	2,666		14,339		1,159		8,273		

Notes:
Sample excludes children not residing with their biological mother, born outside the U.S., with unidentified state of residence, or with a twin sibling. Sample further excludes children who turned six years old more than a year prior to their state's kindergarten cutoff date. ADHD, hearing problems, and communications problems are based on diagnosis by Spring of kindergarten year. Health scale is a measure of 1 through 5 that a parent provides in response to a question on the child's general health assessment, with 1 representing the best health, and 5 representing the poorest health.

6.2 Baseline Results

The results of our DID specifications in equations (1) and (2) are presented in Table 2 for overweight, ADHD and health scale (top panel), and for hearing-related outcomes (bottom panel). For each outcome of interest, the estimate for α_1 in equations (1) or (2) is reported when the regression is estimated with state fixed effects but without control variables, and with state fixed effects and control variables. As can be seen, the estimate for α_1 is relatively stable as we transition from specifications that exclude the child-level control variables to specifications that include them. This is with the exception of when the dependent variable is the indicator variable for frequent ear infections. This alleviates to some extent concerns that the results are being driven by differences between California and other states that are correlated with our outcomes of interest (except for possibly with regards to frequent ear infections).

Table 2 presents evidence that children born after the introduction of PFL in California had lower risks of being overweight, or being diagnosed with ADHD, hearing problems, or communication problems by kindergarten. Their parents were more likely to assess their child's overall health more positively and less likely to report a history of frequent ear infections. The reductions are sizable - a 4.1, 0.7, 2.4, 1.1 and 2.7 percentage point reduction in overweight, ADHD, hearing problems, communication problems and frequent ear infections, respectively, among California children born after PFL's introduction. For the health scale, reductions indicate that parents are reporting better overall health for their children, and we observe a decrease of 0.056 of a standard deviation for the Z-score constructed for this measure. With the exception of communication problems, these effects are very similar in magnitude to the simple differences in means computed in Table 1.

6.3 Testing for Pre-Existing Trends

The results of Table 2 crucially depend on the absence of pre-existing differential trends for California children in our health outcomes of interest. We present results for placebo tests using the ECLS-K1999 and ECLS-B data to evaluate whether any differential changes in children's outcomes are observed for California children in the ECLS-B data in comparison to the ECLS-K1999 data. Our objective is to rule out pre-existing health improvements prior to PFL's introduction.

The set of results at the top of Table 3 are probit or OLS results using the ECLS-B and ECLS-K1999 data. As can be seen, all coefficient estimates are not statistically significant, with the exception of when the dependent variable is health scale or communication problems. For health scale, the point estimate is positive, which is actually counter evidence of pre-existing trends of improvements in this health outcome. Our results on communications problems are not robust to the placebo analysis, as we do see a statistically significant negative coefficient, which is evidence of pre-existing trends of this condition's improvement prior to PFL's introduction in California. This may also explain the deviation of the estimated effect from the simple difference-in-means computed using Table 1.

We also note that we cannot test for pre-existing trends for frequent ear infections, as this outcome

Table 2: Changes in Health Outcomes for California Children Born After PFL's Introduction

	Overweight		ADHD		Health Scale (Z-Score)	
	State FE, No Controls	State FE & Controls	State FE, No Controls	State FE & Controls	State FE, No Controls	State FE & Controls
California*PostPFL	-0.042***	-0.041***	-0.010***	-0.007***	-0.048**	-0.056***
	(0.008)	(0.008)	(0.002)	(0.002)	(0.021)	(0.017)
Mean of Dependent Variable	0.287	0.287	0.015	0.015	-0.003	-0.003
Observations	26,437	26,435	26,379	26,349	26,437	26,437
	Hearing Problems		Communication Problems		Frequent Ear Infections	
	State FE, No Controls	State FE & Controls	State FE, No Controls	State FE & Controls	State FE, No Controls	State FE & Controls
California*PostPFL	-0.027**	-0.024***	-0.013***	-0.011**	-0.008	-0.027***
	(0.004)	(0.004)	(0.005)	(0.004)	(0.008)	(0.008)
Mean of Dependent Variable	0.037	0.037	0.076	0.076	0.256	0.256
Observations	26,379	26,377	26,437	26,435	17,469	17,468

Notes:
The coefficient estimates presented for California*PostPFL represent α_1 from equations (1) and (2) for health scale z-score (OLS model) and all other dependent variables (probate model), respectively. The sample is all ECLS surveys (K1999, ECLS-B, and K2011), with the exception of frequent ear infections as a dependent variable which is only the two ECLS-K surveys. Sample excludes children not residing with their biological mother, born outside the U.S., with unidentified state of residence, or with a twin sibling. Sample further excludes children who turned six years old more than a year prior to their state's kindergarten cutoff date or did not turn six by their state's kindergarten cutoff date. ADHD, hearing problems, and communications problems diagnosis are based on diagnosis by Spring of kindergarten year. Overweight is an indicator that the child's BMI is above the 85^{th} percentile based on their age in months and sex during kindergarten. Frequent ear infections is an indicator variable based on parents' reporting of frequent ear infections in the kindergarten questionnaire. Health scale is a measure of 1 through 5 that a parent provides in response to a question on the child's general health assessment, with 1 representing the best health, and 5 representing the poorest health. The sample excludes children who had missing values for any of the first five dependent variables presented. Despite this, sample sizes vary for the different dependent variables in the probit regression, due to some control variables in the regression predicting failure perfectly, and thus the need to exclude observations with certain values for these control variables when running probit regressions. Heteroskedasticity-robust standard errors clustered at the state level are in parenthesis. *** p<0.01, ** p<0.05, * p<0.1

variable is not available in the ECLS-B data, and so these results should be viewed with this caveat.

Because the placebo analysis utilizes only two, rather than three, ECLS samples, we must verify that smaller sample sizes are not driving the null effects found. The bottom panel of Table 3 provides our DID analysis with the ECLS-B (pre) and ECLS-K2011 (post) data - i.e. a DID using only two ECLS samples, with sample sizes even smaller than when using the ECLS-K1999 and ECLS-B samples. We would expect to find statistically significant health improvements for our outcomes, largely corresponding to the results in Table 2. Indeed, this is exactly what we observe in the bottom panel of Table 3, suggesting that small sample sizes are not causing the lack of statistical significance in the placebo tests.

As discussed in Section 5.2, there are potential concerns that the declines in hearing problems observed in California following PFL's introduction, relative to other states, can be attributed to a differential trend in the penetration of hearing screening programs in California instead of by PFL itself. Since the bottom panel of Table 3 utilizes only the ECLS-B and ECLS-K2011 samples, we can use it to present robustness checks for the hearing problems results that control for the percent of newborns undergoing hearing screenings in each child's state and year of birth. The hearing problem results presented in Table 3 confirm that the statistically significant decrease in hearing problems persists even with this additional control, thereby alleviating concerns that our results for hearing problems are related to hearing screening policies in California.

6.4 Differential Effects based on Child Characteristics

We next ask whether the improvements in health outcomes observed in Table 2 vary depending on child characteristics. We examine differential effects based on the following: quartile of child's socioeconomic ranking, mother's education, whether English is child's first or second language, and sex of child.

As discussed in Section 4.2, ADHD diagnosis rates are very low in kindergarten, such that for some of our sub-groups in the differential effects analysis, the number of California children with positive ADHD diagnosis in the pre- or post-treatment periods is less than 10. Because of this, a differential effects analysis with kindergarten ADHD data may be very sensitive to outliers and, hence, less credible, and so we instead use first grade ADHD diagnosis rates from both the ECLS-K datasets, as ECLS-B only surveys children through kindergarten. [25]

Table 4 presents means and standard deviations of our health outcomes by each of the child characteristics. We use these means to calculate the magnitude of the effect estimated at the child characteristic level and to examine general differences in health outcomes based on child characteristics. All our outcomes are based on all three ECLS surveys, with the exception of ADHD and frequent ear infections, which are based only on the two ECLS-K surveys. The number of observations for each child characteristic both within, and outside of, California is reported in the last two columns of each section in Table 4. There are at least several

[25]Nevertheless, the differential effects analysis using all three ECLS samples and kindergarten ADHD outcomes was very consistent with the first grade ADHD differential effects analysis. Improvements were driven by children from the lowest SES quartile and with mothers having less than a high school diploma. Significance levels were lower than those in the first-grade analysis, although p-values were lower than 0.1.

