Journal of Health Economics xxx (xxxx) xxx

Contents lists available at ScienceDirect

Journal of Health Economics

journal homepage: www.elsevier.com/locate/jhe

Negative effects of long parental leave on maternal health: Evidence from a substantial policy change in Austria

Caroline Chuard[1]

University of St. Gallen, Swiss Institute for Empirical Economic Research, Varnbühlstrasse 14, 9000 St. Gallen, Switzerland

ARTICLE INFO

JEL classification:
I18
J13
J16
J18

Keywords:
Parental leave
Maternal health
Mental health
Regression discontinuity

ABSTRACT

I study the effect of parental leave duration on maternal health in the short- to medium-run leveraging variation in parental leave duration induced by an Austrian policy reform in the year 2000. Using rich administrative data and a regression discontinuity framework, I find that long parental leave of 2.5 years instead of 1.5 years deteriorates maternal health. Worse mental health mainly drives this effect. Alternative channels such as differential fertility, long-term employment effects, a change in disposable income or alternative mode of childcare are unlikely to be of main importance. There is substantial heterogeneity with longer leave spells being less harmful for mothers with unhealthy babies proxied by low birth weight.

1. Introduction

Governments around the world seek to improve female well-being with the use of a broad set of gender equalization policies. At the core of these policies is parental leave. Parental leave allows mothers to recover from childbirth and enables women to combine a professional career and motherhood. While most developed countries acknowledge the importance of parental leave policies to enhance female empowerment and gender equality, there is, nevertheless, substantial variation in parental leave policies across the globe. The difference across countries in leave length has even widened over the last decades. Most European countries increase paid parental leave duration to levels of usually around a year. The United States, instead, offers only 12 weeks of unpaid leave. This is the ground for an ongoing debate on optimal leave policies.

The existing literature has done great advances in evaluating the direct impact of leave policies on child development and maternal labor market outcomes as recently reviewed by Olivetti and Petrongolo (2017) and Rossin-Slater (2018). However, although maternal health and well-being have been recognized as important outcomes – if not the most important ones – only few studies have quantified the effect on these outcomes. This information gap implicitly sets the value of mothers to zero, a fact that has also been lately pointed out by a study in the prestigious *Science* journal (Evans et al., 2021). Furthermore, it remains an open question, even when recognizing that introducing parental leave is beneficial for maternal health (Bütikofer et al., 2021), whether more leave is always better (Baker and Milligan, 2008b). Mechanisms for improving or deteriorating maternal health might vary along the leave duration. While physical relaxation from giving birth and breastfeeding have been shown to be dominant drivers for

E-mail address: caroline.chuard@unisg.ch.

[1] The author acknowledges funding from the UBS International Center of Economics in Society at the University of Zurich, Switzerland, the Jacobs Center for Productive Youth Development, and the Basic Research Fund of the University of St.Gallen. I am grateful to Maarten Lindeboom, two anonymous referees, Elisabetta Aurino, Michael Baker, Pietro Biroli, Patrick Chuard-Keller, Beatrix Eugster, Sébastien Fontenay, Johannes Kunz, Maya Rossin-Slater, Hannes Schwandt, Giuseppe Sorrenti, Tímea Virágh, Josef Zweimüller, and seminar participants at the University of Zurich, Mini-Workshop CDI, the 10th ifo Dresden Workshop on Labor Economics and Social Policy, the Essen Health Conference 2021, the ESPE 2021 conference, the EALE 2021 conference, and the joint doctoral Workshop on Applied Microeconomics of the University of Konstanz and the University of St. Gallen. All remaining errors are my own.

https://doi.org/10.1016/j.jhealeco.2023.102726
Received 25 January 2022; Received in revised form 20 December 2022; Accepted 2 January 2023
Available online 4 January 2023
0167-6296/© 2023 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

improving health when leave is short, these forces are unlikely to continue to exist after a certain amount of months on leave (Chung et al., 2007; Dagher et al., 2014). Instead, depreciation of human capital could start to play a role if mothers stay away from the labor market for too long.

Obviously, these are non-trivial questions to answer. On the one hand, only few countries have experimented with longer leave times. On the other hand, this type of question requires long time series of health data, which are often not available. Thus, Austria – having experienced three parental leave reforms from 1990 to 2000 and providing rich administrative data – offers a unique opportunity to study the effect of parental leave duration on maternal health in the short- (<5 years) to medium-run (>5 years).

In this study, I focus mainly on the reform in 2000, where maternal leave substantially increased from 1.5 years to 2.5 years, well beyond leave levels in many other countries. With strict enforcement of this policy change by July 1, this is an optimal setting to employ a regression discontinuity design. Furthermore, given a retrospective introduction for the reform, there is no need to worry about sorting around the thresholds. This eliminates any possibility of manipulating conception or delivery dates.[2]

Given the data availability, I evaluate the policy changes with maternal health outcomes from 1998 to 2007 for mothers who reside in Upper Austria. This allows me to observe mothers from 2 years pre-birth up to 7 years post-birth. I observe health outcomes in the Statutory Health Insurance Fund and provide information on all outpatient doctor visits, including prescriptions for medications and all inpatient hospitalizations, including diagnoses.

I find a negative impact of long leave of 2.5 years compared to leave of 1.5 years on maternal health. Specifically, long leave is substantially worse for maternal outpatient health so that it increases accumulated outpatient costs, the number of outpatient doctor visits, and the number of prescriptions by 0.990, 1.177, and 1.536 of a standard deviation, respectively.

In a dynamic analysis, where I follow mothers year by year, I show how effects emerge over time. Specifically, I document that these effects become significant only few years after birth. This may explain why studies that focus on outcomes measured relatively shortly after birth do not find any significant effects. As effects continue to accumulate over time, this suggests that the full extent of the reforms may not even be completely materialized at the end of my observation period.

An analysis of the type of medications prescribed and the type of doctors visited reveals that outpatient mental health may explain the results. Specifically, medications prescribed for the nervous system and against depression follow the described pattern. However, hospitalizations with a mental health issue, and thus likely more serious health issues, do not occur. Diagnoses and medications related to cardiovascular health are not significantly affected.

While I primarily study maternal health as outcome in this paper, I show that total fertility and labor market outcomes measured as being employed 5 and 10 years after giving birth, as well as daily wages are not affected for this sample of mothers. This aligns with the original work by Lalive and Zweimüller (2009), Lalive et al. (2013), and Kleven et al. (2020) who initiated this line of research when evaluating the same Austrian reforms with respect to maternal labor market outcomes and fertility.

In a mediation analysis, these types of additional outcomes, therefore, cannot explain the health effects. Also additional channels such as disposable income, alternative mode of childcare, or paternal involvement in child rearing may only be of secondary importance in explaining the negative health effects. If anything, the increased uptake of paternity leave could have a dampening effect on the negative health effects of mothers.

A heterogeneity analysis reveals interesting additional insights. According to the results, mothers with unhealthy babies as proxied by low birth weight are harmed less by longer leave than mothers with healthy babies. These types of *unhealthy* newborns may need extra time with their mothers and thus, a shorter leave may also be stressful for their mothers and deteriorate maternal health.

Relationship to Literature and Contribution. This article contributes in various dimensions to the literature on parental leave and family health.[3] First, I can study maternal health prior to birth and up to 7 years after birth. This allows me to understand the transition from an absence of significant effects in the short-run to large effects in the long-run. Baker and Milligan (2008b), for example, who evaluate an increase in parental leave in Canada from 6 to 12 months on maternal health up to 2 years after giving birth, find no significant effects. Instead, Beuchert et al. (2016), who study a reform in Denmark that increased maternity leave by on average 32 days, find small positive effects on the health of mothers who can be followed up to 5 years after giving birth. But even more so, the study by Bütikofer et al. (2021), which analyzes the introduction and several extensions of paid parental leave in Norway, reports large and positive effects on several health measures at age 40. With my study, I can confirm within the same country that the timing of health measurement matters.

Second, the substantial increase of parental leave from 1.5 to 2.5 years, well beyond typical levels of leave duration in other countries allows to explore the health gradient regarding parental leave duration. While Bütikofer et al. (2021) showed diminishing marginal returns to extensions of parental leave duration, I add to this finding by showing that these returns can even become negative. This highlights the importance of two key influencing factors: the initial level of parental leave and the extent to which parental leave is increased.

Third, while many works focus on self-reported health and health promoting behavior in survey data (Baker and Milligan, 2008b; Bütikofer et al., 2021; Chatterji and Markowitz, 2005, 2012), I can use very rich administrative data containing also new outcome variables. As such, I extend the analysis with detailed objective outcome measures both in the inpatient and in the outpatient sector, which are reported by the doctors themselves to the health insurance fund. For example, in the literature on the health effects of

[2] For more detailed information on the date of announcement see Lalive et al. (2013).

[3] In addition to the literature mentioned here, several public health studies document the correlation between maternity leave and mother's health as extensively summarized by Aitken et al. (2015), Andres et al. (2016), Staehelin et al. (2007), and most recently by Van Niel et al. (2020).

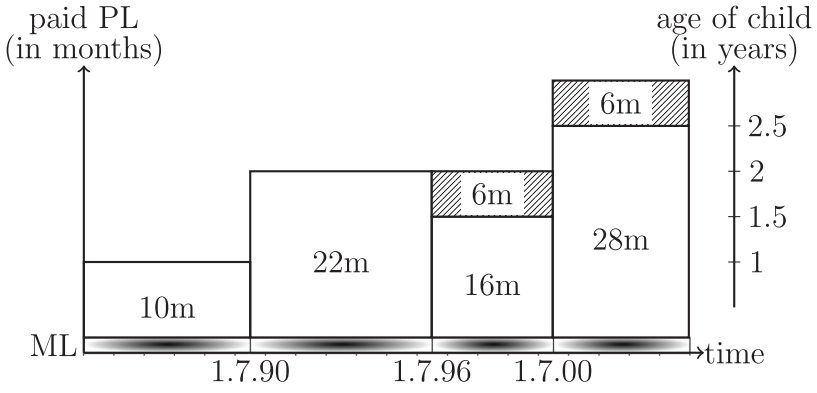


Fig. 1. Overview Austrian parental leave reforms. *Notes:* This figure shows an overview of the parental leave reforms starting from July 1990 to July 2000. The shaded area indicates the two months of mandatory maternity leave. Individual rectangles indicate the duration of post-birth parental leave, while dashed rectangles refer to those months of parental leave that are additionally and exclusively available to the second parent.

early retirement subjective and objective health measures lead to different results (Johnston and Lee, 2009) because of the classical justification bias. While such a bias may not exist in this setting, adding the dimension of objective health measures still completes the picture. Furthermore, the detailed level of the data allow me to study channels driving the main effects and to analyze the heterogeneous impact by characteristics of the child, such as the gender, the birth order, and the health of the newborn.

Finally, my study is also closely related to the literature on the evaluation of parental leave policies and other outcomes, such as child development and maternal labor market outcomes.[4] All of these margins have to be considered by policy makers in order to find the optimal parental leave design.

The rest of the paper is structured as follows. I start by describing the institutional setting where I both explain the Austrian family policy system and the health care system. In Section 3, I describe the empirical method and in Section 4 the data. Section 5 presents the results. I start by showing the baseline results for the reform 2000. I then look more deeply at the dynamics, study channels explaining the baseline results, analyze heterogeneous effects, conduct several sensitivity analyses, and provide some suggestive evidence on optimal leave policy by including results on the reforms in 1990 and 1996. A conclusion follows in Section 6.

2. Institutional background

In this section, I first describe the Austrian family policy system and the reforms on parental leave duration that I exploit in this paper. Then, I shortly discuss the Austrian health care system.

2.1. Austrian family policy system

Family policies have a long history in Austria. Traditionally, a familistic approach with relatively little professional childcare institutions available best describes the Austrian childcare system (Dörfler and Blum, 2014). This led to placing a high value to the families themselves and to protecting pregnant women early on. Pregnant employed women were not allowed to work during the first four weeks after giving birth, while getting a financial compensation during this time period as early as in 1888. Many changes followed the next century: introducing paid and mandatory pre-birth leave, extending paid and mandatory maternity leave, and introducing initially unpaid, ultimately paid voluntary leave up to the first birthday of the child.

Fig. 1 summarizes the most recent reforms in the Austrian parental leave system. Up to 30th of June 1990, mothers were entitled to eight weeks both pre- and post-birth mandatory maternity leave. While on maternity leave mothers are entitled to the full wage which they earned in the last quarter prior to giving birth. After mandatory leave, eligible mothers could choose to go on voluntary leave, which lasted at maximum up to the first birthday of the child. Mothers were guaranteed job protection and they were paid a flat benefit (indexed to wage inflation). The median replacement rate was about 40 percent of net earnings. First-time mothers over the age of 25 were eligible for parental leave if they had worked for at least 52 weeks in the 2 years preceding birth. The eligibility criteria are shorter for younger mothers and higher-order births.

As of 1st of July 1990, voluntary leave got extended up to the second birthday of the child. It was then defined as parental leave, so that fathers could take up to 6 months of the full parental leave duration. This policy reform resulted from parliamentary

[4] The same Austrian reforms that I use in this study are additionally to already summarized research also used in the following studies which evaluate other outcomes: Danzer and Lavy (2018) and Danzer et al. (2020) focus on cognitive outcomes of the affected children and Chuard (2020) studies the effect of maternal employment during pregnancy on newborn health. Maternal labor market outcomes are also studied in other country settings by Berger et al. (2005), Baker and Milligan (2008a), Rossin-Slater et al. (2013), Schönberg and Ludsteck (2014), and Bana et al. (2020) and child development is also studied among others by Ruhm (2000), Berger et al. (2005), Tanaka (2005), Baker and Milligan (2010), Liu and Skans (2010), Rasmussen (2010), Rossin (2011), Dustmann and Schönberg (2012), Baker and Milligan (2015), Carneiro et al. (2015), and Dahl et al. (2016).

procedural requests which wanted to introduce paternal leave. However, because of the flat benefit structure almost no fathers were taking up parental leave, which resulted in an increase of maternal leave from 1 to 2 years.

