The Effect of Prenatal Maternity Leave on Short and Long-term Child Outcomes

by

Alexander AHAMMER

Martin HALLA

Nicole SCHNEEWEIS

The Effect of Prenatal Maternity Leave on Short and Long-term Child Outcomes[†]

Alexander Ahammer[a,b], Martin Halla[a,b,c,d], Nicole Schneeweis[a,b,c,e]

[a]*Johannes Kepler University Linz, Austria*
[b]*Christian Doppler Laboratory Aging, Health, and the Labor Market, Linz*
[c]*IZA, Institute for the Study of Labor, Bonn*
[d]*GÖG, Austrian Public Health Institute, Vienna*
[e]*CEPR, Centre for Economic Policy Research, London*

forthcoming in:
Journal of Health Economics

Last update: October 16, 2019

ABSTRACT

Maternity leave policies are designed to safeguard the health of pregnant workers and their unborn children. We evaluate a maternity leave extension in Austria which increased mandatory prenatal leave from 6 to 8 weeks. We exploit that the assignment to the extended leave was determined by a cutoff date. We find no evidence for significant effects of this extension on children's health at birth or long-term health and labor market outcomes. Subsequent maternal health and fertility are also unaffected. We conclude that employment during the 33rd and 34th week of gestation is not harmful for expecting mothers (without major problems in pregnancy) and their unborn children.

JEL Classification: J13, I18, J28, I13, J83, J88.
Keywords: Maternity leave, fetal origins hypothesis, infant health, birth outcomes, birth weight, long-term child outcomes, fertility.

I. Introduction

Developed countries have special regulations in place to address the safety and health of pregnant workers and their unborn children. One important element of these regulations is *maternity leave* (ML). This is the temporary employment-protected period of absence for women around the time of childbirth and should be distinguished from parental leave.[1] There is considerable variation in ML arrangements across countries in terms of pre- and postnatal durations, obligation to take ML, job-protection, and income support.[2] The median duration of paid and job-protected ML in OECD member countries listed in Table 1 amounts to 16.5 weeks, 6 weeks of which can be taken prior to birth. In this paper, we are interested in *prenatal* ML. We evaluate the impact of maternal employment during pregnancy on child and maternal outcomes. Despite the popular belief that prenatal ML is beneficial to the infant and mother, the empirical evidence on the impact of prenatal ML is limited and existing policies are not evidence-based.

[Table 1]

We evaluate a prenatal ML extension in Austria. Until 1973, statutory ML prohibited employment 6 weeks before to (usually) 6 weeks after delivery. The reform in 1974 increased mandatory prenatal ML from 6 to 8 weeks. All other ML regulations (such as the associated transfer payments) remained unaffected by the reform. Our estimation strategy exploits that the assignment to the extended leave was determined by a cutoff date. This gives rise to a fuzzy regression discontinuity design (RDD), which provides us with a *local average treatment effect* (LATE) that identifies the causal effect of an extended *prenatal* ML duration due to being assigned to the new regulations.

Our research design has a number of interesting features. *First,* assigned and non-assigned mothers, while having different prenatal ML durations, were both mandated to the *same postnatal* ML duration and eligible for the same parental leave. This allows us to cleanly identify the effect of variation in prenatal ML, not only on birth outcomes, but also on post-birth outcomes. Since the reform took place in 1974, we are able to study its long-run effects on children and mothers up to 40 years after birth. This is important, since the fetal origins hypothesis stresses that (health) effects of prenatal events may remain latent for many years (Almond and Currie, 2011*a*,*b*). To the best of our knowledge, none of the existing studies on prenatal ML or employment during pregnancy examine potential long-run impacts. *Second*, to check the robustness of our results, we can additionally use information on unaffected non-working mothers, who

---

[1]The leave that often follows ML and allows one or both parents to remain at home to care for young children is usually called parental leave (see *OECD Family database*, "Child-related leave: PF2.1 Key characteristics of parental leave systems," updated: October, 2017). We follow this semantic convention throughout the paper.

[2]Currently, 32 states have ratified the *Maternity Protection Convention* issued by the *International Labour Organization* (ILO), which mandates, among others, at least 14 weeks of ML and an entitlement to cash and medical benefits.

are not eligible for ML. This second source of variation extends our RDD with a difference-in-differences (DiD) approach. The DiD component differences out potential seasonal effects and accounts for any unobserved characteristics that follow a seasonal pattern between children born in different months. Thus, the combination of these two sources of variation ensures a clean identification of treatment effects. *Third*, we can rely on high-quality administrative data sources covering the universe of all births in Austria. The *Austrian Social Security Database* (ASSD) provides information on the mother's eligibility for ML, her actual leave duration, and her return to work behavior. Since we observe the actual duration of prenatal ML for each mother, we are able to not only estimate an *Intention to Treat effect* (ITT) as previous studies in this literature, but we can also identify a LATE of prenatal ML. The *Austrian Birth Register* comprises several outcomes to assess children's health at birth, and enables us to closely track subsequent maternal fertility. The ASSD further allows to assess children's long-term human capital outcomes (up to 40 years of age) and maternal mortality. For a subsample of mothers and children, we also have data on long-term health outcomes (i.e, health care utilization between 25 and 40 years after birth). *Fourth*, the institutional setting promotes a clear interpretation of our results. We analyze the intensive margin of prenatal ML. More specifically, our LATE captures a reduction of *in utero* exposure to maternal employment in the 33rd and 34th week of gestation for a group of mothers without major problems in this stage of pregnancy.

We consider this estimate to be informative in two ways. It contributes to the literature on the consequences of maternal behavior, in particular employment, in the third trimester of pregnancy on short and long-term health and human capital outcomes. Furthermore, it is informative for designing prenatal ML policies and assessing existing ML regulations. The Austrian ML legislation is comparable to that of other OECD member countries. Among the 16 countries given in Table 1, the median duration of paid and job-protected prenatal ML is 6 weeks, and in many countries, such as Austria, the Czech Republic, Germany, Greece, Hungary or Italy, ML is mandatory for a period of 6 or more weeks prior to birth.

There are several potential mechanisms through which extended prenatal ML could alter the health of pregnant workers and their unborn children. First, the extended absence from work while enjoying job-protection and full income replacement should reduce expecting mother's psychological and physiological stress level. The mandatory nature of ML should reinforce this channel, since pregnant women do not have to justify their leave-taking towards their employer and co-workers. Nevertheless, it is also possible that a longer absence from work increases mental stress for some groups of women. Wüst (2015) argues that exclusion from employment can have adverse effects on psychological stress. Second, certain groups of workers could benefit from a reduction in specific occupational exposures.[3] For women whose counterfactual home environment is healthier than their job environment, an extended prenatal ML should

---

[3]Examples are second-hand tobacco smoke in the hospitality industry (Bharadwaj et al., 2014), chemicals in certain branches of manufacturing (Chen et al., 2000; Snijder et al., 2012), anaesthetic gases and antineoplastic drugs in the medical sector (Lawson et al., 2012), low levels of radiation in the aviation industry, or shift work (Bonzini et al., 2011) and noise.

have positive effects. At the same time, it cannot be ruled out that, for some women, the counterfactual home environment is less beneficial. In this case, an increase in prenatal ML may have negative effects. In our research design, we can abstract from self-selection into ML with respect to the relative quality of the work versus home environment, since ML is mandatory. Finally, the modified allocation of time (i.e., substituting work with leisure) may also lead to healthier behavior. Expecting mothers may have more time to rest, to follow a healthy diet, or to do necessary prenatal medical check-ups.

The existing literature provides evidence for the importance of these mechanisms. The fetal origins hypothesis and supporting empirical evidence emphasize a number of factors in the prenatal environment that are important for later child and adult outcomes (Almond and Currie, 2011a,b). Maternal stress is one such factor. Most studies distinguish the effects of prenatal stress by pregnancy trimester of exposure. The reform we consider in this paper has the potential to reduce maternal stress in the third trimester (more specifically in the 33rd and 34th week of pregnancy). Multiple studies provide evidence that prenatal stress has adverse effects on birth outcomes throughout pregnancy. For instance, Black et al. (2016) find negative effects of stress induced by the death of the mother's parent during pregnancy on birth outcomes with similar effects across all trimesters of exposure.[4] Persson and Rossin-Slater (2018), studying an equivalent treatment with a focus on long-run mental health outcomes, confirm this pattern. Although we do not directly observe maternal stress (e.g., with cortisol levels), a reduction in stress in the third trimester is likely an important causal channel of our treatment. Regarding healthier behavior during pregnancy, a number of factors (such as nutrition and physical activity) are discussed. While causal evidence is lacking for some of these determinants, the importance of prenatal check-ups is documented in design-based studies. For example, Evans and Lien (2005) exploit a 1992 bus strike in Pennsylvania, which led to a sharp decline in prenatal care visits among women pregnant at that time. They conclude that prenatal check-ups reduce maternal smoking and enhance birth weight.[5]

We find no evidence for an impact of the prenatal ML extension on children's health at birth. The estimated treatment effects are statistically insignificant and precisely estimated zero effects. This finding is consistent across subsamples of mothers who are expected to be more vulnerable, such as blue-collar workers. In line with this zero effect on children's health outcomes in the short-run, we also find no evidence for significant effects on long-run health and labor market outcomes. Treated and untreated children have statistically indistinguishable labor market and health outcomes up to the age of 40. Thus, there is also no evidence for latent effects that manifest later in life. Our analysis of subsequent maternal fertility and health also does not reveal any significant effects of the reform. Treated and untreated mothers do not significantly

---

[4]This finding is consistent with previous studies on the effects of prenatal exposure to stressful events such as armed conflicts (Mansour and Rees, 2012) or hurricanes (Currie and Rossin-Slater, 2013). Earlier papers using landmine explosions (Camacho, 2008) and a large earthquake (Torche, 2011) find the strongest effects in the first trimester.

[5]Sonchak (2015) finds similar effects of prenatal care on birth weight for disadvantaged white mothers.

differ in their completed fertility and the timing of subsequent births. The same holds true for maternal health measures up to 40 years after birth. We therefore conclude that the reform had no measurable effects on children and mothers.

The political justification for this reform was to improve the health of pregnant workers and their children. Our evaluation provides no evidence for any impact of the extension from 6 to 8 weeks of prenatal ML. In contrast, the reform has clear cost. It has increased public spending on transfer payments for prenatal ML by one-third and additional cost for firms cannot be ruled out. Importantly, some women may prefer to work during this period, but are not allowed to. While our results must be interpreted within the scope of the Austrian setting, we conclude more generally that mandatory prenatal ML starting in the 35th week of gestation is sufficient for pregnant workers without problems in pregnancy. It should be emphasized that we do not interpret our results as a general argument against (mandatory) prenatal ML. Quite the contrary, we consider our finding to be valuable for designing an optimal prenatal ML policy.

Our findings add to a small stock of empirical evidence on prenatal ML and employment during pregnancy on child and maternal outcomes.[6] So far, only a handful of design-based papers provide evidence on the effects of prenatal ML.[7] With regards to the United States, there are two studies available. Rossin (2011) evaluates the effects of 12 weeks unpaid but job-protected ML introduced by the *The Family Medical Leave Act* (FMLA) in 1993. This policy allows mothers to take a leave during their pregnancy and/or after childbirth. The author's identification is based on variation in FMLA policies across states and variation in firm coverage. She finds that unpaid ML led to small increases in birth weight, decreases in the likelihood of a premature birth, and substantial decreases in infant mortality. These effects are present only for children of highly educated and married mothers, who were most able to take advantage of unpaid leave. The coefficients on infant mortality are larger in magnitude than on birth outcomes, suggesting that the effects of FMLA are more likely to stem from leave taking after birth. Stearns (2015) evaluates the effect of state-based access to paid ML on health at birth. She exploits the fact that five states were required to start providing wage replacement benefits to pregnant women in the year 1978 through their *Temporary Disability Insurance* (TDI) programs. Eligible women could access this *de facto* paid ML in the period immediately before or after birth. Based on state-level data she implements a difference-in-differences approach, which suggests that access to six weeks of paid ML lowered rates of low birth weight and preterm births by around 3 and 7 percent, respectively. In contrast to Rossin (2011), the effects were driven by disadvantaged African American and unmarried mothers.

Wüst (2015) uses Danish data to study the effect of maternal employment during pregnancy

---

[6]In contrast, the effects of maternal employment after childbirth and during the first years of a child's life is extensively studied. In particular, there are a number of design-based papers on the effect of different postnatal maternity and parental leave durations on child outcomes available (Liu and Skans, 2010; Baker and Milligan, 2010; Rasmussen, 2010; Baker and Milligan, 2015; Dustmann and Schönberg, 2012; Carneiro et al., 2015; Dahl et al., 2016; Danzer et al., 2017)

[7]The evidence from observational studies on the effects of working conditions on pregnancy outcomes is summarized by two meta-analyses (Mozurkewich et al., 1999; Palmer et al., 2013).

on birth outcomes. She focuses on the pregnancy weeks 12 and 30. To account for selection into employment she exploits variation across pregnancies and compares outcomes of mothers' consecutive children. She finds that mothers who are employed (either in week 12 or 30) are *less* likely to have a preterm birth. As a potential explanation for this finding she discusses maternal stress caused by not working in a country with a particularly high female employment rate. Del Bono et al. (2012) provide structural parameters of the production functions of birth weight and fetal growth. Applying family-fixed effects models and GMM techniques to American and British survey data, the authors estimate effects of maternal employment during pregnancy. In contrast to Wüst (2015), they find that work interruptions, especially in the last 3 months of pregnancy, are beneficial for birth outcomes. Most recently, Chuard (2018) exploits an Austrian parental leave reform which has indirect effects on prenatal employment for subsequent births. She uses variation in prenatal employment prior to the 32nd week of pregnancy and finds no effects on birth outcomes.

The remainder of the paper is organized as follows: In Section II, we present our research design. We provide details on the ML system, the reform in the year 1974, and other relevant aspects of the institutional setting. We describe our data sources and present our estimation strategy. In Section III, we present our estimation results along with a number of robustness checks. Section IV concludes and discusses potential policy implications.

II. Research design

II.1. Institutional background

In this section, we summarize the institutional background and describe the ML system before and after the 1974 reform. To enhance the understanding of the context we first provide information on female labor force participation. Finally, we describe changes in the public prenatal care program over time.

II.1.1. Female labor force participation

Throughout the 1970s, labor force participation rates remained quite constant in Austria. Among women between 15 and 60 years of age the rate was around 55 percent. The equivalent male rate amounted to roughly 85 percent. The highest female participation rate among all age groups in 1971 was 62.4 percent for those aged 20 to 29 (Butschek, 1974). Our estimation sample is dominated by this age group. In comparison, the rate for women aged 30 to 39 was only 50.9 percent (Butschek, 1974). This significant reduction was due to women leaving the labor force when they married or had their first child.

Female employment rates in Austria have been remarkably similar to the other OECD member countries (see Appendix Table A.1). Between the 1970s and the 2010s rates had increased from 47 percent to 67 percent. The equivalent median employment rates among OECD member

countries are 48 and 62 percent. While female employment in Austria was characterized by a high share of full-time employment in the 1970s (around 85 percent), it had gone down to only 55 percent in the 2010s. The median values among OECD member countries show a similar, yet less pronounced, pattern.

II.1.2. Maternity leave system and its reform in 1974

In 1957, Austria introduced a legislation which mandated 12 weeks of paid job-protected ML. This prohibited pregnant women from working 6 weeks before and 6 weeks after birth. The beginning of the prenatal ML was determined based on the doctor's estimation of the date of delivery. The prenatal ML could be started earlier if the mother's or the child's health was at risk due to the work environment. The latter had to be certified by either the chief medical officer of the *Regional Health Insurance Fund* or by an occupational physician of the *Labour Inspectorate*. The postnatal ML was regularly extended for all nursing mothers to 8 weeks and for nursing mothers with premature births to 12 weeks. Since 1962, *all* mothers experiencing a premature birth have a mandatory postnatal ML of 12 weeks. Since 1968, women who end up with a shorter prenatal ML due to an earlier (but not preterm) birth are compensated with a longer postnatal ML.[8] The last major reform of the ML system took place in 1974, which extended the compulsory ML duration to 16 weeks. Since then, pregnant women are prohibited from working 8 weeks before the delivery and usually 8 weeks after the delivery.

Assignment to the extended prenatal ML was determined by the cutoff date April 1, 1974 and was phased-in as follows: All pregnant women who commenced their prenatal leave on April 1, 1974 or later were assigned to up to 8 weeks, while all women who started their leave earlier were assigned to 6 weeks. Thus, the child's due date determined the number of prenatal leave days:

(1) Due date on May 13 or earlier: These mothers were assigned to 6 weeks of prenatal leave (42 days, between April 1 and May 13).

(2) Due date between May 14 to May 26: These mothers were assigned to prenatal leave durations that were gradually increased from 43 to 55 days.

(3) Due date on May 27 or later: These mothers were assigned to 8 weeks of prenatal leave (56 days between April 1 and May 27).

Panel A of Figure 1 depicts the relationship between the assignment to the extended prenatal ML and its actual length. We use the actual birth date as a proxy for the expected due date, since we cannot observe the latter (to be discussed in more detail below). The figure plots the average prenatal leave duration by birth date. Until the end of April we observe a constant

---

[8]For mothers with an earlier birth — but not a premature birth (fewer than 37 weeks of gestation) — the mandatory postnatal leave duration is extended. Note that for these mothers the maximum postnatal ML duration is less than 12 weeks.

mean of about 6.3 weeks (or 44.2 days). Throughout May we see a steady increase in the average prenatal leave duration, which reflects the stepwise increase as specified by the reform. Starting from the end of May, the reform was in full effect. For mothers who gave birth in June, we observe an average prenatal leave duration of 7.9 weeks (or 55.3 days). In our estimation analysis below we will focus on children born in April and June, which represent the groups of 'non-assigned' (N) and 'assigned' (A) mothers, respectively. We exclude mothers who gave birth in May from our analysis. Thus, we focus on the jump in the average prenatal ML duration from 6.3 to 7.9 weeks.

[Figure 1]

We now turn to postnatal leave regulations. The new rules applied to all mothers who gave birth on February 18, 1974 or later, since these mothers were still on ML by April 1. For women who gave birth earlier, the old regulations applied. Panel B of Figure 1 depicts the *postnatal* ML duration. Most importantly, all women who gave birth in April or June were assigned to the extended postnatal leave duration. We can see that the average duration is constant at about 8.8 weeks (or 61.5 days) starting from April. Thus, assigned and non-assigned mothers — while having differential average *prenatal* ML durations — do not differ in their *postnatal* ML durations. This feature of the reform allows us to cleanly identify the effect of variation in the prenatal ML duration also in the case of post-birth outcomes.[9]

During ML mothers receive a transfer payment that amounts to 100 percent of the average net earnings of the preceding 13 weeks (*Wochengeld*). Furthermore, they cannot be dismissed by their employer until 4 months after delivery. After ML most mothers were eligible for parental leave until the child's first birthday. The eligibility criteria for parental leave and the associated transfer payments did not differ for non-assigned and assigned mothers.

II.1.3. Public prenatal care

In the 1970s, infant mortality was comparably high in Austria, amounting to about 21 deaths of infants under the age of 1 per 1,000 live births (see Appendix Table A.1). This was above the median across countries and well above the figures for the U.S. and many other European countries. This is somewhat surprising, since Austria already had a Bismarckian welfare system in place, which provided almost universal access to high-quality healthcare.[10] In order to improve perinatal health outcomes, the *Austrian Federal Ministry of Health* launched the first

---

[9]Appendix Figure A.1 plots the average prenatal and postnatal ML duration for a wider window, ranging from January 1973 to December 1975. It shows that both durations had been constant before and after the reform. Note that the reform led to a comparably smaller increase in the postnatal ML duration. This can partly be explained by nursing mothers and mother with a premature birth, who were assigned to 8 and 12 weeks of postnatal ML already before the reform.