Table 3: Testing for Pre-Existing Improvements in California Children Health Outcomes and Results with only ECLS-B and ECLS-K2011

	Overweight	ADHD	Health Scale	Hearing Problems	Communication Problems
			ECLS-K1999 & ECLS-B		
California*PostPFL	-0.009	0.002	0.04*	0.014	-0.019**
	(0.012)	(0.004)	(0.020)	(0.010)	(0.008)
Mean of Dependent Variable	0.284	0.013	0.003	0.032	0.076
Observations	16,998	16,441	17,005	16,948	16,998
			ECLS-B & ECLS-K2011		
California*PostPFL	-0.044***	-0.008*	-0.083***	-0.026**	-0.012*
	(0.012)	(0.004)	(0.017)	(0.012)	(0.007)
Mean of Dependent Variable	0.306	0.020	0.001	0.060	0.086
Observations	14,080	14,013	14,081	14,022	14,080

Notes:
For sample restrictions and details on health outcomes, see Table 1. The ECLS sample used or specification is indicated at the top above each panel. The coefficient estimates presented for California*PostPFL represent α_1 from equations (1) and (2) for health scale Z-score (OLS model) and all other dependent variables (probit model), respectively. In the specification for the ECLS-K1999 and ECLS-B sample analysis, the PostPFL indicator is a placebo variable and is thus assigned to California observations from the ECLS-B sample. Results for hearing problems in the bottom panel (with the ECLS-B and ECLS-K2011 samples) control for the percent of newborns reported having had hearing screenings a the year of birth and state level. Heteroskedasticity-robust standard errors clustered at the state level are in parenthesis. *** p<0.01, ** p<0.05, * p<0.1

hundred observations in California for each characteristic, thus alleviating concern about lack of statistical power in the differential effects analysis.

According to Table 4, children are worse off in terms of overweight, ADHD and their health scale as their socioeconomic status or mother's education decreases. However, hearing problems, communication problems and frequent ear infections do not vary substantially based on socioeconomic status or mother's education.

Table 4 shows that children are worse off - or have higher diagnosis rates - for diagnosis-based health conditions - ADHD, hearing problems, communication problems, and frequent ear infections - when English is their first language. For overweight and the health scale - which are not diagnosis-based - the opposite is true.[26,27] This suggests reduced awareness or access to health professionals among immigrant populations, and that diagnosis rates for children with English as a second language are not reflecting the actual prevalence of these health conditions.[28]

Because some health outcomes vary substantially by sex, the last breakdown is by this characteristic. This is particularly so for ADHD - boys are nearly three times more likely than girls to be diagnosed with ADHD,[29] and this is also observed in our data. Communication problems are also substantially more prevalent among boys than they are among girls in our data. Lastly, parents report poorer health scales for their sons than for their daughters.

Differential effects in child health outcomes in response to the introduction of California's PFL may be driven by several factors. First, as explained in Section 3, California's PFL program increased leave-taking following the birth of a child primarily for mothers with lower socioeconomic status, who did not have any paid leave benefits from their employer prior to the policy and could not afford to take leave from work at no pay. Thus, we would expect greater effects among children coming from lower socioeconomic backgrounds. Nevertheless, the greatest effect does not necessarily have to be for the lowest socioeconomic ranking. This is because leave-taking is still associated with a 45% reduction in salary, it is not necessarily protected leave (which may offset the incentive to take leave for the lowest socioeconomic status mothers),[30] and the mother had to be working prior to giving birth. Thus, the most disadvantaged mothers may still be unable to take full advantage of the new policy.

Second, a higher prevalence of adverse health outcomes may lead to an estimated effect of greater magnitude. This would be consistent with an improvement in health for the same percent of children across a certain characteristic, which would translate into a larger percentage point reduction in the probability of

[26]English as a second language is most notably relevant for the Hispanic population, which is the ethnicity with the highest rates of child overweight and obesity in the U.S. Source: http://www.cdc.gov/obesity/data/childhood.html

[27]Recall that overweight is not parent-reported but rather based on BMI measurements during the survey.

[28]Elder (2010) and Evans et al. (2010) are two studies showing that an environment with greater access to professionals aware of ADHD - in their case the start of elementary school - substantially increases ADHD diagnosis rates.

[29]Source: http://www.cdc.gov/nchs/data/series/sr_10/sr10_258.pdf

[30]The paid leave provided under PFL is only protected if the employee is eligible for FMLA benefits, which are available to less than 60% of the workforce in the U.S. (see Section 3). Because FMLA covers only employees working for larger employers, and wages at larger employers are generally higher (Bayard and Troske (1999); Brown and Medoff (1989)), this results in the leave being protected for a greater proportion of advantaged mothers.

the adverse health outcome for population segments with greater prevalence. Furthermore, a statistically significant effect may be easier to detect for health outcomes with greater prevalence. Thus, despite the greater expected effect on children from lower socioeconomic backgrounds, we may also have a greater effect for hearing problems, communication problems, frequent ear infections, and ADHD among children with English as a first language, which is actually correlated with a higher socioeconomic ranking. We may also expect greater effects among male children as opposed to female children, given the higher prevalence of these adverse health outcomes among this population.

If children with English as a second language are not sufficiently diagnosed with certain health conditions, then we may not be able to detect an improvement in these health outcomes because the diagnosis rates observed in the data for this population are not reflective of the actual prevalence. Under such circumstances, null findings for PFL's effect on children with English as a second language do not necessarily indicate that there was no actual improvement in these health conditions among these children.[31]

Table 5 presents the results for our analysis of differential effects in response to PFL's introduction based on child characteristics. The table presents the estimates for β_1^j for $j \epsilon \{1, ..., n\}$ in equation (3). For all dependent variables, with the exception of the health scale variable, the regressions are linear probability models (see discussion on choice of econometric models in Section 5).

According to Table 5, for overweight and ADHD, decreases are driven by children from the lowest socioeconomic ranking and with mothers having less than a high school diploma. For overweight, reductions are 9.7 to 14.7 percentage points, while for ADHD, these are 3.1 to 3.7 percentage point reductions. Improvements in the health scale (0.134-0.144 of a standard deviation) and the decrease in hearing problems (2.5-4.5 percentage points) are driven by children from up to the third quartile of socioeconomic ranking and for mothers with relatively low education levels. For communication problems and frequent ear infections, we do not observe any statistically significant effects when the effect is broken down by either socioeconomic ranking or mother's education. Overall, the majority of our health outcomes exhibit improvements primarily for disadvantaged children, in terms of socioeconomic ranking and their mother's education.

We note that a statistically significant 6.5 percentage points reduction in overweight is also observed for children with mothers having a bachelor degree. Moreover, the estimated 2.9 percentage point decrease in the probability of being overweight is close to being statistically significant for children with English as a first language, with a p-value of 0.18. Thus, despite the greatest and most highly statistically significant reductions for overweight being observed among the least advantaged children, there is also some evidence of a reduction in overweight among higher SES children.

Diagnosis-dependent conditions - ADHD, hearing problems, communication problems, and frequent ear infections - decrease more among children with English as a first language, as opposed to children with

[31]We note that a greater prevalence of undiagnosed but present health conditions among certain population segments may also exist for other population breakdowns among our diagnosis-based health conditions. Although the pervalence of hearing problems, communication problems and frequent ear infections does not differ substantially based on socioeconomic ranking or mother's education levels, this may still reflect under-reporting of these conditions among certain population segments if, for example, children from lower socioeconomic status should have a higher prevalence of these health conditions.