Out of budgetary constraints, six of the available 24 months of parental leave were exclusively reserved for fathers for births that occurred after 1st of July 1996. This effectively reduced maternal leave from 2 to 1.5 years, leaving the flat benefit structure and job protection unchanged. As of 1st of July 2000, parental leave got extended to 36 months (reserving six months to fathers). This reform was implemented in 2002, but was retroactively put in place for all employed mothers that gave birth after 1st of July 2000. Thus comparing employed mothers who gave birth before 30th of June 2000 versus after, leads to a difference in parental leave duration of 1 year. As a result, increasing parental leave from 18 to 30 months available to mothers.

Due to the substantial increase of the last reform from an already very generous level of leave duration from 1.5 years to 2.5 years, well beyond classical levels of leave duration in other countries, and the availability of data surrounding the policy reform itself, I will focus in this paper on the reform in 2000. Later in Section 5.6, I will get back to the reforms in 1990 and 1996.

2.2. Austrian health care system

Austria's health insurance system is organized as a Bismarck model, characterized by universal health care coverage financed jointly by employers and employees. Depending on occupation and place of residence, individuals are insured with 1 out of 9 regional health insurance funds. This mandatory health insurance, based on employment and residency, leaves no choice for individuals regarding the provider and the insurance package. Ultimately, every employee ends up with the same health insurance coverage characterized by a standardized catalogue of health care benefits.

The Upper Austrian Sickness Fund covers all inpatient and outpatient health services and the health care system is of high quality.[5] Patients are free to consult a medical specialist, but an initial visit to a general practitioner for a referral to a specialist is recommended.

Altogether, I argue that the Austrian health care system allows me to infer the employees general health status from the measure of health costs and health care utilization. This is both the result of there being a low entry barrier to use the Austrian health care system and the fact that pre- and post-reform mothers are not facing any different entry barriers to enter the system.[6] While the typical *time-argument*, i.e. if you have more time you might visit the doctor more, could induce mothers to use the health care system more when facing longer parental leave, this effect is less worrisome in this context given that health care outcomes are also observed for the years when mothers are already back in the labor market.

3. Empirical method

In this paper, I assess the average causal effect of duration of parental leave on maternal health. To do this, the identification strategy relies on exogenous policy-induced variation in leave duration based on the described reform. Therefore, I first test whether the change in maximum leave duration induced by the policy reform affected actual behavior of mothers (first stage). I then exploit this sharp discontinuous jump in leave duration at the policy reform thresholds in a regression discontinuity design. Mothers before and after the cutoff are supposed to be similar except for the parental leave duration they are entitled to. I estimate the following reduced form equation for the reform sample on eligible women using a meaningful bandwidth of 60 days[7]:

$$Y_{it} = \beta_0 + \beta_1 T_i + \beta_2 R_i + \beta_3 T_i * R_i + \epsilon_{it}, \tag{1}$$

with Y_{it} the health outcomes of mother i, t years after giving birth to her child; T_i an indicator for giving birth after the policy change (such that $T = 1$ for July-August mothers and $T = 0$ for May-June mothers); R_i the number of days from the policy reform (the so-called rating or running variable), and an interaction of the two latter variables to allow for separate trends on each side of the cutoff.[8] Here in this reduced form equation, β_1 identifies the net effect of changing parental leave length on maternal health.

As systematic differences in maternal characteristics by season of birth have been shown to exist in the US (Currie and Schwandt, 2013), I will additionally estimate a difference-in-differences regression discontinuity (RD-DiD) design.[9] Specifically, I interact the

[5] According to OECD statistics, roughly 10.4% of GDP were spent on health care in 2019 ranking it among the top 10 countries. At the same time, life expectancy at birth was with 82 years ranked in the top 20 countries.

[6] Furthermore, using a regression discontinuity design to assess the impact on overall maternal health measured by health care utilization, ensures the same type of measure on both sides of the cutoff. Specifically, using health care utilization as an objective proxy for actual health is only problematic if doctors would treat pre- and post-reform patients (mothers) differently. This is highly unlikely, especially considering that insurance data is observed up to seven years after the policy reform where prescribing doctors will not recover the duration of parental leave of their patients.

[7] I estimate local linear regressions with a triangular kernel (as suggested in Hahn et al., 2001) and separate trends on both sides of the cutoff. Because of the concerns raised by Kolesár and Rothe (2018), I do not cluster standard errors by birth day but use heteroskedastic-robust standard errors as suggested in Lee and Lemieux (2010). In the robustness analysis, I vary the bandwidth and the functional form. Generally, the optimal bandwidth selector for each outcome as suggested by Calonico et al. (2018) and shown in Table A.1 reports optimal bandwidths in the range of 35 to 58 days.

[8] In a robustness analysis, I also include control variables, which given the RD setting, however, do not affect results.

[9] Such seasonalities exist in the Austrian context too. However, the pattern is much more noisy when compared to the US context. The small seasonalities as shown in Fig. A.1 in Austria point toward a positive selection of mothers in early spring, meaning that mothers are slightly more likely to be married, of high SES, and less likely to be foreign when giving birth to their child from March to June. These months translate into a conception date between July and October, which are exactly the *favorable* conception months, as also shown by Currie and Schwandt (2013). However, these seasonalities are much less pronounced when studying a single year (such as the reform year in 2000) and also when studying a shift from a single day to another one instead of analyzing monthly changes. Graphical results are available upon request.

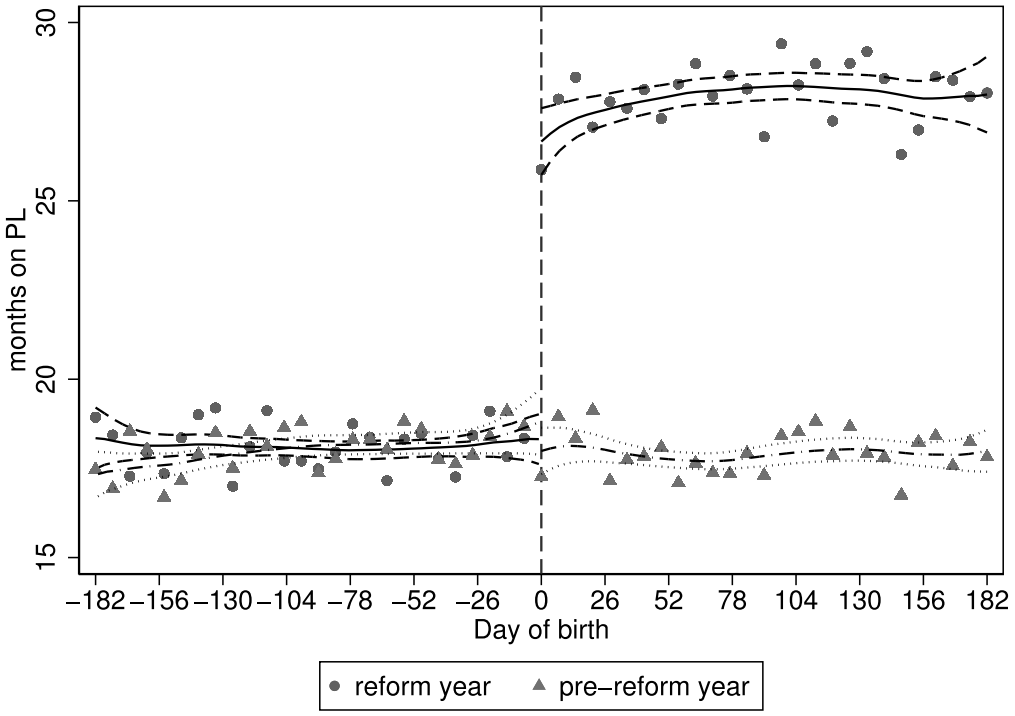


Fig. 2. Discontinuity in actual parental leave duration. *Notes:* This figure shows the average amount of months eligible mothers spend on leave by the birthday (organized in one-week bins) of the child relative to the policy change in July 2000 (normalized to zero and indicated by the vertical dashed line). Dots refer to the reform year and, therefore, depict a discontinuity in July when the policy reforms were implemented. Instead, triangles refer to the preceding non-reform year (1999) and are, therefore, stable across the year. The solid/dash-dotted lines represent fitted values from a local linear regression in the reform/preceding non-reform year and the dashed/dotted lines show the corresponding 95 percent confidence interval.

coefficient of interest with the pre-reform year. This approach allows to control for differences in outcomes between mothers who gave birth after versus before the reforms in July that are unrelated to the reform itself. However, because of the structure of the data, i.e. a limited time window from 1998 to 2007 and health effects that depend on the time since mothers gave birth to their child, I have to disregard one year of information with the RD-DiD setting. For this reason, and also because the standard RD design estimates a local average treatment effect (LATE) where maternal characteristics at the daily level are unlikely to change discontinuously at the threshold, I define the standard RD as my baseline model.

Identification. There are two identifying assumptions in the baseline setting, and a third one in the DiD extension: (1) There has to be a sharp treatment discontinuity at the cutoff (Lee and Lemieux, 2010); (2) Instead, there should be no discontinuity in potential outcomes at the cutoff (Lee and Lemieux, 2010); (3) Due to adding the difference-in-differences structure, I have to assume that month-of-birth effects do not vary across years.

Assumption (1) is fulfilled, as shown in Fig. 2 and in Table 1. Specifically, Fig. 2 depicts how the policy change in maximum leave duration affected actual behavior, i.e. the months on parental leave. Dots describe weekly averages in the reform year 2000 and triangles depict the non-reform year 1999 preceding the reform. Post-reform mothers who give birth in July show a significant jump in months on parental leave. For example, July 2000 mothers stay on leave for almost 27 months versus 18 months as the June 2000 mothers (+46%). Generally, mothers almost fully exploit the available parental leave duration and, thus, show a high norm for maternal care (Kleven et al., 2020). In the non-reform year, there is no discontinuity. Table 1 additionally shows that taking out the seasonality by including the pre-reform year does not strongly affect the point estimate.

There are standard procedures to validate the untestable assumption (2). I test for discontinuities in non-outcome variables, which I report in Table A.2. Both maternal and child characteristics evolve smoothly across the cutoff. As a result, RD estimates are likely to generate unbiased estimates. In order to test assumption (2), I also show the density of the running variable and implement the recommended McCrary test for manipulation in Fig. A.2 (McCrary, 2008). None of the manipulation testing procedures suggest that there is systematic manipulation of mothers to be on either side of the cutoff. This result is also not surprising since the 2000 reform was announced after implementation so that systematic sorting around the threshold was impossible.

Finally, assumption (3) seems relatively mild and reasonable, as I am only using the pre-reform year. It is very unlikely that month-of-birth effects change over such a short time horizon.

Table 1
Reform effects on parental leave length.

	Parental leave length	
	(1)	(2)
Post	8.335***	8.997***
	(0.623)	(0.397)
RD	✓	
RD-DiD		✓
Mean of Dep. Var.	18.162	18.162
Observations	3187	6482

Notes: This table provides parameter estimates for the coefficient β_1 of Eq. (1) for the reform sample. The sample includes all PL eligible mothers within a bandwidth of 60 days around the policy cutoff in July 2000. Column (1) estimates standard RD estimates, while column (2) incorporates a DiD layer by including the pre-reform year to difference out seasonality which has nothing to do with the reform itself. Heteroskedastic-robust standard errors are shown in parentheses and significance levels are indicated by */**/*** for the 10%/5%/1% level.

4. Data

My empirical analysis uses three separate administrative data sets: The Austrian Social Security Database (ASSD), the data from the Statutory Health Insurance Fund (SHIF) of Upper Austria, and the Austrian Birth Register (ABR). The social security database includes daily information on the full working history of the universe of private sector employees, birth events, and parental leave spells.[10] The data also contains information on individual characteristics, such as the gender and age, and on job characteristics, such as yearly earnings per employer. Based on an individual identifier the data can be linked to health outcomes for private sector employees living in the state of Upper Austria from 1998 to 2007.[11] The data set includes information on all covered health services, such as health care utilization in the outpatient sector (i.e. doctor visits and prescriptions) and diagnoses in the inpatient sector. All expenditures are reported in real prices expressed in year 2000 euros. Finally, I complement these data with additional information on maternal characteristics at the time of birth (such as her education and marital status), the child's health at birth (gestational length, birth weight, and Apgar scores), and additional birth characteristics (sex of child and birth order) from a matching procedure with the Austrian Birth Register.

In order to analyze the parental leave reforms and the effect on health, I consider the following outcome measures:

(i) *Outpatient health care expenditures* measure costs for all consultations with doctors and costs for prescribed medical drugs.
(ii) *Outpatient doctor visits* report the distinct number of days a patient saw a doctor in the outpatient sector.
(iii) *Prescriptions* inform about the type of medication according to the Anatomical Therapeutic Chemical (ATC) Classification System. With the help of this pharmaceutical coding system, I create variables for the total number of packages consumed of prescriptions in general, and an indicator for having ever been prescribed any of the following specific medications: antidepressants, general nervous system medication, analgesics, and cardiovascular system medication.
(iv) *Days of hospitalizations* give detailed information about each inpatient hospital visit.
(v) *Inpatient diagnoses* for hospitalizations are classified according to the ninth revision of the International Classification of Diseases and Related Health Problems (ICD-9). I will put a special emphasis on general mental disorders, depression, and diseases of the circulatory system. The previous literature on where to expect long-term effects of parental leave guides these choices.

Given the various dimensions of the studied health outcomes, I use z-scores for reporting the health effects by subtracting the mean and dividing by the standard deviation within each year of observation and reform sample. Furthermore, I use accumulated annual measures at each point in time and at the end of the observed time window to have less noisy measures for health care utilization, and to get information about health dynamics. In my baseline results, I only report estimates for the last point in time, i.e. accumulated outcomes over ten years.[12] This means that I will show results for mothers who gave birth to their first child 7 years ago.

In this study, I focus on private sector employed mothers excluding civil servants, self-employed, and farmers due to data restrictions in the ASSD. This limits the sample to roughly 85 percent of the overall Austrian population. I include mothers between the ages of 15 to 45 with a single (instead of multiple) birth of any birth order. Furthermore, I restrict my sample to those mothers

[10] The main collection purpose of this data is to document an employee's working history in order to calculate old age social security benefits.

[11] Upper Austria is one out of nine regions in Austria. It is with one sixth of Austria's total population the third most populous state.