[10]Patients hold mandatory health insurance administered through 9 *Regional Health Insurance Funds* ("Gebiets-krankenkassen"), which cover private employees and their dependents, and 16 social security institutions that provide health insurance for specific occupational groups such as farmers, civil servants, and self-employed persons.

nationwide prenatal screening program in 1974. This so-called *Mother-Child-Pass Examination Program* (MCPEP) initially advocated pregnant mothers to participate in four prenatal screenings (in pregnancy weeks ≤ 16, 19, 27 and 37) and in one postnatal examination (in the first week after birth). Over time the aim and scope of the MCPEP has expanded substantially (Halla et al., 2016). Before the introduction of the MCPEP women could consult their gynaecologist for the same medical examinations. The essential feature of the MCPEP was the newly introduced financial incentive along with an information campaign. Mothers received 8,000 Austrian schillings (1,427.7 in 2018 euros) if they participated in at least one prenatal and the one postnatal examination. All mothers in our estimation sample were already exposed to the MCPEP and its financial incentives were offered equally to assigned and non-assigned mothers.[11]

II.2. Data

We construct our main data set by combining three administrative data sources. These are matched based on social security numbers and other unique identifiers for earlier years. The *Austrian Social Security Database* (ASSD) includes administrative records to verify pension claims and is structured as a matched employer–employee data set.[12] For each individual we observe on a daily basis where she is employed, along with her occupation, experience, and tenure. Information on earnings is provided per year and per employer. The limitations of the data are top-coded wages and the lack of information on (contracted) working hours (Zweimüller et al., 2009). We draw information from the ASSD to measure eligibility and the actual duration of prenatal and postnatal ML. The ASSD also allows us to construct outcome variables in the domains of subsequent fertility, human capital outcomes, and mortality. Furthermore, we use mothers' labor market histories to construct sample stratification variables.

The *Austrian Birth Register* (ABR) includes all live births in Austria with individual-level information on birth characteristics such as date, place of birth, birth weight, and birth length. This information is complemented by maternal socioeconomic characteristics such as age, marital status, occupation, and religious denomination. The ABR has two drawbacks. First, we do not observe information on parity for the early birth cohorts we consider. We will use subsamples of young versus older mothers to approximate a comparison between first births and higher-order parities. Second, the ABR does not provide information on the expected due date for births before 1984. We solve this problem by using the actual birth date as a proxy. This introduces some measurement error in our assignment variable. Fortunately, this affects only cases where the actual birth date deviates substantially from the expected due date (i.e., either

---

[11]The only difference which has to be noted is that assigned mothers were already in pregnancy week 19 at time of the introduction of the MCPEP. The first prenatal screening according to the MCPEP was already scheduled for week 16. Thus, it is possible that assigned mothers were more likely to participate in this first prenatal screening. There is no data on the actual participation rates in this prenatal screening available for this period.

[12]Note that we remove self-employed mothers from our analysis, since this group is exposed to somewhat different ML regulations. This exclusion does not affect the generalizability of our results, since less than 3 percent of all women in the 1974 labor force were self-employed.

very early or very late births). Using the distribution of gestational lengths from the year 1984, we can approximate the average measurement error in the assignment variable in our sample. It amounts to 0.6 days only. Appendix A.2 provides a detailed discussion of the source and nature of the measurement error. Below, we also a provide robustness test based on inverse measurement error weighting and a subsample with very little measurement error that confirms our findings.

Finally, we use information provided in the *Upper Austrian Sickness Fund* database to construct long-term health outcomes for mothers and their children. This database includes information on healthcare expenditures for all private employees and their dependents in Upper Austria starting in the year 1998. It covers roughly one million members representing 75 percent of the population in Upper Austria (see also footnote 10).

II.3. Estimation strategy

Our treatment variable is the actual prenatal ML duration in weeks M. Assignment into treatment, A, depends on the expected due date. We consider all eligible women who gave birth in June 1974 as assigned, $A_i = 1$, and those who gave birth in April 1974 as not assigned, $A_i = 0$. We disregard mothers who gave birth in May, when the reform was phased in and 28 cases of multiple births. While we have seen before that the relationship between assignment and treatment is strong, it is not fully deterministic. Hence, we set up a *fuzzy* RDD where assignment into treatment is used as an instrumental variable (IV) for the endogenous treatment variable. This design can be translated into a *two-stage least squares* (2SLS) setup with the following first stage estimation of the prenatal ML duration M_i:

$$M_i = \alpha_0 + \alpha_1 A_i + \mathbf{x}_i \boldsymbol{\gamma}' + \eta_i, \tag{1}$$

where $\mathbf{x}$ is a vector of control variables comprising information on the mother's age, citizenship, religious denomination, place of residence, the child's sex, and the child's legitimacy status; and η is a stochastic error term. In the second stage, we then use the exogenous variation $\widehat{M}$ to explore its effect on the respective outcome variable Y:

$$Y_i = \beta_0 + \varphi_{\text{rdd}} \cdot \widehat{M}_i + \mathbf{x}_i \boldsymbol{\delta}' + \varepsilon_i. \tag{2}$$

We estimate robust standard errors allowing for heteroskedasticity of unknown form. Alternative methods (analytical, clustered on the birth date level, and wild bootstrapped on the same cluster-level) yield very similar results.

II.3.1. Identifying assumptions

Four conditions need to hold for $\hat{\varphi}_{\text{rdd}}$ to be informative. First, assignment A must predict the actual prenatal ML duration M. Second, mothers must not precisely manipulate their

child's expected date of birth around the cutoff. Third, assignment must not be correlated with any outcome-determining factor. Fourth, for each mother, the prenatal ML duration must be longer (or equal) in the case of assignment as compared to the counterfactual situation of non-assignment. We now discuss these assumptions in detail.

The first condition is testable. We have already shown the distinctive jump in the prenatal ML duration at the cutoff (see Panel A of Figure 1). This condition also holds in our regression framework. Table 2 shows corresponding first stage estimates for the whole sample as well as for several subsamples of women.

[Table 2]

We obtain an $\hat{\alpha}_1$ of 1.586, implying that assignment increases the average prenatal ML duration by 1.6 weeks or 11 days. The estimated coefficient is highly statistically significant with an F-statistics of about 757. There are two sources of non-compliance, which explain why the first stage coefficient is below the 2 weeks extension of mandatory prenatal ML. First, the actual birth date may deviate from the expected due date, which shortens or prolongs the actual prenatal ML duration.[13] Second, as described above, the prenatal ML can start earlier if the mother's or the child's health is at risk. This leads to one-sided non-compliance. We see that non-assigned women had an average prenatal ML duration of 6.3 weeks, while assigned women were on average 7.9 weeks on prenatal ML. This indicates that before the reform, women were more likely to start their ML early due to medical reasons. This has a mechanical component: the shorter the mandatory prenatal ML duration is, the higher the likelihood of medical complications before the start of ML. Our data from 1974 do not allow us to distinguish between the two sources of non-compliance, since we do not observe gestational length. Utilizing data from the year 1984, we find that around 15 percent of all mothers started their ML early due to medical reasons.

Table 2 shows first stage estimates across subsamples. The coefficients vary somewhat between different groups of mothers. The largest difference occurs between very young (< 21) and older mothers (≥ 29), for whom we obtain coefficients of 1.33 and 1.72, respectively. For mothers with a low and high income, we estimate first stage coefficients of 1.49 and 1.68, respectively. The first stage coefficient is 1.68 for white collar workers and 1.52 for women in a blue collar occupation. This difference is statistically not significant. Overall, while our first stage estimates are strong and quite stable for all groups of mothers, we conclude that the compliance rate tends to be stronger for women with more favorable characteristics. Despite these differences in compliance, our heterogeneity analysis below reveals no significantly different treatment effects across groups.

---

[13]Note that most children are born between the 39th to 41st week of pregnancy. In 1984 (the first year we have information on gestational length in the birth register), 79.5 percent are born during this three-week period, while only 4.6 percent are born later and around 16 percent are born earlier. Thus, even if women start their ML on time, the probability of a shorter leave is larger as compared to a longer leave.

The inability to precisely manipulate assignment into treatment is the key identifying assumption behind any RDD. Public discussion about the potential reform of the ML system started in December 1973. The earliest media coverage we found is a newspaper article published on December 13, 1973. This reports that the Socialist-led government plans to extend maternity leave without providing any details.[14] The bill was submitted on February 5, 1974. The legislative proposal underwent a preliminary deliberation by the *Committee on Social Affairs* of the *National Council* on February 22, 1974. The bill was then passed by the National Council on March 6, 1974 and approved by the Federal Council on March 14, 1974. It became effective on April 1, 1974.[15] This timing rules out that parents adjusted their conception behavior. If more women got pregnant due to the reform, those children would had been born in September 1974 the earliest. Furthermore, the manipulation of the due date of an ongoing pregnancy is unlikely. The expected due date is determined by the doctor at the beginning of pregnancy. The manipulation of delivery dates is also unlikely, since there was no real incentive to do this.[16]

[Figure 2]

Figure 2 shows the average number of births per day around the cutoff date. There is no evidence of manipulation. More formally, the manipulation test by Frandsen (2017) for discrete running variables confirms this. We cannot reject the null hypothesis that there is no shift in the discontinuity at the birthday cutoff.[17] Thus, there is no evidence of manipulations in birth dates.

Whether the assignment is correlated with any outcome-determining factor is not testable. Importantly, we are not aware of any change or reform (except for MCPEP, see footnote 11) that coincides with the timing of the reform under consideration. It is reassuring that none of our covariates change discontinuously around the cutoff. Figure 3 plots the daily averages of all covariates and other pre-determined variables between January and September 1974. In line with this, we observe essentially unchanged results if we use a specification without any covariates (see Appendix Table A.2).

[Figure 3]

---

[14]We have scanned four major newspapers (*Neue Kronen Zeitung*, *Die Presse*, *Salzburger Nachrichten*, *Oberösterreichische Nachrichten*) in the period from November 1973 through March 1974 for all articles discussing maternity leave. We found a total of five articles.

[15]The signed law was published in the Federal Law Gazette on March 29, 1974 (see *Bundesgesetzblatt* 59/1974).

[16]If there is any potential manipulation here, women might have tried to deliver their child a bit earlier. A later delivery is not possible anyways. Theoretically, women might want to get their child before the due date to take advantage of a longer postnatal leave (e.g., if the child was delivered 1 week earlier, women were mandated a postnatal leave of 1 additional week). Since the Austrian system allows women to take parental leave subsequent to ML, this incentive should be negligible. While we have no data on the type of delivery for the year 1974, we perform a sensitivity check and control for the day of the week of the delivery in our RDD estimations. This partly controls for planned cesarian sections since these are normally not scheduled on weekends. The results are robust.

[17]The test requires picking a parameter k for the maximal degree of nonlinearity in the birth date probability mass function still considered to be compatible with no manipulation. Following Frandsen (2017), we perform the test for $k = 0$, $k = 0.01$, and $k = 0.02$; for which we obtain p-values of 0.729, 0.741, and 0.744.

In our setting, the monotonicity assumption implies that there is no mother who would have a shorter prenatal ML duration if assigned, but a longer duration if not being assigned: $M_i(A_i = 1) \geq M_i(A_i = 0)$. Given that ML is mandatory in Austria, we regard a violation of this assumption as unlikely.

Finally, it is instructive to ask the question whether our research design has sufficient statistical power to rule out meaningful effects. Consider the outcome birth weight. We obtain here a minimum detectable effect size (MDES) of only 16 grams per one additional week of prenatal ML. Appendix A.3 provides details for this calculation. To put this quantity in perspective, an increase in the birth weight of an average newborn by 16 grams is equivalent to an increase by only 0.4 percent or 0.03 standard deviations. This effect size seems (medically) irrelevant. Put differently, 'relevant' reform effects which would pass a cost-benefit threshold are well-identified by our research design.

II.3.2. Non-working mothers, an additional control group

To check the robustness of our results, we use information on unaffected non-working mothers.[18] While these mothers clearly differ (in their observable characteristics) from working mothers, they are useful since they were never eligible for ML.[19] The reform had by definition no impact on non-working mothers, hence they serve as an additional control group. This second source of variation can either be used to complement or substitute our RDD approach. In the case where we use non-working mothers to extend our RDD approach, we gain a difference-in-differences (DiD) component. This differences out any potential seasonal effects between children born in April and June.[20]

To translate this combined approach into a regression framework, we extend our first stage estimation with a binary variable, W, capturing the mother's employment status at the time of birth, and an interaction between this variable and the assignment variable A:

$$M_i = \theta_0 + \theta_1 A_i + \theta_2 W_i + \theta_3 (A_i \times W_i) + \mathbf{x}_i \boldsymbol{\zeta}' + u_i, \tag{3}$$

where the latter is again equal to one for all women who gave birth in June, irrespective of their employment status, and zero otherwise. For non-working mothers the ML reform did not affect allocation of time. We impute $M_i = 40$ if i was not working at time of birth. The specific value chosen has no impact on the estimation results. Instead of using the assignment variable A as an exclusion restriction, in this approach we use the interaction term $A \times W$ to identify the effect

---

[18]Working mothers who gave birth in April and June 1973 could in principle serve as an alternative control group. Unfortunately, we have incomplete information to link observations across time for cohorts before 1974.

[19]Non-working mothers were on average 2.8 years older at the time of birth and more likely to be married and Catholic. Appendix Table A.3 provides further descriptive statistics for working and non-working mothers, who gave birth in April or June 1974.

[20]There is evidence that birth month is related to socioeconomic characteristics of mothers and later educational attainment of children. See, for example, Buckles and Hungerman (2013) or Schneeweis and Zweimüller (2014).

of the ML extension on the respective outcome Y in the second stage estimation:

$$Y_i = \rho_0 + \varphi_{\text{rdd-did}} \cdot \widehat{M}_i + \rho_1 A_i + \rho_2 W_i + \mathbf{x}_i \iota' + v_i, \tag{4}$$

where our alternative treatment effect of interest is $\hat{\varphi}_{\text{rdd-did}}$.

In the case of using this approach to substitute our RDD analysis, we identify the effects of the reform solely based on the DiD component (thus, we do no exploit the RDD in a 2SLS setup). Now the identification strategy is identical to a simple DiD approach,

$$Y_i = \lambda_0 + \lambda_1 A_i + \lambda_2 W_i + \varphi_{\text{did}} \cdot (A_i \times W_i) + \mathbf{x}_i \boldsymbol{\omega}' + w_i, \tag{5}$$

where the treatment effect of interest is equal to $\hat{\varphi}_{\text{did}}$. The identifying assumption is that the trends in the outcome variables would have been the same for these two groups of mothers (working and non-working) in the absence of the reform. While this assumption is untestable, we are confident that it holds, since we do not observe any differences in pre-trends (see Appendix Figure A.2.)

II.3.3. Outcome variables

We consider various short- and long-term outcomes for both the child and the mother. We are interested in health at birth outcomes, as well as long-term labor market and health outcomes of the child. To infer on effects on the mother, we examine her subsequent fertility and health outcomes. In Table 3, we provide descriptive statistics for our outcome variables and covariates, separately for assigned and non-assigned mothers.

[Table 3]

Health at birth   To construct health at birth outcomes we use information on birth weight and length. We consider both as continuous variables measured in logs. Additionally we construct a binary variable indicating low birth weight. This is equal to one if birth weight is lower than 2,500 grams, and zero else. Average birth weight in the sample is 3,256 grams, and roughly 6% of all children had low birth weight. According to the medical literature, intrauterine growth retardation can either start early or late in pregnancy leading to symmetrically or asymmetrically growth restricted newborns, respectively.[21] Since the ML reform affected pregnant women in their 33rd and 34th week of gestation, we examine the asymmetric growth restriction. We

---

[21]Symmetric growth restricted fetuses have a proportionally small body, with small weight and length. The causes are genetic factors, maternal diseases, infections or other toxic environmental effects occurring in early gestation. Asymmetric growth restriction is associated with small weight but normal length and is typically caused by risk factors in the last phase of gestation (after week 32). Common risk factors are poor maternal nutrition, placental insufficiency, preeclampsia or chronic hypertension in late pregnancy (Lin and Santolaya-Forgas, 1998; Valsamakis et al., 2006). Approximately 70-80 percent of growth restricted newborns can be classified as asymmetrically growth restricted (Lin and Santolaya-Forgas, 1998).

construct a binary variable combining the birth weight and the child's *Ponderal* index, $PI = kg/m^3$ (Landmann et al., 2006). We define growth as asymmetrically restricted if birth weight is lower than 2,500 grams and the Ponderal index is in the lowest quartile of its sample distribution. In our sample, about 4% of children have an asymmetric growth restriction according to this definition. Finally, we study premature births. The information on pregnancy duration had not been recorded in the ABR until 1984. We exploit the regulation which mandates a postnatal ML duration of 12 weeks for mothers who experienced a premature birth. This regulation did not change with the reform under consideration. Thus, we are able to identify preterm births in the ASSD. We find 6% of all births to be premature in our data.[22]

Children's long-term outcomes   We consider children's long-term labor market and health outcomes. About one-third of all children in our sample can be uniquely matched to their mother in the ASSD and can be included in our 2SLS estimations.[23] Fortunately, the availability of the data link seems to be idiosyncratic. It is not correlated with our IV and should therefore not bias our results: The share of children for whom we have information on labor market outcomes is similar for assigned (32.7%) and non-assigned (32.4%) mothers. Since we observe the assignment status A_i and long-term outcomes for all children, we can provide a reduced form estimate based on the full sample. We analyze employment, occupation, and wages between 25 and 40 years of age.

Moreover, using the database from the *Upper Austrian Sickness Fund*, we are able to construct long-term health outcomes for all children employed in the private sector in Upper Austria in the period between 1999 and 2014. Over this period (during which children were between 25 and 40 years of age) we aggregate health care spending in the outpatient sector and the days spent in hospital for the 511 children remaining in our sample.

Maternal outcomes   For mothers our outcome variables include measures of subsequent fertility and health. In the former domain we consider a potential effect of the reform on the tempo and quantum of fertility. To capture the tempo we consider the time until the mother's next birth in logs. About half of the mothers had at least one further birth. The average duration until the next birth was 4.36 years. We employ two measures for completed fertility. First, we use a binary variable indicating whether the mother gave birth at least once more, as well as the number of subsequent births. Furthermore, we study mortality to capture effects on the mother's health. We construct two binary variables indicating whether the mother survived at least 20 and 40 years after giving birth. The average survival rates at these two points in time are 99% and 92%, respectively. In the same manner as for children, we aggregate health care spending

---

[22]Note that women with multiple births have also 12 weeks of mandatory postnatal ML. We dropped these mothers from the analysis.

[23]Two-thirds of children cannot be uniquely matched to their mother in the ASSD. This link in the administrative data was not comprehensively available for early cohorts. For these children we do not observe their mothers' employment status W_i nor their treatment M_i..

in the outpatient sector and the days spent in hospital for a subsample of 1,117 mothers, who were insured with the *Upper Austrian Sickness Fund* between 1999 and 2014.

III. Estimation results

III.1. Health at birth outcomes

We present our main estimation results in Table 4. In Panel A, we summarize our RDD estimates $\hat{\varphi}_{\text{rdd}}$. Each coefficient is obtained by estimating the fuzzy RDD outlined in section II.3 via 2SLS. In columns (1) to (5), we consider the birth weight in logs, a low birth weight indicator, an indicator for an asymmetric growth restriction, birth length in logs, and an indicator for a premature birth as outcome variables. Across outcomes we find no statistically significant effects of the reform. Interestingly, for a majority of the estimates, the signs point to an adverse effect of the reform. However, most estimates are very close to zero and precisely estimated.

The point estimate on log birth weight suggests a 0.4 percent reduction for a one-week extension of prenatal ML. This change in birth weight is tiny and its potential effect on later life outcomes is negligible. Black et al. (2007) exploit variation in birth weight across twins in Norway and show that a one percent increase in birth weight increases earnings by only 0.12 percent. Likewise, Figlio et al. (2014) estimate an associated increase in test scores of only 0.4 percent of its standard deviation for twins in Florida.