Table 4: Means of Health Outcomes by Child Characteristics

	ECLS-K1999, ECLS-B, and ECLS-K2011 Samples						ECLS-K1999 and ECLS-K2011 Samples			
	Overweight, indicator	Health Scale, Z-Score	Hearing Problems, indicator	Communication Problems, indicator	Number of Observations CA	Number of Observations Other States	ADHD, indicator	Frequent Ear Infections, indicator	Number of Observations CA	Number of Observations Other States
SES Quartile 1	0.329	0.318	0.038	0.067	1,191	5,214	0.038	0.215	721	3,104
	(0.470)	(1.139)	(0.191)	(0.251)			(0.190)	(0.411)		
SES Quartile 2	0.322	0.036	0.040	0.086	806	5,807	0.035	0.260	449	3,740
	(0.467)	(0.987)	(0.195)	(0.280)			(0.184)	(0.439)		
SES Quartile 3	0.268	-0.117	0.041	0.077	864	5,795	0.029	0.273	513	4,112
	(0.443)	(0.909)	(0.197)	(0.266)			(0.167)	(0.446)		
SES Quartile 4	0.232	-0.232	0.030	0.074	964	5,796	0.018	0.278	627	4,171
	(0.422)	(0.847)	(0.172)	(0.261)			(0.131)	(0.448)		
Mother's Education Less than High School	0.329	0.380	0.035	0.062	802	2,559	0.027	0.198	482	1,467
	(0.470)	(1.176)	(0.183)	(0.242)			(0.161)	(0.398)		
Mother's Education High School Diploma	0.304	0.088	0.035	0.078	886	5,910	0.032	0.265	518	3,845
	(0.460)	(1.025)	(0.183)	(0.269)			(0.176)	(0.441)		
Mother's Education Some College	0.301	-0.030	0.040	0.079	1,079	7,452	0.036	0.267	634	5,067
	(0.459)	(0.964)	(0.196)	(0.269)			(0.186)	(0.443)		
Mother's Education Bachelor	0.236	-0.206	0.036	0.075	672	4,189	0.020	0.267	452	3,029
	(0.425)	(0.859)	(0.185)	(0.264)			(0.138)	(0.443)		
Mother's Education Graduate	0.244	-0.242	0.039	0.079	386	2,500	0.023	0.262	224	1,718
	(0.430)	(0.833)	(0.193)	(0.269)			(0.149)	(0.440)		
English Second Language	0.348	0.370	0.031	0.040	1,517	2,773	0.006	0.146	818	1,459
	(0.476)	(1.141)	(0.175)	(0.197)			(0.078)	(0.353)		
English First Language	0.275	-0.075	0.038	0.083	2,299	19,816	0.033	0.275	1,491	13,667
	(0.447)	(0.949)	(0.192)	(0.276)			(0.178)	(0.446)		
Male	0.297	0.047	0.042	0.099	1,911	11,337	0.043	0.276	1,151	7,625
	(0.457)	(1.022)	(0.201)	(0.299)			(0.181)	(0.447)		
Female	0.277	-0.053	0.032	0.052	1,914	11,275	0.015	0.240	1,159	7,502
	(0.448)	(0.967)	(0.176)	(0.222)			(0.176)	(0.427)		

Notes:
For sample restrictions and details on health outcomes, see Table 1. Number of observations is the number for each group of children with relevant characteristic with non-empty values for all health outcomes excluding frequent ear infections. A lower health scale indicates a higher overall health assessment for the child. Standard deviations are in parenthesis.

English as a second language. Statistically significant decreases of 1.6, 2.6 and 5.9 percentage points are observed for ADHD, hearing problems, and frequent ear infections, respectively. Most of the point estimates for English as a second language are negative, but none are close to being statistically significant. As discussed above, it is possible that this differential effect in favor of children with English as a first language is due to substantially higher diagnosis rates of these adverse health conditions at the baseline, in comparison to children with English as a second language. Moreover, these results do not necessarily rule out a health benefit for children with English as a second language, as there may be undiagnosed cases among these children, as discussed above. Lastly, we note that a differential effect in favor of children with English as a first language is not necessarily reflective of a greater effect for children with high socioeconomic ranking. In our data, roughly 16 and 30 percent of children with English as a first language are at the bottom and top quartiles of socioeconomic ranking, respectively. For the two outcomes that are not diagnosis-dependent - overweight and health scale - we do see an improvement primarily for children with English as a second language.

In terms of gender, the decrease appears to be driven by male rather than by female children. This is consistent both with evidence that the male population is more susceptible to many of these health conditions and with the mechanical explanation provided above that when the baselines are higher, if the percent improvements across subgroups are similar, subgroups with higher baselines will experience a greater magnitude in their improvements.

When the decreases are statistically significant, they are sizable relative to the means of the population segments observed in Table 4. Some of the average estimated decreases in hearing problems and ADHD may be too large, due to the use of a linear probability model with an indicator dependent variable. We note that all effects estimated are intent-to-treat (ITT). As such, the actual treatment effect on children whose mothers extended their maternity leave duration is likely larger. Furthermore, as labor force participation rates decrease with lower socioeconomic status, and for most health outcomes the effects are greater for children from lower socioeconomic status, the differences in the effects between higher and lower socioeconomically-ranked children are likely larger than the estimates presented in Table 5.

In addition to the results in Table 5, which present point estimates for the third-order interaction terms from equation (3), we also estimated differential changes in child health outcomes following PFL's introduction by running separate DID regressions (as specified in equations (1) and (2)) for each sub-group (i.e., each child characteristic). This can address concerns about the specification in equation (3), by allowing the effect of the control variables to vary across subgroups. Furthermore, because these are standard DID regressions, probit regressions can be estimated when the dependent variable is an indicator variable, rather than linear probability models. The results from these regressions can be found in Table 9 in the Appendix.

Table 5: Differential Changes in Health Outcomes for California Children Born After PFL's Introduction

Dependent Variable	Overweight	ADHD	Health Scale (Z-Score)	Hearing Problems	Communication Problems	Frequent Ear Infections
Socioeconomic Status Q1	-0.097***	-0.031**	-0.138**	-0.011	0.012	0.016
	(0.031)	(0.014)	(0.067)	(0.013)	(0.018)	(0.037)
Socioeconomic Status Q2	0.009	-0.026	-0.144*	-0.045***	-0.021	-0.067
	(0.037)	(0.017)	(0.079)	(0.015)	(0.022)	(0.044)
Socioeconomic Status Q3	-0.017	0.001	0.078	-0.025*	-0.021	-0.020
	(0.036)	(0.016)	(0.077)	(0.015)	(0.021)	(0.041)
Socioeconomic Status Q4	-0.040	0.001	-0.007	-0.010	-0.004	-0.059
	(0.033)	(0.015)	(0.072)	(0.014)	(0.020)	(0.037)
Mother's Education Less	-0.147***	-0.037**	-0.080	-0.016	0.017	0.049
than High School	(0.039)	(0.018)	(0.084)	(0.016)	(0.023)	(0.047)
Mother's Education High	-0.016	-0.005	-0.134*	-0.017	-0.011	-0.030
School Diploma	(0.037)	(0.017)	(0.080)	(0.016)	(0.022)	(0.044)
Mother's Education Some	0.012	-0.026*	-0.103	-0.042***	-0.029	-0.054
College	(0.032)	(0.015)	(0.069)	(0.013)	(0.019)	(0.038)
Mother's Education	-0.065*	-0.013	0.048	-0.020	0.002	-0.066
Bachelor	(0.039)	(0.017)	(0.083)	(0.016)	(0.023)	(0.043)
Mother's Education	-0.002	0.024	-0.003	0.006	-0.012	-0.032
Graduate	(0.051)	(0.024)	(0.109)	(0.021)	(0.030)	(0.061)
English Second Language	-0.078**	0.009	-0.180***	-0.012	-0.004	0.013
	(0.031)	(0.015)	(0.066)	(0.013)	(0.018)	(0.038)
English First Language	-0.029	-0.016*	0.020	-0.026***	-0.012	-0.059**
	(0.022)	(0.010)	(0.046)	(0.009)	(0.013)	(0.024)
Male	-0.064***	-0.024**	-0.038	-0.027***	-0.015	-0.049*
	(0.024)	(0.011)	(0.052)	(0.010)	(0.014)	(0.028)
Female	-0.019	0.001	-0.070	-0.019*	0.001	-0.013
	(0.024)	(0.011)	(0.052)	(0.010)	(0.014)	(0.028)