[12] In the RD-DiD setting outcomes are accumulated over nine years. Health outcomes are *only* available from 1998 to 2007. Thus, when I estimate an RD-DiD, I lose 1 year of information for each individual. This arises because I want mothers to be comparable in both samples (the pre-reform year and the reform year) which means to have the same amount of years since giving birth. As an example, for the 2000 reform sample, I observe mothers 2 years prior to giving birth, while I only start observing them 1 year prior to giving birth for the 1999 sample. Therefore, I have to drop the first year for the 2000 sample and the last year for the 1999 sample.

Table 2
Descriptive statistics.

	Mothers in		Eligible	Reform sample	
	Austria (1)	Upper Austria (2)	for PL (3)	+/− 60 days (4)	+ pre-reform yr. (5)
Panel A: Maternal characteristics at birth					
Eligible for pl (in %)	78.14	79.26	100	100	100
Months on PL	17.66	18.18	19.82	22.80	20.44
	(8.754)	(8.798)	(7.058)	(8.809)	(8.263)
Age of mother	27.19	27.12	26.94	28.28	28.24
	(5.002)	(4.812)	(4.760)	(4.865)	(4.928)
Married (in %)	70.99	69.86	66.95	62.35	62.00
Mandatory schooling only (in %)	22.49	25.19	23.80	16.13	17.54
High SES (in %)	44.65	43.84	45.81	52.37	50.82
Foreign (in %)	12.92	9.689	9.111	11.92	12.23
Panel B: Child characteristics					
Birth weight	3325.6	3351.2	3339.3	3365.9	3368.0
	(519.3)	(510.0)	(510.3)	(526.1)	(530.2)
Low birth weight (in %)	4.434	3.906	4.062	4.801	4.674
Gestation length (in weeks)	39.74	39.75	39.74	39.67	39.68
	(1.759)	(1.703)	(1.718)	(1.850)	(1.856)
Preterm (in %)	4.061	3.757	3.846	4.612	4.443
Healthy Apgar score (in %)	95.78	96.18	96.07	95.89	96.07
Girl (in %)	48.67	48.72	48.72	49.64	49.01
Birth order	1.736	1.761	1.651	1.672	1.683
	(0.864)	(0.870)	(0.797)	(0.820)	(0.823)
Panel C: Maternal health characteristics (annual measures)					
Outpatient costs (in € 2000)		236.5	245.5	260.0	263.3
		(518.8)	(534.0)	(467.8)	(486.8)
Outpatient doctor visits		6.550	6.767	7.360	7.443
		(7.194)	(7.221)	(6.721)	(7.334)
Medication (# of packages)		3.280	3.377	3.278	3.418
		(7.163)	(6.916)	(4.802)	(5.605)
Hosp. (in nights)		1.041	1.052	1.619	1.637
		(4.604)	(4.660)	(4.811)	(4.791)
Observations	761,588	163,835	129,848	3187	6482

Notes: Column (1) covers all births in Austria from 1989–2000, while column (2) focuses on those mothers residing in Upper Austria for the entire period from 1998–2007. Column (3) restricts column (2) to mothers who were eligible for PL, and columns (4–5) show the descriptives for the samples used in the RD and RD-DiD analysis, respectively. Standard deviations are shown in parentheses if necessary. PL stands for parental leave and SES for socioeconomic status. A high-SES mothers is defined as a mother who has completed at least high school or who finished apprenticeship training or intermediate vocational school and earns a wage above the median wage; Low birth weight/preterm is a dummy measure equaling 1 if the birth weight/gestation length is below 2500 g/37 weeks; Healthy Apgar score is a dummy measure equaling 1 if the 5 min Apgar score is at least 7 out of 10.

that can be observed over the entire time window from 1998 to 2007 (around 95 percent of the remaining mothers). There are several reasons why mothers could leave the sample – mothers can become self-employed or mothers can move out of Upper Austria.

Descriptives. Table 2 reports summary statistics. The first column presents means and standard deviations for the overall sample of mothers in Austria who have a live birth between 1989 to 2000. The second column restricts the sample to those mothers which reside in Upper Austria from 1998 to 2007. They make up roughly one fifth of the births from the overall population and seem to be representative of Austria as none of the observable characteristics differ from the sample in column (1). The third column restricts column (2) to the 79% of mothers who are eligible for parental leave. Thus, the months on PL are rising significantly when moving from column (2) to column (3). However, for the rest of the variables, eligible mothers are relatively comparable to all mothers in Upper Austria. Finally, columns (4) and (5) are used in the RD and RD-DiD analysis, respectively. They restrict column (3) to the 60 days prior to and post the policy reform in 2000 for column (4) and including the pre-reform year 1999 for column (5).

In my baseline specification, I focus on column (4) – the reform sample for eligible mothers in Upper Austria. In this sample, mothers are on average 28 years old when giving birth to their child, 16% of them have mandatory schooling only, and 12% are foreigners. Their children weigh on average 3366 g at birth, while 5% of them are being born with a low birth weight or preterm. The average mother in the reform sample needs 3 prescribed medications, visits doctors in the outpatient sector 7 times, spends 2 nights in hospital, and accumulates EUR 260 outpatient health care expenditures per year. Reform mothers as reported in column (4) confirm general population trends when compared to overall mothers in Austria from 1989–2000; they tend to be older at birth, are less likely to be married, and are more likely to be better educated.

Table 3
Reform effects on general health outcomes.

	Outpatient costs (1)	Outpatient doctor visits (2)	Medication (# packages) (3)	Days of hospitalization (4)
Post	0.990**	1.177**	1.536***	0.212
	(0.403)	(0.538)	(0.552)	(0.298)
Mean of Dep. Var.	0.131	0.038	−0.070	0.075
Observations	3187	3187	3187	3187

Notes: This table provides RD parameter estimates for the coefficient β_1 of Eq. (1) on general health outcomes. Health outcomes are standardized and accumulated over time and shown for the last observable year in 2007. The sample includes all PL eligible mothers within a bandwidth of 60 days around the policy cutoff. Heteroskedastic-robust standard errors are shown in parentheses and significance levels are indicated by */**/*** for the 10%/5%/1% level.

5. Results

In this section, I start by reporting the baseline results. As I can exploit the panel structure of the data and follow mothers over time, I then report health dynamics to understand how the baseline results develop over time and at what time after birth significant effects emerge. In a next step, I explore mechanisms driving the baseline results, such as specific causes and prescriptions; and alternative channels such as fertility and spacing, labor market outcomes, disposable income, alternative mode of care, and paternal uptake of parental leave. These results will be completed by a heterogeneity analysis and several robustness checks, including covariates in baseline estimations, playing with different bandwidth choices and functional forms, and reporting placebo estimates. Finally, by additionally considering the reforms in 1990 and 1996, I will provide suggestive evidence for optimal policy.

5.1. Baseline results for the 2000 reform

I present estimates on different general maternal health outcomes in Table 3 and in Fig. 3.[13] The general health outcomes that I consider are: accumulated health costs in the total outpatient sector, accumulated doctor visits in the outpatient sector, accumulated packages of prescribed medications, and the accumulated number of days of hospitalizations (inpatient) at the end of the observation period (all standardized). RD estimates, based on Eq. (1), report the difference in health outcomes for mothers giving birth to their child after the policy change in 2000.

The 2000 reform significantly increases overall outpatient costs by 0.990 of a standard deviation, outpatient doctor visits by 1.177 of a standard deviation, and the number of prescribed medication packages by 1.536 of a standard deviation, respectively. While the sign on the days of hospitalizations point toward a positive impact too, this effect is not significantly different from zero. There are several explanations for this: (1) There truly is no effect (e.g., because the increased contact with health care professionals in the outpatient sector can prevent worse health, such as hospitalizations in the inpatient sector); or (2) the positive effect cannot be statistically proven (e.g., because the number of hospitalizations is already very low in the baseline, with an average of 2 days per year versus about 7 doctor visits and 3 prescribed packages of medication per year, so that it is very difficult to detect a significant difference here).

In the RD-DiD specification as reported in Table A.3, where I control for seasonality, the direction of the effects stays the same but looses its significance for the two first outcomes. However, because of the availability of the data, the RD-DiD specification lacks one year of information and with the increase of the effect over time, as I will show later, this might explain the non-significant effects in some of the cases.

5.2. Health dynamics

Fig. 4 reports the dynamics of the general health outcomes over time and shows an estimate for each observed year for the RD specification. Outcome measures are accumulated up to the stated year. Thus, the last dot in each subfigure corresponds to the estimate reported in Table 3 before. Confidence intervals are shown with dashed lines.

For the 2000 reform, one can observe mothers already two years *before* giving birth. This enables a further test for the appropriateness of the RD specification, as pre- versus post-reform mothers should not differ prior to giving birth, i.e. prior to being affected by the policy change. As can be seen, this test holds true for outpatient doctor visits, prescriptions, and hospitalizations. Furthermore, one has to keep in mind that pre- versus post-reform mothers are on leave for 1.5 versus 2.5 years, respectively. As such, it is interesting to see whether the health outcomes diverge while post-reform mothers are still on leave and pre-reform mothers started working already, i.e. in the years 1.5 to 2.5 after giving birth. While this does not seem to be the entire driver of the overall pattern, it is true that for outpatient doctor visits and prescriptions, there is a somewhat steeper increase in the effect in

[13] Table A.3 shows RD-DiD estimates. RD-DiD estimates additionally control for seasonality unrelated to the policy change, as I difference out seasonality effects by including the pre-reform year.

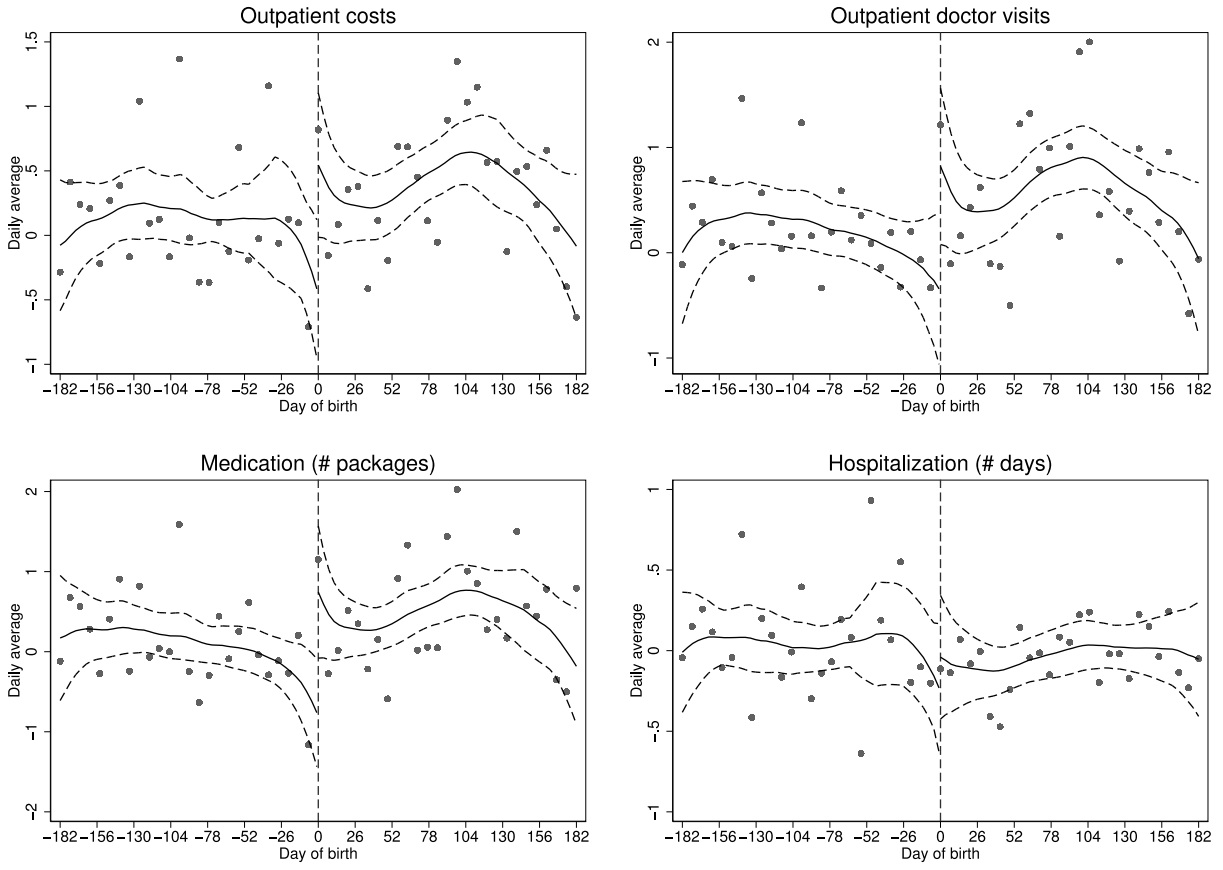


Fig. 3. RD plots on general health outcomes. *Notes:* This figure shows the average standardized accumulated health utilization (black dots) by the birth day (organized in one-week bins) of the child relative to the policy change in July 2000 (normalized to zero and indicated by the vertical dashed line). The solid lines represent fitted values from a local linear regression in the reported reform year and the dashed lines show the corresponding 95 percent confidence interval.

the years 1 to 2 after giving birth. At this point in time post-reform mothers are still on leave and might have more time to see a doctor. However, for all outpatient measures, especially, for doctor visits and prescribed medications, the effect increases over time and still has a steep increase in the last year observed. While the effect on hospitalizations is relatively variable, one can suppose that more doctor visits and prescriptions prevent a higher rate of hospitalization.

The general take-away from looking at the dynamics is that health effects accumulate over time. Only looking at the time in which some mothers are still on leave and others started working does not fully capture all the health effects of parental leave reforms. This can also explain the effects that have already been found in the literature. Baker and Milligan (2008b), for example, focused on health outcomes 1 year after birth and could not detect any significant effects. However, this may just have been a too short time window. Instead, the strong and positive effects of longer leave reported by Bütikofer et al. (2021) at 40 may capture the fact that the effects on maternal health increase over time and may even be higher at older ages. Mothers analyzed with the 2000 reform, for example, are on average 35 years old in my analysis. These mothers still show a steep and positive slope of the health effects, suggesting that health effects are still accumulating over time at that age.