To further put our point estimates into perspective, we compare them to socioeconomic gradients in birth outcomes. Consider the estimate on low birth weight. We obtain a positive but tiny coefficient that amounts to a 0.03 percentage points increase for a one-week extension of prenatal ML. In comparison, the corresponding gradient between mothers with high and low income is 2 percentage points. Details are provided in Appendix Table A.4. The same gradient is observed between married and unmarried mothers. Thus, our coefficient amounts to only 1.5 percent of these gradients. The coefficient for premature birth is somewhat larger and amounts to 0.5 percentage points for a one-week extension of prenatal ML. The respective socioeconomic gradients with respect to income and marital status are 4 and 5 percentage points. Our estimated coefficient amounts to 12.5 percent and 10 percent of these gradients. For this outcome, we cannot rule out meaningful effects.

Our findings are not sensitive to the sample choice with respect to the window around the cutoff date (see Appendix Figure A.3). We provide two types of robustness checks with respect to the measurement error in our assignment variable. First, we weight our observations to account for the measurement error (see Appendix A.2 for details). Panel B of Table A.5 summarizes results based on inverse measurement error weighting. Second, we restrict the analysis to births in the first half of April and the second half of June (see Panels B to G in Appendix Table A.5). In this subsample, the measurement error in our assignment variable is negligible. Across methods we find very similar results as compared to our baseline estimates.

[Table 4]

For comparison we provide naïve OLS estimates in Panel B of Table 4. Assuming the prenatal ML duration be exogenous, we simply regress each outcome on the mother's actual leave duration in weeks. To avoid capturing the reform's effect with our OLS estimates, we restrict the sample to births in the pre-treatment period. The OLS estimates comprise a potential causal effect as well as different sources of endogeneity. Regarding the latter, it is useful to distinguish between an early start date and a late end date of ML. An early start may either reflect a curative intervention due to diagnosed health problems in pregnancy, or a preventative intervention by risk-averse mothers and doctors. The former would lead to a negative association between ML duration and health at birth outcomes, while the effect of the latter is ambiguous. Assuming that health-conscious mothers have better health outcomes, the association would be positive. A late end date captures a mechanical relationship between a longer ML and a longer gestation. Since a longer gestation is associated with higher birth weight and higher length, this mechanical component contributes to a positive association between ML duration and health at birth.

We find statistically significant positive OLS estimates for all outcome variables except the asymmetric growth restriction. The source of the positive correlation between longer ML and health at birth is unclear and most likely reflects the dominance of the mechanical component. The OLS estimates should not be interpreted causally.

III.1.1. Interpretation of estimation results

When interpreting these results, we have to keep in mind that we cannot observe how assigned mothers in fact spend the additional free time obtained through the ML extension. This is not unique to our research design. Every paper analyzing ML (or parental leave) reforms shares the feature that estimated effects have a reduced-form character in this respect. It is informative to interpret the results in regards to the Austrian institutional setting. In our specific case of a *mandatory* leave, the problem is alleviated to some degree, since we should have mainly one-sided non-compliance (to borrow the language of RCTs). We expect assigned mothers to be compliant, in the sense that they do not work in the 33rd and 34th week of gestation. In contrast, non-assigned mothers may be non-compliant and do not work. As discussed above, the institutional setting offers the possibility of early ML whenever the mother's or the child's health is at risk due to work. This source of non-compliance is captured by our estimation strategy and does not complicate the interpretation of our results.

However, expecting mothers — as any other employee — are always entitled to sick leave if supported by a medical certificate. This source of non-compliance is not captured by our estimation strategy, and potentially complicates the interpretation of our results. In an extreme scenario, a non-assigned mother could go on sick leave in her 33rd week of pregnancy until her prenatal ML starts. Thus, she would *de facto* stop working 8 weeks before her due date, just as an assigned mother. Fortunately, we can assess the importance of sick leaves prior to prenatal

ML for all blue-collar workers.[24] It turns out that only 5.1% of all non-assigned mothers were on sick leave prior to their ML. Most importantly, this share was comparable for assigned mothers (4.2%). The difference between these two shares is not statistically significant ($p = 0.1684$, $n = 3{,}962$).

This clarifies the interpretation of our results. First, our LATE is most likely driven by counterfactual comparisons of mothers without major problems in this stage of pregnancy. Second, treated mothers were (in contrast to non-treated mothers) indeed *not* exposed to work in the 33rd and 34th week of gestation. Thus, our LATE captures a reduction of *in utero* exposure to maternal employment in the 33rd and 34th week of gestation for a group of mothers without major problems in pregnancy.

III.1.2. Treatment effect heterogeneity

We stratify our sample according to different characteristics of mothers and repeat our estimation analysis for each subsample. We distinguish mothers by occupational collar, age, and earnings. The occupational collar (blue versus white collar) is highly correlated with the job task (manual labor versus office work). One might expect women performing manual labor to benefit more from an increase in prenatal leave duration. The stratification by age (less than 21 years, between 21 and 28, and 29 years and older) is not only interesting *per se*, but also allows us to infer on parity to some extent. In the subsample of the youngest mothers, the vast majority of cases are presumably first births. A sample split by earnings (below versus above the sample median) considers more general differences between socioeconomic backgrounds.

[Figure 4]

Figure 4 graphically summarizes our RDD estimates for these subsamples, along with our baseline estimates. We focus here on three outcome variables (birth weight, length, and premature birth). The general finding is a zero effect in each stratum. The same holds for all other outcome variables (see Panel A of Appendix Figure A.4). The estimated treatment effects are (with one exception)[25] all statistical insignificant. The widths of the 95% confidence intervals vary somewhat and are, as expected, larger for smaller subsamples. We conclude that the reform had no effects on birth outcomes, not even for children born to more vulnerable women, or to those exposed to more exhausting working conditions.[26]

---

[24]Sick workers receive their compensation from two sources: First, workers continue to receive their salaries from firms. Second, after a certain period, they receive also public sickness benefits. The ASSD contains only information on sick leaves once the public sickness benefits are paid (Halla et al., 2015). Until September 1974, blue-collar workers received public sickness benefits already after 5 days of sick leave. This allows us to observe their sick leave with little error. In contrast, white-collar workers received public sickness benefits, depending on their tenure, only after several weeks. Thus, for white-collar workers we observe only long sick leave spells.

[25]For older mothers, we obtain a significant positive effect on premature birth.

[26]We have also estimated quantile treatment effects (QTEs) for our continuous outcome variables. Across outcomes we do not find any evidence for significant QTEs (see Appendix Figure A.5).

III.1.3. Robustness checks using non-working mothers as a control group

To check robustness of our results on health at birth, we augment our fuzzy RDD model from above with a DiD component. As in equation (4), we introduce non-working mothers, who were unaffected by the reform, as a control group.[27] This allows us to subtract the estimated RDD effect for non-working mothers from the equivalent effect for working mothers, which will difference out any seasonal effects. The means of our observables are very similar across the two subsamples (see Appendix Table A.3). Panel C of Table 4 summarizes estimation results from this alternative specification. The estimated effects are similar to those obtained by our RDD estimations above (see Panel A). All point estimates are precisely estimated and statistically insignificant.[28]

Panel D of Table 4 presents estimation results from a simple DiD specified in equation (5), where we compare pre- and post-reform effects of working and non-working mothers. This specification shows also economical and statistical zero-effects across all outcomes.[29]

III.2. Children's long-term outcomes

In a next step, we examine children's long-term outcomes. Despite the lack of evidence that the reform altered children's birth outcomes, one should not jump to the conclusion that the reform had no impact on children at all. Several medical and epidemiological studies postulate the idea that events *in utero* might indeed affect the infant, but these effects remain latent for many years (Almond and Currie, 2011*b*). Panel A of Table 5 summarizes our RDD estimates for long-term human capital and health outcomes. We mean-centered all continuous outcome variables measured in levels. In columns (1) to (3), we consider labor market outcomes at the age of 40. At this point in time, about 84% were in a regular employment, and, among those, about 70% were employed as a white-collar worker. Average daily wages amounted to € 119. In the remaining columns we consider effects on health. In particular, we examine aggregate spending in the outpatient sector (column 4, in € 1,000), and the aggregate days spent in hospital (column 5). Both variables refer to the period in which children were between 25 and 40 years of age. We observe an average spending in the outpatient sector of about € 1,830 and an average of 9.2 days spent in hospital.

[Table 5]

The analysis of long-term outcomes confirms our conclusion derived from the analysis of birth

---

[27]Note that we cannot perform this analysis on the premature birth outcome, since this variable is only observed for working mothers.

[28]The estimated first stage coefficient from equation (3), $\hat{\theta}_3$, is 1.586 (0.058). The Kleinbergen-Paap *rk* Wald *F*-statistic is 756.6, and the partial r^2 is 0.093.

[29]Note that the DiD estimator captures the average treatment effect of the reform, which extended compulsory prenatal ML duration by two weeks. In contrast, the LATE captures the effect of one additional week on prenatal ML due to assignment. To ensure arithmetic comparability of these two estimates, the DiD estimate has to be divided by two.

outcomes. Across outcomes, we do not observe any statistically significant effects of the reform. This applies to human capital outcomes as well as to health outcomes. For labor market outcomes our point estimates are close to zero. For health outcomes, we have fewer observations and our estimates are rather imprecise. Thus, we cannot rule out meaningful effects for the latter. To assess the robustness of these findings we also examine labor market outcomes at various stages over the life cycle. Figure 5 summarizes RDD estimates for our three outcome variables at the ages of 25, 30, 35 and 40. Again, neither estimate is economically or statistically significant.

[Figure 5]

As explained in Section II.3.3 (see footnote 23), we have fewer observations available for children's long-term outcomes due to a missing data link. Although our estimates are derived from sufficiently strong first stages (see Panel A of Table 5), we additionally provide reduced form estimates based on the full population of children in the ASSD. In Panel C of Table 5 we compare children born in April 1974 with those born in June 1974. For the analysis of labor market outcomes, we have now at least 11,000 observations. For our health outcomes, the number of observations has increased to almost 3,300. Across outcomes we find statistically insignificant reduced form estimates. For comparison, Panel B provides the reduced form estimates for our smaller sample of children who can be linked to the ABR. These reduced form estimates are also statistically insignificant. This analysis confirms our findings from Panel A.[30]

III.3. Maternal outcomes

So far, we have provided evidence that the ML extension had no significant effects on children's outcomes, neither at the time of birth nor in the long-run. It is still possible that extended ML, while having no discernible effects on children, has altered the physiological or psychological well-being of mothers. For example, a potential reduction in maternal stress prior to birth might have changed pregnancy and birth experiences for mothers, which in turn may have reduced the number of pregnancy complications, obstetric labor complications, or health problems in the postpartum period. Since we do not observe maternal health at the time of birth, we focus on long-term outcomes. We examine two informative outcome dimensions. First, we consider mothers' subsequent fertility behavior. If the extended leave has improved well-being of mothers, we would see an increased quantum or tempo of fertility. Mothers may either be more willing or more able to conceive and deliver a further child. On the other hand, if extended leave has reduced maternal well-being, subsequent fertility might have decreased. Accordingly, we estimate the effect of the reform on completed fertility, and — conditional on having another

---

[30]We also run our health at birth and maternal outcomes regressions based on a sample of children for whom we observe at least one of the long-term outcomes from Table 5. The results are in line with our previous findings (see Appendix Table A.6).

birth — on the duration until the next birth. Second, we examine mothers' long-term health. We examine mortality for the full sample of mothers, and health-care utilization in the inpatient and outpatient sector for the subsample of mothers in Upper Austria.

[Table 6]

Panel A of Table 6 summarizes our RDD estimates for these maternal outcomes. Panel B provides OLS estimates for comparison, where the sample is again restricted to the pre-treatment period. In columns (1) to (3), we focus on fertility. About half of *all* mothers have at least one further birth, with an average duration to their next birth of about 4.36 years. The average total number of subsequent births is 0.7. Across columns, we do not find evidence for any significant effects of prenatal ML on subsequent maternal fertility. The point estimates are neither statistically nor economically significant. In columns (4) to (5), we examine mortality. For 20-year survival, we find a clear zero-effect. For 40-year survival, we obtain a marginally significant negative effect. This suggests that an additional week of ML decreases the probability of being alive after 40 years by 0.7 percentage points, corresponding to a reduction of roughly 0.76% of the sample mean. It should be emphasized that the estimated effect is only significant at the 10 percent level. Finally, in columns (6) and (7), we present estimates for outpatient expenses and hospital days (both variables aggregated for the time period of 25 to 40 years after birth). Again, none of the coefficients is statistically or economically significant.

[Figure 6]

In Figure 6 (and Panel B of Appendix Figure A.4), we additionally study whether certain socioeconomic groups respond differently to the reform. Again we stratify mothers by their occupational collar, age, and income. Most of these subgroups resemble the baseline, with coefficients being close to zero and insignificant. In terms of fertility, we find that low income mothers experience a small increase in subsequent fertility due to the reform. Overall, the ML reform in 1974 had no substantial impact on maternal outcomes.

IV. Conclusions

We have analyzed the impact of a reform of the Austrian ML legislation in the year 1974, which has increased the compulsory prenatal ML duration from 6 to 8 weeks. The political justification for this reform was to improve the health of pregnant workers and their unborn children. Extended leave was determined by a cutoff date, which allows us to implement an RDD. Since assigned and non-assigned mothers, while experiencing different durations of prenatal ML were both exposed to the same postnatal ML, we are able to study not only potential effects on birth outcomes, but also on post-birth outcomes. The reform took place in 1974, this allows us to study long-term outcomes for children and mothers of up to 40 years after birth. Furthermore, in

our administrative data, we observe the actual duration of ML. We are able to not only estimate an ITT but also a LATE. Our LATE captures a reduction of the *in utero* exposure to maternal employment in the 33rd and 34th week of gestation for a group of mothers without major problems in pregnancy. We find no evidence for a significant effect on children's health at birth or on their long-term health and human capital outcomes. Subsequent maternal health and fertility also remain unaffected. The estimated treatment effects are statistically insignificant and precisely estimated zero effects. This finding is broadly consistent across subsamples of mothers. The reform has increased public spending on transfer payments for prenatal ML by one-third. In the first calendar year with full implementation (1975), the additional transfers due to reform amounted to approximately € 2.7 million (in 2018 Euros).[31] The reform has restricted female workers in their freedom to work without producing any measurable benefits.

While our findings must be interpreted taking into consideration the prevailing Bismarckian healthcare system in Austria, we conclude more generally that mandatory prenatal ML starting in the 35th week of gestation is sufficient for pregnant workers without major problems in pregnancy. It should be emphasized that we do not interpret our results as a general argument against (mandatory) prenatal ML. Quite the contrary, we consider our findings to be valuable for designing an optimal prenatal ML policy in places without comprehensive legislation in place. Finally, we suggest reassessing existing ML legislations with long compulsory durations and reducing either the extent of obligation or the duration. This is relevant for many OECD countries, since their ML legislation is comparable to the Austrian case.

---

[31] See government bill RV 1033 BlgNR 13. GP

REFERENCES

Almond, Douglas and Janet Currie (2011*a*), Human Capital Development Before Age Five, *in* O.Ashenfelter and D.Card, eds, 'Handbook of Labor Economics, Vol. 4b, Ch. 15', North Holland, pp. 1315–1486.

Almond, Douglas and Janet Currie (2011*b*), 'Killing Me Softly: The Fetal Origins Hypothesis', *Journal of Economic Perspectives* **25**(3), 153–72.

Baker, Michael and Kevin Milligan (2010), 'Evidence from Maternity Leave Expansions of the Impact of Maternal Care on Early Child Development', *Journal of Human Resources* **45**, 1–32.

Baker, Michael and Kevin Milligan (2015), 'Maternity Leave and Children's Cognitive and Behavioral Development', *Journal of Population Economics* **28**, 373–391.

Bharadwaj, Prashant, Julian V. Johnsen and Katrine V. Løken (2014), 'Smoking Bans, Maternal Smoking and Birth Outcomes', *Journal of Public Economics* **115**, 72–93.

Black, Sandra E., Paul J. Devereux and Kjell G. Salvanes (2007), 'From the Cradle to the Labor Market? The Effect of Birth Weight on Adult Outcomes', *Quarterly Journal of Economics* **122**(1), 409–439.

Black, Sandra E., Paul J. Devereux and Kjell G. Salvanes (2016), 'Does Grief Transfer across Generations? Bereavements during Pregnancy and Child Outcomes', *American Economic Journal: Applied Economics* **8**(1), 193–223.

Bonzini, M, K T. Palmer, D Coggon, M Carugno, A Cromi and M M. Ferrario (2011), 'Shift Work and Pregnancy Outcomes: A Systematic Review with Meta-Analysis of Currently Available Epidemiological Studies', *BJOG An International Journal of Obstetrics and Gynaecology* **118**, 1429–1437.

Buckles, Kasey S. and Daniel M. Hungerman (2013), 'Season of Birth and Lator Outcomes: Old Questions, New Answers', *Review of Economics and Statistics* **95**, 711–724.

Butschek, Felix (1974), 'Erwerbstätigkeit in Österreich 1961 bis 1980', *WIFO-Monatsberichte* **47**(10), 479–490.

Camacho, Adriana (2008), 'Stress and Birth Weight: Evidence from Terrorist Attacks', *American Economic Review: Papers and Proceedings* **98**(2), 511–515.

Carneiro, Pedro, Katrine V. Løken and Kjell G. Salvanes (2015), 'A Flying Start? Maternity Leave Benefits and Long Run Outcomes of Children', *Journal of Political Economy* **123**, 365–412.

Chen, Dafang, Sung-Il Cho, Changzhong Chen, Xiaobin Wang, Andrew I. Damokosh, Louise Ryan, Thomas J. Smith, David C. Christiani and Xiping Xu (2000), 'Exposure to Benzene, Occupational Stress, and Reduced Birth Weight', *Occupational and Environmental Medicine* **57**, 661–667.

Chuard, Caroline (2018), Womb at Work: The Missing Impact of Maternal Employment on Newborn Health, Working Paper 301, University of Zurich.

23

Currie, Janet and Maya Rossin-Slater (2013), 'Weathering the Storm: Hurricanes and Birth Outcomes', *Journal of Health Economics* **32**(3), 487–503.

Dahl, Gordon, Katrine V. Løken, Magne Mogstad and Kari Salvanes (2016), 'What Is the Case for Paid Maternity Leave?', *Review of Economics and Statistics* **98**(4), 655–670.

Danzer, Natalia, Martin Halla, Nicole Schneeweis and Martina Zweimüller (2017), Parental Leave, (In)formal Childcare and Long-term Child Outcomes , IZA Discussion Paper 10812, Institute for the Study of Labor, Bonn, Germany.

Deke, John and Lisa Dragoset (2012), Statistical Power for Regression Discontinuity Designs in Education: Empirical Estimates of Design Effects Relative to Randomized Controlled Trials, Working paper, Mathematica Policy Research.

Del Bono, Emilia, John Ermisch and Marco Francesconi (2012), 'Intrafamily Resource Allocations: A Dynamic Structural Model of Birth Weight', *Journal of Labour Economics* **30**(3), 657–706.

Dustmann, Christian and Uta Schönberg (2012), 'Expansions in Maternity Leave Coverage and Children's Long-Term Outcomes', *American Economic Journal: Applied Economics* **4**(3), 190–224.

Eurofound (2015), Maternity Leave Provisions in the EU Member States: Duration and Allowances, Publications Office of the European Union, European Union, Luxembourg.

Evans, William N. and Diana S. Lien (2005), 'The Benefits of Prenatal Care: Evidence from the PAT Bus Strike', *Journal of Econometrics* **125**(1), 207–239.

Figlio, David, Jonathan Guryan, Krzysztof Karbownik and Jeffrey Roth (2014), 'The Effects of Poor Neonatal Health on Children's Cognitive Development', *American Economic Review* **104**(12), 3921–3955.

Frandsen, Brigham R. (2017), Party Bias in Union Representation Elections: Testing for Manipulation in the Regression Discontinuity Design when the Running Variable is Discrete, *in* M. D.Cattaneo and J. C.Escanciano, eds, 'Regression Discontinuity Designs: Theory and Applications', Advances in Econometrics, Emerald Publishing, pp. 281–315.

Halla, Martin, Gerald J. Pruckner and Thomas Schober (2016), 'Cost Savings of Developmental Screenings: Evidence from a Nationwide Program', *Journal of Health Economics* **49**, 120–135.