Notes:
For sample restrictions and details on health outcomes, see Tables 1 and 4. All coefficient estimates are from linear probability models (with the exception of the health scale dependent variable). The table presents coefficient estimates on the third-order interaction terms from equation (3). For each characteristic breakdown, a separate regression was run. Education levels are mother's educational attainment. SES Quartile is the Quartile of the child's socioeconomic ranking. A lower health scale indicates improved health outcomes. Number of observations ranges from 26,405 to 26,437, with the exception of ADHD and frequent ear infections, which only use the ECLS-K surveys and have a range of observations from 17,471 to 17,473. Heteroskedasticity-robust standard errors clustered at the state level are in parenthesis. *** p<0.01, ** p<0.05, * p<0.1

Table 6: The Synthetic Control Method and Only the ECLS-K Samples - Child Health in California in Various DID Specifications

	Overweight	ADHD	Health Scale	Hearing Problems	Communication Problems
	Synthetic Control Method				
California*PostPFL	-0.052**	-0.013**	-0.124**	-0.028***	0.001
	(0.016)	(0.004)	(0.028)	(0.005)	(0.008)
Mean of Dependent Variable	0.304	0.009	0.115	0.0284	0.0549
Number of States	6	6	5	5	5
Observations	8,344	8,344	7,920	7,920	7,473
	ECLS-K1999 & ECLS-K2011				
California*PostPFL	-0.043***	-0.010***	-0.041**	-0.018***	-0.020***
	(0.008)	(0.003)	(0.020)	(0.005)	(0.007)
Mean of Dependent Variable	0.275	0.029	-0.023	0.033	0.081
Observations	17,468	17,434	17,470	17,417	17,468

Notes:
For sample restrictions and details on health outcomes, see Table 1. The ECLS sample used or specification is indicated at the top of each panel. All health outcome dependent variables are as reported for the kindergarten ECLS surveys, with the exception of the ADHD result in the bottom panel (ECLS-K1999 & ECLS-K2011 samples), which is from the first grade survey (report of ADHD diagnosis by the Spring first grade survey). The coefficient estimates presented for California*PostPFL represent α_1 from equations (1) and (2) for health scale Z-score (OLS model) and all other dependent variables (probit model), respectively. This is with the exception of the synthetic control method results, which are from linear probability models for the indicator dependent variables, due to the need to weigh each observation in the regression analysis. Heteroskedasticity-robust standard errors clustered at the state level are in parenthesis. *** p<0.01, ** p<0.05, * p<0.1

7 Robustness Checks

We present two robustness checks. First, we use a variation of the synthetic control method, which allows us to define a more comparable control group of children based on their state of residence during kindergarten. Second, we verify that our results hold when evaluating our DID framework using the two ECLS-K surveys rather than all three ECLS surveys. The robustness check results are presented in Table 6. Table 6 shows the α_1 coefficient estimate from our DID equations (1) or (2) for linear probability (OLS for health scale) and probit models, respectively.

7.1 The Synthetic Control Method

The first robustness check is a variation of the synthetic control method. The synthetic control method is intended to address concerns that the control group in our DID specification is not comparable to our treatment group, i.e. California. If this is true, then the estimated effect of a policy change within a DID framework may be driven by differential trends in the treated unit that are due to different pre-trends in comparison to the control units. The synthetic control method addresses this concern by constructing a new control group, more similar to the treatment unit, based on pre-trends in the outcome of interest and control variables.

The new control group is children from a subset of the states used in the original control group. Weights are calculated for each state in the new set of control states, and these will correspondingly be assigned to children in the data from these states. The weights are constructed with the objective of minimizing the distance in quantifiable pre-treatment characteristics between California and the new weighted set of control states.

As such, for the construction of the synthetic controls (i.e. the weights for the control states), we first compare trends at the aggregated state-by-survey level over the two pre-treatment ECLS surveys - ECLS-K1999 and ECLS-B - for our different outcomes of interest and child, mother and household characteristics. The synthetic control method selects the weighted combination of states that best matches California on each health outcome of interest and the control variables for the pre-treatment surveys. This results in synthetic weights for each state. The weighted sum of the states creates a synthetic control state that minimizes distances between it and California for the outcome of interest and the control variables prior to treatment. Following Fitzpatrick (2008) and Courtemanche and Zapata (2014), who also applied this method to individual data, we then multiply these weights assigned at the state-level to individual children observed in each survey, based on their state of residence. Due to the need to run regressions with each observation having an assigned weight, the results presented for the synthetic control method with indicator dependent variables are linear probability models rather than probit models.

The results are presented in the top panel of Table 6, and are quite similar to our base analysis results, providing reassurance that our base results are not driven by using observations from states that may not be highly comparable to California. This is with the exception of when communication problems is the dependent variable, thus as in the pre-trends check presented in Table 3, our results for communication problems are not robust to the synthetic control method.

7.2 ECLS-K Surveys

The bottom panel of Table 6 presents the DID results for health outcomes using only the ECLS-K surveys. We emphasize that our results (including the differential effects presented in Table 5) are robust to regressions using only ECLS-K data, without ECLS-B data. This alleviates concern generated by the fact that sampling procedures substantially differed between the ECLS-K surveys and ECLS-B.

We take advantage of presenting the ECLS-K results to examine the coefficient estimate from the baseline specification - equation (2) - when using ADHD first grade data. See discussion in Section 4.2 on why first grade ADHD data is preferable over kindergarten ADHD data. All coefficient estimates suggest a statistically significant improvement in health for children residing in California and born following California's PFL. [32]

8 Discussion

8.1 Magnitudes Estimated

Because the magnitude of our estimated effects may seem large, we locate them within existing estimates from the related epidemiological literature on the effects of breastfeeding on child health outcomes. This comparison is not straightforward as the estimates are not entirely comparable. As discussed in Section 2.2, PFL could have improved child health outcomes through several channels, breastfeeding being only one of these (the others including reduced maternal stress and delaying child care outside the home). Thus, one may ex-ante expect our estimated effects to be larger than those found in the epidemiological literature. However, the epidemiological studies do not definitively establish causation. Because not accounting for confounding factors related to breastfeeding often increases the magnitude of the relationship between breastfeeding and positive outcomes, this could make our PFL estimates smaller in magnitude than the estimates from the epidemiological literature.[33] Lastly, as discussed in Section 4.2, our estimates represent intention-to-treat effects rather than treatment-on-the-treated effects, while the epidemiological estimates directly identify mothers' exact breastfeeding duration/initiation. We are able to address this last concern by scaling our ITT estimates to obtain an implied TOT estimate. Rossin-Slater et al. (2013) calculate that 0.596 of mothers had worked any hours in the previous year, which is the best estimate available for the proportion of mothers who were eligible for PFL, and we use this proportion to obtain implied TOT estimates from our ITT estimates.

For the reasons discussed above we expect our implied TOT estimates to be roughly in the range of estimates examining breastfeeding alone, with the caveat that these two sets of results are estimating the effect of two different, but related, treatments. We estimate the effect of PFL which increased mean paid maternity leave taken by 3 to 5 weeks, from an initial mean of three weeks (Rossin-Slater et al. (2013); Baum and Ruhm (2016)), while the epidemiological studies discussed examine varying degrees of breastfeeding. These include the effect of a binary treatment (ever breastfed vs never breastfed), as well as the effect of an increased duration of breastfeeding from 2 to 12 weeks. Even though the results are not completely

[32] We also ran the analysis with just the two ECLS-K surveys using the ADHD kindergarten outcomes. A 0.4 percentage point reduction is estimated (mean is 1.3%) with a p-value of 0.054.

[33] We note that the epidemiological studies are not purely correlational as they do attempt to deal with the presence of confounding factors in various ways, including the use of control groups (such as siblings) to account for unobserved factors and the use of regression controls to account for observed factors.

analogous, the comparison provides context for our seemingly large effects and shows they are comparable to estimates from other literature.

The results in this study suggest a moderate effect of the PFL program on overweight: those in the treatment group have a 31% lower odds of being overweight compared to those not treated.[34] This is similar to the estimated protective effect of breastfeeding on the risk of being overweight found by two separate studies; Tulldahl et al. (1999) find that Swedish adolescents who were exclusively breastfed for at least two months had a 34% lower odds of being overweight than those who were not, and similarly Hediger et al. (2001) find that for U.S. children aged between 3 and 5, breastfeeding reduces the odds of having a BMI between the 85^{th} and 94^{th} percentile by 37% compared to those who were never breastfed.