5.3. Channels

In a next step, I examine drivers of the general health effects. With the analysis of specific health outcomes, I want to shed light on potential mechanisms. Furthermore, a differential impact of the policy reforms on fertility, spacing, and maternal employment could drive the health effects. Therefore, I present a mediation analysis to understand the relative importance of these mechanisms. I, then, study the role of disposable income and alternative mode of care before I analyze the impact of paternal leave. While only few fathers can be linked to mothers, I cannot comprehensively examine whether paternal uptake of parental leave drives maternal health effects. Nevertheless, I show some suggestive evidence that the involvement of fathers, if anything, prevents mothers from even worse health outcomes.

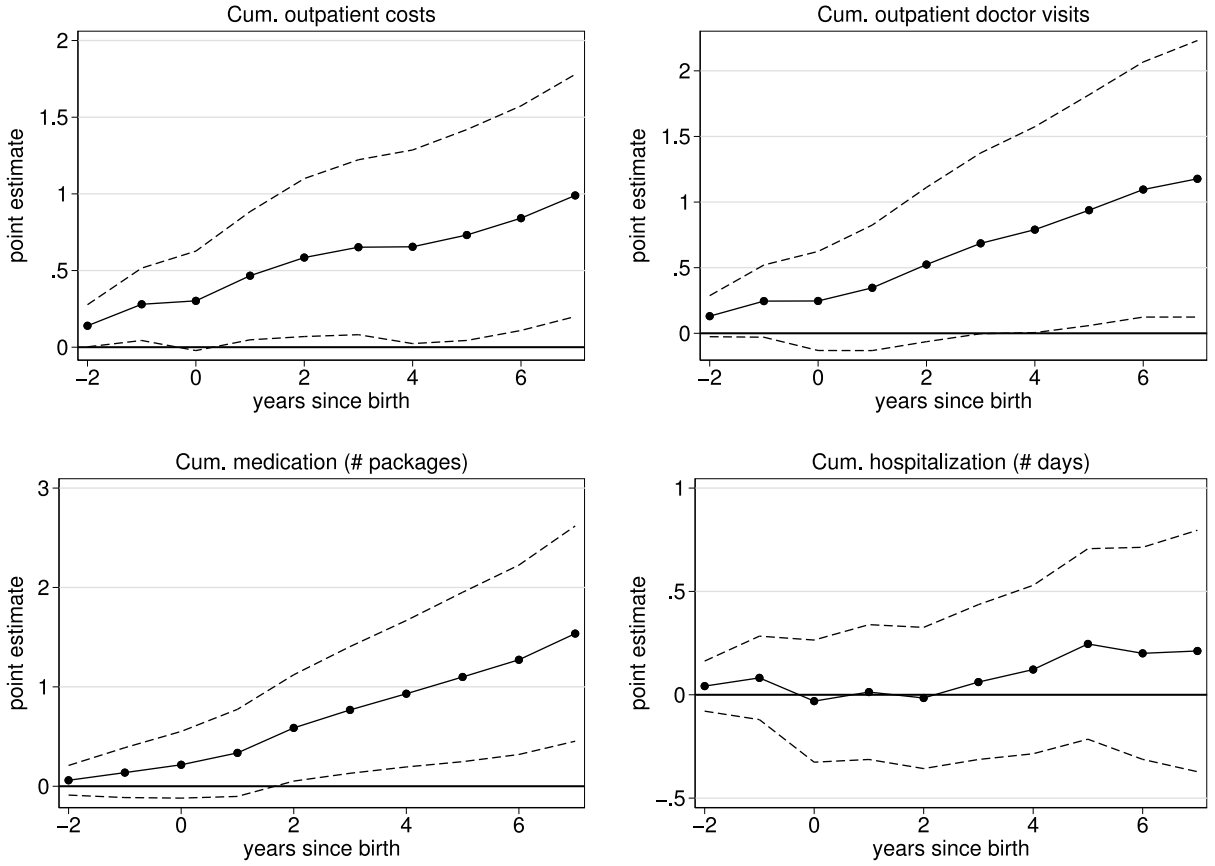


Fig. 4. Health dynamics. *Notes:* This figure shows RD parameter estimates for the coefficient β_1 of Eq. (1) on different standardized accumulated health outcomes. Each dot represents a separate estimation on accumulating the health outcome up to this point in time. The sample includes all PL eligible mothers within a bandwidth of 60 days around the policy cutoff. Dashed lines report 95 percent confidence intervals.

Maternal mental and cardiovascular health. Tables 4–5 report detailed health outcomes. Specifically, I show detailed effects on maternal mental and cardiovascular health. These health domains are based on the findings of the previous literature. For example, postpartum mental health is a common concern in the public health literature. Estimates from systematic reviews for the prevalence of depressive episodes postpartum range from 11.9 percent to 19.2 percent with slightly lower prevalence in Europe than other parts of the world (Gavin et al., 2005; Woody et al., 2017; Shorey et al., 2018). As such, the likelihood of depressive episodes is significantly higher in the postpartum period than during other periods of a woman's life (Shorey et al., 2018) which poses a significant health risk not only for the mother herself but for other family members, too. Importantly, Shorey et al. (2018) also state that postpartum depression prevalence is increasing beyond six months postpartum and that mothers with depressive postpartum symptoms have a higher likelihood of recurrence of depressive episodes in the future (Miller, 2002), posing a high risk for long-term maternal health.

Furthermore, pregnancy has often been found to act as contributor to overweight in women. While average weight retention postpartum ranges from 0.5 to 3 kg, it may be as high as 17.7 kg (Gore et al., 2003). Obviously, this poses a significant health risk for mothers given the association of being overweight or obese with health risks such as cardiovascular diseases and diabetes (Rooney et al., 2005).

Table 4 shows estimates on mental health outcomes both in the inpatient (mental disorders and depression) as well as in the outpatient (prescriptions for antidepressants, nervous system medication, and analgesics) sector. Table 5 presents cardiovascular health outcomes, such as general hospital diagnoses categorized as circulatory system diseases and prescriptions related to the cardiovascular system.

The estimates on maternal mental health confirm the previously found results. The increase in parental leave duration to a very generous level of 2.5 years in the year 2000 significantly deteriorated maternal mental health. The latter effects are large in magnitude: the number of prescriptions for anti-depression drugs rises by 0.936 of a standard deviation and the one for nervous system medication by 0.955 of a standard deviation.[14] Although prescriptions for mental health issues increase significantly,

[14] The effects are robust to adjusted inference for multiple hypothesis testing and remain significant at the 10 percent level. Specifically, I use the method proposed by Romano and Wolf (2005), which is an iterative procedure that controls for the type I error rate within a family of outcomes at a fixed level of

Table 4
Estimates on mental health outcomes.

	Inpatient health		Outpatient health		
	Mental disorders (1)	Depression (2)	Antidep. drugs (3)	Nervous sys. drugs (4)	Analgesics (5)
Post	0.184	−0.072	0.936**	0.955**	0.478
	(0.311)	(0.290)	(0.470)	(0.447)	(0.362)
Mean of Dep. Var.	−0.006	−0.029	−0.353	−0.326	−0.080
Observations	3187	3187	3187	3187	3187

Notes: This table provides RD parameter estimates for the coefficient β_1 of Eq. (1) on mental health outcomes. The first two columns refer to health outcomes in the inpatient sector (hospitalizations) while columns (3)–(5) measure health outcomes in the outpatient sector (prescriptions). Health outcomes are standardized and accumulated over time and shown for the last observable year in 2007. The sample includes all PL eligible mothers within a bandwidth of 60 days around the policy cutoff. Heteroskedastic-robust standard errors are shown in parentheses and significance levels are indicated by */**/*** for the 10%/5%/1% level.

Table 5
Estimates on cardiovascular health outcomes.

	Diseases of circulatory system (1)	Cardiovascular drugs (2)
Post	0.116	0.566
	(0.273)	(0.461)
Mean of Dep. Var.	−0.068	−0.064
Observations	3187	3187

Notes: This table provides RD parameter estimates for the coefficient β_1 of Eq. (1) on cardiovascular health outcomes. Column (1) refers to health outcomes in the inpatient sector (hospitalizations) while column (2) measures health outcomes in the outpatient sector (prescriptions). Health outcomes are standardized and accumulated over time and shown for the last observable year in 2007. The sample includes all PL eligible mothers within a bandwidth of 60 days around the policy cutoff. Heteroskedastic-robust standard errors are shown in parentheses and significance levels are indicated by */**/*** for the 10%/5%/1% level.

hospitalizations for mental health issues are not affected. As for general health measures, this means, that more intensive access to outpatient healthcare may prevent more severe health issues that would be observable in the inpatient sector.

For cardiovascular health outcomes, which are reported in Table 5, there appears to be no systematic effect. There are several possible reasons for not finding an effect. While there is limited causal evidence on the impact of breastfeeding on maternal health, some non-experimental public health papers (Chowdhury et al., 2015; Nguyen et al., 2017; Feltner et al., 2018) show a positive impact of longer breastfeeding duration on maternal health via a reduction in maternal weight and blood pressure. However, it is important to keep in mind that the 2000 reform on the duration of parental leave in this context was well above the WHO-recommended duration of breastfeeding of 6 months. Thus, the cardiovascular health channel might play a smaller role in the presence of generous parental leave (Lackey et al., 2021). Furthermore, given that mothers in my sample are around 35 years old on average when outcomes are measured, it may be that more time must elapse for cardiovascular effects to emerge. In particular, it is generally true that negative cardiovascular outcomes do not occur until older ages, so the mothers in my sample may be too young to have experienced significant effects.

Fertility, spacing, and employment. Previous studies (Kleven et al., 2020; Lalive et al., 2013) analyzing the same Austrian reform in the year 2000, did not detect impacts on completed fertility and maternal labor market outcomes in the medium to long run. Specifically, if fertility were to be affected positively, post-reform mothers would spend more time in the hospital to give birth which would affect one of the outcome measures directly. Thus, redoing the same exercise for the sample of mothers observable in the health data, I corroborate these previous findings as shown in Table A.4. Neither total fertility nor labor market outcomes such as employment status and daily wages measured 10 years after giving birth are affected by any of the policy changes. Unsurprisingly, adding fertility and employment status at the end of the observation period, thus, testing the strength of the mechanisms by adding them one at a time as reported in columns (2)–(3) in Table A.5 and jointly in column (4) in a *horse race*, does not alter the previously stated results (column 1) for outpatient costs, outpatient doctor visits, and medications in size and significance. While the effect on days of hospitalization half in size, they are still not significantly different from zero.[15]

significance. The family of outcomes, in this specific setting, are inpatient vs. outpatient health measures. Thus, I group mental disorders and depression into the family of inpatient health measures and anti-depression drugs, nervous system drugs, and analgesics into the family of outpatient health.

[15] I report all results in reduced form as a net effect of the policy change on maternal health. Doing a two-stage least squares (2SLS) regression analysis – where I insert the prediction of parental leave duration in months (second stage) on maternal health from a regression of the policy changes on parental leave duration (first stage) – needs the additional assumption that all effects on maternal health work entirely through the change in parental leave duration. This

Moreover, not only is completed fertility unaffected by the 2000 reform, but the same is true for the interval between two births (birth spacing). Given the retrospective introduction of the 2000 reform, the age difference from the focal child to an older sibling could by construction not have been affected. Additionally, as shown in previous work the timing to a younger sibling was also not altered (Chuard, 2020).

Disposable income. Since parental leave benefits are flat, they essentially increase disposable income for some mothers while decreasing it for others. For mothers who would have stayed home anyway when their child is between 1.5–2.5 years old disposable income increases. As a result of the reform, mothers who chose to return to work after 1.5 years either had a higher or lower disposable income depending on their earnings. For a similar reform in 1990, Danzer et al. (2020) estimate the share of mothers who benefit financially from the reform to equal 65 percent of the full sample. Thus, to check for a differential impact, I conduct a simple subgroup analysis by studying the effect for mothers who earn below or above median wage prior to giving birth. The latter group would be expected to benefit less from the reform due to the flat benefit structure and higher initial wages. However, even mothers with high wages who decide to go back to work need to arrange childcare so that the overall effect on disposable income is less clear.

Results can be found in Table A.7. While results are not significantly different for the two subgroups, the sign of the interaction term for mothers earning above the median is in fact even negative, and, thus, pointing towards a less detrimental impact on maternal health for longer leave. However, this effect is relatively small in size and not significantly different from zero which suggests that disposable income may only play a minor role. This finding is (1) in line with Danzer et al. (2020) and (2) not surprising given the previously stated results. While there were no effects in the medium to long run on maternal employment, long-term disposable income cannot be affected. Health dynamics, instead, showed accumulating effects over time even after the official end of parental leave. Thus, when mothers are already back at work and on their unaffected employment path, health is still impacted which cannot be explained by disposable income since it is unaffected in the long run.

Alternative mode of care. Previous to the reform, parents could arrange informal childcare, send their child to formal childcare, or decide to stay at home. In regions without nurseries mothers have to rely on informal care if they decide to go back to work. Danzer et al. (2020) – studying the 1990 reform which also increased parental leave by 1 year – show that children in regions without formal childcare available, where mothers most likely switch from informal care to maternal care, benefit in various dimensions but especially in terms of health outcomes.

The proportion of 0–2 years old children in institutional childcare facilities in Upper Austria was in the year 2000 with 4.1 percent below the overall Austrian average of 7.7 percent, which would imply a larger shift from informal to maternal care.[16] Thus, transferring the results of Danzer et al. (2020) to this study and assuming mothers value their child's health positively would suggest that the negative effects on maternal health could even be worse if health outcomes of children were to be positively impacted at the same time. Obviously, the findings of Danzer et al. (2020) might not translate to the 2000 reform, since the harmful impact on maternal health as found in this study could also negatively impact their child directly.

Paternal leave. With the data sets at hand, only a fraction of mothers can be linked to their partner. A connection between partners arises if one partner has ever been insured via the other, for example, because of being out of the labor force. The set of fathers that can be linked is, thus, not representative and equals roughly 14 percent for the 2000 reform. For none of these fathers, there is a significant jump in the likelihood of going on parental leave as can be seen in Table A.8.