Halla, Martin, Susanne Pech and Martina Zweimüller (2015), The Effect of Statutory Sick Pay Regulations on Workers' Health, IZA Discussion Paper 9091, Institute for the Study of Labor (IZA), Bonn, Germany.

Landmann, Eva, Irwin Reiss, Björn Misselwitz and Ludwig Gortner (2006), 'Ponderal Index for Discrimination between Symmetric and Asymmetric Growth Restriction: Percentiles for Neonates from 30 Weeks to 43 Weeks of Gestation', *The Journal of Maternal-Fetal and Neonatal Medicine* **19**(3), 157–160.

Lawson, Christina C., Carissa M. Rocheleau, Elizabeth A. Whelan, Eileen N. Lividoti Hibert, Barbara Grajewski, Donna Spiegelman and Janet W. Rich-Edwards (2012), ' Occupational Exposures Among Nurses and Risk of Spontaneous Abortion', *American Journal of Obstetrics & Gynecology* **206**(4), 327.e1–327.e8.

Lin, Chin-Chu and Joaquin Santolaya-Forgas (1998), 'Current Concepts of Fetal Growth Restriction: Part I. Causes, Classification, and Pathophysiology', *Obstetrics & Gynecology* **92**(6), 1044–1055.

Liu, Qian and Oskar Nordström Skans (2010), 'The Duration of Paid Parental Leave and Children's Scholastic Performance', *B. E. Journal of Economic Analysis and Policy (Contributions)* **10**(1), Article 3.

Mansour, Hani and Daniel I. Rees (2012), 'The Effect of Prenatal Stress on Birth Weight: Evidence from the Al-Aqsa Intifada', *Journal of Development Economics* **99**(1), 190–199.

Mozurkewich, Ellen L., Barbara Luke and Fredric M. Wolf (1999), 'Working Conditions and Pregnancy Outcome', *Obstetrics & Gynecology* **95**(4).

OECD (2017), OECD Family Database, Database, Organization for Economic Cooperation and Development, Paris.

Palmer, Keith T., Matteo Bonzini, E Clare Harris, Cathy Linaker and Jens Peter Bonde (2013), 'Work Activities and Risk of Prematurity, Low Birthweight and Pre-eclampsia: An Updated Review with Meta-Analysis', *Occupational and Environmental Medicine* **70**(4), 213–222.

Persson, Petra and Maya Rossin-Slater (2018), 'Family Ruptures, Stress, and the Mental Health of the Next Generation', *American Economic Review* **108**(4–5), 1214–1252.

Rasmussen, Astrid (2010), 'Increasing the Length of Parents' Birth-Related Leave: The Effect on Children's Long-Term Educational Outcomes', *Labour Economics* **17**(1), 91–100.

Rossin, Maya (2011), 'The Effects of Maternity Leave on Children's Birth and Infant Health Outcomes in the United States', *Journal of Health Economics* **30**, 221–239.

Schneeweis, Nicole and Martina Zweimüller (2014), 'Early Tracking and the Misfortune of Being Young', *Scandinavian Journal of Economics* **116**(2), 394–428.

Schochet, Peter Z. (2009), 'Statistical Power for Regression Discontinuity Designs in Education Evaluations', *Journal of Educational and Behavioral Statistics* **34**(2), 238–266.

Snijder, Claudia A., Nel Roeleveld, Egbert te Velde, Eric A.P. Steegers, Hein Raat, Albert Hofman, Vincent W.V. Jaddoe, Alex Burdorf Claudia A. Snijder, Nel Roeleveld, Egbert te Velde, Eric A.P. Steegers, Hein Raat, Albert Hofman, Vincent W.V. Jaddoe and Alex Burdorf (2012), 'Occupational Exposure to Chemicals and Fetal Growth: The Generation R Study ', *Human Reproduction* **27**(3), 910–920.

Sonchak, Lyudmyla (2015), 'Medicaid Reimbursement, Prenatal Care and Infant Health', *Journal of Health Economics* **44**, 10–24.

Stearns, Jenna (2015), 'The Effects of Paid Maternity Leave: Evidence from Temporary Disability Insurance', *Journal of Health Economics* **43**, 85–102.

Torche, Florencia (2011), 'The Effect of Maternal Stress on Birth Outcomes: Exploiting a Natural Experiment', *Demography* **48**(4), 1473–1491.

Valsamakis, George, Christina Kanaka-Gantenbein, Ariadne Malamitsi-Puchner and George Mastorakos (2006), 'Causes of Intrauterine Growth Restriction and the Postnatal Development of the Metabolic Syndrome', *Annals of the New York Academy of Sciences* **1092**, 138–147.

Wüst, Miriam (2015), 'Maternal Employment During Pregnancy and Birth Outcomes: Evidence From Danish Siblings', *Health Economics* **24**(6), 711–725.

Zweimüller, Josef, Rudolf Winter-Ebmer, Rafael Lalive, Andreas Kuhn, Jean-Philippe Wuellrich, Oliver Ruf and Simon Büchi (2009), The Austrian Social Security Database (ASSD), Working Paper 0901, The Austrian Center for Labor Economics and the Analysis of the Welfare State, University of Linz.

Figure 1 — Average pre- and postnatal ML durations by birth date of the child.
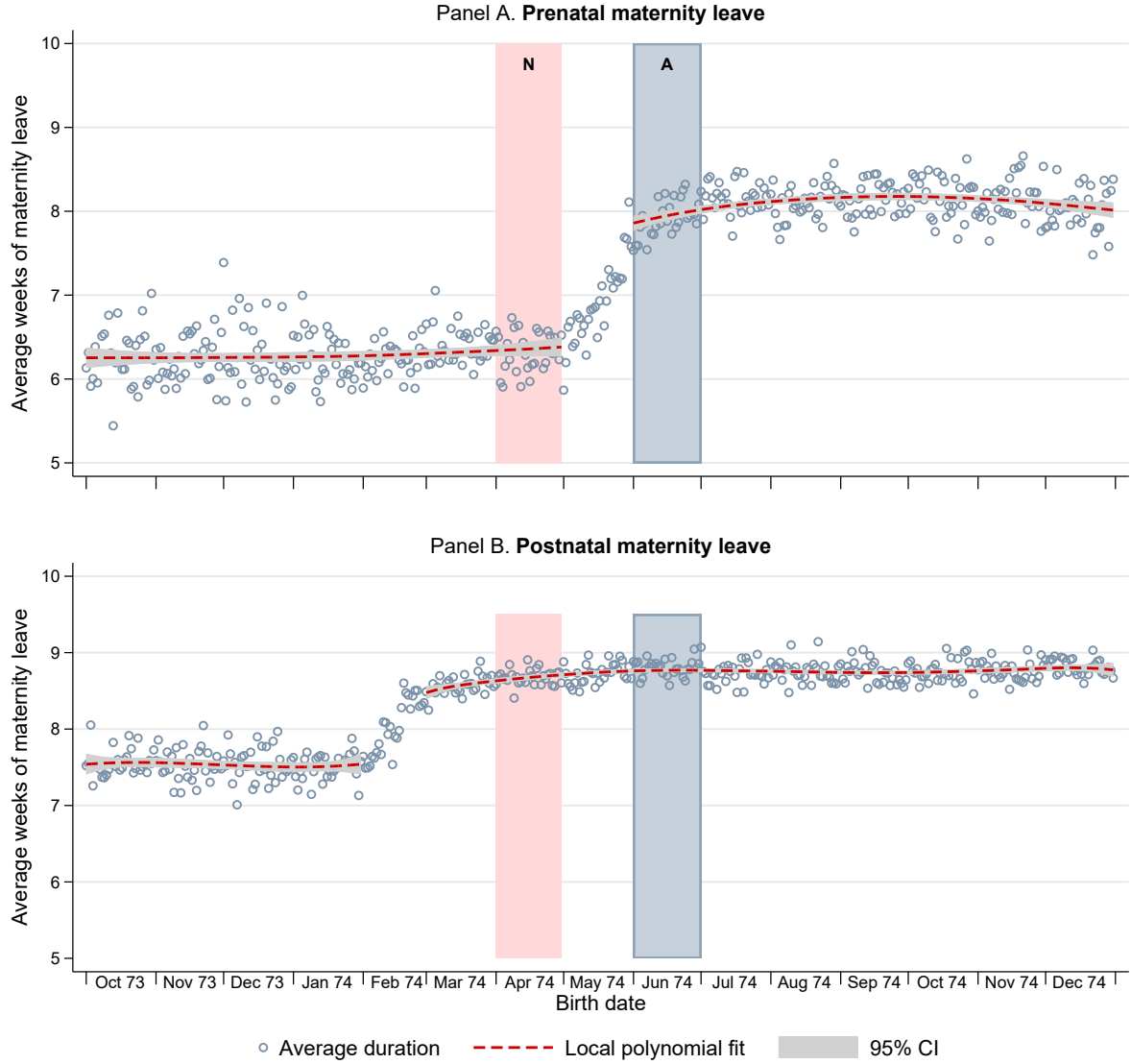
Panel A. **Prenatal maternity leave**

Panel B. **Postnatal maternity leave**

Notes: These graphs show the average prenatal (Panel A) and postnatal (Panel B) ML durations by birth date of the child between October 1973 and December 1974. Separate local quadratic fits for the pre-extension period (Panel A: until April 30, 1974; Panel B: until January 31, 1974) and the post-extension period (Panel A: starting with June 1, 1974; Panel B: starting with March 1, 1974) are depicted by the scattered lines. The red-shaded area highlights the subset of non-assigned births ('N'), which we use in our estimation analysis. These mothers were assigned to 6 weeks of prenatal ML. The framed blue-shaded area highlights the subset of assigned births ('A'), which we use in our estimation analysis. These mothers were assigned to 8 weeks of prenatal ML. Assigned and non-assigned mothers were on postnatal ML for 8 weeks. Mothers who gave birth in May (during which the reform was phased-in) are excluded form our estimation analysis.

FIGURE 2 — Density of assignment variable (number of daily births).
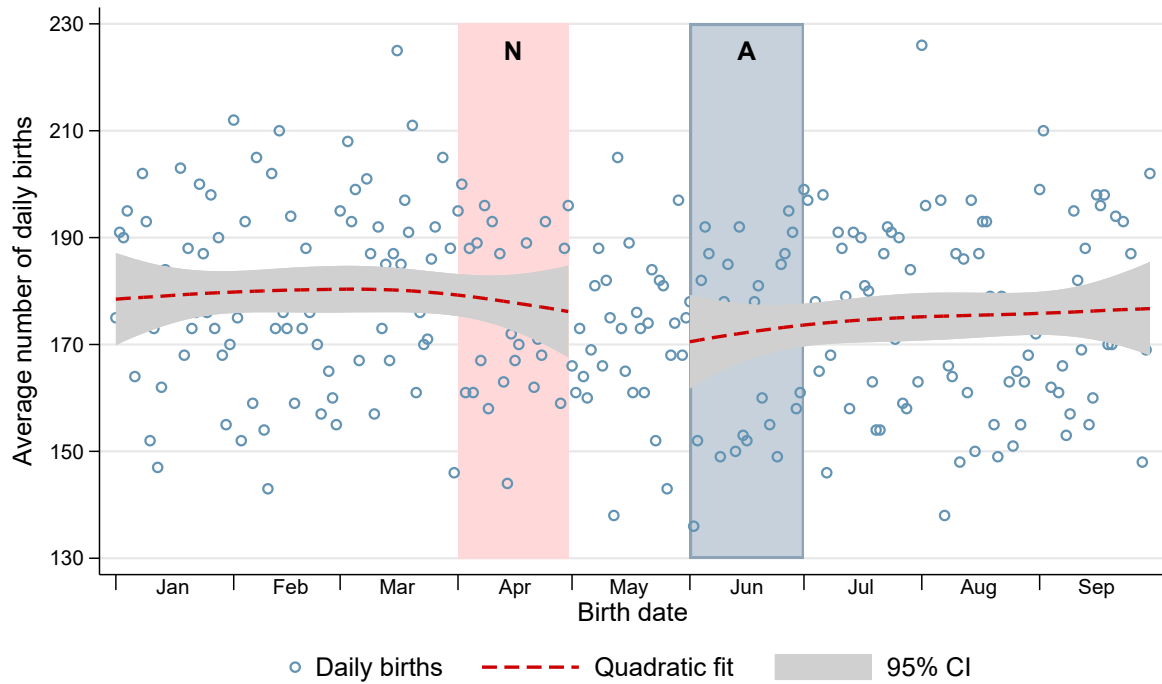
Notes: This figure depicts the average number of daily births between January and September 1974. Separate quadratic fits for the pre-treatment (until April 30, 1974) and the post-treatment period (starting with June 1, 1974) are depicted by the scattered lines. The red-shaded area ▨ highlights the subset of non-assigned births ('N'), which we use in our estimation analysis. These mothers were assigned to 6 weeks of prenatal ML. The framed blue-shaded area ▨ highlights the subset of assigned births ('A'), which we use in our estimation analysis. These mothers were assigned to 8 weeks of prenatal ML. Mothers who gave birth in May (during which the reform was phased-in) are excluded form our estimation analysis.

FIGURE 3 — Daily averages of covariates and sample stratification variables.
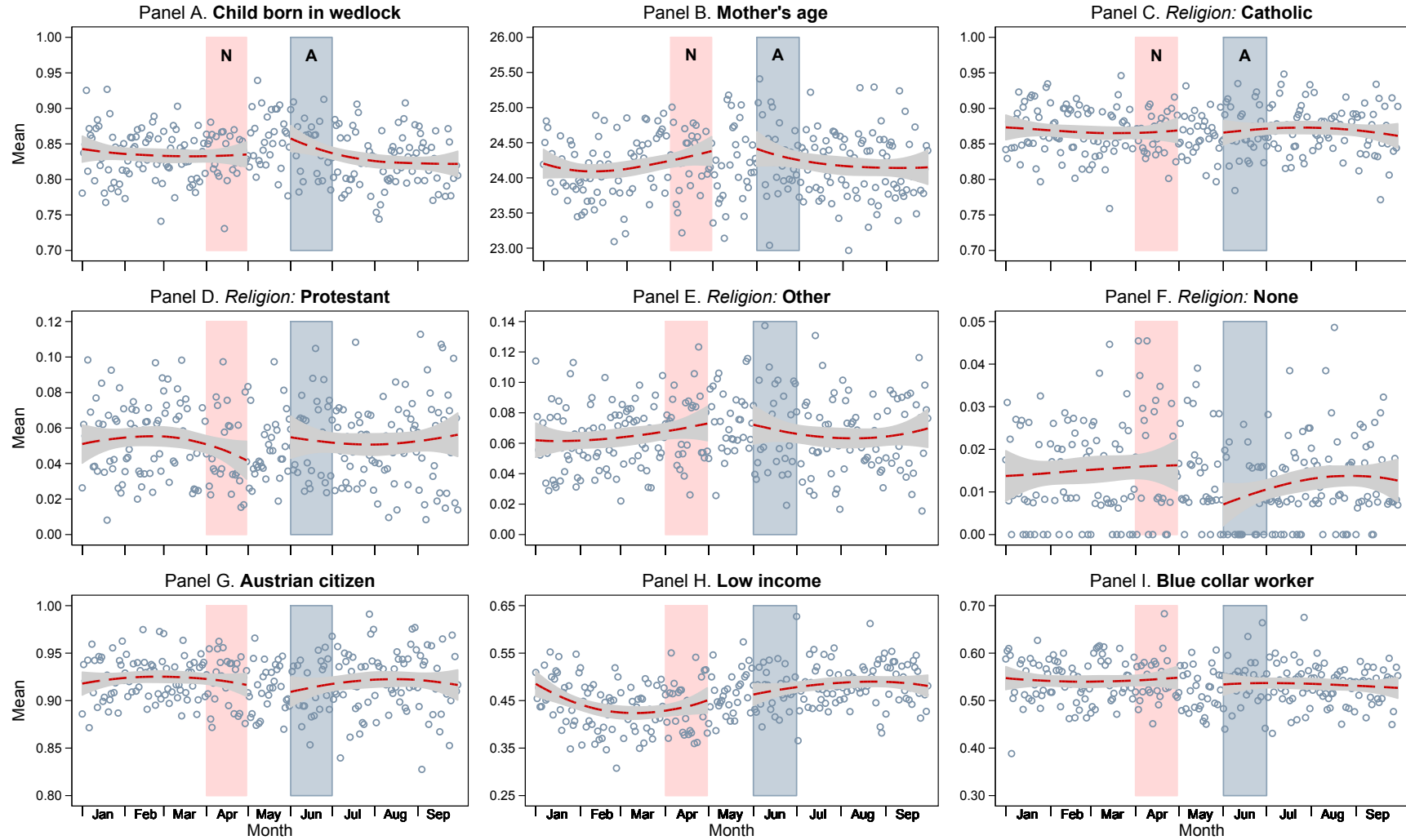
Notes: In these graphs, we plot daily averages for several covariates and sample stratification variables in our data. Separate quadratic fits for the pre-treatment (until April 30, 1974) and the post-treatment period (starting with June 1, 1974) are depicted by the scattered line. The red-shaded area ▨ highlights the subset of non-assigned births ('N'), which we use in our estimation analysis. These mothers were assigned to 6 weeks of prenatal ML duration. The framed blue-shaded area ▨ highlights the subset of assigned births ('A'), which we use in our estimation analysis. These mothers were assigned to 8 weeks of prenatal ML. Mothers who gave birth in May (during which the reform was phased-in) are excluded form our estimation analysis. We observe no significant discontinuities at the cutoff.

FIGURE 4 — Heterogeneous treatment effects for health at birth outcomes.
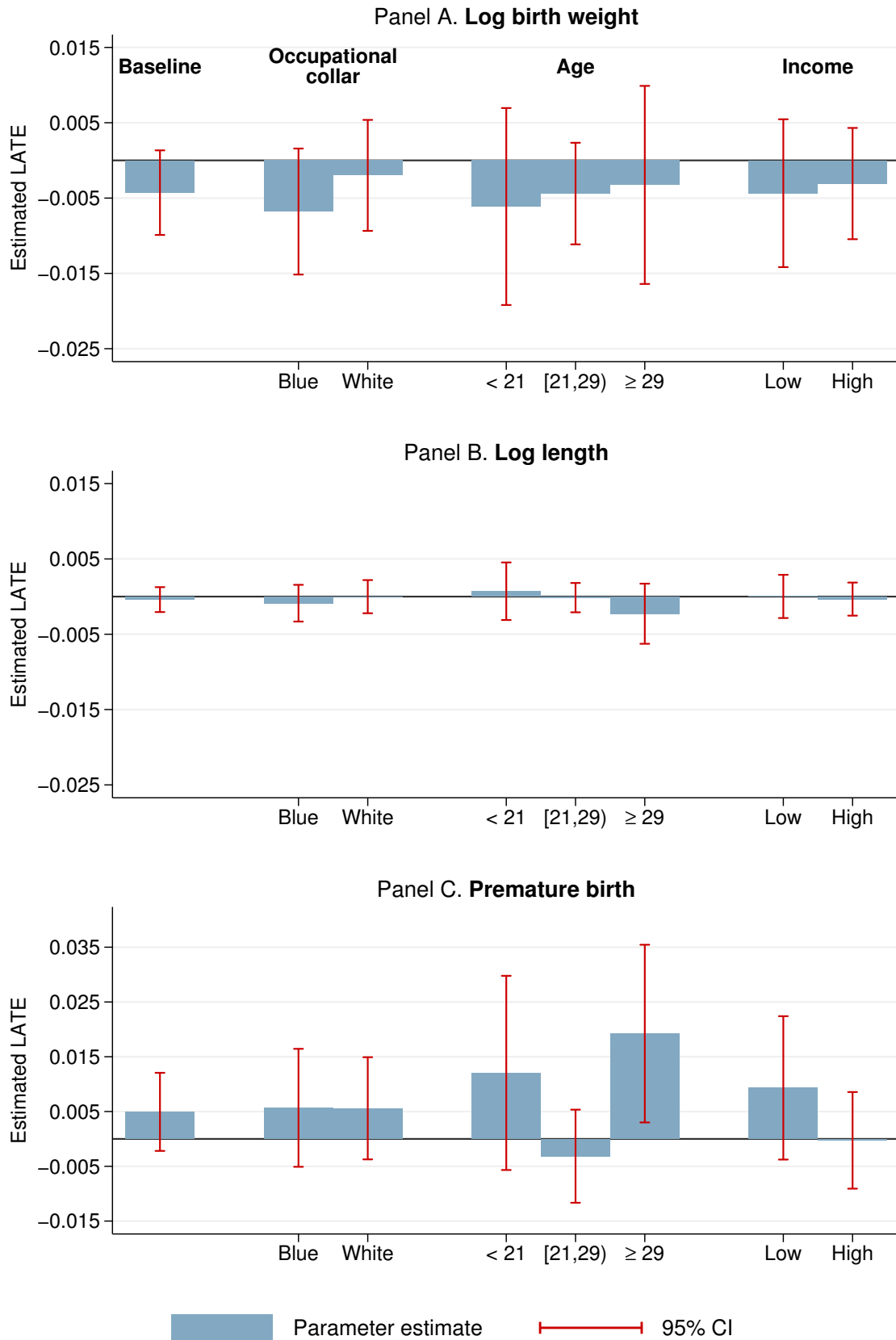
Notes: This figure summarizes fuzzy RDD estimates (obtained via 2SLS) of extending compulsory prenatal ML duration on health at birth for different subsamples. The duration of prenatal ML is instrumented with the assignment to a reform which extended compulsory leave by two weeks. Corresponding first stage estimates are summarized in Table 2. Further details are provided in the notes to Table 4. Graphs for the full set of health at birth outcomes are given in Panel A of Appendix Figure A.4.