The epidemiological literature suggests that children who were breastfed for at least 12 weeks had a 44% lower risk of ADHD compared to those who were only breastfed for 2 weeks (Mimouni-Bloch et al. (2013)), and 68% lower odds of ADHD compared to those were not (Julvez et al. (2007)), while our estimated effects suggest that California's PFL lowered both the risk and odds of ADHD by 58%.[35] The results in this paper suggest that those in the treatment group have an 18% reduction in the risk of frequent ear infections compared to those not treated. This is significantly lower than the effects found by Uhari et al. (1996) who compute a pooled estimate indicating that breastfeeding reduces the risk of frequent ear infections by 52%.

Thus, the magnitudes of our estimated effects (of increased maternal leave) are generally close to the estimates in the epidemiological literature examining the effect of breastfeeding initiation and duration on child health outcomes. The notable exception is that our estimate of the reduction in ear infections is significantly lower than the estimates in the epidemiological literature. which suggests that the latter might not be adequately accounting for all confounding factors. Overall, the size of our results are close to those in the epidemiological literature which validates, at least to some extent, the reliability of our estimates.

8.2 Maternal Employment as Suggested Mechanisms

Our data enables us to examine patterns in maternal employment following California's PFL, which can potentially reveal underlying mechanisms for improved health outcomes. As discussed in Section 2.2, we are only aware of evidence linking maternal employment to child overweight, among all our health outcomes (Anderson et al. (2003); Ruhm (2008)).

Table 7 presents estimates from equation (3) on three variables representing maternal employment while the child is in kindergarten: an indicator as to whether the mother was employed; the number of "usual" working hours per week the mother reported; and an indicator as to whether the mother reported working

[34] Since the epidemiological literature focuses on odds ratios and risk ratios, and not treatment effects, we convert our estimated treatment effect to odds ratios as follows: $OR = \frac{(\bar{y}+\alpha_1)/(1-\bar{y}-\alpha_1)}{\bar{y}/(1-\bar{y})}$, where $\bar{y}$ =mean of the dependent variable, and α_1 =estimated treatment effect, and to risk ratios as follows: $RR = \frac{\bar{y}+\alpha_1}{\bar{y}}$, where $\bar{y}$ =mean of the dependent variable, and α_1 =implied treatment-on-the-treated, obtained by scaling the estimated intention-to-treat effect by 0.596 (an estimate of the proportion of mothers eligible for PFL).

[35] This estimate of 58% refers to the first grade ADHD results from Table 6, which are preferred since they are more precise as ADHD diagnosis rates increase substantially after kindergarten. Furthermore, Mimouni-Bloch et al. (2013) evaluate a sample of children aged 6-12 years and so it may be misleading to compare their estimates with our kindergarten estimates.

full-time.[36] The results very clearly suggest a decrease in employment among mothers from a high socioeconomic ranking - from the highest socioeconomic quartile, with a bachelor or graduate degree, and with English as a first language. For children from a lower socioeconomic ranking, the results in Table 7 suggest an increase in maternal employment.[37,38]

In our differential effects analysis in Table 5, the results suggest that children from higher socioeconomic status experience a decrease in the risk of overweight following PFL's introduction, in addition to the large and statistically significant decreases in overweight among children from lower socioeconomic ranking. The results of both Table 5 and Table 7 are consistent with Anderson et al. (2003) and Ruhm (2008), who show that decreased labor force participation among advantaged mothers reduces their children's probability of being overweight. This suggests a potential mechanism for the decreases in overweight among high socioeconomically-ranked children observed in Table 5 - they are driven by mothers' lower labor force participation rates.[39]

We emphasize that caution must be exercised when interpreting the results in Table 7 as suggestive of mechanisms for child health improvements following California's PFL. A reverse direction of causality is also plausible - i.e. better child health increases maternal employment (Corcnan et al. (2005)).

9 Conclusions

Despite the fact that one of the objectives of extended maternity leave is health improvement for both mother and child, very few studies have actually evaluated the effects of longer maternity leave mandates on maternal or child health outcomes.[40] To the best of our knowledge, this study is the first to evaluate changes in children's health outcomes following the introduction of the first paid family leave program in the U.S.: California's PFL.

We use a simple difference-in-differences framework with data from the Early Childhood Longitudinal Studies, which document child outcomes for kindergarten children who were born before and after the introduction of PFL. Our results suggest that following PFL's introduction, California children experienced lower probabilities of overweight, ADHD, hearing problems and frequent ear infections, although we cannot test for pre-trends or the robustness of the specification for frequent ear infections. We also find that parents' overall assessment of their children's health improved. While some specifications suggest a

[36] The last two variables are for the entire population rather than being conditional on working.

[37] We note that these results do not necessarily contradict the results of Rossin-Slater et al. (2013) and Baum and Ruhm (2016), according to which maternal employment 1-3 years after birth increased following PFL, as these studies did not evaluate differential effects by mothers' characteristics, and in particular for Baum and Ruhm (2016), it is likely that their sample is more representative of less advantaged mothers as their use of NLSY-97 data results in a sample of relatively young mothers.

[38] We note that the results in Table 7 are robust to all alternative specifications presented in this paper in Section 7.

[39] Anderson et al. (2003) and Ruhm (2008) do not find that maternal labor force participation affects children from less advantaged backgrounds, so Table 7's results of increased labor force participation among disadvantaged mothers do not contradict Table 5's findings on better overweight outcomes among disadvantaged children. Thus, improved overweight outcomes among disadvantaged children following California's PFL are likely due to other factors attributed to PFL's introduction (e.g., breastfeeding) rather than changes in maternal labor force participation.

[40] Notable exceptions include: Ruhm (2000) and Rossin (2011) on birth outcomes and infant/child mortality and Chatterji and Markowitz (2005) on maternal health.

Table 7: Maternal Employment following California's PFL - Differential Effects

Dependent Variable	Maternal Employment	Mother Working Hours	Mother Working Full Time
Socioeconomic Status Q1	0.072*	2.486*	0.049
	(0.037)	(1.499)	(0.039)
Socioeconomic Status Q2	0.001	-1.332	-0.036
	(0.043)	(1.743)	(0.045)
Socioeconomic Status Q3	-0.031	-0.749	-0.012
	(0.042)	(1.710)	(0.044)
Socioeconomic Status Q4	-0.106***	-5.176***	-0.106***
	(0.038)	(1.553)	(0.040)
Mother's Education Less	0.048	0.911	0.015
than High School	(0.046)	(1.879)	(0.049)
Mother's Education High	0.016	0.364	0.003
School Diploma	(0.046)	(1.886)	(0.049)
Mother's Education Some	0.027	0.329	-0.008
College	(0.037)	(1.499)	(0.039)
Mother's Education	-0.118***	-3.618**	-0.040
Bachelor	(0.044)	(1.800)	(0.047)
Mother's Education	-0.130**	-8.317***	-0.206***
Graduate	(0.057)	(2.321)	(0.060)
English Second Language	0.031	0.828	0.030
	(0.038)	(1.519)	(0.040)
English First Language	-0.031	-1.641	-0.038
	(0.025)	(0.998)	(0.026)
Male	-0.031	-1.940*	-0.053*
	(0.028)	(1.143)	(0.030)
Female	-0.006	-0.635	-0.012
	(0.028)	(1.144)	(0.030)

Notes:
For sample restrictions, see Table 1. The dependent variable stated at the top of each column is a measure of maternal employment as of when the child starts kindergarten. All coefficient estimates are from linear probability models (with the exception of the dependent variable for number of weekly hours employed). The table presents coefficient estimates of the third-order interaction terms from equation (3). SES Quartile is the Quartile of the child's socioeconomic ranking. For each characteristic breakdown, a separate regression was run. Number of observations is 23,832-23,855. Heteroskedasticity-robust standard errors clustered at the state level are in parenthesis. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

decrease in the probability of communication problems following PFL's introduction, these results are not robust to several alternative specifications, including the placebo analysis and the synthetic control method.

The results suggest that improvements in health outcomes are driven primarily by children from disadvantaged backgrounds - i.e. with lower socioeconomic ranking and with mothers with lower educational attainment. While we also observe substantial health improvements for children with English as a first language, we stress that this is not necessarily reflective of higher socioeconomic ranking, as 16% of children with English as a first language in our sample ranked at the lowest socioeconomic quartile. We also note that the null findings among children with English as a second language for diagnosis-dependent conditions do not necessarily indicate a lack of effect on these children's actual health, since low diagnosis levels among children of immigrants may be due to their reduced access to health professionals. Our findings of a greater effect among more disadvantaged children are consistent with the fact that PFL primarily benefited less advantaged mothers, who could not afford to take unpaid leave following the birth of a child.