Looking at the general male population and the absolute number of men starting a paternal leave spell in a given month, as depicted in Fig. 5 for the years 2000–2004, shows a slightly different pattern. While paternal uptake of parental leave is low in the early 2000s (below 3 percent of the overall births), there has been a steady increase over the years. This might have led to a larger involvement of fathers in raising their children. Other studies have shown that an introduction of paternal leave – even for brief periods – improves marital stability and child outcomes and increases fathers' involvement in childcare and housework thus reducing sex specialization (Cools et al., 2015; Fontenay and Tojerow, 2020; Olafsson and Steingrimsdottir, 2020; Patnaik, 2019; Persson and Rossin-Slater, 2019; Tamm, 2019). With this second set of data I cannot directly link mothers and fathers, so that one can only hypothesize that the results found in the literature would translate to better maternal health in the Austrian setting. This would, however, mean that the detrimental impact on maternal health as found in this paper would even be worse without paternal uptake of parental leave.

5.4. Heterogeneity

In this section, I study whether the effects of the reform varied with characteristics of the child such as the birth order (first birth), the gender of the child, and the health of the newborn (born with low birth weight or preterm). I report results for general health outcomes in Table 6. Interestingly, the first stage, i.e. the policy effect on the duration of parental leave does not vary by these characteristics except for first-time mothers choosing a significantly shorter PL duration as can be seen in Fig. A.3. As such, results can mostly be interpreted as the impact of the specific group considered and not as the result of a different first stage. This first-stage result is interesting in itself, as it highlights again the strong social norm for maternal care (i.e. almost fully exploiting policy-induced parental leave duration).

---

is a strong and non-testable assumption. With the small and non-significant effects on the most likely affected alternative outcomes as reported in Table A.4, I also show results from a 2SLS regression in Table A.6. The interpretation of the results remains essentially unchanged.

[16] Data retrieved from Statistik Austria: Kindertagesheimstatistik 2019/2020.

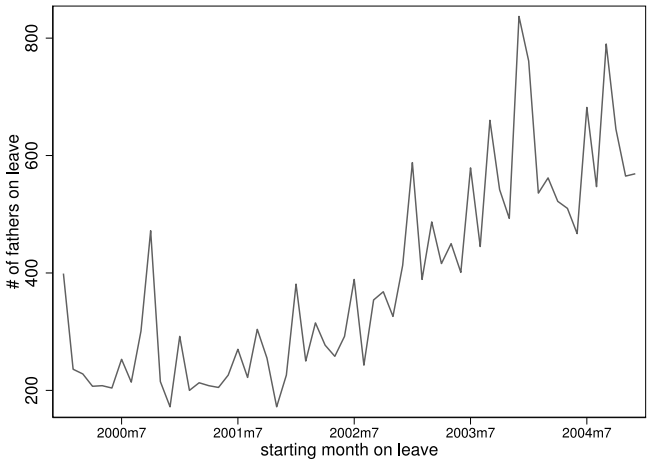


Fig. 5. Paternal uptake of parental leave. *Notes:* This figure shows the total number of fathers taking up parental leave (any duration) by the starting month of each parental leave spell. The sample includes the full set of males observable in the Austrian Social Security Database and not only those living in Upper Austria.

Specifically, I augment Eq. (1) by interacting all parameters with one of the heterogeneous subgroups, including a dummy for the group itself, and report only the parameters of interest, i.e. the effect of the policy change and the interaction term. The interaction term reports the additional effect (added to the baseline RD estimate) for the stated heterogeneous group, i.e. for first-time mothers in the first panel. Given the already relatively small sample size and some interactions with unlikely events (for example a low birth weight baby), results have to be interpreted with care.

There is no significant difference for first-time mothers compared to non-first-time mothers and the interaction term of the policy effect of first-time mothers is relatively noisy oscillating around zero. Because of the differential first-stage for this specific heterogeneous group, i.e. first-time mothers spending less time on parental leave, reported interacted effects measure a reduced form effect so that longer leave might actually be worse for first-time mothers if these were to spend the same amount of months on leave as non-first-time mothers. With little formal childcare available, finding childcare arrangements for two or more children might be more difficult and, thus, stressful for non-first-time mothers, so that these might be harmed less by a longer parental leave duration.

When it comes to heterogeneity by gender of the child, there is an interesting pattern too. The 2000 reform, significantly increases all health domains for mothers of boys and girls, with a less negative (though not significantly different) impact on outpatient health of mothers of girls. This suggests, that the turning point where leave has negative effects on maternal health is earlier for mothers of boys than for those of girls. This lines up well with research by Baker (2021) who shows that boys aged 0–1 receive more hostile parenting and are more likely to be viewed as difficult by their parents. In combination with longer leave a more confrontational parenting style for parents of young boys (Baker, 2021) might be worse for maternal health. However, looking at inpatient health, the effect is reverse. Mothers of girls score significantly lower when it comes to hospitalization, thus, suggesting that the less intense contact with healthcare professionals in the outpatient sector may have actually led to more severe conditions requiring hospitalization.

Next, I examine the effect of the reform by the health of the newborn. I create an indicator for the child being born with low birth weight or preterm, meaning with less than 2500 g or prior to 37 weeks of gestation as defined by the WHO. A low birth weight baby will generally be put in neonatal care to gain weight. Several studies have shown that health at birth (and especially low birth weight) predicts later life outcomes of children (Almond et al., 2018; Carneiro et al., 2015; Schwandt, 2018), so that *unhealthy* babies may demand an extra amount of maternal care which may last into later childhood. As such, one would expect mothers from unhealthy babies to benefit more from longer leave, which is what the sign of the point estimates confirms. Mothers with unhealthy babies, proxied by a low birth weight, show a significant relative decline in all outpatient health measures compared to mothers with healthy babies. While the effect for preterm births is not significant, the direction of the effect stays the same. This difference in significance of the two proxies might arise since preterm babies might be developed enough without the need of intensive care.

In all, the heterogeneity analysis reveals insights which are of great importance for policy makers. The results suggest that mothers with unhealthy babies are harmed less from longer leave compared to mothers with healthy babies as proxied by the birth weight of the baby. Interestingly, there were several additional family policy reforms in the 2000s. These reforms allowed parents to choose from a variety of options with different levels of leave duration and benefits. When such a choice is available it might be beneficial to inform parents about potential pros and cons of each choice and, especially, allow mothers to choose their preferred parental leave duration after birth when the health of the newborn has already materialized.

Table 6
Heterogeneity of 2000 reform effects.

	Outpatient costs (1)	Outpatient doctor visits (2)	Medication (# packages) (3)	Days of hospitalization (4)
Panel A: First child				
RD	1.161*	1.068	1.796**	−0.164
	(0.593)	(0.785)	(0.857)	(0.444)
Interaction	−0.370	0.206	−0.572	0.768
	(0.801)	(1.074)	(1.089)	(0.595)
Panel B: Girl				
RD	1.385**	1.582**	2.045**	−0.432
	(0.568)	(0.754)	(0.800)	(0.448)
Interaction	−0.829	−0.883	−1.080	1.270**
	(0.804)	(1.069)	(1.096)	(0.605)
Panel C: Low birth weight				
RD	1.257***	1.496***	1.849***	0.409
	(0.410)	(0.555)	(0.576)	(0.272)
Interaction	−0.044**	−0.052***	−0.052***	−0.027
	(0.018)	(0.020)	(0.016)	(0.023)
Panel D: Preterm				
RD	1.071***	1.167**	1.689***	0.258
	(0.407)	(0.541)	(0.569)	(0.305)
Interaction	−0.016	0.002	−0.030	−0.009
	(0.023)	(0.032)	(0.023)	(0.014)
Observations	3187	3187	3187	3187

Notes: This table provides RD parameter estimates for the coefficient β_1 of Eq. (1) and an interaction with the stated heterogeneous group for the reform sample in the year 2000. The sample includes all PL eligible mothers within a bandwidth of 60 days around the policy cutoff. Health outcomes are standardized and accumulated over time up to 2007, so that birth occurred 7 years ago for the reform in 2000. Heteroskedastic-robust standard errors are shown in parentheses and significance levels are indicated by */**/*** for the 10%/5%/1% level.

5.5. Robustness

I conduct several robustness analyses. I start by showing Table A.9 with estimates on general health outcomes, where I additionally control for maternal characteristics, such as her age at birth, marital status, origin, and high-SES status and child characteristics, such as weight at birth, gestational length, Apgar score 5 min after birth, sex of the child, and birth order of the child.[17] In a standard RD setting, including these controls should not matter, especially as these controls varied smoothly around the cutoff. The results are essentially unchanged to the previously stated ones and thus confirm the validity of the RD in this context.

In Fig. A.4, I report the robustness to different choices of bandwidths. The reported results show how estimates on the general health outcome measures vary from choosing a bandwidth of 30 days up to 120 days, including my baseline results with a bandwidth of 60 days. The sign of the effects, and thus the interpretation of each policy change, stays constant over all choices of bandwidths. With larger bandwidths, the effects generally loose in significance. However, with larger bandwidths a linear approximation might also be less precise to describe the data correctly. Furthermore, as Table A.1 shows, the optimal bandwidth selector as suggested by Calonico et al. (2018) prefers shorter bandwidths in the range of the selected one in the baseline regressions.

Table A.10 varies the functional form of the RD from a linear model with the same trend on both sides of the cutoff, to a quadratic model with a differential trend on both sides of the cutoff. The interpretation of the results remains the same, while the most flexible specification (i.e. quadratic polynomial with differential trend on both sides of the cutoff) provides the largest and most significant results. Moving to even higher-order polynomials such as controlling for third or fourth degree polynomials is not recommended as Gelman and Imbens (2019) show that this can lead to nonsensical results.

Finally, I report the results of a placebo analysis where I assume a reform in July for any year from 1989 to 1999 where no reform has taken place. This sensitivity check is reported in Table A.11 and shows no significant effect on any of the health outcome measures except for a marginally significant reduction of outpatient costs in 1993 and prescriptions in 1994. However, reporting two estimates significant is within the expected range for a table with estimates for 4 outcomes × 9 reforms.

5.6. Optimal policy

In the analysis so far, I have focused on the 2000 reform because of the sharp increase in the duration of parental leave from 1.5 years to 2.5 years from an already generous baseline. In a variation of the previous analysis, I now want to study the reforms in

[17] High-SES mothers are characterized based on a combination of educational and wage criteria. A high-SES mothers is defined as a mother who has completed at least high school or who finished apprenticeship training or intermediate vocational school and earns a wage above the median wage.

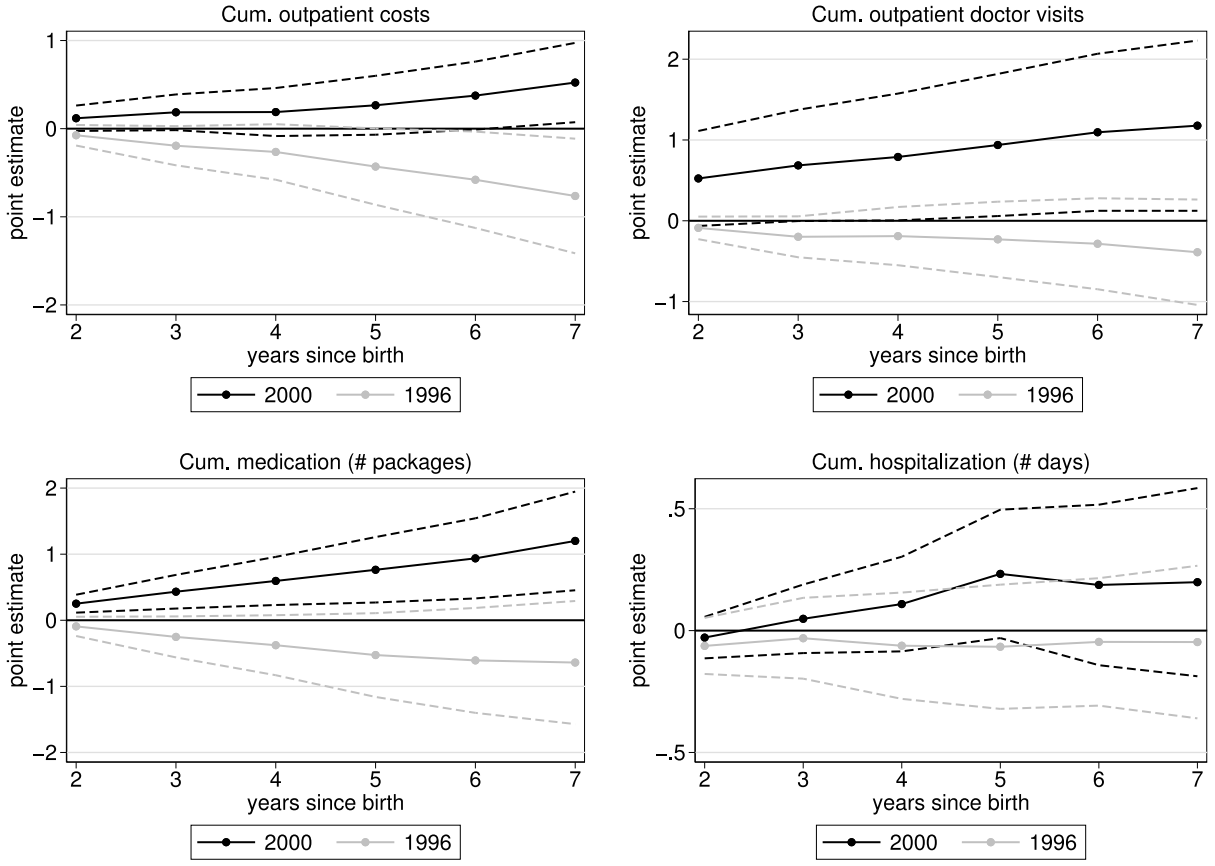


Fig. 6. Comparing the reforms directly: 1996 vs. 2000 reform. *Notes:* This figure shows RD parameter estimates for the coefficient β_1 of Eq. (1) on different standardized accumulated health outcomes. Each dot represents a separate estimation on accumulating the health outcome up to this point in time. The sample includes all PL eligible mothers within a bandwidth of 60 days around the policy cutoff. Dashed lines report 95 percent confidence intervals. RD point estimates for the 2000 reform are reported in black, those for the 1996 reform in gray.

1990 and 1996 also mentioned in Section 2.1. However, due to data availability from 1998 to 2007 and dynamic effects over time since birth, the three reforms can only be directly compared with certain adjustments. To do so, I focus on the years 8 to 11 after giving birth for the first two reforms and the years 2 to 7 for the two latter reforms. This adjustment enables having two overlapping time windows so that one is also able to draw suggestive conclusions about the health gradient.