FIGURE 5 — Estimated treatment effects on children's labour market outcomes over the life cycle.
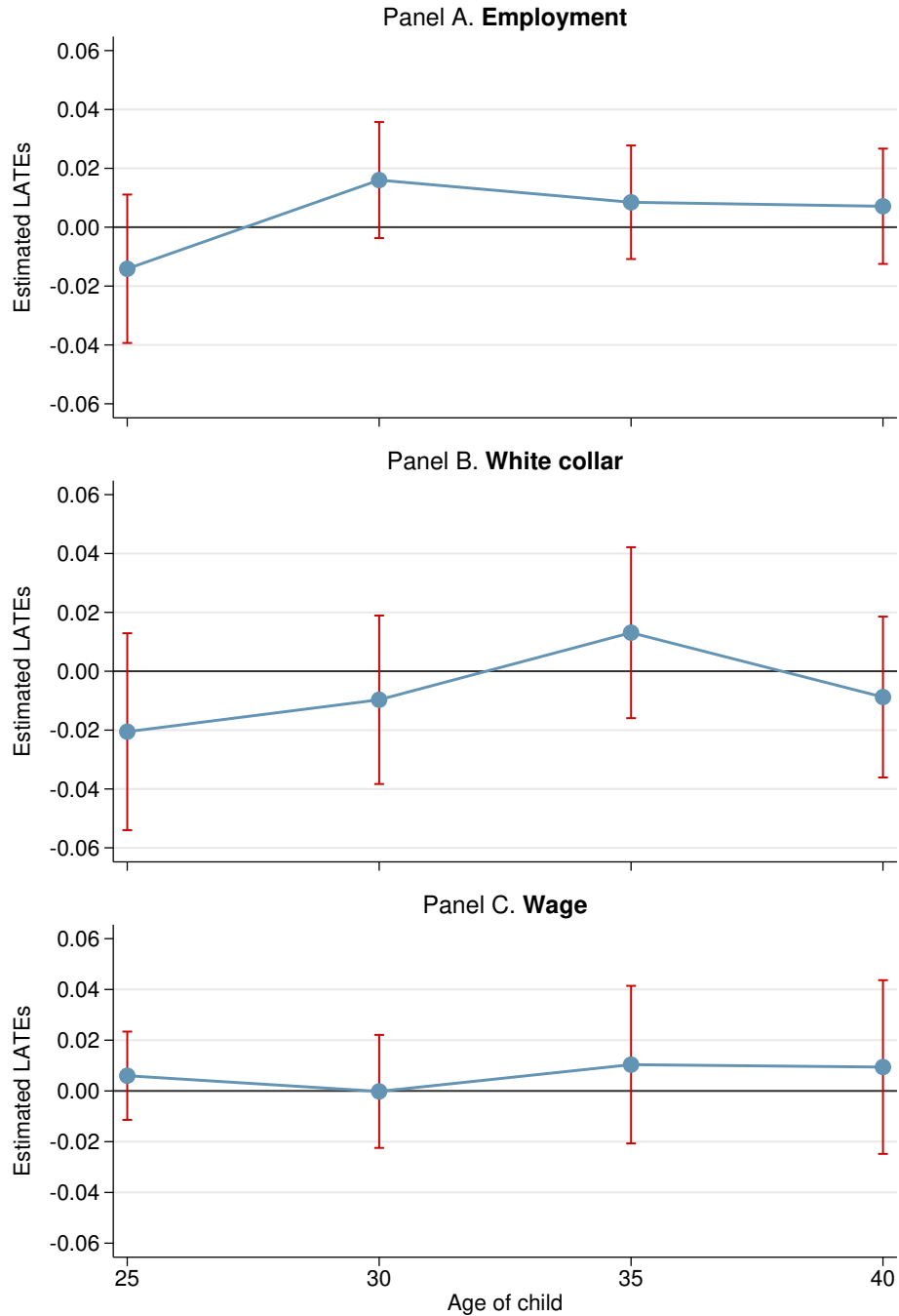
Notes: This figure presents fuzzy RDD estimates of extending ML duration by two weeks on long-term labor market outcomes over the child's life cycle (at ages 25, 30, 35, and 40); where the respective outcome is regressed on prenatal ML duration (in weeks), instrumented by a reform assignment indicator. Each dot represents a separate regression. The sample in each regression consists of children born to working mothers giving birth in April and June 1974, who could uniquely be tracked in our administrative data and for whom we had data on the respective outcome variable. The outcome in Panel A is 'employed,' a binary variable indicating whether the child was in employment at a certain age. The outcome in Panel B is 'white collar,' a binary variable indicating whether the child worked in a white-collar job (conditional on being employed) at a certain age. The outcome in Panel C is 'wage,' the daily wage in € 100 (conditional on being employed) at a certain age (the coefficients on wage are divided by its sample mean, see Table 3). In each specification, we control for a binary variable indicating whether the child was born in wedlock, child sex, the mother's religion, whether the mother is an Austrian citizen, the province a mother lives in, and flexibly for age of the mother (separate dummies for every value of age between 20 and 34, and two additional categories indicating whether age is lower than 20 or higher than 34). Error bars indicate the 95 percent confidence interval and are based on robust standard errors.

FIGURE 6 — Heterogeneous treatment effects for subsequent maternal outcomes.
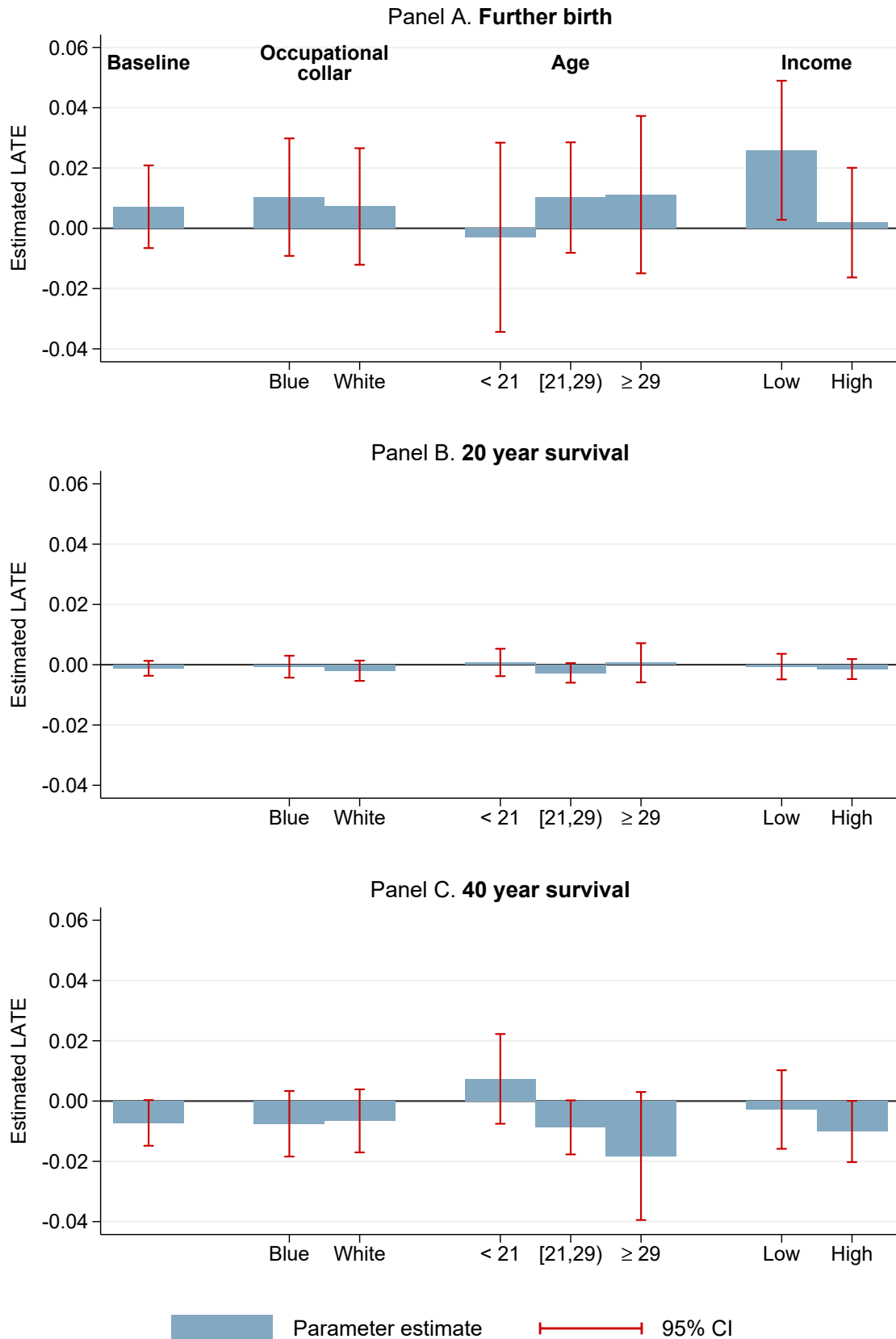
Notes: This figure summarizes fuzzy RDD estimates (obtained via 2SLS) of extending compulsory prenatal ML duration on subsequent maternal fertility and mortality for different subsamples. The duration of prenatal ML is instrumented with the assignment to a reform which extended compulsory leave by two weeks. Corresponding first stage estimates are summarized in Table 2. Further details are provided in the notes to Table 6. Panel A of Appendix Figure A.4 provides equivalent figures for *all* maternal outcomes.

TABLE 1 — Paid and job-protected ML across OECD-countries (OECD, 2017; Eurofound, 2015).

Country	Time period[a]	Maternity leave (weeks)[b]	Of which can be taken before birth[c]
Austria	1970s	12, 16 (1974)	6, 8 (1974)
	1990s	16	8
	2010s	16	8
Belgium	1970s	14	6
	1990s	14, 15 (1991)	6, 7 (1991)
	2010s	15	6
Canada	1970s	15 (1971), 17 (1976)	8 (1971)
	1990s	17	8
	2010s	17	8
Czech Republic	1970s	26	8
	1990s	28	8
	2010s	28	8
Denmark	1970s	14	8
	1990s	18	4
	2010s	18	4
France	1970s	14, 16 (1978)	6
	1990s	16	6
	2010s	16	6
Germany	1970s	14	6
	1990s	14	6
	2010s	14	6
Greece	1970s	12	6
	1990s	15, 16 (1993)	7, 8 (1993)
	2010s	17	8
Hungary	1970s	12	6
	1990s	12, 24 (1992)	6, 4 (1992)
	2010s	24	4
Italy	1970s	14.7, 21.7 (1972)	8.7
	1990s	21.7	8.7
	2010s	21.7	8.7
Japan	1970s	12	6
	1990s	14	6
	2010s	14	6
Mexico	1970s	4.3, 12 (1975)	6 (1975)
	1990s	12	6
	2010s	12	6
Spain	1970s	12, 14 (1976)	6, 8 (1976)
	1990s	16	10
	2010s	16	10
UK	1970s	18	11
	1990s	18	11
	2010s	39	11
USA[d]	1970s	0	0
	1990s	0	0
	2010s	0	0
Median across countries	1970s	14	6
	1990s	16	6
	2010s	16.5	6

Notes: [a]Time periods cover the years 1970-1979, 1990-1999 and 2010-2014; [b]Federal regulations reported (Canada: ML is 17 weeks, including 2 weeks that are not paid; Denmark: job-protection introduced in 1989; France: ML is 26 weeks (8 before birth) for a third or later-born child). [c]Prenatal ML is mandatory in Austria, Belgium, Czech Republic, France, Germany, Greece, Hungary, Italy and Mexico (Austria: 8 weeks mandatory since 1974; Belgium: 1 week mandatory since 1991; Czech Republic: ML starts at least 6 weeks before due date; France: 2 weeks mandatory; Germany: 6 weeks mandatory; Greece: 8 weeks mandatory; Hungary: 2 weeks mandatory; Italy: mandatory leave reduced from 8.7 to 4.3 weeks upon medical proof in 2000; Mexico: 2 weeks mandatory). [d]Unpaid, job-protected ML of up to 12 weeks introduced in 1993.

TABLE 2 — First-stage regressions.

	N	Coef.	Std. err.	F-statistic[†]	Shea's r^{2}[‡]	p-value[§]
Baseline	7,350	1.586***	(0.058)	756.6	0.093	
Subsamples						
Mother's age						
< 21	1,935	1.333***	(0.105)	160.4	0.076	0.013
[21, 29)	3,959	1.669***	(0.079)	444.6	0.101	—
≥ 29	1,456	1.724***	(0.139)	153.6	0.097	0.968
Income						
Low	3,654	1.493***	(0.081)	337.4	0.085	—
High	3,108	1.684***	(0.074)	522.6	0.109	0.073
Occupation						
Blue collar	3,941	1.520***	(0.079)	373.7	0.087	—
White collar	3,344	1.675***	(0.086)	375.4	0.100	0.190

Notes: This table gives first-stage statistics of RDD regressions for the baseline sample consisting of all working mothers giving birth in April and June 1974 and the subsamples of our heterogeneity analysis. We provide the first-stage coefficient (obtained from a regression of ML duration on the assignment variable) along with its standard error, the F-statistic of the first-stage, and the partial r^2 of the first-stage. Each row represents a separate first-stage regression.

[†] Kleinbergen-Paap rK Wald F-statistic
[‡] Shea's partial R^2
[§] Test for the equality of first stage coefficients with respect to the base group (—)

TABLE 3 — Descriptive statistics, assigned vs. non-assigned mothers.

	Assigned		Non-assigned			
	N	Mean	N	Mean	Diff.	p-value[a]
Prenatal maternity leave (in weeks)	3,629	7.92	3,721	6.33	-1.59	0.000
Health at birth outcomes						
Birth weight (in grams)	3,629	3246.29	3,721	3267.40	21.11	0.087
Birth weight is below 2,500 grams	3,629	0.06	3,721	0.06	-0.00	0.915
Asymmetric growth restriction[b]	3,629	0.04	3,721	0.04	0.00	0.973
Length (in cm)	3,629	50.41	3,721	50.44	0.03	0.622
Premature birth	3,629	0.07	3,721	0.06	-0.01	0.200
Long-term child outcomes						
Employed at age 40	1,189	0.84	1,206	0.83	-0.01	0.501
White collar employee at age 40[c]	1,000	0.69	1,002	0.69	0.01	0.677
Daily wage at age 40 (in €100)[c]	781	1.20	778	1.19	-0.02	0.564
Agg. physician expenses (in €1,000)[d]	255	1.81	256	1.86	0.05	0.800
Agg. hospital days[d]	255	8.13	256	10.26	2.13	0.313
Maternal outcomes						
Number of next births	3,629	0.71	3,721	0.69	-0.02	0.301
Further birth	3,629	0.50	3,721	0.49	-0.01	0.451
Time to next birth (in years)[e]	1,803	4.34	1,816	4.37	0.03	0.774
20 year survival probability	3,629	0.99	3,721	0.99	0.00	0.363
40 year survival probability	3,629	0.92	3,721	0.93	0.01	0.080
Agg. physician expenses (in €1,000)[d]	551	5.02	566	5.11	0.09	0.674
Agg. hospital days[d]	551	35.87	566	33.48	-2.39	0.495
Sample stratification variables						
Blue collar worker	3,604	0.54	3,681	0.54	0.00	0.718
Below median income in 1973	3,340	0.48	3,422	0.44	-0.03	0.007
Covariates						
Age at birth	3,629	24.31	3,721	24.27	-0.04	0.745
Child born in wedlock	3,629	0.85	3,721	0.84	-0.01	0.285
Child is male	3,629	0.50	3,721	0.51	0.01	0.230
Religion						
Catholic	3,629	0.87	3,721	0.87	-0.00	0.792
Protestant	3,629	0.05	3,721	0.05	-0.01	0.113
Other religion	3,629	0.07	37,21	0.07	0.00	0.766
No religion	3,629	0.01	3,721	0.02	0.01	0.002
Mother is Austrian citizen	3,629	0.91	3,721	0.92	0.01	0.383

Notes: This table presents summary statistics for our treatment (ML duration), as well as our outcome, sample stratification, and control variables. Statistics are provided separately for both assigned mothers (i.e., mothers giving birth in June 1974) and non-assigned mothers (giving birth in April 1974). The population includes only mothers who had been working and had been eligible for ML.

[a] The p-value corresponds to a t-test for the difference in sample means.
[b] Asymmetric growth restriction is defined as having low birth weight *and* a low Ponderal index ($PI = kg/m^3$).
[c] Conditional on being employed.
[d] Aggregated over the period of 25 to 40 years after birth.
[e] Conditional on having a further birth.

TABLE 4 — Estimated treatment effects on health at birth outcomes.

	(1) Log birth weight	(2) Low birth weight	(3) Asymmetric growth restr.	(4) Log length	(5) Premature birth
Panel A. RDD					
Prenatal maternity leave	−0.004	0.0003	−0.00006	−0.0004	0.005
	(0.003)	(0.003)	(0.003)	(0.001)	(0.004)
No. of observations	7,350	7,350	7,350	7,350	7,350
Mean of outcome	5.77	0.06	0.04	3.92	0.06
Kleinbergen-Paap *rK* Wald *F*-statistic	756.63	756.63	756.63	756.63	756.63
Panel B. OLS (only pre-treatment period)					
Prenatal maternity leave	0.006***	−0.006***	−0.002	0.001**	−0.011***
	(0.002)	(0.002)	(0.002)	(0.001)	(0.003)
No. of observations	3,721	3,721	3,721	3,721	3,721
Mean of outcome	5.77	0.06	0.04	3.92	0.06
Panel C. RDD-DiD					
Prenatal maternity leave	−0.002	0.0005	−0.001	0.0002	
	(0.005)	(0.006)	(0.005)	(0.001)	
No. of observations	10,424	10,424	10,424	10,424	
Mean of outcome	5.78	0.05	0.03	3.92	
Kleinbergen-Paap *rK* Wald *F*-statistic	754.64	754.64	754.64	754.64	
Panel D. DiD					
Assigned × working	−0.004	0.0008	−0.002	0.0003	
	(0.008)	(0.009)	(0.007)	(0.002)	
No. of observations	10,424	10,424	10,424	10,424	
Mean of outcome	5.78	0.05	0.03	3.92	

Notes: This table summarizes estimated effects of extending compulsory prenatal ML duration on health at birth. Panel A summarizes fuzzy RDD estimates (obtained via 2SLS), where the duration of prenatal ML is instrumented by the assignment to a reform that extended compulsory leave by two weeks. Corresponding first stage estimates are summarized in Table 2. Panel B summarizes OLS estimates for the pre-treatment period, where the prenatal ML duration is used as an explanatory variable. Panel D are difference-in-differences estimates which compare pre- and post-reform outcomes between working and non-working mothers. In Panel C we combine these two sources of variation in regression discontinuity difference-in-differences estimators. In Panel A the sample consists of working mothers giving birth in April and June 1974, in panel B the sample is restricted to women giving birth in April 1974. In panels C and D we extend the sample from A with non-working mothers giving birth in April and June 1974. Each cell represents a separate estimation. The outcomes 'birth weight' and 'length' (columns 1 and 4) are continuous variables specified in logs, while 'low birth weight' (column 2), 'asymmetric growth restriction' (column 3), and 'premature birth' (column 5) are binary variables indicating whether birth weight is below 2,500 grams, whether both birth weight is low and the Ponderal index is in the lowest quarter of its sample distribution, and whether the child was born prematurely, respectively. In each specification we control for a binary variable indicating whether the child was born in wedlock, child sex, the mother's religion, whether the mother is an Austrian citizen, the province the mother lives in, and flexibly for mother's age (separate dummies for every value of age between 20 and 34, and two additional categories indicating whether age is lower than 20 or higher than 34). Note that coefficients in panels A, B, and C correspond to a one week increase in ML, while coefficients in panel D indicate the ITT of a two week increase in ML. Robust standard errors are in parentheses, stars indicate statistical significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

TABLE 5 — Estimated treatment effects on children's long-term outcomes.