Improvements in health outcomes appear to be driven primarily by male children, as opposed to female children. This is important, as male children are generally more susceptible to these adverse health conditions, in particular ADHD.

Our DID framework is very simple and only measures an intent-to-treat (ITT) effect of California's PFL. As such, the results cannot guarantee that it is indeed the PFL program that generated these improved health outcomes among California children, as opposed to other factors. Nevertheless, we emphasize that any other factor or combination of factors that would compromise the notion that California's PFL is driving these results would have to affect this specific combination of health outcomes - all negatively associated with increased breastfeeding, and other potential implications of increased parental leave benefits - at this specific timing, and be driven by children coming from lower socioeconomic backgrounds. With this in mind, and given the findings showing no pre-existing time trends in these health outcomes and that the results still hold when using a more narrowly-defined set of control states, based on the synthetic control method, the overall evidence presented in the paper is suggestive of positive health benefits resulting from extended paid maternity leave benefits.

Two of the health outcomes analyzed in this paper - overweight and ADHD - are among the most serious and common child health problems in the U.S. today. According to the CDC, approximately 17.7% of children ages 6-11 were obese in the U.S. during 2011-2012. In addition, approximately 9.5% of children ages 4-17 had been diagnosed with ADHD as of 2011-2013.[41]

Overweight and ADHD are frequently health conditions lasting through adulthood, and in the case of obesity, they are associated with additional and even more severe health problems, such as cardiovascular complications, diabetes, or some types of cancer. As such, the economic and social costs of these conditions are quite high in terms of treatment and medication, labor market participation, and the social well-being of these children, and finding policy measures that may reduce their prevalence can have important impli-

[41]Sources: http://www.cdc.gov/obesity/data/childhood.html and http://www.cdc.gov/NCH's/data/databriefs/db201.htm

cations. The prevalence of these conditions is greater among children with lower income levels,[42] and the results of this paper raise the question of whether this gap in child obesity or ADHD rates based on income can be at least partially attributed to less parental leave take-up among low-income mothers following the birth of a child. If this is true, then paid parental leave policies may be one effective channel for narrowing the inequality between children born to high and low-income families in terms of their health outcomes, as well as much longer-term outcomes through adulthood (Currie (2009)). We hope that this paper is only the beginning in attempting to uncover the answers behind these important issues.

[42]See http://www.cdc.gov/obesity/downloads/pednssfactsheet.pdf for obesity and Pastor, Reuben, Duran and Hawkins (2015) for ADHD.

References

Abadie, A., Diamond, A. and Hainmueller, J. (2010). Synthetic control methods for comparative case studies: Estimating the effect of california's tobacco control program, *Journal of the American Statistical Association* **105**(490): 493–505.

Abadie, A. and Gardeazabal, J. (2003). The economic costs of conflict: A case study of the basque country, *American Economic Review* **93**(1): 113–132.

Ai, C. and Norton, E. C. (2003). Interaction terms in logit and probit models, *Economics Letters* **80**(1): 123–129.

Anderson, P. M., Butcher, K. F. and Levine, P. B. (2003). Maternal employment and overweight children, *Journal of health economics* **22**(3): 477–504.

Appelbaum, E. and Milkman, R. (2011). Leaves that pay: Employer and worker experiences with paid family leave in california, *Technical report*, Center for Economic and Policy Research.

Arenz, S., Ruckerl, R., Koletzko, B. and von Kries, R. (2004). Breast-feeding and childhood obesity - a systematic review, *International journal of obesity* **28**(10): 1247–1256.

Baker, M. and Milligan, K. (2008). Maternal employment, breastfeeding, and health: Evidence from maternity leave mandates, *Journal of Health Economics* **27**(4): 871–887.

Baker, M. and Milligan, K. (2010). Evidence from maternity leave expansions of the impact of maternal care on early child development, *Journal of Human Resources* **45**(1): 1–32.

Bartel, A., Rossin-Slater, M., Ruhm, C., Stearns, J. and Waldfogel, J. (2015). Paid family leave, fathers' leave-taking, and leave-sharing in dual-earner households, *Technical report*, Working Paper.

Baum, C. L. (2003). Does early maternal employment harm child development? an analysis of the potential benefits of leave taking, *Journal of Labor Economics* **21**(2): 409–448.
URL: *http://www.jstor.org/stable/10.1086/345563*

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

Bayard, K. and Troske, K. R. (1999). Examining the employer-size wage premium in the manufacturing, retail trade, and service industries using employer-employee matched data, *American Economic Review* pp. 99–103.

Berger, L. M., Hill, J. and Waldfogel, J. (2005). Maternity leave, early maternal employment and child health and development in the us*, *The Economic Journal* **115**(501): F29–F47.
URL: *http://dx.doi.org/10.1111/j.0013-0133.2005.00971.x*

Brown, C. and Medoff, J. (1989). The employer size-wage effect, *Journal of Political Economy* **97**(5): 1027–1059.
URL: *http://www.jstor.org/stable/1831885*

Carlson, K. (2015). Fear itself: The effects of distressing economic news on birth outcomes, *Journal of Health Economics* **41**: 117 – 132.
URL: *http://www.sciencedirect.com/science/article/pii/S0167629615000272*

Carneiro, P., Loken, K. V. and Salvanes, K. G. (2015). A flying start? maternity leave benefits and long-run outcomes of children, *Journal of Political Economy* **123**(2): 365–412.
URL: *http://www.jstor.org/stable/10.1086/679627*

Cawley, J. (2010). The economics of childhood obesity, *Health Affairs* **29**(3): 364–371.

Chatterji, P. and Markowitz, S. (2005). Does the length of maternity leave affect maternal health?, *Southern Economic Journal* **72**: 16–41.

Chonmaitree, T., Trujillo, R., Jennings, K., Alvarez-Fernandez, P., Patel, J. A., Loeffelholz, M. J., Nokso-Koivisto, J., Matalon, R., Pyles, R. B., Miller, A. L. et al. (2016). Acute otitis media and other complications of viral respiratory infection, *Pediatrics* p. e20153555.

Conrad, P. and Bergey, M. R. (2014). The impending globalization of adhd: Notes on the expansion and growth of a medicalized disorder, *Social Science and Medicine* **122**(2): 31–43.

Copper, R. L., Goldenberg, R. L., Das, A., Elder, N., Swain, M., Norman, G., Ramsey, R., Cotroneo, P., Collins, B. A., Johnson, F., Jones, P. and Meier, A. (1996). The preterm prediction study: Maternal stress is associated with spontaneous preterm birth at less than thirty-five weeks' gestation, *American Journal of Obstetrics and Gynecology* **175**(5): 1286 – 1292.
URL: *http://www.sciencedirect.com/science/article/pii/S000293789670042X*

Corcnan, H., Noonan, K. and Reichman, N. E. (2005). Mothers' labor supply in fragile families: The role of child health, *Eastern Economic Journal* **31**(4): 601–616.

Courtemanche, C. J. and Zapata, D. (2014). Does universal coverage improve health? the massachusetts experience, *Journal of Policy Analysis and Management* **33**(1): 36–69.

Currie, J. (2009). Healthy, wealthy, and wise: Socioeconomic status, poor health in childhood, and human capital development, *Journal of Economic Literature* **47**(1): 87–122.
URL: *http://www.aeaweb.org/articles.php?doi=10.1257/jel.47.1.87*

Currie, J. and Rossin-Slater, M. (2015). Early-life origins of life-cycle well-being: Research and policy implications, *Journal of Policy Analysis and Management* **34**(1): 208–242.

Currie, J., Stabile, M., Manivong, P. and Roos, L. L. (2010). Child health and young adult outcomes, *Journal of Human Resources* **45**(3): 517–548.

Curry, A. and Gaffney, M. (2010). Directors of speech and hearing programs in state health and welfare agencies (dshpshwa): overview and summary of 1999-2004 dshpshwa data, *Technical report*, Center for Disease Control and Prevention.