Results for the comparison of the 1996 and 2000 reforms are reported in Fig. 6. Outpatient measures diverge for the two reforms and increase over time. This means that the 1996 reform – reducing maternal leave from 2 years to 1.5 years – actually improves maternal health and that the 2000 reform, as already mentioned, deteriorates maternal health. For the direct comparison of the 1990 and 1996 reforms, overlapping time windows are short and only available 8 to 11 years after giving birth. Thus, effects are relatively small and noisy as reported in Fig. A.5.

In order to say something about the optimal parental leave duration, it is obviously very important to mention here again that these results show reduced form effects where potential other channels as studied in Section 5.3 may vary too. For example, this is especially true for the 1996 reform which increased paternal involvement in child rearing and may, thus, explain part of the positive effect on maternal health. Nevertheless, these direct comparisons provide suggestive evidence for a hump-shaped relationship between parental leave duration and maternal health with optimal leave durations in the range of 1 to 2 years.

6. Conclusion

I exploit an Austrian reform in the year 2000 that changed the length of maternal leave drastically from 1.5 to 2.5 years to estimate the effect on maternal health in the short- to the medium-run. There was strict enforcement: mothers giving birth on June 30 were subject to the old policy regime and mothers giving birth on July 1 to the new one. This is the optimal setting for using a regression discontinuity design, where I additionally control for seasonality by estimating it in differences to the preceding non-reform year. I can evaluate maternal health from 1998 to 2007. This allows me to estimate the impact of parental leave duration from 2 years pre-birth up to 7 years post-birth.

The results provide evidence for negative returns to maternal health for *long* parental leave of 2.5 instead of 1.5 years. Negative returns are measured each with an increase in outpatient health care expenditures, more doctor visits in the outpatient sector, and more prescribed packages of medication. The drivers for the observed pattern are multifold. In more detail, it is especially maternal mental health which explains the results and not cardiovascular health. Furthermore, changes in fertility, spacing between births, long-run maternal labor market outcomes, disposable income, and alternative mode of childcare may only play a minor role in explaining the negative effects. A general increasing trend in paternal uptake of parental leave may mitigate negative effects and act as buffer for even worse maternal health outcomes. Generally, effects accumulate over time, i.e. up to 7 years after giving birth. Thus, one would expect even stronger effects in the very long-run.

This paper adds to the findings of Bütikofer et al. (2021) who state that there are diminishing returns to maternity leave length. The findings in this paper show that the returns can even become negative with too long parental leave duration. In combination with the results on child development as reviewed by Rossin-Slater (2018) – extensions in existing paid leave policies not affecting any measures of child well-being – this study helps policy makers for designing optimal parental leave.

Altogether, the heterogeneity analysis shows that it could also be important to take birth outcomes, such as the baby being born with low birth weight, into account when informing parents about potential positive and negative effects of different parental leave schemes. However, for a complete picture, the direct effect on fathers and spillovers from mothers to fathers should be studied more in future research. Failing to measure the impact on paternal outcomes implicitly sets their value at zero and hinders gender equality.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix

See Figs. A.1–A.5 and Tables A.1–A.11.

Table A.1
Optimal bandwidth selection according to Calonico et al. (2018).

Dependent variable	Parental leave length (1)	Outpatient costs (2)	Doctor visits (3)	Medication (# packages) (4)	Days of hospitalization (5)
Optimal bandwidth	45.606	35.023	45.105	57.507	42.389

Notes: This table shows optimal bandwidths (in days) according to Calonico et al. (2018) for the different types of outcomes using the command *bwselect* in STATA. Other default settings for the command rdbwselect are: p(1) for local linear regression, triangular for the kernel, and mserd for one common MSE-optimal bandwidth selector for the RD treatment effect estimator. RD-results with the respective optimal bandwidth are graphically reported in Fig. A.4.

Table A.2
Covariate balance table.

	Maternal age (1)	High SES (2)	Married (3)	Foreign (4)	Birth weight (5)	Gestation length (6)	5 min Apgar (7)	Girl (8)	Birth order (9)
Panel A: RD estimates									
Post	−0.340	0.016	−0.015	−0.029	24.280	−0.024	0.019	−0.059	0.052
	(0.362)	(0.039)	(0.038)	(0.025)	(40.659)	(0.140)	(0.049)	(0.039)	(0.070)
Mean of Dep. Var.	28.299	0.520	0.631	0.119	3354.802	39.678	9.685	0.501	1.727
Observations	3187	3187	3187	3187	3187	3187	3183	3187	3187
Panel B: RD-DiD estimates									
Post ∗ reform	−0.255	0.023	−0.013	−0.028*	12.490	−0.044	0.002	−0.041	0.012
	(0.244)	(0.026)	(0.025)	(0.017)	(26.830)	(0.094)	(0.034)	(0.026)	(0.045)
Mean of Dep. Var.	28.299	0.520	0.631	0.119	3354.802	39.678	9.685	0.501	1.727
Observations	6482	6482	6482	6482	6482	6482	6473	6482	6482

Notes: This table provides parameter estimates for the coefficient β_1 of Eq. (1) for the 2000 reform. The sample includes all PL eligible mothers within a bandwidth of 60 days around each policy cutoff. Panel A estimates standard RD estimates, while Panel B incorporates a DiD layer by including the pre-reform year to difference out seasonality which has nothing to do with the reform itself. Heteroskedastic-robust standard errors are shown in parentheses and significance levels are indicated by */**/*** for the 10%/5%/1% level.

Table A.3
Reform effects on general health outcomes controlling for seasonality.

	Outpatient costs (1)	Outpatient doctor visits (2)	Medication (# packages) (3)	Days of hospitalization (4)
Post ∗ reform	0.452	0.512	0.943***	−0.163
	(0.280)	(0.324)	(0.331)	(0.195)
Mean of Dep. Var.	0.147	0.051	−0.087	0.145
Observations	6482	6482	6482	6482

Notes: This table provides RD-DiD parameter estimates for the coefficient β_1 of Eq. (1) on general health outcomes. Therefore estimates in this table incorporate a DiD layer by including the pre-reform year to difference out seasonality which has nothing to do with the reform itself. Health outcomes are standardized and accumulated over time and shown for the last observable year in 2007. The sample includes all PL eligible mothers within a bandwidth of 60 days around the policy cutoff. Heteroskedastic-robust standard errors are shown in parentheses and significance levels are indicated by */**/*** for the 10%/5%/1% level.

Table A.4
Reform effects on alternative outcomes.

	Total fertility	Employment status		Daily wages	
		5 years post birth	10 years post birth	5 years post birth	10 years post birth
	(1)	(2)	(3)	(4)	(5)
Post	0.081	−0.052	−0.043	−1.979	−4.556
	(0.071)	(0.035)	(0.029)	(2.325)	(2.981)
Mean of Dep. Var.	2.333	0.736	0.864	24.943	35.732
Observations	3187	3187	3187	3187	3187

Notes: This table provides RD parameter estimates for the coefficient β_1 of Eq. (1) on alternative outcomes. The first column reports the effect on total fertility, columns (2)–(3) show the effect on the employment status 5 or 10 years after giving birth, and columns (4)–(5) report the effect on daily wages 5 or 10 years after giving birth, respectively. The sample includes all PL eligible mothers within a bandwidth of 60 days around the policy cutoff. Heteroskedastic-robust standard errors are shown in parentheses and significance levels are indicated by */**/*** for the 10%/5%/1% level.

Table A.5
Reform effects on general health outcomes – Mediation analysis.

	Baseline estimation (1)	Incl. total fertility (2)	Incl. empl. status (3)	Incl. both (4)
Panel A: Dependent variable outpatient costs				
Post	0.990**	0.986**	0.922**	0.923**
	(0.403)	(0.402)	(0.402)	(0.402)
Mean of Dep. Var.	0.131	0.131	0.131	0.131
Panel B: Dependent variable outpatient doctor visits				
Post	1.177**	1.171**	1.118**	1.117**
	(0.538)	(0.539)	(0.537)	(0.538)
Mean of Dep. Var.	0.038	0.038	0.038	0.038
Panel C: Dependent variable medication (# packages)				
Post	1.536***	1.539***	1.491***	1.495***
	(0.552)	(0.552)	(0.551)	(0.550)
Mean of Dep. Var.	−0.070	−0.070	−0.070	−0.070
Panel D: Dependent variable days of hospitalization				
Post	0.212	0.188	0.110	0.099
	(0.298)	(0.297)	(0.290)	(0.290)
Mean of Dep. Var.	0.075	0.075	0.075	0.075
Controlling for total # of kids		✓		✓
Controlling for empl. status			✓	✓
Observations	3187	3187	3187	3187

Notes: This table provides RD parameter estimates for the coefficient β_1 of Eq. (1) for the reform sample in the year 2000. The sample includes all PL eligible mothers within a bandwidth of 60 days around the policy cutoff. Health outcomes are standardized and accumulated over time up to 2007, so that birth occurred 7 years ago for the reform in 2000. Column (1) repeats the baseline estimates from Table 3, column (2) additionally controls for the total fertility, column (3) controls for the employment status at the last observable point in time, and column (4) controls for both. Heteroskedastic-robust standard errors are shown in parentheses and significance levels are indicated by */**/*** for the 10%/5%/1% level.

Table A.6
Reform effects on general health outcomes – Two-stage least squares.

	Outpatient costs (1)	Outpatient doctor visits (2)	Medication (# packages) (3)	Days of hospitalization (4)
Months on PL	0.119**	0.141**	0.184***	0.025
	(0.048)	(0.064)	(0.067)	(0.036)
Mean of Dep. Var.	0.131	0.038	−0.070	0.075
Observations	3187	3187	3187	3187

Notes: This table provides 2SLS parameter estimates on general maternal health outcomes using the results from Table 1 as first stage. The sample includes all PL eligible mothers within a bandwidth of 60 days around the policy cutoff. Health outcomes are standardized and accumulated over time up to 2007. Heteroskedastic-robust standard errors are shown in parentheses and significance levels are indicated by */**/*** for the 10%/5%/1% level.

Table A.7
Heterogeneity of reform effects by below/above median wage.

	Outpatient costs (1)	Outpatient doctor visits (2)	Medication (# packages) (3)	Days of hospitalization (4)
RD	1.088**	1.754**	1.578**	0.491
	(0.535)	(0.783)	(0.759)	(0.334)
Interaction	−0.113	−0.962	0.057	−0.489
	(0.805)	(1.083)	(1.104)	(0.597)
Mean of Dep. Var.	0.131	0.038	−0.070	0.075
Observations	3141	3141	3141	3141

Notes: This table provides RD parameter estimates for the coefficient β_1 of Eq. (1) and an interaction with the stated heterogeneous group of above median wage for the reform sample in the year 2000. The sample includes all PL eligible mothers within a bandwidth of 60 days around the policy cutoff. Health outcomes are standardized and accumulated over time up to 2007, so that birth occurred 7 years ago for the reform in 2000. Heteroskedastic-robust standard errors are shown in parentheses and significance levels are indicated by */**/*** for the 10%/5%/1% level.

Table A.8
Reform effects on paternal parental leave uptake.

	Paternal leave uptake	
	(1)	(2)
Post	0.058	0.041
	(0.059)	(0.037)
RD	✓	
RD-DiD		✓
Mean of Dep. Var.	0.062	0.062
Observations	448	931

Notes: This table provides parameter estimates for the coefficient β_1 of Eq. (1) for the reform sample. The sample includes all fathers who can be linked (only a small fraction) to PL eligible mothers in Upper Austria within a bandwidth of 60 days around each policy cutoff. Column (1) estimates standard RD estimates, while column (2) incorporates a DiD layer by including the pre-reform year to difference out seasonality which has nothing to do with the reform itself. Heteroskedastic-robust standard errors are shown in parentheses and significance levels are indicated by */**/*** for the 10%/5%/1% level.

Table A.9
Reform effects on general health outcomes including covariates.

	Outpatient costs (1)	Outpatient doctor visits (2)	Medication (# packages) (3)	Days of hospitalization (4)
Panel A: RD estimates				
Post	1.020**	1.188**	1.521***	0.160
	(0.407)	(0.537)	(0.556)	(0.302)
Add. controls	✓	✓	✓	✓
Mean of Dep. Var.	0.131	0.038	−0.070	0.075
Observations	3183	3183	3183	3183
Panel B: RD-DiD estimates				
Post ∗ reform	0.455	0.482	0.918***	−0.197
	(0.280)	(0.323)	(0.332)	(0.194)
Add. controls	✓	✓	✓	✓
Mean of Dep. Var.	0.147	0.051	−0.087	0.145
Observations	6473	6473	6473	6473

Notes: This table provides parameter estimates for the coefficient β_1 of Eq. (1) including covariates for the reform sample in the year 2000. All columns additionally control for predetermined maternal characteristics (age at birth, marital status, high SES, and origin) and child characteristics (gestational length, weight at birth, Apgar score, sex, and birth order of the child). The sample includes all PL eligible mothers within a bandwidth of 60 days around the policy cutoff. Panel A estimates standard RD estimates, while Panel B incorporates a DiD layer by including the pre-reform year to difference out seasonality which has nothing to do with the reform itself. Health outcomes are standardized and accumulated over time up to 2007, so that birth occurred 7 years ago for the reform in 2000. Heteroskedastic-robust standard errors are shown in parentheses and significance levels are indicated by */**/*** for the 10%/5%/1% level.

Table A.10
Reform effects on general health outcomes – Robustness to different functional forms.

	Outpatient costs (1)	Outpatient doctor visits (2)	Medication (# packages) (3)	Days of hospitalization (4)
Panel A: Linear				
Post	0.523	0.897**	1.000**	−0.083
	(0.351)	(0.426)	(0.475)	(0.233)
Panel B: Linear interaction				
Post	0.990**	1.177**	1.536***	0.212
	(0.403)	(0.538)	(0.552)	(0.298)
Panel C: Quadratic				
Post	0.629	1.070*	1.341*	−0.160
	(0.460)	(0.568)	(0.689)	(0.285)
Panel D: Quadratic interaction				
Post	1.629***	1.518**	2.217***	0.302
	(0.597)	(0.774)	(0.782)	(0.453)
Mean of Dep. Var.	0.131	0.038	−0.070	0.075
Observations	3187	3187	3187	3187

Notes: This table provides parameter estimates for the coefficient β_1 of Eq. (1) for the reform sample in the year 2000. The sample includes all PL eligible mothers within a bandwidth of 60 days around the policy cutoff. Panel A reports RD estimates with the same linear trend on both sides of the cutoff; Panel B allows for a differential linear trend on both sides of the cutoff (baseline estimates); Panel C controls for the same quadratic trend on both sides of the cutoff; and Panel D allows this quadratic trend to vary on both sides of the cutoff. Health outcomes are standardized and accumulated over time up to 2007, so that birth occurred 7 years ago for the reform in 2000. Heteroskedastic-robust standard errors are shown in parentheses and significance levels are indicated by */**/*** for the 10%/5%/1% level.