	Labour market outcomes at age 40			Health outcomes between age 25 and 40	
	(1) Employed	(2) White collar	(3) Wage[†]	(4) Outpatient expenses[†]	(5) Hospital days[†]
Panel A. RDD					
Prenatal maternity leave	0.007	−0.005	−0.001	−0.026	−0.154
	(0.010)	(0.014)	(0.013)	(0.069)	(0.149)
No. of observations	2,395	2,002	1,559	511	511
Mean of outcome	0.84	0.69	1.00	1.00	1.00
Kleinbergen-Paap *rK* Wald *F*-statistic	205.76	177.61	130.33	48.66	48.66
Panel B. Matched sample reduced form estimates					
Born in June 1974	0.011	−0.007	−0.001	−0.042	−0.249
	(0.015)	(0.021)	(0.020)	(0.115)	(0.243)
No. of observations	2,395	2,002	1,559	511	511
Mean of outcome	0.84	0.69	1.00	1.00	1.00
Panel C. Full population reduced form estimates					
Born in June 1974	0.005	0.0007	−0.0005	0.04	−0.09
	(0.005)	(0.008)	(0.008)	(0.033)	(0.090)
No. of observations	15,450	13,838	11,023	3,287	3,287
Mean of outcome	0.90	0.54	1.00	1.00	1.00

Notes: This table presents fuzzy RDD estimates of extending ML duration by two weeks on long-term child outcomes in Panel A, where the respective outcome is regressed on prenatal ML duration (in weeks), instrumented by a reform-assignment indicator. Each column represents a separate regression. The sample in each column consists of children born to working mothers giving birth in April and June 1974, who could uniquely be tracked in the our administrative data and for whom we had data on the respective outcome variable. The outcome 'employed' is a binary variable indicating whether the child was in employment at age 40, 'white collar' is a binary variable indicating whether the child worked in a white-collar job at age 40, conditional on being employed, 'wage' is the daily wage in € 100 at age 40, conditional on being employed, 'outpatient expenses' are aggregated physician expenses between age 25 and 40 in € 1,000, and 'hospital days' is the aggregate number of days spent in hospital between age 25 and 40. In each specification, we control for a binary variable indicating whether the child was born in wedlock, child sex, the mother's religion, whether the mother is an Austrian citizen, the province a mother lives in, and flexibly for age of the mother (separate dummies for every value of age between 20 and 34, and two additional categories indicating whether age is lower than 20 or higher than 34). Compulsory ML was extended by two weeks due to the reform, hence coefficients have to be multiplied by a factor of 2. Additionally, in Panel B, we present reduced form estimates for the sample of matched children, and in Panel C we present results for the entire population of children born in April or June 1974 observed at age 40 in our data, irrespective of whether we can match them to their mothers as in Panels A and B. Regressing the respective outcome on a binary variable indicating whether the child was born in June 1974 (as opposed to April in the same year) gives us a reduced form estimate of the reform. In Panel B we only control for the child's sex, and whether the child could be identified in the birth register. Robust standard errors are in parentheses, stars indicate statistical significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

[†] Coefficients are divided by the respective variable's sample mean (see Table 3).

TABLE 6 — Estimated treatment effects on subsequent maternal outcomes.

	(1) No. of next births[†]	(2) Further birth	(3) Log time to next birth[‡]	(4) 20 year survival	(5) 40 year survival	(6) Outpatient expenses[†,§]	(7) Hospital days[†,§]
Panel A. RDD							
Prenatal maternity leave	0.023	0.007	−0.002	−0.001	−0.007*	−0.016	0.035
	(0.018)	(0.007)	(0.016)	(0.001)	(0.004)	(0.023)	(0.057)
No. of observations	7,350	7,350	3,619	7,350	7,350	1,117	1,117
Mean of outcome	1.00	0.49	7.10	0.99	0.92	1.00	1.00
Kleinbergen-Paap *rK* Wald *F*-statistic	756.63	756.63	366.73	756.63	756.63	185.00	185.00
Panel B. OLS (only pre-treatment period)							
Prenatal maternity leave	0.005	0.003	−0.01	−0.00002	−0.0009	−0.02	−0.0007
	(0.007)	(0.003)	(0.008)	(0.000)	(0.002)	(0.011)	(0.027)
No. of observations	3,721	3,721	1,816	3,721	3,721	566	566
Mean of outcome	1.00	0.49	7.11	0.99	0.93	1.00	1.00

Notes: This table presents estimated treatment effects of extending compulsory ML duration by two weeks on different subsequent maternal fertility and health outcomes. Each cell represents a separate regression. The sample in Panel A consists of working mothers giving birth in April and June 1974, in Panel B the sample is restricted to women giving birth in April 1974. 'No. of next births' (column 1) is a count variable measuring the number of children the mother has given birth to subsequently, 'further birth' (column 2) is a binary variable indicating whether the mother gave birth at least one more time, and 'time to next birth' (column 3) is the number of days passed until the mother gave birth again in logs, conditional on having another child. The outcomes '20 year survival' and '40 year survival' (columns 4 and 5) are binary variables indicating whether the mother was still alive 20 and 40 years after birth, respectively, 'outpatient expenses' are aggregated physician expenses (25 to 40 years after giving birth) in € 1,000, and 'hospital days' is the aggregate number of days spent in hospital (25 to 40 years after giving birth). In each specification we control for a binary variable indicating whether the child was born in wedlock, child sex, the mother's religion, whether the mother is an Austrian citizen, the province a mother lives in, and flexibly for age of the mother (separate dummies for every value of age between 20 and 34, and two additional categories indicating whether age is lower than 20 or higher than 34). Panel A presents fuzzy RDD estimates obtained via 2SLS where duration of ML is instrumented by assignment to the reform, panel B are OLS estimates where ML duration is used as an explanatory variable. Compulsory ML was extended by two weeks, hence coefficients in panel A have to be multiplied by a factor of 2. Robust standard errors are in parentheses, stars indicate statistical significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

[†] Coefficients are divided by the respective variable's sample mean (see Table 3).
[‡] Time to next birth is conditional on giving birth again, thus the sample includes only mothers who had another child.
[§] Health outcomes are only available for a subset of mothers living in Upper Austria.

A. Online Appendix

This Online Appendix (not for publication) provides additional material discussed in the unpublished manuscript 'The Effect of Prenatal Maternity Leave on Short and Long-term Child Outcomes' by Alexander Ahammer, Martin Halla, and Nicole Schneeweis. Appendix A provides additional figures and tables, Appendix A.2 discusses in more detail the measurement error in our instrumental variable, and Appendix A.3 provides details on the ex-post calculation of a minimum detectable effect size.

A.1. Additional figures and tables

FIGURE A.1 — Average pre- and postnatal ML durations between 1973 and 1975.
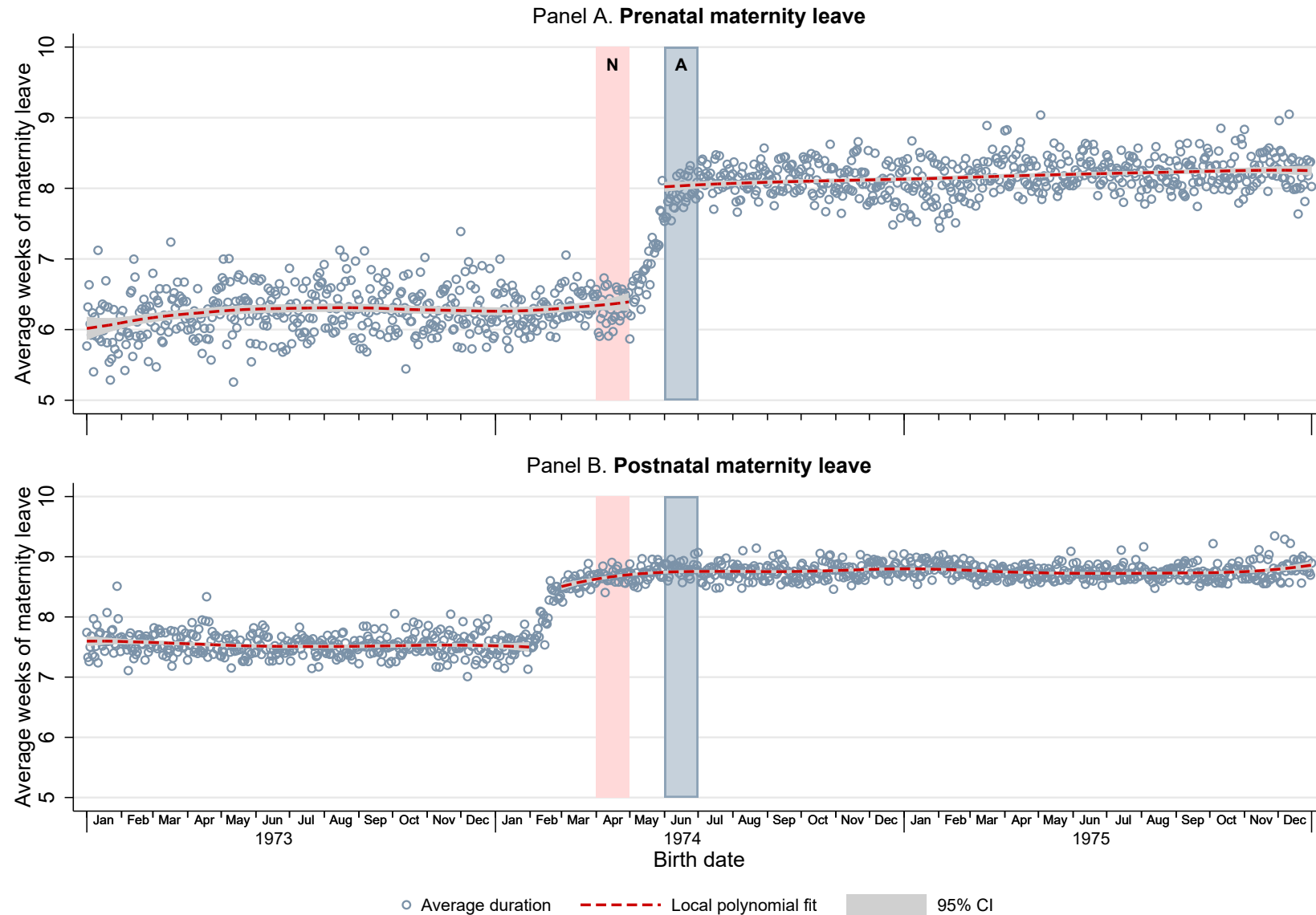
Panel A. **Prenatal maternity leave**

Panel B. **Postnatal maternity leave**

○ Average duration　－－－－ Local polynomial fit　▨ 95% CI

Notes: These graphs show the average prenatal (Panel A) and postnatal (Panel B) ML durations by birth date of the child between January 1973 and December 1975. Separate local quadratic fits for the pre-extension period (Panel A: until April 30, 1974; Panel B: until January 31, 1974) and the post-extension period (Panel A: starting with June 1, 1974; Panel B: starting with March 1, 1974) are depicted by the scattered line. The red-shaded area ▨ highlights the subset of non-assigned births ('N'), which we use in our estimation analysis. These mothers were assigned to 6 weeks of prenatal ML. The framed blue-shaded area ▨ highlights the subset of assigned births ('A'), which we use in our estimation analysis. These mothers were assigned to 8 weeks of prenatal ML. Assigned and non-assigned mothers were assigned to 8 weeks of postnatal ML. Mothers who gave birth in May (during which the reform was phased-in) are excluded form our estimation analysis.

A.2

FIGURE A.2 — Event studies, differences in outcomes between working and non-working mothers.
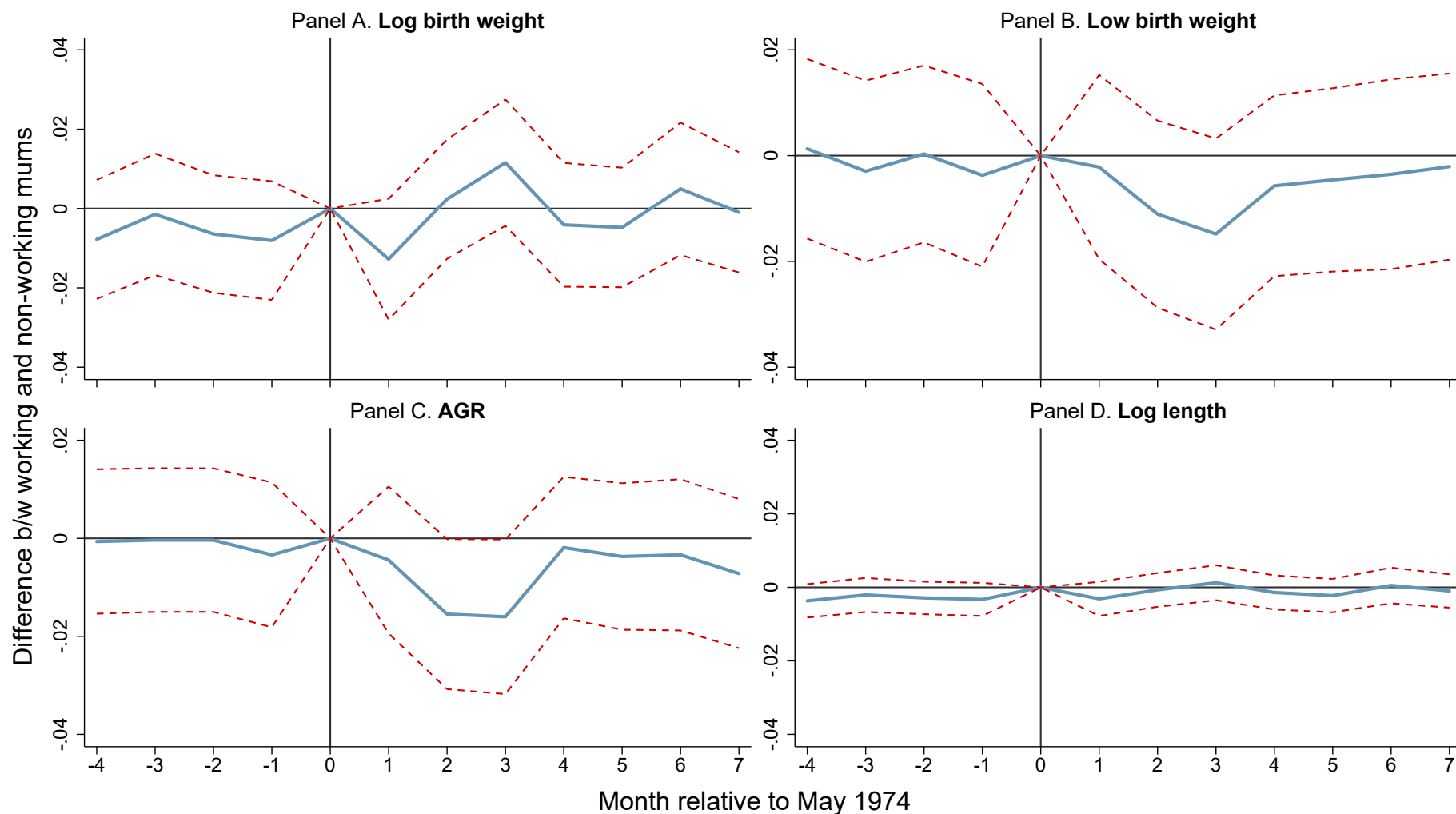
Notes: These graphs plot the differences of our outcome variables between working and non-working mothers in 1974. In each panel, we regress the outcome on a dummy variable indicating whether the mother was working in the respective month, which is interacted with a set of dummies indicating the calendar month relative to May 1974. We plot the coefficients on these interaction terms along with their 95 percent confidence intervals (based on robust standard errors). To aid visual comparison of estimates, we plot event studies based on continuous (panels A and D) and binary (panels B and C) variables on separate scales.

FIGURE A.3 — Health at birth RDD estimates with wider windows around the cutoff date.
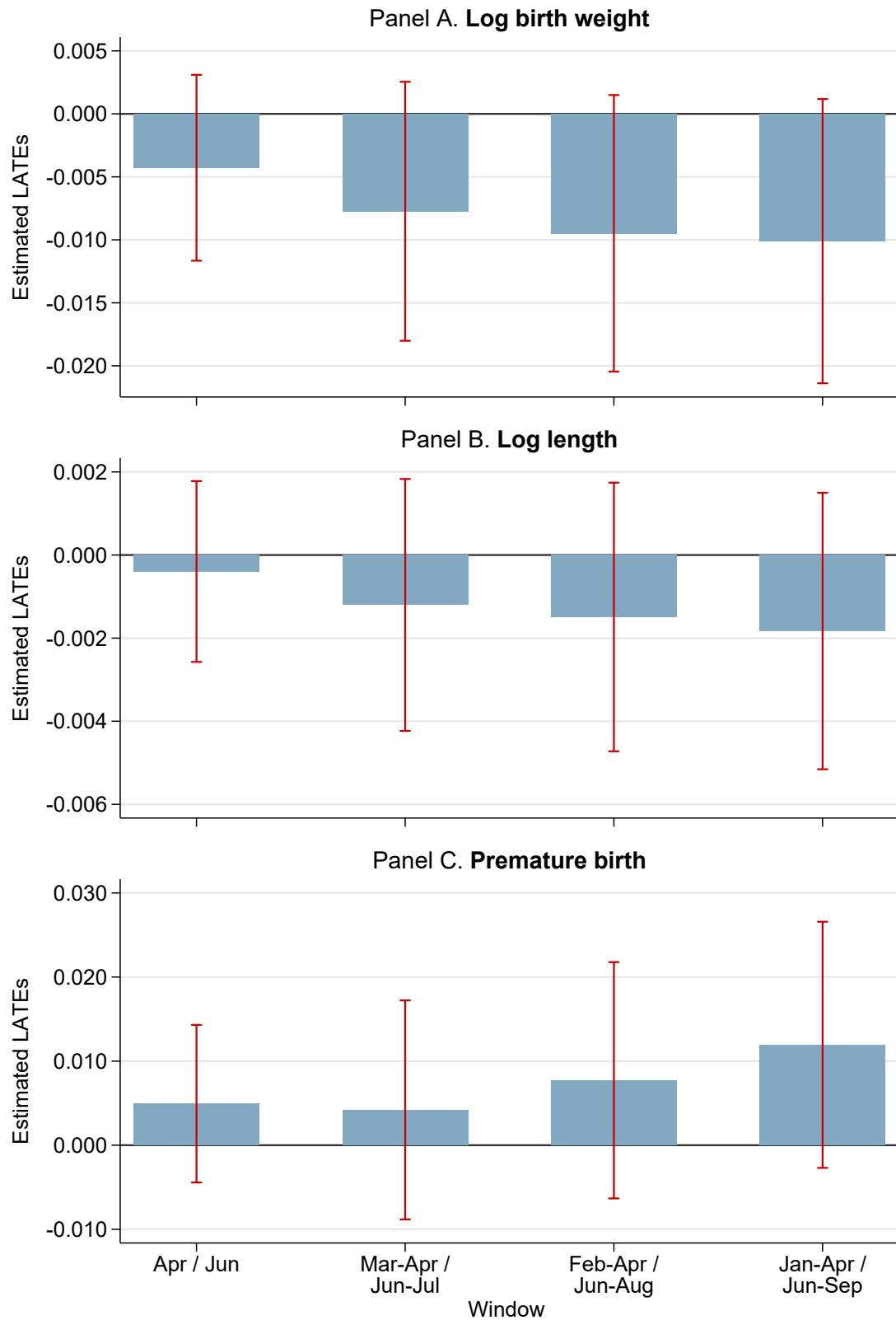
Notes: In this graph we summarize the results of performing our RDD estimations on different samples based on the window around the cutoff date we consider. The leftmost bar (*'Apr / Jun'*) resembles our baseline results from Table 4, where we essentially compare mothers giving birth in June 1974 to mothers giving in April 1974. For the estimate depicted by the second bar from the left (*'Mar–Apr / Jun–Jul'*) we extend the window by one month on each side of the cutoff, comparing mothers giving birth in June and July 1974 to those giving birth in March and April 1974. In order to obtain the estimates depicted by the next two bars, we continue to extend the window by another 1 and 2 months, respectively, on each side of the cutoff. Similar graphs for other outcomes we consider are available upon request.