Danzer, N. and Lavy, V. (2017). Paid parental leave and children's schooling outcomes, *The Economic Journal* .

Duffy, L. C., Faden, H., Wasielewski, R., Wolf, J. and Krystofik, D. (1997). Exclusive breastfeeding protects against bacterial colonization and day care exposure to otitis media, *Pediatrics* **100**(4): e7–e7.

Duncan, B., Ey, J., Holberg, C. J., Wright, A. L., Martinez, F. D. and Taussig, L. M. (1993). Exclusive breastfeeding for at least 4 months protects against otitis media, *Pediatrics* **91**(5): 867–872.

Dustmann, C. and Schonberg, U. (2012). Expansions in maternity leave coverage and children's long-term outcomes, *American Economic Journal: Applied Economics* **4**(3): 190–224.
URL: *http://www.jstor.org/stable/23269736*

Elder, T. E. (2010). The importance of relative standards in adhd diagnoses: evidence based on exact birth dates, *Journal of health economics* **29**(5): 641–656.

Evans, W. N., Morrill, M. S. and Parente, S. T. (2010). Measuring inappropriate medical diagnosis and treatment in survey data: The case of adhd among school-age children, *Journal of health economics* **29**(5): 657–673.

Fitzpatrick, M. D. (2008). Starting school at four: The effect of universal pre-kindergarten on children's academic achievement, *The BE Journal of Economic Analysis & Policy* **8**(1).

Fletcher, J. M. (2014). The effects of childhood adhd on adult labor market outcomes, *Health economics* **23**(2): 159–181.

Hediger, M. L., Overpeck, M. D., Kuczmarski, R. J. and Ruan, W. J. (2001). Association between infant breastfeeding and overweight in young children, *Jama* **285**(19): 2453–2460.

Huang, R. and Yang, M. (2015). Paid maternity leave and breastfeeding practice before and after california's implementation of the nation's first paid family leave program, *Economics & Human Biology* **16**: 45–59.

Julvez, J., Ribas-Fito, N., Forns, M., Garcia-Esteban, R., Torrent, M. and Sunyer, J. (2007). Attention behaviour and hyperactivity at age 4 and duration of breast-feeding, *Acta Paediatrica* **96**(6): 842–847.

Lalive, R. and Zweimüller, J. (2009). How does parental leave affect fertility and return to work? evidence from two natural experiments, *The Quarterly Journal of Economics* **124**(3): 1363–1402.

Lichtman-Sadot, S. (2014). The value of postponing pregnancy: California's paid family leave and the timing of pregnancies, *The BE Journal of Economic Analysis & Policy* **4.4**: 1467–1499.

Liu, Q. and Nordstrom, S. (2009). The duration of paid parental leave and children's scholastic performance, *The B.E. Journal of Economic Analysis and Policy* **10**(1).

Mark Curtis, E., Hirsch, B. T. and Schroeder, M. C. (2016). Evaluating workplace mandates with flows versus stocks: An application to california paid family leave, *Southern Economic Journal* .

Marom, T., Tan, A., Wilkinson, G. S., Pierson, K. S., Freeman, J. L. and Chonmaitree, T. (2014). Trends in otitis media - related health care use in the united states, 2001-2011, *JAMA pediatrics* **168**(1): 68–75.

Mimouni-Bloch, A., Kachevanskaya, A., Mimouni, F. B., Shuper, A., Raveh, E. and Linder, N. (2013). Breastfeeding may protect from developing attention-deficit/hyperactivity disorder, *Breastfeeding Medicine* **8**(4): 363–367.

Owen, C. G., Martin, R. M., Whincup, P. H., Smith, G. D. and Cook, D. G. (2005). Effect of infant feeding on the risk of obesity across the life course: a quantitative review of published evidence, *Pediatrics* **115**(5): 1367–1377.

Paradise, J. L., Rockette, H. E., Colborn, D. K., Bernard, B. S., Smith, C. G., Kurs-Lasky, M. and Janosky, J. E. (1997). Otitis media in 2253 pittsburgh-area infants: prevalence and risk factors during the first two years of life, *Pediatrics* **99**(3): 318–333.

Pastor, P. N., Duran, C. and Reuben, C. (2015). Quickstats: percentage of children and adolescents aged 5-17 years with diagnosed attention-deficit/hyperactivity disorder (adhd), by race and hispanic ethnicity-national health interview survey, united states, 1997-2014, *MMWR Morb Mortal Wkly Rep* **64**(33): 925.

Pastor, P. N., Reuben, C. A., Duran, C. R. and Hawkins, L. D. (2015). Association between diagnosed adhd and selected characteristics among children aged 4-17 years: United states, 2011-2013., *National Center for Health Statistics data brief* **201**.

Persson, P. and Rossin-Slater, M. (2015). Family ruptures and intergenerational transmission of stress, *Technical report*, Stanford University.

Rasmussen, A. W. (2010). Increasing the length of parents' birth-related leave: The effect on children's long-term educational outcomes, *Labour Economics* **17**(1): 91 – 100.
URL: *http://www.sciencedirect.com/science/article/pii/S0927537109000785*

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**(2): 221 – 239.
URL: *http://www.sciencedirect.com/science/article/pii/S0167629611000117*

Rossin-Slater, M., Ruhm, C. J. and Waldfogel, J. (2013). The effects of california's paid family leave program on mothers' leave-taking and subsequent labor market outcomes, *Journal of Policy Analysis and Management* **32**(2): 224–245.
URL: *http://dx.doi.org/10.1002/pam.21676*

Ruhm, C. J. (1998). The economic consequences of parental leave mandates: Lessons from europe, *The Quarterly Journal of Economics* **113**(1): 285–317.

Ruhm, C. J. (2000). Parental leave and child health, *Journal of Health Economics* **19**(6): 931 – 960.
URL: *http://www.sciencedirect.com/science/article/pii/S0167629600000473*

Ruhm, C. J. (2008). Maternal employment and adolescent development, *Labour Economics* **15**(5): 958–983.

Schonberg, U. and Ludsteck, J. (2014). Expansions in maternity leave coverage and mothers' labor market outcomes after childbirth, *Journal of Labor Economics* **32**(3): 469–505.
URL: *http://www.jstor.org/stable/10.1086/675078*

Tanaka, S. (2005). Parental leave and child health across oecd countries, *The Economic Journal* **115**(501): F7–F28.
URL: *http://dx.doi.org/10.1111/j.0013-0133.2005.00970.x*

Tulldahl, J., Pettersson, K., Andersson, S. W. and Hulthén, L. (1999). Mode of infant feeding and achieved growth in adolescence: early feeding patterns in relation to growth and body composition in adolescence, *Obesity research* **7**(5): 431–437.

Uhari, M., Mäntysaari, K. and Niemelä, M. (1996). Meta-analytic review of the risk factors for acute otitis media, *Clinical Infectious Diseases* **22**(6): 1079–1083.

U.S. Department of Health and Human Services (2011). Women's health usa 2011, *Technical report*, Health Resources and Services Administration, Maternal and Child Health Bureau, Rockville, Maryland.
URL: *http://www.mchb.hrsa.gov/whusa11/more/downloads/pdf/w11.pdf*

U.S. Department of Labor (2012). National compensation survey: Employee benefits in the united states, march 2012, *Technical report*, U.S. Bureau of Labor Statistics.

Van-Giezen, R. W. (2013). Paid leave in private industry over the past 20 years, *Pay & Benefits, Bureau of Labor Statistics* **2**(18).
URL: *http://www.bls.gov/opub/btn/volume-2/paid-leave-in-private-industry-over-the-past-20-years.htm*

Appendix

A Placebo Tests: Did Maternal and Child Characteristics Change Differentially in California?

We test whether maternal and child characteristics in California changed substantially between the two ECLS-K samples used, relative to these characteristics in all other states, by estimating a variation of our DID model in equation (1) with child, maternal or household characteristics used as dependent variables (and no control variables). The coefficient estimate presented is that of α_1 in equation (1). As no child-level controls are used in these regressions, we collapse the ECLS data to the state-survey level. The results are presented in Table 8 - first when just using the two ECLS-K surveys (first 3 columns) followed by using all ECLS surveys (including ECLS-B) for the analysis (last 3 columns). Table 8 presents evidence that child, maternal or household characteristics did not change substantially in California after the introduction of PFL. A total of 34 regressions are run with 17 characteristics in two different samples. The coefficient estimates in 3 of the 34 regressions are statistically significant at the 10% level or less, which is consistent with 10% of tests being significant by chance.