Table A.11
Placebo reforms.

	Outpatient costs (1)	Outpatient doctor visits (2)	Medication (# packages) (3)	Days of hospitalization (4)
Panel A: RD estimates for non-reform year 1989				
Post	−0.726	−0.777	−0.167	−0.566
	(0.555)	(0.555)	(0.491)	(0.476)
Panel B: RD estimates for non-reform year 1991				
Post	0.406	−0.021	0.133	0.061
	(0.495)	(0.490)	(0.479)	(0.246)
Panel C: RD estimates for non-reform year 1992				
Post	0.265	0.245	0.399	0.128
	(0.584)	(0.527)	(0.551)	(0.287)
Panel D: RD estimates for non-reform year 1993				
Post	−0.677**	−0.578	−0.236	−0.282
	(0.344)	(0.503)	(0.466)	(0.311)
Panel E: RD estimates for non-reform year 1994				
Post	−0.432	−0.697	−0.846*	−0.521
	(0.329)	(0.516)	(0.471)	(0.387)
Panel F: RD estimates for non-reform year 1995				
Post	−0.584	−0.115	−0.207	−0.470
	(0.418)	(0.499)	(0.567)	(0.407)
Panel G: RD estimates for non-reform year 1997				
Post	−0.037	−0.316	−0.524	−0.765
	(0.800)	(0.547)	(0.580)	(0.465)
Panel H: RD estimates for non-reform year 1998				
Post	−0.260	−0.833	−0.903	−0.562
	(0.458)	(0.538)	(0.678)	(0.914)
Panel I: RD estimates for non-reform year 1999				
Post	−0.246	0.333	−0.154	−0.032
	(0.485)	(0.521)	(0.531)	(0.265)

Notes: This table provides RD parameter estimates for the coefficient β_1 of Eq. (1) for all non-reform samples from 1989 to 1999. The sample includes all PL eligible mothers within a bandwidth of 60 days around the placebo policy cutoff in July. Health outcomes are standardized and accumulated over time up to 2007. Heteroskedastic-robust standard errors are shown in parentheses and significance levels are indicated by */**/*** for the 10%/5%/1% level.

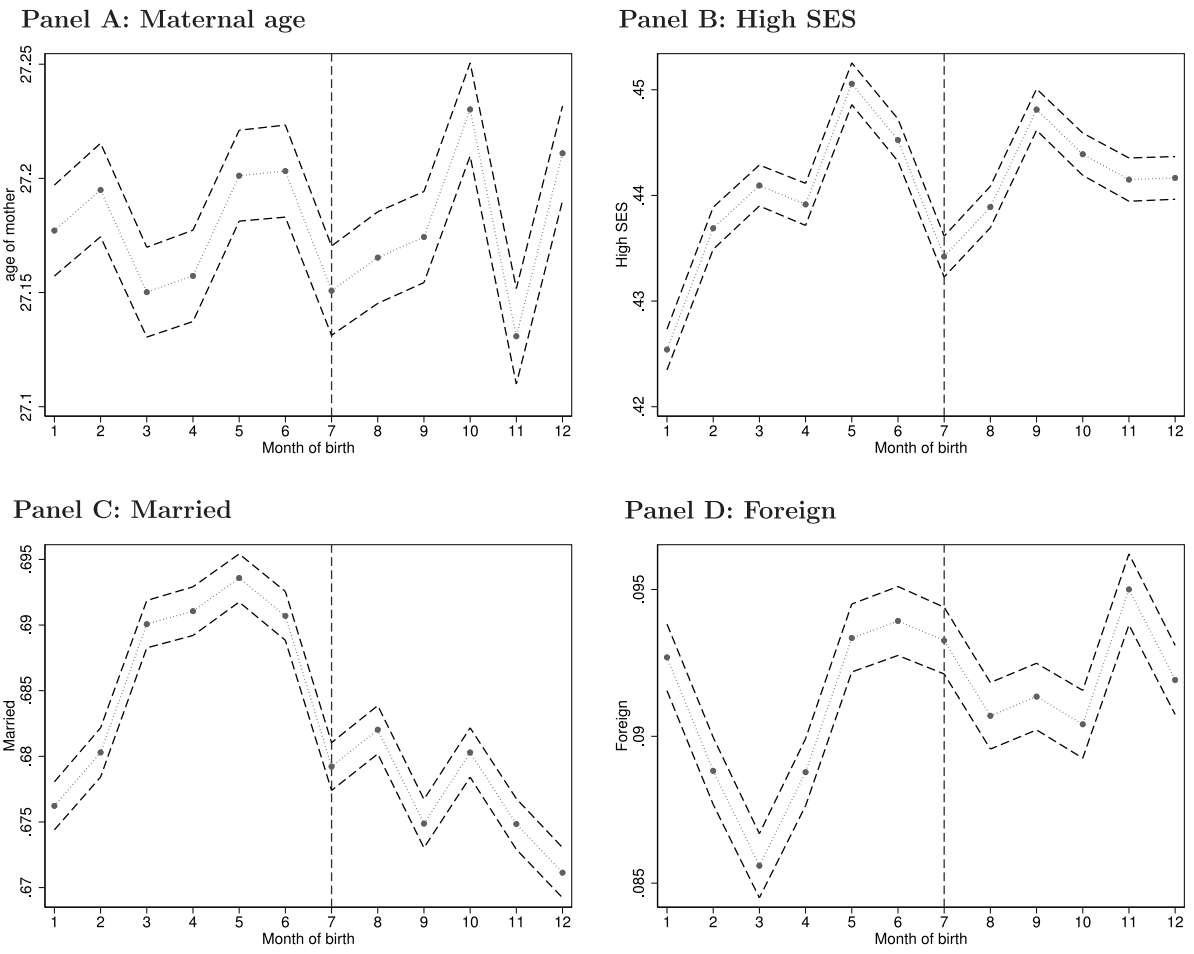


Fig. A.1. Seasonality in maternal characteristics. *Notes:* This figure shows averages for different predetermined maternal characteristics by month of birth of the child. The sample covers all mothers in Upper Austria (as described in Table 2 in column 2). Averages are indicated by black dots, while dashed lines correspond to the 95 percent confidence interval. The vertical dashed line shows the month of July – the month in which all studied reforms take place.

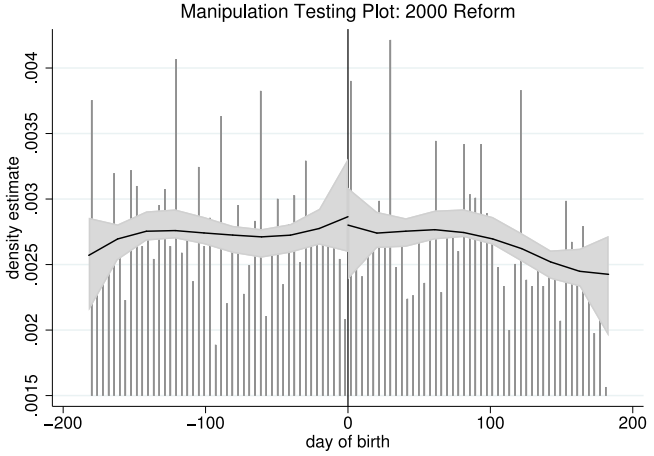


Fig. A.2. McCrary density test. *Notes:* This figure shows the density of the running variable (i.e. the day of birth of a child) for the 2000 reform. Furthermore it includes the manipulation test as suggested by McCrary (2008) in order to prove that there is no selection of mothers to be on either side of the cutoff. The sample includes all PL eligible mothers in Upper Austria.

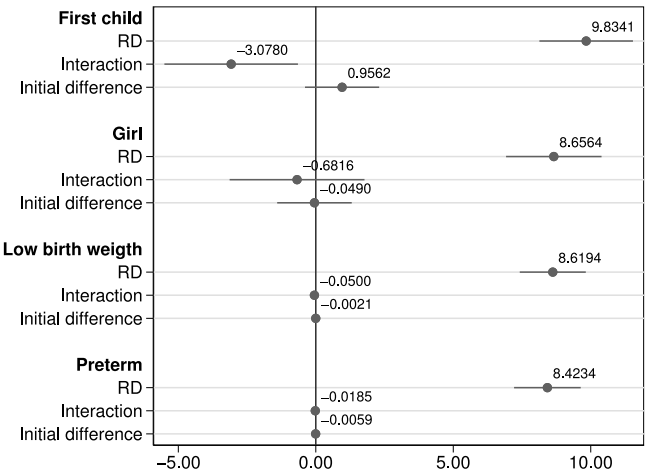


Fig. A.3. Heterogeneity in parental leave duration by child characteristics. *Notes:* This figures shows RD parameter estimates for the coefficient β_1 of Eq. (1), an interaction with the stated heterogeneous group, and an estimate for the initial difference in parental leave duration taken by the stated heterogeneous group. The sample includes all PL eligible mothers within a bandwidth of 60 days around the policy cutoff. 95% confidence intervals are shown with the vertical lines.

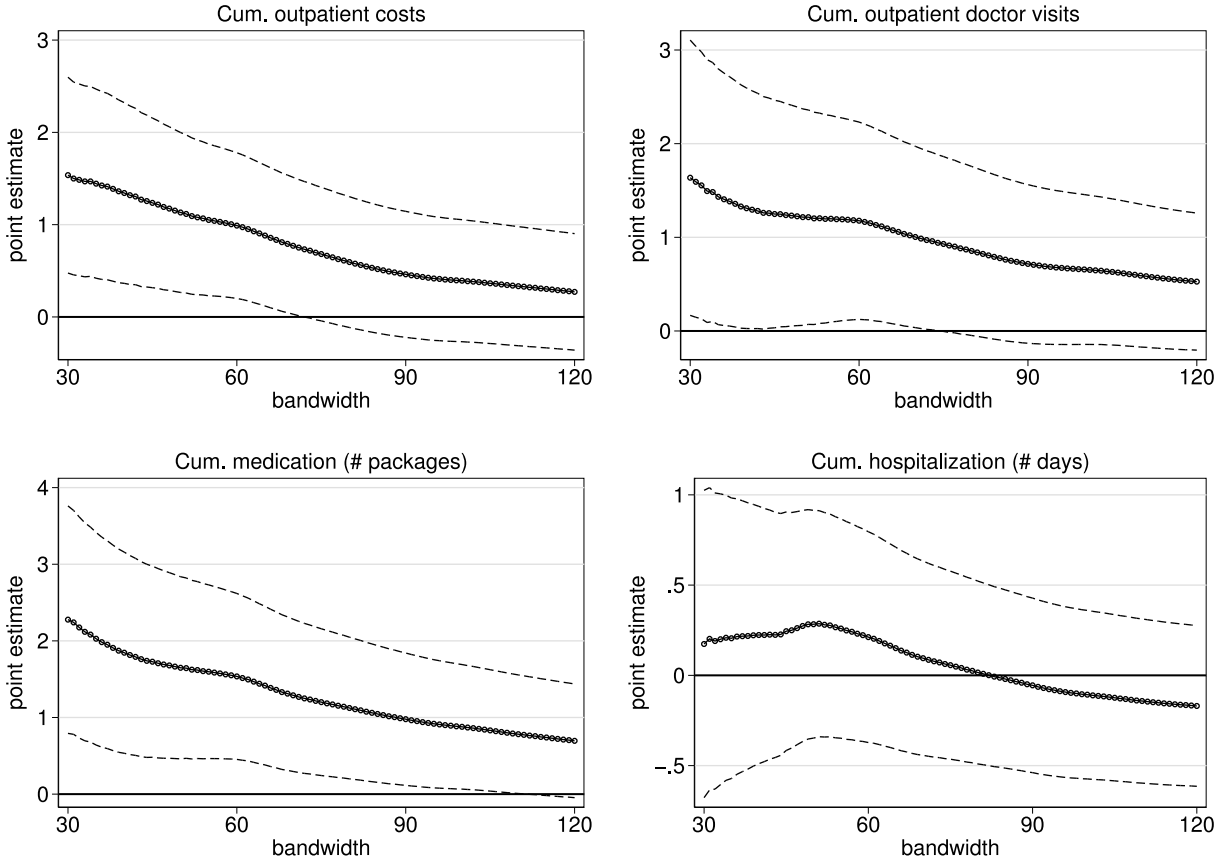


Fig. A.4. Robustness to different choices of bandwidth. *Notes:* This figures shows RD parameter estimates for the coefficient β_1 of Eq. (1) on different standardized accumulated health outcomes. The sample includes all PL eligible mothers. Each hollow circle represents a separate estimation, varying the bandwidth from 30 to 120 days around the policy cutoff. Dashed lines report 95 percent confidence intervals.