FIGURE A.4 — Heterogeneity analysis for the full set of outcome variables.
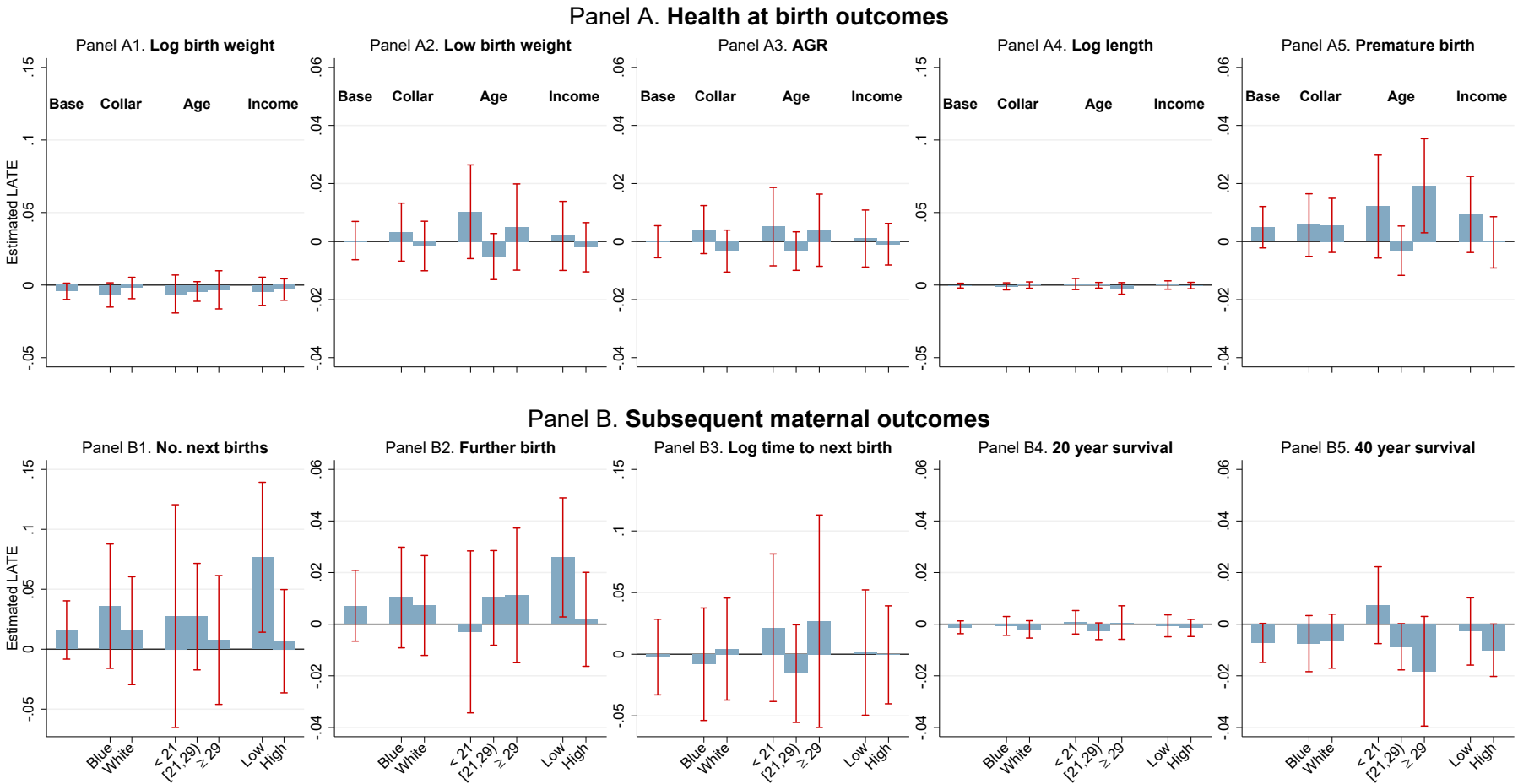
Notes: In this figure we extend the heterogeneity analyses from the main manuscript by all remaining outcome variables. Please see section II.3.3 in the paper for a description on each variable. For the number of next births (panel B1), all coefficients are divided by the sample mean of the number of next births. To aid the visual comparison of estimates, continuous and binary outcomes are on separate scales.

A.5

FIGURE A.5 — Unconditional quantile treatment effects.
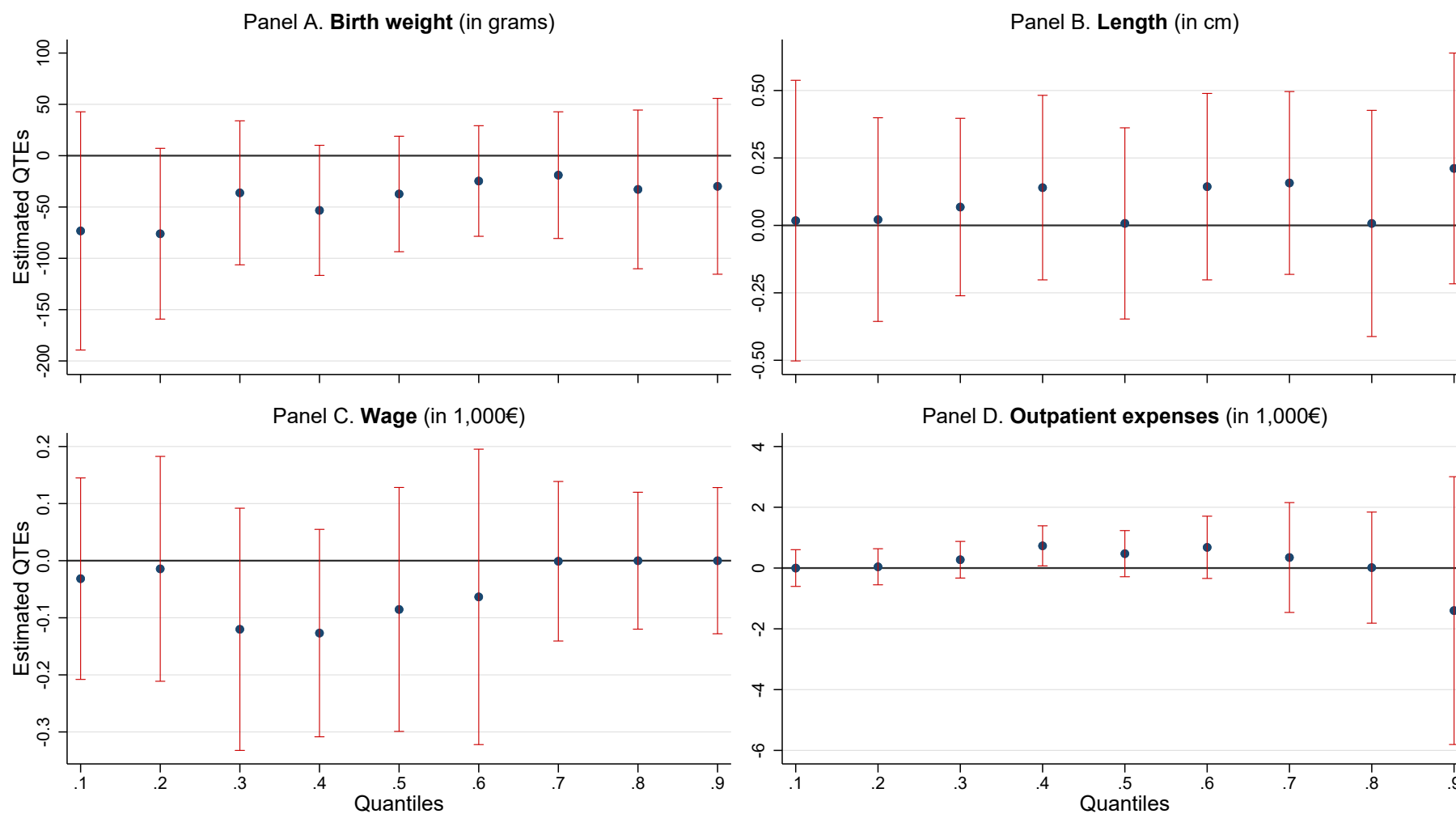
Notes: Unconditional weighted quantile treatment effects are estimated using the Frölich & Melly (2013, *JBES*) estimator, which allows for endogeneity in the treatment variable. We use a global logit to estimate the suggested complier weights without any local covariate smoothing (however, note that under local smoothing results were very similar; our bandwidth choices for these analyses were $h = \{0.1, 0.5, 1, 10, 50, 100, 500, 1000, 5000, 7500\}$). In estimating the complier weights, we control for a binary variable indicating whether the child was born in wedlock, child sex, the mother's religion, whether the mother is an Austrian citizen, the province a mother lives in, and flexibly for age of the mother (separate dummies for every value of age between 20 and 34, and two additional categories indicating whether age is lower than 20 or higher than 34). Note that birth weight and length are rounded in our data, thus they are in fact discrete realizations of continuous processes. Because our quantile estimator performed badly for these two outcomes, we added white Gaussian noise to the variables prior to estimation. Error bars indicate 95% confidence intervals.

TABLE A.1 — Related variables across OECD-countries.

Country	Time period[a]	Fertility rate[b]	Female employment[c]	% full-time[d]	Infant mortality[e]	Maternal mortality[f]
Austria	1970s	1.9	47.4	84.6	21.0	20.9
	1990s	1.4	57.3	74.6	6.0	4.5
	2010s	1.4	66.9	54.9	3.4	3.0
Belgium	1970s	1.9	39.8	-	16.6	-
	1990s	1.6	47.3	69.7	6.7	6.5
	2010s	1.8	57.2	57.3	3.5	5.0
Canada	1970s	1.9	44.9	75.6	14.7	10.7
	1990s	1.6	61.6	71.6	6.0	3.9
	2010s	1.6	69.6	73.0	4.9	5.8
Czech Republic	1970s	2.2	71.8	-	18.9	20.0
	1990s	1.5	61.7	89.7	7.7	11.3
	2010s	1.5	59.4	90.0	2.6	6.3
Denmark	1970s	1.8	60.0	-	10.9	-
	1990s	1.8	69.5	64.1	5.7	12.2
	2010s	1.7	71.2	63.2	3.6	4.9
France	1970s	2.1	48.1	-	14.0	20.3
	1990s	1.7	55.0	72.2	5.8	10.6
	2010s	2.0	61.6	69.6	3.5	6.6
Germany	1970s	1.6	48.9		18.5	38.2
	1990s	1.3	55.9	66.4	5.6	6.3
	2010s	1.4	68.9	53.6	3.4	4.5
Greece	1970s	2.3	31.3	-	23.7	23.4
	1990s	1.3	38.1	91.5	7.8	3.0
	2010s	1.4	45.9	88.3	3.5	2.9
Hungary	1970s	2.1	-	-	31.0	30.1
	1990s	1.6	47.4	91.5	11.8	13.0
	2010s	1.3	51.9	91.0	4.9	11.4
Italy	1970s	2.2	32.1	87.7	22.6	31.9
	1990s	1.3	37.2	87.6	6.5	4.6
	2010s	1.4	47.1	69.4	2.9	2.5
Japan	1970s	2.0	52.6	-	10.4	33.3
	1990s	1.5	60.4	-	4.1	7.6
	2010s	1.4	66.7	-	2.2	4.2
Mexico	1970s	6.0	26.5	-	63.9	163.3
	1990s	3.0	37.0	61.8	26.8	82.3
	2010s	2.2	46.6	65.0	13.3	41.3
Spain	1970s	2.8	30.7	-	19.9	22.7
	1990s	1.2	32.8	85.1	6.0	3.5
	2010s	1.3	51.8	75.9	3.0	3.1
UK	1970s	2.0	53.0	-	15.9	17.4
	1990s	1.8	62.2	55.8	6.4	6.7
	2010s	1.9	66.6	57.0	4.0	6.2
USA	1970s	1.9	49.2	73.0	16.4	14.4
	1990s	2.0	65.4	73.9	7.9	8.0
	2010s	1.9	63.5	73.8	6.0	-
Median across countries	1970s	2.0	47.8	75.6	17.2	21.6
	1990s	1.6	56.8	71.8	6.3	6.6
	2010s	1.5	62.0	69.0	3.5	4.9

Notes: Data are drawn from the OECD Family Database 2017, the OECD Health Statistics 2019, and the OECD Labor Force Statistics 2019.
[a]Time periods cover the years 1970-1979, 1990-1999 and 2010-2014; [b]Number of children per women aged 15-59; [c]Civilian female employment as percent of female population aged 15-64; [d]Female full-time employment as percent of female employment; [e]Deaths below age 1 per 1,000 live births; [f]Maternal deaths per 100,000 live births;
Mean values for denoted time periods reported; yearly gaps in the data were linearly interpolated in 32 cases; data for some countries and years are missing (Fertility rate: Canada 2013-2014; Female employment rate: Austria 2014, Czech Republic 1970-1974, Denmark 2014, France 2014, Greece 1970, 2013, 2014, Hungary 1970-1991, Japan 2014, Sweden 2014; Female full-time employment: Austria 1970-1972, Belgium 1970s, Canada 1970-1975, Czech Republic 1970-1992, Denmark 1970s, France 1970s, Germany 1970s, Greece 1970s, Hungary 1970-1994, Italy 1970-1972, Japan 1970-2014, Mexico 1970-1990, Spain 1970s, Sweden 1970-1975, UK 1970s; Infant mortality: Mexico 1970; Maternal mortality: Belgium 1970s, Czech Republic 1970-1977, Denmark 1970-1996, France 2014, USA 2010s).

TABLE A.2 — Treatment effects estimated without control variables.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A. Health at birth outcomes							
	Log birth weight	Low birth weight	Asymmetric growth restr.	Log length	Premature birth		
Prenatal maternity leave	−0.005	0.0004	−0.00009	−0.0005	0.005		
	(0.003)	(0.003)	(0.003)	(0.001)	(0.004)		
No. of observations	7,350	7,350	7,350	7,350	7,350		
Kleinbergen-Paap rK Wald F-statistic	752.77	752.77	752.77	752.77	752.77		
Panel B. Long-term child outcomes							
	Employed	White collar	Wage[†]	Outpatient expenses[†]	Hospital days[†]		
Prenatal maternity leave	0.007	−0.006	0.010	−0.017	−0.139		
	(0.010)	(0.014)	(0.018)	(0.068)	(0.139)		
No. of observations	2,395	2,002	1,559	511	511		
Kleinbergen-Paap rK Wald F-statistic	199.07	172.90	130.08	64.88	64.88		
Panel C. Maternal outcomes							
	No. of next births[†]	Further birth	Log time to next birth[‡]	20 year survival	40 year survival	Outpatient expenses[†,§]	Hospital days[†,§]
Prenatal maternity leave	0.019	0.006	−0.004	−0.001	−0.007*	−0.010	0.041
	(0.019)	(0.007)	(0.016)	(0.001)	(0.004)	(0.024)	(0.060)
No. of observations	7,350	7,350	3,619	7,350	7,350	1,117	1,117
Kleinbergen-Paap rK Wald F-statistic	752.77	752.77	368.18	752.77	752.77	171.81	171.81

Notes: This table replicates the fuzzy RDD estimations for extending compulsory prenatal ML duration from Table 4 (Panel A, health at birth outcomes), Table 5 (panel B, long-term child outcomes), and Table 6 (maternal outcomes), but without any covariates used in the regression. As in the baseline specifications, coefficients correspond to a one week increase in ML. Robust standard errors are in parentheses, stars indicate statistical significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

[†] Coefficients are divided by the respective variable's sample mean (see Table 3).

[‡] Time to next birth is conditional on giving birth again, thus the sample includes only mothers who had another child.

[§] Health outcomes are only available for a subset of mothers living in Upper Austria.

TABLE A.3 — Descriptive statistics, working vs. non-working mothers.

	Working		Non-working			
	N	Mean	N	Mean	Diff.	p-value[a]
Prenatal maternity leave (in weeks)[b]	7350	7.11	3074	40.00	32.89	0.000
Health at birth outcomes						
Birth weight (in grams)	7350	3256.98	3074	3351.40	94.42	0.000
Birth weight is below 2,500 grams	7350	0.06	3074	0.04	−0.02	0.000
Asymmetric growth restriction[c]	7350	0.04	3074	0.03	−0.01	0.002
Length (in cm)	7350	50.42	3074	50.94	0.52	0.000
Covariates						
Age at birth	7350	24.29	3074	27.07	2.78	0.000
Child born in wedlock	7350	0.84	3074	0.92	0.08	0.000
Child is male	7350	0.51	3074	0.53	0.02	0.045
Religion						
Catholic	7350	0.87	3074	0.92	0.05	0.000
Protestant	7350	0.05	3074	0.05	−0.00	0.973
Other religion	7350	0.07	3074	0.02	−0.05	0.000
No religion	7350	0.01	3074	0.01	−0.00	0.034
Mother is Austrian citizen	7350	0.92	3074	0.98	0.06	0.000

Notes: This table presents summary statistics for our treatment (ML duration); as well as our outcome and control variables used for RDD-DiD analyses. The sample is comprised of mothers giving birth in April and June 1974. Statistics are provided separately for both working and non-working mothers, where working status is assessed at time of birth.

[a] The p-value corresponds to a t-test for the difference in sample means.
[b] ML duration is assumed to be 40 weeks for non-working mothers. The specific value chosen has no impact on the estimation results.
[c] Asymmetric growth restriction is defined as having low birth weight *and* a low Ponderal index ($PI = kg/m^3$).

TABLE A.4 — Socioeconomic gradients in health at birth outcomes

	N	Low birth weight	AGR	Premature birth
		Selected outcome variables		
Panel A. Occupation				
White collar	3,344	0.05	0.03	0.05
Blue collar	3,941	0.06	0.04	0.07
Difference		−0.01	−0.01	−0.02
p-value		0.009	0.159	0.004
Panel B. Income				
Above 75th pct.	1,690	0.05	0.03	0.06
Below 25th pct.	1,689	0.07	0.04	0.09
Difference		−0.02	0.00	−0.04
p-value		0.008	0.466	0.000
Panel C. Austrian citizenship				
Yes	6,186	0.05	0.03	0.06
No	1,164	0.09	0.06	0.09
Difference		−0.04	−0.03	−0.04
p-value		0.000	0.000	0.000
Panel D. Born in wedlock				
Yes	6,741	0.05	0.04	0.06
No	609	0.07	0.05	0.11
Difference		−0.02	−0.01	−0.05
p-value		0.066	0.220	0.000
Panel E. Age at birth				
Age > 35	260	0.09	0.06	0.08
Age < 20	1,256	0.06	0.04	0.08
Difference		0.02	0.01	0.00
p-value		0.192	0.346	0.913

Notes: This table provides sample means for our binary health at birth outcomes for different socioeconomic groups. The sample is based on mothers giving birth in April and June 1974.