B Differential Effects following California's PFL - DID Regressions Run Separately for Each Subgroup

We test whether the results presented in Table 5 are robust to DID regressions that are run separately for each subgroup, rather than regressions with third-order interaction terms based on child characteristic breakdowns for the entire sample. Thus, our regressions take the form of equations (1) or (2) when the dependent variable is not an indicator variable or is an indicator variable, respectively.

Most of the results - presented in Table 9 - are very similar to those presented in Table 5, with similar magnitudes estimated. Most estimates have greater statistical significance than those presented in Table 5, due to a substantial decrease in the standard errors for the point estimates. However, there are a few instances with coefficient estimates that are positive and statistically significant - overweight for mothers with some college, health scale for children from the third socioeconomic quartile, communication problems for the lowest socioeconomic quartile or with mothers having less than a high school diploma, and frequent ear infections for mothers with less than a high school diploma. We do not have an interpretation for these results, which are not consistent with Table 5's results and do not correspond to improvements in health outcomes following California's PFL. We note that these positive and significant point estimates do not exhibit a consistent pattern and appear rather sporadic. We further note that overall 77 coefficient estimates are presented in Table 9, out of which 5 are positive and statistically significant and 37 are negative and

Table 8: Correlation between the Introduction of PFL and Child/Mother/Household Characteristics

	ECLS-K1999, ECLS-B, and ECLS-K2011			ECLS-K1999 & ECLS-K2011		
	CA*PostPFL	Mean of Dependent Variable	R-squared	CA*PostPFL	Mean of Dependent Variable	R-squared
Mother's Education - Grade 12 or Less	0.0236	0.124	0.490	-0.0133	0.102	0.927
	(0.0303)			(0.0184)		
Mother's Education - High School Diploma	0.00181	0.243	0.533	-3.62e-05	0.239	0.749
	(0.0202)			(0.0292)		
Mother's Education - Some College	-0.0341	0.313	0.417	-0.0224	0.323	0.582
	(0.0258)			(0.0308)		
Mother's Education - Bachelor	0.0308	0.203	0.485	0.0323	0.217	0.626
	(0.0312)			(0.0470)		
Mother's Education - Graduate or more	-0.0221	0.117	0.594	0.00344	0.120	0.778
	(0.0148)			(0.0198)		
Child White	-0.0633	0.533	0.665	0.0240	0.620	0.783
	(0.0450)			(0.0570)		
Child Black	-0.0235	0.131	0.787	-0.0517	0.122	0.830
	(0.0218)			(0.0339)		
Child Hispanic	0.0105	0.151	0.825	-0.0229	0.140	0.889
	(0.0204)			(0.0292)		
Child Other Race	0.0763**	0.184	0.599	0.0505	0.118	0.786
	(0.0329)			(0.0388)		
Child Female	0.0247	0.493	0.409	-0.0451	0.485	0.674
	(0.0256)			(0.0312)		
Mother Married at Birth	-0.0289	0.691	0.517	0.0403	0.733	0.719
	(0.0316)			(0.0326)		
Mother's Age	0.475	32.99	0.689	1.050**	33.42	0.809
	(0.346)			(0.435)		
English Second Language	-0.00234	0.106	0.706	-0.00155	0.0836	0.858
	(0.0186)			(0.0181)		
Age in Mos. at Start of Kindergarten	-0.382	65.88	0.659	-0.277	65.86	0.732
	(0.241)			(0.414)		
Number of Older Siblings	-0.0862**	0.865	0.589	-0.0559	0.893	0.759
	(0.0409)			(0.0430)		
Household Income	-1,988	52469	0.645	620.8	55065	0.763
	(3,222)			(4,627)		
Socioeconomic Status Ranking	-0.00248	0.0451	0.665	0.0828	0.0968	0.741
	(0.0607)			(0.0826)		

Notes:
Sample excludes children not residing with their biological mother, born outside the U.S., with unidentified state of residence, or with a twin sibling. Sample further excludes children who turned six years old more than a year prior to their state's kindergarten cutoff date or before their state's kindergarten cutoff. States not appearing in either ECLS-K2011 or both ECLS-B and ECLS-K1999 are excluded (8 states and Washington, D.C.). Children with one of the six health outcome variables missing are also excluded. Coefficient estimates are for α_1 from the DID regression specification reported in equation (1), with the listed child/mother/household characteristic as the dependent variable. The data is collapsed at the state-survey level. Number of observations is 120 for the first three columns (all ECLS data) and 77 in the last three columns (ECLS-K1999 and ECLS-K2011 data). Heteroskedasticity-robust standard errors clustered at the state level are in parenthesis. *** p<0.01, ** p<0.05, * p<0.1

statistically significant.

Table 9: Child Health Outcomes following California's PFL by Child Characteristics (DID Regressions)

Dependent Variable	Overweight	ADHD	Health Scale (Z-Score)	Hearing Problems	Communication Problems	Frequent Ear Infections
Socioeconomic Status Q1	-0.103***	-0.032***	-0.124***	(0.008)	0.017*	0.006
	(0.015)	(0.008)	(0.039)	(0.010)	(0.010)	(0.022)
Socioeconomic Status Q2	0.018	-0.018**	-0.150***	-0.061***	-0.037***	-0.045***
	(0.017)	(0.008)	(0.032)	(0.007)	(0.008)	(0.015)
Socioeconomic Status Q3	-0.017	-0.004	0.063**	-0.024**	-0.022**	-0.014
	(0.012)	(0.006)	(0.024)	(0.011)	(0.009)	(0.017)
Socioeconomic Status Q4	-0.027**	-0.002	-0.020	-0.015**	(0.010)	-0.059***
	(0.013)	(0.005)	(0.029)	(0.007)	(0.009)	(0.015)
Mother's Education Less	-0.142***	-0.023**	-0.109*	-0.031**	0.020*	0.072**
than High School	(0.023)	(0.011)	(0.060)	(0.012)	(0.012)	(0.032)
Mother's Education High	-0.015	-0.014	-0.143***	-0.029***	-0.017*	-0.051***
School Diploma	(0.014)	(0.009)	(0.028)	(0.010)	(0.009)	(0.019)
Mother's Education Some	0.020*	-0.020***	-0.107***	-0.047***	-0.036***	-0.052***
College	(0.011)	(0.006)	(0.022)	(0.006)	(0.006)	(0.015)
Mother's Education	-0.051***	-0.014**	0.052	-0.006	-0.004	-0.066***
Bachelor	(0.016)	(0.006)	(0.034)	(0.009)	(0.010)	(0.021)
Mother's Education	0.008	0.008	-0.037	-0.005	-0.025	0.019
Graduate	(0.017)	(0.014)	(0.046)	(0.012)	(0.017)	(0.030)
English Second Language	-0.071***	N/A	-0.184***	-0.044***	0.001	0.004
	(0.016)		(0.039)	(0.017)	(0.010)	(0.024)
English First Language	-0.031***	-0.014***	0.017	-0.028***	-0.017***	-0.042***
	(0.009)	(0.003)	(0.016)	(0.005)	(0.005)	(0.007)
Male	-0.061***	-0.022***	-0.045**	-0.028***	-0.022**	-0.062***
	(0.012)	(0.005)	(0.020)	(0.007)	(0.009)	(0.013)
Female	-0.017*	0.004	-0.067***	-0.020***	0.002	0.008
	(0.010)	(0.004)	(0.021)	(0.006)	(0.004)	(0.011)

Notes:
For sample restrictions and details on health outcomes, see Tables 1 and 5. Each coefficient estimate is from a separate regression with observations only from the subgroup listed for that row. The coefficient estimates presented are for α_1 from equations (1) or (2) for health scale Z-score (OLS model) and all other dependent variables (probit model), respectively. There were no observations from the ESL California population with non-zero ADHD values in the ECLS-K1999. As such, the regression with the ESL population and ADHD as the dependent variable could not be estimated in a probit model. *** p<0.01, ** p<0.05, * p<0.1