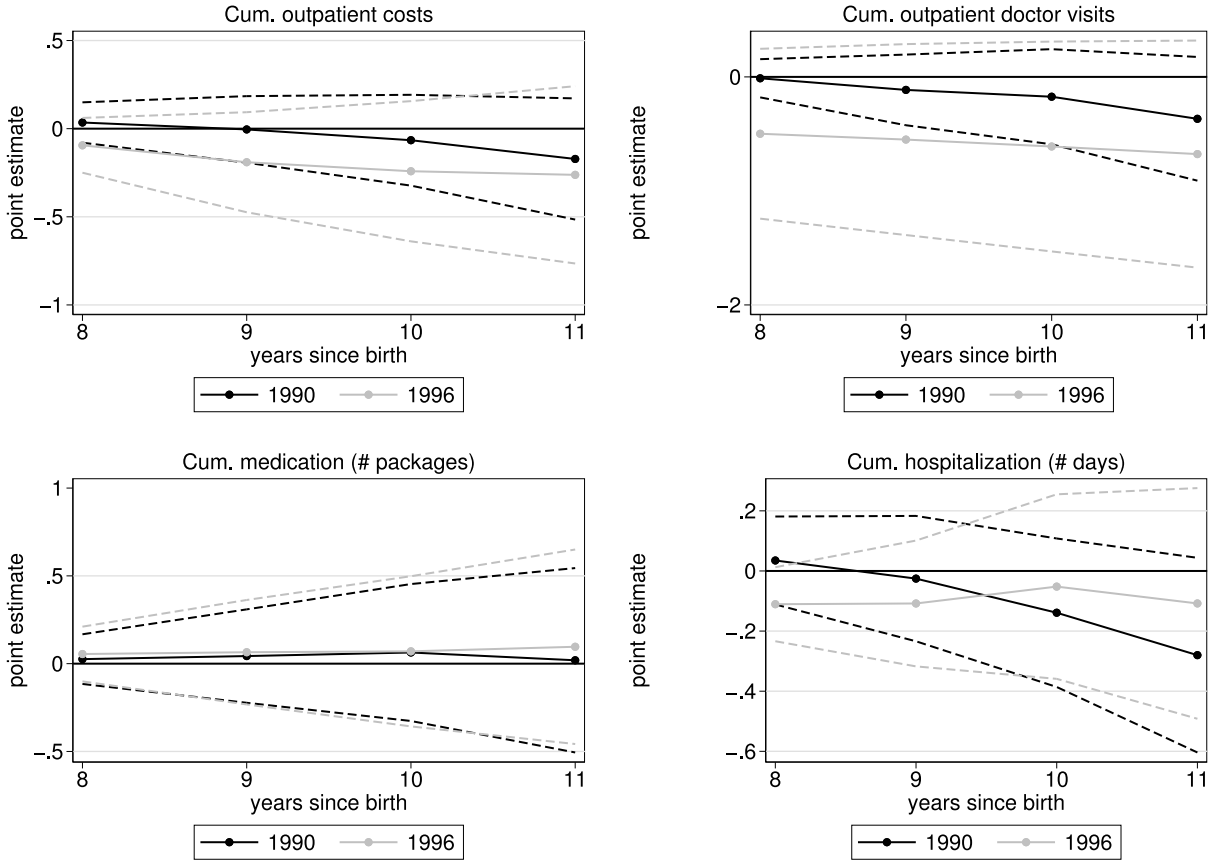


Fig. A.5. Comparing the reforms directly: 1990 vs. 1996 reform. *Notes:* This figures shows RD parameter estimates for the coefficient β_1 of Eq. (1) on different standardized accumulated health outcomes. Each dot represents a separate estimation on accumulating the health outcome up to this point in time. The sample includes all PL eligible mothers within a bandwidth of 60 days around the policy cutoff. Dashed lines report 95 percent confidence intervals. RD point estimates for the 1990 reform are reported in black, those for the 1996 reform in gray.

References

Aitken, Z., Garrett, C.C., Hewitt, B., Keogh, L., Hocking, J.S., Kavanagh, A.M., 2015. The maternal health outcomes of paid maternity leave: a systematic review. Soc. Sci. Med. 130, 32–41.

Almond, D., Currie, J., Duque, V., 2018. Childhood circumstances and adult outcomes: act II. J. Econ. Lit. 56 (4), 1360–1446.

Andres, E., Baird, S., Bingenheimer, J.B., Markus, A.R., 2016. Maternity leave access and health: a systematic narrative review and conceptual framework development. Matern. Child Health J. 20 (6), 1178–1192.

Baker, M., 2021. Gender Differences in (some) Formative Inputs to Child Development. Technical report, National Bureau of Economic Research.

Baker, M., Milligan, K., 2008a. How does job-protected maternity leave affect mothers' employment? J. Labor Econ. 26 (4), 655–691.

Baker, M., Milligan, K., 2008b. Maternal employment, breastfeeding, and health: evidence from maternity leave mandates. J. Health Econ. 27 (4), 871–887.

Baker, M., Milligan, K., 2010. Evidence from maternity leave expansions of the impact of maternal care on early child development. J. Hum. Resour. 45 (1), 1–32.

Baker, M., Milligan, K., 2015. Maternity leave and children's cognitive and behavioral development. J. Popul. Econ. 28 (2), 373–391.

Bana, S.H., Bedard, K., Rossin-Slater, M., 2020. The impacts of paid family leave benefits: regression kink evidence from California administrative data. J. Policy Anal. Manage. 39 (4), 888–929.

Berger, L.M., Hill, J., Waldfogel, J., 2005. Maternity leave, early maternal employment and child health and development in the US. Econ. J. 115 (501), F29–F47.

Beuchert, L.V., Humlum, M.K., Vejlin, R., 2016. The length of maternity leave and family health. Labour Econ. 43, 55–71.

Bütikofer, A., Riise, J., Skira, M.M., 2021. The impact of paid maternity leave on maternal health. Am. Econ. J.: Econ. Policy 13 (1), 67–105.

Calonico, S., Cattaneo, M.D., Farrell, M.H., Titiunik, R., 2018. Regression discontinuity designs using covariates. Rev. Econ. Stat..

Carneiro, P., Løken, K.V., Salvanes, K.G., 2015. A flying start? Maternity leave benefits and long-run outcomes of children. J. Polit. Econ. 123 (2), 365–412.

Chatterji, P., Markowitz, S., 2005. Does the length of maternity leave affect maternal health? Southern Econ. J. 72 (1), 16–41.

Chatterji, P., Markowitz, S., 2012. Family leave after childbirth and the mental health of new mothers. J. Mental Health Policy Econ. 15 (2), 61.

Chowdhury, R., Sinha, B., Sankar, M.J., Taneja, S., Bhandari, N., Rollins, N., Bahl, R., Martines, J., 2015. Breastfeeding and maternal health outcomes: a systematic review and meta-analysis. Acta Paediatr. 104, 96–113.

Chuard, C., 2020. Womb at work: the missing impact of maternal employment on newborn health. J. Health Econ. 73, 102342.

Chung, M., Raman, G., Chew, P., Magula, N., Trikalinos, T., Lau, J., et al., 2007. Breastfeeding and maternal and infant health outcomes in developed countries. Evid. Technol. Assess. (Full Rep.) 153 (153), 1–186.

Cools, S., Fiva, J.H., Kirkebøen, L.J., 2015. Causal effects of paternity leave on children and parents. Scand. J. Econ. 117 (3), 801–828.

Currie, J., Schwandt, H., 2013. Within-mother analysis of seasonal patterns in health at birth. Proc. Natl. Acad. Sci. 110 (30), 12265–12270.
Dagher, R.K., McGovern, P.M., Dowd, B.E., 2014. Maternity leave duration and postpartum mental and physical health: implications for leave policies. J. Health Politics Policy Law 39 (2), 369–416.
Dahl, G.B., Løken, K.V., Mogstad, M., Salvanes, K.V., 2016. What is the case for paid maternity leave? Rev. Econ. Stat. 98 (4), 655–670.
Danzer, N., Halla, M., Schneeweis, N., Zweimüller, M., 2020. Parental leave, (in) formal childcare and long-term child outcomes. J. Hum. Resour. 0619–10257R1.
Danzer, N., Lavy, V., 2018. Paid parental leave and children's schooling outcomes. Econ. J. 128 (608), 81–117.
Dörfler, S., Blum, S., 2014. Europäische Kinderbetruungskulturen im Vergleich: Jüngste Entwicklungen in der Vorschulischen Betreuung in Deutschland, Frankreich, Österreich und Schweden. Technical report, Österreichisches Institut für Familienforschung an der Universität Wien.
Dustmann, C., Schönberg, U., 2012. Expansions in maternity leave coverage and children's long-term outcomes. Am. Econ. J.: Appl. Econ. 4 (3), 190–224.
Evans, D.K., Jakiela, P., Knauer, H.A., 2021. The impact of early childhood interventions on mothers. Science 372 (6544), 794–796.
Feltner, C., Weber, R.P., Stuebe, A., Grodensky, C.A., Orr, C., Viswanathan, M., 2018. Breastfeeding programs and policies, breastfeeding uptake, and maternal health outcomes in developed countries. Comp. Eff. Rev. 210.
Fontenay, S., Tojerow, I., 2020. Work Disability after Motherhood and How Paternity Leave Can Help. Technical report, IZA Discussion Paper.
Gavin, N.I., Gaynes, B.N., Lohr, K.N., Meltzer-Brody, S., Gartlehner, G., Swinson, T., 2005. Perinatal depression: a systematic review of prevalence and incidence. Obstet. Gynecol. 106 (5 Part 1), 1071–1083.
Gelman, A., Imbens, G., 2019. Why high-order polynomials should not be used in regression discontinuity designs. J. Bus. Econom. Statist. 37 (3), 447–456.
Gore, S.A., Brown, D.M., West, D.S., 2003. The role of postpartum weight retention in obesity among women: a review of the evidence. Ann. Behav. Med. 26 (2), 149–159.
Hahn, J., Todd, P., Van der Klaauw, W., 2001. Identification and estimation of treatment effects with a regression-discontinuity design. Econometrica 69 (1), 201–209.
Johnston, D.W., Lee, W.-S., 2009. Retiring to the good life? The short-term effects of retirement on health. Econom. Lett. 103 (1), 8–11.
Kleven, H., Landais, C., Posch, J., Steinhauer, A., Zweimüller, J., 2020. Do Family Policies Reduce Gender Inequality? Evidence from 60 Years of Policy Experimentation. Technical report, National Bureau of Economic Research.
Kolesár, M., Rothe, C., 2018. Inference in regression discontinuity designs with a discrete running variable. Amer. Econ. Rev. 108 (8), 2277–2304.
Lackey, K.A., Fehrenkamp, B.D., Pace, R.M., Williams, J.E., Meehan, C.L., McGuire, M.A., McGuire, M.K., 2021. Breastfeeding beyond 12 months: is there evidence for health impacts? Annu. Rev. Nutr. 41.
Lalive, R., Schlosser, A., Steinhauer, A., Zweimüller, J., 2013. Parental leave and mothers' careers: the relative importance of job protection and cash benefits. Rev. Econom. Stud. 81 (1), 219–265.
Lalive, R., Zweimüller, J., 2009. How does parental leave affect fertility and return to work? Evidence from two natural experiments. Q. J. Econ. 124 (3), 1363–1402.
Lee, D.S., Lemieux, T., 2010. Regression discontinuity designs in economics. J. Econ. Lit. 48 (2), 281–355.
Liu, Q., Skans, O.N., 2010. The duration of paid parental leave and children's scholastic performance. BE J. Econ. Anal. Policy 10 (1).
McCrary, J., 2008. Manipulation of the running variable in the regression discontinuity design: a density test. J. Econometrics 142 (2), 698–714.
Miller, L.J., 2002. Postpartum depression. JAMA 287 (6), 762–765.
Nguyen, B., Jin, K., Ding, D., 2017. Breastfeeding and maternal cardiovascular risk factors and outcomes: a systematic review. PLoS One 12 (11), e0187923.
Olafsson, A., Steingrimsdottir, H., 2020. How does daddy at home affect marital stability? Econ. J. 130 (629), 1471–1500.
Olivetti, C., Petrongolo, B., 2017. The economic consequences of family policies: lessons from a century of legislation in high-income countries. J. Econ. Perspect. 31 (1), 205–230.
Patnaik, A., 2019. Reserving time for daddy: the consequences of fathers' quotas. J. Labor Econ. 37 (4), 1009–1059.
Persson, P., Rossin-Slater, M., 2019. When dad can stay home: fathers' workplace flexibility and maternal health. Technical report, National Bureau of Economic Research.
Rasmussen, A.W., 2010. Increasing the length of parents' birth-related leave: the effect on children's long-term educational outcomes. Labour Econ. 17 (1), 91–100.
Romano, J.P., Wolf, M., 2005. Stepwise multiple testing as formalized data snooping. Econometrica 73 (4), 1237–1282.
Rooney, B.L., Schauberger, C.W., Mathiason, M.A., 2005. Impact of perinatal weight change on long-term obesity and obesity-related illnesses. Obstet. Gynecol. 106 (6), 1349–1356.
Rossin, M., 2011. The effects of maternity leave on children's birth and infant health outcomes in the United States. J. Health Econ. 30 (2), 221–239.
Rossin-Slater, M., 2018. Maternity and family leave policy. In: Averett, S., Argys, L., Hoffman, S. (Eds.), The Oxford Handbook of Women and the Economy. Oxford University Press, New York.
Rossin-Slater, M., Ruhm, C.J., Waldfogel, J., 2013. The effects of California's paid family leave program on mothers' leave-taking and subsequent labor market outcomes. J. Policy Anal. Manage. 32 (2), 224–245.
Ruhm, C.J., 2000. Parental leave and child health. J. Health Econ. 19 (6), 931–960.
Schönberg, U., Ludsteck, J., 2014. Expansions in maternity leave coverage and mothers' labor market outcomes after childbirth. J. Labor Econ. 32 (3), 469–505.
Schwandt, H., 2018. The Lasting Legacy of Seasonal Influenza: In-Utero Exposure and Labor Market Outcomes. CEPR Discussion Paper.
Shorey, S., Chee, C.Y.I., Ng, E.D., Chan, Y.H., San Tam, W.W., Chong, Y.S., 2018. Prevalence and incidence of postpartum depression among healthy mothers: a systematic review and meta-analysis. J. Psychiatr. Res. 104, 235–248.
Staehelin, K., Bertea, P.C., Stutz, E.Z., 2007. Length of maternity leave and health of mother and child–a review. Int. J. Public Health 52 (4), 202–209.
Tamm, M., 2019. Fathers' parental leave-taking, childcare involvement and labor market participation. Labour Econ. 59, 184–197.
Tanaka, S., 2005. Parental leave and child health across OECD countries. Econ. J. 115 (501).
Van Niel, M.S., Bhatia, R., Riano, N.S., de Faria, L., Catapano-Friedman, L., Ravven, S., Weissman, B., Nzodom, C., Alexander, A., Budde, K., et al., 2020. The impact of paid maternity leave on the mental and physical health of mothers and children: a review of the literature and policy implications. Harv. Rev. Psychiatry 28 (2), 113–126.
Woody, C., Ferrari, A., Siskind, D., Whiteford, H., Harris, M., 2017. A systematic review and meta-regression of the prevalence and incidence of perinatal depression. J. Affect. Disord. 219, 86–92.