A.10

TABLE A.5 — Estimated treatment effects on health at birth outcomes with the sample being restricted to the first half of April and the second half of June.

	(1) Log birth weight	(2) Low birth weight	(3) Asymmetric growth restr.	(4) Log length	(5) Premature birth
Panel A. Baseline RDD					
Prenatal maternity leave	−0.004	0.0003	−0.00006	−0.0004	0.005
	(0.003)	(0.003)	(0.003)	(0.001)	(0.004)
No. of observations	7,350	7,350	7,350	7,350	7,350
Mean of outcome	5.77	0.06	0.04	3.92	0.06
Kleinbergen-Paap rK Wald F-statistic	756.63	756.63	756.63	756.63	756.63
Panel B. RDD with observations weighted with inverse measurement error					
Prenatal maternity leave	−0.003	−0.002	−0.002	−0.0002	0.007
	(0.003)	(0.004)	(0.003)	(0.001)	(0.004)
No. of observations	7,350	7,350	7,350	7,350	7,350
Mean of outcome	5.77	0.05	0.04	3.92	0.07
Kleinbergen-Paap rK Wald F-statistic	524.68	524.68	524.68	524.68	524.68
Panel C. RDD with half measurement error					
Prenatal maternity leave	−0.003	0.001	−0.002	−0.001	0.006
	(0.004)	(0.004)	(0.004)	(0.001)	(0.005)
No. of observations	3,777	3,777	3,777	3,777	3,777
Mean of outcome	5.77	0.06	0.04	3.92	0.06
Kleinbergen-Paap rK Wald F-statistic	433.15	433.15	433.15	433.15	433.15
Panel D. Baseline RDD-DiD					
Prenatal maternity leave	−0.002	0.0005	−0.001	0.0002	
	(0.005)	(0.006)	(0.005)	(0.001)	
No. of observations	10,424	10,424	10,424	10,424	
Mean of outcome	5.78	0.05	0.03	3.92	
Kleinbergen-Paap rK Wald F-statistic	754.64	754.64	754.64	754.64	
Panel E. RDD-DiD with half measurement error					
Prenatal maternity leave	0.002	0.001	−0.003	0.001	
	(0.006)	(0.007)	(0.006)	(0.002)	
No. of observations	5,372	5,372	5,372	5,372	
Mean of outcome	5.78	0.05	0.04	3.92	
Kleinbergen-Paap rK Wald F-statistic	432.95	432.95	432.95	432.95	
Panel F. Baseline DiD					
Assigned × working	−0.004	0.0008	−0.002	0.0003	
	(0.008)	(0.009)	(0.007)	(0.002)	
No. of observations	10,424	10,424	10,424	10,424	
Mean of outcome	5.78	0.05	0.03	3.92	
Panel G. DiD with half measurement error					
Assigned × working	0.003	0.001	−0.005	0.001	
	(0.010)	(0.013)	(0.010)	(0.003)	
No. of observations	5,372	5,372	5,372	5,372	
Mean of outcome	5.78	0.05	0.04	3.92	

Notes: This table summarizes RDD estimates where each observation giving birth on day t is weighted by $1/(|ME_t| + 0.1)$, where ME_t is the average measurement error in days given in Tables A.7 and A.8. Additionally, we provide RDD, RDD-DiD, and DiD estimates equivalent to Table 4, but based on samples restricted to the first half of April (April 1–April 15) and the second half of June (June 16–June 30) in Panels C, E, and G. Additionally, we include the corresponding baseline estimates from Table 4 in Panels A, D, and F. Robust standard errors are in parentheses, stars indicate statistical significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

TABLE A.6 — Treatment effects estimated based on a sample of children for whom at least one long-term outcome is available.

	(1)	(2)	(3)	(4)	(5)		
Panel A. Health at birth outcomes							
	Log birth weight	Low birth weight	Asymmetric growth restr.	Log length	Premature birth		
Prenatal maternity leave	−0.005	0.005	0.003	−0.000	0.007		
	(0.004)	(0.005)	(0.004)	(0.001)	(0.006)		
No. of observations	2,400	2,400	2,400	2,400	2,400		
Kleinbergen-Paap *rK* Wald *F*-statistic	206.21	206.21	206.21	206.21	206.21		
Panel B. Maternal outcomes							
	No. of next births[†]	Further birth	Log time to next birth[‡]	20 year survival	40 year survival	Outpatient expenses[†,§]	Hospital days[†,§]
Prenatal maternity leave	0.002	0.009	−0.001	−0.001	−0.012*	0.011	0.014
	(0.033)	(0.013)	(0.027)	(0.001)	(0.007)	(0.046)	(0.120)
No. of observations	2,400	2,400	1,247	2,400	2,400	423	423
Kleinbergen-Paap *rK* Wald *F*-statistic	206.21	206.21	109.91	206.21	206.21	35.52	35.52

Notes: This table replicates the fuzzy RDD estimations for extending compulsory prenatal ML duration from Table 4 (Panel A, health at birth outcomes) and Table 6 (Panel B, maternal outcomes) on a subset of children for whom we observe at least one of the long-term outcomes from Table 5. As in the baseline specifications, coefficients correspond to a one week increase in ML. Robust standard errors are in parentheses, stars indicate statistical significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

[†] Coefficients are divided by the respective variable's sample mean (see Table 3).
[‡] Time to next birth is conditional on giving birth again, thus the sample includes only mothers who had another child.
[§] Health outcomes are only available for a subset of mothers living in Upper Austria.

A.2. Measurement error in the instrumental variable

We have some measurement error (ME) in our instrumental variable (IV).[1] A ME in the instrumental variable weakens the first stage. Fortunately, our first stage is still sufficiently strong. However, there is a potential correlation between the ME in the IV and the error term in the equation of interest. To discuss how this correlation could bias our results, we proceed in four steps. First, we describe the source and nature of the ME. Second, we discuss the potential extent of the ME in our main sample. Third, we describe the potential correlation between the ME and the error term. Fourth, we report results based on a subsample with very little ME. We will conclude that the ME in our IV is too small to cause any significant bias.

Source and nature of measurement error.) The ME results from the fact that we determine the assignment to the extended prenatal ML based on the actual birth date, while in reality the expected due date was decisive. A ME, however, only occurs in the case of either very early or very late births:

1.) April births: presumably assigned to 42 days (=6 weeks) of ML:

   a.) Kids who were born in April and were late arrivers (with due dates in March or earlier) are classified correctly

   b.) Kids who were born on time or up to 14 days earlier are classified correctly[2]

   c.) *Only kids who were born 21 or more days earlier* have a ME in their assignment variable.

2.) June births: presumably assigned to 56 days (=8 weeks) of ML:

   a.) Kids who were born in June but were early arrivers (with due dates in July or later) are classified correctly

   b.) Kids who were born on time or up to 7 days after their due date are also classified correctly[3]

   c.) *Only kids who were born 14 or more days later* have a ME in their assignment variable.

In two tables below we show in detail the ME in the assigned days by birth date. Appendix Table A.7 shows for June births the ME by birth date for births arriving $0, 7, 14, \ldots, 56$ days early. Equivalently, Appendix A.8 shows for April births the ME in the assigned days by birth date for births arriving $0, 7, 14, 28$ days late. In these tables, the ME is calculated as actual assignment (number of assigned days of prenatal ML based on the due date) minus presumed assignment (56 or 42 days). Here we have to keep in mind that children with an expected due date in May were phased-in. For example, a child born on June 1 might have been born 7 days after the due date, which is May 25. The actual assignment is 54 days, the measured assignment is 56 days (= 8 weeks), the ME is therefore minus 2 days. In our estimation strategy, we use a binary IV (equal to one for an assignment of 8 weeks of prenatal ML, zero for an assignment of 6 weeks of prenatal ML). Thus, for this fictitious child born on June 1, whose expected due date was on May 25, we assume $A_i = 1$, despite the actual leave being only 54 (and not 56 days).

---

[1]Note that we have a precise measure of our endogenous treatment variable the duration of prenatal maternity leave. Our administrative data allows us to observe the exact starting and end date. Thus, we do <u>not</u> face the problem of attenuation bias in the structural equation of our RDD-IV estimation approach.

[2]Minor exception: Children born on April 30, who are born 14 days early (see Appendix Table A.8).

[3]Minor exception: Children born on June 1 and June 2, who are born 7 days later (see Appendix Table A.7).

TABLE A.7 — Measurement error in the assignment variable by birth date: June births.

	Born x days later					
$x =$	0	7	14	21	28	
Birth date						**Average ME**[†]
June 1	0	-2	-9	-14	-14	-0.701
June 2	0	-1	-8	-14	-14	-0.527
June 3	0	0	-7	-14	-14	-0.353
June 4	0	0	-6	-13	-14	-0.308
June 5	0	0	-5	-12	-14	-0.263
June 6	0	0	-4	-11	-14	-0.218
June 7	0	0	-3	-10	-14	-0.173
June 8	0	0	-2	-9	-14	-0.128
June 9	0	0	-1	-8	-14	-0.083
June 10	0	0	0	-7	-14	-0.039
June 11	0	0	0	-6	-13	-0.034
June 12	0	0	0	-5	-12	-0.029
June 13	0	0	0	-4	-11	-0.024
June 14	0	0	0	-3	-10	-0.019
June 15	0	0	0	-2	-9	-0.014
June 16	0	0	0	-1	-8	-0.009
June 17	0	0	0	0	-7	-0.004
June 18	0	0	0	0	-6	-0.003
June 19	0	0	0	0	-5	-0.003
June 20	0	0	0	0	-4	-0.002
June 21	0	0	0	0	-3	-0.002
June 22	0	0	0	0	-2	-0.001
June 23	0	0	0	0	-1	-0.001
June 24	0	0	0	0	0	0.000
June 25	0	0	0	0	0	0.000
June 26	0	0	0	0	0	0.000
June 27	0	0	0	0	0	0.000
June 28	0	0	0	0	0	0.000
June 29	0	0	0	0	0	0.000
June 30	0	0	0	0	0	0.000

Share of births in the year 1984:

| 0.525 | 0.134 | 0.040 | 0.005 | 0.0005 |

Average ME across all birth dates[‡]**= –0.098 days**

Notes: This table shows the measurement error in the assignment variable in days by birth date for births arriving 0, 7, 14, 21, and 28 days later. The ME is calculated as actual assignment (number of assigned days of prenatal ML based on the due date) minus measured assignment (56 days). For example, a child born on June 1 might have been born 7 days after the due date, which is May 25 (the phasing in period). The actual assignment is 54 days, the measured assignment is 56 days, the ME is therefore –2 days.

[†] The 'Average ME by birth date ' is calculated assuming a distribution of gestational length as in the year 1984. For this year the *Austrian Birth Register* includes information on gestational length.

[‡] The 'Average ME across all birth dates' is an unweighted average of the 'average ME by birth date'.

				Born x days earlier						
$x =$	0	7	14	21	28	35	42	49	56	
Birth date										**Average ME**[†]
April 1	0	0	0	0	0	0	0	7	14	0.193
April 2	0	0	0	0	0	0	1	8	14	0.204
April 3	0	0	0	0	0	0	2	9	14	0.215
April 4	0	0	0	0	0	0	3	10	14	0.225
April 5	0	0	0	0	0	0	4	11	14	0.236
April 6	0	0	0	0	0	0	5	12	14	0.247
April 7	0	0	0	0	0	0	6	13	14	0.257
April 8	0	0	0	0	0	0	7	14	14	0.268
April 9	0	0	0	0	0	1	8	14	14	0.285
April 10	0	0	0	0	0	2	9	14	14	0.301
April 11	0	0	0	0	0	3	10	14	14	0.318
April 12	0	0	0	0	0	4	11	14	14	0.334
April 13	0	0	0	0	0	5	12	14	14	0.351
April 14	0	0	0	0	0	6	13	14	14	0.367
April 15	0	0	0	0	0	7	14	14	14	0.384
April 16	0	0	0	0	1	8	14	14	14	0.413
April 17	0	0	0	0	2	9	14	14	14	0.443
April 18	0	0	0	0	3	10	14	14	14	0.473
April 19	0	0	0	0	4	11	14	14	14	0.502
April 20	0	0	0	0	5	12	14	14	14	0.532
April 21	0	0	0	0	6	13	14	14	14	0.562
April 22	0	0	0	0	7	14	14	14	14	0.592
April 23	0	0	0	1	8	14	14	14	14	0.644
April 24	0	0	0	2	9	14	14	14	14	0.697
April 25	0	0	0	3	10	14	14	14	14	0.750
April 26	0	0	0	4	11	14	14	14	14	0.803
April 27	0	0	0	5	12	14	14	14	14	0.856
April 28	0	0	0	6	13	14	14	14	14	0.908
April 29	0	0	0	7	14	14	14	14	14	0.961
April 30	0	0	1	8	14	14	14	14	14	1.069

Share of births in the year 1984:

	0.525	0.136	0.075	0.033	0.020	0.010	0.007	0.004	0.012	

Average ME across all birth dates[‡]**= 0.480**

Notes: This table shows the measurement error in the assignment variable in days by birth date for births arriving 0, 7, 14, 21, 28, 35, 42, 49, and 56 days earlier. The ME is calculated as actual assignment (number of assigned days on prenatal ML based on the due date) minus measured assignment (42 days). For example, a child born on April 30 might have been born 14 days prior to the due date, which is May 14 (the phasing in period). The actual assignment is 43 days, the measured assignment is 42 days, the ME is therefore 1 day.

[†] The 'Average ME by birth date ' is calculated assuming a distribution of gestational length as in the year 1984. For this year the *Austrian Birth Register* includes information on gestational length.

[‡] The 'Average ME across all birth dates' is an unweighted average of the 'average ME by birth date'.

Potential extent of the measurement error.) In order to assess the extent of the ME, one would need to know how many children are in each cell of Appendix Tables A.7 and A.8. Since we do not observe gestational length for children born in 1974, we cannot calculate these numbers. However, we can use the distribution of gestational lengths from the year 1984 (the first year where we observe this variable) to derive an approximation. Here we have to assume that the reform had no impact on gestational length; for which we indeed find no evidence.[4] The shares are listed in the second to last row of Appendix Table A.7 and A.8, respectively. It turns out that the share of children with any ME is small, and the share of children with a large ME is almost negligible. In the case of June births, a large ME occurs only for children, who fulfill two criteria: i. they are born in the beginning of June, and ii. they are born many days after the expected-due date. Late births are very rare.

The very right column lists for each birth date an 'Average ME by birth date'. This is calculated assuming the distribution of gestational length as in the year 1984. To provide an idea for the extent of the overall ME in our main estimation sample, we calculate an 'Average ME across all birth dates'. This is an unweighted average of the 'average ME by birth date'. For June births we have average ME of minus 0.098 days. For April births we apply an equivalent procedure and we find an average ME of 0.480 days. Using these averages, we can contrast the measured (or presumed) assigned days of prenatal ML with the average actual assigned days of prenatal ML for our two groups (see Table A.9). We see now that the average ME in our IV is only 0.6 days or 4 percent.

TABLE A.9 — Extent of ME in IV in main sample

Group	Measured (or presumed) assigned days of prenatal ML	Average actual assigned days of prenatal ML
June births	56	55.90
April births	42	42.48
Difference	14	13.42

Correlations of measurement error.) We have seen that our ME is essentially a function of the expected due date and the difference between the expected due date and the actual birth day Since there is no evidence for sorting, the variation across calendar days should not cause a correlation between the ME and the error term in the structural equation. In contrast, large differences between the expected due date and the actual birth day are not innocuous. Very early (and very late) births are expected to lead to different (health) outcomes. Since we are not able to control for gestational length in our estimations, there could be a correlation between the ME in our IV and the structural error term. While this could bias our results, we expect the overall

---

[4]Given that there is an extended postnatal ML of 12 weeks for mothers who experienced a preterm birth (gestation length is below 37 full weeks), we are able to study whether the increase in prenatal ML had any effects on the proportion of women with preterm births. This regulation was the same before and after the reform. Our results show that this is not the case. Thus, if gestational length has changed due to the reform, it has not changed at the margin of 37 weeks. Are other changes , i.e. at other margins, plausible? Given that we find no effect on birth weight or birth length either — which is strongly correlated with gestational length — we consider also other changes as highly unlikely.

effect, given the small extent of the ME, to be negligible. Still, this is an issue which requires further attention. Below, we suggest a test based on a subsample, with very little ME.

Robustness tests.) We exploit that we can approximate the average ME by the date of birth and implement two strategies. First, we weight our observations by $1/(|ME_t| + 0.1)$. This is the inverse of the absolute value of the average measurement error in days, ME_t, as given in Tables A.7 and A.8. We add 0.1 to account for the fact that there are birth dates with zero days of ME. The results summarized in Panel B of Table A.5 are very similar to the baseline estimates in Panel A. Second, we use a subsample with a lower average ME. We restrict the analysis to births between June 16 and June 30 and April 1 and April 15, where the ME amounts to only minus 0.002 and 0.279 days, respectively. In this subsample, the average ME in our IV is reduced to 0.28 days or 2 percent (see Table A.10 below). Despite reducing the ME in our assignment variable by half, the estimation results do not change as compared to the baseline estimations (compare Panels A and C in Appendix Table A.5). We conclude that the extent of the ME is indeed so small that our main estimation results are not significantly biased.

TABLE A.10 — Extent of ME in IV in robustness sample

Group	Measured (or presumed) assigned days of prenatal ML	Average actual assigned days of prenatal ML
June 16 to 30 births	56	55.998
April 1 to 15 births	42	42.279
Difference	14	13.719

In our other estimation models (RDD-DiD and DiD=reduced form), the ME in the assignment variable leads to different econometric issues. The bottom line is still that the ME is so small such that the consequences for the estimates are negligible. This is supported by an equivalent robustness test, see Panels D to G in Appendix Table A.5).

A.3. Ex-post calculation of minimum detectable effect size

To calculate minimum detectable effect sizes, we follow Schochet (2009); Deke and Dragoset (2012). These papers show that a valid minimum detectable effect size (MDES) for fuzzy regression discontinuity designs can be derived by correcting the MDES of a randomized controlled trial by the degree of imperfect compliance. In our case, perfect compliance would imply that all non-assigned mothers had a prenatal ML duration of 6 weeks, while all assigned mothers had 8 weeks. In contrast, we have sample means of 6.33 and 7.92 weeks, respectively. This gives a difference in treatment of 1.59 weeks (instead of 2 weeks as in the case of full compliance).

- In a first step, we calculate the MDES required under a randomized controlled trial. We assume a power level of 0.8, a significance level of 0.05, and we use the sample mean of non-assigned mothers (3,267.40 grams), as well as the standard deviations (524.34 and 532.96) of both groups. We also account for the fact that we do not have the same number of observations in both groups (but 3,721 and 3,629 mothers, respectively). Based on these parameters we obtain a MDES$^{\text{RCT}}$ of 30.67 grams.

- In a second step, we have to account for the non-compliance. Following Deke and Dragoset (2012), we calculate the MDES by multiplying the MDES$^{\text{RCT}}$ by the square root of the so-called *design effect*. The latter is given by $1/(1 - \rho^2)$, where ρ is the correlation between treatment status and the continuous assignment variable.[5] This gives us a MDES of 32 grams for our reform of a prenatal ML extension by two weeks.

Is an increase in birth weight of 32 grams per two additional weeks (or 16 grams per one additional week) of prenatal ML a large quantity? An increase in the birth weight of an average newborn by 16 grams is equivalent to an increase by only 0.4 percent (or 0.03 standard deviations). For other outcomes, we obtain equally small MDES. For instance, for birth length the MDES is equal to only 0.17 cm per two additional weeks of prenatal ML.

Notably, we have used here (unconditional) standard deviations. In contrast, in our estimates we include covariates, which reduces the variation in the outcome variables. Thus, our statistical power is in fact somewhat larger, reducing the MDES.

---

[5]Note, in our paper we use a binary assignment variable. If we use a binary assignment variable for the calculation of ρ, we obtain essentially the same result.